Chapter 1

Fourier Series

Basic definitions and examples of Fourier series are given in Section 1. In Section 2 we prove the fundamental Riemann-Lebesgue lemma and discuss the Fourier series from the mapping point of view. Pointwise and uniform convergence of the Fourier series of a function to the function itself under various regularity assumptions are studied in Section 3. As an application, it is shown that every continuous function can be approximated by polynomials in a uniform manner in Section 4. In Section 5 the L^2-theory of Fourier series is discussed. In the two appendices basic facts on series of functions and sets of measure zero are present respectively.

1.1 Definition and Examples

In the previous power series has been studied. Now we come to Fourier series.

First of all, a **trigonometric series** is a series of functions of the form

$$\sum_{n=0}^{\infty}(a_n \cos nx + b_n \sin nx), \quad a_n, b_n \in \mathbb{R}.$$

As $\cos 0x = 1$ and $\sin 0x = 0$, we always set $b_0 = 0$ and express the series as

$$a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx).$$

It is called a **cosine series** if all b_n's vanish and **sine series** if all a_n's vanish. Trigonometric series form an important class of series of functions. In Mathematical Analysis II, we studied the convergence of the series of functions. We recall

- Uniform convergence implies pointwise convergence of a series of functions,

- Absolute convergence implies pointwise convergence of a series of functions,
- Weierstrass M-Test for uniform and absolute convergence (see Appendix I).

For instance, using the fact $|\cos nx|, |\sin nx| \le 1$, Weierstrass M-Test tells us that a trigonometric series is uniformly and absolutely convergent when its coefficients satisfy

$$\sum_n |a_n|, \quad \sum_n |b_n| < \infty \ ,$$

and this is the case when $|a_n|, |b_n| \le Cn^{-s}, \forall n \ge 1$, for some constant C and $s > 1$. Since the partial sums are continuous functions and uniform convergence preserves continuity, the infinite series

$$\varphi(x) \equiv a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

is a continuous function on $\mathbb{R}$. We claim that φ is also of period 2π. For, by pointwise convergence, we have

$$\begin{aligned} \varphi(x + 2\pi) &= \lim_{n\to\infty} \sum_{k=0}^{n} \big(a_k \cos(kx + 2k\pi) + b_k \sin(kx + 2k\pi)\big) \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} (a_k \cos kx + b_k \sin kx) \\ &= \varphi(x), \end{aligned}$$

hence it is 2π-periodic.

Recall that a power series is associated to a function which is smooth at a certain point. Indeed, it is given by the Taylor's series at this point. Let the point be x_0 and f is smooth in an open interval containing x_0, this series is given by

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n \ , \quad c_n = \frac{f^{(n)}(x_0)}{n!} \ .$$

Similarly, there is a trigonometric series associated to an integrable function. It is called the Fourier series of the function. Let us define it now.

Given a 2π-periodic function which is Riemann integrable function f on $[-\pi, \pi]$, its **Fourier series** or **Fourier expansion** is the trigonometric series given by

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos ny \, dy, \quad n \ge 1 \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin ny \, dy, \quad n \ge 1 \qquad \text{and} \\ a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) \, dy. \end{aligned} \tag{1.1}$$

Note that a_0 is the average of the function over the interval. From this definition we gather two basic information. First, the Fourier series of a function involves the integration of the function over an interval, hence any modification of the values of the function over a subinterval, not matter how small it is, may change the Fourier coefficients a_n and b_n. This is unlike power series which only depend on the local properties (derivatives of all order at a designated point). We may say Fourier series depend on the global information but power series only depend on local information. Second, recalling from the theory of Riemann integral, we know that two integrable functions which are equal almost everywhere have the same integral. (We will see the converse is also true, namely, two functions with the same Fourier series are equal almost everywhere.) In Appendix II we recall the concept of a measure zero set and some of its basic properties. Therefore, the Fourier series of two such functions are the same. In particular, the Fourier series of a function is completely determined with its value on the open interval $(-\pi, \pi)$, regardless its values at the endpoints.

The motivation of the Fourier series comes from the belief that for a "nice function" of period 2π, its Fourier series converges to the function itself. In other words, we have

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \ , \quad \forall x \in \mathbb{R}. \tag{1.2}$$

Whenever this holds, the coefficients a_n, b_n are given by (1.1). A formal argument proceeds as follows. Multiply (1.2) by $\cos mx$ and then integrate over $[-\pi, \pi]$. Using the relations

$$\int_{-\pi}^{\pi} \cos nx \cos mx \, dx = \begin{cases} \pi, & n = m \\ 0, & n \neq m \end{cases} ,$$

$$\int_{-\pi}^{\pi} \cos nx \sin mx \, dx = 0 \quad (n, m \geq 1), \qquad \text{and}$$

$$\int_{-\pi}^{\pi} \cos nx \, dx = \begin{cases} 2\pi, & n = 0 \\ 0, & n \neq 0 \end{cases} ,$$

we arrive at the expression of a_n, $n \geq 0$, in (1.2). Similarly, by multiplying (1.2) by $\sin mx$ and then integrate over $[-\pi, \pi]$, one obtain the expression of b_n, $n \geq 1$, in (1.2) after using

$$\int_{-\pi}^{\pi} \sin nx \sin mx \, dx = \begin{cases} \pi, & n = m \\ 0, & n \neq m \end{cases} .$$

Of course, (1.2) arises from the hypothesis that every sufficiently nice function of period 2π is equal to its Fourier expansion. The study of under which "nice conditions" this could happen is one of the main objects in the theory of Fourier series.

We can associate a Fourier series for any integrable function on $[-\pi, \pi]$. As the right hand side of (1.2) consists of 2π-periodic functions, it is natural to extend its left hand side, that is, the function f itself, as a 2π-periodic function. The extension is straightforward. First of all, the real line can be expressed as the disjoint union of intervals $((2n-1)\pi, (2n+$

$1)\pi]$, $n \in \mathbb{Z}$. Each number x belongs to one and exactly one such interval. Let $\tilde{f}(x) = f(x - 2n\pi)$ where n is the unique integer satisfying $(2n-1)\pi < x \leq (2n+1)\pi$. It is clear that $\tilde{f}$ is equal to f on $(-\pi, \pi]$. As the original function is defined on $[-\pi, \pi]$, apparently an extension in strict sense is possible only if $f(-\pi) = f(\pi)$. Since the function value at one point does not change the Fourier series, from now on it will be understood that the extension of a function to a 2π-periodic function refers to the extension for the restriction of this function on $(-\pi, \pi]$. Note that for the 2π-periodic extension of a continuous function on $[-\pi, \pi]$ has a jump discontinuity at $\pm\pi$ when $f(\pi) \neq f(-\pi)$. It is is continuous on $\mathbb{R}$ if and only if $f(-\pi) = f(\pi)$. In the following we will not distinguish f with its extension $\tilde{f}$.

We will use

$$f(x) \sim a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

to denote the fact that the right hand side of this expression is the Fourier series of f. Note that in general $\sim$ cannot be replaced by $=$.

Example 1.1 We consider the function $f_1(x) = x$. Its extension is a piecewise smooth function with jump discontinuities at $(2n+1)\pi, n \in \mathbb{Z}$. As f_1 is odd and $\cos nx$ is even,

$$\pi a_n = \int_{-\pi}^{\pi} x \cos nx \, dx = 0, \quad n \geq 0,$$

and

$$\begin{aligned} \pi b_n &= \int_{-\pi}^{\pi} x \sin nx \, dx \\ &= -x \frac{\cos nx}{n} \Big|_{-\pi}^{\pi} + \int_{-\pi}^{\pi} \frac{\cos nx}{n} \, dx \\ &= (-1)^{n+1} \frac{2\pi}{n}. \end{aligned}$$

Therefore,

$$f_1(x) \sim 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nx.$$

Since f_1 is an odd function, it is reasonable to see that no cosine functions are involved in its Fourier series. How about the convergence of this Fourier series? Although the coefficients decay like $O(1/n)$ as $n \to \infty$, its convergence is not clear at this moment. On the other hand, this Fourier series is equal to 0 at $x = \pm\pi$ but $f_1(\pm\pi) = \pi$. So, one thing is sure, namely, the Fourier series is not always equal to its function. It is worthwhile to observe that the bad points $\pm\pi$ are precisely the discontinuity points of f_1.

Notation The big O and small $\circ$ notations are very convenient in analysis. We say a sequence $\{x_n\}$ satisfies $x_n = O(n^s)$ means that there exists a constant C independent of

n such that $|x_n| \le Cn^s$ as $n \to \infty$, in other words, the growth (resp. decay $s \ge 0$) of $\{x_n\}$ is not faster (resp. slower $s < 0$) the s-th power of n. On the other hand, $x_n = \circ(n^s)$ means $|x_n|n^{-s} \to 0$ as $n \to \infty$.

Example 1.2 Next consider the function $f_2(x) = x^2$. Unlike the previous example, its 2π-periodic extension is continuous on $\mathbb{R}$. After performing integration by parts, the Fourier series of f_2 is seen to be

$$f_2(x) \equiv x^2 \sim \frac{\pi^2}{3} - 4\pi \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} \cos nx.$$

As f_2 is an even function, this is a cosine series. The rate of decay of the Fourier coefficients is like $O(1/n^2)$. Using Weierstrass M-test, this series converges uniformly to a continuous function. Later we will see that this continuous function is equal to f_2, but at this stage we do not know.

We list more examples of Fourier series of functions and leave them for you to verify.

(a) $f_3(x) \equiv |x| \sim \dfrac{\pi}{2} - \dfrac{4}{\pi} \displaystyle\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cos(2n-1)x,$

(b) $f_4(x) = \begin{cases} 1, & x \in [0, \pi] \\ -1, & x \in (-\pi, 0) \end{cases} \sim \dfrac{4}{\pi} \displaystyle\sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)x,$

(c) $f_5(x) = \begin{cases} x(\pi - x), & x \in [0, \pi] \\ x(\pi + x), & x \in (-\pi, 0) \end{cases} \sim \dfrac{8}{\pi} \displaystyle\sum_{n=1}^{\infty} \frac{1}{(2n-1)^3} \sin(2n-1)x.$

Let $\{c_n\}_{-\infty}^{\infty}$ be a bisequence of complex numbers. (A bisequence is a map from $\mathbb{Z}$ to $\mathbb{C}$.) A (complex) trigonometric series is the infinite series associated to the bisequence $\{c_n e^{inx}\}_{-\infty}^{\infty}$ and is denoted by $\sum_{-\infty}^{\infty} c_n e^{inx}$. To be in line with the real case, it is said to be convergent at x if

$$\lim_{n\to\infty} \sum_{k=-n}^{n} c_n e^{nix}$$

exists. Now, a complex Fourier series can be associated to a complex-valued function. Let f be a 2π-periodic complex-valued function which is integrable on $[-\pi, \pi]$. Its Fourier series is given by the series

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx},$$

where the Fourier coefficients c_n are defined to be

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad n \in \mathbb{Z}.$$

Here for a complex function f, its integration over some $[a,b]$ is defined to be

$$\int_a^b f(x)dx = \int_a^b f_1(x)dx + i\int_a^b f_2(x)dx,$$

where f_1 and f_2 are respectively the real and imaginary parts of f. It is called integrable if both real and imaginary parts are integrable. The same as in the real case, formally the expression of c_n is obtained as in the real case by first multiplying the relation

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

with e^{imx} and then integrating over $[-\pi,\pi]$ with the help from the relation

$$\int_{-\pi}^{\pi} e^{imx}e^{-inx}\,dx = \begin{cases} 2\pi, & n = m \\ 0, & n \neq m \end{cases} .$$

When f is of real-valued, there are two Fourier series, that is, the real and the complex ones. To relate them it is enough to observe the Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$, so for $n \geq 1$

$$\begin{aligned} 2\pi c_n &= \int_{-\pi}^{\pi} f(x)e^{-inx}dx \\ &= \int_{-\pi}^{\pi} f(x)\left(\cos nx - i\sin nx\right)dx \\ &= \int_{-\pi}^{\pi} f(x)\cos nx dx - i\int_{-\pi}^{\pi} f(x)\sin nx dx \\ &= \pi(a_n - ib_n)\ . \end{aligned}$$

we see that

$$c_n = \frac{1}{2}(a_n - ib_n), \quad n \geq 1, \quad c_0 = a_0\ .$$

By a similar computation, we have

$$c_n = \frac{1}{2}(a_{-n} + ib_{-n}), \quad n \leq -1\ .$$

It follows that $c_{-n} = \overline{c_n}$ for all n. In fact, the converse is true, that is, a complex Fourier series is the Fourier series of a real-valued function if and only if $c_{-n} = \overline{c_n}$ holds for all n. Indeed, letting

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

be the Fourier series of f, it is straightforward to verify that

$$\overline{f(x)} \sim \sum_{n=-\infty}^{\infty} d_n e^{inx}, \quad d_n = \overline{c_{-n}}\ .$$

Hence when f is real-valued, $\overline{f} = f$ so $c_n = \overline{c_{-n}}$ holds. The complex form of Fourier series sometimes makes expressions and computations more elegant. We will use it whenever it makes things simpler.

We have been working on the Fourier series of 2π-periodic functions. For functions of $2T$-period, their Fourier series are not the same. They can be found by a scaling argument. Let f be $2T$-periodic. The function $g(x) = f(Tx/\pi)$ is a 2π-periodic function. Thus,

$$f\left(\frac{Tx}{\pi}\right) = g(x) \sim a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

where $a_0, a_n, b_n, n \geq 1$ are the Fourier coefficients of g. By a change of variables, we can express everything inside the coefficients in terms of f, $\cos n\pi x/T$ and $\sin n\pi x/T$. The result is

$$f(x) \sim a_0 + \sum_{n=1}^{\infty}\left(a_n \cos \frac{n\pi}{T}x + b_n \sin \frac{n\pi}{T}x\right),$$

where

$$\begin{aligned} a_n &= \frac{1}{T}\int_{-T}^{T} f(y)\cos\frac{n\pi}{T}y\,dy, \\ b_n &= \frac{1}{T}\int_{-T}^{T} f(y)\sin\frac{n\pi}{T}y\,dy, \quad n \geq 1, \qquad \text{and} \\ a_0 &= \frac{1}{2T}\int_{-T}^{T} f(y)\,dy. \end{aligned}$$

It reduces to (1.1) when T is equal to π. Can you give a reasonable definition of the Fourier series of an integrable function on $[a, b]$?

1.2 Riemann-Lebesgue Lemma

From the examples of Fourier series of functions in the previous section we see that the coefficients decay to 0 eventually. We will show that this is generally true. This is the content of the following result.

Theorem 1.1 (Riemann-Lebesgue Lemma). *For $f \in R[a, b]$,*

$$\int_a^b f(x)\cos nx\,dx, \quad \int_a^b f(x)\sin nx\,dx \to 0\ , \qquad \text{as } n \to \infty\ .$$

In particular, taking $[a, b] = [-\pi, \pi]$, we have $a_n, b_n \to 0$ as $n \to \infty$.

To prepare for the proof, we study how to approximate an integrable function by step functions. Let $a_0 = a < a_1 < \cdots < a_N = b$ be a partition of $[a, b]$. A **step function** s satisfies $s(x) = s_j$, $\forall x \in (a_j, a_{j+1}]$, $\forall j \geq 0$. The value of s at a is not important, but for definiteness let's set $s(a) = s_0$. We can express a step function in a better form by introducing the characteristic function χ_E for a set $E \subset \mathbb{R}$:

$$\chi_E = \begin{cases} 1, & x \in E, \\ 0, & x \notin E. \end{cases}$$

Then,

$$s(x) = \sum_{j=0}^{N-1} s_j \chi_{I_j}, \quad I_j = (a_j, a_{j+1}], \quad j \geq 1,\ I_0 = [a_0, a_1].$$

Lemma 1.2. *For every step function s, there exists some constant C independent of n such that*

$$\left| \int_a^b s(x) \cos nx dx \right|, \quad \left| \int_a^b s(x) \sin nx dx \right| \leq \frac{C}{n}, \quad \forall n \geq 1\ .$$

Proof. Let $s(x) = \sum_{j=0}^{N-1} s_j \chi_{I_j}$. We have

$$\begin{aligned} \int_a^b s(x) \cos nx dx &= \int_a^b \sum_{j=0}^{N-1} s_j \chi_{I_j} \cos nx\, dx \\ &= \sum_{j=0}^{N-1} s_j \int_{a_j}^{a_{j+1}} \cos nx\, dx \\ &= \frac{1}{n} \sum_{j=0}^{N-1} s_j (\sin n a_{j+1} - \sin n a_j). \end{aligned}$$

It follows that

$$\left| \int_a^b s(x) \cos nx dx \right| \leq \frac{C}{n}, \quad \forall n \geq 1, \qquad C = 2 \sum_{j=0}^{N-1} |s_j|.$$

Clearly a similar estimate holds for the other case. □

Lemma 1.3. *Let $f \in R[a, b]$. Given $\varepsilon > 0$, there exists a step function s such that $s \leq f$ on $[a, b]$ and*

$$0 \leq \int_a^b (f - s) < \varepsilon.$$

Proof. As f is integrable, it can be approximated from below by its Darboux lower sums. In other words, for $\varepsilon > 0$, we can find a partition $a = a_0 < a_1 < \cdots < a_N = b$ such that

$$0 \leq \int_a^b f - \sum_{j=0}^{N-1} m_j (a_{j+1} - a_j) < \varepsilon,$$

where $m_j = \inf\{f(x) : x \in [a_j, a_{j+1}]\}$. It follows that

$$0 \le \int_a^b (f-s) < \varepsilon$$

after setting

$$s(x) = \sum_{j=0}^{N-1} m_j \chi_{I_j}, \quad I_j = (a_j, a_{j+1}], \quad j \ge 1,\ I_0 = [a_0, a_1].$$

□

Now we prove Theorem 1.1. For $\varepsilon > 0$, we can find s as constructed in Lemma 1.3 such that $0 \le f - s$ and

$$0 \le \int_a^b (f-s) < \frac{\varepsilon}{2}.$$

By Lemma 1.2, there exists some n_0 such that

$$\left| \int_a^b s(x) \cos nx dx \right| < \frac{\varepsilon}{2},$$

for all $n \ge n_0$. Therefore,

$$\begin{aligned}
\left| \int_a^b f(x) \cos nx dx \right| &\le \left| \int_a^b (f-s) \cos nx\, dx \right| + \left| \int_a^b s(x) \cos nx dx \right| \\
&\le \int_a^b |f-s| + \frac{\varepsilon}{2} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\ .
\end{aligned}$$

The same argument applies when $\cos nx$ is replaced by $\sin nx$. The proof of Riemann-Lebesgue Lemma is completed.

It is useful to bring in a "mapping" point of view between functions and their Fourier series. Let $R_{2\pi}$ be the collection of all 2π-periodic complex-valued functions integrable on $[-\pi, \pi]$ and $\mathcal{C}$ consisting of all complex-valued bisequences $\{c_n\}$ satisfying $c_n \to 0$ as $n \to \pm\infty$. The Fourier series sets up a mapping Φ from $R_{2\pi}$ to $\mathcal{C}$ by sending f to $\{\hat{f}(n)\}$ where, to make things clear, we have let $\hat{f}(n) = c_n$, the n-th Fourier coefficient of f. When real-functions are considered, restricting to the subspace of $\mathcal{C}$ given by those satisfying $c_{-n} = \overline{c_n}$, Φ maps all real functions into this subspace. Perhaps the first question we ask is: Is Φ one-to-one? Clearly the answer is no, for two functions which differ on a set of measure zero have the same Fourier coefficients. However, we have the following result (which will be established later):

Uniqueness Theorem. *The Fourier series of two functions in $R_{2\pi}$ coincide if and only if they are equal almost everywhere.*

Thus Φ is essentially one-to-one. We may also study how various properties in $R_{2\pi}$ and $\mathcal{C}$ correspond under Φ. In fact, there are obvious and surprising ones. Some of them are listed below and more can be found in the exercise. Observe that both $R_{2\pi}$ and $\mathcal{C}$ carry the structure of a vector space over $\mathbb{C}$.

Property 1. Φ is a linear map. Observe that both $R_{2\pi}$ and $\mathcal{C}$ form vector spaces over $\mathbb{R}$ or $\mathbb{C}$. The linearity of Φ is clear from its definition.

Property 2. When $f \in R_{2\pi}$ is k-th differentiable on $\mathbb{R}$ and all derivatives up to k-th order belong to $R_{2\pi}$, $\hat{f^k}(n) = (in)^k \hat{f}(n)$ for all $n \in \mathbb{Z}$. See Proposition 1.4 below for a proof. This property shows that differentiation turns into the multiplication of a factor $(in)^k$ under Φ. This is amazing!

Property 3. Every translation in $\mathbb{R}$ induces a "translation operation" on functions defined on $\mathbb{R}$. More specifically, for $a \in \mathbb{R}$, set $f_a(x) = f(x+a)$, $x \in \mathbb{R}$. Clearly f_a belongs to $R_{2\pi}$. We have $\hat{f_a}(n) = e^{ina}\hat{f}(n)$. This property follows directly from the definition. It shows that a translation in $R_{2\pi}$ turns into the multiplication of a factor e^{ina} under Φ.

Proposition 1.4. *Let f be a 2π-periodic function which is differentiable on $[-\pi, \pi]$ with $f' \in R_{2\pi}$. If*

$$f'(x) \sim a_0' + \sum_{n=1}^{\infty} \big(a_n' \cos nx + b_n' \sin nx\big),$$

then $a_0' = 0$, $a_n' = nb_n$, and $b_n' = -na_n$. In complex notation, $\hat{f}'(n) = in\hat{f}(n)$.

Proof. We compute

$$\begin{aligned}
\pi a_n' &= \int_{-\pi}^{\pi} f'(y) \cos ny \, dy \\
&= f(y) \cos ny \big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f(y)(-n \sin ny) \, dy \\
&= n \int_{-\pi}^{\pi} f(y) \sin ny \, dy \\
&= \pi n b_n .
\end{aligned}$$

Similarly,

$$\begin{aligned}\pi b_n' &= \int_{-\pi}^{\pi} f'(y)\sin ny\,dy \\ &= f(y)\sin ny\big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f(y)n\cos ny\,dy \\ &= -n\int_{-\pi}^{\pi} f(y)\cos ny\,dy \\ &= -\pi n a_n.\end{aligned}$$

□

Property 2 links the regularity of the function to the rate of decay of its Fourier coefficients. This is an extremely important property. When f is a 2π-periodic function whose derivatives up to k-th order belong to $R_{2\pi}$, applying Riemann-Lebesgue lemma to $f^{(k)}$ we know that $\widehat{f^{(k)}}(n) = \circ(1)$ as $n \to \infty$. By Property 2 it follows that $\hat{f}(n) = \circ(n^{-k})$, that is, the Fourier coefficients of f decay faster that n^{-k}. Since $\sum_{n=1}^{\infty} n^{-2} < \infty$, an application of Weierstrass M-test establishes the following result: The Fourier series of f converges uniformly provided f, f' and f'' belong to $R_{2\pi}$. Therefore, the function

$$g(x) \equiv a_0 + \sum_{n=1}^{\infty}(a_n\cos nx + b_n\sin nx)$$

is a continuous 2π-periodic function. Using its uniform convergence, we see that the Fourier coefficients of g are given by a_n and b_n, the same as f. By the Uniqueness Theorem stated earlier we conclude that g is equal to f, that is, the Fourier series of f is equal to f provided $f, f', f'' \in R_{2\pi}$. A more general result will be proved in the next section.

1.3 Convergence of Fourier Series

In this section we study the convergence of the Fourier series of a function to the function itself. Recall that the series $a_0 + \sum_{n=1}^{\infty}(a_n\cos nx + b_n\sin nx)$, or $\sum_{n=-\infty}^{\infty} c_n e^{inx}$, where a_n, b_n, c_n are the Fourier coefficients of a function f converges to f at x means that the n-th partial sum

$$(S_n f)(x) = a_0 + \sum_{k=1}^{n}(a_k\cos kx + b_k\sin kx)$$

or

$$(S_n f)(x) = \sum_{k=-n}^{n} c_k e^{ikx}$$

converges to $f(x)$ as $n \to \infty$.

We start by expressing the partial sums in closed form. Indeed,

$$\begin{aligned}
(S_n f)(x) &= a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f + \sum_{k=1}^{n} \frac{1}{\pi} \int_{-\pi}^{\pi} f(y)(\cos ky \cos kx + \sin ky \sin ky)\ dy \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \Big(\frac{1}{2} + \sum_{k=1}^{n} \cos k(y-x) \Big) f(y)\ dy \\
&= \frac{1}{\pi} \int_{x-\pi}^{x+\pi} \Big(\frac{1}{2} + \sum_{k=1}^{n} \cos kz \Big) f(x+z)\ dz \\
&= \frac{1}{\pi} \int_{-\pi}^{\pi} \Big(\frac{1}{2} + \sum_{k=1}^{n} \cos kz \Big) f(x+z)\ dz\ ,
\end{aligned}$$

where in the last step we have used the fact that the integrals over any two periods are the same. Using the elementary formula

$$\cos\theta + \cos 2\theta + \cdots + \cos n\theta = \frac{\sin\left(n+\frac{1}{2}\right)\theta - \sin\frac{1}{2}\theta}{2\sin\frac{\theta}{2}}, \quad \theta \neq 0,$$

we obtain

$$\frac{1}{2} + \sum_{k=1}^{n} \cos k\theta = \frac{\sin(n+\frac{1}{2})\theta}{2\sin\frac{\theta}{2}}.$$

Noting that by the L'Hospital Rule,

$$\lim_{\theta\to 0} \frac{\sin(n+\frac{1}{2})\theta}{2\sin\frac{\theta}{2}} = \frac{2n+1}{2}\ ,$$

we introduce the **Dirichlet kernel** D_n by

$$D_n(z) = \begin{cases} \dfrac{\sin\left(n+\frac{1}{2}\right)z}{2\pi\sin\frac{1}{2}z}, & z \neq 0 \\ \dfrac{2n+1}{2\pi}, & z = 0. \end{cases}$$

(In fact, there are infinitely many Dirichlet kernels indexed by n, but usually people refer them as one.) It is a continuous, 2π-periodic function. We have successfully expressed the partial sums of the Fourier series in the following closed form:

$$(S_n f)(x) = \int_{-\pi}^{\pi} D_n(z) f(x+z)\, dy,$$

Taking $f \equiv 1$, we have $S_n f = 1$ for all n. Hence

$$1 = \int_{-\pi}^{\pi} D_n(z)\, dz.$$

We have arrived at the fundamental relation

$$(S_n f)(x) - f(x) = \int_{-\pi}^{\pi} D_n(z)(f(x+z) - f(x))\, dz. \tag{1.3}$$

In order to show $S_n f(x) \to f(x)$, it suffices to show the right hand side of (1.3) tends to 0 as $n \to \infty$.

The Dirichlet kernel plays a crucial role in the study of the convergence of Fourier series. We list some of its properties as follows.

Property I. $D_n(z)$ is an even, continuous, 2π-periodic function vanishing at $z = 2k\pi/(2n+1), -n \le k \le n$, on $[-\pi, \pi]$.

Property II. D_n attains its maximum value $(2n+1)/2\pi$ at 0.

Property III.

$$\int_{-\pi}^{\pi} D_n(z)dz = 1 \ .$$

Property IV. For every $\delta > 0$,

$$\int_0^\delta |D_n(z)|dz \to \infty, \qquad \text{as } n \to \infty.$$

Only the last property needs a proof. Indeed, for each n we can fix an N such that $\pi N < (2n+1)\delta/2 \le (N+1)\pi$, so $N \to \infty$ as $n \to \infty$. We compute

$$\begin{aligned}
\int_0^\delta |D_n(z)|dz &= \int_0^\delta \frac{|\sin(n+\frac{1}{2})z|}{2\pi|\sin\frac{z}{2}|}\, dz \\
&\ge \frac{1}{\pi}\int_0^{(n+\frac{1}{2})\delta} \frac{|\sin t|}{t}\, dt \\
&\ge \frac{1}{\pi}\int_0^{N\pi} \frac{|\sin t|}{t} dt \\
&= \frac{1}{\pi}\sum_{k=1}^N \int_{(k-1)\pi}^{k\pi} \frac{|\sin t|}{t}\, dt \\
&= \frac{1}{\pi}\sum_{k=1}^N \int_0^\pi \frac{|\sin s|}{s+(k-1)\pi}\, ds \\
&\ge \frac{1}{\pi}\sum_{k=1}^N \int_0^\pi \frac{|\sin s|}{\pi k}\, ds \\
&= \frac{2}{\pi^2}\sum_{k=1}^N \frac{1}{k}, \quad \text{as } \int_0^\pi |\sin s|\, ds = 2, \\
&\to \infty,
\end{aligned}$$

as $N \to \infty$.

To elucidate the effect of the kernel, we fix a small $\delta > 0$ and split the integral into two parts:

$$\int_{-\pi}^{\pi} \chi_A(z) D_n(z)(f(x+z) - f(x))\, dz,$$

and

$$\int_{-\pi}^{\pi} \chi_B(z) D_n(z)(f(x+z) - f(x))\, dz,$$

where $A = (-\delta, \delta)$ and $B = [-\pi, \pi] \setminus A$. The second integral can be written as

$$\int_{-\pi}^{\pi} \frac{\chi_B(z)(f(x+z) - f(x))}{2\pi \sin \frac{z}{2}} \sin(n + 1/2) z\, dz.$$

As $|\sin z/2|$ has a positive lower bound on B, the function

$$\frac{\chi_B(z)(f(x+z) - f(x))}{2\pi \sin \frac{z}{2}}$$

belongs to $R[-\pi, \pi]$ and the second integral tends to 0 as $n \to \infty$ in view of Riemann-Lebesgue lemma. The trouble lies on the first integral. It can be estimated by

$$\int_{-\delta}^{\delta} |D_n(z)||f(x+z) - f(x)| dz.$$

In view of Property IV, No matter how small δ is, this term may go to ∞ so it is not clear how to estimate this integral.

The difficulty can be resolved by imposing a further regularity assumption on the function. First a definition. A function f defined on $[a, b]$ is called **Lipschitz continuous** at $x \in [a, b]$ if there exist L and δ such that

$$|f(y) - f(x)| \le L\,|y - x|\,, \quad \forall y \in [a, b],\ |y - x| \le \delta. \tag{1.4}$$

Here both L and δ depend on x. Incidentally, we point out that if $f \in C[a, b]$ is Lipschitz continuous at x, there exists some L' such that

$$|f(y) - f(x)| \le L'|y - x|, \quad \forall y \in [a, b].$$

In fact, this comes from (2.4) if $|y - x| \le \delta$. For y satisfying $|y - x| > \delta$, we have

$$|f(y) - f(x)| \le \frac{|f(y)| + |f(x)|}{\delta} |y - x|,$$

hence we could take

$$L' = \max \left\{ L, \frac{2M}{\delta} \right\}\ ,$$

where $M = \sup\{|f(y)| :\ y \in [a, b]\}$.

Theorem 1.5. *Let f be a 2π-periodic function integrable on $[-\pi,\pi]$. Suppose that f is Lipschitz continuous at x. Then $\{S_n f(x)\}$ converges to $f(x)$ as $n \to \infty$.*

Note that we have identified f with its 2π-periodic extension $\tilde{f}$. When $x = \pm\pi$, f is Lipschitz continuous at x means $\tilde{f}$ is Lipschitz continuous at x.

Proof. Let Φ_δ be a cut-off function satisfying (a) $\Phi_\delta \in C(\mathbb{R})$, $\Phi_\delta \equiv 0$ outside $(-\delta,\delta)$, (b) $0 \le \Phi_\delta \le 1$ and (c) $\Phi_\delta = 1$ on $(-\delta/2, \delta/2)$. We write

$$\begin{aligned}
(S_n f)(x) - f(x) &= \int_{-\pi}^{\pi} D_n(z)(f(x+z) - f(x))\,dz \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin(n+\frac{1}{2})z}{\sin\frac{z}{2}}(f(x+z) - f(x))\,dz \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_\delta(z)\frac{\sin(n+\frac{1}{2})z}{\sin\frac{z}{2}}(f(x+z) - f(x))\,dz \\
&\quad + \frac{1}{\pi}\int_{-\pi}^{\pi} (1-\Phi_\delta(z))\frac{\sin(n+\frac{1}{2})z}{\sin\frac{z}{2}}(f(x+z) - f(x))\,dz \\
&\equiv I + II\ .
\end{aligned}$$

By our assumption on f, there exists $\delta_0 > 0$ such that

$$|f(x+z) - f(x)| \le L\,|z|\,, \quad \forall\,|z| < \delta_0.$$

(If you use the discussion following (1.4), this δ_0 is not needed.) Using $\sin\theta/\theta \to 1$ as $\theta \to 0$, there exists δ_1 such that $2|\sin z/2| \ge |z/2|$ for all $z, |z| < \delta_1$. For $z, |z| < \delta \equiv \min\{\delta_0, \delta_1\}$, we have $|f(x+z) - f(x)|/|\sin z/2| \le 4L$ and

$$\begin{aligned}
|I| &\le \frac{1}{2\pi}\int_{-\delta}^{\delta} \Phi_\delta(z)\frac{\left|\sin(n+\frac{1}{2})z\right|}{\left|\sin\frac{z}{2}\right|}\,|f(x+z) - f(x)|\,dz \\
&\le \frac{1}{2\pi}\int_{-\delta}^{\delta} 4L\,dz \\
&= \frac{4\delta L}{\pi}.
\end{aligned} \tag{1.5}$$

For $\varepsilon > 0$, we further restrict and fix one δ so that

$$\frac{4\delta L}{\pi} < \frac{\varepsilon}{2}. \tag{1.6}$$

After fixing δ, we turn to the second integral

$$\begin{aligned}
II &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{(1-\Phi_\delta(z))(f(x+z) - f(x))}{\sin\frac{z}{2}}\sin\left(n+\frac{1}{2}\right)z\,dz \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{(1-\Phi_\delta(z))(f(x+z) - f(x))}{\sin\frac{z}{2}}\left(\cos\frac{z}{2}\sin nz + \sin\frac{z}{2}\cos nz\right)\,dz \\
&\equiv \int_{-\pi}^{\pi} F_1(x,z)\sin nz\,dz + \int_J F_2(x,z)\cos nz\,dz,
\end{aligned}$$

where

$$F_1(x,z) = \frac{1}{2\pi}\frac{(1-\Phi_\delta(z))(f(x+z)-f(x))}{\sin\frac{z}{2}}\cos\frac{z}{2}\ ,$$

and

$$F_2(x,z) = \frac{1}{2\pi}\frac{(1-\Phi_\delta(z))(f(x+z)-f(x))}{\sin\frac{z}{2}}\sin\frac{z}{2}\ .$$

As $1-\Phi_\delta(z) = 0$ on $[-\delta/2,\delta/2]$, these two functions vanish outside the two intervals $[-\pi,-\delta/2]$ and $[\delta/2,\pi]$. Now $|\sin z/2|$ has a positive lower bound on these two intervals, so F_1 and F_2 are integrable on $[-\pi,\pi]$. By Riemann-Lebesgue Lemma, for $\varepsilon > 0$, there is some n_0 such that

$$\left|\int_{-\pi}^{\pi} F_1 \sin nz\, dz\right|,\ \left|\int_{-\pi}^{\pi} F_2 \cos nz\, dz\right| < \frac{\varepsilon}{4}, \quad \forall n \geq n_0. \tag{1.7}$$

Putting (1.5), (1.6) and (1.7) together,

$$|S_n f(x) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon, \quad \forall n \geq n_0.$$

We have shown that $S_n f(x)$ tends to $f(x)$ whenever f is Lipschitz continuous at x. □

We leave some remarks concerning this proof. First, the cut-off function Φ_δ can be replaced by $\chi_{[-\delta,\delta]}$ without affecting the proof. (However, it will be needed in the proof of Theorem 1.7.) Second, the regularity condition Lipschitz continuity is used to kill off the growth of the kernel at x. Third, this method used in this proof is a standard one. It will appear in many other places.

A careful examination of it reveals a convergence result for functions with jump discontinuity after using the evenness of the Dirichlet kernel.

Theorem 1.6. *Let f be a 2π-periodic function integrable on $[-\pi,\pi]$. Suppose at some $x \in [-\pi,\pi]$, $\lim_{y\to x_+} f(y)$ and $\lim_{y\to x_-} f(y)$ exist and there are $\delta > 0$ and constant L such that*

$$|f(y) - f(x^+)| \leq L(y-x), \quad \forall y, \quad 0 < y - x < \delta,$$

and

$$|f(y) - f(x^-)| \leq L(x-y), \quad \forall y, \quad 0 < x - y < \delta.$$

Then $\{S_n f(x)\}$ converges to $(f(x^+)+f(x^-))/2$ as $n \to \infty$.

Again, note that we have identified f with $\tilde{f}$. Here $f(x^+)$ and $f(x^-)$ stand for $\lim_{y\to x^+} f(y)$ and $\lim_{y\to x^-} f(y)$ respectively. We leave the proof of this theorem as an exercise.

A function f defined on $[a,b]$ is called to satisfy the **Lipschitz condition** if there exists an L such that

$$|f(x) - f(y)| \leq L|x-y|\ , \quad \forall x,y \in [a,b]\ .$$

(In some books this condition is called “a Lipschitz condition”. Frankly speaking, I don't know the difference.) When f satisfies the Lipschitz condition, it is Lipschitz continuous everywhere. It is better to call this condition uniformly Lipschitz continuous. Every continuously differentiable function on $[a, b]$ satisfies the Lipschitz condition. In fact, by the fundamental theorem of calculus, for $x, y \in [a, b]$,

$$\begin{aligned} |f(y) - f(x)| &= \Big| \int_x^y f'(t)dt \Big| \\ &\leq M|y - x|, \end{aligned}$$

where $M = \sup\{|f'(t)| : t \in [a, b]\}$. Similarly, every continuous, piecewise C^1-function satisfies the Lipschitz condition, see exercise.

Now, we have a theorem on the uniform convergence of the Fourier series of a function to the function itself.

Theorem 1.7. *Let f a 2π-periodic function satisfying the Lipschitz condition on $\mathbb{R}$. Its Fourier series converges to f uniformly as $n \to \infty$.*

Proof. Observe that when f is Lipschitz continuous on $\mathbb{R}$, δ_0 and δ_1 can be chosen independent of x and (1.5), (1.6) holds uniformly in x. In fact, δ_0 only depends on L, the constant appearing in the Lipschitz condition. Thus the theorem follows if n_0 in (1.7) can be chosen uniformly in x. This is the content of the lemma below. We apply it by taking $F(x, y)$ to be $F_1(x, z)$ or $F_2(x, z)$ with y replaced by z.

□

Lemma 1.8. *Let $F(x, y)$ be periodic in y and $F \in C([-\pi, \pi] \times [-\pi, \pi])$. For any fixed x,*

$$g(n, x) \equiv \frac{1}{2\pi} \int_{-\pi}^{\pi} F(x, y)e^{-iny}\, dy \to 0$$

uniformly in x as $n \to \infty$.

Proof. We need to show that for every $\varepsilon > 0$, there exists some n_0 independent of x such that

$$|g(n, x)| < \varepsilon, \quad \forall n \geq n_0.$$

Observe that

$$\begin{aligned} g(n, x) &= \int_{-\pi}^{\pi} F(x, y)e^{-iny}\, dy \\ &= \int_{-\pi-\frac{\pi}{n}}^{\pi-\frac{\pi}{n}} F\left(x, z + \frac{\pi}{n}\right) e^{-in(z+\frac{\pi}{n})}\, dz \quad y = z + \frac{\pi}{n}, \\ &= -\int_{-\pi}^{\pi} F\left(x, z + \frac{\pi}{n}\right) e^{-inz}\, dz \quad (F \text{ is } 2\pi\text{-periodic}). \end{aligned}$$

We have

$$g(n,x) = \frac{1}{2}\int_{-\pi}^{\pi} \left(F(x,y) - F\left(x, y+\frac{\pi}{n}\right)\right) e^{-iny}\,dy.$$

As $F \in C([-\pi,\pi]\times[-\pi,\pi])$, it is uniformly continuous in $[-\pi,\pi]\times[-\pi,\pi]$. For $\varepsilon > 0$, there exists a δ such that

$$|F(x,y) - F(x',y')| < \frac{\varepsilon}{\pi} \quad \text{if } |x-x'|\,,\ |y-y'| < \delta.$$

We take n_0 so large that $\pi/n_0 < \delta$. Then, using $|e^{-iny}| = 1$,

$$\begin{aligned} |g(n,x)| &\le \frac{1}{2}\int_{-\pi}^{\pi} \left|F(x,y) - F\left(x, y+\frac{\pi}{n}\right)\right|\,dy \\ &< \frac{\varepsilon}{2\pi}\int_{-\pi}^{\pi} dy \\ &= \varepsilon, \qquad \forall n \ge n_0. \end{aligned}$$

□

Example 1.3. We return to the functions discussed in Examples 1.1 and 1.2. Indeed, $f_1(x) = x$ is smooth except at $n\pi$. According to Theorem 1.5, the series

$$2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nx$$

converges to x for every $x \in (-\pi,\pi)$. On the other hand, we observed before that the series tend to 0 at $x = \pm\pi$. As $f_1(\pi_+) = -\pi$ and $f(\pi_-) = \pi$, we have $f_1(\pi_+) + f(\pi_-) = 0$, which is in consistency with Theorem 1.5. In the second example, $f_2(x) = x^2$ is continuous, 2π-periodic. By Theorem 1.7, its Fourier series

$$\frac{\pi^2}{3} - 4\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} \cos nx$$

converges to x^2 uniformly on $[-\pi,\pi]$.

So far we have been working on the Fourier series of 2π-periodic functions. It is clear that the same results apply to the Fourier series of $2T$-periodic functions for arbitrary positive T.

We have shown the convergence of the Fourier series under some additional regularity assumptions on the function. But the basic question remains, that is, is the Fourier series of a continuous, 2π-periodic function converges to itself? It turns out the answer is negative. A not-so-explicit example can be found in Stein-Shakarchi and an explicit but complicated one was given by Fejér (see Zygmund "Trigonometric Series"). You may google for more. In fact, using the uniform boundedness principle in functional analysis, one can even show that "most" continuous functions have divergent Fourier series. The situation is very much like in the case of the real number system where transcendental numbers are uncountable while algebraic numbers are countable despite the fact that it is difficult to establish a concrete number is transcendental.

1.4 Weierstrass Approximation Theorem

As an application of the uniform convergence theorem of the last section, we now prove a theorem of Weierstrass concerning the approximation of continuous functions by polynomials. First we consider how to approximate a continuous function by continuous, piecewise linear functions. A continuous function on $[a, b]$ is **piecewise linear** if there exists a partition $a = a_0 < a_1 < \cdots < a_n = b$ such that f is linear on each subinterval $[a_j, a_{j+1}]$.

Proposition 1.9. *Let f be a continuous function on $[a, b]$. For every $\varepsilon > 0$, there exists a continuous, piecewise linear function g such that $\|f - g\|_\infty < \varepsilon$.*

Recall that $\|f - g\|_\infty = \sup\{|f(x) - g(x)| : \ x \in [a, b]\}$.

Proof. As f is uniformly continuous on $[a, b]$, for every $\varepsilon > 0$, there exists some δ such that $|f(x) - f(y)| < \varepsilon/2$ for $x, y \in [a, b], |x - y| < \delta$. We partition $[a, b]$ into subintervals $I_j = [a_j, a_{j+1}]$ whose length is less than δ and define g to be the piecewise linear function satisfying $g(a_j) = f(a_j)$ for all j. For $x \in [a_j, a_{j+1}]$, g is given by

$$g(x) = \frac{f(a_{j+1}) - f(a_j)}{a_{j+1} - a_j}(x - a_j) + f(a_j).$$

We have

$$\begin{aligned}
|f(x) - g(x)| &= |f(x) - \frac{f(a_{j+1}) - f(a_j)}{a_{j+1} - a_j}(x - a_j) + f(a_j)| \\
&\le |f(x) - f(a_j)| + |\frac{f(a_{j+1}) - f(a_j)}{a_{j+1} - a_j}(x - a_j)| \\
&\le |f(x) - f(a_j)| + |f(a_{j+1}) - f(a_j)| \\
&< \varepsilon,
\end{aligned}$$

and the result follows. □

Next we study how to approximate a continuous function by trigonometric polynomials (or, equivalently, finite Fourier series).

Proposition 1.10. *Let f be a continuous function on $[0, \pi]$. For $\varepsilon > 0$, there exists a trigonometric polynomial h such that $\|f - h\|_\infty < \varepsilon$.*

Proof. First we extend f to $[-\pi, \pi]$ by setting $f(x) = f(-x)$ (using the same notation) to obtain a continuous function on $[-\pi, \pi]$ with $f(-\pi) = f(\pi)$. By the previous proposition, we can find a continuous, piecewise linear function g such that $\|f - g\|_\infty < \varepsilon/2$. Since $g(-\pi) = f(-\pi) = f(\pi) = g(\pi)$, g can be extended as the Lipschitz continuous, 2π-periodic function. By Theorem 1.7, there exists some N such that $\|g - S_N g\|_\infty < \varepsilon/2$. Therefore, $\|f - S_N g\|_\infty \le \|f - g\|_\infty + \|g - S_N g\|_\infty < \varepsilon/2 + \varepsilon/2 = \varepsilon$. The proposition follows after noting that every finite Fourier series is a trigonometric polynomial (see Exercise). $\square$

Theorem 1.11 (**Weierstrass Approximation Theorem**). *Let $f \in C[a, b]$. Given $\varepsilon > 0$, there exists a polynomial p such that $\|f - p\|_\infty < \varepsilon$.*

Proof. Consider $[a, b] = [0, \pi]$ first. Extend f to $[-\pi, \pi]$ by reflection as before and, for $\varepsilon > 0$, fix a trigonometric polynomial h such that $\|f - h\|_\infty < \varepsilon/2$. This is possible due to the previous proposition. Now, we express h as a finite Fourier series $a_0 + \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx)$. Using the fact that

$$\cos\theta = \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n)!}, \quad \text{and } \sin\theta = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}\theta^{2n-1}}{(2n-1)!},$$

where the convergence is uniform on $[-\pi, \pi]$, each $\cos nx$ and $\sin nx, n = 1, \cdots, N$, can be approximated by polynomials. Putting all these polynomials together we obtain a polynomial $p(x)$ satisfying $\|h - p\|_\infty < \varepsilon/2$. It follows that $\|f - p\|_\infty \le \|f - h\|_\infty + \|h - p\|_\infty < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

When f is continuous on $[a, b]$, the function $\varphi(t) = f(\frac{b-a}{\pi}t + a)$ is continuous on $[0, \pi]$. From the last paragraph, we can find a polynomial $p(t)$ such that $\|\varphi - p\|_\infty < \varepsilon$ on $[0, \pi]$. But then the polynomial $q(x) = p(\frac{\pi}{b-a}(x - a))$ satisfies $\|f - q\|_\infty = \|\varphi - p\|_\infty < \varepsilon$ on $[a, b]$. $\square$

1.5 Mean Convergence of Fourier Series*

This section, which will be covered in MATH3093, is for optional reading.

In Section 2 we studied the uniform convergence of Fourier series. Since the limit of a uniformly convergent series of continuous functions is again continuous, we do not expect results like Theorem 1.6 applies to functions with jumps. In this section we will measure the distance between functions by a norm weaker than the uniform norm. Under the new L^2-distance, you will see that every integrable function is equal to its Fourier expansion almost everywhere.

Recall that there is an inner product defined on the n-dimensional Euclidean space called the Euclidean metric

$$\langle x, y\rangle_2 = \sum_{j=1}^{n} x_j y_j, \qquad x, y \in \mathbb{R}^n.$$

With this inner product, one can define the concept of orthogonality and angle between two vectors. Likewise, we can also introduce a similar product on the space of integrable functions. Specifically, for $f, g \in R[-\pi, \pi]$, the L^2**-product** is given by

$$\langle f, g\rangle_2 = \int_{-\pi}^{\pi} f(x)g(x)\,dx.$$

The L^2-product behaves like the Euclidean metric on $\mathbb{R}^n$ except at one aspect, namely, the condition $\langle f, f\rangle_2 = 0$ does not imply $f \equiv 0$. This is easy to see. In fact, when f is equal to zero except at finitely many points, then $\langle f, f\rangle_2 = 0$. From Appendix II $\langle f, f\rangle_2 = 0$ if and only if f is equal to zero except on a set of measure zero. This minor difference with the Euclidean inner product will not affect our discussion much, except more caution is needed when we proceed. Parallel to the Euclidean case, we define the L^2-norm of an integrable function f to be

$$\|f\|_2 = \sqrt{\langle f, f\rangle_2},$$

and the L^2**-distance** between two integrable functions f and g by $\|f - g\|_2$. (When f, g are complex-valued, one should define the L^2-product to be

$$\langle f, g\rangle_2 = \int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx\ ,$$

so that $\langle f, f\rangle_2 \geq 0$. We will be restricted to real functions in this section.) One can verify that the triangle inequality

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$$

holds. We can also talk about $f_n \to f$ in L^2-sense, i.e., $\|f_n - f\|_2 \to 0$, or equivalently,

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |f_n - f|^2 = 0, \qquad \text{as } n \to \infty\ .$$

This is a convergence in an average sense. It is not hard to see that when $\{f_n\}$ tends to f uniformly, $\{f_n\}$ must tend to f in L^2-sense. In fact, we have the inequality

$$\|f\|_2^2 \leq 2\pi\|f\|_\infty^2\ ,$$

which means

$$\|f - g\|_2 \leq \sqrt{2\pi}\|f - g\|_\infty\ ,$$

so uniform convergence is stronger than L^2-convergence. A moment's reflection will show that the converse is not always true. Hence convergence in L^2-sense is weaker than uniform convergence. We will discuss various metrics and norms in Chapter 2.

Our aim in this section is to show that the Fourier series of every integrable function converges to the function in the L^2-sense.

Just like the canonical basis $\{e_1, \ldots, e_n\}$ in $\mathbb{R}^n$, the functions

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin nx \right\}_{n=1}^{\infty}$$

forms an "orthonormal basis" in $R[-\pi, \pi]$, see Section 1.1. In the following we denote by

$$E_n = \left\langle \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin nx \right\rangle_{j=1}^{n}$$

the $(2n+1)$-dimensional vector space spanned by the first $2n+1$ trigonometric functions.

We start by considering the general situation. Let $\{\phi_n\}_{n=1}^{\infty}$ be an orthonormal set (not neccessarily a basis) in $R[a, b]$, i.e.,

$$\int_a^b \phi_n \phi_m = \delta_{nm}, \quad \forall n, m \geq 1.$$

Let

$$\mathcal{S}_n = \langle \phi_1, \ldots, \phi_n \rangle$$

be the n-dimensional subspace spanned by $\phi_1, \ldots, \phi_n$. For a general $f \in R[a, b]$, we consider the minimization problem

$$\inf \left\{ \|f - g\|_2 : g \in \mathcal{S}_n \right\}. \tag{1.10}$$

From a geometric point of view, this infimum gives the L^2-distance from f to the finite dimensional subspace $\mathcal{S}_n$.

Proposition 1.12. *Let $f \in R[a, b]$. The followings hold:*

(a)

$$\|f - h\|_2 \leq \|f - g\|_2 \quad \forall g \in \mathcal{S}_n \ ,$$

where $h = \sum_{j=1}^n \alpha_j \phi_j$, $\alpha_j = \langle f, \phi_j \rangle$, and equality holds only if $g = h$.

(b)

$$\langle f, h \rangle = \|h\|_2^2 \ .$$

Proof. To minimize $\|f - g\|_2$ is the same as to minimize $\|f - g\|_2^2$. Every g in $\mathcal{S}_n$ can be written as $g = \sum_{j=1}^n \beta_j \phi_j$, $\beta_j \in \mathbb{R}$. We have

$$\begin{aligned} \|f - g\|_2^2 &= \int_{-\pi}^{\pi} \Big(f - \sum_{j=1}^n \beta_j \phi_j \Big)^2 \\ &= \int_{-\pi}^{\pi} f^2 - 2 \sum_{j=1}^n \beta_j \alpha_j + \sum_{j=1}^n \beta_j^2. \end{aligned}$$

When $g = h$, we have

$$\|f - h\|_2^2 = \int_{-\pi}^{\pi} f^2 - \sum_{j=1}^{n} \alpha_j^2 \ .$$

Therefore,

$$\|f - h\|_2^2 \le \|f - g\|_2^2$$

is the same as

$$\int_{-\pi}^{\pi} f^2 - \sum_{j=1}^{n} \alpha_j^2 \le \int_{-\pi}^{\pi} f^2 - 2\sum_{j=1}^{n} \beta_j \alpha_j + \sum_{j=1}^{n} \beta_j^2 \ .$$

But this follows readily from the inequality

$$\sum_{j=1}^{n} (\beta_j - \alpha_j)^2 \ge 0 \ .$$

It is also clear that the equality holds if and only if $\beta_j = \alpha_j$ for all j, that is, $g = h$. (a) is established.

To prove (b), we note that $th \in \mathcal{S}_n$ for all $t \in \mathbb{R}$ and the function $\lambda(t) \equiv \|f - th\|_2^2$ attains its minimum at $t = 1$. Therefore,

$$0 = \lambda'(1) = -2\langle f, h\rangle + 2\|h\|_2^2 \ .$$

□

Given an orthonormal set $\{\phi_n\}_{n=1}^{\infty}$, one may define the "Fourier series" of an L^2-function f with respect to the orthnormal set $\{\phi_n\}$ to be the series $\sum_{n=1}^{\infty} \langle f, \phi_n\rangle \, \phi_n$ and set $P_n f = \sum_{j=1}^{n} \langle f, \phi_j\rangle \, \phi_j$. Proposition 1.14 asserts that the distance between f and $\mathcal{S}_n$ is equal to $\|f - P_n f\|_2$. The function $P_n f$ is sometimes called the orthogonal projection of f on $\mathcal{S}_n$. Indeed, one can verify that it satisfies

$$\langle f - P_n f, g\rangle_2 = 0 \ , \quad \forall g \in \mathcal{S}_n \ ,$$

so $f - P_n f$ is orthogonal to $\mathcal{S}_n$. (Indeed, this inequality comes from Proposition 1.14 and $\mu'(0) = 0$ where $\mu(t) \equiv \|P_n f - f + tg)\|_2^2$.)

As a special case, taking $\{\phi_n\} = \left\{1/\sqrt{2\pi}, \cos nx/\sqrt{\pi}, \sin nx/\sqrt{\pi}\right\}$ and $\mathcal{S}_{2n+1} = E_n$, a direct computation shows that $P_{2n+1} f = S_n f$, where $S_n f$ is the n-th partial sum of the Fourier series of f. Thus we can rewrite Proposition 1.14 in this special case as

Corollary 1.13. *For $f \in R_{2\pi}$, for each $n \ge 1$,*

$$\|f - S_n f\|_2 \le \|f - g\|_2 \ ,$$

and

$$\langle f, S_n f\rangle = \|S_n f\|_2^2 \ ,$$

for all g of the form

$$g = c_0 + \sum_{k=1}^{n} (c_j \cos kx + d_j \sin kx), \quad c_0, c_k, d_k \in \mathbb{R}.$$

Here is the main result of this section.

Theorem 1.14. *For every $f \in R_{2\pi}$,*

$$\lim_{n\to\infty} \|S_n f - f\|_2 = 0.$$

Proof. Let $f \in R[-\pi, \pi]$. We further assume $f \geq 0$. For $\varepsilon > 0$, we can find a step function $s \geq 0$ such that $s \leq f$ and $\int_{-\pi}^{\pi}(f - s) < \varepsilon^2/16M$ where $M = \sup_x f(x)$. Then

$$\|f - s\|_2 \leq \sqrt{M \int_{-\pi}^{\pi} (f - s)} = \frac{\varepsilon}{4} \ .$$

Next we modify s near its points of discontinuity to get a continuous, piecewise linear function g_1 satisfying

$$\|s - g_1\|_2 < \frac{\varepsilon}{4}.$$

In case $g_1(\pi) \neq g_1(-\pi)$, we modify this function near π to get a new, piecewise linear function g satisfying $g(\pi) = g(-\pi)$ and

$$\|g - g_1\|_2 < \frac{\varepsilon}{4} \ .$$

Now g is a continuous, piecewise linear (hence piecewise C^1-), 2π-periodic function. Appealing to Theorem 1.7, we can find some n_1 such that

$$\|g - S_n g\|_\infty < \frac{\varepsilon}{4\sqrt{2\pi}} \ , \quad \forall n \geq n_1 \ .$$

It implies

$$\|g - S_n g\|_2 \leq \sqrt{2\pi}\|g - S_n g\|_\infty < \frac{\varepsilon}{4} \ .$$

Putting things together, we have, for all $n \geq n_1$,

$$\begin{aligned}
\|f - S_n f\|_2 &\leq \|f - S_n g\|_2 \quad \text{(by Corollary 1.15)} \\
&\leq \|f - s\|_2 + \|s - g_1\|_2 + \|g_1 - g\|_2 + \|g - S_n g\|_2 \\
&< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} \\
&= \varepsilon \ .
\end{aligned}$$

We have proved the theorem for non-negative functions. In the general case, we use the relation $f = f^+ - f^-$ and the triangle inequality to get

$$\begin{aligned} \|f - S_n f\|_2 &= \|f^+ - f^- - S_n f^+ + S_n f^-\|_2 \\ &\le \|f^+ - S_n f^+\|_2 + \|f^- - S_n f^-\|_2 \ . \end{aligned}$$

□

Note that the use of Corollary 1.15 is the key to the proof in this theorem. As an application we have the following result concerning the uniqueness of the Fourier expansion.

Corollary 1.15. *(a) Suppose that f_1 and f_2 in $R_{2\pi}$ have the same Fourier series. Then f_1 and f_2 are equal almost everywhere.*

(b) Suppose that f_1 and f_2 in $C_{2\pi}$ have the same Fourier series. Then f_1 is equal to f_2 everywhere.

Proof. Let $f = f_2 - f_1$. The Fourier coefficients of f all vanish, hence $S_n f = 0$, for all n. By Theorem 1.16, $\|f\|_2 = \lim_{n\to\infty} \|f - S_n f\|_2 = 0$. From Appendix II we know that f^2, hence f, must vanish almost everywhere. In other words, f_2 is equal to f_1 almost everywhere. (a) holds. To prove (b), letting f be continuous and assuming that it is not equal to zero at some x_0, by continuity it is non-zero for all points near x_0. We can find some small $\delta > 0$ such that $f^2(x) \ge f^2(x_0)/2$ for all $x \in (x_0 - \delta, x_0 + \delta)$. But then

$$\begin{aligned} \int_{-\pi}^{\pi} f^2 &\ge \int_{x_0-\delta}^{x_0+\delta} f^2 \\ &\ge \frac{f^2(x_0)}{2} \times 2\delta > 0 \ , \end{aligned}$$

contradicting $\|f\|_2 = 0$. Hence f must vanish identically. □

Another interesting consequence of Theorem 1.16 is the Parseval's identity.

Corollary 1.16 (Parseval's Identity). *For every $f \in R_{2\pi}$,*

$$\|f\|_2^2 = 2\pi a_0^2 + \pi \sum_{n=1}^{\infty} \left(a_n^2 + b_n^2\right),$$

where a_n and b_n are the Fourier coefficients of f.

Proof. Making use of Corollary 1.15 and the relations such as $\langle f, \cos nx/\sqrt{\pi}\rangle_2 = \sqrt{\pi}a_n, n \geq 1$,

$$\begin{aligned}\langle f, S_n f\rangle_2 &= \|S_n f\|_2^2 \\ &= 2\pi a_0^2 + \pi \sum_{j=1}^{n}(a_j^2 + b_j^2).\end{aligned}$$

By Theorem 1.16,

$$\begin{aligned}0 = \lim_{n\to\infty} \|f - S_n f\|_2^2 &= \lim_{n\to\infty} \left(\|f\|_2^2 - 2\langle f, S_n f\rangle_2 + \|S_n f\|_2^2\right) \\ &= \lim_{n\to\infty} \left(\|f\|_2^2 - \|S_n f\|_2^2\right) \\ &= \|f\|_2^2 - \lim_{n\to\infty} \|S_n f\|_2^2 \\ &= \|f\|_2^2 - \left[2\pi a_0^2 + \pi \sum_{n=1}^{\infty} \left(a_n^2 + b_n^2\right)\right].\end{aligned}$$

□

In general, an orthonormal set $\{\phi_n\}$ in $R[a, b]$ is called **complete** if

$$\|f - \sum_{k=1}^{n}\langle f, \phi_k\rangle\|_2^2 \to 0, \qquad \text{as } n \to \infty\ ,$$

for every f. Whevever this happens, the proof above shows that the general Parseval's Identity

$$\|f\|_2^2 = \sum_{n=1}^{\infty}\langle f, \phi_n\rangle^2$$

holds. Our main theorem asserts that $\{1/\sqrt{2\pi}, \cos nx/\sqrt{\pi}, \sin nx/\sqrt{\pi}\}$ forms a complete orthonormal set in $R[-\pi, \pi]$. It plays the role like the canonical basis $\{e_1, \cdots, c_n\}$ in the Eulcidean space $\mathbb{R}^n$.

The norm of f can be regarded as the length of the "vector" f. Parseval's Identity shows that the square of the length of f is equal to the sum of the square of the length of the orthogonal projection of f onto each one-dimensional subspace spanned by the sine and cosine functions. This is an infinite dimensional version of the ancient Pythagoras theorem. It is curious to see what really comes out when you plug in some specific functions. For instance, we take $f(x) = x$ and recall that its Fourier series is given by $\sum 2(-1)^{n+1}/n \sin nx$. Therefore, $a_n = 0, n \geq 0$ and $b_n = 2(-1)^{n+1}/n$ and Parseval's identity yields Euler's summation formula

$$\frac{\pi^2}{6} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots .$$

You could find more interesting identities by applying the same idea to other functions.

Appendix I Series of Functions

This appendix serves to refresh your memory after the long, free summer.

A (real) sequence is a mapping φ from $\mathbb{N}$ to $\mathbb{R}$. For $\varphi(n) = a_n$, we usually denote the sequence by $\{a_n\}$ rather than φ. This is a convention. We say the sequence is convergent if there exists a real number a satisfying, for every $\varepsilon > 0$, there exists some n_0 such that $|a_n - a| < \varepsilon$ for all $n, n \geq n_0$. When this happens, we write $a = \lim_{n\to\infty} a_n$.

An (infinite) series is always associated with a sequence. Given a sequence $\{x_n\}$, set $s_n = \sum_{k=1}^n x_k$ and form another sequence $\{s_n\}$. This sequence is the infinite series associated to $\{x_n\}$ and is usually denoted by $\sum_{k=1}^\infty x_k$. The sequence $\{s_n\}$ is also called the sequence of n-th partial sums of the infinite series. By definition, the infinite series is convergent if $\{s_n\}$ is convergent. When this happens, we denote the limit of $\{s_n\}$ by $\sum_{k=1}^\infty x_k$, in other words, we have

$$\lim_{n\to\infty} \sum_{k=1}^n x_k = \sum_{k=1}^\infty x_k .$$

So the notation $\sum_{k=1}^\infty x_k$ has two meanings, first, it is the notation for an infinite series and, second, the limit of its partial sums (whenever it exists).

When the target $\mathbb{R}$ is replaced by $\mathbb{C}$, we obtain a sequence or a series of complex numbers, and the above definitions apply to them after replacing the absolute value by the complex absolute value or modulus.

Let $\{f_n\}$ be a sequence of real- or complex-valued functions defined on some non-empty E on $\mathbb{R}$. It is called convergent pointwisely to some function f defined on the same E if for every $x \in E$, $\{f_n(x)\}$ converges to $f(x)$ as $n \to \infty$. Keep in mind that $\{f_n(x)\}$ is sequence of real or complex numbers, so its convergence has a valid meaning. A more important concept is the uniform convergence. The sequence $\{f_n\}$ is uniformly convergent to f if, for every $\varepsilon > 0$, there exists some n_0 such that $|f_n(x) - f(x)| < \varepsilon$ for all $n \geq n_0$. In notation $f_n \rightrightarrows f$. Equivalently, uniform convergence holds if, for every $\varepsilon > 0$, there exists some n_1 such that $\|f_n - f\|_\infty < \varepsilon$ for all $n \geq n_1$. Here $\|f\|_\infty$ denotes the sup-norm of f on E.

An (infinite) series of functions is the infinite series given by $\sum_{k=1}^\infty f_k(x)$ where f_k are defined on E. Its convergence and uniform convergence can be defined via its partial sums $s_n(x) = \sum_{k=1}^n f_k(x)$ as in the case of sequences of numbers.

Among several criteria for uniform convergence, the following test is the most useful one.

Weierstrass M-Test. *Let $\{f_k\}$ be a sequence of functions defined on some $E \subset \mathbb{R}$. Suppose that there exists a sequence of non-negative numbers, $\{a_k\}$, such that* *(a)* $|f_k(x)| \le a_k$ *for all* $k \ge 1$*, and*

(b) $\sum_{k=1}^{\infty} a_k$ *is convergent.*

Then $\sum_{k=1}^{\infty} f_k$ *converges uniformly and absolutely on* E*.*

Also, the following "exchange theorem".

Exchange Theorem. *Let* $s_n = \sum_{k=1}^{n} f_k$ *be uniformly convergent to* $\sum_{k=1}^{\infty} f_k$ *on some* $E \subset \mathbb{R}$*. Then*

(a) $\sum_{k=1}^{\infty} f_k \in C(E)$ if $f_k \in C(E)$ *for all* k*.*

(b) *If* E *is an interval and* f_k*'s are differentiable with* $\sum_{k=1}^{n} f_k' \rightrightarrows \sum_{k=1}^{\infty} f_k'$*, then* $\sum_{k=1}^{\infty} f_k$ *is also differentiable and*

$$\Big(\sum_{k=1}^{\infty} f_k\Big)' = \sum_{k=1}^{\infty} f_k' \ .$$

Appendix II Sets of Measure Zero

Let E be a subset of $\mathbb{R}$. It is called of measure zero, or sometimes called a null set, if for every $\varepsilon > 0$, there exists a (finite or infinite) sequence of intervals $\{I_k\}$ satisfying (1) $E \subset \cup_{k=1}^{\infty} I_k$ and (2) $\sum_{k=1}^{\infty} |I_k| < \varepsilon$. (When the intervals are finite, the upper limit of the summation should be changed accordingly.) Here I_k could be an open, closed or any other interval and its length $|I_k|$ is defined to the $b - a$ where $a \le b$ are the endpoints of I_k.

The empty set is a set of measure zero from this definition. Every finite set is also null. For, let $E = \{x_1, \cdots, x_N\}$ be the set. For $\varepsilon > 0$, the intervals $I_k = (x_1 - \varepsilon/(4N), x_k + \varepsilon/(4N))$ clearly satisfy (1) and (2) in the definition.

Next we claim that every countable set is also of measure zero. Let $E = \{x_1, x_2, \cdots\}$ be a countable set. We choose

$$I_k = \left(x_k - \frac{\varepsilon}{2^{k+2}}, x_k + \frac{\varepsilon}{2^{k+2}}\right) \ .$$

Clearly, $E \subset \cup_{k=1}^{\infty} I_k$. On the other hand,

$$\begin{aligned} \sum_{k=1}^{\infty} |I_k| &= \sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} \\ &= \frac{\varepsilon}{2} \\ &< \varepsilon \ . \end{aligned}$$

We conclude that every countable set is a null set.

There are uncountable sets of measure zero. For instance, the Cantor set which plays an important role in analysis, is of measure zero. Here we will not go into this.

The same trick in the above proof can be applied to the following situation.

Proposition A.1. *The union of countably many null sets is a null set.*

Proof. Let $E_k, k \geq 1$, be sets of measure zero. For $\varepsilon > 0$, there are intervals satisfying $\{I_j^k\}, E_k \subset \cup_j I_j^k$, and $\sum_j |I_j^k| < \varepsilon/2^k$. It follows that $E \equiv \cup_k E_k \subset \cup_{j,k} I_j^k = \cup_k \cup_j I_j^k$ and

$$\sum_k \sum_j |I_j^k| < \sum_k \frac{\varepsilon}{2^k} = \varepsilon.$$

□

The concept of a null set comes up naturally in the theory of Riemann integration. A theorem of Lebsegue asserts that a bounded function is Riemann integrable if and only if its discontinuity set is null. The following result is used in the uniqueness assertion on Fourier series. I provide a proof here, but you may just take it for granted.

Proposition A.2. *Let f be a non-negative integrable function on $[a, b]$. Then $\int_a^b f = 0$ if and only if f is equal to 0 except on a null set. Consequently, two integrable functions f, g satisfying*

$$\int_a^b |f - g| = 0,$$

if and only if f is equal to g except on a null set.

Proof. We set, for each $k \geq 1$, $A_k = \{x \in [a, b] : \ f(x) > 1/k\}$. It is clear that

$$\{x : \ f(x) > 0\} = \bigcup_{k=1}^{\infty} A_k \ .$$

By Proposition A.1., it suffices to show that each A_k is null. Thus let us consider A_{k_0} for a fixed k_0. Recall from the definition of Riemann integral, for every $\varepsilon > 0$, there exists a partition $a = x_1 < x_2 < \cdots < x_n = b$ such that

$$0 \leq \sum_{k=1}^{n-1} f(z_k)|I_k| = \left| \sum_{k=1}^{n-1} f(z_k)|I_k| - \int_a^b f \right| < \frac{\varepsilon}{k_0} \ ,$$

where $I_k = [x_k, x_{k+1}]$ and z_k is an arbitrary tag in $[x_j, x_{j+1}]$. Let $\{k_1, \cdots, k_m\}$ be the index set for which I_{k_j} contains a point z_{k_j} from A_{k_0}. Choosing the tag point to be z_{k_j}, we have $f(z_{k_j}) > 1/k_0$. Therefore,

$$\frac{1}{k_0} \sum_{k_j} |I_{k_j}| = \sum_{k_j} f(z_{k_j})|I_{k_j}| \leq \sum_{k=1}^{n-1} f(z_k)|I_k| < \frac{\varepsilon}{k_0},$$

so

$$\sum_{k_j} |I_{k_j}| < \varepsilon.$$

We have shown that A_{k_0} is of measure zero.

Conversely, assume that the set $Z = \{x \in [a, b] : f(x) \neq 0\}$ is of measure zero. For $\varepsilon > 0$, there are open intervals I_k whose union covers Z and $\sum_k |I_k| < \varepsilon$. Without loss of generality we may assume these I_k are mutually disjoint. Let $[c, d] \subset (a, b)$. Since $[c, d]$ is a compact set and I_k's cover $[c, d]$, there exist a finite subcover. Re-index these finite intervals as $(a_1, b_1), (a_2, b_2), \cdots, (a_n, b_n)$ where $a_1 < c$ and $d < b_n$. The points $c, b_1, a_2, b_2, \cdots, a_n, d$ form a partition of $[c, d]$. Since f vanishes away from $[c, b_1], [a_2, b_2], \cdots [a_n, d]$, the Riemann sum $R = \sum f(x_j)\Delta x_j$ satisfies $|R| \leq M \sum \Delta x_j \leq M\varepsilon$, where $M = \sup |f|$. Since ε can be arbitrarily small, we conclude that

$$\int_c^d f(x)dx = 0 \ ,$$

for all $a < c < d < b$. Now, given $\varepsilon > 0$, fix c, d so that $M(c - a + b - d) < \varepsilon$. Then

$$\begin{aligned} \left| \int_a^b f(x)dx \right| &\leq \left| \int_a^c f(x)dx \right| + \left| \int_c^d f(x)dx \right| + \\ &\quad \left| \int_d^b f(x)dx \right| \\ &\leq M(c-a) + M(b-d) \\ &< \varepsilon \ , \end{aligned}$$

done. □

A property holds **almost everywhere** if it holds except on a null set. For instance, this proposition asserts that the integral of a non-negative function is equal to zero if and only if it vanishes almost everywhere.

Comments on Chapter 1. Historically, the relation (2.2) comes from a study on the one-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

where $u(x, t)$ denote the displacement of a string at the position-time (x, t). Around 1750, D'Alembert and Euler found that a general solution of this equation is given by

$$f(x - ct) + g(x + ct)$$

where f and g are two arbitrary twice differentiable functions. However, D. Bernoulli found that the solution could be represented by a trigonometric series. These two different ways of representing the solutions led to a dispute among the mathematicians at that time, and it was not settled until Fourier gave many convincing examples of representing functions by trigonometric series in 1822. His motivation came from heat conduction. After that, trigonometric series have been studied extensively and people call it Fourier series in honor of the contribution of Fourier. Nowadays, the study of Fourier series has matured into a branch of mathematics called harmonic analysis. It has equal importance in theoretical and applied mathematics, as well as other branches of natural sciences and engineering.

The book by R.T. Seely, “An Introduction to Fourier Series and Integrals”, W.A. Benjamin, New York, 1966, is good for further reading.

In some books the Fourier series of a function is written in the form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

instead of

$$a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

so that the formula for a_0 is the same as the other a_n's (see (2.1)). However, our notation has the advantage that a_0 has a simple meaning, i.e., it is the average of the function over a period.

Concerning the convergence of a Fourier series to its function, we point out that an example of a continuous function whose Fourier series diverges at some point can be found in Stein-Sharachi. More examples are available by googling. The classical book by A. Zygmund, “Trigonometric Series” (1959) reprinted in 1993, contains most results before 1960. After 1960, one could not miss to mention Carleson's sensational work in 1966, whose result implies that the Fourier series of every function in $R_{2\pi}$ converges to the function itself almost everywhere.

There are several standard proofs of the Weierstrass approximation theorem, among them Rudin's proof in “Principles” by expanding an integral kernel and Bernstein's proof based on binomial expansion are both worth reading. Recently the original proof of Weierstrass by the heat kernel is available on the web. It is nice to take a look too. In Chapter 3 we will reproduce Rudin's proof and then discuss Stone-Weierstrass theorem, a far reaching generalization of Weierstrass approximation theorem.

The aim of this chapter is to give an introduction to Fourier series. It will serve the purpose if your interest is aroused and now you consider to take our course on Fourier analysis in the future. Not expecting a thorough study, I name Stein-Shakarchi as the only reference.