Monographs in Mathematical Economics 2

Toru Maruyama

Fourier Analysis of Economic Phenomena

Springer

Monographs in Mathematical Economics

Volume 2

More information about this series at http://www.springer.com/series/13278

Toru Maruyama

Fourier Analysis of Economic Phenomena

Springer

Toru Maruyama
Professor Emeritus
Keio University
Tokyo, Japan

ISSN 2364-8279 ISSN 2364-8287 (electronic)
Monographs in Mathematical Economics
ISBN 978-981-13-2729-2 ISBN 978-981-13-2730-8 (eBook)
https://doi.org/10.1007/978-981-13-2730-8

Library of Congress Control Number: 2018958313

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Fourier analysis as a branch of mathematical analysis was born from ingenious attempt to express a function as a weighted sum of trigonometric waves. In particular, abstract theory addressing a similar problem on topological groups is sometimes called harmonic analysis. However, Fourier analysis and harmonic analysis are almost synonymous, and we do not distinguish between these two names in this book.

The interaction between Fourier analysis and other fields of mathematics has a long and remarkable history. The role of Fourier analysis in physics has also been impossible to exaggerate. However, the situation is quite different regarding its relation with economics. Indeed, Fourier analysis has not been so familiar to economists, except for specialists of time series analysis. Therefore, we have to confess that the extent of the interaction between Fourier analysis and economics is quite restricted.

Nevertheless, the narrowness of the interaction realm is a completely different matter from its significance. As a peculiar example of significant interaction between them, I would like to choose the studies into periodic economic phenomena as the main materials for this book.

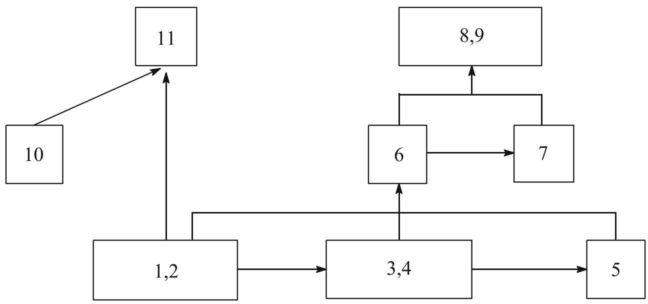
Not a few economic variables show behaviors over time which enjoy periodic or almost periodic regularities. Trade cycles or business fluctuations are typical examples of periodic economic phenomena. Many economic theorists since the latter half of the nineteenth century have incessantly endeavored to examine periodic fluctuations of economic activities. In the process of my research into trade cycles, two classical works impressed on me the close connection between dynamic economic theory and Fourier analysis. The first one is N. Kaldor's theory of trade cycles based upon nonlinear investment functions. The second is the theory of periodic behaviors of moving average processes due to E. Slutsky. Although their works are more or less rudimentary and devoid of mathematical rigor, it is indeed surprising that their results can be basically justified by more exact mathematical reasoning. Rather, I should say, they invoked several new developments of mathematics. It is remarkable that the works of Kaldor and Slutsky share a common mathematical background of Fourier analysis. That is why I selected these two as key figures for this book.

It is the object of this book to describe the entire mathematical skeleton which supports their economic theories exploring the periodic behaviors of an economy.

This book is not intended to be a well-balanced collection of economic applications or something like a convenient cookbook. What I expect the readers to find in the book is a systematic body of carefully constructed mathematical theories which are hidden in the background of the well-known traditional economic doctrines mentioned above.

I do not intend to go into details of many economic theories developed more recently. If I dared to try it, I would have to spend many pages on the explanations of concepts and reasonings peculiar to economics. It would eventually obscure the main theme of the book. The attitude I chose is to restrict myself to a rather narrow field of applications familiar to many economists and to give full details of mathematical materials instead.

The Kaldorian theory of trade cycles is discussed in Chap. 11. The Hopf bifurcation theorem plays a key role in the existence proof of periodic solutions of the fundamental equation of the Kaldor model. A natural approach to Hopf's theory is explored from the viewpoint of Fourier analysis. The classical results (Chaps. 1 and 2) on the convergence problem of Fourier series as well as the Carleson–Hunt theorem are effectively made use of. The theory of Fredholm operators provides another prerequisite. We also give a systematic exposition of the topic in Chap. 10.



Logical Connections between Chapters

The theory of periodic or almost periodic stationary processes, initiated by Slutsky, is discussed in Chaps. 8 and 9. The most essential tool is provided by the Herglotz–Bochner theorem on the spectral representation of positive definite functions. It is prepared in Chap. 6 in the framework of generalized harmonic analysis fortified by L. Schwartz's theory of distributions. Stone's theorem on one-parameter groups of unitary operators (Chap. 7) is crucial for the spectral representation of stationary processes.

Chapters 1, 2, 3, 4, and 5 may be regarded as a menu for a standard basic course on Fourier analysis. The contents of these chapters are more or less commonsense in this mathematical discipline. Classical Fourier series (Chaps. 1 and 2); Fourier transforms on $\mathfrak{L}^1$, $\mathfrak{L}^2$, $\mathfrak{S}$ (space of rapidly decreasing functions), and $\mathfrak{S}'$ (space of tempered distributions) (Chaps. 3 and 4); and the summation method and spectral synthesis (Chap. 5) are the contents of the menu.

This book is the English edition of the Japanese version, published in 2017 by Chisen Shokan, Tokyo. There is no significant difference between them, except one point. The last section of Chap. 8 in the English edition is not included in the Japanese version. This section is a brief exposition of E. Slutsky's original work, and it was added on the advice of the series editor in charge.

Since I had started to learn Fourier analysis in his class, the late Professor Tatsuo Kawata continuously gave me kind encouragement as well as rich advice. Reading through the manuscript of this book, I am impressed by Professor Kawata's decisive influence upon me, both in the motivation and in the way of thinking.

The materials of the book are based upon my lectures given during the graduate course in economics at Keio University and the graduate course in mathematical sciences at the University of Tokyo. The critical comments from my former students, including Messrs. H. Kawabi, Y. Hosoya, and C. Yu Chaowen, were always beneficial.

The kind and detailed comments by the anonymous referee deserve special mention. I revised, though not sufficiently enough, the earlier manuscript according to the advice of the referee.

I have to express my sincere gratitude to Professor S. Kusuoka for the tedious work of checking the manuscript as the editor in charge. I am very much indebted to him for his kind advice concerning probability theory, which I am not familiar with very well. I would also like to devote my thanks to Professors A. Ioffe, L. Nirenberg, and P.H. Rabinowitz for their suggestive comments on my work.

I incorporate in the text several of my articles published elsewhere; I appreciate the generous permissions of the American Mathematical Society, Operations Research Society of Japan, and Society of Mathematical Economics.

Finally, I express my deepest thanks to the members of my Dream Team, Mmes. Y. Hagiwara, Y. Katsuragi, T. Noma, and H. Yusa for their efficient cooperation.

Tokyo, Japan
March 2018

Toru Maruyama

Contents

Chapter 1
Fourier Series on Hilbert Spaces

Let $e_1, e_2, \cdots, e_l$ be the standard basis of an l-dimensional Euclidean space consisting of l unit vectors. Then any vector x can be expressed as

$$x = \sum_{i=1}^{l} c_i e_i$$

and such an expression is determined uniquely. The coefficients $c_1, c_2, \cdots, c_l$ are computed as $c_i = \langle x, e_i \rangle$ (inner product).

Is a similar expression possible for an infinite dimensional vector space (i.e. pre-Hilbert space or Hilbert space)? It is the theory of Fourier series which answers this question.

In this chapter, we start by defining a Fourier series on an abstract Hilbert space and go on to discuss its fundamental properties.

1.1 Hilbert Spaces

Let $\mathfrak{H}$ be a complex vector space.[1,2] A function $\langle \cdot, \cdot \rangle : \mathfrak{H} \times \mathfrak{H} \to \mathbb{C}$ is called an inner product on $\mathfrak{H}$ if it satisfies the following four conditions for any elements x, y, x_1, and x_2 of $\mathfrak{H}$:

(i) $\langle x, x \rangle \geqq 0$; $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.
(ii) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (conjugate complex).

[1] A systematic theory of functional analysis in the framework of Hilbert spaces is discussed in many textbooks in this discipline. For instance, Halmos [4] and Schwartz [8] are very well-written classics. See also Lax [6] Chap. 6 and Maruyama [7] Chap. 3.

[2] In the case of a real vector space $\mathfrak{H}$, an inner product is a real-valued function $\langle \cdot, \cdot \rangle : \mathfrak{H} \times \mathfrak{H} \to \mathbb{R}$ such that (i)–(iv) are satisfied. Of course, (ii) should be rewritten as $\langle x, y \rangle = \langle y, x \rangle$.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_1

(iii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$; $\alpha \in \mathbb{C}$.
(iv) $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$.

It follows from the above axioms that:

(a) $\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle$.
(b) $\langle x, \alpha y \rangle = \bar{\alpha} \langle x, y \rangle$; $\alpha \in \mathbb{C}$.

Given an inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{H}$, we define a function $\| \cdot \| : \mathfrak{H} \to \mathbb{R}$ by

$$\|x\| = \sqrt{\langle x, x \rangle}, \quad x \in \mathfrak{H}. \tag{1.1}$$

Then a couple of important inequalities immediately follow:

(I) **Schwarz inequality** $|\langle x, y \rangle| \leqq \|x\| \cdot \|y\|$;
(II) **triangular inequality** $\|x + y\| \leqq \|x\| + \|y\|$.

We recognize that the function $\| \cdot \|$ is a norm on $\mathfrak{H}$, taking account of (II).

A normed vector space endowed with the norm (1.1) defined by an inner product is called a **pre-Hilbert space**. A complete pre-Hilbert space is called a **Hilbert space**.

We briefly pick up a few basic facts concerning a Hilbert space.

1° (parallelogram law) (i) In a pre-Hilbert space $\mathfrak{H}$, the equality

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \tag{1.2}$$

holds good for any $x, y \in \mathfrak{H}$,

(ii) Conversely, if $(\mathfrak{H}, \| \cdot \|)$ is a normed space which satisfies (1.2), then it is a pre-Hilbert space.

2° (Pythagorean theorem) Two elements x, y of a Hilbert space $\mathfrak{H}$ are said to be **orthogonal** if $\langle x, y \rangle = 0$. The notation $x \perp y$ means that these two elements are orthogonal.

If x and y in $\mathfrak{H}$ are orthogonal, we have

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

More generally, the relation

$$\|x_1 + x_2 + \cdots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2$$

holds good if $x_1, x_2, \cdots, x_n$ in $\mathfrak{H}$ are mutually orthogonal, i.e. $x_i \perp x_j$ $(i \neq j)$.

3° (orthogonality) Two elements x and y of a Hilbert space $\mathfrak{H}$ are orthogonal if and only if

$$\|x\| \leqq \|x + \lambda y\| \quad \text{for all} \quad \lambda \in \mathbb{C}.$$

4° (Minimum distance theorem) Let C be a nonempty closed convex set in a Hilbert space $\mathfrak{H}$. Then there exists, for any $z \in \mathfrak{H}$, a unique element $x \in C$ such that

$$\|x - z\| = \inf_{y \in C} \|y - z\|.$$

Definition 1.1 Let $\mathfrak{H}$ be a pre-Hilbert space and M be its nonempty subset. Then the set

$$M^{\perp} = \{x \in \mathfrak{H} \,|\, x \perp y \quad \text{for all} \quad y \in M\}$$

is called the **orthogonal complement** of M.

5° (orthogonal decomposition) Let $\mathfrak{H}$ be a Hilbert space and $\mathfrak{M}$ be its closed subspace. Then for any $z \in \mathfrak{H}$, there exist some $x \in \mathfrak{M}$ and $y \in \mathfrak{M}^{\perp}$ such that $z = x + y$. Such x and y are determined uniquely.

The following theorem due to F. Riesz is the most important foundation in the theory of Hilbert spaces. It asserts that any continuous linear functional on a Hilbert space can be represented as an inner product.

Theorem 1.1 (F. Riesz) *Let $\mathfrak{H}$ be a Hilbert space.*

(i) *If we define an operator $\Lambda_y : \mathfrak{H} \to \mathbb{C}$ for each $y \in \mathfrak{H}$ by*

$$\Lambda_y(x) = \langle x, y \rangle,$$

then $\Lambda_y \in \mathfrak{H}'$ and $\|\Lambda_y\| = \|y\|$. ($\mathfrak{H}'$ is the dual space of $\mathfrak{H}$.)

(ii) *Conversely, for each $\Lambda \in \mathfrak{H}'$, there exists uniquely $y_\Lambda \in \mathfrak{H}$ such that*

$$\Lambda(x) = \langle x, y_\Lambda \rangle \quad \text{for all} \quad x \in \mathfrak{H}.$$

Proof (i) We have, by the Schwarz inequality, that

$$|\Lambda_y(x)| = |\langle x, y \rangle| \leqq \|x\| \cdot \|y\|.$$

Consequently, Λ_y is a bounded linear functional with

$$\|\Lambda_y\| \leqq \|y\|. \tag{1.3}$$

On the other hand, since

$$\Lambda_y(y) = \langle y, y \rangle = \|y\|^2,$$

it follows that

$$\|\Lambda_y\| \geqq \|y\|. \tag{1.4}$$

Hence $\|\Lambda\| = \|y\|$ by (1.3) and (1.4).

(ii) Assume $\Lambda \in \mathfrak{H}'$. The theorem is trivial in the case of $\Lambda \equiv 0$. So, without loss of generality, we may assume $\Lambda \not\equiv 0$. $\mathfrak{M} = \mathrm{Ker}\Lambda$ is a nonempty closed subspace and $\mathfrak{M} \subsetneq \mathfrak{H}$. For any $z \in \mathfrak{H}\backslash\mathfrak{M}$, there exists a unique $x \in \mathfrak{M}$ such that

$$\|x - z\| = \inf_{y \in \mathfrak{M}} \|y - z\| \tag{1.5}$$

(by 4°). Since $\mathfrak{M}$ is a vector subspace,

$$x + \lambda w \in \mathfrak{M}$$

for any $w \in \mathfrak{M}$ and $\lambda \in \mathbb{C}$. We have, by (1.5),

$$\|x - z\| \leqq \|x + \lambda w - z\|.$$

Hence we can conclude, by 3°, that

$$(x - z) \perp w \quad \text{for any} \quad w \in \mathfrak{M}. \tag{1.6}$$

Here $x - z \notin \mathfrak{M}$, since $x \in \mathfrak{M}$ and $z \notin \mathfrak{M}$, and so $\Lambda(x - z) \neq 0$. If we define

$$u = \frac{x - z}{\Lambda(x - z)},$$

then $\Lambda(u) = 1$. It follows that

$$\Lambda(v - \Lambda(v)u) = \Lambda(v) - \Lambda(v)\Lambda(u) = 0 \tag{1.7}$$

for all $v \in \mathfrak{H}$. $u \perp w$ for all $w \in \mathfrak{M}$ by (1.6), and $v - \Lambda(v)u \in \mathfrak{M}$ by (1.7). Therefore we have

$$\langle v - \Lambda(v)u, u\rangle = \langle v, u\rangle - \Lambda(v)\|u\|^2 = 0.$$

Consequently, we can conclude that

$$\Lambda(v) = \left\langle v, \frac{u}{\|u\|^2} \right\rangle \quad \text{for all} \quad v \in \mathfrak{H}.$$

Defining y_Λ by

$$y_\Lambda = \frac{u}{\|u\|^2},$$

we arrive at our desired result. The uniqueness of y_Λ is obvious. □

1.2 Orthonormal Systems

We start by discussing the theory of Fourier series on real or complex Hilbert spaces. $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ denote an inner product and a norm of a Hilbert space, respectively.

Definition 1.2 A subset Φ of a Hilbert space $\mathfrak{H}$ is called an **orthogonal system** in $\mathfrak{H}$ if $\langle\varphi,\varphi'\rangle = 0$ for any two distinct elements φ and φ' of Φ. In particular, Φ is said to be an **orthonormal system** if $\|\varphi\| = 1$ for all $\varphi \in \Phi$.

It is convenient to keep in mind the following rule of calculation. That is, if $\Phi = \{\varphi_1, \varphi_2, \cdots, \varphi_n\}$ is an orthonormal system, then

$$\Big\| \sum_{i=1}^{n} c_i\varphi_i \Big\|^2 = \Big\langle \sum_{i=1}^{n} c_i\varphi_i, \sum_{i=1}^{n} c_i\varphi_i \Big\rangle = \sum_{i=1}^{n} c_i\overline{c_i} = \sum_{i=1}^{n} |\ c_i\ |^2 \tag{1.8}$$

for any of the complex numbers $c_1, c_2, \cdots, c_n$, where $\overline{c_i}$ is the conjugate complex number of c_i.

It follows that

$$\sum_{i=1}^{n} c_i\varphi_i = 0 \quad \Longrightarrow \quad c_i = 0 \quad \text{for all} \quad i.$$

Therefore vectors which form an orthonormal system are linearly independent.

We next show some typical examples of orthonormal systems. Although we have no chance to make use of Examples 1.4, 1.5, and 1.6, we briefly give proofs.[3]

Example 1.1 The system of unit vectors $e_1 = (1, 0, \cdots, 0)$, $e_2 = (0, 1, 0, \cdots, 0)$, $\cdots$, $e_l = (0, \cdots, 0, 1)$ in $\mathbb{C}^l$ (or $\mathbb{R}^l$) forms an orthonormal system.

Example 1.2 The system of functions

$$\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}}\cos x,\ \frac{1}{\sqrt{\pi}}\sin x,\ \cdots,\ \frac{1}{\sqrt{\pi}}\cos nx,\ \frac{1}{\sqrt{\pi}}\sin nx,\ \cdots;\ n = 1, 2, \cdots$$

forms an orthonormal system in $\mathfrak{L}^2([-\pi,\pi],\mathbb{C})$ or $\mathfrak{L}^2([-\pi,\pi],\mathbb{R})$.

Example 1.3 The system of functions

$$\frac{1}{\sqrt{2\pi}}e^{inx}; \quad n = 0, \pm 1, \pm 2, \cdots$$

forms an orthonormal system in $\mathfrak{L}^2([-\pi,\pi],\mathbb{C})$.

[3] See Folland [3] Chap. 6, Terasawa [10] pp. 145–149, pp. 414–416, Yosida [11] Chap. 1, §3 and Chap. 2, §2.

Example 1.4 The system of functions

$$\sqrt{\frac{2n+1}{2}}P_n(x); \quad n = 0, 1, 2, \cdots$$

forms an orthonormal system in $\mathfrak{L}^2([-1, 1], \mathbb{C})$, where

$$P_n(x) = \frac{1}{2^n n!} \cdot \frac{d^n}{dx^n}(x^2 - 1)^n \qquad \text{(Legendre polynomial)}.$$

$P_n(x)$ is a polynomial of degree n, and the term of the highest degree is $(2n)!/2^n(n!)^2 \cdot x^n$.[4]

For any function $f : [-1, 1] \to \mathbb{R} \in \mathfrak{C}^n([-1, 1], \mathbb{R})$, we obtain, by repeating integration by parts n times, that[5]

$$2^n n!\langle f, P_n\rangle = \int_{-1}^{1} f(x)\frac{d^n}{dx^n}(x^2 - 1)^n dx = (-1)^n \int_{-1}^{1} \frac{d^n}{dx^n}f(x) \cdot (x^2 - 1)^n dx. \tag{1.9}$$

Applying this formula to $f = P_m (m < n)$, we have

$$2^n n!\langle P_m, P_n\rangle = 0 \quad (m < n),$$

since the n-th derivative of P_m is zero. The same result also holds good for $m > n$ by a similar reasoning. Hence it follows that

$$\langle P_m, P_n\rangle = 0 \quad \text{if} \quad m \neq n.$$

This proves the orthogonality of P_m and P_n $(m \neq n)$.

We next show that $\|P_n\| = 1$. The n-th derivative of $P_n(x)$ is computed as

$$\frac{d^n}{dx^n}P_n(x) = \frac{(2n)!}{2^n \cdot n!} = \frac{1 \cdot 2 \cdot \cdots \cdot (2n)}{(2 \cdot 1)(2 \cdot 2) \cdot \cdots \cdot (2 \cdot n)} = 1 \cdot 3 \cdot 5 \cdots (2n - 1).$$

Applying (1.9) to $f = P_n$, we obtain

[4]The term of the highest degree $= (1/2^n n!)\{(2n)(2n - 1) \cdots (n + 1)\}x^n$.

[5]The first process of integration by parts is as follows:

$$f(x)\frac{d^{n-1}}{dx^{n-1}}(x^2 - 1)^n\Big|_{-1}^{1} - \int_{-1}^{1} f'(x)\frac{d^{n-1}}{dx^{n-1}}(x^2 - 1)^n dx$$

$$= -\int_{-1}^{1} f'(x)\frac{d^{n-1}}{dx^{n-1}}(x^2 - 1)^n dx.$$

$$
\begin{aligned}
2^n n!\langle P_n, P_n\rangle &= (-1)^n \int_{-1}^{1} (1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1))(x^2-1)^n dx \\
&= (1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1)) \int_{-1}^{1} (1-x^2)^n dx.
\end{aligned}
\tag{1.10}
$$

The following computation is based upon the beta integration formula.[6]

$$
\begin{aligned}
\int_{-1}^{1} (1-x^2)^n dx &= 2\int_{0}^{1} (1-x^2)^n dx = \int_{0}^{1} (1-y)^n y^{-\frac{1}{2}} dy \\
&= \frac{\Gamma(n+1)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(n+\frac{3}{2}\right)} = \frac{n!\sqrt{\pi}}{\Gamma\left(n+\frac{3}{2}\right)} \\
&= \frac{n!}{\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\cdots\cdots\left(n+\frac{1}{2}\right)} = \frac{2^{n+1}\cdot n!}{1\cdot 3\cdot 5\cdot\cdots\cdot(2n+1)},
\end{aligned}
\tag{1.11}
$$

where Γ is the gamma function. Equations (1.10) and (1.11) imply

$$
\|P_n\|^2 = \frac{2}{2n+1}.
$$

[6] We denote by $\mathbb{C}_+$ the complex half-plane $\mathcal{R}ez > 0$. The function $(\Gamma : \mathbb{C}_+ \to \mathbb{C})$ defined by

$$
\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx \tag{t}
$$

is called the **gamma function**. Γ is analytic on $\mathbb{C}_+$. Sometimes the domain of Γ is taken to be $(0,\infty)$ rather than $\mathbb{C}_+$. In the text above, we also follow this policy. The function $B : \mathbb{C}_+ \times \mathbb{C}_+ \to \mathbb{C}$ defined by

$$
B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx
$$

is called the **beta function**. (Note that the integral is convergent.) The following formulas hold good (cf. Cartan [1] Chap. V, §3 and Takagi [9] Chap. 5, §68):

1° $\Gamma(s+1) = s\Gamma(s)$.
We also obtain $\Gamma(n+1) = n!$ for all $n \in \mathbb{N}$ by induction.
2° $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$.
3° $\Gamma(s)\Gamma(1-s) = \pi/\sin \pi s$.
If $s = 1/2$, in particular, we have $\Gamma(1/2) = \sqrt{\pi}$.
4° $\Gamma\left(n+\frac{1}{2}\right) = \frac{1}{2}\left(\frac{3}{2}\right)\cdots\left(n-\frac{1}{2}\right)\sqrt{\pi} \quad (n = 0, 1, 2, \cdots)$.

Hence we have finally that

$$\left\| \sqrt{\frac{2n+1}{2}} P_n(x) \right\|_2 = 1.$$

Example 1.5 The system of functions

$$\frac{1}{\sqrt{2^n n! \sqrt[4]{\pi}}} H_n(x) e^{-\frac{x^2}{2}}; \quad n = 0, 1, 2, \cdots$$

forms an orthonormal system in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$, where

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \qquad \text{(Hermite polynomial)}.$$

$H_n(x)$ is a polynomial of n-th degree, and its term of highest degree is $(2x)^n$.

Applying integration by parts to any $f \in \mathfrak{C}^n(\mathbb{R}, \mathbb{R})$, we have

$$\int_{-\infty}^{\infty} f(x) H_n(x) e^{-x^2} dx = (-1)^n \int_{-\infty}^{\infty} f(x) \frac{d^n}{dx^n} e^{-x^2} dx = \int_{-\infty}^{\infty} \frac{d^n}{dx^n} f(x) \cdot e^{-x^2} dx. \tag{1.12}$$

It follows from (1.12) that

$$\int_{-\infty}^{\infty} H_m(x) H_n(x) e^{-x^2} dx = 0$$

for $m < n$. We also get the same result for $m > n$. In other words, H_m and H_n ($m \neq n$) are orthogonal with respect to the measure $e^{-x^2} dx$ with the weight e^{-x^2}. If we define $\tilde{H}_n(x) = H_n(x) e^{-x^2/2}$, then $\tilde{H}_m$ and $\tilde{H}_n$ ($m \neq n$) are orthogonal in the usual sense, i.e.

$$\int_{-\infty}^{\infty} H_m(x) e^{-\frac{x^2}{2}} \cdot H_n(x) e^{-\frac{x^2}{2}} dx = 0 \quad \text{if} \quad m \neq n.$$

Setting $f = H_n = (2x)^n + \cdots$ in (1.12), we have $(d^n/dx^n) f(x) = 2^n n!$. Hence

$$\| H_n(x) e^{-\frac{x^2}{2}} \|^2 = 2^n n! \int_{-\infty}^{\infty} e^{-x^2} dx = 2^n n! \sqrt{\pi}.$$

It follows finally that

$$\left\| \frac{1}{\sqrt{2^n n! \sqrt[4]{\pi}}} H_n(x) e^{-\frac{x^2}{2}} \right\|_2 = 1.$$

Example 1.6 The system of functions

$$\frac{1}{n!}L_n(x)e^{-\frac{x}{2}}; \quad n = 0, 1, 2, \cdots$$

forms an orthonormal system in $\mathfrak{L}^2((0, \infty), \ \mathbb{C})$, where

$$L_n(x) = e^x \frac{d^n}{dx^n}(x^n e^{-x}) \qquad \text{(Laguerre polynomial).}$$

$L_n(x)$ is a polynomial of degree n. The term of the highest degree is $(-1)^n x^n$.
For any $f \in \mathfrak{C}^n([0, \infty), \mathbb{R})$, integration by parts gives

$$\begin{aligned} \int_0^\infty f(x)L_n(x)e^{-x}dx &= \int_0^\infty f(x) \cdot e^x \frac{d^n}{dx^n}(x^n e^{-x}) \cdot e^{-x}dx \\ &= (-1)^n \int_0^\infty \frac{d^n}{dx^n} f(x) \cdot (x^n e^{-x})dx. \end{aligned} \tag{1.13}$$

It follows that

$$\int_0^\infty L_m(x)L_n(x)e^{-x}dx = 0$$

if $m < n$. The same result also holds good for $m > n$. If we define $\tilde{L_n}(x) = L_n(x)e^{-x/2}$, $\tilde{L}_m$ and $\tilde{L}_n$ $(m \neq n)$ are orthogonal, i.e.

$$\int_0^\infty L_m(x)e^{-\frac{x}{2}} \cdot L_n(x)e^{-\frac{x}{2}}dx = 0 \qquad \text{if} \quad n \neq m.$$

Apply (1.13) for $f = L_n$. Then we have

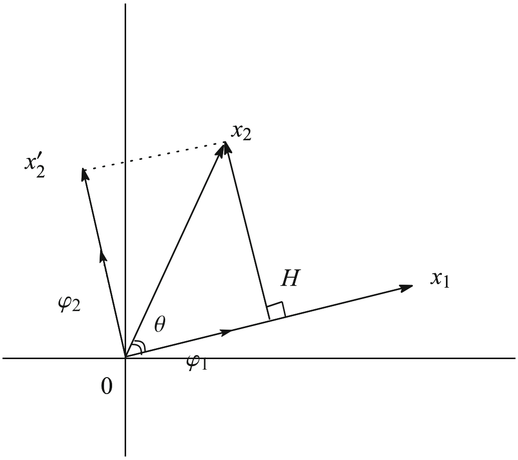
$$\begin{aligned} \left\| \frac{1}{n!}L_n(x)e^{-\frac{x}{2}} \right\|_2^2 &= \frac{(-1)^n}{(n!)^2} \int_0^\infty (-1)^n \cdot x^n e^{-x}dx \\ &= \frac{1}{(n!)^2} \int_0^1 x^n e^{-x}dx = \frac{1}{(n!)^2}\Gamma(n+1) = 1, \end{aligned}$$

taking account of $d^n/dx^n(L_n) = (-1)^n$.

We have illustrated some concrete examples of orthonormal systems in various Hilbert spaces. Here arises a question: does there exist an orthonormal system in a general Hilbert space? The next theorem gives a positive answer to this problem.

Theorem 1.2 (Schmidt's orthogonalization) *Let $\{x_m\}$ be a linearly independent sequence in a Hilbert space $\mathfrak{H}$. Then we can construct an orthonormal system $\{\varphi_n\}$ of the form*

Fig. 1.1 Orthogonalization



$$\varphi_1 = c_{11}x_1, \quad \varphi_2 = c_{21}x_1 + c_{22}x_2, \quad \cdots$$
$$\varphi_n = c_{n1}x_1 + c_{n2}x_2 + \cdots + c_{nn}x_n, \quad \cdots$$

by choosing suitable complex numbers c_{nm}.

Let us start by illustrating the intuitive idea of the proof (See Fig. 1.1).

Since $\{x_n\}$ is linearly independent, $x_n \neq 0$ for all n. So we can define $\varphi_1 = x_1/\|x_1\|$. Draw a line spanned by x_1. Let H be the intersection of this line and another one which runs through the top of the vector x_2 and is orthogonal to the first. The length of $0H$ is

$$\|x_2\| \cos\theta = \|x_2\| \cdot \|\varphi_1\| \cos\theta = \langle x_2, \varphi_1 \rangle,$$

where θ is the angle between φ_1 and x_2. The vector x_2' defined by

$$x_2' = x_2 - \langle x_2, \varphi_1 \rangle \varphi_1$$

is the vector connecting the top of x_2 and H. x_2' is clearly orthogonal to φ_1. If we adjust the length by $\varphi_2 = x_2'/\|x_2'\|$, φ_1 is orthogonal to φ_2 and $\|\varphi_1\| = \|\varphi_2\| = 1$. We have only to repeat this process.

Proof (of Theorem 1.2) Define $\{x_n'\}$ and $\{\varphi_n\}$ by

$$x_1' = x_1, \quad \varphi_1 = x_1'/\|x_1'\|$$
$$x_2' = x_2 - \langle x_2, \varphi_1 \rangle \varphi_1, \quad \varphi_2 = x_2'/\|x_2'\|$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$x_{n+1}' = x_{n+1} - \sum_{k=1}^{n} \langle x_{n+1}, \varphi_k \rangle \varphi_k, \quad \varphi_{n+1} = x_{n+1}'/\|x_{n+1}'\|.$$

We note that each φ_n is a linear combination of $\{x_1, x_2, \cdots, x_n\}$ and each x_n is a linear combination of $\{\varphi_1, \varphi_2, \cdots, \varphi_n\}$. Such a construction as above is possible because each x'_n is not zero. It can be confirmed as follows. First of all, it is obvious that $x_1 = x'_1 \neq 0$ since $\{x_n\}$ is linearly independent. Assume next that $x'_k \neq 0$ $(1 \leqq k \leqq n)$ and $x'_{n+1} = 0$. Then

$$x_{n+1} = \sum_{k=1}^{n} \langle x_{n+1}, \varphi_k \rangle \varphi_k.$$

Since each φ_k is a linear combination of $\{x_1, x_2, \cdots, x_k\}$, x_{n+1} would be a linear combination of $\{x_1, x_2, \cdots, x_n\}$. It is a contradiction to the linear independence of $\{x_1, x_2, \cdots, x_{n+1}\}$. Hence we must have $x'_{n+1} \neq 0$.

It is clear that $\|\varphi_n\| = 1$ for all n.

Finally, we show the orthogonality of $\{\varphi_n\}$. To start with, $\langle \varphi_1, \varphi_2 \rangle = 0$ since

$$\langle x'_2, \varphi_1 \rangle = \langle x_2, \varphi_1 \rangle - \langle x_2, \varphi_1 \rangle \langle \varphi_1, \varphi_1 \rangle = \langle x_2, \varphi_1 \rangle - \langle x_2, \varphi_1 \rangle = 0.$$

Assume next that $\varphi_1, \varphi_2, \cdots, \varphi_n$ are orthogonal to each other. Then it is easy to show that

$$\langle \varphi_{n+1}, \varphi_j \rangle = 0 \quad \text{for} \quad j = 1, 2, \cdots, n. \tag{1.14}$$

In fact, we have by simple calculations that

$$\langle x'_{n+1}, \varphi_j \rangle = \langle x_{n+1}, \varphi_j \rangle - \sum_{k=1}^{n} \langle x_{n+1}, \varphi_k \rangle \langle \varphi_k, \varphi_j \rangle = \langle x_{n+1}, \varphi_j \rangle - \langle x_{n+1}, \varphi_j \rangle = 0,$$

which implies (1.14). □

Theorem 1.2 tells us that any Hilbert space of infinite dimension has an orthonormal system which consists of at least countable vectors.

1.3 Fourier Series

The unit vectors $e_1 = (1, 0, \cdots, 0)$, $e_2 = (0, 1, 0, \cdots, 0)$, $\cdots$, $e_l = (0, \cdots, 0, 1)$ form an orthonormal system in $\mathbb{R}^l$. Any element $x \in \mathbb{R}^l$ can be expressed as a linear combination of $e_1, e_2, \cdots$, and e_l :

$$x = \sum_{i=1}^{l} x_i e_i. \tag{1.15}$$

Such an expression is unique. Is a similar expression possible in a general Hilbert space?

Definition 1.3 Let $\mathfrak{H}$ be a Hilbert space and $\{\varphi_1, \varphi_2, \cdots, \varphi_n, \cdots\}$ be an orthonormal system. For any $x \in \mathfrak{H}$, $\langle x, \varphi_n \rangle$ $(n = 1, 2, \cdots)$ is called the n-th **Fourier coefficient** of x with respect to $\{\varphi_n\}$, and

$$\sum_{n=1}^{\infty} \langle x, \varphi_n \rangle \, \varphi_n \tag{1.16}$$

is the **Fourier series** of x with respect to $\{\varphi_n\}$.

Of course, the series (1.16) may or may not be convergent. In any case, the expression

$$x \sim \sum_{n=1}^{\infty} \langle x, \ \varphi_n \rangle \, \varphi_n$$

tells us that (1.16) is the Fourier series of x with respect to $\{\varphi_n\}$.

Now we have to answer the basic question: does the series (1.16) converge to x in the norm, i.e.

$$\left\| x - \sum_{n=1}^{p} \langle x, \ \varphi_n \rangle \, \varphi_n \right\| \to 0 \quad \text{as} \quad p \to \infty?$$

The answer to this question is given by Theorem 1.6. However, we need some more preparation before we arrive at the answer.

In the case where the system of functions explained in Example 1.2 is chosen as an orthonormal system in $\mathfrak{L}^2([-\pi, \pi], \mathbb{C})$, the Fourier series of $f \in \mathfrak{L}^2([-\pi, \pi], \mathbb{C})$ is given by

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{1.17}$$

where

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx; \quad n = 0, 1, 2, \cdots,$$
$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx; \quad n = 1, 2, \cdots.$$

If the orthogonal system is as in Example 1.3, the Fourier series of $f \in \mathfrak{L}^2([-\pi, \pi], \mathbb{C})$ is given by

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}, \tag{1.18}$$

where

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx; \quad n = 0, \pm 1, \pm 2, \cdots .$$

This is called the **Fourier series in complex form**.[7]

We now proceed to investigate the meaning of Fourier coefficients from the viewpoint of approximation theory. Suppose that $\varphi_1, \varphi_2, \cdots, \varphi_n$ be an orthonormal system in a Hilbert space $\mathfrak{H}$. Consider a problem to approximate any element x of $\mathfrak{H}$ by a linear combination of φ_j's. We wish to find the coefficients $c_1, c_2, \cdots, c_n$ so as to minimize

$$\left\| x - \sum_{i=1}^{n} c_i \varphi_i \right\|.$$

Theorem 1.3 (Best approximation by Fourier coefficients) *Let $\varphi_1, \varphi_2, \cdots, \varphi_n$ be an orthonormal system in a Hilbert space $\mathfrak{H}$. For any $x \in \mathfrak{H}$, the inequality*

$$\left\| x - \sum_{i=1}^{n} \langle x,\ \varphi_i\rangle\ \varphi_i \right\| \leqq \left\| x - \sum_{i=1}^{n} c_i \varphi_i \right\|$$

holds good.

Proof

$$\begin{aligned} J &\equiv \left\| x - \sum_{i=1}^{n} c_i \varphi_i \right\|^2 \\ &= \langle x - \sum_{i=1}^{n} c_i \varphi_i,\ x - \sum_{i=1}^{n} c_i \varphi_i \rangle \\ &= \langle x,\ x\rangle - \sum_{i=1}^{n} c_i \langle \varphi_i,\ x\rangle - \sum_{i=1}^{n} \bar{c}_i \langle x,\ \varphi_i\rangle + \sum_{i,j=1}^{n} c_i \bar{c}_j \langle \varphi_i,\ \varphi_j\rangle \end{aligned} \tag{1.19}$$

[7] More rigorously, the Fourier coefficient corresponding to $(1/\sqrt{2\pi})e^{inx}$ is

$$\frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} f(x)e^{-inx}dx,$$

and the Fourier series is

$$\sum_{n=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} f(x)e^{-inx}dx \cdot \frac{1}{\sqrt{2\pi}}e^{inx}.$$

$$= \|x\|^2 - \sum_{i=1}^{n} c_i \langle \varphi_i,\ x\rangle - \sum_{i=1}^{n} \bar{c}_i \langle x,\ \varphi_i\rangle + \sum_{i=1}^{n} |\ c_i\ |^2 \qquad \text{(cf. (1.8) on page 5)}$$

$$= \|x\|^2 - \sum_{i=1}^{n} c_i \langle \varphi_i,\ x\rangle - \sum_{i=1}^{n} \bar{c}_i \langle x,\ \varphi_i\rangle + \sum_{i=1}^{n} |\ c_i\ |^2 + \sum_{i=1}^{n} \langle \varphi_i,\ x\rangle\langle x,\ \varphi_i\rangle$$

$$- \sum_{i=1}^{n} \langle \varphi_i,\ x\rangle\langle x,\ \varphi_i\rangle$$

$$= \|x\|^2 + \sum_{i=1}^{n} |\ c_i - \langle x,\ \varphi_i\rangle\ |^2 - \sum_{i=1}^{n} |\ \langle x,\ \varphi_i\rangle\ |^2 .$$

Hence J attains its minimum when $c_i = \langle x, \varphi_i\rangle$ $(i = 1, 2, \cdots, n)$. □

According to Theorem 1.2, there exists an orthonormal system $\{\varphi_1, \varphi_2, \cdots, \varphi_n, \cdots\}$ consisting of countably infinite number of vectors in a Hilbert space of infinite dimension. If we substitute $\langle x, \varphi_i\rangle$ as c_i $(i = 1, 2, \cdots, n)$ in (1.19), then

$$\|x\|^2 - \sum_{i=1}^{n} |\ \langle x, \varphi_i\rangle\ |^2 \geqq 0$$

since $J \geqq 0$. Passing to the limit as $n \to \infty$, we have an important result.

Theorem 1.4 (Bessel's inequality) *If $\mathfrak{H}$ is a Hilbert space of infinite dimension and $\{\varphi_1, \varphi_2, \cdots, \varphi_n, \cdots\}$ is an orthonormal system in $\mathfrak{H}$, then we have*

$$\sum_{i=1}^{\infty} |\ \langle x,\ \varphi_n\rangle\ |^2 \leqq \|x\|^2 .$$

Corollary 1.1 (Riemann–Lebesgue lemma)

$$\lim_{n\to\infty} \langle x,\ \varphi_n\rangle = 0.$$

1.4 Completeness of Orthonormal Systems

We now examine the conditions under which $x \in \mathfrak{H}$ can be represented in the form of Fourier series based upon these preparations. A sufficient number of vectors must be included in an orthonormal system in order to represent x as a linear combination of vectors in an orthonormal system even in the case of an l-dimensional Euclidean space. In this sense, the menu of the orthonormal system is required to be sufficiently rich in order to represent $x \in \mathfrak{H}$ (Hilbert space) in the form of Fourier series.

Definition 1.4 Let Φ be an orthonormal system in a Hilbert space $\mathfrak{H}$. Φ is said to be **complete (as an orthonormal system)** if there is no orthonormal system which contains Φ as a proper subset.

An orthonormal system Φ in a Hilbert space $\mathfrak{H}$ is complete if and only if

$$\langle x,\ \varphi\rangle = 0 \quad \text{for all} \quad \varphi \in \Phi \Longrightarrow x = 0.$$

(Proof is almost obvious.)

All the orthonormal systems shown in Examples 1.1–1.6 above are complete. Here we shall only show the completeness of [8]

$$\Phi = \left\{ \frac{1}{\sqrt{2\pi}} e^{inx} \ ;\ n = 0, \pm 1, \pm 2, \cdots \right\}$$

in $\mathfrak{L}^2([-\pi, \pi],\ \mathbb{C})$ (in Example 1.3).

By Theorem 1.3, we obtain

$$\left\| f - \sum_{j=-n}^{n} \frac{1}{\sqrt{2\pi}} c_j e^{ijx} \right\|_2 \geqq \left\| f - \sum_{j=-n}^{n} \frac{1}{\sqrt{2\pi}} \alpha_j e^{ijx} \right\|_2 = \| f \|_2^2 - \sum_{j=-n}^{n} \frac{1}{2\pi} \mid \alpha_j \mid^2, \tag{1.20}$$

for any $f \in \mathfrak{L}^2([-\pi, \pi], \mathbb{C})$ and any complex numbers c_j $(j = 0, \pm 1, \pm 2, \cdots)$, where $\|\cdot\|_2$ is the $\mathfrak{L}^2$-norm and α_j's are Fourier coefficients:

$$\alpha_j = \left\langle f\ ,\ \frac{1}{\sqrt{2\pi}} e^{ijx} \right\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) e^{-ijx} dx.$$

If f is continuous with $f(-\pi) = f(\pi)$, in particular, the left-hand side of (1.20) can be arbitrarily small by choosing c_j and n in a suitable manner (Weierstrass approximation theorem). Hence the subspace span Φ which is spanned by Φ is $\mathfrak{L}^2$-dense in

$$\mathfrak{M} = \{ f \in \mathfrak{L}^2([-\pi,\ \pi]\ ,\ \mathbb{C}) |\ f \text{ is continuous},\ f(-\pi) = f(\pi) \}.$$

Since $\mathfrak{M}$ is dense in $\mathfrak{L}^2([-\pi,\ \pi],\ \mathbb{C})$, span Φ is dense in $\mathfrak{L}^2([-\pi,\ \pi],\ \mathbb{C})$. Consequently, the left-hand side of (1.20) can be arbitrarily small by suitable choice of c_j's. Thus if

$$\alpha_j = \left\langle f, \frac{1}{\sqrt{2\pi}} e^{ijx} \right\rangle = 0 \quad \text{for all} \quad j = 0, \pm 1, \cdots$$

for $f \in \mathfrak{L}^2([-\pi,\ \pi],\ \mathbb{C})$, then $f = 0$. This proves the completeness of Φ.[9]

[8] See Folland [3] Chap. 6 and Kawata [5] pp. 33–36 for other orthonormal systems.

[9] We acknowledge Yosida [12] p. 88 for the proof here.

The next theorem confirms the existence of a complete orthonormal system in a general Hilbert space.

Theorem 1.5 (Existence of orthonormal system) *For any orthonormal system Φ in a Hilbert space, there exists a complete orthonormal system which contains Φ.*

Proof We denote by $\mathcal{O}$ the family of all the orthonormal systems in $\mathfrak{H}$ which contain Φ as a subset. Define a partial ordering $\prec$ on $\mathcal{O}$ by

$$\Phi \prec \Phi' \Longleftrightarrow \Phi \subset \Phi'.$$

Let $\{\Phi_\alpha\}$ be a chain (i.e. totally ordered subfamily) of $\mathcal{O}$. Then we must have

$$\cup_\alpha \Phi_\alpha \in \mathcal{O} \quad \text{and} \quad \Phi_\alpha \prec \cup_\alpha \Phi_\alpha \quad \text{for all } \alpha.$$

Hence the partial ordering structure $(\mathcal{O}, \prec)$ is inductive. Zorn's lemma confirms the existence of a maximal element in $\mathcal{O}$ with respect to $\prec$. □

We can now state and prove the basic theorem concerning the Fourier expansion of $x \in \mathfrak{H}$.

Theorem 1.6 (Fundamental theorem of Fourier series expansion) *The following five statements for an orthonormal system $\Phi = \{\varphi_1, \varphi_2, \cdots, \varphi_n, \cdots\}$ in a Hilbert space $\mathfrak{H}$ are equivalent to each other:*

(i) *Φ is complete.*
(ii) *For any $x, y \in \mathfrak{H}$,*

$$\langle x, \varphi_n \rangle = \langle y, \varphi_n \rangle \quad \text{for all } n \Longrightarrow x = y.$$

(iii) *Every $x \in \mathfrak{H}$ can be represented as*

$$x = \sum_{n=1}^{\infty} \langle x, \varphi_n \rangle \, \varphi_n,$$

in the sense that

$$\lim_{p\to\infty} \left\| x - \sum_{n=1}^{p} \langle x, \varphi_n \rangle \varphi_n \right\| = 0.$$

(iv) *For any $x, y \in \mathfrak{H}$,*

$$\langle x, y \rangle = \sum_{n=1}^{\infty} \alpha_n \overline{\beta_n},$$

where

$$\alpha_n = \langle x, \varphi_n \rangle, \quad \beta_n = \langle y, \varphi_n \rangle; \quad n = 1, 2, \cdots .$$

(v) *For any* $x \in \mathfrak{H}$.

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x , \varphi_n \rangle|^2 . \qquad \textit{(}\textbf{Parseval's equality}\textit{)}$$

Most of the proof seems to be easy. So we show only the step (ii) $\Longrightarrow$ (iii) in detail.

By Bessel's inequality, we obtain

$$\sum_{n=1}^{p} |\langle x, \varphi_n \rangle|^2 \leqq \|x\|^2 < \infty$$

for any $x \in \mathfrak{H}$. If we denote by S_p a partial sum of the Fourier series, i.e.

$$S_p = \sum_{n=1}^{p} \langle x, \varphi_n \rangle \varphi_n ,$$

it follows that ($p < q$)

$$\|S_q - S_p\|^2 = \sum_{n=p+1}^{q} |\langle x, \varphi_n \rangle|^2 \longrightarrow 0 \quad \text{as} \quad p, q \longrightarrow \infty.$$

Hence the sequence $\{S_p\}$ is Cauchy in $\mathfrak{H}$, and so it has a limit $y \in \mathfrak{H}$, i.e.

$$y = \sum_{n=1}^{\infty} \langle x, \varphi_n \rangle \varphi_n .$$

Since

$$\langle y, \varphi_n \rangle = \langle x, \varphi_n \rangle \quad \text{for all} \quad n$$

for this y, (ii) implies that $x = y$. This proves (iii).

However, we have to make sure that we do not have a positive answer to the question: "Is there any complete orthonormal system which consists of countable vectors?" If the answer is yes, any vector can be expanded by its Fourier series. The next theorem fills this gap, fortunately.

Theorem 1.7 (orthonormal system consisting of countable vectors) *If a Hilbert space $\mathfrak{H}$ is separable, there exists a complete orthonormal system which consists of countable vectors.*

Proof The theorem is clear if $\dim \mathfrak{H} < \infty$, so we may assume $\dim \mathfrak{H} = \infty$ without loss of generality.

Let $D = \{x_1, x_2, \cdots\}$ be a countable dense subset of $\mathfrak{H}$. We denote by φ_1 the first element of D which is not zero (that is, the nonzero element with the smallest suffix). Next let φ_2 be the first element of D which does not belong to span $\{\varphi_1\}$. Repeating this process, we define φ_n as the first element of D which does not belong to $\text{span}\{\varphi_1, \varphi_2, \cdots, \varphi_{n-1}\}$. Since we assume that $\dim \mathfrak{H} = \infty$, we get a linearly independent family $\{\varphi_1, \varphi_2, \cdots, \varphi_n, \cdots\}$ consisting of countable (distinct) vectors. Then we can prove that

$$\overline{\text{span}\{\varphi_n; n = 1, 2, \cdots\}} = \overline{D} = \mathfrak{H}. \tag{1.21}$$

Assume that (1.21) is not true. Then there must exist some $x_p \in D$ such that

$$x_p \in \mathfrak{H} \setminus \overline{\text{span}\{\varphi_n\}}.$$

x_p is, of course, different from φ_1. Since x_p is not chosen as φ_{n+1} for any n, x_p is not the first element of D which does not belong to $\text{span}\{\varphi_1, \varphi_2, \cdots, \varphi_n\}$. Hence there is no element of $\{x_p, x_{p+1}, \cdots\}$ which is chosen as φ_{n+1}. This holds good for all n. So $\{\varphi_1, \varphi_2, \cdots\}$ must be chosen from $\{x_1, x_2, \cdots, x_{p-1}\}$ (elements of D whose suffixes are smaller than p). But it contradicts the fact that $\{\varphi_1, \varphi_2, \cdots\}$ has countably infinite distinct elements. Thus we can conclude that (1.21) holds good.

Applying the Schmidt orthogonalization (Theorem 1.2) to $\{\varphi_n\}$, we obtain an orthonormal system $\{\widetilde{\varphi}_n\}$ in $\mathfrak{H}$.

Finally, we proceed to show the completeness of $\{\widetilde{\varphi}_n\}$. Since each $\widetilde{\varphi}_n$ is a linear combination of $\{\varphi_1, \varphi_2, \cdots, \varphi_n\}$ and each φ_n is a linear combination of $\{\widetilde{\varphi}_1, \widetilde{\varphi}_2, \cdots, \widetilde{\varphi}_n\}$ (as remarked in the proof of Theorem 1.2), we have

$$\text{span}\{\varphi_n; n = 1, 2, \cdots\} = \text{span}\{\widetilde{\varphi}_n; n = 1, 2, \cdots\}$$

and both of these sets are dense in $\mathfrak{H}$. If we assume that $\{\widetilde{\varphi}_n\}$ is not complete, there must exist some nonzero $x \in \mathfrak{H}$ such that

$$\langle x, \widetilde{\varphi}_n \rangle = 0 \quad \text{for all} \ \ n. \tag{1.22}$$

Since $\text{span}\{\widetilde{\varphi}_n\}$ is dense in $\mathfrak{H}$, there exists some sequence $\{\xi_p\}$ of finite linear combinations of $\{\widetilde{\varphi}_n\}$ such that $\xi_p \to x$ (as $p \to \infty$). It follows from (1.22) that

$$\langle x, \xi_p \rangle = 0 \quad \text{for all} \ \ p.$$

Passing to the limit, we must have

$$\lim_{p\to\infty} \langle x, \xi_p \rangle = \langle x, x \rangle = \|x\|^2 = 0,$$

which contradicts $x \neq 0$. Thus we can conclude that $\{\widetilde{\varphi}_n\}$ is complete. □

Corollary 1.2 (separable Hilbert space $\cong$ l_2) *Any separable Hilbert space is isomorphic to l_2.*

Proof Let $\{\varphi_n\}$ be a complete orthonormal system which consists of countable vectors. (The existence of such a system is guaranteed by Theorem 1.7.) Thanks to Theorem 1.6, we have

$$x = \sum_{n=1}^{\infty} \langle x, \varphi_n \rangle \, \varphi_n$$

and

$$\|x\|^2 = \sum_{n=1}^{\infty} |\, \langle x, \varphi_n \rangle \,|^2 < \infty.$$

If we define an operator T by

$$T : x \longmapsto \{\langle x, \varphi_1 \rangle, \langle x, \varphi_2 \rangle, \cdots\}, \quad x \in \mathfrak{H},$$

T is an isometric linear operator of the form

$$T : \mathfrak{H} \longrightarrow l_2.$$

Since T is isometric, it is injective.[10]

Furthermore, T is surjective. In order to show it, let $\{\alpha_n\}$ be any element of l_2. Since

$$S_p = \sum_{n=1}^{p} \alpha_n \varphi_n; \quad p = 1, 2, \cdots$$

is a Cauchy sequence, there exists some $x \in \mathfrak{H}$ such that

$$x = \sum_{n=1}^{\infty} \alpha_n \varphi_n$$

[10]The injectivity of T is also verified by Theorem 1.6 (ii).

and

$$\alpha_n = \langle x, \varphi_n \rangle; \quad n = 1, 2, \cdots .$$

This proves the surjectivity of T.

We conclude that $T : \mathfrak{H} \to l_2$ is an isometric isomorphism. □

Finally, we would like to add some remarks.

Remark 1.1

1° Any separable Hilbert space has a complete orthonormal system which consists of countable vectors. Hence any element can be expanded by Fourier series by using this system. However, the situation is different in the case of a general Hilbert space, because it may be impossible to construct a complete system consisting of countable vectors. (See Dudley [2] pp. 126–131, Yosida [12] pp. 86–88.)

2° How can we compute Fourier coefficients when we consider $\mathfrak{L}^2([-l, l], \mathbb{C})$ instead of $\mathfrak{L}^2([-\pi, \pi], \mathbb{C})$ as a concrete Hilbert space? If we change the variable $t \in [-l, l]$ to $x \in [-\pi, \pi]$ by

$$x = \frac{\pi t}{l}, \quad \text{i.e.} \;\; t = \frac{lx}{\pi},$$

$f(t) \in \mathfrak{L}^2([-l, l], \mathbb{C})$ can be transformed to a function $\tilde{f}(x) \in \mathfrak{L}^2([-\pi, \pi], \mathbb{C})$ by

$$\tilde{f}(x) \equiv f\left(\frac{lx}{\pi}\right).$$

The Fourier coefficients of $\tilde{f}$ are given by

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} \tilde{f}(x) \cos nx \; dx \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{lx}{\pi}\right) \cos nx \; dx \\ &= \frac{1}{\pi} \int_{-l}^{l} f(t) \cos \frac{n\pi t}{l} \cdot \frac{\pi}{l} \, dt \\ &= \frac{1}{l} \int_{-l}^{l} f(t) \cos \frac{n\pi t}{l} \, dt \; ; \;\; n = 0, 1, 2, \cdots . \end{aligned}$$

Similarly, we have

$$b_n = \frac{1}{l}\int_{-l}^{l} f(t)\sin\frac{n\pi t}{l}dt; \quad n = 1, 2, \cdots .$$

Thus the Fourier series of f is given in the form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}\Big(a_n \cos\frac{n\pi t}{l} + b_n \sin\frac{n\pi t}{l}\Big).$$

References

1. Cartan, H.: Théorie élémentaires des fonctions analytiques d'une ou plusieurs variables complexes. Hermann, Paris (1961) (English edn.) Elementary Theory of Analytic Functions of One or Several Complex Variables. Addison Wesley, Reading (1963)
2. Dudley, R.M.: Real Analysis and Probability. Wadsworth and Brooks, Pacific Grove (1988)
3. Folland, G.B.: Fourier Analysis and its Applications. American Mathematical Society, Providence (1992)
4. Halmos, P.R.: Introduction to a Hilbert Space and the Theory of Spectral Multiplicity. Chelsea, New York (1951)
5. Kawata, T.: Fourier Kaiseki (Fourier Analysis). Sangyo Tosho, Tokyo (1975) (Originally published in Japanese)
6. Lax, P.D.: Functional Analysis. Wiley, New York (2002)
7. Maruyama, T.: Kansu Kaisekigaku (Functional Analysis). Keio Tsushin, Tokyo (1980) (Originally published in Japanese)
8. Schwartz, L.: Analyse hilbertienne. Hermann, Paris (1979)
9. Takagi, T.: Kaiseki Gairon (Treatise on Analysis), 3rd edn. Iwanami Shoten, Tokyo (1961) (Originally published in Japanese)
10. Terasawa, K.: Shizen Kagakusha no tame no Sugaku Gairon (Treatise on Mathematics for Natural Scientists). Iwanami Shoten, Tokyo (1954) (Originally published in Japanese)
11. Yosida, K.: Lectures on Differential and Integral Equations. Interscience, New York (1960)
12. Yosida, K.: Functional Analysis, 3rd edn. Springer, New York (1971)

Chapter 2
Convergence of Classical Fourier Series

We have discussed basic contents of the theory of Fourier series on a general Hilbert space. We now proceed to the classical problem concerning the Fourier series expansion of an integrable function with respect to the trigonometric functions. If we choose $\mathfrak{L}^2([-\pi, \pi], \mathbb{C})$ as a Hilbert space and

$$\frac{1}{\sqrt{2\pi}}, \ \frac{1}{\sqrt{\pi}}\cos x, \ \frac{1}{\sqrt{\pi}}\sin x, \ \cdots, \ \frac{1}{\sqrt{\pi}}\cos nx, \ \frac{1}{\sqrt{\pi}}\sin nx, \ \cdots; \\ n = 1, 2, \cdots$$

as a complete orthonormal system, the Fourier series of $f \in \mathfrak{L}^2([-\pi, \pi], \mathbb{C})$ is given in the form

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

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx \, dx, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx \, dx.$$

This Fourier series converges to f in $\mathfrak{L}^2$-norm.

However, it is necessary for us to investigate convergence criteria of Fourier series other than $\mathfrak{L}^2$-convergence: pointwise convergence, almost everywhere convergence, uniform convergence, and so on. This chapter is devoted to these important topics.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_2

2.1 Dirichlet Integrals

Let $f : \mathbb{R} \to \mathbb{C}$ (or $\mathbb{R}$) be a 2π-periodic function which is integrable on $[-\pi, \pi]$. We are now following the traditional setting rather than the framework of a Hilbert space. However,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx$$

make sense for any $f \in \mathfrak{L}^1([-\pi, \pi], \mathbb{C})$ by Hölder's inequality. Therefore we can consider the formal series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{2.1}$$

for $f \in \mathfrak{L}^1([-\pi, \pi], \mathbb{C})$ in a way similar to that for $\mathfrak{L}^2([-\pi, \pi], \mathbb{C})$. We call (2.1) the Fourier series of f, although $\mathfrak{L}^1([-\pi, \pi], \mathbb{C})$ is not a Hilbert space.

We start by investigating the convergence of (2.1) at a specific point $x \in [-\pi, \pi]$. It is the so-called Dirichlet integral which provides us with an effective tool. A partial sum $S_n(x)$ of (2.1) at x can be computed as follows:

$$\begin{aligned} S_n(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \cos 0 \cdot t \, dt + \sum_{k=1}^{n} \Big(\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos kt \, dt \Big) \cos kx \\ &\qquad + \sum_{k=1}^{n} \Big(\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin kt \, dt \Big) \sin kx \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} f(t) dt + \frac{1}{\pi} \sum_{k=1}^{n} \int_{-\pi}^{\pi} f(t) \cos kt \cos kx \, dt \\ &\qquad + \frac{1}{\pi} \sum_{k=1}^{n} \int_{-\pi}^{\pi} f(t) \sin kt \sin kx \, dt \qquad (2.2) \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \Big\{ \frac{1}{2} + \sum_{k=1}^{n} \cos k(t - x) \Big\} dt. \end{aligned}$$

We have to know the formula

$$\frac{1}{2} + \cos u + \cos 2u + \cdots + \cos nu = \frac{\sin \dfrac{2n+1}{2} u}{2 \sin \dfrac{u}{2}} \tag{2.3}$$

(assuming $\sin(u/2) \neq 0$).

In fact, summing up each side of

$$\begin{aligned}
&\sin\frac{u}{2} = \frac{1}{2}\cdot 2\sin\frac{u}{2},\\
&\sin\frac{3u}{2} - \sin\frac{u}{2} = \cos u\cdot 2\sin\frac{u}{2},\\
&\dots\dots\dots\dots\dots\dots\\
&\sin\frac{2n+1}{2}u - \sin\frac{2n-1}{2}u = \cos nu\cdot 2\sin\frac{u}{2},
\end{aligned}$$

we obtain

$$\sin\frac{2n+1}{2}u = 2\sin\frac{u}{2}\left(\frac{1}{2} + \cos u + \cdots + \cos nu\right),$$

from which (2.3) immediately follows.

We obtain, by (2.2) and (2.3), that

$$S_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\frac{\sin\frac{2n+1}{2}(t-x)}{2\sin\frac{t-x}{2}}dt. \tag{2.4}$$

The formula (2.4) can be rewritten in the form

$$S_n(x) = \frac{1}{\pi}\int_{-\pi-x}^{\pi-x} f(x+z)\frac{\sin\frac{2n+1}{2}z}{2\sin\frac{z}{2}}dz \tag{2.5}$$

by changing variables: $z = t - x$. Since the entire integrand of (2.5) is 2π-periodic,

$$S_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+z)\frac{\sin\frac{2n+1}{2}z}{2\sin\frac{z}{2}}dz. \tag{2.6}$$

The integral appearing in (2.6) is called the **Dirichlet integral**, and the function $D_n(z)$ defined by

$$D_n(z) \equiv \frac{1}{2\pi}\cdot\frac{\sin\frac{2n+1}{2}z}{\sin\frac{z}{2}} \tag{2.7}$$

is called the **Dirichlet kernel**. Then $S_n(x)$ can be expressed as

$$S_n(x) = \int_{-\pi}^{\pi} f(x+z)D_n(z)dz. \tag{2.8}$$

Remark 2.1 Of course, the integrand of (2.4) can not be defined at t such that $2^{-1}(t-x) = m\pi$ $(m \in \mathbb{Z})$ because $\sin 2^{-1}(t-x) = 0$ for such t. However, the set $\{t \in [-\pi, \pi] | 2^{-1}(t-x) = m\pi, m \in \mathbb{Z}\}$ is of measure zero, so it does not influence the integral. Rigorously speaking, the Dirichlet integral (2.7) can not be defined at z such that $z/2 = m\pi$ $(m \in \mathbb{Z})$. But we may assign any values at these points.

It is clear by (2.3) that

$$\int_{-\pi}^{\pi} D_n(z)dz = 1. \tag{2.9}$$

Hence by (2.8) and (2.9), the relation

$$S_n(x) - c = \int_{-\pi}^{\pi} (f(x+z) - c)D_n(z)dz \tag{2.10}$$

holds good for any c. In particular, it is obvious that

$$S_n(x) - f(x) = \int_{-\pi}^{\pi} (f(x+z) - f(x))D_n(z)dz \tag{2.10$'$}$$

for $c = f(x)$. Since $D_n(z)$ is an even function, $S_n(x)$ can be reexpressed as

$$S_n(x) = \int_0^{\pi} [f(x+z) + f(x-z)]D_n(z)dz = 2\int_0^{\pi} \phi(z)D_n(z)dz,$$

where

$$\phi(z) = \frac{1}{2}\big[f(x+z) + f(x-z)\big].$$

Consequently, we have

$$S_n(x) - c = 2\int_0^{\pi} [\phi(z) - c]D_n(z)dz \tag{2.11}$$

for any c. In the special case of $c = f(x)$, (2.11) becomes

$$S_n(x) - f(x) = 2\int_0^{\pi} \left\{\frac{f(x+z) + f(x-z)}{2} - f(x)\right\} D_n(z)dz. \tag{2.11$'$}$$

Thus we obtain the following theorem.

Theorem 2.1 (convergence at a point) *Let $f : \mathbb{R} \to \mathbb{C}$ (or $\mathbb{R}$) be a 2π-periodic function which is integrable on $[-\pi, \pi]$. The sequence of partial sums $S_n(x)$ of the Fourier series (2.1) converges to $f(x)$ at a point x if and only if the integral on the right-hand side of (2.10′) or (2.11′) converges to* 0.

Of course, $S_n(x)$ converges to c if and only if the integral on the right-hand side of (2.10) or (2.11) converges to 0.

Similar results also follow when the family of functions

$$\frac{1}{\sqrt{2\pi}} e^{inx}; \quad n = 0, \pm 1, \pm 2, \cdots$$

is chosen as an orthonormal system. Let $f : \mathbb{R} \to \mathbb{C}$ be 2π-periodic and integrable on $[-\pi, \pi]$. Consider the formal Fourier series

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}, \tag{2.1′}$$

where

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt; \quad n = 0, \pm 1, \pm 2, \cdots .$$

Computing the partial sums $S_n(x)$, we have

$$S_n(x) = \frac{1}{2\pi} \sum_{k=-n}^{n} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt \cdot e^{ikx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{k=-n}^{n} e^{ik(x-t)} dt. \tag{2.2′}$$

The sum appearing in (2.2′) can be rewritten as

$$\sum_{k=-n}^{n} e^{ik(x-t)} = \sum_{k=-n}^{n} e^{-ikz} = \sum_{k=-n}^{n} e^{ikz}$$

by changing variables: $z = t - x$. This is exactly equal to $2\pi \cdot D_n(z)$, where $D_n(z)$ is the Dirichlet kernel (2.7); i.e.,

$$\begin{aligned} \frac{1}{2\pi} \sum_{k=-n}^{n} e^{ikz} &= \frac{1}{2\pi} \frac{e^{-inz} - e^{inz+iz}}{1 - e^{iz}} \\ &= \frac{1}{2\pi} \cdot \frac{e^{-i(\frac{2n+1}{2})z} - e^{i(\frac{2n+1}{2})z}}{e^{-iz/2} - e^{iz/2}} \end{aligned} \tag{2.7′}$$

$$= \frac{1}{2\pi} \cdot \frac{\sin \frac{2n+1}{2} z}{\sin \frac{z}{2}}$$

$$= D_n(z).$$

Hence the partial sums (2.2′) of the Fourier series in complex form can be expressed as

$$S_n(x) = \int_{-\pi}^{\pi} f(x+z) D_n(z) dz. \tag{2.8′}$$

We can verify, by a similar argument, that Theorem 2.1 also holds good for the convergence of Fourier series in complex form.

2.2 Dini, Jordan Tests

We now provide more convenient criteria of convergence of a Fourier series. The next lemma establishes a result analogous to the Riemann–Lebesgue lemma (Corollary 1.1, p. 14) for a Fourier series on a Hilbert space in the case of $\mathfrak{L}^1$.

Lemma 2.1 *For any* $\varphi \in \mathfrak{L}^1([a, b], \mathbb{R})$*,*

$$\lim_{p \to \infty} \int_a^b \varphi(x) \sin px dx = 0, \qquad \lim_{p \to \infty} \int_a^b \varphi(x) \cos px dx = 0,$$
$$\lim_{p \to \infty} \int_a^b \varphi(x) e^{ipx} dx = 0.$$

Proof It is enough for us to check only the first one. We can prove the second quite similarly. And the last one immediately follows from the other two.

Assume, in addition, that φ is of the class $\mathfrak{C}^1$ for the moment. Then we obtain, by integration by parts, that

$$\int_a^b \varphi(x) \sin px dx = -\varphi(x) \frac{\cos px}{p} \Big|_a^b + \int_a^b \varphi'(x) \frac{\cos px}{p} dx$$
$$= -\varphi(b) \frac{\cos pb}{p} + \varphi(a) \frac{\cos pa}{p} + \int_a^b \varphi'(x) \frac{\cos px}{p} dx \longrightarrow 0$$
$$\text{as} \quad p \longrightarrow \infty.$$

We depend upon the continuity of φ' (and so the boundedness of $\varphi'(x) \cdot \cos px$) to conclude that the third term tends to 0.

We now consider the general case: $\varphi \in \mathfrak{L}^1([a,b],\mathbb{R})$. Since $\mathfrak{C}^1([a,b],\mathbb{R})$ is dense in $\mathfrak{L}^1([a,b],\mathbb{R})$, there exists, for any $\varepsilon > 0$, some $\varphi_\varepsilon \in \mathfrak{C}^1$ such that $\|\varphi - \varphi_\varepsilon\|_1 < \varepsilon/2$. ($\|\cdot\|_1$ denotes the $\mathfrak{L}^1$-norm.) Applying the above result to this φ_ε, we must have

$$\left|\int_a^b \varphi_\varepsilon(x)\sin px\ dx\right| < \frac{\varepsilon}{2}$$

for sufficiently large p. Hence we obtain the following evaluation for large p:

$$\begin{aligned}
&\left|\int_a^b \varphi(x)\sin px\ dx\right| \\
&\qquad \leqq \left|\int_a^b [\varphi(x)-\varphi_\varepsilon(x)]\sin px\ dx\right| + \left|\int_a^b \varphi_\varepsilon(x)\sin px\ dx\right| \\
&\qquad \leqq \|\ \varphi - \varphi_\varepsilon\ \|_1 \cdot \|\ \sin px\ \|_\infty + \frac{\varepsilon}{2} \\
&\qquad\qquad\qquad (\text{Hölder's inequality, } \|\cdot\|_\infty \text{ is the essential sup norm}) \\
&\qquad \leqq \|\ \varphi - \varphi_\varepsilon\ \|_1 \cdot 1 + \frac{\varepsilon}{2} \\
&\qquad \leqq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&\qquad = \varepsilon.
\end{aligned}$$

□

The same result is also valid for $\varphi \in \mathfrak{L}^1([a,b],\mathbb{C})$.[1]

Corollary 2.1 *For any $\varphi \in \mathfrak{L}^1([-\pi,\pi],\mathbb{C})$, the Fourier coefficients $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ tend to 0 as $n \to \infty$.*

We now proceed to the convergence problem of a Fourier series at a point. Assumptions imposed on f are still the same.

Making use of (2.11) in the previous section and writing $\xi(z) = \phi(z) - c$, we have

$$\begin{aligned}
S_n(x) - c &= 2\int_0^\pi \xi(z) D_n(z)\ dz \\
&= \frac{1}{\pi}\int_0^\pi \frac{\sin \dfrac{2n+1}{2} z}{\sin \dfrac{z}{2}} \xi(z)\ dz
\end{aligned}$$

[1] We owe this result to Kolmogorov–Fomin [14] pp. 387–388. For more general results, see Kawata [11] pp. 63–64.

$$= \frac{1}{\pi}\int_0^\delta \frac{\sin\dfrac{2n+1}{2}z}{\dfrac{z}{2}}\xi(z)\,dz + \frac{1}{\pi}\int_\delta^\pi \frac{\sin\dfrac{2n+1}{2}z}{\sin\dfrac{z}{2}}\xi(z)\,dz \tag{2.12}$$

$$+ \frac{1}{\pi}\int_0^\delta \left(\sin\frac{2n+1}{2}z\right)\left(\frac{1}{\sin\dfrac{z}{2}} - \frac{1}{\dfrac{z}{2}}\right)\xi(z)\,dz$$

$$\equiv I_1 + I_2 + I_3 \quad (0 < \delta < \pi).$$

It can be readily verified that $I_2 \to 0$ as $n \to \infty$. To see this, we have to observe that

$$\int_\delta^\pi \sin\left(nz + \frac{z}{2}\right)\frac{\xi(z)}{\sin\dfrac{z}{2}}dz$$

$(\xi(z)(\sin(z/2))^{-1}$ is integrable on $[\delta, \pi]$, and so this integral makes sense)

$$= \int_\delta^\pi \sin nz \cos\frac{z}{2}\,\frac{\xi(z)}{\sin\dfrac{z}{2}}\,dz + \int_\delta^\pi \cos nz \sin\frac{z}{2}\,\frac{\xi(z)}{\sin\dfrac{z}{2}}\,dz,$$

where

$$\cos\frac{z}{2}\frac{\xi(z)}{\sin\dfrac{z}{2}}, \qquad \sin\frac{z}{2}\frac{\xi(z)}{\sin\dfrac{z}{2}}$$

are both integrable on $[\delta, \pi]$. Hence, by Lemma 2.1, $I_2 \to 0$ as $n \to \infty$.

How about I_3? Since

$$\frac{1}{\sin\dfrac{z}{2}} - \frac{1}{\dfrac{z}{2}} = \frac{\dfrac{z}{2} - \sin\dfrac{z}{2}}{\dfrac{z}{2}\sin\dfrac{z}{2}} = \frac{\dfrac{z}{2} - \left\{\dfrac{z}{2} + O(z^3)\right\}}{\dfrac{z}{2}\sin\dfrac{z}{2}} = O(z) \qquad (\text{as } z \to 0),$$

we obtain

$$\left(\frac{1}{\sin\dfrac{z}{2}} - \frac{1}{\dfrac{z}{2}}\right)\xi(z) \;\in\; \mathfrak{L}^1\big([0,\ \delta],\ \mathbb{C}\big).$$

So, again by Lemma 2.1, we conclude that $I_3 \to 0$ as $n \to \infty$.

There remains only to check I_1.

Theorem 2.2 (convergence at a point) *Let $f : \mathbb{R} \to \mathbb{C}$ (or $\mathbb{R}$) be a function which is 2π-periodic and integrable on $[-\pi, \pi]$. The sequence of partial sums of the Fourier series of f converges to some number c at a point x if and only if*

$$\int_0^\delta \frac{\sin \frac{2n+1}{2} z}{z} \xi(z)\, dz \;\to\; 0 \;(as\; n \to \infty)$$

for some $\delta \in (0, \pi)$, where

$$\xi(z) = \frac{1}{2}\big[f(x+z) + f(x-z)\big] - c.$$

Thus the answer to the question whether a Fourier series converges or not at a point x depends upon the behavior of f in a neighborhood $(x - \delta,\ x + \delta)$ of x (δ may be arbitrarily small). We call this the **local property** of the convergence of a Fourier series.

The next theorem immediately follows from Theorem 2.2 and Lemma 2.1.

Theorem 2.3 (Dini test) *Let $f : \mathbb{R} \to \mathbb{C}$ (or $\mathbb{R}$) be a function which is 2π-periodic and integrable on $[-\pi, \pi]$. We write $\xi(z) = (1/2)[f(x+z) + f(x-z)] - c$ for some number c. If*

$$\int_0^\delta \frac{\mid \xi(z) \mid}{z}\, dz \;<\; \infty \tag{2.13}$$

for some $\delta > 0$, then $S_n(x) \to c$ as $n \to \infty$.

The proof is almost obvious. The condition (2.13) is called the **Dini condition**.

Remark 2.2 It is easy to check that the Dini condition is satisfied if and only if

$$\int_{-\delta}^\delta \left|\frac{f(x+z) - c}{z}\right| dz \;<\; \infty \tag{2.14}$$

for some $\delta > 0$. (The assumptions imposed on f are the same as above.)[2]

We state another convergence criterion due to Jordan.

Theorem 2.4 (Jordan test) *Let $f : \mathbb{R} \to \mathbb{C}$ (or $\mathbb{R}$) be a function which is 2π-periodic and integrable on $[-\pi, \pi]$. If f is of bounded variation on some neighborhood $[x - \delta,\ x + \delta]$ ($\delta > 0$), then*

$$S_n(x) \longrightarrow \frac{1}{2}\big[f(x+0) + f(x-0)\big] \quad (as\; n \to \infty).$$

---

[2]The requirement (2.14) is called the Dini condition in Kolmogorov–Fomin [14], Vol.II, p. 388.

Lemma 2.2 *There exists some constant C such that*

$$\Big|\int_a^b \frac{\sin t}{t}dt\Big| \leqq C \tag{2.15}$$

for any $a < b$. (C does not depend upon a and b.)

Proof We assume first that $0 \leqq a < b$.

The case $1 \leqq a < b$: There exists some $\eta \in [a, b]$ which satisfies

$$\int_a^b \frac{\sin t}{t}dt = \frac{1}{a}\int_a^\eta \sin t\ dt = \frac{1}{a}(\cos a - \cos \eta)$$

by the second mean-valued theorem of integration.[3] Hence

$$\Big|\int_a^b \frac{\sin t}{t}dt\Big| \leqq 2.$$

The case $0 \leqq a < b \leqq 1$:

$$0 \leqq \int_a^b \frac{\sin t}{t}dt \leqq \int_a^b dt \leqq 1.$$

The case $0 \leqq a \leqq 1 < b$:

$$\Big|\int_a^b \frac{\sin t}{t}dt\Big| \leqq \Big|\int_a^1\Big| + \Big|\int_1^b\Big| \leqq 1 + 2 = 3.$$

[3]**Weierstrass-type**. Let $g : [a, b] \to \mathbb{R}$ be integrable and $h : [a, b] \to \mathbb{R}$ be bounded and monotone. Then there exists some $\eta \in [a, b]$ such that

$$\int_a^b g(x)h(x)dx = h(a)\int_a^\eta g(x)dx + h(b)\int_\eta^b g(x)dx$$

(cf. Stromberg [16] pp. 328–329, Takagi [17] pp. 287–288).

Bonnet-type. Let $g : [a, b] \to \mathbb{R}$ be integrable and $h : [a, b] \to \mathbb{R}$ be nonnegative and nonincreasing. Then there exists some $\eta \in [a, b]$ such that

$$\int_a^b g(x)h(x)dx = h(a)\int_a^\eta g(x)dx.$$

If h is nonnegative and nondecreasing, then there exists some $\eta \in [a, b]$ such that

$$\int_a^b g(x)h(x)dx = h(b)\int_\eta^b g(x)dx$$

(cf. Kawata [10] I, p. 18, Stromberg [16] p. 334).

By the above observations, (2.15) holds good for $C = 3$ in the case $0 \leqq a < b$. Hence in the general case (not necessarily $0 \leqq a < b$), $C = 6$ (for instance) satisfies (2.15). (C is independent of a and b.)[4] □

Proof of Theorem 2.4. Without loss of generality, we may assume that f is real-valued.

Since f is of bounded variation on $[x-\delta,\ x+\delta]$, $f(x+0)$ and $f(x-0)$ certainly exist. Specifying $c = (1/2)\ [f(x+0) + f(x-0)]$, we write $\xi(z) = (1/2)\big[f(x+z) + f(x-z)\big] - c$. Then ξ is of bounded variation on $[-\delta, \delta]$ and $\xi(0+) = 0$. There exists a pair of nonnegative and nondecreasing functions $h_1, h_2 : [-\delta,\ \delta] \to \mathbb{R}$ which satisfies $h_1(0+) = h_2(0+) = 0$, and

$$\xi(z) = h_1(z) - h_2(z);\ \ z \in [-\delta,\ \delta]. \tag{2.16}$$

For any $\varepsilon > 0$, there exists some $\delta > 0$ such that

$$0 < z \leqq \delta \Rightarrow 0 \leqq h_1(z),\ \ h_2(z) < \varepsilon. \tag{2.17}$$

Reviewing the calculation of (2.12) on page 30, we should recall that $S_n(x) - c$ is split into three parts:

$$S_n(x) - c = I_1 + I_2 + I_3.$$

We established there that $I_2 + I_3 \to 0$ as $n \to \infty$. Hence

$$\mid I_2 + I_3 \mid < \varepsilon \quad \text{for all} \quad n \geqq n_0$$

for a sufficiently large $n_0 \in \mathbb{N}$. Consequently,

$$\begin{aligned} \mid S_n(x) - c \mid &\leqq \frac{1}{\pi}\left| \int_0^\delta \frac{\sin \dfrac{2n+1}{2} z}{\dfrac{z}{2}} \xi(z) dz \right| + \varepsilon \\ &\leqq \frac{2}{\pi} \sum_{i=1}^{2} \left| \int_0^\delta \frac{\sin \dfrac{2n+1}{2} z}{z} h_i(z) dz \right| + \varepsilon \end{aligned} \tag{2.18}$$

for large n. The magnitude of the integral

[4] We acknowledge Kawata [11] pp. 83–84 for the proof. It can be proved that $C = \pi/2 + 4/\pi \doteqdot 2.8$ satisfies (2.15). See fine calculations by Kawata [10] I, pp. 67–69.

$$\int_0^{\delta} \frac{\sin \frac{2n+1}{2} z}{z} h_i(z) dz$$

is evaluated as follows. By the second mean-value theorem of integration and (2.17), we have

$$\left| \int_0^{\delta} \frac{\sin \frac{2n+1}{2} z}{z} h_i(z) dz \right| = h_i(\delta) \left| \int_{\theta_i}^{\delta} \frac{\sin \frac{2n+1}{2} z}{z} dz \right| \leqq \varepsilon \left| \int_{\theta_i}^{\delta} \frac{\sin \frac{2n+1}{2} z}{z} dz \right| \tag{2.19}$$

for certain $\theta_i \in [0, \delta]$. The final integration is reduced to

$$\left| \int_{2^{-1}(2n+1)\theta_i}^{2^{-1}(2n+1)\delta} \frac{\sin w}{w} dw \right| \tag{2.20}$$

by changing variables: $w = 2^{-1}(2n+1)z$. The value of (2.20) is bounded by some absolute constant C, by Lemma 2.2. Hence we obtain

$$| S_n(x) - c | \leqq \frac{4}{\pi} \varepsilon C + \varepsilon = \left(\frac{4C}{\pi} + 1 \right) \varepsilon \quad \text{for all} \quad n \geqq n_0 \tag{2.21}$$

by (2.18) and (2.19). This completes the proof.[5] □

2.3 Almost Everywhere Convergence: Historical Survey

In the previous section, we studied some necessary and sufficient conditions as well as striking sufficient conditions for a Fourier series to be convergent at a point. However, it is natural to ask if we can expect a Fourier series of a function with nice properties (continuity, integrability, and so on) to be automatically convergent everywhere. Is there any class of functions which guarantees pointwise convergence or almost everywhere convergence? This is a serious problem in the long history of the theory of Fourier series.

First of all, P. du Bois Reymond, in 1876, constructed an example of a continuous function, the Fourier series of which diverges at a point. An optimistic expectation as stated above was doomed to failure by this discovery. We show its basic ideas by following several steps.[6] Let x be any point of $[-\pi, \pi]$.

(i) We first prove that

[5] The proof here is due to Kawata [11] p. 85.

[6] Kolmogorov–Fomin [14] II, pp. 390–391.

$$\int_{-\pi}^{\pi} |D_n(z)|dz \longrightarrow \infty \quad (\text{as } n \to \infty),$$

where $D_n(z)$ is the Dirichlet kernel.

By the definition of $D_n(z)$, we have

$$|D_n(z)| = \frac{\left|\sin \dfrac{2n+1}{2} z\right|}{2\pi \left|\sin \dfrac{z}{2}\right|}.$$

It is easy to see that

$$\left|\sin \frac{2n+1}{2} z_k\right| = 1$$

for any z_k such that

$$\frac{2n+1}{2} z_k = \left(k + \frac{1}{2}\right)\pi, \quad k \in \mathbb{Z}.$$

If such a z_k is in $[-\pi, \pi]$, we have $-n-1 \leqq k \leqq n$. Here we make use only of $0 \leqq k \leqq n$. Let I_k be a sufficiently small interval with center z_k $(0 \leqq k \leqq n)$ such that

$$\left|\frac{2n+1}{2} z - \frac{2n+1}{2} z_k\right| = \left|\frac{2n+1}{2} z - \frac{2k+1}{2}\pi\right| < \frac{\pi}{3},$$

that is,

$$|z - z_k| < \frac{2\pi}{3(2n+1)}.$$

The length of I_k is equal to $4\pi/3(2n+1)$ and

$$\left|\sin \frac{2n+1}{2} z\right| > \frac{1}{2}$$

on I_k. On the other hand, the values of $\sin(z/2)$ on I_k satisfy

$$\sin \frac{z}{2} < \frac{z}{2} < \frac{1}{2} \cdot \frac{\dfrac{2k+1}{2}\pi + \dfrac{\pi}{3}}{\dfrac{2n+1}{2}} < \frac{k+1}{2n+1}\pi.$$

Hence we have

$$\int_{\bigcup_{k=0}^{n} I_k} |D_n(z)|dz \geqq \frac{1}{2\pi}\sum_{k=0}^{n}\frac{1}{2}\frac{1}{\frac{k+1}{2n+1}\pi}\cdot\frac{4\pi}{3(2n+1)}$$

$$= \frac{1}{3\pi}\sum_{k=0}^{n}\frac{1}{k+1} \to \infty \quad \text{as} \quad n \to \infty.$$

This proves (i).

(ii) Define an operator $\Lambda_n : \mathfrak{C}([-\pi, \pi], \mathbb{C}) \to \mathbb{C}$ by

$$\Lambda_n(f) = \int_{-\pi}^{\pi} f(x+z)D_n(z)dz\ ;\ \ f \in \mathfrak{C}\big([-\pi, \pi],\ \mathbb{C}\big).$$

Then Λ_n is a continuous linear functional on $\mathfrak{C}([-\pi, \pi], \mathbb{C})$. $\{\Lambda_n | n \in \mathbb{N}\}$ is not bounded in the operator norm by (i) and the resonance theorem.[7] Therefore it is not bounded in the weak topology.

(iii) Since $\{\Lambda_n\}$ is not weakly convergent by (ii), there exists some $f \in \mathfrak{C}([-\pi, \pi], \mathbb{C})$ for which

$$\lim_{n\to\infty}\int_{-\pi}^{\pi} f(x+z)D_n(z)dz$$

does not exist.

(iv) Thus du Bois Reymond's conclusion immediately follows.

A lot of trials followed the work of du Bois Reymond. For instance, G.H. Hardy, in 1913, gave a illuminating result that the series

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}\frac{1}{\log(\mid n \mid +2)}$$

is convergent almost everywhere for any $f \in \mathfrak{L}^1([-\pi, \pi], \mathbb{C})$, where c_n's are Fourier coefficients in complex form.

Depending upon the accumulation of these experimental results, N.N. Lusin, in 1915, raised an important question explicitly – "Is there any $f \in \mathfrak{L}^2([-\pi, \pi], \mathbb{C})$, the Fourier series of which diverges on a set of positive measure?"

[7] **Banach–Steinhaus resonance theorem**: Let $\mathfrak{X}$ be a Banach space, $\mathfrak{Y}$ a normed space, and $\{T_\alpha | \alpha \in A\}$ a family of bounded linear operators of $\mathfrak{X}$ into $\mathfrak{Y}$. If $\{T_\alpha x | \alpha \in A\}$ is bounded in $\mathfrak{Y}$ for each $x \in \mathfrak{X}$, then $\{T_\alpha | \alpha \in A\}$ is uniformly bounded, i.e. $\sup_{\alpha\in A} \|T_\alpha\| < \infty$. (The converse is obvious.) cf. Dunford–Schwartz [2] pp. 52–53, Maruyama [15] pp. 344–345, Yosida [18] p. 69.

Lusin's clarification of the problem promoted and invoked further researches very much. In particular, the qualitative difference between $\mathfrak{L}^1$ and $\mathfrak{L}^p$ $(p > 1)$ was highlighted in relation with Lusin's problem. One of the most remarkable results was obtained by A.N. Kolmogorov in 1923. It surprised the mathematical world by showing an example of $f \in \mathfrak{L}^1([-\pi, \pi], \mathbb{C})$, the Fourier series of which diverges everywhere.

The mathematical research exhibited a great jump in 1960s. The decades 1920s–1930s may be called the period of classical accumulation which leads to that of harvest in 1960s.

Y. Katznelson proved:
"One of the following statements holds good for $\mathfrak{L}^p([-\pi, \pi], \mathbb{C})$, $1 \leqq p < \infty$:

a. the Fourier series of every $f \in \mathfrak{L}^p$ converges almost everywhere,
b. there exists some $f \in \mathfrak{L}^p$, the Fourier series of which diverges everywhere."

Furthermore, J-P. Kahane and Y. Katznelson also proved the same result for the space $\mathfrak{C}([-\pi, \pi], \mathbb{C})$.[8]

And the best result came last – L. Carleson established the following decisive result in 1966.

Carleson's theorem *The Fourier series of any $f \in \mathfrak{L}^2([-\pi, \pi], \mathbb{C})$ converges to f almost everywhere.*

R.A. Hunt extended this result to the case of any $p > 1$ in the following year. Thus Lusin's problem was solved completely. Carleson's theorem will play a crucial role in Chaps. 9 and 11 in this book. However, we can find no easy way leading to the proof of this theorem. So it is very regrettable that we have to state the theorem without its proof.[9]

2.4 Uniform Convergence

We state and prove the condition which guarantees the uniform convergence of a Fourier series. In this case, functions to be expanded by Fourier series must be continuous.

Theorem 2.5 (uniform convergence) *Assume that $f : \mathbb{R} \to \mathbb{C}$ (or $\mathbb{R}$) is 2π-periodic and absolutely continuous. If, in addition, $f' \in \mathfrak{L}^2([-\pi, \pi], \mathbb{C})$, then the Fourier series of f converges uniformly to f on $\mathbb{R}$.*

[8] The classical works cited here are Hardy [3], Kolmogorov [12, 13], Katznelson [9] and Kahane–Katznelson [7]. See also Hardy–Rogozinski [4]. Zygmund [19] is the best treatise in the discipline.

[9] The classical works are Carleson [1] and Hunt [5]. Jørsboe and Melbro [6] devoted their entire volume to a clear-cut proof of Carleson's theorem.

Proof Since f is absolutely continuous, it is differentiable almost everywhere. We denote by $a_n^{'}$ and $b_n^{'}$ the Fourier coefficients (with respect to trigonometrical functions) of f:

$$a_n' = \frac{1}{\pi}\int_{-\pi}^{\pi} f'(x)\cos nx; \quad n = 0, 1, 2, \cdots,$$

$$b_n' = \frac{1}{\pi}\int_{-\pi}^{\pi} f'(x)\sin nx; \quad n = 0, 1, 2, \cdots.$$

We obtain, by integration by parts,[10] that

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx = \frac{1}{\pi}\left\{ f(x)\frac{\sin nx}{n}\Big|_{-\pi}^{\pi} - \frac{1}{n}\int_{-\pi}^{\pi} f'(x)\sin nx dx \right\} = -\frac{b_n'}{n}.$$

Similarly, we get

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx = \frac{a_n'}{n}.$$

Hence the following evaluation holds good:

$$\sum_{n=1}^{\infty}\Big(\mid a_n \mid + \mid b_n \mid \Big) \leqq \sum_{n=1}^{\infty}\Big(\frac{\mid b_n' \mid}{n} + \frac{\mid a_n' \mid}{n}\Big), \tag{2.22}$$

where

$$2\frac{\mid a_n' \mid}{n} \leqq \mid a_n' \mid^2 + \frac{1}{n^2}, \quad 2\frac{\mid b_n' \mid}{n} \leqq \mid b_n' \mid^2 + \frac{1}{n^2}. \tag{2.23}$$

Applying Bessel's inequality (Theorem 1.4, p. 14) to $f' \in \mathfrak{L}^2\ ([-\pi, \pi], \mathbb{C})$, we have

$$\sum_{n=1}^{\infty}\Big(\mid a_n' \mid^2 + \mid b_n' \mid^2 \Big) < \infty. \tag{2.24}$$

[10] Let u be absolutely continuous and v be integrable on $[a, b]$. Then u, v is also integrable. Denoting by V the indefinite integral of v, we obtain the formula:

$$\int_a^b u(t)v(t)dt = u(b)V(b) - u(a)V(a) - \int_a^b u'(t)V(t)dt.$$

cf. Kato [8] p. 107, Stromberg [16] p. 323.

By (2.22), (2.23) and (2.24), it follows that

$$\sum_{n=1}^{\infty}\Big(\mid a_n \mid + \mid b_n \mid \Big) < \infty. \tag{2.25}$$

Since the absolute value of each term of the Fourier series

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

of f is bounded by the corresponding term of the positive series

$$\frac{\mid a_0 \mid}{2} + \sum_{n=1}^{\infty}\Big(\mid a_n \mid + \mid b_n \mid \Big),$$

the Fourier series of f uniformly converges to some continuous function $\varphi : \mathbb{R} \to \mathbb{C}$.[11]

Finally, both f and φ are square integrable on $[-\pi, \pi]$, since they are continuous. Theorem 1.6 (p. 16) implies that $f = \varphi$ because all the Fourier coefficients of f and φ are identical. That is, the Fourier series of f uniformly converges to f. □[12]

Remark 2.3

1° It is obvious that Theorem 2.5 is valid also for a Fourier series in complex form. The Fourier coefficients c_n are computed as

$$\begin{aligned} c_n &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \\ &= \frac{1}{in} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f'(x) e^{-inx} dx \quad \text{(integration by parts)} \\ &= \frac{1}{in} c_n', \end{aligned}$$

where the c_n''s are the Fourier coefficients of f'. Hence

$$\sum_{n=-\infty}^{\infty} |c_n| \leqq \sum_{n=-\infty}^{\infty} \frac{1}{|n|} |c_n'|.$$

---

[11] cf. Stromberg [16] pp. 141–142, Takagi [17] p. 155, Theorem 39.

[12] The proof given here is due to Kolmogorov–Fomin [14], pp. 391–392.

By the evaluation

$$2\frac{|c_n'|}{n} \leqq |c_n'|^2 + \frac{1}{n^2}$$

and Bessel's inequality, we obtain

$$\sum_{n=-\infty}^{\infty} |c_n| < \infty.$$

The remaining arguments are the same.

2° Assume that f is of $\mathfrak{C}^1$-class on $[-\pi, \pi]$, in particular. Then it is clear that the Fourier series of f uniformly converges to f and that the series which sums up the Fourier coefficients is absolutely convergent.

2.5 Fejér Integral and $(C, 1)$-summability

Consider a series $\sum_{n=0}^{\infty} a_n$, in general. The average of its partial sums

$$S_n = \sum_{k=0}^{n-1} a_k$$

is called the $(C, 1)$**-sum** and is denoted by σ_n; i.e.

$$\sigma_n = \frac{1}{n} \sum_{k=0}^{n-1} S_k.$$

If σ_n converges to σ (as $n \to \infty$), we say that the series $\sum_{n=0}^{\infty} a_n$ is **first summable in the sense of Cesàro** or, briefly, $(C, 1)$**-summable**. More generally, the concept of (C, m)-summability is discussed in Kawata [10] I, p. 12.

Remark 2.4 [13]

(i) If a series $\sum_{n=0}^{\infty} a_n$ converges, it is $(C, 1)$-summable to the same limit.

[13] Takagi [17] p. 9, p. 275. Related results are neatly discussed in Stromberg [16] pp. 473–484.

It can be verified as follows. Assume that $S_n \to S^*$ (as $n \to \infty$). If we define t_n by $S_n = S^* + t_n$ $(n = 1, 2, \cdots)$, then $t_n \to 0$. σ_n can be rewritten as

$$\sigma_n = \frac{1}{n}(S_0 + S_2 + \cdots + S_{n-1}) = S^* + \frac{1}{n}(t_0 + t_1 + \cdots + t_{n-1}).$$

For any $\varepsilon > 0$, there exists some $k_0 \in \mathbb{N}$ such that

$$|t_k| < \varepsilon \quad \text{for all} \quad k \geqq k_0.$$

Defining T by

$$T = \text{Max}\{\,|t_0|, |t_1|, \cdots, |t_{k_{0-1}}|\,\},$$

we obtain the evaluation

$$\left|\frac{1}{n}(t_0 + t_1 + \cdots + t_{n-1})\right| \leqq \frac{1}{n}\{T \cdot k_0 + \varepsilon(n - k_0)\} \leqq \frac{1}{n}T \cdot k_0 + \varepsilon.$$

It follows that $|(1/n)(t_0 + t_1 + \cdots + t_{n-1})| < 2\varepsilon$. Hence $\sigma_n \to S^*$.

(ii) The converse is not necessarily true. We show it by an example. A partial sum of the series $1-1+1-1+\cdots$ is given by $S_{2n} = 0,\ S_{2n+1} = 1$. It must be observed that the series does not converge. However, $\sigma_n \to 1/2$, since $\sigma_{2n} = n/2n = 1/2$ and $\sigma_{2n+1} = (n+1)/(2n+1)$.

Let $f : \mathbb{R} \to \mathbb{C}$ (or $\mathbb{R}$) be a 2π-periodic function which is integrable on $[-\pi, \pi]$. A partial sum of the Fourier series of f is denoted by

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n-1}(a_k \cos kx + b_k \sin kx),$$

and the $(C, 1)$-sum

$$\sigma_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} S_k(x)$$

is also called a **Fejér sum** of f.

As was already shown, $S_k(x)$ can be expressed in the form

$$S_k(x) = \int_{-\pi}^{\pi} f(x+z)D_k(z)dz,$$

where $D_k(x)$ is the Dirichlet kernel. Hence we have

$$\sigma_n(x) = \int_{-\pi}^{\pi} f(x+z)\Big[\frac{1}{n}\sum_{k=0}^{n-1} D_k(z)\Big]dz. \tag{2.26}$$

Defining $K_n(z)$ by

$$K_n(z) = \frac{1}{n}\sum_{k=0}^{n-1} D_k(z), \tag{2.27}$$

we can rewrite (2.26) by

$$\sigma_n(x) = \int_{-\pi}^{\pi} f(x+z)K_n(z)dz. \tag{2.28}$$

$K_n(z)$ is called the **Fejér kernel** and the integral (2.28) is called the **Fejér integral**.

Lemma 2.3 (representation of Fejér kernel)

$$K_n(z) = \frac{1}{2n\pi}\left\{\frac{\sin n\cdot\dfrac{z}{2}}{\sin\dfrac{z}{2}}\right\}^2. \tag{2.29}$$

Proof Multiplying $2\sin^2(z/2)$ to both sides of (2.27), we have

$$\begin{aligned} K_n(z)\cdot 2\sin^2\frac{z}{2} &= \frac{1}{n}\sum_{k=0}^{n-1} D_k(z)2\sin^2\frac{z}{2} = \frac{1}{2n\pi}\sum_{k=0}^{n-1} 2\sin\Big(\frac{2k+1}{2}z\Big)\sin\frac{z}{2} \\ &= \frac{1}{2n\pi}\sum_{k=0}^{n-1}\big[\cos kz - \cos(k+1)z\big] = \frac{1}{2n\pi}(1-\cos nz) = \frac{1}{n\pi}\sin^2\frac{nz}{2}. \end{aligned}$$

The relation (2.29) immediately follows from this. (Of course we are assuming $\sin(z/2) \neq 0$.) □

Lemma 2.4 (properties of Fejér kernel)

(i) $K_n(z) \geqq 0$.

(ii) $\displaystyle\int_{-\pi}^{\pi} K_n(z)dz = 1$.

(iii) *If we define* $M_n(\delta)$ *by* $M_n(\delta) = \sup\limits_{\delta\leqq|z|\leqq\pi} K_n(z) \quad (\delta > 0)$, *then*

$$\lim_{n\to\infty} M_n(\delta) = 0.$$

Proof (i) is obvious. Taking account of

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

we obtain (ii) by (2.27). So we have only to check (iii). By Lemma 2.3, we have

$$0 \leqq K_n(z) = \frac{1}{2n\pi}\left\{\frac{\sin\dfrac{nz}{2}}{\sin\dfrac{z}{2}}\right\}^2.$$

Since it is obvious that

$$\sin^2\frac{\delta}{2} \leqq \sin^2\frac{z}{2}$$

for $|z| \in [\delta, \pi]$, we have

$$0 \leqq K_n(z) \leqq \frac{1}{2n\pi}\cdot\frac{\sin^2\dfrac{nz}{2}}{\sin^2\dfrac{\delta}{2}} \leqq \frac{1}{2n\pi}\cdot\frac{1}{\sin^2\dfrac{\delta}{2}}$$

for such z. Hence

$$M_n(\delta) \leqq \frac{1}{2n\pi}\cdot\frac{1}{\sin^2\dfrac{\delta}{2}} \longrightarrow 0 \quad (\text{as } n \to \infty).$$

□

Theorem 2.6 (Fejér's summation formula) *If a function $f : \mathbb{R} \to \mathbb{C}$ (or $\mathbb{R}$) is continuous and 2π-periodic, then its Fejér sum $\sigma_n(x)$ converges uniformly to f on $\mathbb{R}$.*

Proof Since f is continuous and periodic, it is bounded and uniformly continuous. So there exists $\delta > 0$ for each $\varepsilon > 0$ such that

$$| x' - x'' | \leqq \delta \quad\Longrightarrow\quad | f(x') - f(x'') | \leqq \frac{\varepsilon}{3}.$$

Fix such a positive number $\delta > 0$. It follows that

$$\begin{aligned} f(x) - \sigma_n(x) &= \int_{-\pi}^{\pi} [f(x) - f(x+z)] K_n(z) dz \\ &= \int_{-\pi}^{-\delta} + \int_{-\delta}^{\delta} + \int_{\delta}^{\pi} \\ &\equiv J_1 + J_2 + J_3. \end{aligned} \tag{2.30}$$

Defining $A = \sup_{x \in \mathbb{R}} |f(x)|$ $(< \infty)$, we obtain the evaluations

$$|\ J_1\ | \leqq 2A \int_{-\pi}^{-\delta} K_n(z) dz \leqq 2A \cdot M_n(\delta) \cdot \pi, \tag{2.31}$$

$$|\ J_3\ | \leqq 2A \int_{\delta}^{\pi} K_n(z) dz \leqq 2A \cdot M_n(\delta) \cdot \pi, \tag{2.32}$$

$$|\ J_2\ | \leqq \frac{\varepsilon}{3} \int_{-\delta}^{\delta} K_n(z) dz \leqq \frac{\varepsilon}{3}. \tag{2.33}$$

We observe that (2.31), (2.32), and (2.33) are valid regardless of x. The right-hand sides of (2.31) and (2.32) can be smaller than $\varepsilon/3$ for a sufficiently large n, by Lemma 2.4 (iii). Hence

$$|\ f(x) - \sigma_n(x)\ | \leqq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \quad \text{for all} \quad x \in \mathbb{R}$$

for large n. This proves the theorem. □

We will recapitulate Fejér's summability method in more detail in Chap. 5.

References

1. Carleson, L.: On the convergence and growth of partial sums of Fourier series. Acta Math. **116**, 135–157 (1966)
2. Dunford, N., Schwartz, J.T.: Linear Operators, Part 1. Interscience, New York (1958)
3. Hardy, G.H.: On the summability of Fourier series. Proc. London Math. Soc. **12**, 365–372 (1913)
4. Hardy, G.H., Rogozinski, W.W.: Fourier Series, 3rd edn. Cambridge University Press, London (1956)
5. Hunt, R.A.: On the convergence of Fourier series. In: Proceedings of the Conference on Orthogonal Expansions and Their Continuous Analogues, pp. 234–255. Southern Illinois University Press, Carbondale (1968)
6. Jørsboe, O.G., Melbro, L.: The Carleson-Hunt Theorem on Fourier Analysis. Springer, Berlin (1982)
7. Kahane, J.P., Katznelson, Y.: Sur les ensembles de divergence des séries trigonométriques. Stud. Math. **26**, 305–306 (1966)

8. Kato, T.: Iso Kaiseki (Functional Analysis). Kyoritsu Shuppan, Tokyo (1957) (Originally published in Japanese)
9. Katznelson, Y.: Sur les ensembles de divergence des séries trigonometriques. Stud. Math. **26**, 301–304 (1966)
10. Kawata, T.: Ohyo Sugaku Gairon (Elements of Applied Mathematics). I, II, Iwanami Shoten, Tokyo (1950, 1952) (Originally published in Japanese)
11. Kawata, T.: Fourier Kaiseki (Fourier Analysis). Sangyo Tosho, Tokyo (1975) (Originally published in Japanese)
12. Kolmogorov, A.N.: Une série de Fourier-Lebesgue divergent presque partout. Fund. Math. **4**, 324–328 (1923)
13. Kolmogorov, A.N.: Une série de Fourier-Lebesgue divergent partout. C. R. Acad. Sci. Paris **183**, 1327–1328 (1926)
14. Kolmogorov, A.N., Fomin, S.V.: (English edn.) Introductory Real Analysis. Prentice-Hall, New York (1970) (Japanese edn., translated from 4th edn.) Kansu-kaiseki no Kiso, 4th edn. Iwanami Shoten, Tokyo (1979), (Originally published in Russian). I use the Japanese edition for quotations.
15. Maruyama, T.: Kansu Kaisekigaku (Functional Analysis). Keio Tsushin, Tokyo (1980) (Originally published in Japanese)
16. Stromberg, K.R.: An Introduction to Classical Real Analysis. American Mathematical Society, Providence (1981)
17. Takagi, T.: Kaiseki Gairon (Treatise on Analysis), 3rd edn. Iwanami Shoten, Tokyo (1961) (Originally published in Japanese)
18. Yosida, K.: Functional Analysis, 3rd edn. Springer, New York (1971)
19. Zygmund, A.: Trigonometric Series, 2nd edn. Cambridge University Press, Cambridge (1959)

Chapter 3
Fourier Transforms (I)

The objects of classical theory of Fourier series discussed in the preceding chapter are periodic functions. Is it possible to construct an analogous theory for nonperiodic functions? It is the theory of Fourier transforms which answers this question positively.

3.1 Fourier Integrals

We first make a rough sketch of ideas in order to grasp what is going on. A function $f : \mathbb{R} \to \mathbb{C}$ is assumed to be integrable on each finite interval and to satisfy some condition which assures the convergence of its Fourier series (say, Dini's condition) at each point. The restriction of f to the interval $[-l, l]$ can be expanded by its Fourier series in the form (cf. pp. 20–21)

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi}{l} x + b_n \sin \frac{n\pi}{l} x \right), \tag{3.1}$$

where

$$\begin{aligned} a_n &= \frac{1}{l} \int_{-l}^{l} f(t) \cos \frac{n\pi}{l} t dt; \;\; n = 0, 1, 2, \cdots , \\ b_n &= \frac{1}{l} \int_{-l}^{l} f(t) \sin \frac{n\pi}{l} t dt; \;\; n = 1, 2, \cdots . \end{aligned} \tag{3.2}$$

Substituting (3.2) into (3.1), we obtain

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_3

$$f(x) = \frac{1}{2l}\int_{-l}^{l} f(t)dt + \sum_{n=1}^{\infty}\left\{\frac{1}{l}\int_{-l}^{l} f(t)\cos\frac{n\pi}{l}tdt \cdot \cos\frac{n\pi}{l}x\right\}$$
$$+ \sum_{n=1}^{\infty}\left\{\frac{1}{l}\int_{-l}^{l} f(t)\sin\frac{n\pi}{l}tdt \cdot \sin\frac{n\pi}{l}x\right\}$$
$$= \frac{1}{2l}\int_{-l}^{l} f(t)dt + \sum_{n=1}^{\infty}\frac{1}{l}\int_{-l}^{l} f(t)\cos\frac{n\pi}{l}(t-x)dt$$
$$= \frac{1}{2l}\int_{-l}^{l} f(t)dt + \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{\pi}{l}\int_{-l}^{l} f(t)\cos\frac{n\pi}{l}(t-x)dt. \tag{3.3}$$

If we assume, in addition, that f is integrable on $\mathbb{R}$, then the first term of (3.3) tends to 0 as $l \to \infty$. Furthermore, the summation part $\sum(\cdots)$ of the second term can be regarded as an approximation of the integral on $[0, \infty)$ (with respect to λ) of the function

$$\int_{-l}^{l} f(t)\cos\lambda(t-x)dt. \tag{3.4}$$

Dividing $[0, \infty)$ into intervals of length π/l, we evaluate the values of (3.4) at the right end-point of each interval and sum up these values after the multiplication by π/l. This operation gives $\sum(\cdots)$. (See Fig. 3.1.)

The second term of (3.3) may be expected to converge to

$$\frac{1}{\pi}\int_0^{\infty}\left\{\int_{-\infty}^{\infty} f(t)\cos\lambda(t-x)dt\right\}d\lambda$$

as $l \to \infty$, since the decomposition becomes finer and finer as $l \to \infty$. Although this is not a rigorous argument, we expect the relation

$$f(x) = \frac{1}{\pi}\int_0^{\infty}\left\{\int_{-\infty}^{\infty} f(t)\cos\lambda(t-x)dt\right\}d\lambda \tag{3.5}$$

to be almost correct. If we define a_λ and b_λ by

$$a_\lambda = \frac{1}{\pi}\int_0^{\infty} f(t)\cos\lambda tdt, \quad b_\lambda = \frac{1}{\pi}\int_0^{\infty} f(t)\sin\lambda tdt, \tag{3.6}$$

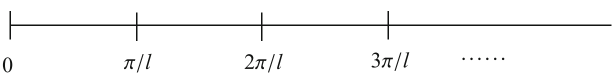


Fig. 3.1 Interpretation of (3.3)

we have the expression

$$f(x) = \int_0^\infty (a_\lambda \cos \lambda x + b_\lambda \sin \lambda x) d\lambda. \tag{3.7}$$

This representation of f in the integral form should be regarded as an analogue to the Fourier series of periodic functions.

The outline of the derivation of (3.7) shown above seems almost correct. However, there remains some dubious reasoning in the limit operation. We now try to give a more rigorous proof of (3.7).

Theorem 3.1 (Fourier integral formula) *Assume that a function* $f : \mathbb{R} \to \mathbb{C}$ *is integrable and satisfies*

$$\int_{-\delta}^{\delta} \left| \frac{f(x+z) - f(x)}{z} \right| dz < \infty \tag{3.8}$$

for each $\delta > 0$ *and each* $x \in \mathbb{R}$*. Then we obtain*

$$f(x) = \frac{1}{\pi} \int_0^\infty \left\{ \int_{-\infty}^\infty f(t) \cos \lambda (t - x) dt \right\} d\lambda. \tag{3.5}$$

Remark 3.1 Note that (3.8) is exactly Dini's condition in the case $c = f(x)$ (cf. p. 31).

Proof If we define $J(A)$ $(A > 0)$ by

$$J(A) = \frac{1}{\pi} \int_0^A \left\{ \int_{-\infty}^\infty f(t) \cos \lambda (t - x) dt \right\} d\lambda, \tag{3.9}$$

we have

$$\int_0^A \left\{ \int_{-\infty}^\infty |f(t) \cos \lambda (t - x)| dt \right\} d\lambda < \infty, \tag{3.10}$$

since f is integrable on $\mathbb{R}$. Hence applying Fubini's theorem, we obtain

$$\begin{aligned} J(A) &= \frac{1}{\pi} \int_{-\infty}^\infty \left\{ \int_0^A f(t) \cos \lambda (t - x) d\lambda \right\} dt \\ &= \frac{1}{\pi} \int_{-\infty}^\infty f(t) \frac{\sin A(t - x)}{t - x} dt \qquad (3.11) \\ &= \frac{1}{\pi} \int_{-\infty}^\infty f(x + z) \frac{\sin Az}{z} dz \quad (\text{changing variables}: \quad z = t - x). \end{aligned}$$

Consequently,

$$\begin{aligned} J(A) - f(x) &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(x+z) - f(x)}{z} \sin Az\, dz \\ &= \frac{1}{\pi} \int_{-N}^{N} \frac{f(x+z) - f(x)}{z} \sin Az\, dz \\ &\qquad + \frac{1}{\pi} \int_{|z| \geqq N} \frac{f(x+z)}{z} \sin Az\, dz - \frac{f(x)}{\pi} \int_{|z| \geqq N} \frac{\sin Az}{z} dz \\ &\equiv I_1 + I_2 + I_3. \end{aligned} \tag{3.12}$$

Taking account of the well-known result

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin Az}{z} dz = 1 \quad (A > 0), \tag{3.13}$$

we get the evaluations

$$| I_2 | < \frac{\varepsilon}{3} \quad , \quad | I_3 | < \frac{\varepsilon}{3}$$

for any $\varepsilon > 0$ if $N > 0$ is sufficiently large. Fix such a large $N > 0$. Then (3.8) and Lemma 2.1 (pp. 28–29) imply $| I_1 | < \varepsilon/3$ for a large A. Thus we obtain

$$| J(A) - f(x) | < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

for a large $A > 0$. This proves (3.5). □

The expression (3.5) is called the **Fourier integral formula**.

Since the integral

$$\int_{-\infty}^{\infty} f(t) \cos \lambda(t - x) dt$$

appearing in (3.5) is an even function in λ, $f(x)$ can be written in the form

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f(t) \cos \lambda(t - x) dt \right\} d\lambda. \tag{3.14}$$

By the integrability of f, the integral

$$\int_{-\infty}^{\infty} f(t) \sin \lambda(t - x) dt$$

exists, and it is an odd function in λ. Hence

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(t)\sin\lambda(t-x)dt\right\}d\lambda = 0, \tag{3.15}$$

where **the integral** (**with respect to** λ) **is in the sense of Cauchy's principal value**; i.e.

$$\lim_{N\to\infty}\int_{-N}^{N}\cdots d\lambda.$$

The operation (3.14)–i(3.15) gives the Fourier integral formula in complex form:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(t)e^{-i\lambda(t-x)}dt\right\}d\lambda. \tag{3.16}$$

If we define $g(\lambda)$ by

$$g(\lambda) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\lambda t}dt, \tag{3.17}$$

(3.16) becomes

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(\lambda)e^{i\lambda x}d\lambda. \tag{3.18}$$

We must be sure that the integral in (3.17) always exists, but the integral in (3.18) exists only in the sense of Cauchy's principal value in general. While (3.17) just defines a function g, (3.18) is a variant of the Fourier integral formula and implies a positive claim that the right-hand side is equal to $f(x)$.

Needless to say, $g(\lambda)$ corresponds to Fourier coefficients and (3.18) corresponds to the Fourier series in the case of periodic functions.[1]

We now proceed to the theory of Fourier transforms based upon the preliminary observations explained above. The function $g(\lambda)$ defined by (3.17) is the Fourier transform of $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$, which can be interpreted as a Fourier coefficient of a function with period ∞.[2]

---

[1] The exposition in this section is primarily due to Kolmogorov–Fomin [6] Chap. 8, §3.

[2] Some basic textbooks on the classical theory of Fourier transforms are Dym–McKean [2], Goldberg [3], Kawata [4, 5], Titchmarsh [10].

3.2 Fourier Transforms on $\mathfrak{L}^1(\mathbb{R}, \mathbb{C})$

Definition 3.1 Let f be an element of $\mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. The function $\hat{f} : \mathbb{R} \longrightarrow \mathbb{C}$ defined by

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-itx} dx$$

is called the **Fourier transform** of f. The mapping $\mathcal{F} : f \longmapsto \hat{f}$ is also called the **Fourier transform**.

Theorem 3.2 (uniform continuity) *The Fourier transform $\hat{f}$ of $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ is uniformly continuous on $\mathbb{R}$.*

Proof Since

$$\hat{f}(t+u) - \hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\{e^{-i(t+u)x} - e^{-itx}\} dx,$$

we obtain the evaluation

$$| \hat{f}(t+u) - \hat{f}(t) | \leqq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} | f(x) || e^{-iux} - 1 | \, dx.$$

The right-hand side is independent of t. Taking account of

$$| f(x) || e^{-iux} - 1 | \leqq 2 | f(x) |,$$

we can apply the dominated convergence theorem to conclude that the right-hand side converges to 0 as $u \to 0$. Hence $\hat{f}$ is uniformly continuous. □

It is quite easy to check the following relations for $f, g \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ and $\alpha \in \mathbb{C}$:

$$(\widehat{f+g})(t) = \hat{f}(t) + \hat{g}(t), \tag{3.19}$$

$$(\widehat{\alpha f})(t) = \alpha \hat{f}(t), \tag{3.20}$$

$$| \hat{f}(t) | \leqq \frac{1}{\sqrt{2\pi}} \| f \|_1 . \tag{3.21}$$

We denote by $\mathfrak{U}(\mathbb{R}, \mathbb{C})$ the space of all the uniformly continuous functions of $\mathbb{R}$ into $\mathbb{C}$, endowed with the usual vector operations and the uniform convergence norm. Then the mapping $\mathcal{F} : f \longmapsto \hat{f}$ is a bounded linear operator of the form

$$\mathcal{F} : \mathfrak{L}^1(\mathbb{R}, \mathbb{C}) \longrightarrow \mathfrak{U}(\mathbb{R}, \mathbb{C}),$$

the operator norm of which is given by

$$\| \mathcal{F} \| \leqq \frac{1}{\sqrt{2\pi}}$$

by (3.21).

More properties of $\hat{f}$ will be discussed later. In particular, it must be remembered that $\hat{f}$ vanishes at infinity.

Remark 3.2 The following formulas hold good for $f, g \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$.

(i) $\hat{\overline{f}}(t) = \overline{\hat{f}(-t)}$. ($\overline{f(\cdot)}$ is the conjugate complex of $f(\cdot)$.)
(ii) If we write $f_y(x) = f(x-y)$, then $\hat{f}_y(t) = \hat{f}(t)e^{-ity}$.
(iii) If we write $\varphi(x) = \lambda f(\lambda x)$ for a real $\lambda \neq 0$, then $\hat{\varphi}(t) = \hat{f}(t/\lambda)$.
(iv) If we define $h(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\infty} g(t)e^{itx}dt$ (integrable),

then

$$(h * f)(x) = \int_{-\infty}^{\infty} g(t)\hat{f}(t)e^{itx}dt,$$

where $h * f$ is the convolution of h and f, i.e.

$$(h * f)(x) = \int_{-\infty}^{\infty} h(x-y)f(y)dy.$$

We now illustrate some examples of Fourier transforms.

Example 3.1 If we define a function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1 & \text{on } (a,\ b), \\ 0 & \text{otherwise,} \end{cases}$$

the Fourier transform of this function is given by

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-itx}dx.$$

Hence it is evaluated as

$$\hat{f}(t) = \frac{b-a}{\sqrt{2\pi}}$$

at $t = 0$, and

$$\hat{f}(t) = \frac{-1}{\sqrt{2\pi}\, it}\ e^{-itx}\Big|_a^b = \frac{1}{\sqrt{2\pi}\, it}\ (e^{-ita} - e^{-itb})$$

at $t \neq 0$. If $a = -b$ in particular, $\hat{f}$ becomes

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}\, it}\ (e^{itb} - e^{-itb}) = \frac{2\sin tb}{\sqrt{2\pi}t}.$$

Example 3.2 The Fourier transform of $f(x) = e^{-\gamma|x|} \quad (\gamma > 0)$ is given by

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\gamma|x|} e^{-itx} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\gamma|x|} (\cos tx - i \sin tx) dx,$$

where

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{-\gamma|x|} \sin tx dx &= \int_0^{\infty} e^{-\gamma x} \sin tx dx + \int_{-\infty}^{0} e^{\gamma x} \sin tx dx \\
&= \int_0^{\infty} e^{-\gamma x} \sin tx dx - \int_{\infty}^{0} e^{-\gamma y} \sin t(-y) dy \\
&\quad \text{(changing variables: } -x = y) \\
&= \int_0^{\infty} e^{-\gamma x} \sin tx dx - \int_0^{\infty} e^{-\gamma y} \sin ty dy \\
&= 0.
\end{aligned}$$

Hence we have

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\gamma|x|} \cos tx dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\gamma x} \cos tx dx \equiv \frac{2}{\sqrt{2\pi}} I.$$

By integration by parts, I can be computed as

$$I = \frac{\gamma}{\gamma^2 + t^2}.$$

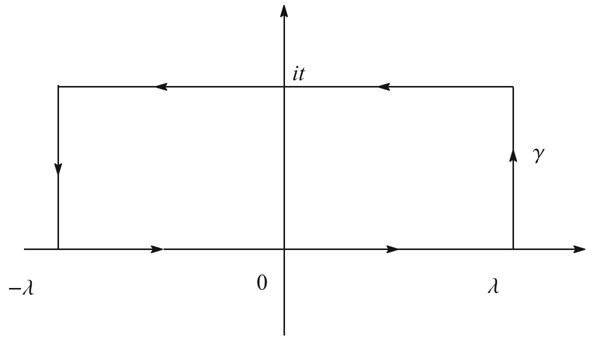
So we obtain

$$\hat{f}(t) = \frac{2}{\sqrt{2\pi}} \cdot \frac{\gamma}{\gamma^2 + t^2}.$$

Example 3.3 Find the Fourier transform of $f(x) = e^{-x^2/2}$.

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} e^{-itx} dx = e^{-t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x+it)^2/2} dx.$$

Fig. 3.2 Closed path γ



Integrating the holomorphic function $e^{-z^2/2}$ along a closed path γ given in Fig. 3.2, we obtain

$$\int_\gamma e^{-z^2/2} dz = 0$$

by Cauchy's theorem.[3] Hence

$$\frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} e^{-(x+it)^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} e^{-x^2/2} dx + \frac{1}{\sqrt{2\pi}} \int_t^0 e^{-(-\lambda+it)^2/2} i dt \tag{3.22}$$
$$+ \frac{1}{\sqrt{2\pi}} \int_0^t e^{-(\lambda+it)^2/2} i dt.$$

The second and the third terms of (3.22) converge to 0 as $\lambda \to \infty$ for a fixed t. In fact, we can verify this fact for the third term by observing

$$\left| \int_0^t e^{-(\lambda+it)^2/2} dt \right| \leqq \int_0^t e^{-(\lambda^2-t^2)/2} dt = e^{-\lambda^2/2} \int_0^t e^{t^2/2} dt \to 0 \quad \text{as } \lambda \to \infty.$$

Consequently, we obtain by the Gauss integration that

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} e^{-itx} dx = e^{-t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = e^{-t^2/2}.$$

Thus we arrive at the conclusion that $f(x) = e^{-x^2/2}$ is a special function which is invariant under the Fourier transform.

[3] cf. Cartan [1] Chap. II, and Takagi [9] pp. 208–210.

It immediately follows that the Fourier transform of $f(x) = e^{-ax^2}$ $(a > 0)$ is given by

$$\hat{f}(t) = \frac{1}{\sqrt{2a}} e^{-t^2/4a}.$$

We now list some basic properties of Fourier transforms.

We start by investigating the convolutions of two functions. Let $f,\ g : \mathbb{R} \longrightarrow \mathbb{C}$ be integrable. The convolution of f and g is defined by

$$h(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy. \tag{3.23}$$

A simple calculation gives

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x-y)g(y)|dx\ dy &= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}|f(z)|dz\right\}|g(y)|dy \\ &= \| f \|_1 \cdot \| g \|_1 < \infty.\end{aligned}$$

Thus we observe that:

(i) $h(x)$ is defined at almost every x, and
(ii) h is integrable.

The definition (3.23) makes sense almost everywhere and $h(x)$ is usually denoted by $(f * g)(x)$.

Theorem 3.3 (Fourier transform of a convolution)

$$(\widehat{f * g})(t) = \sqrt{2\pi}\,\hat{f}(t)\hat{g}(t)$$

for $f,\ g \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$.

Proof The theorem is proved by a direct calculation:

$$\begin{aligned}(\widehat{f * g})(t) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(x-y)g(y)dy\right\}e^{-itx}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(x-y)e^{-itx}dx\right\}g(y)dy \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(w)e^{-itw}dw\right\}g(y)e^{-ity}dy \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(w)e^{-itw}dw \cdot \int_{-\infty}^{\infty} g(y)e^{-ity}dy \\ &= \sqrt{2\pi}\,\hat{f}(t)\hat{g}(t).\end{aligned}$$

□

Theorem 3.4 (Fourier transform of a derivative) *Assume that $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ is absolutely continuous on every finite interval and $f' \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. Then*

$$\widehat{f'}(t) = it\hat{f}(t).$$

Proof Since f is absolutely continuous on any finite interval, $f(x)$ can be expressed as

$$f(x) = f(0) + \int_0^x f'(t)dt.$$

The right-hand side converges as $x \to \pm\infty$, since $f' \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. This limit is 0, since $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. Consequently, we have

$$\begin{aligned}
\widehat{f'}(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(x)e^{-itx}dx \\
&= \frac{1}{\sqrt{2\pi}} \left\{ f(x)e^{-itx}\Big|_{-\infty}^{\infty} + it\int_{-\infty}^{\infty} f(x)e^{-itx}dx \right\} \\
&= it \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-itx}dx \\
&= it\hat{f}(t).
\end{aligned}$$

□

Remark 3.3 The corresponding result for the Fourier coefficients of an absolutely continuous function on $[-\pi, \pi]$ is as follows. Let $f : \mathbb{R} \to \mathbb{C}$ be a 2π-periodic function which is integrable on $[-\pi, \pi]$ and satisfies $\hat{f}(0) = 0$. If we define

$$F(x) = f(0) + \int_0^x f(t)dt,$$

then F is a 2π-periodic continuous function and the Fourier coefficients of F are computed as

$$\hat{F}(n) = \frac{1}{in}\hat{f}(n), \quad n \neq 0.$$

In fact, the continuity (even the absolute continuity) of F is obvious. F is 2π-periodic since

$$F(x + 2\pi) - F(x) = \int_x^{x+2\pi} f(x)dt = \sqrt{2\pi}\hat{f}(0) = 0.$$

Finally, we see by integration by parts that

$$\hat{F}(n) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} F(t) e^{-int} dt = \frac{-1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} F'(t) \frac{1}{-in} e^{-int} dt = \frac{1}{in} \hat{f}(n)$$

$$\text{(assuming } n \neq 0).$$

Repeating the same reasoning as that in Theorem 3.4, we obtain the next corollary.

Corollary 3.1 (Fourier transform of higher derivatives) *If the ν-th derivative $f^{(\nu)}(\nu = 0, 1, \cdots, k-1)$ of $f : \mathbb{R} \to \mathbb{C}$ is absolutely continuous on each finite interval and $f^{(\nu)} \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ $(\nu = 0, 1, \cdots, k-1)$, then*

$$\widehat{f^{(k)}}(t) = (it)^k \hat{f}(t).$$

The next result is an analogue of the Riemann–Lebesgue lemma concerning Fourier coefficients (Corollary 1.1, p. 14).

Theorem 3.5 (Riemann–Lebesgue) *For $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$,*

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

Proof Assume first that $g \in \mathfrak{C}_0^1(\mathbb{R}, \mathbb{C})$, i.e. g is continuously differentiable and the support of g is compact. Then we have, by Theorem 3.4 and (3.21),

$$|\, \widehat{g'}(t) \,| = |\, t\hat{g}(t) \,| \leqq \frac{1}{\sqrt{2\pi}} \parallel g' \parallel_1,$$

which implies

$$\lim_{|t| \to \infty} |\, \hat{g}(t) \,| \leqq \lim_{|t| \to \infty} \frac{1}{|\, t \,|} \cdot \frac{1}{\sqrt{2\pi}} \parallel g' \parallel_1 = 0. \tag{3.24}$$

For any $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ and any $\varepsilon > 0$, there exists some $g \in \mathfrak{C}_0^1(\mathbb{R}, \mathbb{C})$ such that

$$\parallel f - g \parallel_1 < \sqrt{2\pi} \cdot \varepsilon.$$

Again by (3.21), we have

$$|\, \hat{f}(t) - \hat{g}(t) \,| \leqq \varepsilon. \tag{3.25}$$

It follows, from (3.24) and (3.25), that

$$\overline{\lim_{|t| \to \infty}} \,|\, \hat{f}(t) \,| \leqq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that

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

□

This establishes the important result that **the Fourier transform of a function in $\mathfrak{L}^1$ is uniformly continuous and vanishes at infinity.**

The next theorem clarifies how fast $\hat{f}$ vanishes at infinity.

Theorem 3.6 (speed of vanishing at infinity) *Assume that the derivatives $f^{(\nu)}(\nu = 0, 1, \cdots, k-1)$ of a function are absolutely continuous on each finite interval and $f^{(\nu)} \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ $(\nu = 0, 1, \cdots, k-1)$. Then*

$$\hat{f}(t) = o(| t |^{-k}) \quad \textit{as} \quad | t |\to \infty.$$

Proof By Corollary 3.1,

$$\widehat{f^{(k)}}(t) = (it)^k \hat{f}(t) = \frac{1}{(\frac{1}{it})^k} \hat{f}(t).$$

Since $f^{(k)} \in \mathfrak{L}^1$, $\widehat{f^{(k)}}(t) \to 0$ (as $| t |\to \infty$) by Theorem 3.5. Hence we obtain

$$| \widehat{f^{(k)}}(t) |= \frac{1}{| \frac{1}{t} |^k} | \hat{f}(t) |\to 0 \qquad \text{as} \qquad | t |\to \infty.$$

□

This theorem tells us that **the higher the derivatives f admits, the more rapidly $\hat{f}$ vanishes at infinity.** We can also prove a dual result that **the more rapidly f vanishes at infinity, the smoother $\hat{f}$ becomes.** We prepare a lemma.

Lemma 3.1 *For any complex number $a = \alpha + i\beta$,*

$$| e^{ia} - 1 |\leqq \begin{cases} |a|e^{-\beta} & \textit{for} \quad \beta \leqq 0, \\ |a| & \textit{for} \quad \beta \geqq 0. \end{cases}$$

Theorem 3.7 (derivative of Fourier transform) *Assume that f, $xf(x) \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. Then $\hat{f}$ is differentiable and*

$$\frac{d}{dt}\hat{f}(t) = (\widehat{-ixf})(t).$$

Proof First we observe that

$$\frac{\hat{f}(t+h) - \hat{f}(t)}{h} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-itx} \left(\frac{e^{-ihx} - 1}{h} \right) dx.$$

By Lemma 3.1,

$$\left| f(x)e^{-itx}\left(\frac{e^{-ihx}-1}{h}\right)\right| \leqq |\, xf(x)\,| \, .$$

It is clear that

$$f(x)e^{-itx}\left(\frac{e^{-ihx}-1}{h}\right) \longrightarrow -ixf(x)e^{-itx} \qquad \text{as } \ h \to 0.$$

Combining these results and the assumption $xf(x) \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$, we obtain, by the dominated convergence theorem,

$$\begin{aligned}
\lim_{h\to 0} \frac{\hat{f}(t+h)-\hat{f}(t)}{h} &= \lim_{h\to 0} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-itx}\left(\frac{e^{-ihx}-1}{h}\right) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-ixf(x))e^{-itx} dx \\
&= (\widehat{-ixf})(t).
\end{aligned}$$

□

Similarly:

Corollary 3.2 (higher derivatives of Fourier transform) *If $f(x), xf(x), \cdots,$ $x^k f(x)$ are integrable, then $\hat{f}$ is differentiable k-times and*

$$\hat{f}^{(\nu)}(t) = [\widehat{(ix)^\nu f}](t) \ , \ \nu = 0, 1, \cdots, k.$$

We have arrived at a significant result: **the Fourier transform $\hat{f}$ of a function f which is very smooth and vanishes at infinity very rapidly is also very smooth and vanishes at infinity very rapidly.** This observation leads to the idea of the space of rapidly decreasing functions, which is to be discussed in Chap. 4.

But before going on to this topic, we make a digression to a simple application to physics. The heat equation will be discussed in the next section.

3.3 Application: Heat Equation

We interpret $u(x, \ t)$ as the temperature at a point $x \in (-\infty, \ \infty)$ on a line and at time $t \geqq 0$. Given an initial condition $u(x, \ 0) = u_0(x)$, the dynamic process of heat conduction is described by the following partial differential equation, called the **heat equation**:

$$\frac{\partial u}{\partial t}(x,\ t) = \frac{\partial^2 u}{\partial x^2}(x,\ t) \tag{3.26}$$

$$\text{subject to} \quad u(x,\ 0) = u_0(x). \tag{3.27}$$

$u_0(x)$ is a known function, and we suppose that

$$u_0(x), \quad u_0'(x), \quad u_0''(x) \in \mathfrak{L}^1(\mathbb{R}, \mathbb{R}).$$

We would like to find a solution of (3.26) and (3.27) in the space of functions u which satisfy:

(i) $x \mapsto u(x,t),\ x \mapsto u_x(x,t),\ x \mapsto u_{xx}(x,t)$ are integrable on $\mathbb{R}$ for any fixed t $(u_x = \partial u/\partial x,\ u_{xx} = \partial^2 u/\partial x^2)$ and
(ii) for any bounded interval $[0, T]$, there exists some integrable function $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that

$$|\ u_t(x,t)\ | \leqq f(x) \quad \text{for all} \quad t \in [0,\ T] \quad \text{and} \quad x \in \mathbb{R}.$$

The Fourier transforms of both sides of (3.26) (with respect to x) are given by

$$\begin{aligned}\text{Fourier transform of the left-hand side of (3.26)} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_t(x,t) e^{-i\lambda x} dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{\partial}{\partial t} \int_{-\infty}^{\infty} u(x,\ t) e^{-i\lambda x} dx\end{aligned}$$

(making use of differentiation formula under integral[4]), and

$$\text{Fourier transform of the right-hand side of (3.26)} = -\lambda^2 \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x,t) e^{-i\lambda x} dx$$

(by Corollary 3.1).

If we define $v(\lambda, t)$ by

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x,t) e^{-i\lambda x} dx = v(\lambda, t),$$

we have

$$v_t(\lambda,\ t) = -\lambda^2 v(\lambda,\ t). \tag{3.28}$$

[4]cf. Takagi [9] p. 166, Theorem 42. We must be sure that we are dealing with an improper integral here. See also Stromberg [8] p. 380.

It is not difficult to solve (3.28) under the initial condition

$$v(\lambda,\ 0) = \widehat{u_0}(\lambda). \tag{3.29}$$

In fact, (3.28) and (3.29) can be regarded as an ordinary differential equation in t when λ is arbitrarily fixed. The solution is given by

$$v(\lambda,\ t) = e^{-\lambda^2 t}\widehat{u_0}(\lambda). \tag{3.30}$$

Applying the result on p. 56 to the case $a = 1/4t$, we can observe that $e^{-\lambda^2 t}$ is the Fourier transform of

$$x \mapsto \frac{1}{\sqrt{2t}} e^{-x^2/4t}.$$

If we denote by $\mathcal{F}$ the mapping $f \mapsto \hat{f}$, we obtain, by (3.30), that

$$\begin{aligned} v(\lambda, t) &= \mathcal{F}\left(\frac{1}{\sqrt{2t}} e^{-x^2/4t}\right)(\lambda) \cdot \mathcal{F}(u_0)(\lambda) \\ &= \frac{1}{\sqrt{2\pi}} \cdot \mathcal{F}\left(\frac{1}{\sqrt{2t}} e^{-x^2/4t} * u_0(x)\right)(\lambda) \qquad \text{(by Theorem 3.3).} \end{aligned}$$

Hence $v(\lambda, t)$ is the Fourier transform of

$$u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} e^{-(x-\xi)^2/4t} \cdot u_0(\xi) d\xi = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} e^{-\xi^2/4t} \cdot u_0(x - \xi) d\xi. \tag{3.31}$$

The function $u(x, t)$ thus obtained is the solution of (3.26) and (3.27), and is called the **Poisson formula** of the heat equation. Check that (3.31) is really the solution.[5]

References

1. Cartan, H.: Théorie élémentaires des fonctions analytiques d'une ou plusieurs variables complexes. Hermann, Paris (1961) (English edn.) Elementary Theory of Analytic Functions of One or Several Complex Variables. Addison Wesley, Reading (1963)
2. Dym, H., McKean, H.P.: Fourier Series and Integrals. Academic Press, New York (1972)
3. Goldberg, R.R.: Fourier Transforms. Cambridge University Press, Cambridge (1961)
4. Kawata, T.: Fourier Henkan to Laplace Henkan (Fourier Transforms and Laplace Transforms). Iwanami Shoten, Tokyo (1957) (Originally published in Japanese)

[5]The exposition here is basically due to Kolmogorov–Fomin [6] Chap. 8, §4, 6°. See also Schwartz [7] pp. 211–215 and pp. 311–315.

5. Kawata, T.: Fourier Kaiseki (Fourier Analysis). Sangyo Tosho, Tokyo (1975) (Originally published in Japanese)
6. Kolmogorov, A.N., Fomin, S.V.: (English edn.) Introductory Real analysis. Prentice-Hall, New York (1970) (Japanese edn. translated from 4th edn.) Kansu-kaiseki no Kiso, 4th edn. Iwanami Shoten, Tokyo (1979) (Originally published in Russian). I use the Japanese edition for quotations.
7. Schwartz, L.: Méthods mathématique pour les sciences physiques. Hermann, Paris (1965)
8. Stromberg, K.R.: An Introduction to Classical Real Analysis. American Mathematical Society, Providence (1981)
9. Takagi, T.: Kaiseki Gairon (Treatise on Analysis), 3rd edn. Iwanami Shoten, Tokyo (1961) (Originally published in Japanese)
10. Titchmarsh, E.C.: Introduction to the Theory of Fourier Integrals, 2nd edn. Oxford, London (1948)

Chapter 4
Fourier Transforms (II)

In the previous chapter, we observed a peculiar relation between the smoothness and the rapidity of vanishing at infinity of a function f, as well as its Fourier transform $\hat{f}$. Based upon this observation, we introduce an important function space $\mathfrak{S}$, which is invariant under the Fourier transforms. We then proceed to $\mathfrak{L}^2$-theory of Fourier transforms due to M. Plancherel. As a simple application of Plancherel's theory, we discuss how to solve integral equations of convolution type. Finally, a tempered distribution is defined as an element of $\mathfrak{S}'$, and its Fourier transform is examined in detail.

4.1 Fourier Transforms of Rapidly Decreasing Functions

As we have already remarked, it is conjectured that if a function f is very smooth and vanishes at infinity very fast, its Fourier transform $\hat{f}$ enjoys the same properties. We now provide a rigorous proof of this conjecture.

Definition 4.1 A smooth (differentiable infinitely many times) function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is said to be **rapidly decreasing** at infinity if it satisfies

$$\sup_{x \in \mathbb{R}} \mid x^n f^{(m)}(x) \mid < \infty$$

for all $n,\ m = 0, 1, 2, \cdots$. We denote by $\mathfrak{S}(\mathbb{R})$ the set of all the rapidly decreasing functions. ($f^{(m)}$ denotes the m-th derivative of f.)

The condition $\sup_{x \in \mathbb{R}} \mid x^n f^{(m)}(x) \mid < \infty$ means that every derivative of f decreases to 0 more rapidly than any power of $1/ \mid x \mid$ as $\mid x \mid \to \infty$.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_4

We often denote by $\mathfrak{D}(\mathbb{R})$ the set of all the smooth functions $f : \mathbb{R} \longrightarrow \mathbb{C}$ with compact support. It is obvious that $\mathfrak{D} \subset \mathfrak{S}$. The function $f(x) = e^{-x^2}$ is an example which is contained in $\mathfrak{S}$ but not in $\mathfrak{D}$.[1]

It is quite easy to verify the following facts.

1° A function $f \in \mathfrak{E}(\mathbb{R})$ belongs to $\mathfrak{S}(\mathbb{R})$ if and only if

$$\sup_{x\in\mathbb{R}} |P(x) f^{(m)}(x)| < \infty$$

for all the polynomials $P(x)$ and all the nonnegative integers $m \geqq 0$.

2°

$$\lim_{|x|\to\infty} P(x) f^{(m)}(x) = 0$$

for any $f \in \mathfrak{S}(\mathbb{R})$ and any polynomial $P(x)$.

In fact, there exists some $C > 0$ such that

$$\sup_{x\in\mathbb{R}} |(1 + |x|^2) P(x) f^{(m)}(x)| \leqq C < \infty$$

by 1°, since $(1 + |x|^2) P(x)$ is a polynomial. Hence

$$|P(x) f^{(m)}(x)| \leqq \frac{C}{1 + |x|^2}$$

for all $x \in \mathbb{R}$. It implies that $\lim_{|x|\to\infty} P(x) f^{(m)}(x) = 0$. Note that the convergence to 0 is uniform with respect to m.

Theorem 4.1 (properties of $\mathfrak{S}$)

(i) *If f is an element of $\mathfrak{S}$, $x^n f^{(m)}(x)$ is bounded and integrable for any integers $n, m \geqq 0$.*

(ii) *If f is an element of $\mathfrak{S}$, $(x^n f(x))^{(m)}$ is bounded and integrable for any integers $n, m \geqq 0$.*

Proof (i) By the definition of $\mathfrak{S}$, there exists some $M > 0$ such that

$$\sup_{x\in\mathbb{R}} | x^{n+2} f^{(m)}(x) | \leqq M < \infty.$$

Consequently,

$$| x^n f^{(m)}(x) | \leqq \frac{M}{x^2} \quad \text{for all} \quad x \in \mathbb{R}.$$

[1] Where the domain and the range are clear enough, we write just $\mathfrak{S}$, $\mathfrak{D}$ and so on, for brevity.

(i) immediately follows from this.
(ii) is clear from (i) and the Leibnitz formula. □

$\mathfrak{S}$ is a vector space, the topology of which is defined by the family of seminorms

$$p_{n,m}(f) = \sup_{x\in\mathbb{R}} |\, x^n f^{(m)}(x)\, | \ ; \quad n,\ m = 0, 1, 2, \cdots .$$

Thus $\mathfrak{S}$ is a metrizable locally convex Hausdorff topological vector space. A sequence $\{f_\nu\}$ converges to 0 (in the above topology) if and only if

$$x^n f_\nu^{(m)}(x) \to 0 \quad \text{as} \quad \nu \to \infty \quad \text{(uniformly)}^2$$

for all $n, m = 0, 1, 2, \cdots$.

In the case of $n = m = 0$, in particular, we have

$$p_{0,0}(f) = \sup_{x\in\mathbb{R}} |f(x)| \quad \text{(uniform-norm).}$$

So the convergence of $\{f_\nu\}$ to f in the topology of $\mathfrak{S}$ implies the uniform convergence.

We shall use the following notations of some function spaces:

$\mathfrak{C}_\infty(\mathbb{R}, \mathbb{C})$: the space of continuous functions ($\mathbb{R} \to \mathbb{C}$) vanishing at infinity.

$\mathfrak{C}_0(\mathbb{R}, \mathbb{C})$: the space of continuous functions ($\mathbb{R} \to \mathbb{C}$) with compact support.

$\mathfrak{C}_0^\infty(\mathbb{R}, \mathbb{C}) = \mathfrak{D}(\mathbb{R})$: the space of infinitely differentiable functions ($\mathbb{R} \to \mathbb{C}$) with compact support.

We now list several peculiar relations between these function spaces from topological viewpoint. The first three results concern the uniform convergence:

1° $\mathfrak{C}_0$ is dense in $\mathfrak{C}_\infty$ with respect to the uniform convergence topology.
2° $\mathfrak{D}$ is dense in $\mathfrak{C}_0$ with respect to the uniform convergence topology.
3° $\mathfrak{D}$ is dense in $\mathfrak{S}$ with respect to the topology defined above (consequently also with respect to the uniform convergence topology).

Proof Let f be any element of $\mathfrak{S}$. We choose $\psi \in \mathfrak{D}$ so as to satisfy

$$\psi(x) = 1 \quad \text{for} \quad |x| \leqq 1.$$

[2] See Schwartz [11], Bourbaki [1] and Grothendieck [4] for the theory of locally convex spaces. See also Appendix B in this book.

Define a function f_ε, for any $\varepsilon > 0$, by

$$f_\varepsilon(x) = f(x)\psi(\varepsilon x).$$

Then $f_\varepsilon \in \mathfrak{D}$ and

$$D^\alpha(f_\varepsilon(x) - f(x)) = D^\alpha[f(x)(\psi(\varepsilon x) - 1)]$$

is a linear combination of functions of the form

$$\begin{aligned} &D^\beta f(x)\varepsilon^\gamma D^\gamma \psi(\varepsilon x) \qquad (\beta + \gamma = \alpha,\ \gamma > 0), \\ &(D^\alpha f(x))(\psi(\varepsilon x) - 1). \end{aligned}$$

Hence $f_\varepsilon \to f$ as $\varepsilon \to 0$ by the definition of the topology of $\mathfrak{S}$. □

The following two facts concern the relationships with $\mathfrak{L}^p(\mathbb{R}, \mathbb{C})$. First:

4° $\mathfrak{C}_0$ is dense in $\mathfrak{L}^p(\mathbb{R}, \mathbb{C})$ $(p \geqq 1)$ with respect to $\mathfrak{L}^p$-topology.[3]

Combining 2° and 4°, we obtain the next one:

5° $\mathfrak{D}$ is dense in $\mathfrak{L}^p(\mathbb{R}, \mathbb{C})$ $(p \geqq 1)$ with respect to $\mathfrak{L}^p$-topology. (Since $\mathfrak{D} \subset \mathfrak{S}$, $\mathfrak{S}$ is also dense in $\mathfrak{L}^p(\mathbb{R}, \mathbb{C})$.)

Theorem 4.2 is a basic result which verifies that $\mathfrak{S}$ is invariant under the Fourier transform.

Theorem 4.2 (Fourier transform on $\mathfrak{S}$) *The Fourier transform on $\mathfrak{S}$ is an automorphism on $\mathfrak{S}$. The inverse transform is given by*

$$\tilde{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{ixt}dt\ ,\ \ f \in \mathfrak{S}.$$

*(This is called the **Fourier inversion formula**.)*

Proof Since $f \in \mathfrak{S}$ is an integrable function, we can define its Fourier transform $\hat{f}$ by

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-itx}dx\ ,\ \ f \in \mathfrak{S}. \tag{4.1}$$

$\hat{f}$ is differentiable infinitely many times. In fact, the n-th formal derivative of the integrand of (4.1) is

$$(-i)^n x^n f(x)e^{-itx}, \tag{4.2}$$

[3] A more general theorem is explained in Maruyama [9] pp. 232–233.

the absolute value of which is $| \; x^n f(x) \; |$ (independent of t). Since $x^n f(x)$ is integrable by Theorem 4.1, the integral of (4.2) over $\mathbb{R}$ is uniformly convergent in t. Hence, by applying the formula of differentiation under an (improper) integral,[4] we obtain

$$\hat{f}^{(n)}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-i)^n x^n f(x) e^{-itx} dx. \tag{4.3}$$

Thus $\hat{f}$ is differentiable arbitrarily many times.

By Theorem 3.4 (or Corollary 3.1, p. 58),

$$i^m t^m \hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f^{(m)}(x) e^{-itx} dx. \tag{4.4}$$

It follows that

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-itx} \frac{d^m}{dx^m} (x^n f(x)) dx & \\ &= i^m t^m \cdot \widehat{x^n f}(t) \\ &= i^m t^m \cdot \hat{f}^{(n)}(t) \cdot (-i)^{-n} \qquad \text{(by (4.3))} \\ &= i^{m+n} t^m \hat{f}^{(n)}(t), \end{aligned} \tag{4.5}$$

which gives

$$| \; t^m \hat{f}^{(n)}(t) \; | \leqq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |(x^n f(x))^{(m)}| dx. \tag{4.6}$$

Since $(x^n f(x))^{(m)}$ is integrable by Theorem 4.1(ii), the right-hand side of (4.6) is finite, and so

$$\sup_{t \in \mathbb{R}} | \; t^m \hat{f}^{(n)}(t) \; | < \infty.$$

This holds good for any $m, n = 0, 1, 2, \cdots$. We can conclude that $\hat{f} \in \mathfrak{S}$.

The continuity of the Fourier transform $\mathcal{F} : \mathfrak{S} \to \mathfrak{S}$ is easily seen from (4.6). In fact, if $\{f_\nu\}$ is a sequence in $\mathfrak{S}$ which converges to 0, then it satisfies, by the Leibnitz formula:

$$\sup_{x \in \mathbb{R}} | \; (x^\nu f_\nu(x))^{(m)} \; | \to 0 \quad \text{as} \quad \nu \to \infty$$

[4] Takagi [13] p. 166.

for any $m, n = 0, 1, 2, \cdots$. Hence we have, by (4.6),

$$\sup_{t \in \mathbb{R}} | t^m \hat{f}_\nu^{(n)}(t) | \to 0 \quad \text{as} \quad \nu \to \infty,$$

which implies that $\{\hat{f}_\nu\}$ converges to 0 in $\mathfrak{S}$. This proves the continuity of the Fourier transform $\mathcal{F}$.

We define the operator $\tilde{\mathcal{F}}$ which associates with each $f \in \mathfrak{S}$ the function

$$\tilde{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{ixt} dt.$$

The difference between $\mathcal{F}$ and $\tilde{\mathcal{F}}$ is just the sign of the exponential. Obviously,

$$\tilde{\mathcal{F}}(f)(x) = \mathcal{F}(f)(-x).$$

It is clear that $\tilde{\mathcal{F}}$ is a continuous linear operator of $\mathfrak{S}$ into itself.

Finally, we have to show that $\tilde{\mathcal{F}}$ is the inverse of $\mathcal{F}$. It is enough to show that

$$\tilde{\mathcal{F}} \circ \mathcal{F} = I \quad ; \quad \text{i.e. } \tilde{\hat{f}} = f \quad \text{for all} \quad f \in \mathfrak{S}, \tag{4.7}$$

$$\mathcal{F} \circ \tilde{\mathcal{F}} = I \quad ; \quad \text{i.e. } \hat{\tilde{f}} = f \quad \text{for all} \quad f \in \mathfrak{S}, \tag{4.8}$$

where I is the identity operator on $\mathfrak{S}$. We prove only (4.7) ((4.8) can be proved similarly), which is equivalent to

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(t) e^{ixt} dt = f(x) \quad \text{for all} \quad f \in \mathfrak{S}. \tag{4.7$'$}$$

For any $f, \; g \in \mathfrak{S}$ and any $\varepsilon > 0$, we have

$$\begin{aligned}
\int_{-\infty}^{\infty} g(\varepsilon t) \hat{f}(t) e^{ixt} dt &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\varepsilon t) \left\{ \int_{-\infty}^{\infty} f(y) e^{-ity} dy \right\} e^{ixt} dt \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(\varepsilon t) e^{-it(y-x)} dt \right\} f(y) dy \qquad (4.9) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(\varepsilon t) e^{-i(\varepsilon t)\frac{y-x}{\varepsilon}} dt \right\} f(y) dy.
\end{aligned}$$

Changing the variables by $w = \varepsilon t$, $z = (y - x)/\varepsilon$, (4.9) is rewritten as

$$\int_{-\infty}^{\infty} g(\varepsilon t)\hat{f}(t)e^{ixt}dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} g(w)e^{-iwz}dw\right\} f(x+\varepsilon z)dz$$

$$= \int_{-\infty}^{\infty} \hat{g}(z) f(x+\varepsilon z)dz. \tag{4.10}$$

In a particular case of $g(t) = e^{-t^2/2}$, we obtain

$$g(0)\int_{-\infty}^{\infty} \hat{f}(t)e^{ixt}dt = f(x)\int_{-\infty}^{\infty} \hat{g}(z)dz \tag{4.11}$$

by passing to the limit $\varepsilon \downarrow 0$ and the dominated convergence theorem. We note that $g(0) = 1$ and

$$\int_{-\infty}^{\infty} \hat{g}(z)dz = \sqrt{2\pi}$$

by Example 3.3 (p. 54). Therefore

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \hat{f}(t)e^{ixt}dt = f(x).$$

This completes the proof.[5] □

Henceforth we denote the operator $\tilde{\mathcal{F}} : f \longmapsto \tilde{f}$ $(\mathfrak{S} \to \mathfrak{S})$ by $\mathcal{F}^{-1}$ and call it the **inverse Fourier transform**.

Remark 4.1 It is clear that $\mathfrak{S} = \{\hat{f} | f \in \mathfrak{S}\}$ is dense in $\mathfrak{C}_\infty$ (by 1° on p. 67). Furthermore, if we write

$$\mathfrak{S}_c \equiv \{f \in \mathfrak{S} | \text{the support of } \hat{f} \text{ is compact}\},$$

$\mathfrak{S}_c$ and $\{\hat{f} | f \in \mathfrak{S}_c\}$ are dense in $\mathfrak{S}$ (and hence in $\mathfrak{C}_\infty$) with respect to the uniform convergence topology.

[5] The proof of (4.7) here is due to Yosida [16], p. 147. Although this approach is a little bit technical, I adopt it because of its simplicity. There are various other approaches. For instance, see Kawata [6] Chap. 11, Treves [14] Theorem 25.1 or Maruyama [8] Theorem 4.13.

The proof is as follows. If f is any element of $\mathfrak{S}$, then $\hat{f} \in \mathfrak{S}$. Hence there exists a sequence $\{g_n\}$ of $\mathfrak{D}$ such that $g_n \to \hat{f}$ (in $\mathfrak{S}$-topology) (cf. 3° on p. 67). Let $f_n \in \mathfrak{S}$ be the inverse Fourier transform of g_n; i.e. $\hat{f_n} = g_n$. Then $f_n \in \mathfrak{S}_c$. $\{f_n\}$ uniformly converges to f, since $f_n \to f$(in $\mathfrak{S}$-topology) by the continuity of the inverse Fourier transform. This proves that $\mathfrak{S}_c$ is dense in $\mathfrak{S}$.

The denseness of $\{\hat{f} | f \in \mathfrak{S}_c\}$ in $\mathfrak{S}$ can be proved exactly in the same way.

4.2 Fourier Transforms on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$

It may not be possible to define the Fourier transform of a square-integrable function, since $\mathfrak{L}^2(\mathbb{R}, \mathbb{C}) \not\subset \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. However, the Plancherel theorem guarantees the existence of the Fourier transform of a function in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ in a suitable sense.

Theorem 4.3 (Plancherel) *The Fourier transform $\mathcal{F}$ on $\mathfrak{S}$ can be uniquely extended to a bounded linear operator of $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ into itself. This is called the Fourier transform on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ and is denoted by $\mathcal{F}_2$. Similarly, the inverse Fourier transform $\mathcal{F}^{-1}$ can be uniquely extended as a bounded linear operator of $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ into itself, denoted by $\mathcal{F}_2^{-1}$.*

The following properties hold good for $f, \; g \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$:

(i) $\langle \mathcal{F}_2 f \;, \; g \rangle = \langle f, \mathcal{F}_2^{-1} g \rangle$. *($\langle \cdot, \cdot \rangle$ is the inner product in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$.)*
(ii) $\| \mathcal{F}_2 f \|_2 = \| f \|_2$.
(iii) $\mathcal{F}_2 : \mathfrak{L}^2(\mathbb{R}, \mathbb{C}) \to \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ *is a bijection.*
(iv) $\mathcal{F}_2^{-1}$ *is the inverse of* $\mathcal{F}_2$.
(v) $(\mathcal{F}_2 f)(t) = \underset{r\to\infty}{\text{l.i.m.}} \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-r}^{r} f(x) e^{-itx} dx.$ (l.i.m. is the limit with respect to $\mathfrak{L}^2$-norm.)
(vi) $(\mathcal{F}_2^{-1} f)(x) = \underset{r\to\infty}{\text{l.i.m.}} \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-r}^{r} f(t) e^{ixt} dt.$

Proof We denote by $\mathcal{F}$ and $\mathcal{F}^{-1}$ the Fourier transform and the inverse Fourier transform on $\mathfrak{S}$, respectively. Then it is easy to check that

$$\begin{aligned} \langle \mathcal{F} f, \; g \rangle &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f(x) e^{-itx} dx \right\} \overline{g(t)} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \left\{ \overline{\int_{-\infty}^{\infty} g(t) e^{ixt} dt} \right\} dx \\ &= \langle f, \; \mathcal{F}^{-1} g \rangle \quad \text{for all} \quad f, g \in \mathfrak{S}. \end{aligned} \tag{4.12}$$

If $g = \mathcal{F} f$, in particular, (4.12) implies that

$$\| \mathcal{F} f \|_2 = \| f \|_2 \quad \text{for all} \quad f \in \mathfrak{S}. \tag{4.13}$$

Hence $\mathcal{F} : \mathfrak{S} \to \mathfrak{S}$ is an isometric (with respect to $\mathfrak{L}^2$-norm) operator. $\mathfrak{S}$ being dense in $\mathfrak{L}^2$ (by 5° on p. 68),[6] $\mathcal{F}$ can be uniquely extended to a bounded linear operator $\mathcal{F}_2$ on $\mathfrak{L}^2$, preserving the norm. Similarly, $\mathcal{F}^{-1} : \mathfrak{S} \to \mathfrak{S}$ admits the unique extension to a bounded linear operator $\mathcal{F}_2^{-1}$ on $\mathfrak{L}^2$. (Note that whether $\mathcal{F}_2^{-1}$ is the inverse of $\mathcal{F}_2$ has not been verified yet. It is proved in (iv).)

(i) Suppose that $f, g \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. Since $\mathfrak{S}$ is dense in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$, there are sequences $\{f_n\}, \{g_n\}$ in $\mathfrak{S}$ such that

$$\underset{n}{\text{l.i.m.}} f_n = f, \quad \underset{n}{\text{l.i.m.}} g_n = g.$$

According to (4.12), we have

$$\langle \mathcal{F} f_n, g_n \rangle = \langle f_n, \mathcal{F}^{-1} g_n \rangle \quad \text{for all} \quad n.$$

Hence (taking account of $\mathcal{F} = \mathcal{F}_2, \mathcal{F}^{-1} = \mathcal{F}_2^{-1}$ on $\mathfrak{S}$) we obtain

$$\langle \mathcal{F}_2 f_n, g_n \rangle = \langle f_n, \mathcal{F}_2^{-1} g_n \rangle \quad \text{for all} \quad n. \tag{4.14}$$

Passing to the limit of (4.14) as $n \to \infty$, we get (i) by the continuities of $\mathcal{F}_2, \mathcal{F}_2^{-1}$ and the inner product.

(ii) Obvious from (4.13).

(iii) The injectiveness of $\mathcal{F}_2$ immediately follows from (ii). For any $g \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$, there exists a sequence $\{g_n\}$ in $\mathfrak{S}$ which satisfies l.i.m.$g_n = g$. Taking account of the fact that $\mathcal{F} : \mathfrak{S} \to \mathfrak{S}$ is an automorphism, we define $f_n = \mathcal{F}^{-1} g_n$. Since $\{f_n\}$ is a Cauchy sequence in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ by (ii), there exists some $f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ such that $\underset{n}{\text{l.i.m.}} f_n = f$. By the continuity of $\mathcal{F}_2$, we have

$$g_n = \mathcal{F}_2 f_n = \mathcal{F} f_n \to \mathcal{F}_2 f = g.$$

This shows that $\mathcal{F}_2$ is a surjection.

(iv) We have to show that $\mathcal{F}_2^{-1} \circ \mathcal{F}_2 = \mathcal{F}_2 \circ \mathcal{F}_2^{-1} = I$. For any $f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$, there exists a sequence $\{f_n\}$ in $\mathfrak{S}$ which converges to f (in $\mathfrak{L}^2$). Then

$$(\mathcal{F}_2^{-1} \circ \mathcal{F}_2) f_n = (\mathcal{F}^{-1} \circ \mathcal{F}) f_n = f_n. \tag{4.15}$$

Passing to the limit, (4.15) gives

$$(\mathcal{F}_2^{-1} \circ \mathcal{F}_2) f = f.$$

[6] A direct proof is also possible. Since $f \in \mathfrak{S}$ is bounded and integrable, $f \cdot f$ is integrable (i.e. $f \in \mathfrak{L}^2$) by Hölder's inequality. We have only to approximate $f \in \mathfrak{L}^2$ by a simple function φ with compact support and to approximate φ by a smooth function with compact support.

Similarly for $(\mathcal{F}_2 \circ \mathcal{F}_2^{-1})f = f$.

(v) Let $f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. We define $f_r(x)$ by

$$f_r(x) = \begin{cases} f(x) & \text{if} \quad |x| \leqq r, \\ 0 & \text{if} \quad |x| > r. \end{cases}$$

Then there exists a sequence $\{\varphi_n\}$ in $\mathfrak{S}$ such that:

a. $\{\varphi_n\}$ converges to f_r in $\mathfrak{L}^2$, and

b. there exists a bounded interval J containing $[-r, \ r]$ which satisfies supp $\varphi_n \subset J$ for all n.

Since $\{\varphi_n\}$ converges to f_r in $\mathfrak{L}^1$,

$$\left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi_n(x) e^{-itx} dx - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f_r(x) e^{-itx} dx \right|$$
$$\leqq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |\varphi_n(x) - f_r(x)| \, dx \to 0$$
$$\text{as} \quad n \to \infty.$$

Consequently, we obtain

$$\mathcal{F}_2 f_r(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f_r(x) e^{-itx} dx = \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} f(x) e^{-itx} dx,$$

by extension by continuity. Thus (v) follows from $\| f_r - f \|_2 \to 0$ and the continuity of $\mathcal{F}_2$.

(vi) can be proved similarly. □

Remark 4.2

1° It is clear that

$$\langle \mathcal{F}_2 f, \mathcal{F}_2 g \rangle = \langle f, g \rangle$$

for any $f, g \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$.

2° For any $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C}) \cap \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$, the Fourier transform $\hat{f}$ in the usual sense and the Fourier transform $\mathcal{F}_2 f$ in the sense of Plancherel coincide. In fact, since $\| f_r - f \|_1 \to 0$, we have $\| \hat{f}_r - \hat{f} \|_\infty \to 0$ by (3.21) on p. 52. On the other hand, $\| \hat{f}_r - \mathcal{F}_2 f \|_2 \to 0$ as shown above. Hence $\hat{f} = \mathcal{F}_2 f$.[7]

[7] The proof here is due to Dunford–Schwartz [2] III, pp. 1988–1989. See also Sect. 4.4 in this chapter and Schwartz [12] Chap. VII for the connection with the theory of distributions.

4.3 Application: Integral Equations of Convolution Type

In this section, we consider how to solve an integral equation of convolution type:

$$f(x) = g(x) + \int_{-\infty}^{\infty} K(x-y) f(y) dy, \tag{4.16}$$

where the functions $g, K : \mathbb{R} \to \mathbb{C}$ are given. We wish to find a function f which satisfies (4.16). This problem provides a typical sample of applications of Plancherel's theorem. We start with a rough sketch of ideas.

The Fourier transforms of both sides of (4.16) give

$$\hat{f}(t) = \hat{g}(t) + \sqrt{2\pi}\,\hat{f}(t) \cdot \hat{K}(t)$$

from which $\hat{f}$ is expressed as

$$\hat{f}(t) = \frac{\hat{g}(t)}{1 - \sqrt{2\pi}\,\hat{K}(t)}.$$

Hence we have, by the inverse Fourier transform,

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\hat{g}(t)}{1 - \sqrt{2\pi}\,\hat{K}(t)} e^{ixt} dt.$$

This f seems to be a solution of (4.16).

However, we have to admit that there are several obscure points in the above reasoning. First of all, it is required that the Fourier transforms in a certain sense can be defined for g, K and f. In particular, we have to apply the inverse Fourier transform at the final step. Is it possible? We have to specify some suitable function spaces which permit these operations. Furthermore, it is very possible that $1 - \sqrt{2\pi}\,\hat{K}(t) = 0$, which is an obstacle to computing $\hat{f}(t)$.

Taking account of the difficulties stated above, we assume a couple of conditions:

(i) $g \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$.
(ii) $K \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$; $\quad |\hat{K}(t)| \leqq C < \frac{1}{\sqrt{2\pi}}$ for all $t \in \mathbb{R}$.

We would like to find a solution of (4.16) in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ under these conditions. We use the notation $\widehat{\ }$ for denoting Fourier transforms of functions of either $\mathfrak{L}^1$ or $\mathfrak{L}^2$, because no confusion may occur.

Under the conditions (i) and (ii), we have

$$\frac{\hat{g}(t)}{1 - \sqrt{2\pi}\,\hat{K}(t)} \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C}).$$

Hence, by Theorem 4.3,

$$f(x) \equiv \underset{r\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} \frac{\hat{g}(t)}{1-\sqrt{2\pi}\hat{K}(t)} e^{ixt} dt$$

can be defined and $f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. We can also define a function h by

$$h(x) = \int_{-\infty}^{\infty} K(x-y) f(y) dy$$

and this h belongs to $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$.[8]

The Fourier transform of h in the sense of Plancherel is given by

$$\hat{h}(t) = \underset{r\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} \left\{ \int_{-\infty}^{\infty} K(x-y) f(y) dy \right\} e^{-itx} dx = \sqrt{2\pi}\hat{K}(t)\hat{f}(t). \tag{4.17}$$

This can be proved as follows. If we define

$$f_A = f \cdot \chi_{[-A,A]},$$

$f_A \in \mathfrak{L}^1 \cap \mathfrak{L}^2$ and

$$\underset{A\to\infty}{\text{l.i.m.}} f_A = f. \tag{4.18}$$

It follows from Theorem 4.3 (continuity of Fourier transform) that

$$\underset{A\to\infty}{\text{l.i.m.}} \hat{f}_A = \hat{f} \quad \text{(in the sense of Plancherel).}$$

Define a function $h_A(x)$ by

$$h_A(x) = \int_{-\infty}^{\infty} K(x-y) f_A(y) dy.$$

Then the Fourier transform formula of the convolution of $\mathfrak{L}^1$-functions gives

$$\hat{h}_A = \sqrt{2\pi}\hat{K} \cdot \hat{f}_A.$$

[8] In general, it holds good that $u * v \in \mathfrak{L}^p(\mathbb{R}, \mathbb{C})$ and $\|u * v\|_p \leqq \|u\|_p \cdot \|v\|_1$ for any $u \in \mathfrak{L}^p(\mathbb{R}, \mathbb{C})$ $(1 \leqq p \leqq \infty)$ and $v \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. See Maruyama [9] pp. 235–236.

Since[9]

$$\underset{A\to\infty}{\text{l.i.m.}} h_A = h, \tag{4.19}$$

we obtain, applying Theorem 4.3 again,

$$\underset{A\to\infty}{\text{l.i.m.}} \hat{h}_A = \underset{A\to\infty}{\text{l.i.m.}} \sqrt{2\pi}\hat{K}\cdot\hat{f}_A = \hat{h}. \tag{4.20}$$

By the boundedness of $\hat{K}$, we can conclude

$$\hat{h} = \sqrt{2\pi}\hat{K}\cdot\hat{f}.$$

This proves (4.17).

Since

$$\hat{f}(t) = \frac{\hat{g}(t)}{1-\sqrt{2\pi}\hat{K}(t)},$$

we obtain

$$\hat{h}(t) = \frac{\sqrt{2\pi}\hat{K}(t)\hat{g}(t)}{1-\sqrt{2\pi}\hat{K}(t)}.$$

Consequently,

$$(\widehat{g+h})(t) = \hat{g}(t) + \frac{\sqrt{2\pi}\hat{g}(t)\hat{K}(t)}{1-\sqrt{2\pi}\hat{K}(t)} = \frac{\hat{g}(t)}{1-\sqrt{2\pi}\hat{K}(t)} = \hat{f}(t).$$

---

[9]The relation (4.19) can be verified as follows. In general, if $u_r \in \mathfrak{L}^2(\mathbb{R},\mathbb{C})$, and $v \in \mathfrak{L}^1(\mathbb{R},\mathbb{C})$, and $\underset{r\to\infty}{\text{l.i.m.}} u_r = u$, then

$$\underset{r\to\infty}{\text{l.i.m.}} \int v(x-z)u_r(z)dz = \int v(x-z)u(z)dz. \tag{†}$$

In fact, by the footnote just above,

$$\begin{aligned}\left\| \int v(x-z)u_r(z)dz - \int v(x-z)u(z)dz \right\|_2 &= \left\| \int [u_r(z)-u(z)]v(x-z)dz \right\|_2 \\ &\leqq \|u_r-u\|_2\cdot\|v\|_1 \to 0 \quad \text{as} \quad r\to\infty,\end{aligned}$$

which establishes (†). Now (4.19) immediately follows from (†).

Taking the inverse Fourier transforms of both sides in the sense of Plancherel, we have

$$f = g + h,$$
$$\text{i.e.} \quad f(x) = g(x) + \int_{-\infty}^{\infty} K(x-y)f(y)dy.$$

This proves that f is a solution of the integral equation (4.16).[10]

4.4 Fourier Transforms of Tempered Distributions

We now proceed to the theory of Fourier transforms of distributions in the sense of L. Schwartz. The readers can find a brief exposition of Schwartz's theory in the appendices of this book.

Let Ω be an open set in $\mathbb{R}$. We denote by $\mathfrak{E}(\Omega)$ the set of all the infinitely differentiable complex-valued functions defined on Ω. For each compact set $K \subset \Omega$ and $m \in \mathbb{N} \cup \{0\}$, we define a seminorm $p_{K,m}$ by

$$p_{K,m}(\varphi) = \sup_{\substack{x \in K \\ s \leqq m}} |D^s \varphi(x)|.$$

The space $\mathfrak{E}(\Omega)$ endowed with the topology by the family of seminorms $\{p_{K,m}\}$ is a locally convex Hausdorff topological vector (or linear) space (LCHTVS). Furthermore, this topology is completely metrizable. So $\mathfrak{E}(\Omega)$ is a Fréchet space (cf. Appendix B, Sect. B.1).

Let K be a compact set in Ω. We denote by $\mathfrak{D}_K(\Omega)$ the set of all the infinitely differentiable functions whose supports are contained in K; i.e.

$$\mathfrak{D}_K(\Omega) = \{\varphi \in \mathfrak{E}(\Omega) | \text{supp}\varphi \subset K\}.$$

$\mathfrak{D}_K(\Omega)$ is a linear subspace of $\mathfrak{E}(\Omega)$. $\mathfrak{D}_K(\Omega)$ is also a LCHTVS, the topology of which is defined by the relative topology induced by $\mathfrak{E}(\Omega)$. It is defined by the family of seminorms

$$p_{K,m}(\varphi) = \sup_{\substack{x \in K \\ s \leqq m}} |D^s \varphi(x)|, \quad m \in \mathbb{N} \cup \{0\}.$$

$\mathfrak{D}_K(\Omega)$ is also a Fréchet space.

---

[10] This section is basically due to Kawata [5]II, Chap. 17. However, the proof of (4.19) given there does not seem correct.

Construct a sequence $\{K_1, K_2, \cdots\}$ of compact sets in Ω which satisfies

(i) $K_n \subset \text{int}.K_{n+1}$, and

(ii) $\bigcup_{n=1}^{\infty} K_n = \Omega$.

We denote by $\mathfrak{D}(\Omega)$ the set of all elements of $\mathfrak{E}(\Omega)$ with compact support. Then it is the union of $\mathfrak{D}_{K_n}(\Omega)$ $(n = 1, 2, \cdots)$; i.e.

$$\mathfrak{D}(\Omega) = \bigcup_{n=1}^{\infty} \mathfrak{D}_{K_n}(\Omega).$$

The topology of each $\mathfrak{D}_{K_n}(\Omega)$ is explained above. On the other hand, the topology of $\mathfrak{D}(\Omega)$ is given as the strict inductive limit topology induced by $\{\mathfrak{D}_{K_n}(\Omega)\}$. This topology is uniquely determined independently of the choice of $\{K_n\}$ (cf. Appendix B, Sect. B.1). Each element of $\mathfrak{D}(\Omega)$ is called a **test function**.

$\mathfrak{D}(\Omega)$ is a LCHTVS, but is not metrizable. A net $\{\varphi_\alpha\}_{\alpha \in A}$ (A is a directed set) in $\mathfrak{D}(\Omega)$ converges to some $\varphi^* \in \mathfrak{D}(\Omega)$ if and only if there exists a compact set $K \subset \Omega$ (independent of α) such that

$$\text{supp}\varphi_\alpha \subset K \quad \text{for all} \quad \alpha \in A \tag{4.21}$$

and

$$D^s \varphi_\alpha \to D^s \varphi^* \quad \text{(uniform convergence)} \tag{4.22}$$

for all $s \in \mathbb{N} \cup \{0\}$.

An element of the dual space $\mathfrak{D}(\Omega)'$ of $\mathfrak{D}(\Omega)$ is called a **distribution** (or a **generalized function**) on Ω.

For instance, a locally integrable function $f : \Omega \to \mathbb{C}$ defines a distribution $T_f : \mathfrak{D}(\Omega) \to \mathbb{C}$ by

$$T_f(\varphi) = \int_\Omega f(x)\varphi(x)dx, \quad \varphi \in \mathfrak{D}(\Omega).$$

We sometimes denote the distribution T_f simply by f if no confusion occurs.

For any distribution T, we define another distribution S by

$$S(\varphi) = -T(\varphi'), \quad \varphi \in \mathfrak{D}(\Omega).$$

S is called the **generalized derivative** or **distributional derivative** of T and denoted by T'. Any distribution admits a distributional derivative of any order. The s-th distributional derivative, denoted by $D^s T$ or $T^{(s)}$, is defined by

$$D^s T(\varphi) = (-1)^s T(D^s \varphi), \quad \varphi \in \mathfrak{D}(\Omega),$$

where $D^s\varphi$ is the s-th derivative of φ in the usual sense. The basic motivation to define the concept of distributional derivative in this way is explained in Appendix B, Sect. B.4.

If we restrict $T \in \mathfrak{S}(\mathbb{R})'$ to $\mathfrak{D}(\mathbb{R}) \subset \mathfrak{S}(\mathbb{R})$, T is continuous on $\mathfrak{D}(\mathbb{R})$ with respect to its topology (defined by the strictly inductive limit). Observe that the identity mapping of $\mathfrak{D}(\mathbb{R})$ into $\mathfrak{S}(\mathbb{R})$ is continuous. The topology of $\mathfrak{S}(\mathbb{R})$ is defined in Sect. 4.1. In fact, if a net $\{\varphi_\alpha\}$ in $\mathfrak{D}(\mathbb{R})$ converges to $0 \in \mathfrak{D}(\mathbb{R})$, then (4.21) and (4.22) must hold good. Hence

$$\sup_{x\in\mathbb{R}} |x^n \varphi_\alpha^{(m)}(x)| = \sup_{x\in K} |x^n \varphi_\alpha^{(m)}(x)| \to 0$$

for any $m = 0, 1, 2, \cdots$. This implies that $\{\varphi_\alpha\}$ converges to 0 in the topology of $\mathfrak{S}(\mathbb{R})$. Thus we conclude that the identity $I : \mathfrak{D}(\mathbb{R}) \to \mathfrak{S}(\mathbb{R})$ is continuous. In other words, the proper topology of $\mathfrak{D}(\mathbb{R})$ is stronger than the relative topology induced by $\mathfrak{S}(\mathbb{R})$. So $T|_{\mathfrak{D}(\mathbb{R})}$ is continuous, that is, $T|_{\mathfrak{D}(\mathbb{R})}$ is a distribution.

Conversely, if a distribution $T \in \mathfrak{D}(\mathbb{R})'$ is continuous with respect to the relative topology of $\mathfrak{D}(\mathbb{R})$ induced by $\mathfrak{S}(\mathbb{R})$, T can be uniquely extended to an element of $\mathfrak{S}(\mathbb{R})'$ since $\mathfrak{D}(\mathbb{R})$ is dense in $\mathfrak{S}(\mathbb{R})$. (cf. 3° on p. 67.)

Based upon the reasoning so far discussed, we can identify $\mathfrak{S}(\mathbb{R})'$ with a subspace of $\mathfrak{D}(\mathbb{R})'$ by identifying $T \in \mathfrak{S}(\mathbb{R})'$ with its unique extension to a distribution $\in \mathfrak{D}(\mathbb{R})'$.

$$\mathfrak{S}(\mathbb{R})' \subset \mathfrak{D}(\mathbb{R})'.$$

Definition 4.2 An element of $\mathfrak{S}(\mathbb{R})'$ is called a **tempered distribution**.

The motivation for this terminology will be made clear soon.

Let us show several examples.

Example 4.1 An integrable function $f : \mathbb{R} \to \mathbb{C}$ defines a tempered distribution. In fact, the functional

$$T_f\varphi = \int_{\mathbb{R}} f(x)\varphi(x)dx, \quad \varphi \in \mathfrak{S}(\mathbb{R}),$$

which makes sense, is continuous with respect to the topology of $\mathfrak{S}(\mathbb{R})$.

Example 4.2 A bounded measurable function $f : \mathbb{R} \to \mathbb{C}$ defines a tempered distribution. In fact, the functional

$$T_f\varphi = \int_{\mathbb{R}} f(x)\varphi(x)dx, \quad \varphi \in \mathfrak{S}(\mathbb{R}),$$

which makes sense because any $\varphi \in \mathfrak{S}(\mathbb{R})$ is integrable, is continuous.

Example 4.3 Assume that a locally integrable function $f : \mathbb{R} \to \mathbb{C}$ satisfies

$$\lim_{|x|\to\infty} |x|^{-k}|f(x)| = 0, \tag{4.23}$$

and so there exists some constant $A > 0$ such that

$$|f(x)| \leqq A|x|^k \quad \text{for large } \ x.$$

Then f defines a tempered distribution. It can be verified as follows. If $\varphi \in \mathfrak{S}(\mathbb{R})$, we have

$$\sup_{x\in\mathbb{R}}|x^{k+2}\varphi(x)| \leqq B \quad \text{for some} \quad B < \infty$$

and so

$$|\varphi(x)| \leqq \frac{B}{|x|^{k+2}} \quad \text{for all} \quad x \in \mathbb{R}.$$

Hence it holds good that

$$|f(x)\varphi(x)| \leqq \frac{A \cdot B}{|x|^2} \quad \text{for large } \ x.$$

Thus the functional

$$T_f\varphi = \int_{\mathbb{R}} f(x)\varphi(x)dx, \quad \varphi \in \mathfrak{S}(\mathbb{R})$$

makes sense, and is continuous on $\mathfrak{S}(\mathbb{R})$.

The concept of tempered distribution is a generalization of a distribution defined by a locally integrable function which is slowly increasing in the sense of (4.23).

Example 4.4 Assume that μ is a Borel measure on $\mathbb{R}$ which satisfies

$$\int_{\mathbb{R}} (1 + |x|^2)^{-k} d\mu < \infty$$

for some nonnegative integer k. Then μ defines a tempered distribution by the relation

$$T_\mu(\varphi) = \int_{\mathbb{R}} \varphi d\mu, \quad \varphi \in \mathfrak{S}(\mathbb{R}).$$

(Such a measure is called a **tempered measure**.)

It is almost obvious because any rapidly decreasing function φ satisfies

$$\varphi(x) = o((1 + |x|^2)^{-k})$$

for sufficiently large $|x|$.

Example 4.5 Let $f : \mathbb{R} \to \mathbb{C}$ be a function of the class $\mathfrak{C}^\infty$. f is said to be **slowly increasing** if there exists some $k \in \mathbb{N} \cup \{0\}$, for each $m \in \mathbb{N} \cup \{0\}$, such that

$$\lim_{|x|\to\infty} \frac{1}{|x|^k} |f^{(m)}(x)| = 0.$$

A slowly increasing function f defines a tempered distribution T_f.

Fix a slowly increasing function f. Let S be an operator which transforms each $\varphi \in \mathfrak{S}$ to $f\varphi$. S is a continuous linear operator of $\mathfrak{S}$ into itself. For a slowly increasing function f and $T \in \mathfrak{S}'$, their **product** fT is defined by

$$fT : \varphi \mapsto T(f\varphi), \quad \varphi \in \mathfrak{S}.$$

fT is an element of $\mathfrak{S}'$.

Example 4.6 A distribution (on $\mathbb{R}$) with compact support is a tempered distribution. It is obvious that $\mathfrak{D}(\mathbb{R}) \subset \mathfrak{S}(\mathbb{R}) \subset \mathfrak{E}(\mathbb{R})$. We have only to observe the following two facts:

1° The proper topology of $\mathfrak{S}(\mathbb{R})$ is stronger than the relative topology of $\mathfrak{S}(\mathbb{R})$ induced by $\mathfrak{E}(\mathbb{R})$.

2° Since $\mathfrak{D}(\mathbb{R})$ is dense in $\mathfrak{E}(\mathbb{R})$ (cf. Lemma C.5, p. 399), $\mathfrak{S}(\mathbb{R})$ is also dense in $\mathfrak{E}(\mathbb{R})$. Hence $\mathfrak{E}(\mathbb{R})'$ can be regarded as a subspace of $\mathfrak{S}(\mathbb{R})'$; i.e. $\mathfrak{E}(\mathbb{R})' \subset \mathfrak{S}(\mathbb{R})' \subset \mathfrak{D}(\mathbb{R})'$.

The derivatives of a tempered distribution are defined in the same way as the usual distributional derivatives. The operator

$$\varphi \mapsto D^p\varphi, \quad \varphi \in \mathfrak{S}(\mathbb{R})$$

is a continuous linear operator of $\mathfrak{S}(\mathbb{R})$ into $\mathfrak{S}(\mathbb{R})$ for any nonnegative integer $p \geqq 0$. Hence, given $T \in \mathfrak{S}(\mathbb{R})'$, it is possible to define an operator

$$\varphi \mapsto T(D^p\varphi), \quad \varphi \in \mathfrak{S}(\mathbb{R})$$

which is continuous. Thus we can define derivatives of tempered distributions as follows.

Definition 4.3 Let p be a nonnegative integer. The p-th **distributional derivative** of $T \in \mathfrak{S}(\mathbb{R})'$ is defined by

$$D^pT(\varphi) = (-1)^pT(D^p\varphi), \quad \varphi \in \mathfrak{S}(\mathbb{R}).$$

If T is a tempered distribution, so is D^pT.

We can now define the Fourier transform of a tempered distribution. We should recall the fact that the Fourier transform $\mathcal{F}$ of $\mathfrak{S}(\mathbb{R})$ into $\mathfrak{S}(\mathbb{R})(\varphi \mapsto \hat{\varphi})$ is an automorphism of $\mathfrak{S}(\mathbb{R})$. So if $T \in \mathfrak{S}(\mathbb{R})'$, the operator $\hat{T}$ defined by

$$\hat{T} : \varphi \mapsto T(\hat{\varphi}), \quad \varphi \in \mathfrak{S}(\mathbb{R})$$

is also a tempered distribution. We call this operator $\hat{T}$ the Fourier transform of T.

Definition 4.4 Let T be a tempered distribution. The tempered distribution $\hat{T}$ defined by

$$\hat{T}(\varphi) = T(\hat{\varphi}), \quad \varphi \in \mathfrak{S}(\mathbb{R})$$

is called the **Fourier transform** of T.

The Fourier transform on $\mathfrak{S}(\mathbb{R})'$ is an automorphism of $\mathfrak{S}(\mathbb{R})'$ into itself. In order to see this clearly, we introduce the concept of dual operators.

Definition 4.5 Let $\mathfrak{X}$ and $\mathfrak{Y}$ be two topological vector (linear) spaces, and $A : \mathfrak{X} \to \mathfrak{Y}$ a continuous linear operator. Then the operator $A' : \mathfrak{Y}' \to \mathfrak{X}'$ defined by

$$A' : y' \mapsto y' \circ A, \quad y' \in \mathfrak{Y}'$$

is called the **dual operator** of A.

Theorem 4.4 *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be a pair of topological vector spaces, and $A : \mathfrak{X} \to \mathfrak{Y}$ a continuous linear operator. Then the dual operator $A' : \mathfrak{Y}' \to \mathfrak{X}'$ is continuous in the strong topology.*[11]

Proof Let $\{y'_\alpha\}$ be a net on $\mathfrak{Y}'$, which converges to 0 in the strong topology. Let B be any bounded set in $\mathfrak{X}$. Then $A(B)$ is bounded in $\mathfrak{Y}$, since A is continuous. Taking account of the relation $A'(y'_\alpha) = y'_\alpha \circ A$, we obtain

$$\sup_{x\in B}|A'(y'_\alpha)(x)| = \sup_{x\in B}|(y'_\alpha \circ A)(x)| = \sup_{y\in A(B)}|y'_\alpha(y)| \to 0.$$

This relation holds good for any bounded set $B \subset \mathfrak{X}$. Hence we have immediately that $y'_\alpha \to 0$ (strong topology) $\Rightarrow A'(y'_\alpha) \to 0$ (strong topology). □

We denote by $\mathfrak{F} : \mathfrak{S}(\mathbb{R})' \to \mathfrak{S}(\mathbb{R})'$ the operator which associates a tempered distribution $T \in \mathfrak{S}(\mathbb{R})'$ with its Fourier transform. $\mathfrak{F}$ is also called the Fourier transform.[12]

A relation between $\mathcal{F} : \mathfrak{S}(\mathbb{R}) \to \mathfrak{S}(\mathbb{R})$ and $\mathfrak{F} : \mathfrak{S}(\mathbb{R})' \to \mathfrak{S}(\mathbb{R})'$ is given by

$$\mathfrak{F}T(\varphi) = \hat{T}(\varphi) = T(\hat{\varphi}) = T(\mathcal{F}\varphi)$$
$$\text{for all} \quad T \in \mathfrak{S}(\mathbb{R})', \; \varphi \in \mathfrak{S}(\mathbb{R}),$$

[11] cf. Appendix B, Sect. B.2 (p. 366).

[12] While the Fourier transform on $\mathfrak{S}(\mathbb{R})$ is denoted by $\mathcal{F}$, the Fourier transform on $\mathfrak{S}(\mathbb{R})'$ is denoted by the German Fraktur letter $\mathfrak{F}$. The different letters are used for the sake of clear distinction.

that is,

$$\mathfrak{F}T = T \circ \mathcal{F} \quad \text{for all} \quad T \in \mathfrak{S}(\mathbb{R})'.$$

In other words, $\mathfrak{F}$ is the dual operator of $\mathcal{F}$:

$$\mathfrak{F} = \mathcal{F}'. \tag{4.24}$$

We define an operator $\mathfrak{F}^{-1} : \mathfrak{S}(\mathbb{R})' \to \mathfrak{S}(\mathbb{R})'$ by

$$\mathfrak{F}^{-1}T(\varphi) = T(\mathcal{F}^{-1}\varphi) = T(\tilde{\varphi})$$

(where $T \in \mathfrak{S}(\mathbb{R})'$ and $\varphi \in \mathfrak{S}(\mathbb{R})$). $\mathfrak{F}^{-1}$ is called the **inverse Fourier transform**. As usual, we sometimes write $\tilde{T}$ instead of $\mathfrak{F}^{-1}T$. It is also obvious that $\mathfrak{F}^{-1}$ **is the dual operator of** $\mathcal{F}^{-1}$.

$$\mathfrak{F}^{-1} = (\mathcal{F}^{-1})'. \tag{4.25}$$

Theorem 4.5 $\mathfrak{F}$ *is an automorphism of* $\mathfrak{S}(\mathbb{R})'$. *The inverse Fourier transform* $\mathfrak{F}^{-1}$ *is its inverse.*

Proof The linearity and the injectivity are clear. The surjectivity can be shown as follows. If T is an element of $\mathfrak{S}(\mathbb{R})'$, then

$$\hat{\tilde{T}}(\varphi) = \tilde{T}(\hat{\varphi}) = T(\tilde{\hat{\varphi}}) = T(\varphi) \quad \text{for all} \quad \varphi \in \mathfrak{S}(\mathbb{R}).$$

This shows that T is the Fourier transform of $\tilde{T} \in \mathfrak{S}(\mathbb{R})'$. Similarly, the inverse of $\mathfrak{F}$ is given by the inverse Fourier transform, since

$$(\mathfrak{F}^{-1} \circ \mathfrak{F})(T)(\varphi) = T(\mathcal{F} \circ \mathcal{F}^{-1})(\varphi) = T(\varphi) \quad \text{for all} \quad T \in \mathfrak{S}(\mathbb{R})',\ \varphi \in \mathfrak{S}(\mathbb{R}).$$

Hence $\mathfrak{F}^{-1} \circ \mathfrak{F} = I$ (identity).

$\mathfrak{F}$ and $\mathfrak{F}^{-1}$ are continuous (in the strong topology) in view of Theorem 4.4. □

Remark 4.3 We use the notation $\check{\varphi}$ (or $\check{T}$) which means

$$\check{\varphi}(x) = \varphi(-x), \quad \check{T}(\varphi) = T(\check{\varphi}).$$

Then the following relations can be verified without any difficulty:

(i) $\hat{\hat{\varphi}} = \check{\varphi}, \quad \varphi \in \mathfrak{S}(\mathbb{R}).$
(ii) $\hat{\hat{T}} = \check{T}, \quad T \in \mathfrak{S}(\mathbb{R})'.$

We now explain several examples of Fourier transforms of tempered distributions.

Example 4.7 The distributions defined by $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ and $\hat{f}$ (usual Fourier transform) are denoted by T_f and $T_{\hat{f}}$, respectively. Then

$$\widehat{T_f} = T_{\hat{f}}.$$

In fact, it follows from

$$\begin{aligned}\widehat{T_f}(\varphi) = T_f(\hat{\varphi}) = \int_{\mathbb{R}} f(x)\hat{\varphi}(x)dx &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)\left\{\int_{\mathbb{R}} e^{-ix\xi}\varphi(\xi)d\xi\right\}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \varphi(\xi)\left\{\int_{\mathbb{R}} e^{-ix\xi} f(x)dx\right\}d\xi = T_{\hat{f}}(\varphi); \quad \varphi \in \mathfrak{S}(\mathbb{R}).\end{aligned}$$

This example illuminates that the concept of Fourier transform of tempered distributions is a generalization of the usual Fourier transform.

Example 4.8 $\hat{\delta} = \dfrac{1}{\sqrt{2\pi}}T_1, \quad \widehat{T_1} = \sqrt{2\pi}\delta.$ ($T_1$is the distribution defined by the function $= 1$.)

By direct computation, we have

$$\hat{\delta}(\varphi) = \delta(\hat{\varphi}) = \hat{\varphi}(0) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} 1 \cdot \varphi(\xi)d\xi = \frac{1}{\sqrt{2\pi}}T_1(\varphi) \quad \text{for all} \quad \varphi \in \mathfrak{S}(\mathbb{R}).$$

This shows the first relation. The second one immediately follows from

$$\delta = \check{\hat{\delta}} = \hat{\hat{\delta}} = \frac{1}{\sqrt{2\pi}}\widehat{T_1}.$$

Example 4.9 We next evaluate $\widehat{e^{int}}$. The relation

$$\widehat{e^{int}}\varphi = \frac{1}{\sqrt{2\pi}}\int\left[\int \varphi(x)e^{-itx}dx\right]e^{int}dt, \quad \varphi \in \mathfrak{S}(\mathbb{R})$$

gives the inverse Fourier transform of $\sqrt{2\pi}\hat{\varphi}$. Hence

$$\widehat{e^{int}}\varphi = \sqrt{2\pi}\delta_n\varphi \qquad (\delta_n \text{ is the Dirac function concentrating at } n),$$

which gives

$$\widehat{e^{int}} = \sqrt{2\pi}\delta_n.$$

Example 4.10 For $T \in \mathfrak{S}(\mathbb{R})'$, we have

$$\widehat{T'} = ix\hat{T}, \tag{4.26}$$

$$\widehat{ixT} = -\hat{T}', \tag{4.27}$$

where ixT (resp. $ix\hat{T}$) is a tempered distribution defined as the product of ix and T (resp. $\hat{T}$).

We recapitulate the formula (4.3):

$$D^p \hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} D^p(e^{-i\xi x} f(x))dx, \quad f \in \mathfrak{S}(\mathbb{R}). \tag{4.28}$$

The relation (4.26) follows from

$$\widehat{T'}(\varphi) = T'(\hat{\varphi}) = -T(\hat{\varphi}') = -T(-\widehat{ix\varphi}) = T(\widehat{ix\varphi}) = ix\hat{T}(\varphi).$$

Readers should recall the formula (4.4):

$$\widehat{D^q f}(\xi) = i^q \xi^q \hat{f}(\xi). \tag{4.29}$$

Making use of this formula, we obtain (4.27) by

$$(\widehat{ixT})(\varphi) = (ixT)(\hat{\varphi}) = T(ix\hat{\varphi}) \underset{(4.29)}{=} T(\widehat{\varphi'}) = \hat{T}(\varphi') = -\hat{T}'(\varphi).$$

We add here several remarks on simple convergence of Fourier transforms of distributions.

1° If a sequence $\{T_n\}$ in $\mathfrak{D}(\mathbb{R})'$ simply converges to some $T \in \mathfrak{D}(\mathbb{R})'$, then the sequence $\{T_n'\}$ of the derivatives (in the sense of distribution) also simply converges to T'. (More generally, the sequence $\{D^p T_n\}$ of the p-th derivatives simply converges to $D^p T$.)

Proof For any $\varphi \in \mathfrak{D}(\mathbb{R})$, we have

$$\lim_{n\to\infty} T_n'\varphi = -\lim_{n\to\infty} T_n\varphi' = -T\varphi' = T'\varphi.$$

□

2° A sequence $\{T_n\}$ in $\mathfrak{S}(\mathbb{R})'$ (the space of tempered distributions) simply converges to $T \in \mathfrak{S}(\mathbb{R})'$ if and only if the sequence $\{\hat{T}_n\}$ of the Fourier transforms simply converges to $\hat{T}$.[13]

Proof Suppose that $\{T_n\}$ simply converges to T. Then

$$\lim_{n\to\infty} \hat{T}_n\varphi = \lim_{n\to\infty} T_n\hat{\varphi} = T\hat{\varphi} = \hat{T}\varphi, \quad \varphi \in \mathfrak{S}(\mathbb{R}).$$

Hence $\{\hat{T}_n\}$ simply converges to $\hat{T}$.

[13] cf. Yosida and Kato [15] pp. 98–99.

Conversely, assume that $\{\hat{T}_n\}$ simply converges to $\hat{T}$. Let ψ be the inverse Fourier transform of $\varphi \in \mathfrak{S}(\mathbb{R})$ (i.e. $\hat{\psi} = \varphi$). Then

$$\lim_{n\to\infty} T_n\varphi = \lim_{n\to\infty} T_n\hat{\psi} = \lim_{n\to\infty} \hat{T}_n\psi = \hat{T}\psi = T\hat{\psi} = T\varphi,$$

which proves that $\{T_n\}$ simply converges to T. □

3° $\mathfrak{F} : \mathfrak{S}(\mathbb{R})' \to \mathfrak{S}(\mathbb{R})'$ is simply continuous.

4°

$$\sum_{k=-n}^{n} c_k\delta_k \text{ simply converges} \Longleftrightarrow \sum_{k=-n}^{n} c_k\widehat{e^{ikx}} \text{ simply converges}$$
$$\Longleftrightarrow \sum_{k=-n}^{n} c_k e^{ikx} \text{ simply converges.}$$

These relations can be confirmed by combining 2° and 3°.

4.5 Fourier Transforms on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ Revisited

We have already studied Plancherel's theory of Fourier transforms on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ in Sect. 4.2. The theory of distributions sheds new light on this topic.[14]

Lemma 4.1 $f \in \mathfrak{L}^p(\mathbb{R}, \mathbb{C})$ $(p \geqq 1)$ *defines a tempered distribution by the relation*

$$T_f(\varphi) = \int_{\mathbb{R}} f(x)\varphi(x)dx; \quad \varphi \in \mathfrak{S}(\mathbb{R}).$$

In order to see this, it is enough to show that the measure $d\mu = |f|dx$ is tempered (cf. Example 4.4). It is verified by applying Hölder's inequality to

$$\int_{\mathbb{R}} (1 + |x|^2)^{-k}|f(x)|dx.$$

We also have to recall the fact that $\mathfrak{S}(\mathbb{R})$ is dense in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ (cf. 5° on p. 68).

Theorem 4.3′ (Plancherel) *Let f be an element of $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. We denote by T_f the distribution defined by f.*

(i) *T_f is a tempered distribution and $\widehat{T_f}$ is defined by some $\hat{f} \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$; i.e.*

$$\widehat{T_f} = T_{\hat{f}} \quad \textit{for some} \quad \hat{f} \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C}).$$

Moreover, $\|\hat{f}\|_2 = \|f\|_2$. ($\|\cdot\|_2$ is the $\mathfrak{L}^2$-norm.)

[14] cf. Yosida [16] pp. 151–155.

(ii) $\hat{f}$ *obtained in* (i) *is expressed as*

$$\hat{f}(x) = \underset{h\uparrow\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{|y|\leqq h} e^{-ixy} f(y)dy.$$

(iii) *The mapping* $f \mapsto \hat{f}$ *is a bijection of* $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ *onto* $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$.
(iv) *For any* $f, g \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$,

$$\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle \quad (\textit{inner product in } \mathfrak{L}^2).$$

The mapping $\mathcal{F}_2 : \mathfrak{L}^2(\mathbb{R}, \mathbb{C}) \to \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ defined by $\mathcal{F}_2 : f \mapsto \hat{f}$ is called the **Fourier transform** on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. $\mathcal{F}_2(f) = \hat{f}$ is also called the Fourier transform of f. The concept of the Fourier transform on $\mathfrak{L}^2$ given in Sect. 4.2 (p. 72) is, of course, equivalent to that just defined above.

Proof of Theorem 4.3′ (i) We saw that $f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ defines a tempered distribution in Lemma 4.2. Consequently, the Fourier transform of T_f can be defined. It follows, then, that

$$\begin{aligned} |\widehat{T_f}(\varphi)| = |T_f(\hat{\varphi})| &= \left| \int_{\mathbb{R}} f(x)\hat{\varphi}(x)dx \right| \\ &\leqq \|f\|_2 \cdot \|\hat{\varphi}\|_2 = \|f\|_2 \cdot \|\varphi\|_2 \quad \text{for all} \quad \varphi \in \mathfrak{S}(\mathbb{R}) \end{aligned} \tag{4.30}$$

by Schwarz's inequality. ($\|\hat{\varphi}\|_2 = \|\varphi\|_2$ comes from (4.13).)

This shows that $\widehat{T_f}$ is an $\mathfrak{L}^2$-continuous linear functional on $\mathfrak{S}(\mathbb{R})$. Since $\mathfrak{S}(\mathbb{R})$ is dense in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$, $\widehat{T_f}$ can be uniquely extended to a continuous linear functional on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. The norm of $\widehat{T_f}$ viewed as an operator on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ satisfies $\|\widehat{T_f}\| \leqq \|f\|_2$ by (4.30). Thus, thanks to Riesz's representation theorem (Theorem 1.1, p. 3), there exists uniquely some $\hat{f} \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ such that

$$\begin{aligned} &\widehat{T_f}(\varphi) = \int_{\mathbb{R}} \hat{f}(x)\varphi(x)dx = T_{\hat{f}}(\varphi), \\ &\text{i.e.} \quad \int_{\mathbb{R}} f(x)\hat{\varphi}(x)dx = \int_{\mathbb{R}} \hat{f}(x)\varphi(x)dx \quad \text{for all} \quad \varphi \in \mathfrak{S}(\mathbb{R}). \end{aligned} \tag{4.31}$$

This $\hat{f}$ satisfies $\|\hat{f}\|_2 = \|\widehat{T_f}\| \leqq \|f\|_2$ as shown above. Hence

$$\|\hat{f}\|_2 \leqq \|f\|_2. \tag{4.32}$$

Similarly, we obtain

$$\|\hat{\hat{f}}\|_2 \leqq \|\hat{f}\|_2 \leqq \|f\|_2. \tag{4.33}$$

On the other hand, we have

$$\int_{\mathbb{R}} \hat{\hat{f}}(x)\varphi(x)dx = \int_{\mathbb{R}} \hat{f}(x)\hat{\varphi}(x)dx = \int_{\mathbb{R}} f(-x)\varphi(x)dx \quad \text{for all} \quad \varphi \in \mathfrak{S}(\mathbb{R})$$

by Remark 4.3 (p. 84) and (4.31). Hence

$$\hat{\hat{f}}(x) = f(-x) \quad \text{a.e.} \tag{4.34}$$

Thus it follows that

$$\|\hat{\hat{f}}\|_2 = \|f\|_2. \tag{4.35}$$

Combining (4.33) and (4.35), we obtain $\|f\|_2 = \|\hat{f}\|_2$.

(ii) Define a function f_h $(h > 0)$ by

$$f_h(x) = \begin{cases} f(x) & \text{for} \quad |x| \leqq h, \\ 0 & \text{for} \quad |x| > h. \end{cases}$$

Then we have (say, by the dominated convergence theorem)

$$\lim_{h\to\infty} \|f_h - f\|_2 = 0.$$

Hence, by (i),

$$\lim_{h\to\infty} \|\widehat{f_h} - \hat{f}\|_2 = 0, \quad \text{i.e.} \ \ \hat{f} = \underset{h\to\infty}{\text{l.i.m.}} \widehat{f_h}. \tag{4.36}$$

It follows from (4.31) that

$$\begin{aligned} \int_{\mathbb{R}} \widehat{f_h}(x)\varphi(x)dx &= \int_{\mathbb{R}} f_h(x)\hat{\varphi}(x)dx \\ &= \int_{|x|\leqq h} f(x)\left\{\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-ixy}\varphi(y)dy\right\}dx \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\left\{\int_{|x|\leqq h} e^{-ixy} f(x)dx\right\}\varphi(y)dy \quad \text{for all} \quad \varphi \in \mathfrak{S}(\mathbb{R}) \end{aligned}$$

(apply Fubini's theorem in view of the integrability of $f(x)$ on $|x| \leqq h$),

which implies

$$\widehat{f_h}(x) = \frac{1}{\sqrt{2\pi}} \int_{|x|\leqq h} e^{-ixy} f(y)dy. \tag{4.37}$$

(ii) is verified by (4.36) and (4.37).

$$\begin{tikzcd} \mathfrak{L}^2 & f \arrow[r, mapsto, "\mathcal{F}_2"] \arrow[d, leftrightarrow] & \hat{f} \arrow[d, leftrightarrow] \\ \mathfrak{S}' & T_f \arrow[r, mapsto, "\mathfrak{F}"'] & T_{\hat{f}} = \widehat{T_f} \end{tikzcd}$$

Fig. 4.1 Fourier transform as automorphism

(iii) Since the Fourier transform $\mathcal{F}_2 : \mathfrak{L}^2(\mathbb{R}, \mathbb{C}) \to \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ is an isometric linear operator, it is obviously injective. (See Fig. 4.1.)

Now, the inverse Fourier transform $\mathfrak{F}^{-1}T_f = \widetilde{T_f}$ of the distribution T_f defined by $f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ is a distribution defined by

$$\tilde{f}(x) = \underset{h\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{|y|\leqq h} e^{ixy} f(y)dy, \tag{4.38}$$

which is an element of $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. (Follow a similar argument in (i) and (ii))[15]

The mapping $\mathcal{F}_2^{-1} : f \mapsto \tilde{f}$ is called the **inverse Fourier transform** on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. ($\tilde{f}$ is also called the inverse Fourier transform of f as in the preceding discussions.) We can show $\|f\|_2 = \|\tilde{f}\|_2$ as in (i). (See Fig. 4.2.)

If we define $\mathfrak{S}_2'(\mathbb{R}, \mathbb{C})$ by

$$\mathfrak{S}_2'(\mathbb{R}, \mathbb{C}) = \{T_f \in \mathfrak{S}'(\mathbb{R}, \mathbb{C}) | f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})\},$$

$\mathfrak{F}$ is an automorphism on $\mathfrak{S}_2'$. Correspondingly, $\mathcal{F}_2$ is an automorphism on $\mathfrak{L}^2$.

(iv) This is easy. □

[15] Since $f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ determines a tempered distribution T_f, it has its inverse $\widetilde{T_f} \in \mathfrak{S}(\mathbb{R})'$. $\widetilde{T_f}$ is a continuous linear functional on $\mathfrak{S}(\mathbb{R})$ (with respect to $\mathfrak{L}^2$-norm). This can be checked by a similar computation to that in (4.30). Hence $\widetilde{T_f}$ can be uniquely extended to a continuous linear functional on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$. By Riesz's theorem, there exists some $\tilde{f} \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ which represents $\widetilde{T_f}$. $\tilde{f}$ is given in a concrete form

$$\tilde{f}(x) = \underset{h\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{|y|\leqq h} e^{ixy} f(y)dy$$

(cf. (4.36) and (4.37)). $\widetilde{T_f} = T_{\tilde{f}}$ is the inverse operator of $\widehat{T_f} = T_{\hat{f}}$ on $\mathfrak{S}(\mathbb{R})$. So it is clear that these are mutually inverse as operators extended to $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$.

Fig. 4.2 Inverse Fourier transform as automorphism

$$\begin{tikzcd}
\mathfrak{L}^2 & f \arrow[r, mapsto, "\mathcal{F}_2^{-1}"] \arrow[d, leftrightarrow] & \tilde{f} \arrow[d, leftrightarrow] \\
\mathfrak{S}' & T_f \arrow[r, mapsto, "\mathfrak{F}^{-1}"'] & T_{\tilde{f}} = \widetilde{T_f}
\end{tikzcd}$$

Theorem 4.6 (Parseval) *The following relations hold true for any f and $g \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$:*

(i) $\displaystyle\int_{\mathbb{R}} f(x)g(x)dx = \int_{\mathbb{R}} \hat{f}(\xi)\hat{g}(-\xi)d\xi.$

(ii) $\displaystyle\int_{\mathbb{R}} \hat{f}(\xi)\hat{g}(\xi)e^{i\xi x}d\xi = \int_{\mathbb{R}} f(y)g(x-y)dy.$

Proof (i) It can easily be verified that

$$\hat{\bar{g}} = \overline{\hat{g}(-\xi)} \tag{4.39}$$

by a simple calculation:

$$\begin{aligned}
\hat{\bar{g}}(\xi) &= \underset{h\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{|x|\leqq h} \bar{g}(x)e^{-i\xi x}dx \\
&= \underset{h\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \overline{\int_{|x|\leqq h} g(x)e^{i\xi x}dx} \\
&= \underset{h\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \overline{\int_{|x|\leqq h} g(x)e^{-i(-\xi x)}dx} \\
&= \overline{\hat{g}(-\xi).}
\end{aligned}$$

It follows from Theorem 4.3′(iv) and (4.39) that

$$\begin{aligned}
&\int_{\mathbb{R}} \hat{f}(\xi)\hat{g}(-\xi)d\xi = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{\bar{g}}}(\xi)d\xi \\
&= \langle \hat{f}, \hat{\bar{g}} \rangle = \langle f, \bar{g} \rangle \qquad \text{(Theorem 4.3′(iv))} \\
&= \int_{\mathbb{R}} f(x)g(x)dx.
\end{aligned}$$

(ii) We try to find the Fourier transform of $g(x-y)$ with respect to y. First we assume that g is an element of $\mathfrak{S}$.

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(x-y)e^{-i\xi y}dy &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(t)e^{-i\xi(x-t)}dt \\ &\qquad \text{(changing variables : } t = x-y) \\ &= e^{-i\xi x}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(t)e^{-i\xi(-t)}dt \\ &= e^{-i\xi x}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(t)e^{-i(-\xi)t}dt \\ &= e^{-i\xi x}\hat{g}(-\xi).\end{aligned}$$

In the general case of $g \in \mathfrak{L}^2(\mathbb{R},\mathbb{C})$, the same relation holds good, since $\mathfrak{S}$ is dense in $\mathfrak{L}^2$ and the Fourier transform is continuous. By (i), we obtain

$$\int_{\mathbb{R}} f(y)g(x-y)dy = \int_{\mathbb{R}} \hat{f}(\xi)\hat{g}(\xi)e^{i\xi x}d\xi.$$

□

4.6 Periodic Distributions

A locally integrable function $f \in \mathfrak{L}^1_{loc.}(\mathbb{R},\mathbb{C})$ defines a distribution T_f by the relation[16,17]

$$T_f(\varphi) = \int_{-\infty}^{\infty} f(x)\varphi(x)dx, \quad \varphi \in \mathfrak{D}(\mathbb{R}).$$

Translating $f(x)$ by $\tau \in \mathbb{R}$, the function $f(x-\tau)$ defines a distribution $T_{f(x-\tau)}$ by

$$\begin{aligned}T_{f(x-\tau)}(\varphi) &= \int_{-\infty}^{\infty} f(x-\tau)\varphi(x)dx \\ &= \int_{-\infty}^{\infty} f(x)\varphi(x+\tau)dx, \quad \varphi \in \mathfrak{D}(\mathbb{R}).\end{aligned}$$

[16] This section is based upon Maruyama [10].

[17] We denote by $\mathfrak{L}^1_{loc.}(\mathbb{R},\mathbb{C})$ the space of locally integrable complex-valued functions defined on $\mathbb{R}$. $\mathfrak{D}(\mathbb{R})$ is the space of test functions. See Sect. 4.4 and Appendix C.

Hence, if f is a periodic function with period τ (or simply, τ-periodic), we must have

$$\int_{-\infty}^{\infty} f(x)\varphi(x)dx = \int_{-\infty}^{\infty} f(x)\varphi(x+\tau)dx$$

for all $\varphi \in \mathfrak{D}(\mathbb{R})$. Generalizing this reasoning, the concept of periodic distribution is defined as follows.[18]

Definition 4.6 A distribution $T \in \mathfrak{D}(\mathbb{R})'$ is called a **periodic distribution** with period τ if

$$T(\varphi(x)) = T(\varphi(x+\tau))$$

for all $\varphi \in \mathfrak{D}(\mathbb{R})$.

We denote by $\mathfrak{D}_\tau(\mathbb{R})'$ the set of τ-periodic distributions. For the sake of simplicity, we assume $\tau = 2\pi$ from now on. $\mathbb{T}$ denotes the torus (cf. Appendix A).

To start with, let us confirm that $\mathfrak{D}_{2\pi}(\mathbb{R})'$ can be identified with $\mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})'$.

Assume that $S \in \mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})'$ is given. We associate with each $\varphi \in \mathfrak{D}(\mathbb{R})$ a new function

$$\tilde{\varphi}(x) = \sum_{n=-\infty}^{\infty} \varphi(x + 2n\pi). \tag{4.40}$$

The value $\tilde{\varphi}(x)$ is defined without any ambiguity because the support of φ is compact and so the right-hand side of (4.40) is actually a finite sum for each x. Since $\tilde{\varphi}$ is 2π-periodic and infinitely differentiable, it can be regarded as an element of $\mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})$. If we define an operator T on $\mathfrak{D}(\mathbb{R})$ by

$$T(\varphi) = S(\tilde{\varphi}), \quad \varphi \in \mathfrak{D}(\mathbb{R}), \tag{4.41}$$

T is a continuous linear functional and so $T \in \mathfrak{D}(\mathbb{R})'$. Since it is obvious that $T(\varphi(x)) = T(\varphi(x+2\pi))$, T is a 2π- periodic distribution; i.e. $T \in \mathfrak{D}_{2\pi}(\mathbb{R})'$.

Conversely, let T be an element of $\mathfrak{D}_{2\pi}(\mathbb{R})'$. It can be shown that $T \in \mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})'$. We prepare a lemma due to Yosida–Kato [15], §7.

Lemma 4.2 *For any $a > 0$, there exists some $\theta \in \mathfrak{D}(\mathbb{R})$ which satisfies*[19]

(i) $\operatorname{supp} \theta = [-a, a]$, *and*

(ii) $\displaystyle\sum_{n=-\infty}^{\infty} \theta(x + na) = 1.$

[18] $\mathfrak{D}(\mathbb{R})'$ denotes the dual space of $\mathfrak{D}(\mathbb{R})$. Each element of $\mathfrak{D}(\mathbb{R})'$ is called a distribution.

[19] The notation $\operatorname{supp} \theta$ means the support of the function θ.

Proof Define a function $\theta : \mathbb{R} \to \mathbb{C}$ by

$$\theta(x) = \begin{cases} \displaystyle\int_{|x|}^{a} \exp\left(-\frac{1}{w(a-w)}\right) dw \Big/ \int_{0}^{a} \exp\left(-\frac{1}{w(a-w)}\right) dw & \text{for} \quad |x| \leqq a, \\ 0 & \text{for} \quad |x| > a. \end{cases}$$

Then it is clear that $\theta \in \mathfrak{D}(\mathbb{R})$ and (i) is satisfied. Computing $\theta(x-a)$ for $|x| \leqq a$, we obtain that

$$\theta(x-a) = \int_{|x-a|}^{a} \exp\left(-\frac{1}{w(a-w)}\right) dw \Big/ \int_{0}^{a} \exp\left(-\frac{1}{w(a-w)}\right) dw.$$

Changing the variable by $z = a - w$, we can rewrite it as

$$\begin{aligned} \theta(x-a) &= -\int_{x}^{0} \exp\left(-\frac{1}{z(a-z)}\right) dz \Big/ \int_{0}^{a} \exp\left(-\frac{1}{w(a-w)}\right) dw \\ &= \int_{0}^{x} \exp\left(-\frac{1}{z(a-z)}\right) dz \Big/ \int_{0}^{a} \exp\left(-\frac{1}{w(a-w)}\right) dw \end{aligned}$$

if $x \in [0, a]$. Consequently, it follows that

$$\theta(x) + \theta(x-a) = 1 \quad \text{for} \quad x \in [0, a].$$

We also obtain the same relation for $x \in [-a, 0)$ by a similar argument.

There exists, for each $x \in \mathbb{R}$, a unique integer $k \in \mathbb{Z}$ such that

$$ka \leqq |x| < (k+1)a.$$

Therefore we must have

$$\sum_{n=-\infty}^{\infty} \theta(x+na) = \theta(x-ka) + \theta(x-(k+1)a) = 1.$$

This proves (ii). □

Let us go back to prove that $T \in \mathfrak{D}_{2\pi}(\mathbb{R})'$ can be regarded as an element of $\mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})'$.

Any $\psi \in \mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})$ can be regarded as a 2π- periodic smooth function defined on $\mathbb{R}$. If $\theta \in \mathfrak{D}(\mathbb{R})$ is a function which satisfies (i) and (ii) of Lemma 4.2 for $a = 2\pi$, then $\psi\theta \in \mathfrak{D}(\mathbb{R})$. Define an operator U on $\mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})$ by

$$U(\psi) = T(\psi\theta). \tag{4.42}$$

It is well-defined in the sense that U does not depend upon the choice of θ. In fact, if $\eta \in \mathfrak{D}(\mathbb{R})$ satisfies

$$\sum_{n=-\infty}^{\infty} \eta(x + 2n\pi) = 0,$$

we have[20]

$$\begin{aligned}
T(\psi\eta) &= T\left(\psi(x)\eta(x) \sum_{n=-\infty}^{\infty} \theta(x + 2n\pi)\right) \\
&= \sum_{n=-\infty}^{\infty} T(\psi(x - 2n\pi)\eta(x - 2n\pi)\theta(x)) \quad \text{(by the periodicity of } T\text{)} \\
&= \sum_{n=-\infty}^{\infty} T(\eta(x - 2n\pi) \cdot \psi(x)\theta(x)) \quad \text{(by the periodicity of } \psi\text{)} \\
&= T\left(\left(\sum_{n=-\infty}^{\infty} \eta(x - 2n\pi)\right)\psi(x)\theta(x)\right) \\
&= 0.
\end{aligned}$$

This confirms that U is well-defined.

U is continuous on $\mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})$. In fact, for any net $\{\psi_\alpha\}$ in $\mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})$ which converges to some $\psi \in \mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})$, we have[21]

$$\psi_\alpha\theta \to \psi\theta \quad \text{in} \quad \mathfrak{D}(\mathbb{R}).$$

Consequently, it follows that

$$U(\psi_\alpha) = T(\psi_\alpha\theta) \to T(\psi\theta) = U(\psi).$$

Thus we establish the chain illustrated by Fig. 4.3.

$$\begin{array}{ccccc}
S & \xmapsto{(4.41)} & T & \xmapsto{(4.42)} & U. \\
\in \mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})' & & \in \mathfrak{D}_{2\pi}(\mathbb{R})' & & \in \mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})'
\end{array}$$

Fig. 4.3 $\mathfrak{C}^\infty(\mathbb{T}, \mathbb{C})'$ and $\mathfrak{D}_{2\pi}(\mathbb{R})'$

---

[20] $\displaystyle\sum_{k=-p}^{p} \theta(x + 2k\pi) \to \sum_{n=-\infty}^{\infty} \theta(x + 2n\pi)$ (in $\mathfrak{C}^\infty$) on supp $\psi\eta$.

[21] We should note that supp $\psi_\alpha\theta \subset$ supp θ.

It is natural to ask if S is identical with U. The answer is positive as confirmed by the following calculation:

$$
\begin{aligned}
U(\psi) &= T(\psi\theta) \\
&= S\left(\sum_{n=-\infty}^{\infty} \psi(x+2n\pi)\theta(x+2n\pi)\right) \\
&= S\left(\sum_{n=-\infty}^{\infty} \psi(x)\theta(x+2n\pi)\right) \quad \text{(by the periodicity of } \psi) \\
&= S\left(\psi(x)\sum_{n=-\infty}^{\infty} \theta(x+2n\pi)\right) \\
&= S(\psi) \quad \text{for any} \quad \psi \in \mathfrak{C}^{\infty}(\mathbb{T}, \mathbb{C}).
\end{aligned}
$$

The reasoning discussed above justifies the following result.

Theorem 4.7 (periodic distribution) *There is a one-to-one correspondence between* $\mathfrak{C}^{\infty}(\mathbb{T}, \mathbb{C})'$ *and* $\mathfrak{D}_{2\pi}(\mathbb{R})'$.

The operators defined by (4.41) and (4.42) are inverse to each other.

Let T be a 2π-periodic distribution; i.e. $T \in \mathfrak{D}_{2\pi}(\mathbb{R})'$. The Fourier coefficients of T are defined by

$$
c_n = \frac{1}{\sqrt{2\pi}} T(e^{-inx}), \quad n \in \mathbb{Z}.
$$

The formal series

$$
\frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} c_n e^{inx}
$$

is called the Fourier series of T.

Remark 4.4 Assume that a trigonometric series

$$
\frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} C_n e^{inx} \tag{4.43}
$$

simply converges to a distribution T; i.e.

$$\frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi}\sum_{k=-p}^{p} C_k e^{ikx}\cdot\varphi(x)dx \to T(\varphi) \quad \text{for any} \quad \varphi\in\mathfrak{D}_{2\pi}(\mathbb{R}). \tag{4.44}$$

If we consider a special case of $\varphi = e^{-inx}$, the left-hand side of (4.44) converges to $\sqrt{2\pi}C_n$. On the other hand, $T(e^{-inx}) = \sqrt{2\pi}c_n$. Hence we must have $C_n = c_n$, and so the series (4.43) is nothing other than the Fourier series of T.

Theorem 4.8 (Fourier coefficients of periodic distribution) *Consider a sequence $\{c_n\}_{n\in\mathbb{Z}}$ of complex numbers. The c_n's are the Fourier coefficients of some 2π-periodic distribution if and only if there exists some $N\in\mathbb{N}\cup\{0\}$ such that*

$$c_n = O(|n|^N)\,; \tag{4.45}$$
$$\textit{i.e.}\quad |c_n| \leqq K|n|^N \quad \textit{for some} \quad K>0.$$

Proof[22] Assume first that $\{c_n\}$ satisfies (4.45). We write formally

$$u(x) = \sum_{n\neq 0}\frac{1}{(in)^{N+2}}c_n e^{inx}, \quad x\in\mathbb{R}. \tag{4.46}$$

Since

$$\sum_{n\neq 0}\frac{1}{|n|^{N+2}}|c_n||e^{inx}| \leqq \sum_{n\neq 0}\frac{K}{n^2}$$

by the assumption (4.45), the right-hand side of (4.46) is absolutely, uniformly convergent. Hence $u(x)$ is a continuous function which satisfies

$$|u(x)| \leqq K\sum_{n\neq 0}\frac{1}{n^2}. \tag{4.47}$$

Consequently, $u(x)$ defines a distribution, and

$$u_n(x) = \sum_{\substack{k=-n\\k\neq 0}}^{n}\frac{1}{(ik)^{N+2}}c_k e^{ikx}$$

[22] See Folland [3], pp. 320–322 and Lax [7], p. 570 for an outline of ideas.

simply converges to $u(x)$ $\big($exactly speaking, to the distribution defined by $u(x)\big)$. In fact, it can be verified by

$$\left|\int_{-\infty}^{\infty}(u_n(x)-u(x))\varphi(x)dx\right| \leqq \int_{-\infty}^{\infty}|u_n(x)-u(x)|\cdot|\varphi(x)|dx \to 0$$

$$\text{as} \quad n\to\infty \quad \text{for any} \quad \varphi\in\mathfrak{D}(\mathbb{R}).$$

It follows that, in view of 1° on p. 86,

$$D^{N+2}u(x)(\varphi) = \lim_{n\to\infty}\sum_{\substack{k=-n\\k\neq 0}}^{n}\frac{1}{(ik)^{N+2}}c_k(ik)^{N+2}e^{ikx}(\varphi) = \lim_{n\to\infty}\sum_{\substack{k=-n\\k\neq 0}}^{n}c_k e^{ikx}(\varphi).$$

Hence we must have

$$\frac{1}{\sqrt{2\pi}}\sum_{k=-n}^{n}c_k e^{ikx} \to \frac{1}{\sqrt{2\pi}}(c_0+D^{N+2}u(x)) \quad \text{simply as} \quad n\to\infty.$$

By Remark 4.4, c_n's are the Fourier coefficient of the 2π- periodic distribution defined by $(1/\sqrt{2\pi})(c_0+D^{N+2}u(x))$.

Let us go over to the converse. Assume that

$$\frac{1}{\sqrt{2\pi}}\sum_{k=-n}^{n}c_k e^{ikx}$$

simply converges to some distribution. If c_n's do not satisfy (4.45), there exists a sequence $\{n_r\}$ of integers such that

$$|c_{n_r}| > |n_r|^r; \quad r=1,2,\cdots. \tag{4.48}$$

Define a pair of functions $\lambda(x)$ and $\varphi(x)$ by

$$\lambda(x) = \begin{cases} e^{-x^2/(1-x^2)} & \text{if} \quad |x|<1, \\ 0 & \text{if} \quad |x|\geqq 1 \end{cases} \tag{4.49}$$

and

$$\varphi(x) = \sum_{r=1}^{\infty}c_{n_r}^{-1}\lambda(x-n_r). \tag{4.50}$$

The right-hand side of (4.50) is a finite sum for each x and $\varphi(x) \in \mathfrak{D}(\mathbb{R})$. By the definition of φ, we have[23]

$$\varphi(n) = \begin{cases} 0 & \text{if} \quad n \neq n_r, \\ c_{n_r}^{-1} & \text{if} \quad n = n_r \end{cases} \qquad (r = 1, 2, \cdots).$$

It follows that

$$\begin{aligned} \sum_{k=-n}^{n} c_k \delta_k(\varphi) &= \int_{-\infty}^{\infty} \left\{ \sum_{k=-n}^{n} c_k \delta_k \right\} \varphi(t) dt \\ &= \sum_{k=-n}^{n} c_k \varphi(k) \qquad (4.51) \\ &= \text{the number of } n_r{}'s \text{ between } -n \text{ and } n. \end{aligned}$$

The right-hand side of (4.51) diverges to ∞ as $n \to \infty$. Consequently, by 4° on p. 87, $\sum_{k=-n}^{n} c_k e^{ikt}$ can not be simply convergent. Contradiction. □

References

1. Bourbaki, N.: Eléments de mathématique; Espaces vectoriels topologiques. Hermann, Paris, Chaps. 1–2 (seconde édition) (1965), Chaps. 3–5 (1964) (English edn.) Elements of Mathematics; Topological Vector Spaces. Springers, Berlin/Heidelberg/New York (1987)
2. Dunford, N., Schwartz, J.T.: Linear Operators, Part 1–3. Interscience, New York (1958–1971)
3. Folland, G.B.: Fourier Analysis and Its Applications. American Mathematical Society, Providence (1992)
4. Grothendieck, A.: Topological Vector Spaces. Gordon and Breach, New York (1973)
5. Kawata, T.: Ohyo Sugaku Gairon (Elements of Applied Mathematics). I, II, Iwanami Shoten, Tokyo (1950, 1952) (Originally published in Japanese)
6. Kawata, T.: Fourier Kaiseki (Fourier Analysis). Sangyo Tosho, Tokyo (1975) (Originally published in Japanese)
7. Lax, P.D.: Functional Analysis. Wiley, New York (2002)
8. Maruyama, T.: Kansu Kaisekigaku (Functional Analysis). Keio Tsushin, Tokyo (1980) (Originally published in Japanese)
9. Maruyama, T.: Sekibun to Kansu-kaiseki (Integration and Functional Analysis). Springer, Tokyo (2006) (Originally published in Japanese)
10. Maruyama, T.: Herglotz-Bochner representation theorem via theory of distributions. J. Operations Res. Soc. Japan **60**, 122–135 (2017)
11. Schwartz, L.: Functional Analysis. Courant Institute of Mathematical Sciences, New York University, New York (1964)

[23] $n \neq n_r \Rightarrow \lambda(n - n_r) = 0$, $n = n_r \Rightarrow \lambda(n_r - n_r) = 1$.

12. Schwartz, L.: Théorìe des distributions. Nouvelle édition, entièrement corrigée, refondue et augmentée, Hermann, Paris (1973)
13. Takagi, T.: Kaiseki Gairon (Treatise on Analysis), 3rd edn. Iwanami Shoten, Tokyo (1961) (Originally published in Japanese)
14. Treves, J.F.: Topological Vector Spaces, Distributions and Kernels. Academic Press, New York (1967)
15. Yosida, K., Kato T.: Ohyo Sugaku (Applied Mathematics). I, Shokabou, Tokyo (1961) (Originally published in Japanese)
16. Yosida, K.: Functional Analysis, 3rd edn. Springer, New York (1971)

Chapter 5
Summability Kernels and Spectral Synthesis

We have discussed the $(C, 1)$-summability of Fourier series in Chap. 2. We recapitulate this topic primarily in the framework of Fourier series in complex form.

The details of the Fejér summability on $[-\pi, \pi]$ and $\mathbb{R}$ are examined. The fundamental results on spectral synthesis based upon the summability technique are the main concerns of this chapter. Similar procedures will be effectively utilized in the proof of the Herglotz–Bochner theorem in Chap. 6.

5.1 Shift Operators

Fixing $x_0 \in \mathbb{R}$, we define the shift operator $\tau_{x_0} : \mathfrak{L}^p(\mathbb{R}, \mathbb{C}) \to \mathfrak{L}^p(\mathbb{R}, \mathbb{C})$[1] by

$$(\tau_{x_0} f)(x) = f(x - x_0), \quad f \in \mathfrak{L}^p, \quad x \in \mathbb{R}. \tag{5.1}$$

Since $\|\tau_{x_0} f\|_p = \|f\|_p$, the operator τ_{x_0} is an isometric automorphism of $\mathfrak{L}^p$ onto itself. Furthermore, $x \mapsto \tau_x$ is a group-homomorphism of $\mathbb{R}$ into Aut $(\mathfrak{L}^p)$[2] because

$$\tau_x \circ \tau_y = \tau_{x+y}; \quad x, y \in \mathbb{R}. \tag{5.2}$$

Theorem 5.1 (continuity theorem)

(i) *The operator* $F : \mathbb{R} \to \mathfrak{C}_0(\mathbb{R}, \mathbb{C})$ *defined by*

$$F : x \mapsto \tau_x f \tag{5.3}$$

for any fixed $f \in \mathfrak{C}_0(\mathbb{R}, \mathbb{C})$ *is uniformly continuous.*

[1] More generally, we can consider a locally compact Hausdorff commutative topological group instead of $\mathbb{R}$, and Haar measure instead of Lebesgue measure.

[2] $\mathrm{Aut}(\mathfrak{L}^p)$ is the group of automorphisms on $\mathfrak{L}^p$.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_5

(ii) *If* $f \in \mathfrak{L}^p(\mathbb{R}, \mathbb{C})$ $(1 \leqq p < \infty)$, *then the operator* $F : \mathbb{R} \to \mathfrak{L}^p(\mathbb{R}, \mathbb{C})$ *defined by (5.3) is uniformly continuous.*

Proof (i) Since $f \in \mathfrak{C}_0$ is uniformly continuous, there exists some $\delta > 0$, for each $\varepsilon > 0$, such that

$$|f(x) - f(y)| < \varepsilon \quad \text{if} \quad |x - y| < \delta.$$

If $|x_0 - y_0| < \delta$,

$$|(x - y_0) - (x - x_0)| < \delta \quad \text{for any} \ \ x.$$

Hence

$$|\tau_{x_0} f(x) - \tau_{y_0} f(x)| = |f(x - x_0) - f(x - y_0)| < \varepsilon$$

for any $x \in \mathbb{R}$. That is,

$$\|\tau_{x_0} f - \tau_{y_0} f\|_\infty \leqq \varepsilon \quad \text{if} \quad |x_0 - y_0| < \delta.$$

(ii) Assume $f \in \mathfrak{L}^p$ $(1 \leqq p < \infty)$. Then for any $\varepsilon > 0$, there exists some $\varphi \in \mathfrak{C}_0$ such that $\|f - \varphi\|_p < \varepsilon/3$. It is trivial that

$$\tau_x f - \tau_y f = \tau_x(f - \varphi) + \tau_x\varphi - \tau_y(f - \varphi) - \tau_y\varphi. \tag{5.4}$$

Taking account of

$$\|\tau_x(f - \varphi)\|_p = \|\tau_y(f - \varphi)\|_p = \|f - \varphi\|_p < \frac{\varepsilon}{3},$$

we obtain

$$\|\tau_x f - \tau_y f\|_p < \frac{2}{3}\varepsilon + \|\tau_x\varphi - \tau_y\varphi\|_p \tag{5.5}$$

by (5.4). If we denote supp φ by K, it is easy to check that

$$\text{supp}(\tau_x\varphi - \tau_y\varphi) \subset (x + K) \cup (y + K),$$

which implies

$$m(\text{supp}(\tau_x\varphi - \tau_y\varphi)) < 2mK$$

(m is the Lebesgue measure). Consequently, it follows that

$$\|\tau_x\varphi - \tau_y\varphi\|_p \leqq \|\tau_x\varphi - \tau_y\varphi\|_\infty \cdot (2mK)^{1/p}.$$

If $\delta > 0$ is small enough,

$$\|\tau_x\varphi - \tau_y\varphi\|_\infty < \frac{\varepsilon}{3(2mK)^{1/p}}, \quad \text{if} \quad |x - y| < \delta \tag{5.6}$$

by (i). Combining (5.4), (5.5) and (5.6), we have

$$\|\tau_x\varphi - \tau_y\varphi\|_p < \varepsilon \quad \text{if} \quad |x - y| < \delta.$$

□

Theorem 5.2 (shift operator and convolution) *If $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ and $k : \mathbb{R} \to \mathbb{C}$ is continuous and integrable, then*[3]

$$\int_{-\infty}^{\infty} k(x)\tau_x f dx = k * f.$$

Proof Assume first that f is continuous and supp$f$ is compact. In this case, the integration on the left-hand side is actually evaluated on some finite interval. Hence

$$\int_{-\infty}^{\infty} k(x)\tau_x f dx = \lim \sum_j (x_{j+1} - x_j)k(x_j)\tau_{x_j} f, \tag{5.7}$$

where the limit is taken with respect to $\mathfrak{L}^1$-norm as the decomposition of the interval of integration becomes finer and finer. On the other hand, we have

$$(k * f)(x) = \lim \sum_j (x_{j+1} - x_j)k(x_j)f(x - x_j) \quad \text{(uniform convergence).} \tag{5.8}$$

Comparing (5.7) and (5.8), the proof is finished in this special case.

We shall now turn to the general case: $f \in \mathfrak{L}^1$. There exists, for any $\varepsilon > 0$, some continuous function g with compact support which satisfies $\|f - g\|_1 < \varepsilon$. Since

$$\int_{-\infty}^{\infty} k(x)\tau_x g dx = k * g$$

[3]This result may seem to be trivial. However, we have to note carefully what it means. The right-hand side is a function defined by

$$(k * f)(y) = \int_{-\infty}^{\infty} k(x)f(y - x)dx.$$

The left-hand side is the Cauchy–Bochner integral of the $\mathfrak{L}^1$-valued function $x \mapsto k(x)\tau_x f$. The theorem asserts that these two are equal. cf. Amann and Escher [1], II pp. 17–23, Maruyama [6] pp. 329–333 for the Cauchy–Bochner integral.

as observed above, it follows that

$$\int_{-\infty}^{\infty} k(x)\tau_x f dx - k * f = \int_{-\infty}^{\infty} k(x)\tau_x (f-g)dx + k*(g-f).$$

Consequently, we obtain

$$\left\| \int_{-\infty}^{\infty} k(x)\tau_x f dx - k * f \right\|_1 \leqq 2\|k\|_1 \cdot \varepsilon.$$

□

We denote by $\mathfrak{L}^p_{2\pi}(\mathbb{R}, \mathbb{C})$ the set of all the measurable complex-valued functions on $\mathbb{R}$ which are 2π-periodic and p-th integrable on $[-\pi, \pi]$. Fixing $x_0 \in \mathbb{R}$, we define an operator $\tau_{x_0} : \mathfrak{L}^p_{2\pi}(\mathbb{R}, \mathbb{C}) \to \mathfrak{L}^p_{2\pi}(\mathbb{R}, \mathbb{C})$ $(p \geqq 1)$ by

$$(\tau_{x_0} f)(x) = f(x - x_0), \;\; f \in \mathfrak{L}^p_{2\pi}, \;\; x \in \mathbb{R}.$$

Clearly, $\tau_{x_0} f \in \mathfrak{L}^p_{2\pi}$. Regarding it as an $\mathfrak{L}^p$-function on $[-\pi, \pi]$, we note

$$\|f\|_{\mathfrak{L}^p([-\pi,\pi],\mathbb{C})} = \|\tau_{x_0} f\|_{\mathfrak{L}^p([-\pi,\pi],\mathbb{C})}.$$

The following two theorems can be shown in a similar manner to Theorems 5.1 and 5.2.

Theorem 5.1′ (continuity theorem)

(i) *The operator* $F : \mathbb{R} \to \mathfrak{C}_{2\pi}(\mathbb{R}, \mathbb{C})$ *defined by*

$$F : x \mapsto \tau_x f \tag{5.9}$$

for any fixed $f \in \mathfrak{C}_{2\pi}(\mathbb{R}, \mathbb{C})$ *is uniformly continuous.* $\mathfrak{C}_{2\pi}(\mathbb{R}, \mathbb{C})$ *denotes the set of all the complex-valued* 2π*-periodic continuous functions on* $\mathbb{R}$*.* $\mathfrak{C}_{2\pi}(\mathbb{R}, \mathbb{C})$ *is endowed with the uniform convergence topology.*

(ii) *If* $f \in \mathfrak{L}^p_{2\pi}(\mathbb{R}, \mathbb{C})$ $(1 \leqq p < \infty)$*, then the operator* $F : \mathbb{R} \to \mathfrak{L}^p_{2\pi}$ *defined by (5.9) is uniformly continuous.*

Theorem 5.2′ *If* $f \in \mathfrak{L}^1_{2\pi}(\mathbb{R}, \mathbb{C})$ *and* $k \in \mathfrak{C}_{2\pi}(\mathbb{R}, \mathbb{C})$*, then*[4]

$$\int_{-\pi}^{\pi} k(x)\tau_x f dx = k * f.$$

[4] See footnote 3 for the interpretation of the result.

5.2 Summability Kernels on $[-\pi, \pi]$

We now go over to the Fejér method, which is quite effective for examining the summability of Fourier series in complex form.[5]

We first give a general definition.

Definition 5.1 A sequence $\{k_n : [-\pi, \pi] \to \mathbb{R}\}$ of continuous functions is called a **summability kernel** if it satisfies the following conditions:

(i) $\displaystyle\int_{-\pi}^{\pi} k_n(x)dx = 1 \quad \text{for all} \quad n = 1, 2, \cdots$

(ii) $\|k_n\|_1 \leqq \text{constant} \quad \text{for all} \quad n = 1, 2, \cdots$

(iii) $\displaystyle\lim_{n\to\infty} \int_{\delta \leqq |x| \leqq \pi} |k_n(x)|dx = 0 \quad \text{for any} \quad \delta \in (0, \pi).$

Lemma 5.1 *Let $\mathfrak{X}$ be a Banach space and a function $\varphi : [-\pi, \pi] \to \mathfrak{X}$ be continuous. Then*

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} k_n(x)\varphi(x)dx = \varphi(0)$$

for any summability kernel $\{k_n\}$ on $[-\pi, \pi]$.

Proof In view of the condition (i) in the definition of summability kernels, it holds good, for any $\delta \in (0, \pi)$, that

$$\begin{aligned} \int_{-\pi}^{\pi} k_n(x)\varphi(x)dx - \varphi(0) &= \int_{-\pi}^{\pi} k_n(x)(\varphi(x) - \varphi(0))dx \\ &= \int_{-\delta}^{\delta} + \int_{\delta \leqq |x| \leqq \pi} = I_1 + I_2. \end{aligned} \tag{5.10}$$

I_1 is evaluated as

$$\|I_1\| \leqq \operatorname*{Max}_{|x| \leqq \delta} \|\varphi(x) - \varphi(0)\| \cdot \|k_n\|_1. \tag{5.11}$$

Choosing $\delta > 0$ sufficiently small for any $\varepsilon > 0$, the right-hand side of (5.11) is less than ε by the continuity of φ and the condition (ii) in the definition. As for I_2, we have

$$\|I_2\| \leqq \operatorname*{Max}_{\delta \leqq |x| \leqq \pi} \|\varphi(x) - \varphi(0)\| \int_{\delta \leqq |x| \leqq \pi} |k_n(x)|dx. \tag{5.12}$$

[5] Katznelson [3] Chap. I, §2 is quite helpful.

By the condition (iii), the right-hand side of (5.12) is less than ε for large n. Equations (5.10), (5.11) and (5.12) imply

$$\int_{-\pi}^{\pi} k_n(x)\varphi(x)dx \to \varphi(0) \quad \text{as} \quad n \to \infty.$$

□

In the special case of $\mathfrak{X} = \mathbb{R}$, the assertion of Lemma 5.1 can be expressed as "the sequence $\{k_n dx\}$ of measures on $[-\pi, \pi]$ w^*-converges to the Dirac measure δ_0".

The operator $F : \mathbb{R} \to \mathfrak{L}^1_{2\pi}$ defined by $F : x \mapsto \tau_x f$ for a fixed $f \in \mathfrak{L}^1_{2\pi}(\mathbb{R}, \mathbb{C})$ is, by Theorem 5.1′, uniformly continuous. Of course, $F(0) = f$. Hence the following theorem follows from Lemma 5.1.

Theorem 5.3 (convergence of $k_n * f_\tau$) *For any $f \in \mathfrak{L}^1_{2\pi}(\mathbb{R}, \mathbb{C})$ and a summability kernel $\{k_n\}$ on $[-\pi, \pi]$,*

$$\int_{-\pi}^{\pi} k_n(x)\tau_x f dx \to f \quad \textit{as} \quad n \to \infty$$

in $\|\cdot\|_1$.

The Fejér kernel

$$K_n(x) = \frac{1}{2n\pi}\left(\frac{\sin\dfrac{nx}{2}}{\sin\dfrac{x}{2}}\right)^2 \tag{5.13}$$

is a special example of summability kernel (cf. Chap. 2, Sect. 2.5, p. 42). K_n can be represented in the form

$$K_n(x) = \frac{1}{2\pi}\sum_{j=-(n-1)}^{n-1}\left(1 - \frac{|j|}{n}\right)e^{ijx}. \tag{5.14}$$

We prove the relation (5.14) by a direct calculation.

First of all, taking account of the elementary fact

$$\sin^2\frac{x}{2} = \frac{1}{2}(1 - \cos x) = -\frac{1}{4}e^{-ix} + \frac{1}{2} - \frac{1}{4}e^{ix},$$

we obtain

$$\begin{aligned}
\sin^2\frac{x}{2}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)e^{ijx} &= \left(-\frac{1}{4}e^{-ix}+\frac{1}{2}-\frac{1}{4}e^{ix}\right)\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)e^{ijx} \\
&\overset{(*)}{=} \frac{1}{n}\left(-\frac{1}{4}e^{-inx}+\frac{1}{2}-\frac{1}{4}e^{inx}\right) \qquad (5.15)\\
&= \frac{1}{n}\sin^2\frac{nx}{2}.
\end{aligned}$$

We had better check the second equality $(*)$, since it may be a little bit troublesome:

$$\begin{aligned}
&\left(-\frac{1}{4}e^{-ix}+\frac{1}{2}-\frac{1}{4}e^{ix}\right)\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)e^{ijx} \\
&\quad = \frac{1}{2}\sum_{j=0}^{n-1}\left(1-\frac{j}{n}\right)e^{ijx}+\frac{1}{2}\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)e^{-ijx} \\
&\quad - \frac{1}{4}\sum_{j=0}^{n-1}\left(1-\frac{j}{n}\right)e^{i(j-1)x}-\frac{1}{4}\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)e^{-i(j+1)x} \\
&\quad - \frac{1}{4}\sum_{j=0}^{n-1}\left(1-\frac{j}{n}\right)e^{i(j+1)x}-\frac{1}{4}\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)e^{-i(j-1)x}.
\end{aligned} \qquad (5.16)$$

Changing the indices $j-1$ and $j+1$ to j, (5.16) is rewritten in the form

$$(5.16) = \underbrace{\boxed{\begin{aligned}&\frac{1}{2}\sum_{j=0}^{n-1}\left(1-\frac{j}{n}\right)e^{ijx}\\ &-\frac{1}{4}\sum_{j=-1}^{n-2}\left(1-\frac{j+1}{n}\right)e^{ijx}\\ &-\frac{1}{4}\sum_{j=1}^{n}\left(1-\frac{j-1}{n}\right)e^{ijx}\end{aligned}}}_{\langle A\rangle}\ \underbrace{\boxed{\begin{aligned}&+\frac{1}{2}\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)e^{-ijx}\\ &-\frac{1}{4}\sum_{j=2}^{n}\left(1-\frac{j-1}{n}\right)e^{-ijx}\\ &-\frac{1}{4}\sum_{j=0}^{n-2}\left(1-\frac{j+1}{n}\right)e^{-ijx}.\end{aligned}}}_{\langle B\rangle} \qquad (5.17)$$

We continue calculations, dividing (5.16) into two blocks $\langle A\rangle$ and $\langle B\rangle$. Note that the three formulas constituting the block $\langle A\rangle$ share common terms e^{ijx}, $j = 1, 2, \cdots, n-2$. The coefficients of these terms in $\langle A\rangle$ are seen to be cancelled:

$$\text{coefficient of } e^{ijx} = \frac{1}{2}\left(1-\frac{j}{n}\right)-\frac{1}{4}\left(1-\frac{j+1}{n}\right)-\frac{1}{4}\left(1-\frac{j-1}{n}\right)=0.$$

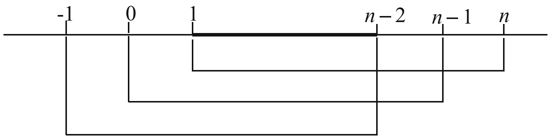


Fig. 5.1 Calculation of (5.17) ⟨A⟩

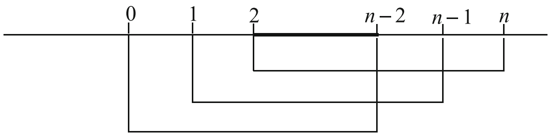


Fig. 5.2 Calculation of (5.17) ⟨B⟩

Hence, the remaining part of ⟨A⟩ is given by

$$\begin{aligned}\langle \mathrm{A}\rangle =&\frac{1}{2}\left(1-\frac{0}{n}\right)e^{i0x}+\frac{1}{2}\left(1-\frac{n-1}{n}\right)e^{i(n-1)x}\\&-\frac{1}{4}\left(1-\frac{-1+1}{n}\right)e^{i(-1)x}-\frac{1}{4}\left(1-\frac{0+1}{n}\right)e^{i0x}\\&-\frac{1}{4}\left(1-\frac{n-1-1}{n}\right)e^{i(n-1)x}-\frac{1}{4}\left(1-\frac{n-1}{n}\right)e^{inx}.\end{aligned}\tag{5.18}$$

Similar arguments apply also to the block ⟨B⟩. The constituent three formulas of ⟨B⟩ share common terms e^{-ijx}, $j=2,3,\cdots,n-2$. The coefficient of these terms in ⟨B⟩ must be zero:

$$\text{coefficient of } e^{-ijx}=\frac{1}{2}\left(1-\frac{j}{n}\right)-\frac{1}{4}\left(1-\frac{j-1}{n}\right)-\frac{1}{4}\left(1-\frac{j+1}{n}\right)=0.$$

Gathering the remaining parts of⟨B⟩, we have

$$\begin{aligned}\langle \mathrm{B}\rangle =&\frac{1}{2}\left(1-\frac{1}{n}\right)e^{-i1x}+\frac{1}{2}\left(1-\frac{n-1}{n}\right)e^{-i(n-1)x}\\&-\frac{1}{4}\left(1-\frac{(n-1)-1}{n}\right)e^{-i(n-1)x}-\frac{1}{4}\left(1-\frac{n-1}{n}\right)e^{-inx}\\&-\frac{1}{4}\left(1-\frac{0+1}{n}\right)e^{-i0x}-\frac{1}{4}\left(1-\frac{1+1}{n}\right)e^{-i1x}.\end{aligned}\tag{5.19}$$

The formula (5.16) is the sum of (5.18) and (5.19). Thus, removing the cancelled parts, we obtain

$$(5.16) = \langle \mathrm{A} \rangle + \langle \mathrm{B} \rangle = \frac{1}{n}\left(-\frac{1}{4}e^{-inx} + \frac{1}{2} - \frac{1}{4}e^{inx}\right).$$

This proves the equality $(*)$ in (5.15). (See Fig. 5.1 and Fig. 5.2.)

The representation (5.14) immediately follows from (5.15) and (5.13).

5.3 Spectral Synthesis on $[-\pi, \pi]$

We have discussed the concept of convolution on $\mathbb{R}$ and its applications in Chap. 3, Sect. 3.2. To start with, a similar concept of convolution on $[-\pi, \pi]$ is introduced.

For any $f, g \in \mathfrak{L}^1_{2\pi}(\mathbb{R}, \mathbb{C})$, we define the **convolution** $h(x)$ **on** $[-\pi, \pi]$ by

$$h(x) = \int_{-\pi}^{\pi} f(x-y)g(y)dy. \tag{5.20}$$

$h(x)$ is denoted by $(f * g)(x)$, the same notation as the usual convolution on $\mathbb{R}$.

(i) $h(x)$ is defined for a.e. $x \in [-\pi, \pi]$.
(ii) $h(x)$ is integrable on $[-\pi, \pi]$.
(iii) The convolution operation $*$ on $[-\pi, \pi]$ is commutative, associative and distributive (with respect to addition).
(iv) The Fourier coefficients of f, g and $f * g$ satisfy the relation

$$\widehat{f * g}(n) = \sqrt{2\pi}\,\hat{f}(n)\hat{g}(n),$$

where $\hat{f}(n)$ is given by

$$\hat{f}(n) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

$\hat{g}(n)$ and $\widehat{f * g}(n)$ are defined similarly.

Theorem 5.4 (special convolution; $e^{inx} * f$**)** *Assume that* $f \in \mathfrak{L}^1_{2\pi}(\mathbb{R}, \mathbb{C})$.

(i) *If* $\varphi(x) = e^{inx}$,

$$(\varphi * f)(x) = \sqrt{2\pi}\,\hat{f}(n)e^{inx}.$$

(ii) *If* $\varphi(x) = \sum_{j=-n}^{n} a_j e^{ijx}$,

$$(\varphi * f)(x) = \sqrt{2\pi} \sum_{j=-n}^{n} a_j \hat{f}(j)e^{ijx}.$$

Proof (i) If $\varphi(x) = e^{inx}$, we have

$$(\varphi * f)(x) = \int_{-\pi}^{\pi} e^{in(x-\xi)} f(\xi) d\xi = e^{inx} \int_{-\pi}^{\pi} f(\xi) e^{-in\xi} d\xi = \sqrt{2\pi} \hat{f}(n) e^{inx}.$$

(ii) is an immediate consequence of (i). □

We denote by $\sigma_n(x)$ the convolution of $f \in \mathfrak{L}^1_{2\pi}(\mathbb{R}, \mathbb{C})$ and K_n on $[-\pi, \pi]$; i.e.

$$\sigma_n(x) = (K_n * f)(x), \qquad x \in [-\pi, \pi]. \tag{5.21}$$

Then $\sigma_n(x)$ can be expressed as

$$\sigma_n(x) = \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) \hat{f}(j) e^{ijx}, \tag{5.22}$$

by (5.14) in the preceding section and Theorem 5.4.

The partial sum $S_n(x)$ of the Fourier series of f in complex form is given by

$$S_n(x) = \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \hat{f}(j) e^{ijx}. \tag{5.23}$$

The number of occurrences of $\hat{f}(j)e^{ijx}$ $(j = 0, \pm 1, \cdots, \pm(n-1))$ in the average of partial sums

$$\frac{1}{n}\{S_0(x) + S_1(x) + \cdots + S_{n-1}(x)\} \tag{5.24}$$

is $n - |j|$. Consequently, we obtain

$$\begin{aligned}(5.24) &= \frac{1}{n \cdot \sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} (n - |j|) \hat{f}(j) e^{ijx} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) \hat{f}(j) e^{ijx} = \sigma_n(x) \\ &\qquad \text{(by (5.22))}.\end{aligned}$$

Theorem 5.5 **($\sigma_n = (C, 1)$-sum)** *For* $f \in \mathfrak{L}^1_{2\pi}(\mathbb{R}, \mathbb{C})$,

$$\sigma_n(x) = \frac{1}{n}\{S_0(x) + S_1(x) + \cdots + S_{n-1}(x)\}. \tag{5.25}$$

That is, $\sigma_n(x)$ is nothing other than the $(C, 1)$-sum of the Fourier series in complex form.

Therefore if we impose an additional assumption that $\sum_{n=-\infty}^{\infty} \hat{f}(n)e^{inx}$ converges, it follows that[6]

$$\sigma_n(x) \to \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{inx} \quad \text{as} \quad n \to \infty.$$

On the other hand, $\sigma_n(x)$ converges to f in $\mathfrak{L}^1$ by (5.21) and Theorem 5.3. So some subsequence $\{\sigma_{n'}\}$ converges to f a.e. However, the original sequence $\{\sigma_n\}$ converges to $(1/\sqrt{2\pi}) \sum \hat{f}(n)e^{inx}$. Consequently, it turns out that $\{\sigma_n\}$ itself converges to f a.e.: $\sigma_n(x) \to f(x)$ a.e.

Thus an important result immediately follows.

Theorem 5.6 (spectral synthesis) *Assume that* $f \in \mathfrak{L}^1_{2\pi}(\mathbb{R}, \mathbb{C})$ *and*

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)| < \infty.$$

Then

$$\sigma_n(x) \to f(x) \quad a.e. \;\; as \;\; n \to \infty;$$

i.e.

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{inx}.$$

This theorem illuminates a procedure to recover the original function f when its Fourier coefficients are given. This operation is called the **spectral synthesis**.

5.4 Summability Kernels on $\mathbb{R}$

We now try to extend the Fejér method of summability and the spectral synthesis on $[-\pi, \pi]$ to similar problems on $\mathbb{R}$.[7]

[6] If a series $\sum_{n=0}^{\infty} a_n$ converges, it is $(C, 1)$-summable to the same limit (cf. Remark 2.4, p. 40).

[7] cf. Katznelson [3] Chap. VI, §1.

Definition 5.2 A family of continuous functions $\{k_\lambda : \mathbb{R} \to \mathbb{R}\}$ $(\lambda \in (0, \infty)$, or $\lambda \in \mathbb{N})$ is called a **summability kernel on** $\mathbb{R}$ if it satisfies:

(i) $\displaystyle\int_{-\infty}^{\infty} k_\lambda(x)dx = 1 \quad \text{for all } \lambda,$

(ii) $\|k_\lambda\|_1 = O(1) \quad \text{as} \quad \lambda \to \infty,$

(iii) $\displaystyle\lim_{\lambda\to\infty} \int_{|x|>\delta} |k_\lambda(x)|dx = 0 \quad \text{for any} \quad \delta > 0.$

If a function $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{R})$ satisfies

$$\int_{-\infty}^{\infty} f(x)dx = 1,$$

a summability kernel can be made based upon f. That is, if we define

$$k_\lambda(x) = \lambda f(\lambda x), \tag{5.26}$$

then $\{k_\lambda\}$ is a summability kernel. In fact, the condition (i) is verified by changing the variables: $y = \lambda x$. (ii) is satisfied, since

$$\|k_\lambda\|_1 = \int_{-\infty}^{\infty} |k_\lambda(x)|dx = \int_{-\infty}^{\infty} |f(y)|dy = \|f\|_1$$

for every $\lambda > 0$. It is also easy to check (iii), since

$$\int_{|x|>\delta} |k_\lambda(x)|dx = \int_{|y|>\lambda\delta} |f(y)|dy \to 0 \quad \text{as} \quad \lambda \to \infty.$$

For instance, if we define functions $A(x)$ and $G(x)$ by

$$A(x) = \frac{1}{2}e^{-|x|}, \quad G(x) = \frac{1}{\sqrt{\pi}}e^{-x^2},$$

the integrals of them over $\mathbb{R}$ are equal to 1. So it is possible to make summation kernels based upon them. The kernels based upon A and G are called the **Abel summability kernel** and the **Gauss summability kernel**, respectively.[8]

We concentrate here on the Fejér kernel on $\mathbb{R}$, which will play a crucial role later on.

---

[8]See Kawata [4] II, pp. 66–73. The Poisson summation kernel is also very useful, although it is not discussed in this book. Helson [2] and Malliavin [5] contain suggestive expositions of the Poisson kernel.

Define a function $K(x)$ by

$$K(x) = \frac{1}{2\pi}\left(\frac{\sin\frac{x}{2}}{\frac{x}{2}}\right)^2. \tag{5.27}$$

The following lemma can be verified by a simple calculation.

Lemma 5.2

$$K(x) = \frac{1}{2\pi}\int_{-1}^{1}(1-|\xi|)e^{i\xi x}d\xi. \tag{5.28}$$

Proof

$$\begin{aligned}
\int_{-1}^{1}(1-|\xi|)e^{i\xi x}d\xi &= \int_{-1}^{1}e^{i\xi x}d\xi - \int_{0}^{1}\xi e^{i\xi x}d\xi - \int_{-1}^{0}(-\xi)e^{i\xi x}d\xi \\
&= \frac{2\sin x}{x} - 2\int_0^1 \xi\cos\xi x d\xi \\
&= \frac{2\sin x}{x} - 2\left\{\frac{\sin\xi x}{x}\xi\Big|_0^1 - \int_0^1\frac{\sin\xi x}{x}d\xi\right\} = -\frac{2}{x^2}(\cos x - 1) \\
&= -\frac{2}{x^2}\left\{1 - 2\sin^2\frac{x}{2} - 1\right\} = \left(\frac{\sin\frac{x}{2}}{\frac{x}{2}}\right)^2.
\end{aligned}$$

This proves (5.28). □

$K(x)$ is clearly integrable. (The integral is equal to 1 as is shown later.) If we define $K_\lambda(x)$ $(\lambda > 0)$ by

$$K_\lambda(x) = \lambda K(\lambda x), \tag{5.29}$$

$\{K_\lambda\}$ satisfies the conditions (ii) and (iii) of summability kernels. It remains to show (i).

By the properties of the Fejér kernel on $[-\pi, \pi]$ (Lemma 2.4, pp. 42–43),

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{(n+1)x}{2}}{\sin\frac{x}{2}}\right)^2 dx = 1 \tag{5.30}$$

for any $\delta \in (0, \pi)$. For $\lambda = n+1$, we have

$$\begin{aligned}
K_{n+1}(x) &= \frac{1}{2\pi(n+1)}\left(\frac{\sin\frac{(n+1)x}{2}}{\frac{x}{2}}\right)^2 \\
&= \frac{1}{2\pi(n+1)}\left(\frac{\sin\frac{x}{2}}{\frac{x}{2}}\right)^2\left(\frac{\sin\frac{(n+1)x}{2}}{\sin\frac{x}{2}}\right)^2.
\end{aligned}$$

Hence

$$K_{n+1}(x) > \frac{1}{2\pi(n+1)}\left(\frac{\sin\delta}{\delta}\right)^2\left(\frac{\sin\frac{(n+1)x}{2}}{\sin\frac{x}{2}}\right)^2$$

for sufficiently small $\delta > 0$.[9] It follows that

$$\begin{aligned}&\left(\frac{\sin\delta}{\delta}\right)^2\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{(n+1)x}{2}}{\sin\frac{x}{2}}\right)^2 \\ &\qquad < \int_{-\delta}^{\delta}K_{n+1}(x)dx \leqq \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{n+1}\left(\frac{\sin\frac{(n+1)x}{2}}{\sin\frac{x}{2}}\right)^2 dx.\end{aligned} \tag{5.31}$$

The integration appearing in the first term of (5.31) satisfies

$$\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{(n+1)x}{2}}{\sin\frac{x}{2}}\right)^2 dx \to 1 \quad \text{as} \quad n\to\infty$$

for arbitrary $\delta \in (0,\pi)$, by Lemma 2.4 (pp. 42–43). The third term of (5.31) is equal to 1 for all n. Consequently, we obtain the evaluation

$$\left(\frac{\sin\delta}{\delta}\right)^2 < \int_{-\delta}^{\delta}K_{n+1}(x)dx \leqq 1$$

for sufficiently large n.

On the other hand, since

$$\int_{-\delta}^{\delta}K_{n+1}(x)dx \to \int_{-\infty}^{\infty}K(x)dx \quad \text{as} \quad n\to\infty,$$

we also get[10]

$$\left(\frac{\sin\delta}{\delta}\right)^2 \leqq \int_{-\infty}^{\infty}K(x)dx \leqq 1.$$

[9] Note that $(\sin z/z)' < 0$ for sufficiently small $z > 0$.

[10] $\int_{-\infty}^{\infty}K(x)dx = \int_{-\infty}^{\infty}K_{n+1}(x)dx = \int_{-\delta}^{\delta}K_{n+1}(x)dx + \int_{|x|>\delta}K_{n+1}(x)dx$. The last term tends to 0 as $n\to\infty$.

This evaluation holds good for any $\delta \in (0, \pi)$. Thus we finally obtain

$$\int_{-\infty}^{\infty} K(x)dx = 1,$$

which implies that the integral of $K_\lambda(x)$ is also equal to 1.

We conclude that $\{K_\lambda\}$ satisfies all the requirements of summability kernels. $\{K_\lambda\}$ is called the **Fejér summability kernel on** $\mathbb{R}$.

5.5 Spectral Synthesis on $\mathbb{R}$: Inverse Fourier Transforms on $\mathfrak{L}^1$

We now try to look for a method of inverse Fourier transforms. Given any integrable function, is it possible to find some function, the Fourier transform of which is exactly equal to it? We already know the positive answer to this question in the frameworks of $\mathfrak{L}^2$ (Plancherel's Theorem 4.3, p. 72), $\mathfrak{S}$(Theorem 4.2, p. 68) and $\mathfrak{S}'$ (Theorem 4.5, p. 84). But how about in the case of $\mathfrak{L}^1$?

The procedure to find some function, the Fourier transform of which is given is called **spectral synthesis**.

A vector-valued integration appearing in the next lemma (which corresponds to Lemma 5.1) is the one in the sense of Cauchy–Bochner.[11]

Lemma 5.3 *Let $\mathfrak{X}$ be a Banach space, $\varphi : \mathbb{R} \to \mathfrak{X}$ a bounded continuous function and $\{k_\lambda\}$ a summability kernel. Then*

$$\lim_{\lambda \to \infty} \int_{-\infty}^{\infty} k_\lambda(x)\varphi(x)dx = \varphi(0).$$

Proof Taking account of the condition (i) of summability kernels, we have

$$\int_{-\infty}^{\infty} k_\lambda(x)\varphi(x)dx - \varphi(0) = \int_{-\infty}^{\infty} k_\lambda(x)(\varphi(x) - \varphi(0))dx = \int_{-\delta}^{\delta} + \int_{|x|>\delta} = I_1 + I_2 \tag{5.32}$$

for any $\delta > 0$.

I_1 can be evaluated as

$$\|I_1\| \leqq \operatorname*{Max}_{|x| \leqq \delta} \|\varphi(x) - \varphi(0)\| \cdot \|k_\tau\|_1. \tag{5.33}$$

Let $\varepsilon > 0$ be any positive number. If we choose $\delta > 0$ sufficiently small, the right-hand side of (5.33) is less than ε. As for I_2, we obtain

[11] See Amann and Escher [1] II, pp. 17–23, Maruyama [6] pp. 329–333.

$$\|I_2\| \leqq \sup_{|x|>\delta} \|\varphi(x) - \varphi(0)\| \int_{|x|>\delta} |k_\lambda(x)| dx. \tag{5.34}$$

The right-hand side of (5.34) is also less than $\varepsilon > 0$ for sufficiently large $\lambda > 0$ by the condition (iii) of summability kernels and the boundedness of φ. The relation (5.32), and the evaluations (5.33), (5.34) imply

$$\int_{-\infty}^{\infty} k_\lambda(x)\varphi(x)dx \to \varphi(0) \quad \text{as} \quad \lambda \to \infty.$$

□

Given $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$, we define an operator $F : \mathbb{R} \to \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ by $F : x \mapsto \tau_x f$. By Theorem 5.1, F is uniformly continuous and bounded. Of course, $F(0) = f$. Thus we obtain the following result from Lemma 5.3.

Theorem 5.7 (convergence of $k_\lambda * f$) *Let $f, g \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ and $\{k_\lambda\}$ a summability kernel. Then*

$$\int_{-\infty}^{\infty} k_\lambda(x)\tau_x f dx \to f \quad \textit{as} \quad \lambda \to \infty$$

in $\|\cdot\|_1$.

Lemma 5.4 *Let $f, g \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. If h is expressed as*

$$h(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H(\xi)e^{i\xi x} dx$$

for some integrable function $H(\xi)$, then it follows that

$$(h * f)(x) = \int_{-\infty}^{\infty} H(\xi)\hat{f}(\xi)e^{i\xi x} d\xi.$$

Proof

$$(h * f)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(\xi)e^{i\xi(x-y)} d\xi \cdot f(y)dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y)e^{-i\xi y} dy \cdot H(\xi)e^{i\xi x} d\xi$$

$$= \int_{-\infty}^{\infty} H(\xi)\hat{f}(\xi)e^{i\xi x} d\xi.$$

□

When we use the Fejér summability kernel, we have

$$K_\lambda(x) = \frac{\lambda}{2\pi}\int_{-1}^{1}(1-|\xi|)e^{i\xi\cdot\lambda x}d\xi = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)e^{i\xi x}d\xi \tag{5.35}$$

(changing variables).

Specity $H(\xi)$ in Lemma 5.4 as

$$H(\xi) = \begin{cases} \frac{1}{\sqrt{2\pi}}\left(1-\frac{|\xi|}{\lambda}\right) & \text{on} \quad [-\lambda,\lambda], \\ 0 & \text{on} \quad [-\lambda,\lambda]^c. \end{cases} \tag{5.36}$$

Then $H \in \mathfrak{L}^1$ and

$$K_\lambda(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}H(\xi)e^{i\xi x}d\xi \tag{5.37}$$

by (5.35). Thanks to Lemma 5.4,

$$(K_\lambda * f)(x) = \int_{-\infty}^{\infty}H(\xi)\hat{f}(\xi)e^{i\xi x}d\xi.$$

This is an analogue of (5.22).

By Theorem 5.7, we have the following result.

Theorem 5.8 (analogue of $\sigma_n \to f$) *For $f \in \mathfrak{L}^1(\mathbb{R},\mathbb{C})$,*

$$\frac{1}{\sqrt{2\pi}}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)\hat{f}(\xi)e^{i\xi x}d\xi \to f \quad \textit{as} \quad \lambda\to\infty \quad \textit{in} \ \|\cdot\|_1.$$

Corollary 5.1 *If $f \in \mathfrak{L}^1(\mathbb{R},\mathbb{C})$ and $\hat{f}(\xi) = 0$ ($\xi \in \mathbb{R}$), then $f = 0$.*

This means that different integrable functions have different Fourier transforms. If we further assume that $\hat{f} \in \mathfrak{L}^1$, then

$$\begin{aligned}
&\lim_{\lambda\to\infty}\frac{1}{\sqrt{2\pi}}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)\hat{f}(\xi)e^{i\xi x}d\xi \\
&\qquad = \lim_{\lambda\to\infty}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\chi_{[-\lambda,\lambda]}(\xi)\left(1-\frac{|\xi|}{\lambda}\right)\hat{f}(\xi)e^{i\xi x}d\xi \\
&\qquad = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\hat{f}(\xi)e^{i\xi x}d\xi \quad \text{(dominated convergence theorem).}
\end{aligned}$$

Thus we obtain the formula of spectral synthesis; i.e. the inverse Fourier transforms on $\mathfrak{L}^1$.

Theorem 5.9 (spectral synthesis) *If both f and $\hat{f}$ are in $\mathfrak{L}^1(\mathbb{R}, \mathbb{C})$, then*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(t) e^{itx} dt.$$

The function space

$$\mathfrak{A} = \{f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C}) \mid \hat{f} \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})\}$$

is called the **Wiener algebra**, since it forms an algebra with respect to the usual operations and the convolution $*$.

Remark 5.1 In order to show that $\mathfrak{A}(\mathbb{R})$ is an algebra, it is enough to verify that $f * g \in \mathfrak{A}(\mathbb{R})$ for any $f, g \in \mathfrak{A}(\mathbb{R})$. Basic properties of $\mathfrak{A}(\mathbb{R})$ are picked up in the following.[12]

1° $f \in \mathfrak{A}(\mathbb{R}) \Leftrightarrow \hat{f} \in \mathfrak{A}(\mathbb{R})$.

Proof By Theorem 5.9,

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(t) e^{itx} dt$$

for $f \in \mathfrak{A}(\mathbb{R})$. If we define $g(x) = f(-x) (\in \mathfrak{L}^1(\mathbb{R}, \mathbb{C}))$,

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(t) e^{-itx} dt.$$

Hence

$$g = \hat{\hat{f}},$$

which implies that $\hat{f} \in \mathfrak{A}(\mathbb{R})$. The converse can be proved similarly. □

2° Define $\|f\|_{\mathfrak{A}(\mathbb{R})} = \|f\|_1 + \|\hat{f}\|_1$. Then

$$\|\hat{f}\|_\infty \leqq \frac{1}{\sqrt{2\pi}} \|f\|_{\mathfrak{A}(\mathbb{R})}.$$

(Obvious from (3.21), p. 52.)

3° $f \in \mathfrak{A}(\mathbb{R}) \Rightarrow f \in \mathfrak{L}^p(\mathbb{R}, \mathbb{C}) \quad (1 \leqq p \leqq \infty)$.

[12] cf. Malliavin [5] Chap. III, §2.

Proof $\int_{-\infty}^{\infty} |f|^p dx \leqq \|f\|_\infty^{p-1} \cdot \|f\|_1.$

$\square$

4° $f, g \in \mathfrak{A}(\mathbb{R}) \Rightarrow f * g \in \mathfrak{A}(\mathbb{R})$.

Proof Since $\|f * g\|_1 \leqq \|f\|_1 \cdot \|g\|_1$ and $\widehat{f * g} = \sqrt{2\pi}\hat{f}\hat{g}$, we have

$$\|\widehat{f * h}\|_1 = \sqrt{2\pi}\|\hat{f}\hat{g}\|_1 \leqq \sqrt{2\pi}\|\hat{f}\|_\infty\|\hat{g}\|_1 < \infty.$$

Hence $f * g \in \mathfrak{A}(\mathbb{R})$. $\square$

Since $H(\xi)$ defined by (5.36) is integrable,

$$H(\xi) = \widehat{K_\lambda}(\xi)$$

by (5.37) and Theorem 5.9. Consequently, it follows that

$$\widehat{K_\lambda * f}(\xi) = \begin{cases} \frac{1}{\sqrt{2\pi}}\left(1 - \frac{|\xi|}{\lambda}\right)\hat{f}(\xi) & \text{on} \quad [-\lambda, \lambda], \\ 0 & \text{on} \quad [-\lambda, \lambda]^c. \end{cases}$$

Combining this result with Theorem 5.8, we obtain the next theorem.

Theorem 5.10 (dense subspace of $\mathfrak{L}^1$) *$\{f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C}) \mid \operatorname{supp}\hat{f}$ is compact$\}$ is dense in $\mathfrak{L}^1$.*

References

1. Amann, H., Escher, J.: Analysis, I, II. Birkhäuser, Basel (2005, 2008)
2. Helson, H.: Harmonic Analysis. Wadsworth and Brooks, Belmont (1991)
3. Katznelson, Y.: An Introduction to Harmonic Analysis, 3rd edn. Cambridge University Press, Cambridge (2004)
4. Kawata, T.: Ohyo Sugaku Gairon (Elements of Applied Mathematics). I, II, Iwanami Shoten, Tokyo (1950, 1952) (Originally published in Japanese)
5. Malliavin, P.: Integration and Probability. Springer, New York (1995)
6. Maruyama, T.: Suri-keizaigaku no Hoho (Methods in Mathematical Economics). Sobunsha, Tokyo (1995) (Originally published in Japanese)

Chapter 6
Fourier Transforms of Measures

So far, we have studied Fourier transforms or Fourier coefficients of functions defined on $\mathbb{R}$ or $\mathbb{T}$. Inverse procedures to recover original functions from given Fourier transforms or Fourier coefficients were also discussed (spectral synthesis). However, there are many functions to which the methods of classical Fourier analysis can not be applied. In this chapter, we develop the theory of Fourier transforms of measures as a similar but new method to overcome such difficulties.

The celebrated theorems, due to G. Herglotz and S. Bochner, to represent positive semi-definite functions in terms of Fourier transforms of some positive measures will be made use of in later chapters. It is one of the aims of this chapter to give rigorous proofs of the Herglotz–Bochner theorems. There are several approaches to them. We present here a couple of ways. One is based on the summability method using the Fejér kernel. Another is due to the theory of Fourier transforms of distributions. Its relation to the representation theorem of one-parameter groups of unitary operators and to Fourier analysis of weakly stationary stochastic processes is discussed in detail in succeeding chapters.

6.1 Radon Measures

To start with, we briefly discuss the relation between the space of continuous functions and its dual space.

Let X be a locally compact Hausdorff topological space. We denote by $\mathfrak{C}_\infty(X, \mathbb{C})$(resp. $\mathfrak{C}_\infty(X, \mathbb{R})$) the set of all the complex-valued (resp. real-valued) continuous functions vanishing at infinity.[1] It is a Banach space with respect to the norm of uniform convergence:

[1] A function f is said to **vanish at infinity** if there exists, for each $\varepsilon > 0$, some compact set $K \subset X$ such that $|f(x)| < \varepsilon$ for all $x \in K^c$.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_6

$$\|f\|_\infty = \sup_{x \in X} |f(x)|, \quad f \in \mathfrak{C}_\infty.$$

Let μ be a complex-valued regular measure on $(X, \mathcal{B}(X))$, the total variation $|\mu|$ of which is finite. $\mathcal{B}(X)$ is the Borel σ-field on X. The completion of such a measure with respect to $|\mu|$ is called a **Radon measure**. The set of all the Radon measures is denoted by $\mathfrak{M}(X)$. $\mathfrak{M}(X)$ is a Banach space, the norm of which is given by the total variation $\|\mu\|_{\mathfrak{M}(X)} = |\mu|$. In particular, the set of all the positive (real-valued) Radon measures is denoted by $\mathfrak{M}_+(X)$.

For any $\mu \in \mathfrak{M}(X)$, we define a linear functional Λ_μ on $\mathfrak{C}_\infty(X, \mathbb{C})$ by

$$\Lambda_\mu f = \int_X f(x) d\mu, \quad f \in \mathfrak{C}_\infty(X, \mathbb{C}).$$

Then Λ_μ is bounded; i.e. $\Lambda_\mu \in \mathfrak{C}_\infty(X, \mathbb{C})'$. Conversely, for any $\Lambda \in \mathfrak{C}_\infty(X, \mathbb{C})'$, there exists a measure $\mu_\Lambda \in \mathfrak{M}(X)$ which satisfies

$$\Lambda f = \int_X f(x) d\mu_\Lambda, \quad f \in \mathfrak{C}_\infty(X, \mathbb{C})$$

and $\|\Lambda\| = |\mu|$. Such a measure μ is uniquely determined. Thus the two Banach spaces $\mathfrak{C}_\infty(X, \mathbb{C})'$ and $\mathfrak{M}(X)$ are isomorphic to each other. This result is called the Riesz–Markov–Kakutani theorem.[2]

Riesz–Markov–Kakutani theorem *Let X be a locally compact Hausdorff topological space. Then $\mathfrak{C}_\infty(X, \mathbb{C})'$ and $\mathfrak{M}(X)$ are isometrically isomorphic as Banach spaces;* i.e.

$$\mathfrak{M}(X) \cong \mathfrak{C}_\infty(X, \mathbb{C})'.$$

6.2 Fourier Coefficients of Measures (1)

The space $\mathfrak{C}(\mathbb{T}, \mathbb{C})$ of complex-valued continuous function on $\mathbb{T}$ is a Banach space with the uniform convergence norm. By the Riesz–Markov–Kakutani theorem, the Banach space $\mathfrak{M}(\mathbb{T})$ of complex-valued Radon measures on $\mathbb{T}$ (norm is given by the total variation) is isomorphic to the dual space of $\mathfrak{C}(\mathbb{T}, \mathbb{C})$. From now on, the measurable space $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ is identified with $([-\pi, \pi), \mathcal{B}([-\pi, \pi)))$. (cf. Appendix A.)

---

[2] Malliavin [7] pp. 96–97, Maruyama [9] pp. 300–315.

Definition 6.1 For $\mu \in \mathfrak{M}(\mathbb{T})$,

$$\hat{\mu}(n) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-inx} d\mu(x),\text{[3]} \quad n \in \mathbb{Z}$$

are called the Fourier coefficients of μ.

Fourier coefficients of a measure are also called **Fourier–Stieltjes** coefficients in order to distinguish them from Fourier coefficients of a function. In particular, the Fourier coefficients of a measure $\mu_f = f dx$ defined by a function $f \in \mathfrak{L}^1(\mathbb{T}, \mathbb{C})$ are given by

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

These are nothing other than usual Fourier coefficients of f.

Theorem 6.1 *For any $f \in \mathfrak{C}(\mathbb{T}, \mathbb{C})$ and $\mu \in \mathfrak{M}(\mathbb{T})$,*

$$\int_{-\pi}^{\pi} f(x) d\mu = \lim_{n \to \infty} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) \hat{f}(j) \hat{\mu}(-j). \tag{6.1}$$

Proof Consider first a trigonometric polynomial

$$P(x) = \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} a_j e^{ijx}$$

as a special case of f. Then we obtain

$$\int_{-\pi}^{\pi} P(x) d\mu = \sum_{j=-(n-1)}^{n-1} a_j \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{ijx} d\mu = \sum_{j=-(n-1)}^{n-1} a_j \hat{\mu}(-j).$$

If we define $\sigma_n(x)$ by

$$\sigma_n(x) = \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) \hat{f}(j) e^{ijx}$$

as in the preceding chapter (p. 110), it follows that

$$\int_{-\pi}^{\pi} \sigma_n(x) d\mu = \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) \hat{f}(j) \hat{\mu}(-j)$$

by the above result.

[3]The interval of integration should be interpreted as $[-\pi, \pi)$.

By Theorem 2.6 (p. 43), $\sigma_n(x)$ uniformly converges to $f(x)$ as $n \to \infty$. Hence

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)d\mu &= \lim_{n\to\infty}\int_{-\pi}^{\pi}\sigma_n(x)d\mu \\ &= \lim_{n\to\infty}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)\hat{f}(j)\hat{\mu}(-j).\end{aligned}$$

□

It is assured that the limit of the right-hand side of (6.1) exists by Theorem 6.1. That is, $(C, 1)$-summability of $\sum \hat{f}(j)\hat{\mu}(j)$ is verified.

Corollary 6.1 *Let $\mu \in \mathfrak{M}(\mathbb{T})$. If $\hat{\mu}(n) = 0$ for all $n \in \mathbb{Z}$, then $\mu = 0$.*

This is obvious, since

$$\hat{\mu}(n) = 0 \quad \text{for all } n \quad \Longrightarrow \int_{-\pi}^{\pi} f(x)d\mu = 0 \quad \text{for all } f \in \mathfrak{C}(\mathbb{T}, \mathbb{C})$$

by Theorem 6.1. Furthermore, the corollary implies the uniqueness of Fourier coefficients of a measure; i.e.

$$\mu \neq \nu \quad \Longrightarrow \quad \hat{\mu} \neq \hat{\nu}.$$

Define a couple of operators $S_n : \mathfrak{C}(\mathbb{T}, \mathbb{C}) \to \mathfrak{C}(\mathbb{T}, \mathbb{C})$ and $S_n^* : \mathfrak{M}(\mathbb{T}) \to \mathfrak{M}(\mathbb{T})$ by

$$S_n : f \mapsto \frac{1}{\sqrt{2\pi}}\sum_{j=-(n-1)}^{n-1}\hat{f}(j)e^{ijx}, \tag{6.2}$$

$$S_n^* : \mu \mapsto \frac{1}{\sqrt{2\pi}}\sum_{j=-(n-1)}^{n-1}\hat{\mu}(j)e^{ijx}dx. \tag{6.3}$$

They are bounded linear operators which associate f and μ with the partial sums of their Fourier series. If we change $\hat{f}(j)e^{ijx}$ and $\hat{\mu}(j)e^{ijx}$ by $\hat{f}(-j)e^{-ijx}$ and $\hat{\mu}(-j)e^{-ijx}$, respectively, there occurs no change in what (6.2) and (6.3) mean. Let us compute $\mu \circ S_n$ and $S_n^*\mu$ at $f \in \mathfrak{C}(\mathbb{T}, \mathbb{C})$, regarding $\mu \in \mathfrak{M}(\mathbb{T})$ as a bounded linear functional on $\mathfrak{C}(\mathbb{T}, \mathbb{C})$.

$$\begin{aligned}(\mu \circ S_n)f &= \int_{-\pi}^{\pi}(S_n f)(x)d\mu = \int_{-\pi}^{\pi}\frac{1}{\sqrt{2\pi}}\sum_{j=-(n-1)}^{n-1}\hat{f}(j)e^{ijx}d\mu \\ &= \frac{1}{\sqrt{2\pi}}\sum_{j=-(n-1)}^{n-1}\hat{f}(j)\int_{-\pi}^{\pi}e^{ijx}d\mu = \sum_{j=-(n-1)}^{n-1}\hat{f}(j)\hat{\mu}(-j).\end{aligned} \tag{6.4}$$

On the other hand, we have

$$
\begin{aligned}
(S_n^*\mu)f &= \int_{-\pi}^{\pi} f(x)\frac{1}{\sqrt{2\pi}}\sum_{j=-(n-1)}^{n-1}\hat{\mu}(j)e^{ijx}dx \\
&= \frac{1}{\sqrt{2\pi}}\sum_{j=-(n-1)}^{n-1}\hat{\mu}(-j)\int_{-\pi}^{\pi} f(x)e^{-ijx}dx \qquad (6.5)\\
&= \sum_{j=-(n-1)}^{n-1}\hat{f}(j)\hat{\mu}(-j).
\end{aligned}
$$

Comparing (6.4) and (6.5), we obtain

$$
(\mu\circ S_n)f = (S_n^*\mu)f \tag{6.6}
$$

for every $f \in \mathfrak{C}(\mathbb{T},\mathbb{C})$ and $\mu \in \mathfrak{M}(\mathbb{T})$. Thus we should observe that S_n^* is the dual operator of S_n.[4]

Similarly, if we define the operators σ_n and σ_n^* by

$$
\sigma_n : f \mapsto \frac{1}{\sqrt{2\pi}}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)\hat{f}(j)e^{ijx}, \tag{6.7}
$$

$$
\sigma_n^* : \mu \mapsto \frac{1}{\sqrt{2\pi}}\sum_{j=-(n-1)}^{n-1}\left(1-\frac{|j|}{n}\right)\hat{\mu}(j)e^{ijx}dx, \tag{6.8}
$$

σ_n^* is the dual operator of σ_n.

Theorem 6.2 *For any $\mu \in \mathfrak{M}(\mathbb{T})$,*

$$
w^*\text{-}\lim_{n\to\infty}\sigma_n^*\mu = \mu. \tag{6.9}
$$

Proof $\sigma_n f(x)$ uniformly converges to $f(x)$ for any $f \in \mathfrak{C}(\mathbb{T},\mathbb{C})$ (cf. Theorem 2.6, p. 43). Hence we have

$$
\begin{aligned}
\mu f &= \int_{-\pi}^{\pi} f(x)d\mu = \lim_{n\to\infty}\int_{-\pi}^{\pi}\sigma_n f(x)d\mu = \lim_{n\to\infty}(\mu\circ\sigma_n)f \\
&= \lim_{n\to\infty}(\sigma_n^*\mu)f = \lim_{n\to\infty}\int_{-\pi}^{\pi} f(x)d(\sigma_n^*\mu).
\end{aligned}
$$

[4] See p. 83 for dual operators.

The fourth equality comes from the fact that σ_n^* is the dual operator of σ_n. The relation (6.9) immediately follows from this. □

6.3 Fourier Coefficients of Measures (2)

We have already studied conditions under which a given sequence $\{a_n\}_{n=-\infty}^{\infty}$ of complex numbers is the one of Fourier coefficients of some periodic distribution (Theorem 4.8, p. 97). In this section, we try to look for certain conditions which assure that the a_n's are Fourier coefficients of a Radon measure on $\mathbb{T} = [-\pi, \pi)$. In particular, under what conditions are the a_n's Fourier coefficients of some positive Radon measure on $\mathbb{T}$? It is the well-known theorem due to G. Herglotz which gives an answer to this question.[5]

Theorem 6.3 *For a given sequence $\{a_n\}_{n=-\infty}^{\infty}$, the following two statements are equivalent.*

(i) *There exists some $\mu \in \mathfrak{M}(\mathbb{T})$ which satisfies $\|\mu\| \leqq C$ (constant) and $a_n = \hat{\mu}(n)$ $(n \in \mathbb{Z})$.*
(ii) *For any trigonometric polynomial P, it holds good that*

$$\left| \sum_{n=-\infty}^{\infty} \hat{P}(n) a_{-n} \right| \leqq C \|P\|_\infty. \tag{6.10}$$

Proof (i)⇒(ii): Let $P(x)$ be a trigonometric polynomial. By Theorem 6.1, we have

$$\int_{-\pi}^{\pi} P(x) d\mu = \lim_{n\to\infty} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) \hat{P}(j)\hat{\mu}(-j), \tag{6.11}$$

where $\hat{P}(j) = 0$ except some finite j's. Hence

$$\text{the right-hand side of (6.11)} = \sum_{n=-\infty}^{\infty} \hat{P}(n)\hat{\mu}(-n) \underset{\text{(i)}}{=} \sum_{n=-\infty}^{\infty} \hat{P}(n) a_{-n}. \tag{6.12}$$

Since $\|\mu\| \leqq C$, (6.11) and (6.12) imply

$$\left| \int_{-\pi}^{\pi} P(x) d\mu \right| = \left| \sum_{n=-\infty}^{\infty} \hat{P}(n) a_{-n} \right| \leqq C \|P\|_\infty.$$

This proves (ii).

[5]For Sects. 6.3–6.8, we are very much indebted to Katznelson [5] Chap. I, §7 and Chap. VI, except for the distribution approach to the Herglotz–Bochner theorem.

(ii)⇒(i): Let $\mathcal{T}$ be the set of all the trigonometric polynomials on $\mathbb{T}$ (with uniform convergence topology). We define a linear functional $\Lambda : \mathcal{T} \to \mathbb{C}$ by

$$\Lambda : P \mapsto \sum_{n=-\infty}^{\infty} \hat{P}(n) a_{-n}. \tag{6.13}$$

Λ is bounded; i.e. $\|\Lambda\| \leqq C$ for some $C > 0$. Λ can be extended uniquely to a bounded linear functional on $\mathfrak{C}(\mathbb{T}, \mathbb{C})$, preserving the norm. We denote this extended functional also by Λ. By the Riesz–Markov–Kakutani theorem, there exists a Radon measure $\mu \in \mathfrak{M}(\mathbb{T})$ which satisfies

$$\Lambda f = \int_{-\pi}^{\pi} f(x) d\mu, \quad f \in \mathfrak{C}(\mathbb{T}, \mathbb{C}).$$

For a special case of $f(x) = e^{inx}$, we have

$$\Lambda e^{inx} = \int_{-\pi}^{\pi} e^{inx} d\mu = \sqrt{2\pi}\, \hat{\mu}(-n). \tag{6.14}$$

On the other hand, (6.13) implies

$$\Lambda e^{inx} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{inx} e^{-inx} dx \cdot a_{-n} = \sqrt{2\pi} \cdot a_{-n}. \tag{6.15}$$

Then $a_n = \hat{\mu}(n)$ follows from (6.14) and (6.15). It is obvious that $\|\mu\| \leqq C$. □

Corollary 6.2 *For a sequence $\{a_n\}_{n=-\infty}^{\infty}$ of complex numbers, the following two statements are equivalent:*

(i) *There exists some $\mu \in \mathfrak{M}(\mathbb{T})$ which satisfies $\|\mu\| \leqq C$ (constant) and $a_n = \hat{\mu}(n)$ $(n \in \mathbb{Z})$.*
(ii) *Let $\sigma_n(x)$ be the $(C, 1)$-sum of a series*

$$\frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} a_n e^{inx},$$

that is,

$$\sigma_n(x) = \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) a_j e^{ijx}. \tag{6.16}$$

Then the measures $\sigma_n dx \in \mathfrak{M}(\mathbb{T})$ are uniformly bounded, i.e.

$$\|\sigma_n dx\| \leqq C' \quad \textit{for all} \quad n \in \mathbb{Z}$$

for some C' (constant).

Remark 6.1 Let $P(x)$ be any trigonometric polynominal

$$P(x) = \sum_{j=-n}^{n} \alpha_j e^{ijx} \quad (\alpha_j \in \mathbb{C}).$$

Then the function the Fourier series of which is $P(x)$ is $P(x)$ itself.[6] This applies to $\sigma_n(x)$ defined above.

Proof of Corollary 6.2 (i)⇒(ii): Assume that there exists some $\mu \in \mathfrak{M}(\mathbb{T})$ such that $\|\mu\| \leqq C$ and $a_n = \hat{\mu}(n)$ $(n \in \mathbb{Z})$. Then

$$\begin{aligned}
\sigma_n(x)dx &= \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) a_j e^{ijx} dx \\
&= \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) \hat{\mu}(j) e^{ijx} dx \\
&= \sigma_n^* \mu.
\end{aligned}$$

That is, $\sigma_n dx = \sigma_n^* \mu$. By Theorem 6.2, $w^*\text{-}\lim_{n\to\infty} \sigma_n^* \mu = \mu$. Hence $\sup_n \|\sigma_n^* \mu\| < \infty$.

(ii)⇒ (i): Conversely, suppose that $\|\sigma_n dx\| \leqq C'$ for all $n \in \mathbb{Z}$. Let P be any trigonometric polynomial. Defining σ_n by (6.16), we can prove that

$$\sum \hat{P}(j) a_{-j} = \lim_{n\to\infty} \int_{-\pi}^{\pi} P(x) \sigma_n(x) dx. \tag{6.17}$$

In fact,

$$\int_{-\pi}^{\pi} P(x)\sigma_n(x)dx = \int_{-\pi}^{\pi} P(x) \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) a_j e^{ijx} dx$$

[6] The Fourier coefficients of $P(x)$ are given by

$$\frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} P(x) e^{-ijx} dx = \sqrt{2\pi}\alpha_j$$

for $|j| \leqq n$, and 0 for $|j| \geqq n+1$. Hence the Fourier series of $P(x)$ is computed as

$$P(x) = \sum_{j=-n}^{n} \sqrt{2\pi}\alpha_j \frac{1}{\sqrt{2\pi}} e^{ijx} = \sum_{j=-n}^{n} \alpha_j e^{ijx}.$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} P(x) \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) a_{-j} e^{-ijx} dx$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) a_{-j} \int_{-\pi}^{\pi} P(x) e^{-ijx} dx$$

$$= \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) a_{-j} \hat{P}(j).$$

$\hat{P}(j) = 0$ except finite number of j's. We obtain (6.17) by letting $n \to \infty$.

(ii) and (6.17) imply

$$\left| \sum \hat{P}(j) a_{-j} \right| \leqq C' \|P\|.$$

Consequently, we have (i) by Theorem 6.3. □

Remark 6.2 Even if we change the statement "$\|\sigma_n dx\| \leqq C'$ for all $n \in \mathbb{Z}$" in Corollary 6.2(i) to "$\|\sigma_n dx\| \leqq C'$ for infinitely many $n \in \mathbb{Z}$", the equivalence of (i) and (ii) holds good.

If we denote indices n for which $\|\sigma_n dx\| \leqq C'$ by n', (6.17) implies

$$\sum \hat{P}(j) a_{-j} = \lim_{n\to\infty} \int_{-\pi}^{\pi} P(x) \sigma_n dx = \lim_{n'\to\infty} \int_{-\pi}^{\pi} P(x) \sigma_{n'} dx.$$

Hence $|\sum \hat{P}(j) \bar{a}_j| \leqq C' \|P\|$.

We should remind the readers of this remark in later discussions in Sect. 6.4. That is, if (6.18) in Lemma 6.2 holds good for infinitely many n, it is equivalent to (i).

6.4 Herglotz's Theorem

Given a sequence $\{a_n\}_{n=-\infty}^{\infty}$ of complex numbers, we pose a question: under what conditions can we represent a_n's as Fourier coefficients of some positive Radon measure? It is Herglotz's theorem which answers this question.

We denote the $(C, 1)$-sum of the Fourier series of a function f by $\sigma_n f$ as in Sect. 6.2.

Lemma 6.1 *If $f \in \mathfrak{L}^1(\mathbb{T}, \mathbb{C})$ is real-valued and $m \leqq f(x)$ a.e., then $m \leqq \sigma_n f(x)$.*

Proof The Fejér kernel K_n satisfies $K_n(x) \geqq 0$ and

$$\int_{-\pi}^{\pi} K_n(x) dx = 1.$$

Hence

$$\sigma_n f(x) - m = \int_{-\pi}^{\pi} K_n(\xi)(f(x-\xi) - m)d\xi \geqq 0.$$

□

Lemma 6.2 *For a sequence* $\{a_n\}_{n=-\infty}^{\infty}$ *of complex numbers, the following two statements are equivalent:*

(i) *There exists a positive Radon measure* $\mu \in \mathfrak{M}(\mathbb{T})$ *such that* $a_n = \hat{\mu}(n)$ $(n \in \mathbb{Z})$.
(ii)

$$\sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) a_j e^{ijx} \geqq 0 \tag{6.18}$$

for any $n \in \mathbb{Z}$ *and* $x \in \mathbb{T}$.

Proof (i)⇒(ii): Suppose that $a_n = \hat{\mu}(n)$ $(n \in \mathbb{Z})$. Then

$$\frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) a_j e^{ijx} = \frac{1}{\sqrt{2\pi}} \sum_{j=-(n-1)}^{n-1} \left(1 - \frac{|j|}{n}\right) \hat{\mu}(j) e^{ijx} \tag{6.18$'$}$$

is the $(C, 1)$-sum of the Fourier series of μ. $\sigma_n^* \mu$ denotes the measure determined by (6.18′) in the manner of Sect. 6.2 (cf.(6.8)). For any $f \in \mathfrak{C}(\mathbb{T}, \mathbb{R})$ with $f \geqq 0$, we have

$$\int_{-\pi}^{\pi} f(x) d(\sigma_n^* \mu) = \int_{-\pi}^{\pi} \sigma_n f(x) d\mu, \tag{6.19}$$

where $\sigma_n f$ is the $(C, 1)$-sum of the Fourier series of f. By $f \geqq 0$ and Lemma 6.1, $\sigma_n f(x) \geqq 0$. Consequently, (6.19) $\geqq 0$. Since this is true for all the continuous functions $f \geqq 0$, we obtain $\sigma_n^* \mu \geqq 0$. (6.18) immediately follows.

(ii)⇒(i): Regarding (6.18) as the $(C, 1)$-sum of the series $\varphi(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}$, we denote it by $\sigma_n \varphi$. Assume that $\sigma_n \varphi(x) \geqq 0$. $\sigma_n \varphi$, viewed as a measure, satisfies

$$\|\sigma_n \varphi\| = \int_{-\pi}^{\pi} \sigma_n \varphi(x) dx = 2\pi a_0.$$

Hence, by Corollary 6.2, there exists a $\mu \in \mathfrak{M}(\mathbb{T})$ such that $a_n = \hat{\mu}(n)$. Since

$$\sigma_n^* \mu(x) = \sigma_n \varphi(x) dx$$

(cf. p. 125), Theorem 6.2 implies that

$$\int_{-\pi}^{\pi} f(x)d\mu = \lim_{n\to\infty}\int_{-\pi}^{\pi} f(x)d(\sigma_n^*\mu) = \lim_{n\to\infty}\int_{-\pi}^{\pi} f(x)\sigma_n\varphi(x)dx \geqq 0 \tag{6.20}$$

holds good for all the continuous functions $f \geqq 0$. So the measure μ must be positive. □

We now define the key concept to state and prove Herglotz's theorem.

Definition 6.2 Let G be either $\mathbb{Z}$ or $\mathbb{R}$.[7] A function $f : G \to \mathbb{C}$ is said to be **positive semi-definite** if

$$\sum_{i,j=1}^{n} f(x_i - x_j)\lambda_i\bar{\lambda}_j \geqq 0 \tag{6.21}$$

for any finite elements $x_1, x_2, \cdots, x_n$ of G and $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$.

The simple properties of a positive semi-definite function can easily be verified.

Lemma 6.3 *Let G be either $\mathbb{Z}$ or $\mathbb{R}$. A positive semi-definite function $f : G \to \mathbb{C}$ satisfies the following properties:*

(i) $f(0) \geqq 0$,
(ii) $\overline{f(x)} = f(-x)$,
(iii) $|f(x)| \leqq f(0)$.

Proof (i) This is an obvious result if we consider a special case where $n = 1$, $x_1 = 0$ and $\lambda_1 = 1$ in the above definition.

(ii) Consider the special case where $n = 2$, $x_1 = 0$, $x_2 = x$, $\lambda_1 = \lambda_2 = 1$. Then

$$2f(0) + f(x) + f(-x) \geqq 0. \tag{6.22}$$

If $\lambda_2 = i$, we have

$$f(0) + f(x)i - f(-x)i - f(0) = f(x)i - f(-x)i \geqq 0. \tag{6.23}$$

Combining (i) and (6.22), we learn that $f(x) + f(-x)$ is real. And (6.23) tells us that $f(x) - f(-x)$ is purely imaginary. Hence

$$\overline{f(x)} = f(-x).$$

[7] G may be, more generally, a commutative group.

(iii) In the case of $|f(x)| = 0$, (iii) is obvious by (i). So we have only to check the case $|f(x)| > 0$. If $n = 2,\ x_1 = 0,\ x_2 = x,\ \lambda_1 = -f(x),\ \lambda_2 = |f(x)|$, then

$$2f(0)|f(x)|^2 - 2|f(x)|^3 \geqq 0.$$

This implies

$$|f(x)| \leqq f(0),$$

since $|f(x)| > 0$. □

Remark 6.3

1° Consider an $(n \times n)$-matrix (α_{ij}), the (i, j)-element α_{ij} of which is $f(x_i - x_j)$ $(1 \leqq i, j \leqq n)$. If f is positive semi-definite, then (α_{ij}) is an Hermitian matrix by Lemma 6.3(ii). (6.21) means that the Hermitian form determined by (α_{ij}) is positive semi-definite.

2°

$$\begin{pmatrix} f(0) & f(-x) \\ f(x) & f(0) \end{pmatrix}$$

is an Hermitian matrix obtained in the case $n = 2,\ x_1 = 0,\ x_2 = x$. The Hermitian form determined by this matrix is positive semi-definite. Thanks to the well-known "sign theorem"of sub-determinants, Lemma 6.3 is almost trivial.[8]

Lemma 6.4 *Assume that a function $f : \mathbb{R} \to \mathbb{C}$ is positive semi-definite and continuous at $t = 0$. Then it has the following properties:*

(i) *f is uniformly continuous on $\mathbb{R}$.*
(ii) *For any continuous function $\zeta : \mathbb{R} \to \mathbb{C}$,*

$$\int_0^T \int_0^T f(s-t)\zeta(s)\overline{\zeta(t)}dsdt \geqq 0, \tag{6.24}$$

$(T > 0)$.

Proof (i) Since $|f(t)| \leqq f(0)$ by Lemma 6.3, $f \equiv 0$ in the case of $f(0) = 0$. Hence it is sufficient to consider the case of $f(0) > 0$. Without loss of generality, we may assume $f(0) = 1$. Specifying the value of t as $t_1 = 0,\ t_2 = t$ and $t_3 = t+h$, we have

[8] See textbooks on linear algebra for Hermitian forms and their signs.

$$\begin{vmatrix} f(0) & f(-t) & f(-t-h) \\ f(t) & f(0) & f(-h) \\ f(t+h) & f(h) & f(0) \end{vmatrix} = 1 - |f(t)|^2 - |f(t+h)|^2$$
$$- |f(h)|^2 + 2\mathcal{R}e\{f(t)f(h)\overline{f(t+h)}\} \geqq 0$$

by the positive semi-definiteness of f (sign theorem of Hermitian forms). Thus it follows that

$$\begin{aligned} |f(t) - f(t+h)|^2 &= |f(t)|^2 + |f(t+h)|^2 - 2\mathcal{R}e\{f(t)\overline{f(t+h)}\} \\ &\leqq 1 - |f(h)|^2 + 2\mathcal{R}e\{f(t)\overline{f(t+h)}[f(h)-1]\} \\ &\leqq 1 - |f(h)|^2 + 2|1 - f(h)| \leqq 4|1 - f(h)|. \end{aligned}$$

Since we are assuming that $f(0) = 1$, f is uniformly continuous.

(ii) Consider two decompositions of $[0, T]$: $0 = s_0 < s_1 < \cdots < s_n = T$ and $0 = t_0 < t_1 < \cdots < t_n = T$. By the continuity of f, the integration in (6.24) is the limit of its Riemann sum

$$\sum_{i,j=1}^{n} f(s_i - t_j)\zeta(s_i)\overline{\zeta(t_j)}(s_i - s_{i-1})(t_j - t_{j-1}).$$

Since f is positive semi-definite, the Riemann sum is $\geqq 0$. Hence (ii) follows. □

A sequence $\{a_n\}_{n=-\infty}^{\infty}$ of complex numbers can be regarded as a function φ : $\mathbb{Z} \to \mathbb{C}$. Therefore we can talk about the positive semi-definiteness of $\{a_n\}$. That is, $\{a_n\}$ is positive semi-definite if

$$\sum_{i,j=1}^{p} a_{n_i - n_j}\lambda_i\bar{\lambda}_j \geqq 0$$

for any finite integers $n_1, n_2, \cdots, n_p$ and $\lambda_1, \lambda_2, \cdots, \lambda_p \in \mathbb{C}$.

Theorem 6.4 (Herglotz)[9] *For a sequence $\{a_n\}_{n=-\infty}^{\infty}$ of complex numbers, the following two statements are equivalent.*

(i) $\{a_n\}_{n=-\infty}^{\infty}$ *is positive semi-definite.*
(ii) *There exists a positive Radon measure $\mu \in \mathfrak{M}(\mathbb{T})$ which satisfies $a_n = \hat{\mu}(n)$ $(n \in \mathbb{Z})$.*

---

[9]Herglotz [3]. Essentially the same result is obtained by Carathéodory [2] and Toeplitz [13]. The proof given here is due to Katznelson [5] Chap. 1, §7.

Proof (ii)⇒(i): Assume that μ is a positive Radon measure on $\mathbb{T}$ such that $a_n = \hat{\mu}(n)$ $(n \in \mathbb{Z})$. Fix a finite number of integers, say $n_1, n_2, \cdots, n_p$. Then for any $n, m \in \{n_1, n_2, \cdots, n_p\}$ and any complex numbers λ_n, λ_m, we have

$$\begin{aligned}\sum_{n,m} a_{n-m}\lambda_n\bar{\lambda}_m &= \sum_{n,m} \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-i(n-m)x} d\mu \cdot \lambda_n\bar{\lambda}_m \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \left|\sum_n \lambda_n e^{-inx}\right|^2 d\mu \geqq 0.\end{aligned}$$

(i)⇒(ii): We write formally

$$\varphi(x) \sim \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} a_n e^{inx}$$

and define

$$\lambda_n = \begin{cases} e^{inx} & \text{if} \quad |n| \leqq N-1, \\ 0 & \text{if} \quad |n| \geqq N \end{cases}$$

for any $N \in \mathbb{N}$. Then

$$\sum a_{n-m}\lambda_n\bar{\lambda}_m = \sum_j C_{j,N} a_j e^{ijx} \tag{6.25}$$

$$\text{where} \quad C_{j,N} = \mathrm{Max}\{0, 2N-1-|j|\}.$$

Hence

$$\sigma_{2N}\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{2N-1} \sum_j C_{j,N} a_j e^{ijx}.$$

By (i) and (6.25), we obtain $\sigma_{2N}\varphi(x) \geqq 0$ for all $N \in \mathbb{N}$. (ii) follows by Lemma 6.2 and Remark 6.2. □

Remark 6.4 The positive measure μ which satisfies (ii) is unique by Corollary 6.1.

So far we have proved the Herglotz theorem having recourse to the summability method. We now turn to discussing an alternative proof based upon the theory of periodic distributions.[10]

The proof of the "if" part is easy and well-known. So it is enough to show the "only if" part.

[10]Due to Maruyama [9]. I acknowledge the priority of Lax [6], pp. 142–143 for the idea of proof.

Assume that $\{a_n\}_{n\in\mathbb{Z}}$ is positive semi-definite. By Lemma 6.3 above, we have

$$|a_n| \leqq a_0 \quad \text{for all} \quad n \in \mathbb{Z}. \tag{6.26}$$

Hence it is obvious that $|a_n| \leqq \text{const.}\ |n|^N$ for some $N \in \mathbb{N} \cup \{0\}$. It follows from Theorem 4.8 (p. 97) that the a_n's are the Fourier coefficients of some 2π-periodic distribution T; i.e.

$$a_n = \frac{1}{\sqrt{2\pi}} T(e^{-inx}). \tag{6.27}$$

Let φ be any element of $\mathfrak{D}_{2\pi}(\mathbb{R})$ and φ_k's its Fourier coefficients. By a simple computation, we have

$$T\varphi = T\left(\frac{1}{\sqrt{2\pi}} \sum \varphi_n e^{inx}\right) = \frac{1}{\sqrt{2\pi}} \sum \varphi_n T(e^{inx}) = \sum \bar{a_n}\varphi_n. \tag{6.28}$$

Since φ is 2π-periodic and smooth, the series summing its Fourier coefficients is absolutely convergent.[11] Taking account of (6.26), we observe that the right-hand side of (6.28) converges.

We now proceed to show that T is positive.

Let $q_N(x)$ be any trigonometric polynomial of order N:

$$q_N(x) = \sum_{n=-N}^{N} \phi_n e^{inx}. \tag{6.29}$$

If we adopt

$$|q_N(x)|^2 = \sum_{n,k=-N}^{N} \phi_n \bar{\phi}_k e^{i(n-k)x}$$

as $\varphi \in \mathfrak{D}_{2\pi}(\mathbb{R})$, (6.28) implies

$$T(|q_N(x)|^2) = \sqrt{2\pi} \sum a_{k-n} \phi_n \bar{\phi}_k \geqq 0. \tag{6.30}$$

Any $q(x) \in \mathfrak{D}_{2\pi}(\mathbb{R})$ can be approximated by a sequence of trigonometric polynomials in $\mathfrak{C}^\infty$-topology. Passing to the limit, (6.30) implies

$$T(|q(x)|^2) \geqq 0 \quad \text{for any} \quad q(x) \in \mathfrak{D}_{2\pi}(\mathbb{R}). \tag{6.31}$$

[11] See Theorem 2.5 (p. 37) and Remark 2.3 (p. 39).

Let $p(x) \in \mathfrak{D}_{2\pi}(\mathbb{R})$ be nonnegative (real-valued) and $q(x) = \sqrt{p(x)}$. Then $q(x) \in \mathfrak{D}_{2\pi}(\mathbb{R})$ and

$$T(p(x)) \geqq 0$$

by (6.31). Consequently, the distribution T is positive and hence it is a positive measure;[12] i.e.

$$T(\varphi) = \int_{\mathbb{T}} \varphi(x) d\mu \quad \text{for all} \quad \varphi \in \mathfrak{C}^{\infty}(\mathbb{T}, \mathbb{C})$$

for some $\mu \in \mathfrak{M}_{+}(\mathbb{T})$. So we must have the desired result:

$$a_n = \frac{1}{\sqrt{2\pi}} T(e^{-inx}) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-inx} d\mu.$$

6.5 Fourier Transforms of Measures

Is any similar representation possible for a positive definite function defined on $\mathbb{R}$ rather than $\mathbb{T}$? The well-known theorem due to S. Bochner provides a positive answer to this question. There are several approaches to its proof. We first discuss the approach based upon the summability method, and then present an alternative proof from the viewpoint of distribution theory. This section is devoted to a few preparatory results for Bochner's theorem.

Definition 6.3 For a measure $\mu \in \mathfrak{M}(\mathbb{R})$,

$$\hat{\mu}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\xi x} d\mu(x)$$

is called the **Fourier transform** of μ.

Another terminology, the **Fourier–Stieltjes transform** is also used in order to distinguish the Fourier transform of a measure from that of a function. It is easy to show that $\hat{\mu}(\xi)$ is uniformly continuous. The Fourier transform of a measure $d\mu_f = f dx$ defined by a function $f \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ is given by

$$\hat{\mu}_f(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\xi x} f(x) dx.$$

It is nothing other than the usual Fourier transform of f.

[12] See Appendix C, Sect. C.3, or Schwartz [12], Chap. 1, §4, Théorème 5. $\mathfrak{M}_{+}(\mathbb{T})$ denotes the set of all the positive Radon measures on $\mathbb{T}$.

Definition 6.4 The **convolution** of a measure $\mu \in \mathfrak{M}(\mathbb{R})$ and a function $\varphi \in \mathfrak{C}_\infty(\mathbb{R}, \mathbb{C})$ is defined by the integration[13]

$$(\mu * \varphi)(x) = \int_{-\infty}^{\infty} \varphi(x-y)d\mu(y).$$

It is clear that

$$\|\mu * \varphi\|_\infty \leqq \|\mu\|_{\mathfrak{M}(\mathbb{R})} \cdot \|\varphi\|_\infty,$$

where $\|\mu\|_{\mathfrak{M}(\mathbb{R})}$ is the norm of μ defined by its total variation $|\mu|$. The convolution operation can also be defined for a pair of a measure $\mu \in \mathfrak{M}(\mathbb{R})$ and a μ-integrable functions $\varphi \in \mathfrak{L}^1_\mu(\mathbb{R}, \mathbb{C})$ in a similar manner.

The following result corresponds to Theorem 6.1.

Theorem 6.5 *For $\mu \in \mathfrak{M}(\mathbb{R})$ and $f \in \mathfrak{A}$ (Wiener algebra),*

$$\int_{-\infty}^{\infty} f(x)d\mu = \int_{-\infty}^{\infty} \hat{f}(\xi)\hat{\mu}(-\xi)d\xi.$$

Proof By the inversion formula (Theorem 5.9, p. 118), we have

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\xi)e^{i\xi x}d\xi.$$

Applying Fubini's theorem, we can deduce the desired result:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)d\mu &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{f}(\xi)e^{i\xi x}d\xi d\mu(x) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{\int_{-\infty}^{\infty} e^{i\xi x}d\mu(x)\right\} \hat{f}(\xi)d\xi \\ &= \int_{-\infty}^{\infty} \hat{f}(\xi)\hat{\mu}(-\xi)d\xi.\end{aligned}$$

□

The result corresponding to Corollary 6.1 immediately follows.[14]

Corollary 6.3 *Let $\mu \in \mathfrak{M}(\mathbb{R})$. If $\hat{\mu}(\xi) = 0$ for every $\xi \in \mathbb{R}$, then $\mu = 0$.*

---

[13] See Sect. 6.7 in this chapter for the convolution of two measures.

[14] "$\hat{\mu}(\xi) = 0$ for a.e. ξ" is sufficient to get $\mu = 0$. However, $\hat{\mu}(\xi)$ is eventually equal to 0 for every ξ, since we know that $\hat{\mu}(\xi)$ is continuous.

Theorem 6.6 *Let $\varphi \in \mathfrak{C}(\mathbb{R}, \mathbb{C})$. Define a function Φ_λ $(\lambda > 0)$ by*

$$\Phi_\lambda(x) = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right) \varphi(\xi) e^{i\xi x} d\xi.$$

Then the following two statements are equivalent:

(i) *There exists some $\mu \in \mathfrak{M}(\mathbb{R})$ such that $\varphi = \hat{\mu}$.*
(ii) *$\Phi_\lambda \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$ for all $\lambda > 0$, and $\|\Phi_\lambda\|_1 = O(1)$ as $\lambda \to \infty$.*

Proof (i)⇒(ii): Writing $\varphi = \hat{\mu}$ ($\mu \in \mathfrak{M}(\mathbb{R})$), we compute the convolution of μ and the Fejér kernel K_λ:

$$\begin{aligned}
(\mu * K_\lambda)(x) &= \int_{-\infty}^{\infty} K_\lambda(x-y) d\mu(y) \\
&= \int_{-\infty}^{\infty} \frac{\lambda}{2\pi} \int_{-1}^{1} (1-|\theta|) e^{i\theta(x-y)\lambda} d\theta \cdot d\mu(y) \quad \text{(by Lemma 5.2, p. 113)} \\
&= \frac{\lambda}{2\pi} \int_{-1}^{1} \underbrace{\left\{ \int_{-\infty}^{\infty} e^{-i\theta\lambda y} d\mu(y) \right\}}_{\sqrt{2\pi}\hat{\mu}(\theta\lambda)} (1-|\theta|) e^{i\theta\lambda x} d\theta \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right) \hat{\mu}(\xi) e^{i\xi x} d\xi \quad \text{(changing variables: } \theta\lambda = \xi) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right) \varphi(\xi) e^{i\xi x} d\xi \quad \text{(by (i))} \\
&= \Phi_\lambda(x).
\end{aligned}$$

Hence (ii) follows from

$$\begin{aligned}
\|\Phi_\lambda\|_1 = \int_{-\infty}^{\infty} |\Phi_\lambda(x)| dx &= \int_{-\infty}^{\infty} |(\mu * K_\lambda)(x)| dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |K_\lambda(x-y)| d|\mu|(y) dx \\
&\leqq \|K_\lambda\|_1 \cdot \|\mu\|_{\mathfrak{M}(\mathbb{R})}.
\end{aligned}$$

(ii)⇒(i): Define a measure μ_λ by $\Phi_\lambda(x)dx$. Since $\{\Phi_\lambda\}_{\lambda>0}$ is $\mathfrak{L}^1$-bounded, $\mu_\lambda = \Phi_\lambda(x)dx$ has a limiting point μ (in w^*-topology) as $\lambda \to \infty$ thanks to Alaoglu's theorem.[15] We can show that $\varphi = \hat{\mu}$ as follows.

[15] cf. Lax [6] pp. 120–121, Maruyama [8] pp. 354–355.

We first show that

$$\int_{-\infty}^{\infty} \varphi(-\xi) g(\xi) d\xi = \int_{-\infty}^{\infty} \hat{\mu}(-\xi) g(\xi) d\xi \tag{6.32}$$

for any $g \in \mathfrak{S}$ (space of rapidly decreasing functions). Let $G \in \mathfrak{S}$ be the inverse Fourier transform of g; i.e. $g = \hat{G}$. Then it follows that

$$\begin{aligned}
\int_{-\infty}^{\infty} g(\xi)\varphi(-\xi) d\xi &= \lim_{\lambda \to \infty} \int_{-\lambda}^{\lambda} g(\xi)\varphi(-\xi)\left(1 - \frac{|\xi|}{\lambda}\right) d\xi \\
&= \lim_{\lambda \to \infty} \int_{-\lambda}^{\lambda} \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(x) e^{-i\xi x} dx \right\} \varphi(-\xi)\left(1 - \frac{|\xi|}{\lambda}\right) d\xi \\
&= \lim_{\lambda \to \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(x) \left\{ \int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right) \varphi(-\xi) e^{-i\xi x} d\xi \right\} dx \\
&= \lim_{\lambda \to \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(x) \left\{ \int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right) \varphi(\xi) e^{i\xi x} d\xi \right\} dx \\
&\qquad \text{(changing the variable } -\xi \ to \ \xi) \\
&= \lim_{\lambda \to \infty} \int_{-\infty}^{\infty} G(x) \Phi_\lambda(x) dx \\
&= \lim_{\lambda \to \infty} \int_{-\infty}^{\infty} G(x) d\mu_\lambda \\
&= \int_{-\infty}^{\infty} G(x) d\mu \qquad (\mu_\lambda \xrightarrow{w^*} \mu) \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\xi) e^{i\xi x} d\xi \cdot d\mu(x) \\
&= \int_{-\infty}^{\infty} g(\xi) \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\xi x} d\mu(x) \right\} d\xi \\
&= \int_{-\infty}^{\infty} g(\xi) \hat{\mu}(-\xi) d\xi.
\end{aligned}$$

This proves (6.32).

Since $\mathfrak{S}$ is dense in $\mathfrak{C}_\infty$, (6.32) tells us that

$$\int_{-\infty}^{\infty} g(\xi)\{\varphi(-\xi) - \hat{\mu}(-\xi)\} d\xi = 0$$

for any $g \in \mathfrak{C}_\infty$. Therefore $\varphi(-\xi) = \hat{\mu}(-\xi)$, that is, $\varphi = \hat{\mu}$. □

Lemma 6.5 *Let $\mu_n, \mu \in \mathfrak{M}(\mathbb{R})$ and $\varphi \in \mathfrak{C}(\mathbb{R}, \mathbb{C})$. If*

$$w^*\text{-} \lim_{n\to\infty} \mu_n = \mu$$

and

$$\lim_{n\to\infty} \hat{\mu}_n(\xi) = \varphi(\xi) \quad for\ each \quad \xi \in \mathbb{R},$$

then $\hat{\mu} = \varphi$.

Proof As in the proof of the previous theorem, pick up any $g \in \mathfrak{S}$. Let $G \in \mathfrak{S}$ be the inverse Fourier transform of g; i.e. $g = \hat{G}$. Then[16]

$$\begin{aligned}
\int_{-\infty}^{\infty} g(\xi)\hat{\mu}(-\xi)d\xi &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} e^{i\xi x} d\mu(x) \right\} g(\xi) d\xi \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(\xi) e^{i\xi x} d\xi \right\} d\mu(x) \\
&= \int_{-\infty}^{\infty} G(x) d\mu(x) \\
&= \lim_{n\to\infty} \int_{-\infty}^{\infty} G(x) d\mu_n(x) \qquad (\text{by } \mu_n \overset{w^*}{\to} \mu) \\
&= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} g(\xi) e^{i\xi x} d\xi \right\} d\mu_n(x) \\
&= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} e^{i\xi x} d\mu_n \right\} g(\xi) d\xi \\
&= \lim_{n\to\infty} \int_{-\infty}^{\infty} g(\xi)\hat{\mu}_n(-\xi) d\xi \\
&= \int_{-\infty}^{\infty} g(\xi)\varphi(-\xi) d\xi \quad (\text{dominated convergence theorem}).
\end{aligned}$$

[16] $g(\xi)\hat{\mu}(-\xi) \to g(\xi)\varphi(-\xi)$. And $\{\mu_n\}$ is bounded (with respect to the operator norm), since $\mu_n \to \mu$ in w^*-topology. Furthermore, since $\|g\|_\infty < \infty$, there exists some $F \in \mathfrak{L}^1$ such that

$$|g(\xi)\hat{\mu}_n(-\xi)| \leqq |g(\xi)| \cdot C \leqq F(\xi).$$

Then apply the dominated convergence theorem.

Hence

$$\int_{-\infty}^{\infty} g(\xi)(\hat{\mu}(-\xi) - \varphi(-\xi))d\xi = 0$$

for any $g \in \mathfrak{S}$. $\hat{\mu} = \varphi$ follows from the continuity of $\hat{\mu}(-\xi) - \varphi(-\xi)$.[17] □

The next theorem corresponds to Theorem 6.3.

Theorem 6.7 *The following statements are equivalent for $\varphi \in \mathfrak{C}(\mathbb{R}, \mathbb{C})$:*

(i) *There exists a $\mu \in \mathfrak{M}(\mathbb{R})$ such that $\varphi = \hat{\mu}$.*
(ii) *There exists a constant C such that*

$$\left| \int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi \right| \leqq C \sup_{x \in \mathbb{R}} |f(x)|$$

for any $f \in \mathfrak{S}$.

Proof (i)⇒(ii): If we choose $C = \|\mu\|$, (ii) is obvious by Theorem 6.5.

(ii)⇒(i): Define a linear functional $\Lambda : \mathfrak{S} \to \mathbb{C}$ by

$$\Lambda : f \mapsto \int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi. \tag{6.33}$$

Then Λ is bounded and $\|\Lambda\| \leqq C$ by (ii). ($\mathfrak{S}$ is endowed with the uniform convergence norm.) Since $\mathfrak{S}$ is dense in $\mathfrak{C}_\infty$, Λ can be extended uniquely as an element of $(\mathfrak{C}_\infty)'$ preserving the norm. The extended functional is denoted by the same notation Λ. By the Riesz–Markov–Kakutani theorem, there exists a unique $\mu \in \mathfrak{M}(\mathbb{R})$ such that

$$\Lambda f = \int_{-\infty}^{\infty} f(x)d\mu, \quad f \in \mathfrak{C}_\infty \tag{6.34}$$

($\|\mu\| \leqq C$). By Theorem 6.5 again, we have

$$\int_{-\infty}^{\infty} f(x)d\mu = \int_{-\infty}^{\infty} \hat{f}(\xi)\hat{\mu}(-\xi)d\xi, \quad f \in \mathfrak{S}. \tag{6.35}$$

---

[17] Assume, on the contrary, that $\hat{\mu}(-\bar{\xi}) - \varphi(-\bar{\xi}) \neq 0$ (say, > 0) for some $\bar{\xi}$. Then there exists a neighborhood U of $\bar{\xi}$ such that $\hat{\mu}(-\xi) - \varphi(-\xi) > 0$ on U. There exists some $g \in \mathfrak{S}$ such that $\operatorname{supp} g \subset U$ and $g(x) > 0$ in the interior of its support. Then clearly

$$\int_{-\infty}^{\infty} g(\xi)(\hat{\mu}(-\xi) - \varphi(-\xi))d\xi > 0.$$

Contradiction.

Combining (6.33) and (6.35), we obtain

$$\int_{-\infty}^{\infty} \hat{f}(\xi)\{\varphi(-\xi) - \hat{\mu}(-\xi)\}d\xi = 0, \quad f \in \mathfrak{S}. \tag{6.36}$$

Taking account of $\mathfrak{S} = \{\hat{f} | f \in \mathfrak{S}\}$ and the continuity of $\varphi(-\xi) - \hat{\mu}(-\xi)$, we conclude that $\varphi = \hat{\mu}$ by (6.36). □

We should note that even if we change the function space $\mathfrak{S}$ appearing in (ii) of Theorem 6.7 to $\mathfrak{S}_c$ or $\mathfrak{C}_0^\infty$, the equivalence of (i) and (ii) is still valid. (See 1°, 2° on p. 67 and Remark 4.1 on p. 71).

Theorem 6.8 *The following two statements are equivalent for $\varphi \in \mathfrak{C}(\mathbb{R}, \mathbb{C})$:*

(i) *There exists some $\mu \in \mathfrak{M}(\mathbb{R})$ such that $\varphi = \hat{\mu}$.*
(ii) *There exists some constant C such that the numerical sequence $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$, for any $\lambda > 0$, is the sequence of the Fourier coefficients of some $\nu_\lambda \in \mathfrak{M}(\mathbb{T})$ with $\|\nu_\lambda\| \leqq C$ (independent of λ).*

Proof (i)⇒(ii): Assume (i). Let E be a measurable set in $\mathbb{T}$ and define $E_n = E + 2\pi n$ $(n \in \mathbb{Z})$, $\tilde{E} = \cup_n E_n$. Then we also define a measure ν on $\mathbb{T}$ by $\nu(E) = \mu(\tilde{E})$.[18] ν satisfies

$$\varphi(n) = \hat{\mu}(n) = \hat{\nu}(n), \quad \|\nu\| \leqq \|\mu\|.$$

If we define a measure μ_λ $(\lambda > 0)$ on $\mathbb{R}$ by

$$\int_{-\infty}^{\infty} f(x)d\mu_\lambda = \int_{-\infty}^{\infty} f(\lambda x)d\mu, \quad f \in \mathfrak{C}_\infty,$$

then $\|\mu_\lambda\| = \|\mu\|$ and

$$\hat{\mu}_\lambda(\xi) = \hat{\mu}(\lambda\xi). \tag{6.37}$$

In fact, (6.37) follows from

$$\hat{\mu}_\lambda(\xi) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\xi x}d\mu_\lambda(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\xi\cdot\lambda x}d\mu(x) = \hat{\mu}(\lambda\xi).$$

[18]For a 2π-periodic function f, we have

$$\int_{-\infty}^{\infty} f(x)d\mu = \int_{-\pi}^{\pi} f(x)d\nu.$$

Hence

$$\varphi(\lambda n) = \hat{\mu}(\lambda n) = \hat{\mu}_\lambda(n).$$

Let ν_λ be the measure on $\mathbb{T}$ which corresponds to μ_λ. Since $\|\nu_\lambda\| \leqq \|\mu_\lambda\| = \|\mu\|$ for every $\lambda > 0$, it is enough to define $C = \|\mu\|$.

(ii)⇒(i): Let f be an element of $\mathfrak{S}_c = \{f \in \mathfrak{S} \mid \text{supp } \hat{f} \text{ is compact}\}$. Since $\hat{f}(\xi)\varphi(-\xi)$ is an element of $\mathfrak{C}_0$, the integration

$$\int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi$$

can be approximated by Riemann sum. That is, for any $\varepsilon > 0$, dividing some interval $[-A, A] \supset$supp $\hat{f}$ into sufficiently small intervals, the length of which is less than λ, we obtain

$$\left|\int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi\right| < \left|\lambda \sum_n \hat{f}(\lambda n)\varphi(-\lambda n)\right| + \varepsilon. \tag{6.38}$$

Define a function $\psi_\lambda(x) \ : \ \mathbb{T} \to \mathbb{C}$ by

$$\psi_\lambda(x) = \sum_{m=-\infty}^{\infty} f\left(\frac{x + 2\pi m}{\lambda}\right). \tag{6.39}$$

Since $f \in \mathfrak{S}_c(\subset \mathfrak{S})$, the series $\psi_\lambda(x)$ is well-defined and continuous. The Fourier coefficients of $\psi_\lambda(x)$ are computed as (cf. Fig. 6.1)

$$\begin{aligned}
\hat{\psi}_\lambda(n) &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \psi_\lambda(x)e^{-inx}dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \sum_{m=-\infty}^{\infty} f\left(\frac{x + 2\pi m}{\lambda}\right) e^{-inx}dx \\
&= \frac{\lambda}{\sqrt{2\pi}} \int_{-\frac{\pi}{\lambda}}^{\frac{\pi}{\lambda}} \sum_{m=-\infty}^{\infty} f\left(x + \frac{2\pi m}{\lambda}\right) e^{-i\lambda nx}dx
\end{aligned}$$

(changing the variable x/λ by x)

$$= \frac{\lambda}{\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} \int_{-\frac{\pi}{\lambda}}^{\frac{\pi}{\lambda}} f\left(x + \frac{2\pi m}{\lambda}\right) e^{-i\lambda nx}dx \tag{6.40}$$

(by uniform convergence of $\sum \cdots$).

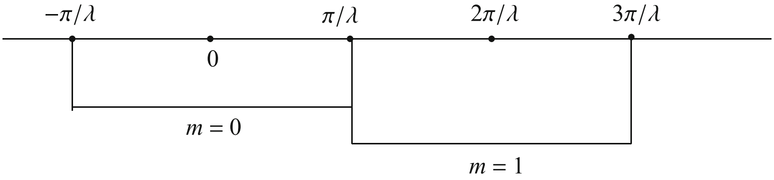


Fig. 6.1 Calculation of (6.40)

The integration of (6.40) is evaluated as

$$\int_{-\frac{\pi}{\lambda}}^{\frac{\pi}{\lambda}} f\left(x + \frac{2\pi m}{\lambda}\right) e^{-i\lambda n x} dx = \int_{-\frac{\pi}{\lambda}+\frac{2\pi m}{\lambda}}^{\frac{\pi}{\lambda}+\frac{2\pi m}{\lambda}} f(u)e^{-i\lambda n u} \cdot e^{2\pi i n m} du$$

$$\text{(changing variables: } u = x + (2\pi m)/\lambda)$$

$$= \int_{-\frac{\pi}{\lambda}+\frac{2\pi m}{\lambda}}^{\frac{\pi}{\lambda}+\frac{2\pi m}{\lambda}} f(u)e^{-i\lambda n u} du \qquad (6.41)$$

$$(e^{2\pi i n m} = 1).$$

Substituting (6.41) into (6.40),[19] we obtain

$$\hat{\psi}_\lambda(n) = \frac{\lambda}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u)e^{-i\lambda n u} du = \lambda \hat{f}(\lambda n). \qquad (6.42)$$

Since $f \in \mathfrak{S}$, we can show that

$$\|\psi_\lambda\|_\infty \leqq \|f\|_\infty + \varepsilon \qquad (6.43)$$

for sufficiently small λ. In fact, there exists some M such that

$$\left| f\left(\frac{x + 2\pi m}{\lambda}\right)\right| \leqq \frac{M}{\left(\frac{x + 2\pi m}{\lambda}\right)^2} \leqq \frac{M}{\pi^2} \cdot \frac{1}{\left(\frac{m}{\lambda}\right)^2} \qquad (6.44)$$

for all $x \in \mathbb{T}$ because f is rapidly decreasing.[20] Consequently, we have

[19] Summing up the upper and lower ends of the integration (6.41) for all $m \in \mathbb{Z}$, we get the integration over $(-\infty, \infty)$. See Fig. 6.1.

[20] $\left(\frac{x + 2\pi m}{\lambda}\right)^2 \geqq \left(\frac{\pi - 2\pi|m|}{\lambda}\right)^2 = \frac{\pi^2}{\lambda^2}(1 - 2|m|)^2 \geqq \frac{\pi^2}{\lambda^2} m^2$, where $|m| \geqq 1$.

$$\begin{aligned}|\psi_\lambda(x)| &\leqq \sum_{m=-\infty}^{\infty}\left|f\left(\frac{x+2\pi m}{\lambda}\right)\right| \\ &\leqq \underset{(m=0)}{\left|f\left(\frac{x}{\lambda}\right)\right|} + \sum_{m\neq 0}\frac{M}{\pi^2}\cdot\frac{\lambda^2}{m^2} \qquad (6.45)\\ &\leqq A \qquad \text{(constant } A \text{ is independent of } x).\end{aligned}$$

Hence we obtain, λ_0 being fixed,

$$\sup_{x\in\mathbb{T}}\sum_{|m|>N_0}\left|f\left(\frac{x+2\pi m}{\lambda_0}\right)\right| < \varepsilon \tag{6.46}$$

for sufficiently large N_0. The evaluation (6.46) holds good for $\lambda \leqq \lambda_0$.

For $|m| \leqq N_0$,

$$\sup_{x\in\mathbb{T}}\left|f\left(\frac{x+2\pi m}{\lambda}\right)\right| \leqq \frac{1}{2N_0+1}\|f\|_\infty \tag{6.47}$$

for sufficiently small $\lambda \leqq \lambda_0$. Choosing $\lambda \leqq \lambda_0$ very small so that both (6.45) and (6.47) hold good, we obtain, by (6.45), (6.46) and (6.47),

$$\|\psi_\lambda\|_\infty \leqq \|f\|_\infty + \varepsilon. \tag{6.48}$$

The evaluation (6.43) immediately follows from this.

By (ii), we can choose $\nu_\lambda \in \mathfrak{M}(\mathbb{T})$ so that

$$\varphi(\lambda n) = \hat{\nu}_\lambda(n) \quad \text{and} \quad \|\nu_\lambda\| \leqq C$$

hold good. The relations $|\hat{\psi}_\lambda(n)| \leqq \sqrt{2\pi}\|\psi_\lambda\|_\infty$, $|\hat{\nu}_\lambda(n)| \leqq (1/\sqrt{2\pi})C$ and (6.42) jointly imply

$$\left|\lambda\sum_{n=-\infty}^{\infty}\hat{f}(\lambda n)\varphi(-\lambda n)\right| = \left|\sum_{n=-\infty}^{\infty}\hat{\psi}_\lambda(n)\hat{\nu}_\lambda(-n)\right| \leqq C\|\psi_\lambda\|_\infty. \tag{6.49}$$

Furthermore, we obtain, by (6.38), (6.43) and (6.49), that

$$\begin{aligned}&\left|\int_{-\infty}^{\infty}\hat{f}(\xi)\varphi(-\xi)d\xi\right| \leqq C\|\psi_\lambda\|_\infty + \varepsilon \\ &\leqq C(\|f\|_\infty + \varepsilon) + \varepsilon = C\|f\|_\infty + (C+1)\varepsilon,\end{aligned}$$

from which it follows that

$$\left| \int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi \right| \leqq C\|f\|_{\infty},$$

since $\varepsilon > 0$ is arbitrary. This holds good for all $f \in \mathfrak{S}_c$. Hence (i) follows by Theorem 6.7 and the remark to it (pp. 141–142). □

6.6 Bochner's Theorem

We have completed all the preparations for proving Bochner's theorem.

Theorem 6.9 *For $\varphi \in \mathfrak{C}^b(\mathbb{R}, \mathbb{C})$, the following two statements are equivalent:*

(i) *There exists some $\mu \in \mathfrak{M}_+(\mathbb{R})$ such that $\varphi = \hat{\mu}$.*
(ii) *For any nonnegative $f \in \mathfrak{A}$ (Wiener algebra),*

$$\int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi \geqq 0. \tag{6.50}$$

Proof (i)⇒(ii): If we assume (i), we have

$$\int_{-\infty}^{\infty} f(x)d\mu = \int_{-\infty}^{\infty} \hat{f}(\xi)\hat{\mu}(-\xi)d\xi \overset{\text{(i)}}{=} \int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi \geqq 0$$

for any nonnegative $f \in \mathfrak{A}$ by Theorem 6.5.

(ii)⇒(i): If there exists some $\mu \in \mathfrak{M}(\mathbb{R})$ such that $\varphi = \hat{\mu}$, (ii) implies

$$\int_{-\infty}^{\infty} f(x)d\mu = \int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi \geqq 0$$

for any nonnegative $f \in \mathfrak{S}$. (Keep in mind that $\mathfrak{S} \subset \mathfrak{A}$.) Since $\mathfrak{S}$ is dense in $\mathfrak{C}_\infty$, μ must be a positive measure.

So it remains only to show the existence of $\mu \in \mathfrak{M}(\mathbb{R})$ which satisfies $\varphi = \hat{\mu}$.

We should recall a couple of properties of the Fejér kernel $K_\lambda(x)$:

1°

$$\frac{1}{\lambda}K_\lambda(x) = K(\lambda x) = \frac{1}{2\pi}\left(\frac{\sin \lambda x/2}{\lambda x/2}\right)^2 \tag{6.51}$$

is nonnegative and converges to $1/2\pi$ as $\lambda \to 0$.[21] This convergence is uniform on any compact set in $\mathbb{R}$.

[21] For the Fejér kernel, see Chap. 5, Sect. 5.4 (pp. 111–115).

2°

$$\widehat{K(\lambda x)}(\xi) = \frac{1}{\lambda}\mathrm{Max}\left\{\frac{1}{\sqrt{2\pi}}\left(1 - \frac{|\xi|}{\lambda}\right), 0\right\}. \tag{6.52}$$

It can be proved as follows. Since $K(\lambda x) = K_\lambda(x)/\lambda$,

$$\widehat{K(\lambda x)}(\xi) = \frac{1}{\lambda}\hat{K}_\lambda(\xi). \tag{6.53}$$

As we have already verified,[22]

$$\begin{aligned}\hat{K}_\lambda(\xi) &= \begin{cases} \frac{1}{\sqrt{2\pi}}\left(1 - \frac{|\xi|}{\lambda}\right) & \text{on} \quad [-\lambda, \lambda], \\ 0 & \text{on} \quad [-\lambda, \lambda]^c \end{cases} \\ &= \mathrm{Max}\left\{\frac{1}{\sqrt{2\pi}}\left(1 - \frac{|\xi|}{\lambda}\right), 0\right\}. \end{aligned} \tag{6.54}$$

(cf. Fig. 6.2.) The relation (6.52) follows from (6.53) and (6.54).

We proceed to show that

$$\lim_{\lambda \to 0} \frac{\sqrt{2\pi}}{\lambda} \int_{-\infty}^{\infty} \hat{K}_\lambda(\xi)\varphi(-\xi)d\xi = \varphi(0). \tag{6.55}$$

Since

$$\frac{1}{\lambda}\int_{-\infty}^{\infty} \hat{K}_\lambda(\xi)d\xi = \int_{-\infty}^{\infty} \widehat{K(\lambda x)}(\xi)d\xi = \frac{1}{\sqrt{2\pi}} \tag{6.56}$$

by (6.54), it follows that

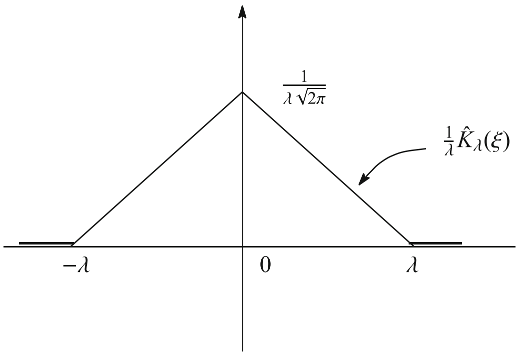
$$\begin{aligned}\left|\int_{-\infty}^{\infty} \frac{\sqrt{2\pi}}{\lambda}\hat{K}_\lambda(\xi)\varphi(-\xi)d\xi - \varphi(0)\right| &= \left|\int_{-\infty}^{\infty} \frac{\sqrt{2\pi}}{\lambda}\hat{K}_\lambda(\xi)\{\varphi(-\xi) - \varphi(0)\}d\xi\right| \\ &\leqq \underbrace{\frac{\sqrt{2\pi}}{\lambda}\|\hat{K}_\lambda\|_1}_{=1} \cdot \|\varphi(-\xi) - \varphi(0)\|_\infty = \|\varphi(-\xi) - \varphi(0)\|_\infty. \end{aligned} \tag{6.57}$$

By the continuity of φ,

$$\sup_{\xi \in [-\lambda, \lambda]} |\varphi(-\xi) - \varphi(0)| \to 0$$

[22] cf. Chap. 5, Sect. 5.5, p. 119.

Fig. 6.2 Graph of $\frac{1}{\lambda}\hat{K}_\lambda(\xi)$



as $\lambda \to 0$. That is, the most right-hand side of (6.57) converges to 0. This proves (6.55).

Finally, taking account of Theorem 6.7 and the remark to it, we wish to prove the existence of a constant C such that

$$\left|\int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi\right| \leqq C\|f\|_\infty \tag{6.58}$$

for any $f \in \mathfrak{C}_0^\infty$. Without loss of generality, we may assume that f is real-valued. (If not, we have only to divide f into real and imaginary parts.) If we choose sufficiently small $\lambda > 0$, for each $\varepsilon > 0$, we have[23]

$$2\pi(\varepsilon + \|f\|_\infty)K(\lambda x) - f(x) \geqq 0 \qquad \text{for all} \quad x \in \mathbb{R}. \tag{6.59}$$

The left-hand side of (6.59) is contained in $\mathfrak{A}$ and nonnegative. So it follows by (ii) that

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi &\leqq (\varepsilon + \|f\|_\infty)\int_{-\infty}^{\infty} \widehat{K(\lambda x)}(\xi)\varphi(-\xi)d\xi \\ &= \frac{1}{\lambda}(\varepsilon + \|f\|_\infty)\int_{-\infty}^{\infty} \hat{K}_\lambda(\xi)\varphi(-\xi)d\xi \\ &= \frac{1}{\sqrt{2\pi}}\frac{\sqrt{2\pi}}{\lambda}(\varepsilon + \|f\|_\infty)\int_{-\infty}^{\infty} \hat{K}_\lambda(\xi)\varphi(-\xi)d\xi \\ &\longrightarrow \frac{\varepsilon + \|f\|_\infty}{\sqrt{2\pi}}\varphi(0) \quad \text{as} \quad \lambda \to 0 \qquad (\text{by } (6.55)).\end{aligned} \tag{6.60}$$

[23] We have already seen that $K(\lambda x) \to 1/2\pi$ (uniformly on any compact set) as $\lambda \to 0$ (cf. p. 146). We can choose λ (independent of $x \in \text{supp } f$) so that (6.59) holds good on supp f. The inequality (6.59) is satisfied, of course, outside supp f.

Hence

$$\int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi \leqq \sqrt{2\pi}(\varepsilon + \|f\|_\infty)\varphi(0). \tag{6.61}$$

By repeating the same arguments for $-\varphi$, we can prove the converse inequality. Thus (6.58) is satisfied for $C = \sqrt{2\pi}\varphi(0)$. □

Combining Theorem 6.8 and Theorem 6.9, we obtain the next result.

Theorem 6.10 *For $\varphi \in \mathfrak{C}^b(\mathbb{R}, \mathbb{C})$, the following two statements are equivalent:*

(i) *There exists some $\mu \in \mathfrak{M}_+(\mathbb{R})$ such that $\varphi = \hat{\mu}$.*
(ii) *For any $\lambda > 0$, there exists some $\nu_\lambda \in \mathfrak{M}_+(\mathbb{T})$ such that the numerical sequence $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$ is the sequence of Fourier coefficients of ν_λ.*

Proof (i)⇒(ii) is proved in the same way as in Theorem 6.8. In the course of the proof, the measure μ_λ is positive. Hence the corresponding ν_λ is also positive.

(ii)⇒(i): Since $\nu_\lambda \in \mathfrak{M}_+(\mathbb{T})$ satisfies

$$\varphi(\lambda n) = \hat{\nu}_\lambda(n), \quad n \in \mathbb{Z},$$

we obtain

$$\varphi(0) = \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} d\nu_\lambda = \frac{1}{\sqrt{2\pi}}\|\nu_\lambda\|.$$

By Theorem 6.8 (for $C = \sqrt{2\pi}\varphi(0)$), there exists some $\mu \in \mathfrak{M}(\mathbb{R})$ which satisfies $\varphi = \hat{\mu}$.

We show that (6.50) is valid for any nonnegative $f \in \mathfrak{A}$. (It is required in order to apply Theorem 6.9.) Without loss of generality, we may assume that supp $\hat{f}$ is compact.[24] Under this assumption, the integration appearing in the left-hand side of (6.50) can be approximated by Riemann sum so that (6.38) in the previous section (p. 143) is satisfied. Furthermore, if we define a function $\psi_\lambda : \mathbb{T} \to \mathbb{C}$ as in (6.39) in the previous section, we have

$$\hat{\psi}_\lambda(n) = \lambda\hat{f}(\lambda n), \quad n \in \mathbb{Z} \tag{6.62}$$

[24] If $f \in \mathfrak{A}$ and $f \geqq 0$, there exists a sequence $\{g_n\}$ in $\mathfrak{C}_0^\infty$ which converges to $\hat{f}$ in $\mathfrak{L}^1$. Let f_n be the inverse Fourier transform of g_n; i.e. $\hat{f}_n = g_n$. Then $\{f_n\}$ is a sequence in $\mathfrak{S}_c$. If

$$\int_{-\infty}^{\infty} \hat{f}_n(\xi)\varphi(-\xi)d\xi = \int_{-\infty}^{\infty} g_n(\xi)\varphi(-\xi)d\xi \geqq 0$$

for each n, we obtain, by passing to the limit,

$$\int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi \geqq 0.$$

as we have already seen (cf. (6.42)). And, by (6.48),

$$\|\psi_\lambda\|_\infty \leqq \|f\|_\infty + \varepsilon \tag{6.63}$$

for sufficiently small $\lambda > 0$. Consequently, $\sum \hat{\psi}_\lambda(n)\hat{\nu}_\lambda(-n)$ converges and, by Theorem 6.1,

$$\lambda \sum \hat{f}(\lambda n)\varphi(-\lambda n) = \sum \hat{\psi}_\lambda(n)\hat{\nu}_\lambda(-n) = \int_{-\pi}^{\pi} \psi_\lambda(x)d\nu_\lambda \geqq 0. \tag{6.64}$$

By Riemann approximation, we obtain

$$\int_{-\infty}^{\infty} \hat{f}(\xi)\varphi(-\xi)d\xi \geqq \lambda \sum \hat{f}(\lambda n)\varphi(-\lambda n) - \varepsilon. \tag{6.65}$$

We get (6.50) from (6.64) and (6.65) by passing to the limit. □

Theorem 6.11 (Bochner)[25] *For $\varphi : \mathbb{R} \to \mathbb{C}$, the following two statements are equivalent:*

(i) *φ is positive semi-definite and continuous at 0.*[26]
(ii) *There exists some $\mu \in \mathfrak{M}_+(\mathbb{R})$ such that $\varphi = \hat{\mu}$.*

Proof (ii)$\Rightarrow$ (i): For any $\xi_1, \xi_2, \cdots, \xi_p \in \mathbb{R}$, $z_1, z_2, \cdots, z_p \in \mathbb{C}$, we have

$$\begin{aligned}\sum_{j,k} \varphi(\xi_j - \xi_k) z_j \bar{z}_k &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sum_{j,k} e^{-i\xi_j x} z_j \cdot e^{i\xi_k x} \bar{z}_k d\mu(x) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left| \sum z_j e^{-i\xi_j x} \right|^2 d\mu(x) \\ &\geqq 0.\end{aligned}$$

Thus φ is positive semi-definite.

Its continuity at 0 is obvious.

(i)$\Rightarrow$(ii): If φ is positive semi-definite, the sequence $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$ is positive semi-definite for any $\lambda > 0$. Hence, by Theorem 6.4 due to Herglotz, there exists some $\nu_\lambda \in \mathfrak{M}_+(\mathbb{T})$ such that

$$\varphi(\lambda n) = \hat{\nu}_\lambda(n), \quad n \in \mathbb{Z}.$$

(ii) immediately follows by Theorem 6.10. □

---

[25] Bochner [1]. μ, which satisfies (ii), is unique by Corollary 6.3.

[26] By Lemma 6.4, a positive semi-definite function is uniformly continuous on $\mathbb{R}$ if it is continuous at 0.

Corresponding to Theorem 6.4, we can give an alternative proof of Bochner's theorem, based upon distribution theory.

Remark 6.5 Since $\mathfrak{S}(\mathbb{R})$ is a dense subspace of $\mathfrak{C}_\infty(\mathbb{R}, \mathbb{C})$ (space of continuous functions vanishing at infinity), any $\mu \in \mathfrak{M}$ (space of Radon measures on $\mathbb{R}$) can be regarded as a tempered distribution; i.e. $\mu \in \mathfrak{S}(\mathbb{R})'$. Here arises a question. Does the usual Fourier transform (Fourier–Stieltjes transform) of μ coincide with its Fourier transform in the sense of distribution? Let us evaluate the Fourier transform $\hat{\mu}(\theta)$ in the sense of distribution for $\theta \in \mathfrak{S}(\mathbb{R})$:

$$\begin{aligned}\hat{\mu}(\theta) = \mu(\hat{\theta}) &= \mu\left[\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \theta(x)e^{-i\xi x}dx\right] \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left[\int_{\mathbb{R}} \theta(x)e^{-i\xi x}dx\right]d\mu(\xi) \\ &= \int_{\mathbb{R}} \theta(x)\left[\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-i\xi x}d\mu(\xi)\right]dx \\ &= \int_{\mathbb{R}} \theta(x)\hat{\mu}(\xi)dx.\end{aligned}$$

($\hat{\mu}(\xi)$is the usual Fourier transform of μ.)

This equality holds good for any $\theta \in \mathfrak{S}(\mathbb{R})$. Consequently, we conclude that the two definitions are identical.

As in Theorem 6.11, (ii)⇒(i) is easy and well-known. So we have only to prove (i)⇒(ii).[27]

Assume (i). Then φ is bounded since

$$|\varphi(x)| \leqq \varphi(0) \quad \text{for all} \quad x \in \mathbb{R} \tag{6.66}$$

by Lemma 6.3. Hence φ defines a tempered distribution ($\in \mathfrak{S}(\mathbb{R}')$). We denote by $\check{\varphi} \in \mathfrak{S}(\mathbb{R})'$ the inverse Fourier transform (as a distribution). By Parseval's theorem, we have[28]

$$\varphi(s) = (\check{\varphi})\hat{}(s) = \check{\varphi}(\hat{s}), \quad \text{for} \quad s \in \mathfrak{S}(\mathbb{R}). \tag{6.67}$$

[27] The basic ideas are given in Lax [6], pp. 144–146. See also Maruyama [10].

[28] Here φ denotes the tempered distribution defined by the bounded function φ. Hence

$$\varphi(s) = \int_{\mathbb{R}} \varphi(x)s(x)dx, \quad s \in \mathfrak{S}(\mathbb{R}).$$

The positive semi-definiteness of φ implies (by Lemma 6.4)

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(x-y)\theta(x)\overline{\theta(y)}dxdy \geqq 0 \quad \text{for any} \quad \theta \in \mathfrak{D}(\mathbb{R}),$$

which can be rewritten as (changing variables: $z = x - y$)

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(z)\theta(x)\overline{\theta(x-z)}dxdz \geqq 0. \tag{6.68}$$

If we define the function $\Theta \in \mathfrak{D}(\mathbb{R})$ by

$$\Theta(z) = \int_{\mathbb{R}} \theta(x)\overline{\theta(x-z)}dx, \tag{6.69}$$

(6.68) can be rewritten as

$$\int_{\mathbb{R}} \varphi(z)\Theta(z)dz \geqq 0. \tag{6.66$'$}$$

By (6.69),

$$\begin{aligned}
\hat{\Theta}(w) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left[\int_{\mathbb{R}} \theta(x)\overline{\theta(x-z)}dx\right]e^{-iwz}dz \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left[\int_{\mathbb{R}} \theta(x)\overline{\theta(u)}dx\right]e^{-iwx}e^{iwu}du \\
&\qquad (\text{changing} \quad \text{variables}: \ u = x - z) \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \theta(x)e^{-iwx}dx \cdot \int_{\mathbb{R}} \bar{\theta}(u)e^{iwu}du \\
&= \hat{\theta}(w)\cdot\sqrt{2\pi}\bar{\hat{\theta}}(w) \\
&= \sqrt{2\pi}|\hat{\theta}(w)|^2.
\end{aligned}$$

It follows that

$$|\hat{\theta}(w)|^2 = \hat{\theta}(w)\bar{\hat{\theta}}(w) = \frac{1}{\sqrt{2\pi}}\hat{\Theta}(w). \tag{6.70}$$

(6.67), (6.66′) and (6.70) imply

$$\varphi(\Theta) \underset{(6.67)}{=} \check{\varphi}(\hat{\Theta}) \underset{(6.70)}{=} \check{\varphi}(\sqrt{2\pi}|\hat{\theta}|^2) \underset{(6.66')}{\geqq} 0. \tag{6.71}$$

The inequality (6.71) holds good for any $\theta \in \mathfrak{D}(\mathbb{R})$. Since $\mathfrak{D}(\mathbb{R})$ is dense in $\mathfrak{S}(\mathbb{R})$, (6.71) is also valid for any $\theta \in \mathfrak{S}(\mathbb{R})$.

Let $p(w) \geqq 0$ be an element of $\mathfrak{D}(\mathbb{R})$ (space of smooth functions with compact support). Then $\sqrt{p(w)}$ is also an element of $\mathfrak{D}(\mathbb{R}) \subset \mathfrak{S}(\mathbb{R})$. Using $\sqrt{p}$ as $|\hat{\theta}|$ in (6.71), we obtain[29]

$$\check{\varphi}(p) \geqq 0 \quad \text{for any} \quad 0 \leqq p \in \mathfrak{D}(\mathbb{R}). \tag{6.72}$$

The result (6.72) tells us that $\check{\varphi}$ is a positive distribution, and so it is a positive measure; i.e.

$$\check{\varphi}(\theta) = \int_{\mathbb{R}} \theta(x) d\mu, \quad \theta \in \mathfrak{D}(\mathbb{R}) \tag{6.73}$$

for some positive measure μ. (cf. Appendix C, pp. 391–392.)

Finally, we claim that $\mu(\mathbb{R}) < \infty$. Let g be an element of $\mathfrak{D}(\mathbb{R})$ which is nonnegative and satisfies

$$g(x) = 1 \quad \text{on} \quad [-1, 1]. \tag{6.74}$$

Define a function g_n by $g_n(x) = g(x/n)$. Let G and G_n be the inverse Fourier transforms of g and g_n, respectively. Taking account of $G_n(y) = nG(ny)$, we apply (6.67) to get

$$\check{\varphi}(g_n) = \int_{\mathbb{R}} \varphi(y) nG(ny) dy. \tag{6.75}$$

Since $\check{\varphi}$ is a positive distribution, the properties ($g \geqq 0$ and (6.74)) of g imply

$$\check{\varphi}(g_n) \geqq \int_{[-n,n]} d\mu.$$

Furthermore, it follows from (6.66) that

$$\int_{\mathbb{R}} \varphi(y) nG(ny) \leqq \varphi(0) \int_{\mathbb{R}} n|G(ny)| dy = \varphi(0) \int_{\mathbb{R}} |G(y)| dy. \tag{6.76}$$

The right-hand side of (6.76) is independent of n. If we write

$$C = \varphi(0) \int_{\mathbb{R}} |G(y)| dy,$$

[29] We note that there exists a unique function in $\mathfrak{S}(\mathbb{R})$, the Fourier transform of which is just $\sqrt{p}$.

we obtain

$$\int_{[-n,n]} d\mu \leqq C,$$

and so $\mu(\mathbb{R}) \leqq C$.

Looking at (6.73), we can conclude that φ is the Fourier transform of the distribution defined by μ; i.e.

$$\varphi(s) = \hat{\mu}(s) \quad \text{for} \quad s \in \mathfrak{D}(\mathbb{R}).$$

This equality holds good for all $s \in \mathfrak{S}(\mathbb{R})$.

6.7 Convolutions of Measures

Let G be either $\mathbb{T}$ or $\mathbb{R}$.[30]

As a final topic of this chapter, we now proceed to N. Wiener's theory to evaluate the discrete part of a Radon measure.

$\mathfrak{M}(G)$ denotes the space of Radon measures on G. Its subset $\mathfrak{M}_d(G)$ is defined by

$$\mathfrak{M}_d(G) = \left\{ \mu \in \mathfrak{M}(G) \;\middle|\; \mu = \sum_{n=1}^{\infty} \alpha_n \delta_{x_n},\ \sum_{n-1}^{\infty} |\alpha_n| < \infty,\ \alpha_n \in \mathbb{C} \right\}.$$

That is, $\mu \in \mathfrak{M}_d(G)$ is a complex measure with finite total variation concentrating at a countable set $\{x_n\}$, each element of which is assigned a mass α_n.

[30] I repeat again the exposition of the concept of the torus. A binary relation $\sim$ on $\mathbb{R}$ is defined by

$$x \sim y \Longleftrightarrow x - y \in 2\pi\mathbb{Z}.$$

$\sim$ is an equivalence relation, and any two numbers are equivalent if and only if the difference between them is a multiple of 2π. Any real number has its equivalent number in $[-\pi, \pi)$.

Given a function $f : [-\pi, \pi) \to \mathbb{C}$, we may not be able to define $f(x + y)$, since $x + y$ may not be in $[-\pi, \pi)$. However, there exists some $z \in [-\pi, \pi)$ which is equivalent to $x + y$. We understand $f(x + y)$ as $f(z)$. A similar convention applies to the Dirac measures on $[-\pi, \pi)$. That is, we understand δ_{x+y} as δ_z. More generally, we can consider $f(x_1 + x_2 + \cdots + x_n)$ or $\delta_{x_1+x_2+\cdots+x_n}$ in the same way.

When we consider a function on $\mathbb{R}$ rather than $[-\pi, \pi)$, such a convention is not necessary.

Exactly speaking, we consider a function defined on the quotient group $\mathbb{R}/2\pi\mathbb{Z}$ modulo $2\pi\mathbb{Z}$. This is called torus and denoted by $\mathbb{T}$. cf. Appendix A. We are indebted to Malliavin [7] Chap. III, §1 for the expositions in this section.

The convolution $\delta_{x_1} * \delta_{x_2}$ of a pair of Dirac measures δ_{x_1} and δ_{x_2} is defined by

$$\delta_{x_1} * \delta_{x_2} = \delta_{x_1+x_2}.$$

Generalizing this concept, we obtain the following definition.

Definition 6.5 Let μ and ν be elements of $\mathfrak{M}(G)$ of the form

$$\begin{aligned} \mu &= \sum_{n=1}^{\infty} \alpha_n \delta_{x_n}, \quad \sum_{n=1}^{\infty} |\alpha_n| < \infty, \\ \nu &= \sum_{n=1}^{\infty} \beta_n \delta_{y_n}, \quad \sum_{n=1}^{\infty} |\beta_n| < \infty. \end{aligned} \tag{6.77}$$

Then the **convolution** $\mu * \nu$ of μ and ν is defined by

$$\mu * \nu = \sum_{n,m} \alpha_n \beta_m \delta_{x_n+y_m}. \tag{6.78}$$

$\mu * \nu$ is, of course, an element of $\mathfrak{M}_d(G)$.

We list some elementary properties of the convolution, the verification of which is easy.

1° $\mu * \nu = \nu * \mu$.
2° $(\mu * \nu) * \theta = \mu * (\nu * \theta)$.
3° $(\mu + \nu) * \theta = \mu * \theta + \nu * \theta$.
4° $\lambda(\mu * \nu) = (\lambda\mu) * \nu = \mu * (\lambda\nu)$ for any $\lambda \in \mathbb{R}$.
5° $\|\mu * \nu\| \leqq \|\mu\| \cdot \|\nu\|$.

In summary, $\mathfrak{M}_d(G)$ is a commutative normed algebra[31] with respect to the convolution operation $*$.

Lemma 6.6 *Let θ be the convolution of μ and $\nu \in \mathfrak{M}_d(G)$. Then*

$$\int_G f(x) d\theta = \int_G \int_G f(x+y) d\mu(x) d\nu(y) \quad \textit{for all} \quad f \in \mathfrak{C}_\infty(G, \mathbb{C}).$$

$$(f \in \mathfrak{C}(G, \mathbb{C}) \textit{ in the case of } G = [-\pi, \pi)).$$

Proof If μ and ν are of the form (6.77), θ is given by

$$\theta = \sum_{n,m} \alpha_n \beta_m \delta_{x_n+y_m}.$$

[31] See Naimark [11], Lax [6] Chaps. 17–20 and Maruyama [8] Chap. 7 for normed algebras.

For $f \in \mathfrak{C}_\infty(G, \mathbb{C})$, we have

$$\int_G \int_G f(x+y)d\mu(x)d\nu(y) = \sum_{n,m} f(x_n + y_m)\alpha_n \beta_m = \int_G f(x)d\theta.$$

$\square$

We now extend the convolution to the entire space $\mathfrak{M}(G)$. Given a pair of μ and $\nu \in \mathfrak{M}(G)$, a linear functional $\Lambda_{\mu,\nu} \in \mathfrak{C}_\infty(G, \mathbb{C})'$ is defined by

$$\Lambda_{\mu,\nu}(f) = \int_G \int_G f(x+y)d\mu(x)d\nu(y), \quad f \in \mathfrak{C}_\infty(G, \mathbb{C}).$$

$\Lambda_{\mu,\nu} \in \mathfrak{C}_\infty(G, \mathbb{C})'$, since we have the evaluation

$$\begin{aligned}|\Lambda_{\mu,\nu}(f)| &\leqq \int_G \int_G |f(x+y)|d|\mu|(x)d|\nu|(y) \\ &\leqq \|f\|_\infty \int_G \int_G d|\mu|(x)d|\nu|(y) \\ &\leqq \|f\|_\infty \|\mu\| \|\nu\|.\end{aligned}$$

So there exists some unique $\theta \in \mathfrak{M}(G)$ such that

$$\Lambda_{\mu,\nu}(f) = \int_G f(x)d\theta, \quad f \in \mathfrak{C}_\infty(G, \mathbb{C})$$

by the Riesz–Markov–Kakutani theorem. We call this θ the convolution of μ and ν and denote it by $\theta = \mu * \nu$.

Similarly to the case of $\mathfrak{M}_d(G)$, we can show the following properties (i)–(v) of the convolution on $\mathfrak{M}(G)$. Furthermore, (vi) can also be verified in addition to them.

Theorem 6.12 *Let $\mu, \nu, \theta \in \mathfrak{M}(G)$. Then the following relations hold good:*

(i) $\mu * \nu = \nu * \mu$.
(ii) $(\mu * \nu) * \theta = \mu * (\nu * \theta)$.
(iii) $(\mu + \nu) * \theta = \mu * \theta + \nu * \theta$.
(iv) $\lambda(\mu * \nu) = (\lambda\mu) * \nu = \mu * (\lambda\nu)$ *for any* $\lambda \in \mathbb{C}$.
(v) $\|\mu * \nu\| \leqq \|\mu\| \cdot \|\nu\|$.
Hence $\mathfrak{M}(G)$ is a Banach algebra with respect to the operation $*$.
(vi) *If two sequences $\{\mu_n\}, \{\nu_n\}$ in $\mathfrak{M}(G)$ converge to μ_0 and ν_0 respectively in w^*-topology, then $\{\mu_n * \nu_n\}$ also converges to $\mu_0 * \nu_0$ in w^*-topology.*

The following basic result holds good for the Fourier transform $\widehat{\mu * \nu}$ of the convolution of μ and $\nu \in \mathfrak{M}(G)$. $\hat{\mu}$ and $\hat{\nu}$ are Fourier transforms of μ and ν, respectively.

Theorem 6.13 *For μ and $\nu \in \mathfrak{M}(G)$,*

$$\widehat{\mu * \nu} = \sqrt{2\pi}\hat{\mu} \cdot \hat{\nu}.$$

Proof Writing $\theta = \mu * \nu$, we obtain

$$\hat{\theta}(t) = \frac{1}{\sqrt{2\pi}} \int_G e^{-itx} d\theta(x) = \frac{1}{\sqrt{2\pi}} \int_G \int_G e^{-it(x+y)} d\mu(x) d\nu(y) = \sqrt{2\pi}\hat{\mu}(t)\hat{\nu}(t).$$

□

A measure contained in $\mathfrak{M}_d(G)$ is called a **discrete** measure. On the other hand, $\mu \in \mathfrak{M}(G)$ is said to be **continuous** if $\mu(\{x\}) = 0$ for all $x \in G$. It is well-known that any $\mu \in \mathfrak{M}(G)$ can be represented as a sum of a discrete measure μ_d and a continuous measure μ_c; i.e.

$$\mu = \mu_d + \mu_c.$$

Such a decomposition is determined uniquely.[32]

We list some elementary rules of the operation $*$ on $\mathfrak{M}(G)$:

1° If $\mu \in \mathfrak{M}(G)$ is continuous, $\mu * \nu$ is also continuous for all $\nu \in \mathfrak{M}(G)$.

2° For $\mu \in \mathfrak{M}(G)$, we define[33] $\mu^\sharp \in \mathfrak{M}(G)$ by

$$\mu^\sharp(E) = \overline{\mu(-E)}.$$

If $\mu = \mu_c + \mu_d$, then

$$\mu^\sharp = \mu_c^\sharp + \mu_d^\sharp.$$

3°

$$\mu * \mu^\sharp = \underbrace{(\mu_c * \mu_c^\sharp + \mu_c * \mu_d^\sharp + \mu_d * \mu_c^\sharp)}_{\text{continuous}} + \underbrace{\mu_d * \mu_d^\sharp}_{\text{discrete}}.$$

4° If $\mu_d = \sum_j a_j \delta_{\tau_j}$, then

$$\mu_d^\sharp = \sum_j \bar{a}_j \delta_{-\tau_j}.$$

[32] Any bounded Borel measure μ on G can be represented as $\mu = \mu_a + \mu_s + \mu_d$, where $\mu_a \ll m$ (Lebesque measure) μ_s is continuous and $\mu_s \perp m$, and μ_d is discrete. Such a decomposition is uniquely determined. cf. Igari [4] p. 137.

[33] $\widehat{\mu^\sharp} = \bar{\hat{\mu}}$.

Consequently,

$$(\mu * \mu^{\sharp})(\{0\}) = \sum_j |a_j|^2.$$

6.8 Wiener's Theorem

We first discuss Wiener's theorem on $[-\pi, \pi)$ and then go over to the case of $\mathbb{R}$.[34]

The following lemma is a simple consequence of 4° above.

Lemma 6.7 *For any $\mu \in \mathfrak{M}(G)$,*

$$\sum |\mu(\tau)|^2 = (\mu * \mu^{\sharp})(\{0\}),$$

where the left-hand side is the sum of countable τ's on which the mass of the discrete μ_d concentrates. Hence μ is continuous if and only if $(\mu * \mu^{\sharp})(\{0\}) = 0$.

Theorem 6.14 *For any $\mu \in \mathfrak{M}([-\pi, \pi))$ and $\tau \in [-\pi, \pi)$,*

$$\mu(\{\tau\}) = \lim_{N\to\infty} \frac{\sqrt{2\pi}}{2N+1} \sum_{j=-N}^{N} \hat{\mu}(j) e^{ij\tau}.$$

Proof We start by showing some preparatory calculations. Define

$$\varphi_N(t) = \sum_{j=-N}^{N} e^{ijt}, \quad t \in [-\pi, \pi). \tag{6.79}$$

If $t = 0$, $\varphi_N(t) = 2N + 1$. If, on the contrary, $t \neq 0$,[35]

$$\sum_{j=0}^{N} e^{ijt} = \frac{e^{i(N+1)t} - 1}{e^{it} - 1}. \tag{6.80}$$

---

[34] We owe this to Katznelson [5] Chap. I, §7 and Chap. VI, §2.

[35]
$$\begin{array}{r l}
 & e^{iNt} + e^{i(N-1)t} + \quad \cdots\cdots \quad + 1 \\ \hline
e^{it} - 1 \,\big) & e^{i(N+1)t} - 1 \\
 & e^{i(N+1)t} - e^{iNt} \\ \hline
 & e^{iNt} - 1 \\
 & e^{iNt} - e^{i(N-1)t} \\ \hline
 & e^{i(N-1)t} - 1 \\
 & \vdots
\end{array}$$

Similarly, we have

$$\sum_{j=-N}^{0} e^{ijt} = \frac{e^{i(N+1)t} - 1}{e^{it} - 1}. \tag{6.81}$$

It follows from (6.80) and (6.81) that

$$\varphi_N(t) = \frac{2\cos(N+1)t - 2}{e^{it} - 1} - 1. \tag{6.82}$$

Taking account of $t \neq 0$, choose a sufficiently small $\varepsilon > 0$ such that $|t| > \varepsilon$. If we define

$$\alpha = \operatorname*{Min}_{|\theta|>\varepsilon} |e^{i\theta} - 1|,$$

we obtain

$$|\varphi_N(t)| \leqq \frac{4}{\alpha} + 1 \tag{6.83}$$

by (6.82).

By the above results, we get by passing to the limit $N \to \infty$

$$\frac{1}{2N+1}\varphi_N(t) \to \begin{cases} 1 & \text{if} \quad t = 0, \\ 0 & \text{if} \quad t \neq 0. \end{cases} \tag{6.84}$$

(Note that $\varphi_N(0) = 2N + 1$.) Of course

$$\left| \frac{1}{2N+1}\varphi_N(t) \right| \leqq 1. \tag{6.85}$$

If we write

$$\varphi_N(\tau - t) = \sum_{j=-N}^{N} e^{ij(\tau - t)} = \sum_{j=-N}^{N} e^{ij\tau} e^{-ijt},$$

we obtain

$$\frac{1}{2N+1}\varphi_N(\tau - t) \to \begin{cases} 1 & \text{if} \quad t = \tau, \\ 0 & \text{if} \quad t \neq \tau, \end{cases} \tag{6.84$'$}$$

$$\left| \frac{1}{2N+1}\varphi_N(\tau - t) \right| \leqq 1 \tag{6.85$'$}$$

by the above calculations. Hence by the dominated convergence theorem,

$$\lim_{N\to\infty}\int_{-\pi}^{\pi}\frac{1}{2N+1}\varphi_N(\tau-t)d\mu(t)=\int_{-\pi}^{\pi}\chi_{\{\tau\}}d\mu=\mu(\{\tau\}). \tag{6.86}$$

The relation (6.86) can be rewritten as

$$\lim_{N\to\infty}\frac{\sqrt{2\pi}}{2N+1}\sum_{j=-N}^{N}\hat{\mu}(j)e^{ij\tau}=\mu(\{\tau\})$$

by definition of φ_N. □

Theorem 6.15 (Wiener) *For $\mu \in \mathfrak{M}([-\pi,\pi))$,*

$$\sum|\mu(\{\tau\})|^2=\lim_{N\to\infty}\frac{2\pi}{2N+1}\sum_{j=-N}^{N}|\hat{\mu}(j)|^2.$$

If μ is continuous, in particular, then the limit appearing on the right-hand side is 0.

Proof By Lemma 6.7, we have

$$\begin{aligned}\sum|\mu(\{\tau\})|^2&=(\mu*\mu^\sharp)(\{0\})\\&=\lim_{N\to\infty}\frac{\sqrt{2\pi}}{2N+1}\sum_{j=-N}^{N}\widehat{(\mu*\mu^\sharp)}(j)e^{ij0}\\&=\lim_{N\to\infty}\frac{2\pi}{2N+1}\sum_{j=-N}^{N}\hat{\mu}(j)\widehat{\mu^\sharp}(j)\quad\text{(Theorem 6.13)}\\&=\lim_{N\to\infty}\frac{2\pi}{2N+1}\sum_{j=-N}^{N}\hat{\mu}(j)\overline{\hat{\mu}(j)}\quad\text{(by footnote 33 on p. 157)}\\&=\lim_{N\to\infty}\frac{2\pi}{2N+1}\sum_{j=-N}^{N}|\hat{\mu}(j)|^2.\end{aligned}$$

□

It is not hard to transform the above results so as to fit with the case $\mathfrak{M}(\mathbb{R})$. Define a function φ_N by

$$\varphi_n(t)=\int_{-n}^{n}e^{i\xi t}d\xi,\quad t\in\mathbb{R}. \tag{6.87}$$

This is clearly an analogue of (6.79). Passing to the limit $n \to \infty$,[36] we have

$$\frac{1}{2n}\varphi_n(t) \to \begin{cases} 1 & \text{if} \quad t = 0, \\ 0 & \text{if} \quad t \neq 0 \end{cases} \tag{6.88}$$

and

$$\left|\frac{1}{2n}\varphi_n(t)\right| \leqq 1. \tag{6.89}$$

If we write

$$\varphi_n(\tau - t) = \int_{-n}^{n} e^{i\xi(\tau - t)} d\xi,$$

we have

$$\frac{1}{2n}\varphi_n(\tau - t) \to \begin{cases} 1 & \text{if} \quad t = \tau, \\ 0 & \text{if} \quad t \neq \tau, \end{cases} \tag{6.88$'$}$$

by the above result ($n \to \infty$), and

$$\left|\frac{1}{2n}\varphi_n(\tau - t)\right| \leqq 1. \tag{6.89$'$}$$

By the dominated convergence theorem,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \frac{1}{2n}\varphi_n(\tau - t) d\mu(t) = \int_{-\infty}^{\infty} \chi_{\{\tau\}} d\mu = \mu(\{\tau\}). \tag{6.90}$$

---

[36] Assume $t \neq 0$. Then

$$\varphi_n(t) = \frac{1}{it} e^{i\xi t}\Big|_{-n}^{n} = \frac{1}{it}(e^{int} - e^{-int}) = \frac{2}{t}\sin nt.$$

Hence

$$\frac{1}{2n}\varphi_n(t) = \frac{1}{nt}\sin t.$$

We obtain (6.88) as $n \to \infty$.

Since

$$\frac{1}{2n}\int_{-\infty}^{\infty}\varphi_n(\tau-t)d\mu(t) = \frac{1}{2n}\int_{-\infty}^{\infty}\int_{-n}^{n} e^{i\xi\tau}\cdot e^{-i\xi t}d\xi d\mu(t)$$
$$= \frac{1}{2n}\int_{-n}^{n} e^{i\xi\tau}\int_{-\infty}^{\infty} e^{-i\xi t}d\mu(t)d\xi$$
$$= \frac{\sqrt{2\pi}}{2n}\int_{-n}^{n}\hat{\mu}(\xi)e^{i\xi\tau}d\xi$$

by the definition of φ_n, we can rewrite (6.90) as

$$\lim_{n\to\infty}\frac{\sqrt{2\pi}}{2n}\int_{-n}^{n}\hat{\mu}(\xi)e^{i\xi\tau}d\xi = \mu(\{\tau\}). \tag{6.90$'$}$$

This is the result which corresponds to Theorem 6.14:

Theorem 6.14′ *For any $\mu \in \mathfrak{M}(\mathbb{R})$ and $\tau \in \mathbb{R}$, we have*

$$\lim_{n\to\infty}\frac{\sqrt{2\pi}}{2n}\int_{-n}^{n}\hat{\mu}(n)e^{i\xi\tau}d\xi = \mu(\{\tau\}).$$

It follows from Theorem 6.14′ that

$$\begin{aligned}\sum|\mu(\{\tau\})|^2 &= (\mu*\mu^\sharp)(\{0\}) \quad \text{(by Lemma 6.7)}\\ &= \lim_{n\to\infty}\frac{\sqrt{2\pi}}{2n}\int_{-n}^{n}(\widehat{\mu*\mu^\sharp})(\xi)e^{i\xi 0}d\xi\\ &= \lim_{n\to\infty}\frac{2\pi}{2n}\int_{-n}^{n}\hat{\mu}(\xi)\widehat{\mu^\sharp}(\xi)d\xi\\ &= \lim_{n\to\infty}\frac{\pi}{n}\int_{-n}^{n}|\hat{\mu}(\xi)|^2d\xi \quad \text{(by footnote 33 on p. 157).}\end{aligned} \tag{6.91}$$

Summing up:

Theorem 6.16 (Wiener) *For $\mu \in \mathfrak{M}(\mathbb{R})$,*

$$\sum|\mu(\{\tau\})|^2 = \lim_{n\to\infty}\frac{\pi}{n}\int_{-n}^{n}|\hat{\mu}(\xi)|^2d\xi.$$

If μ is continuous, the limit appearing on the right-hand side is 0.[37]

[37] The upper and lower ends of the integration can be $\displaystyle\int_{\theta_n}^{\theta_n+n}$ for any sequence $\{\theta_n\}$.

References

1. Bochner, S.: Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse. Math. Ann. **108**, 378–410 (1933)
2. Carathéodory, C.: Über den Variabilitätsbereich der Koefficienten von Potenzreihen die gegebene Werte nicht annemen. Math. Ann. **54**, 95–115 (1907)
3. Herglotz, G.: Über Potenzreihen mit positiven reellen Teil in Einheitskreise. Berichte Verh Säcks. Akad. Wiss. Leipzig. Math.-phys. Kl. **63**, 501–511 (1911)
4. Igari, S.: Jitsu-kaiseki Nyumon (Introduction to Real Analysis). Iwanami Shoten, Tokyo (1996) (Originally published in Japanese)
5. Katznelson, Y.: An Introduction to Harmonic Analysis, 3rd edn. Cambridge University Press, Cambridge (2004)
6. Lax, P.D.: Functional Analysis. Wiley, New York (2002)
7. Malliavin, P.: Integration and Probability. Springer, New York (1995)
8. Maruyama, T.: Kansu Kaisekigaku (Functional Analysis). Keio Tsushin, Tokyo (1980) (Originally published in Japanese)
9. Maruyama, T.: Sekibun to Kansu-kaiseki (Integration and Functional Analysis). Springer, Tokyo (2006) (Originally published in Japanese)
10. Maruyama, T.: Herglotz-Bochner representation theorem via theory of distributions. J. Operations Res. Soc. Japan **60**, 122–135 (2017)
11. Naimark, M.A.: Normed Algebras. Wolters Noordhoff, Groningen (1972)
12. Schwartz, L.: Théorìe des distributions. Nouvelle édition, entièrement corrigée, refondue et augmentée, Hermann, Paris (1973)
13. Toeplitz, O.: Über die Fouriersche Entwickelung positive Funktionen. Rend. di Cinc. Mat. di Palermo **32**, 191–192 (1911)

Chapter 7
Spectral Representation of Unitary Operators

The main topic of this chapter is the spectral representation of unitary operators as well as one-parameter groups of unitary operators. That is, the problem is how to represent such objects by certain analogues of Fourier transforms. We are going to describe theories based upon the Herglotz–Bochner theorem already discussed in Chap. 6. However, the Herglotz–Bochner theorem can be conversely deduced from the spectral representation theorem of unitary operators.

The results obtained here will play a crucial role in the spectral representation theory of weakly stationary stochastic processes to be discussed in the next chapter.

7.1 Lax–Milgram Theorem

Linear spaces appearing in the following are basically complex ones. We have to add one more important fundamental result to the exposition on Hilbert spaces discussed in Chap. 1.

Definition 7.1 Let $\mathfrak{X}$ be a vector space. A function $\Phi : \mathfrak{X} \times \mathfrak{X} \to \mathbb{C}$ is called a **sesquilinear functional** if it satisfies the following conditions, where x, x', y, y' are any elements of $\mathfrak{X}$ and $\alpha \in \mathbb{C}$:

(i) $\Phi(x + x', y) = \Phi(x, y) + \Phi(x', y)$.
(ii) $\Phi(x, y + y') = \Phi(x, y) + \Phi(x, y')$.
(iii) $\Phi(\alpha x, y) = \alpha\Phi(x, y)$.
(iv) $\Phi(x, \alpha y) = \bar{\alpha}\Phi(x, y)$.

Furthermore, if Φ satisfies

$$\Phi(x, y) = \overline{\Phi(y, x)},$$

it is said to be **skew-symmetric**.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_7

An inner product $\langle\cdot,\cdot\rangle$ on a Hilbert space is clearly a skew-symmetric sesquilinear functional.[1]

Remark 7.1 $\Phi(x,x)$ is not necessarily real for a sesquilinear functional Φ. However, if Φ is skew-symmetric, in addition, $\Phi(x,x)$ must be real.

This concept permits us to generalize the Riesz theorem (Theorem 1.1) to represent duals of Hilbert spaces.[2]

Theorem 7.1 (Lax–Milgram) *Let $\mathfrak{H}$ be a Hilbert space. Suppose that a function $B : \mathfrak{H}\times\mathfrak{H}\to\mathbb{C}$ is a sesquilinear functional which satisfies the following two conditions:*

(i) *There exists a constant $\alpha > 0$ such that*

$$|B(x,y)| \leqq \alpha\|x\|\cdot\|y\| \quad \textit{for all} \quad x,y\in\mathfrak{H}.$$

(ii) *There exists a constant $\beta > 0$ such that*

$$|B(y,y)| \geqq \beta\|y\|^2 \quad \textit{for all} \quad y\in\mathfrak{H}.$$

Then there exists a unique $y_\Lambda\in\mathfrak{H}$ for each $\Lambda\in\mathfrak{H}'$ which satisfies

$$\Lambda(x) = B(x,y_\Lambda) \quad \textit{for all} \quad x\in\mathfrak{H}.$$

Proof For a fixed $y\in\mathfrak{H}$, the function $x\mapsto B(x,y)$ is a bounded linear functional on $\mathfrak{H}$. By Riesz's theorem, there exists some unique $z\in\mathfrak{H}$ such that

$$B(x,y) = \langle x,z\rangle \quad \text{for all} \quad x\in\mathfrak{H}. \tag{7.1}$$

If we write

$$Ay = z,$$

(7.1) can be rewritten as

$$B(x,y) = \langle x,Ay\rangle, \quad x,y\in\mathfrak{H}. \tag{7.2}$$

[1] However, we have to remember that an inner product must satisfy one more requirement,

$$\langle x,x\rangle \geqq 0; \quad \langle x,x\rangle = 0 \Leftrightarrow x = 0,$$

besides skew-symmetric sesquilinearity.

[2] The Lax–Milgram theorem has a lot of applications. Among them, its application to the existence theorem of solutions for elliptic partial differential equations is most peculiar. See Evans [4], Chap. 6.

$A : \mathfrak{H} \to \mathfrak{H}$ is a bounded linear operator. It can be verified as follows. In fact, A is linear, since we have, for any $\lambda_1, \lambda_2 \in \mathbb{C}$ and $y_1, y_2 \in \mathfrak{H}$,

$$\begin{aligned}
\langle x, A(\lambda_1 y_1 + \lambda_2 y_2)\rangle &= B(x, \lambda_1 y_1 + \lambda_2 y_2) \qquad \text{(by (7.2))} \\
&= \bar{\lambda}_1 B(x, y_1) + \bar{\lambda}_2 B(x, y_2) \qquad \text{(by sesquilinearity)} \\
&= \bar{\lambda}_1 \langle x, Ay_1\rangle + \bar{\lambda}_2 \langle x, Ay_2\rangle \\
&= \langle x, \lambda_1 Ay_1 + \lambda_2 Ay_2\rangle
\end{aligned}$$

for each $x \in \mathfrak{H}$. Furthermore,

$$\|Ay\|^2 = \langle Ay, Ay\rangle \underset{(7.2)}{=} B(Ay, y) \underset{\text{(i)}}{\leqq} \alpha\|Ay\| \cdot \|y\|.$$

In the case of $\|Ay\| \neq 0$, we obtain

$$\|Ay\| \leqq \alpha\|y\| \quad \text{for all} \quad y \in \mathfrak{H}$$

by dividing both sides by $\|Ay\|$. If $\|Ay\| = 0$, the same result holds good obviously. Hence A is a bounded linear operator.

We next show that A is injective. In fact

$$\beta\|y\|^2 \leqq |B(y, y)| = |\langle y, Ay\rangle| \leqq \|y\| \cdot \|Ay\|$$

by (ii). It follows immediately that

$$\beta\|y\| \leqq \|Ay\|. \tag{7.3}$$

This proves that A is injective.

The inequality (7.3) implies that the image $A(\mathfrak{H})$ of A is a closed subspace of $\mathfrak{H}$. Assume that

$$Ay_n \to z \quad \text{as} \quad n \to \infty$$

for a sequence $\{y_n\}$ in $\mathfrak{H}$. Then $\{y_n\}$ is Cauchy, since

$$\beta\|y_n - y_m\| \leqq \|Ay_n - Ay_m\|$$

by (7.3). Therefore $\{y_n\}$ has a limit $y_0 \in \mathfrak{H}$. Since A is continuous, we have

$$Ay_n \to Ay_0 \quad \text{as} \quad n \to \infty.$$

By the uniqueness of the limit, we must have $z = Ay_0$ and hence $z \in A(\mathfrak{H})$. This proves the closedness of $A(\mathfrak{H})$.[3]

However, it holds good actually that

$$A(\mathfrak{H}) = \mathfrak{H}. \tag{7.4}$$

Suppose that $A(\mathfrak{H}) \subsetneq \mathfrak{H}$ to the contrary. Then there exists some $x \in A(\mathfrak{H})^{\perp} \setminus \{0\}$. But it follows that $x = 0$, since

$$\beta \|x\|^2 \leqq |B(x, x)| = \langle x, Ax \rangle = 0.$$

Contradiction.

Applying Riesz's theorem again to $\Lambda \in \mathfrak{H}'$, there exists a unique $z_\Lambda \in \mathfrak{H}$ such that

$$\Lambda(x) = \langle x, z_\Lambda \rangle \quad \text{for all} \quad x \in \mathfrak{H}. \tag{7.5}$$

By (7.4), there exists some $y_\Lambda \in \mathfrak{H}$ such that $z_\Lambda = Ay_\Lambda$. Hence

$$B(x, y_\Lambda) = \langle x, Ay_\Lambda \rangle = \langle x, z_\Lambda \rangle = \Lambda(x),$$

that is,

$$\Lambda(x) = B(x, y_\Lambda) \quad \text{for all} \quad x \in \mathfrak{H}.$$

Finally, we show the uniqueness of y_Λ. Suppose that there is another $y'_\Lambda \in \mathfrak{H}$ which satisfies

$$\Lambda(x) = B(x, y_\Lambda) = B(x, y'_\Lambda) \quad \text{for all} \quad x \in \mathfrak{H}.$$

Then

$$B(x,\ y_\Lambda - y'_\Lambda) = 0 \quad \text{for all} \quad x \in \mathfrak{H}.$$

---

[3]There is an alternative way to prove $z = Ay_0$. Since (7.2) implies

$$B(x, y_n) = \langle x, Ay_n \rangle \quad \text{for each fixed} \quad x \in \mathfrak{X},$$

we have

$$B(x, y_0) = \langle x, z \rangle \quad \text{for all} \quad x \in \mathfrak{H}$$

by passing to the limit. Hence $z = Ay_0$.

By (ii), we obtain

$$\beta\|y_\Lambda - y'_\Lambda\|^2 \leqq |B(y_\Lambda - y'_\Lambda,\ y_\Lambda - y'_\Lambda)| = 0.$$

This proves $y_\Lambda = y'_\Lambda$. □

Remark 7.2 The skew-symmetricity of B is not assumed in Theorem 7.1. If it is assumed, B itself is an inner product. Consequently, the Lax–Milgram theorem immediately follows through the Riesz theorem applied to B. An important point about the Lax–Milgram theorem is that the skew-symmetricity is not required.

Let $\Phi(x, y)$ be a sesquilinear functional on $\mathfrak{H}$. It is said to be **bounded** if

$$|\Phi(x, y)| \leqq \alpha\|x\| \cdot \|y\| \quad \text{for all} \quad x, y \in \mathfrak{H}$$

for some $\alpha > 0$. The set of all the bounded sesquilinear functionals forms a (complex) vector space under usual operations and a norm on this space is given by

$$\|\Phi\| = \sup_{\substack{x\neq 0\\ y\neq 0}} \frac{|\Phi(x, y)|}{\|x\| \cdot \|y\|}.$$

7.2 Conjugate Operators and Projections

We denote by $\mathcal{L}(\mathfrak{H})$ the space of bounded linear operators ($\mathfrak{H} \to \mathfrak{H}$) on a complex Hilbert space $\mathfrak{H}$.

Fixing $T \in \mathcal{L}(\mathfrak{H})$, we define a function $\Phi : \mathfrak{H} \times \mathfrak{H} \to \mathbb{C}$ by

$$\Phi(x, y) = \langle Tx, y\rangle. \tag{7.6}$$

Then Φ is clearly a sesquilinear functional on $\mathfrak{H}$ and it satisfies

$$|\Phi(x, y)| \leqq \|T\| \cdot \|x\| \cdot \|y\|$$

and so $\|\Phi\| \leqq \|T\|$. On the other hand, since

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \Phi(x, Tx) \leqq \|\Phi\| \cdot \|x\| \cdot \|Tx\|,$$

$\|Tx\| \leqq \|\Phi\| \cdot \|x\|$ if $Tx \neq 0$. Hence $\|T\| \leqq \|\Phi\|$. Thus we prove that

$$\|T\| = \|\Phi\|. \tag{7.7}$$

It should be observed that any bounded sesquilinear functional on $\mathfrak{H}$ can be represented in the form (7.6). It is easily shown by appealing to the Riesz theorem.

Theorem 7.2 *For any bounded sesquilinear functional Φ on $\mathfrak{H}$, there exists uniquely some $T \in \mathcal{L}(\mathfrak{H})$ such that*

$$\Phi(x, y) = \langle Tx, y\rangle \quad x, y \in \mathfrak{H}.$$

Furthermore, $\|T\| = \|\Phi\|$.

Proof Fixing $x \in \mathfrak{H}$, we define $\Phi_x(y) = \Phi(x, y)$. Then $\overline{\Phi_x(y)} \in \mathfrak{H}'$. Hence by the Riesz theorem, there exists uniquely some $Tx \in \mathfrak{H}$ such that

$$\overline{\Phi_x(y)} = \langle y, Tx\rangle, \quad y \in \mathfrak{H}.$$

The operator $T : x \mapsto Tx$ thus defined is clearly linear,[4] and $\|T\| = \|\Phi\|$ by (7.7). □

Theorem 7.3 *For any $T \in \mathcal{L}(\mathfrak{H})$, there exists uniquely some $T^* \in \mathcal{L}(\mathfrak{H})$ which satisfies*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle, \quad x, y \in \mathfrak{H}.$$

It holds good that $\|T\| = \|T^*\|$.

Proof Define a couple of functionals, Φ and $\Psi : \mathfrak{H} \times \mathfrak{H} \to \mathbb{C}$, by

$$\Phi(x, y) = \langle Tx, y\rangle, \quad \Psi(x, y) = \overline{\Phi(y, x)} = \overline{\langle Ty, x\rangle}.$$

These are bounded sesquilinear functionals and

$$\|\Phi\| = \|\Psi\| = \|T\|.$$

By Theorem 7.2, there exists uniquely some $T^* \in \mathcal{L}(\mathfrak{H})$ such that

$$\Psi(x, y) = \langle T^*x, y\rangle, \quad x, y \in \mathfrak{H},$$

and

$$\|T^*\| = \|\Psi\| = \|T\|.$$

This T^* satisfies

$$\langle Tx, y\rangle = \Phi(x, y) = \overline{\Psi(y, x)} = \overline{\langle T^*y, x\rangle} = \langle x, T^*y\rangle, \quad x, y \in \mathfrak{H}.$$

□

[4] Let $x_1, x_2 \in \mathfrak{H}$ and $\alpha_1, \alpha_2 \in \mathbb{C}$. Then, for any $y \in \mathfrak{H}$, $\langle y, \alpha_1 Tx_1 + \alpha_2 Tx_2\rangle = \bar{\alpha_1}\langle y, Tx_1\rangle + \bar{\alpha_2}\langle y, Tx_2\rangle = \bar{\alpha_1}\overline{\Phi(x_1, y)} + \bar{\alpha_2}\overline{\Phi(x_2, y)} = \overline{\Phi(\alpha_1 x_1 + \alpha_2 x_2, y)} = \langle y, T(\alpha_1 x_1 + \alpha_2 x_2)\rangle$. This implies $\alpha_1 Tx_1 + \alpha_2 Tx_2 = T(\alpha_1 x_1 + \alpha_2 x_2)$.

Definition 7.2 The operator T^* defined by Theorem 7.3 is called the **conjugate operator** or **adjoint operator** of T.[5]

The elementary properties of conjugate operators are as follows[6]:

1° $T^{**} = T$.
2° For $S, T \in \mathcal{L}(\mathfrak{H})$ and $\alpha, \beta \in \mathbb{C}$,

$$(\alpha S + \beta T)^* = \bar{\alpha} S^* + \bar{\beta} T^*.$$

3° For $S, T \in \mathcal{L}(\mathfrak{H})$,

$$(ST)^* = T^* S^*.$$

4° $\|T\|^2 = \|T^*T\|$.

Example 7.1 Any $T \in \mathcal{L}(\mathbb{C}^n)$ can be represented by a matrix (t_{ij}). The conjugate operator T^* of T is represented by the conjugate transposed matrix of (t_{ij}).

Example 7.2 Consider the Hilbert space $\mathfrak{L}^2([0,1], \mathbb{C})$. Let $K(x, y)$ be a measurable function of $[0,1] \times [0,1]$ into $\mathbb{C}$ such that

$$\int_0^1 \int_0^1 |K(x, y)|^2 dxdy < \infty.$$

If we define an operator T on $\mathfrak{L}^2([0,1]^2, \mathbb{C})$ by

$$(Tf)(x) = \int_0^1 K(x, y) f(y) dy,$$

then T is a bounded linear operator. The dual operator T^* of T is given by

$$(T^* f)(y) = \int_0^1 \overline{K(x, y)} f(x) dx.$$

If we define a function $K^* : [0,1] \times [0,1] \to \mathbb{C}$ by

$$K^*(y, x) = \overline{K(x, y)},$$

[5] For abstract theory of linear operators on a Hilbert space (theory of C^*-algebra), see Arveson [1], Diximier [2] and Maruyama [7] Chap. 7.

[6] For instance, 4° can be proved as follows: $\|T^*T\| \leqq \|T^*\| \cdot \|T\| = \|T\|^2$. On the other hand,

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle \leqq \|T^*T\| \cdot \|x\|^2$$

for all $x \in \mathfrak{H}$. This implies $\|T\|^2 \leqq \|T^*T\|$.

then T^* can also be expressed as

$$(T^* f)(y) = \int_0^1 K^*(y, x) f(x) dx.$$

Example 7.3 The Fourier transform on $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ in the sense of Plancherel is denoted by $\mathcal{F}_2$. By Theorem 4.3 (p. 72),

$$\mathcal{F}_2^* = \mathcal{F}_2^{-1} \quad \text{(inverse Fourier transform).}$$

Definition 7.3 An operator $T \in \mathcal{L}(\mathfrak{H})$ is called a **symmetric** or **Hermitian operator** if $T = T^*$.[7]

Example 7.4 In Example 7.1 above, $T = T^*$ if $t_{ij} = \overline{t_{ji}}$. In Example 7.2, $T = T^*$ if $K(x, y) = \overline{K(x, y)}$ a.e. x, y.

Theorem 7.4 *The following statements are equivalent for $T \in \mathcal{L}(\mathfrak{H})$:*

(i) T is symmetric.
(ii) The sesquilinear operator Φ defined by $\Phi(x, y) = \langle Tx, y\rangle$ $(x, y \in \mathfrak{H})$ is skew-symmetric.
(iii) $\langle Tx, x\rangle$ is real for all $x \in \mathfrak{H}$.

Proof (i)⇔(ii): This is obvious because

$$\Phi(x, y) = \overline{\Phi(y, x)} \quad \text{for all} \quad x, y \in \mathfrak{H}$$

is equivalent to

$$\langle Tx, y\rangle = \overline{\langle Ty, x\rangle} = \overline{\langle y, T^*x\rangle} = \langle T^*x, y\rangle \quad \text{for all} \quad x, y \in \mathfrak{H}.$$

(i)⇔(iii): If T is symmetric,

$$\overline{\langle x, Tx\rangle} = \langle Tx, x\rangle = \langle x, Tx\rangle$$

for any $x \in \mathfrak{H}$. Hence $\langle Tx, x\rangle$ is real.

Conversely, assume that $\langle Tx, x\rangle$ is real for all $x \in \mathfrak{H}$. Then

$$\langle x + y, T(x + y)\rangle = \langle x, Tx\rangle + \langle y, Ty\rangle + \langle x, Ty\rangle + \langle y, Tx\rangle$$

for any $x, y \in \mathfrak{H}$. Then the left-hand side and the first two terms on the right-hand side are real. So

$$\langle x, Ty\rangle + \langle y, Tx\rangle \equiv \alpha \tag{7.8}$$

[7]For the concept of a self-adjoint operator which is closely related with that of a symmetric operator, see Lax [6] Chap. 32.

is also real. Substituting iy instead of y, we observe similarly that

$$i\langle x, Ty\rangle - i\langle y, Tx\rangle \equiv \beta \tag{7.9}$$

is real. By (7.8) and (7.9), we have

$$2\langle x, Ty\rangle = \alpha - i\beta, \tag{7.10}$$

$$2\langle y, Tx\rangle = \alpha + i\beta, \quad \alpha, \beta \in \mathbb{R}. \tag{7.11}$$

A simple calculation

$$2\langle Tx, y\rangle = 2\overline{\langle y, Tx\rangle} \underset{(7.11)}{=} \alpha - i\beta \underset{(7.10)}{=} 2\langle x, Ty\rangle$$

gives a conclusion

$$\langle Tx, y\rangle = \langle x, Ty\rangle, \quad x, y \in \mathfrak{H},$$

that is, $T = T^*$. □

Remark 7.3 We used the imaginary unit i in the course of the proof (iii)⇒(i). This is an essential point. This argument can not be applied to a real Hilbert space. In fact, $\langle Tx, x\rangle$ is always real in a real Hilbert space. However, a linear operator is not necessarily symmetric.[8]

Combining Theorems 7.2 and 7.4, we get the following corollary.

Corollary 7.1 *For any bounded and skew-symmetric sesquilinear functional* $\Phi : \mathfrak{H} \to \mathbb{C}$*, there exists uniquely some symmetric operator* $T \in \mathcal{L}(\mathfrak{H})$ *such that*

$$\Phi(x, y) = \langle Tx, y\rangle, \quad x, y \in \mathfrak{H}.$$

Furthermore, it holds good that $\|T\| = \|\Phi\|$.

Let $\mathfrak{M}$ be a closed subspace of a Hilbert space $\mathfrak{H}$. We have already seen the orthogonal decomposition; that is, any $z \in \mathfrak{H}$ can be uniquely represented as

$$z = x + y, \quad x \in \mathfrak{M}, \quad y \in \mathfrak{M}^{\perp}.$$

The linear operator $P : \mathfrak{H} \to \mathfrak{M}$ which associates each $z \in \mathfrak{H}$ with x in the decomposition is called a **projection** or **projection operator** of $\mathfrak{H}$ into $\mathfrak{M}$.

We list a few elementary properties of a projection.

[8] This remark is due to Kato [5] p. 257.

1° Let P be the projection of $\mathfrak{H}$ into the closed subspace $\mathfrak{M}$. $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are defined by

$$\mathfrak{M}_1 = \{x \in \mathfrak{H} | Px = x\}, \quad \mathfrak{M}_2 = \{Px | x \in \mathfrak{H}\}.$$

Then $\mathfrak{M} = \mathfrak{M}_1 = \mathfrak{M}_2$.

2° Let P be the projection of $\mathfrak{H}$ into the closed subspace $\mathfrak{M}$. If $\|Px\| = \|x\|$ for $x \in \mathfrak{H}$, then $Px = x$.

Proof Since $x = Px + (x - Px)$, $Px \in \mathfrak{M}$ and $x - Px \in \mathfrak{M}^{\perp}$, we have

$$\|x\|^2 = \|Px\|^2 + \|x - Px\|^2.$$

By assumption $\|Px\| = \|x\|$. Hence $\|x - Px\| = 0$. □

Theorem 7.5 *Let $\mathfrak{H}$ be a Hilbert space and $\mathfrak{M}$ its closed subspace. The projection $P : \mathfrak{H} \to \mathfrak{M}$ satisfies the following properties:*

(i) $P^2 = P$ *(idempotency),*
(ii) $P^* = P$ *(symmetricity).*

If $\mathfrak{M} \neq \{0\}$, then $\|P\| = 1$.

Proof Let $z = x + y$ $(x \in \mathfrak{M},\ y \in \mathfrak{M}^{\perp})$ be the orthogonal decomposition of $z \in \mathfrak{H}$. $P^2 = P$ follows from

$$P^2 z = Px = x = Pz.$$

Since

$$\|Pz\|^2 = \|x\|^2 \leqq \|x\|^2 + \|y\|^2 = \|z\|^2,$$

we obtain $\|P\| \leqq 1$. In the case of $\mathfrak{M} \neq \{0\}$, there exists some $z \in \mathfrak{M} \setminus \{0\}$. We must have $Pz = z$. Hence it follows that $\|P\| = 1$.

Finally, we show $P = P^*$. Decompose $x, y \in \mathfrak{H}$ in the form

$$x = x_1 + x_2, \quad x_1 \in \mathfrak{M}, \quad x_2 \in \mathfrak{M}^{\perp},$$

$$y = y_1 + y_2, \quad y_1 \in \mathfrak{M}, \quad y_2 \in \mathfrak{M}^{\perp}.$$

$\langle Px, y\rangle$ can be calculated as

$$\langle Px, y\rangle = \langle x_1, y\rangle = \langle x_1, y_1\rangle = \langle x, y_1\rangle = \langle x, Py\rangle.$$

This holds good for any $x, y \in \mathfrak{H}$. Hence $P = P^*$. □

The next theorem is the converse assertion of Theorem 7.5.

Theorem 7.6 *Assume that $P \in \mathcal{L}(\mathfrak{H})$ satisfies (i) $P^2 = P$ and (ii) $P^* = P$. Then P is the projection of $\mathfrak{H}$ into*

$$\mathfrak{M} = \{x \in \mathfrak{H} | Px = x\}.$$

Proof By (i), we obtain

$$P(Px) = Px \tag{7.12}$$

for all $x \in \mathfrak{H}$. We also obtain, by (ii),

$$\langle y, x - Px\rangle = \langle y, x\rangle - \langle Py, x\rangle = 0 \tag{7.13}$$

for all $x \in \mathfrak{H}$ and $y \in \mathfrak{M}$. It follows from (7.12) and (7.13) that

$$Px \in \mathfrak{M}, \quad x - Px \in \mathfrak{M}^{\perp}.$$

Taking account of

$$x = Px + (x - Px),$$

we get the desired result. □

In sum, a projection operator on a Hilbert space can be characterized as an idempotent and symmetric operator.

7.3 Unitary Operators

We start with the definition and some properties of unitary operators.

Definition 7.4 An operator $T \in \mathcal{L}(\mathfrak{H})$ is called a **unitary operator** if it satisfies $TT^* = T^*T = I$ (identity operator).

The Fourier transform and the inverse Fourier transform are typical examples of this category. The set of unitary operators in $\mathcal{L}(\mathfrak{H})$ forms a group with respect to the composition of operators. It is called the **unitary group**.

Theorem 7.7 *The following three statements are equivalent for an operator $T \in \mathcal{L}(\mathfrak{H})$:*

(i) *T is a unitary operator.*
(ii) *$T(\mathfrak{H}) = \mathfrak{H}$ and*

$$\langle Tx, Ty\rangle = \langle x, y\rangle$$

for all $x, y \in \mathfrak{H}$.

(iii) $T(\mathfrak{H}) = \mathfrak{H}$ *and*

$$\|Tx\| = \|x\|$$

for all $x \in \mathfrak{H}$.

Proof (i)⇒(ii): Assume T is unitary. Then T is invertible and T^* is the inverse, since $TT^* = I$ by definition. Hence T is a bijection. By a simple calculation, we have

$$\langle Tx, Ty \rangle = \langle x, T^*Ty \rangle = \langle x, y \rangle \quad \text{for all} \quad x, y \in \mathfrak{H}.$$

(ii)⇒(iii): Obvious.
(iii)⇒(i): Assuming (iii), we obtain

$$\langle T^*Tx, x \rangle = \langle Tx, Tx \rangle = \|Tx\|^2 = \|x\|^2 = \langle x, x \rangle$$

for all $x \in \mathfrak{H}$; i.e.

$$\langle (T^*T - I)x, x \rangle = 0 \quad \text{for all} \quad x \in \mathfrak{H}. \tag{7.14}$$

Hence $T^*T = I$. (See the remark below.) Since T is a bijection of $\mathfrak{H}$ onto itself by assumption, there exists the inverse operator T^{-1} of T. Thus we must have $T^{-1} = T^*$, which implies

$$T^*T = TT^* = I.$$

□

$T \in \mathcal{L}(\mathfrak{H})$ **is unitary if and only if it is an isometric isomorphism on** $\mathfrak{H}$.

Remark 7.4 The deduction of $T^*T = I$ from (7.14) in the proof of Theorem 7.7 is based upon the following general principle:

$$\langle Sx, x \rangle = 0 \quad \text{for all} \quad x \in \mathfrak{H} \Rightarrow S = 0 \tag{7.15}$$

for $S \in \mathcal{L}(\mathfrak{H})$. In fact, according to (7.15),

$$\langle S(x + y), x + y \rangle = \langle Sx, y \rangle + \langle Sy, x \rangle = 0, \tag{7.16}$$

$$\langle S(x + iy), x + iy \rangle = -i\langle Sx, y \rangle + i\langle Sy, x \rangle = 0 \tag{7.17}$$

for $x, y \in \mathfrak{H}$. Multiplying (7.17) by i and adding it to (7.16), we have

$$2\langle Sx, y \rangle = 0 \quad \text{for all} \quad x, y \in \mathfrak{H}.$$

This implies that

$$Sx = 0 \quad \text{for all} \quad x \in \mathfrak{H}.$$

In the proof of Theorem 7.7, $T^*T - I$ plays the role of S.

7.4 Resolution of the Identity

In this section, we discuss a general theory of the resolution of the identity, which is to take part in the spectral representation problem of unitary operators, the main topic in the next section.

Definition 7.5 Let $(\Omega, \mathcal{E})$ be a measurable space and $\mathfrak{H}$ a Hilbert space. A mapping

$$E : \mathcal{E} \to \mathcal{L}(\mathfrak{H})$$

is called a **resolution of the identity** on $\mathcal{E}$ if it satisfies the following conditions:

1° $E(\emptyset) = 0,\ E(\Omega) = I.$
2° $E(A)$ is a projection for each $A \in \mathcal{E}$.
3° $E(A \cap B) = E(A)E(B)$ for $A, B \in \mathcal{E}$.
4° $A, B \in \mathcal{E},\ A \cap B = \emptyset \Rightarrow E(A \cup B) = E(A) + E(B).$
5° If we define, for each $x, y \in \mathfrak{H}$, a function $E_{x,y} : \mathcal{E} \to \mathbb{C}$ by

$$E_{x,y}(A) = \langle\, E(A)x, y \,\rangle,\ \ A \in \mathcal{E},$$

then $E_{x,y}$ is a complex measure on $(\Omega, \mathcal{E})$.

The next result is an easy consequence of the definition.

Theorem 7.8 (properties of the resolution) *A resolution E of the identity satisfies the following properties, where $A, B \in \mathcal{E}$ and $x \in \mathfrak{H}$:*

(i) $E_{x,x}(A) = \|E(A)x\|^2$.
(ii) *$E_{x,x}$ is a (positive) measure on $(\Omega, \mathcal{E})$ and*

$$E_{x,x}(\Omega) = \|x\|^2.$$

(iii) $E(A)E(B) = E(B)E(A)$.
(iv) *If $A \cap B = \emptyset$, the images of $E(A)$ and $E(B)$ are orthogonal to each other.*
(v) *E is finitely additive.*

Proof (i) By 2°, $E(A)$ is a projection. Hence

$$E_{x,x}(A) = \langle E(A)x, x \rangle = \langle E(A)^2 x, x \rangle = \langle E(A)x, E(A)x \rangle = \|E(A)x\|^2.$$

(ii) Obvious by $E(\Omega) = I$ and (i).
(iii) By 3°,

$$E(A)E(B) = E(A \cap B) = E(B)E(A).$$

(iv) Suppose that $A \cap B = \emptyset$. Then

$$\begin{aligned} \langle E(A)x, E(B)y \rangle &= \langle x, E(A)E(B)y \rangle \quad (\text{by } E(A) = E(A)^*) \\ &\underset{3^\circ}{=} \langle x, E(A \cap B)y \rangle = \langle x, E(\emptyset)y \rangle \underset{1^\circ}{=} \langle x, 0 \rangle = 0 \end{aligned}$$

for any $x, y \in \mathfrak{H}$.
(v) is also clear by 4°.

□

Remark 7.5

1° E is not necessarily countably additive. Let $A_n \in \mathcal{E}$ be a sequence of mutually disjoint measurable sets. Does a series

$$\sum_{n=1}^{\infty} E(A_n) \tag{7.18}$$

of operators converge strongly? Since each $E(A_n)$ is a projection, its norm is 0 or 1 (cf. Theorem 7.5). Hence the sequence of partial sums

$$\sum_{n=1}^{N} E(A_n); \quad N = 1, 2, \cdots$$

of (7.18) can not be Cauchy except for the case where all but at most finite terms of (7.18) are zero. This observation shows that the series (7.18) does not strongly converge in general except for some special case.

Some more detailed explanation may be required for the above argument. If we define $\mathfrak{M}_n = \{x \in \mathfrak{H} | E(A_n)x = x\}$ for each operator $E(A_n)$, then $E(A_n)$ is a projection to $\mathfrak{M}_n$. In the case of $\mathfrak{M}_n \neq \{0\}$, $\|E(A_n)\| = 1$. Suppose that $\mathfrak{M}_n \neq \{0\}$ for infinitely many n, and say, $\mathfrak{M}_{k_0} \neq \{0\}$, $m < k_0 < n$. Consider partial sums

$$\sum_{k=1}^{m} E(A_k), \quad \sum_{k=1}^{n} E(A_k)$$

of (7.18) and suppose that $x_0 \in \mathfrak{M}_{k_0}$, $\|x_0\| = 1$. Since $\langle E(A_k)x_0, E(A_{k'})x_0\rangle = 0$, we obtain

$$\left\| \sum_{k=m+1}^{n} E(A_k)x_0 \right\|^2 = \sum_{k=m+1}^{n} \|E(A_k)x_0\|^2 \geqq \|E(A_{k_0})x_0\| = \|x_0\| = 1.$$

Hence $\left\| \sum_{k=m+1}^{n} E(A_k) \right\| \geqq 1$. There exist infinitely many indices which satisfy a similar condition to k_0 above. Thus we observe that the sequence of partial sums of (7.18) can not be Cauchy.

2° However, $E(\cdot)x$ is σ-additive for any fixed $x \in \mathfrak{H}$.

To show this, we should recall the following general principle.[9]

Let $\{x_n\}$be a sequence in a Hilbert space $\mathfrak{H}$ such that $x_n \perp x_m$ for $n \neq m$. Then the following statements are equivalent:

(i) $\sum_{n=1}^{\infty} x_n$ converges strongly.

(ii) $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$.

(iii) $\sum_{n=1}^{\infty} \langle x_n, y \rangle$ converges for any $y \in \mathfrak{H}$.

We now go back to 2°. Let $A_n \in \mathcal{E}$ $(n = 1, 2, \cdots)$ be mutually disjoint measurable sets. By Theorem 7.8 (iv),

$$\langle E(A_n)x, E(A_m)x \rangle = 0 \quad \text{if} \quad n \neq m.$$

Since $\langle E(\cdot)x, y \rangle$ is a complex measure by 5° in the definition, we obtain

$$\left\langle E\left(\bigcup_{n=1}^{\infty} A_n\right)x, y \right\rangle = \sum_{n=1}^{\infty} \langle E(A_n)x, y \rangle \quad \text{for all} \quad y \in \mathfrak{H}. \tag{7.19}$$

Hence, by (i)⇔(iii) in the general result,

$$\sum_{n=1}^{\infty} E(A_n)x \quad \text{(strong convergence)}$$

exists. We can rewrite the right-hand side of (7.19) as

---

[9] See Yosida [9] p. 69; Maruyama [7] p. 168.

$$\begin{aligned}
\sum_{n=1}^{\infty}\langle E(A_n)x, y\rangle &= \lim_{N\to\infty}\left\langle \sum_{n=1}^{N} E(A_n)x, y\right\rangle \\
&= \left\langle \lim_{N\to\infty}\sum_{n=1}^{N} E(A_n)x, y\right\rangle \quad \text{(by continuity of inner product)} \\
&= \left\langle \sum_{n=1}^{\infty} E(A_n)x, y\right\rangle \quad \text{(by strong convergence of the first term).}
\end{aligned}$$

Since

$$\left\langle E\left(\bigcup_{n=1}^{\infty} A_n\right)x, y\right\rangle = \left\langle \sum_{n=1}^{\infty} E(A_n)x, y\right\rangle$$

for any $y \in \mathfrak{H}$, we obtain

$$E\left(\bigcup_{n=1}^{\infty} A_n\right)x = \sum_{n=1}^{\infty} E(A_n)x.$$

This proves the σ-additivity of $E(\cdot)x$.

7.5 Spectral Representation of Unitary Operators

Let $\mathfrak{H}$ be a (complex) Hilbert space and $U \in \mathcal{L}(\mathfrak{H})$ a unitary operator. If we define a sequence $\{a_n\}_{n\in\mathbb{Z}}$ by

$$a_n = \langle U^n x, x \rangle, \quad n \in \mathbb{Z}$$

for a fixed $x \in \mathfrak{H}$, then it is positive semi-definite; that is, for any sequence $\{z_n\}_{n\in\mathbb{Z}}$ of complex numbers all of which are zero except finite elements,

$$\sum_{n,m} a_{n-m} z_n \bar{z}_m \geqq 0.$$

It can be verified by[10]

$$\sum_{n,m} a_{n-m} z_n \bar{z}_m = \sum_{n,m}\langle U^{n-m} x, x \rangle z_n \bar{z}_m$$

[10]Note that $U^* = U^{-1}$.

$$
\begin{aligned}
&= \sum_{n,m} \langle\, U^n x, U^m x \,\rangle z_n \bar{z}_m = \left\langle \sum_n z_n U^n x, \sum_m z_m U^m x \right\rangle \\
&= \left\| \sum_n z_n U^n x \right\|^2 \geqq 0.
\end{aligned}
$$

Hence, by Herglotz's theorem (Theorem 6.4, p. 133), a_n's can be represented as Fourier coefficients of some positive Radon measure $m_x \in \mathfrak{M}_+(\mathbb{T})$ on $\mathbb{T}$:

$$
a_n = \langle\, U^n x, x \,\rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} dm_x(\theta). \tag{7.20}
$$

We have to remark here that the measure m_x depends upon x. It is clear that

$$
\frac{1}{\sqrt{2\pi}} m_x(\mathbb{T}) = \|x\|^2 \tag{7.21}
$$

by setting $n = 0$.

We now try to represent $\langle\, U^n x, y \,\rangle$ in a similar way for fixed $x, y \in \mathfrak{H}$. As is well-known, $\langle\, U^n x, y \,\rangle$ can be expressed as a linear combination of[11]

$$
\langle\, U^n(x \pm y), x \pm y \,\rangle, \quad \langle\, U^n(x \pm iy), x \pm iy \,\rangle.
$$

Hence by (7.20), we obtain a representation

$$
\langle\, U^n x, y \,\rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} dm_{x,y}(\theta), \tag{7.22}
$$

$$
\text{where} \quad 4m_{x,y} = m_{x+y} - m_{x-y} + i m_{x+iy} - i m_{x-iy}.
$$

Needless to say, $m_{x,x} = m_x$.

Theorem 7.9 *Let $U \in \mathfrak{L}(\mathfrak{H})$ be a unitary operator. Then there exists a Borel complex measure $m_{x,y}$ for any fixed $x, y \in \mathfrak{H}$, such that*

$$
\langle U^n x, y \rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} dm_{x,y}(\theta).
$$

Such a measure $m_{x,y}$ is determined uniquely.

Theorem 7.10 *The measure $m_{x,y}$ obtained above has the following properties for any $S \in \mathcal{B}(\mathbb{T})$:*

(i) $(x, y) \mapsto m_{x,y}(S)$ *is sesquilinear.*
(ii) $(x, y) \mapsto m_{x,y}(S)$ *is skew-symmetric.*

---

[11] $\langle\, U^n x, y \,\rangle = \frac{1}{4}\{\langle\, U^n(x+y), x+y \,\rangle - \langle\, U^n(x-y), x-y \,\rangle + i\langle\, U^n(x+iy), x+iy \,\rangle - i\langle\, U^n(x-iy), x-iy \,\rangle\}$.

(iii) $\frac{1}{\sqrt{2\pi}} |m_{x,y}(S)| \leqq \|x\| \cdot \|y\|$ *for any* $S \in \mathcal{B}(\mathbb{T})$.

Proof (i) For $\alpha, \beta \in \mathbb{C}$ and $x_1, x_2 \in \mathfrak{H}$,

$$
\begin{aligned}
\langle U^n(\alpha x_1 + \beta x_2), y \rangle = \alpha \langle U^n x_1, y \rangle + \beta \langle U^n x_2, y \rangle \\
&= \frac{\alpha}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} dm_{x_1,y} + \frac{\beta}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} dm_{x_2,y} \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} d(\alpha m_{x_1,y} + \beta m_{x_2,y}),
\end{aligned}
$$

which implies

$$m_{\alpha x_1 + \beta x_2, y} = \alpha m_{x_1,y} + \beta m_{x_2,y}.$$

Similarly, we have

$$m_{x, \alpha y_1 + \beta y_2} = \bar{\alpha} m_{x,y_1} + \bar{\beta} m_{x,y_2}.$$

(ii) By a simple computation,

$$
\begin{aligned}
\langle U^n y, x \rangle = \langle y, U^{-n} x \rangle = \overline{\langle U^{-n} x, y \rangle} \\
&= \overline{\frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{in\theta} dm_{x,y}} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} \overline{dm_{x,y}}.
\end{aligned}
$$

Similarly,

$$\langle U^n y, x \rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} dm_{y,x}.$$

Therefore $\overline{m_{x,y}} = m_{y,x}$.

(iii) Define a function $\varphi_S : \mathfrak{H} \times \mathfrak{H} \to \mathbb{C}$ by

$$\varphi_S(x, y) = m_{x,y}(S)$$

for fixed $S \in \mathcal{B}(\mathbb{T})$. φ_S is a skew-symmetric, sesquilinear functional as we have already checked. It is also clear that

$$\varphi_S(x, x) = m_{x,x}(S) = m_x(S) \geqq 0$$

and $\varphi_S(0, 0) = 0$ for $x = 0$. By Schwarz's inequality,[12]

[12] φ_S is not necessarily an inner product because "$x \neq 0 \Rightarrow \varphi_S(x, x) > 0$" is not fulfilled in general. But all the requirements other than this are satisfied. Hence we can apply Schwarz's inequality here. The condition "$x \neq 0 \Rightarrow \langle x, x \rangle \neq 0$" is not necessary for Schwarz's inequality.

$$|m_{x,y}(S)| \leqq m_x(S)^{1/2} \cdot m_y(S)^{1/2}. \tag{7.23}$$

The inequality (7.23) implies

$$\frac{1}{\sqrt{2\pi}} |m_{x,y}(S)| \leqq \|x\| \cdot \|y\|,$$

since $\frac{1}{\sqrt{2\pi}} m_x(\mathbb{T}) = \|x\|^2$, $\frac{1}{\sqrt{2\pi}} m_y(\mathbb{T}) = \|y\|^2$ (by (7.21)) and $m_x(\cdot)$ as well as $m_y(\cdot)$ are measures. □

Thus $\varphi_S(x, y)$ is a skew-symmetric sesquilinear functional on $\mathfrak{H} \times \mathfrak{H}$ which is bounded in the sense of

$$|\varphi_S(x, y)| \leqq \sqrt{2\pi} \|x\| \cdot \|y\|. \tag{7.24}$$

Thanks to Corollary 7.1, we obtain the next theorem.[13]

Theorem 7.11 *There exists a symmetric operator $E(S) \in \mathcal{L}(\mathfrak{H})$, for each $S \in \mathcal{B}(\mathbb{R})$, such that*

$$\frac{1}{\sqrt{2\pi}} m_{x,y}(S) = \langle E(S)x, y \rangle, \quad x, y \in \mathfrak{H}.$$

It is determined uniquely and satisfies $\|E(S)\| = \frac{1}{\sqrt{2\pi}} \|\varphi_S\|$.

Theorem 7.12 *The operator $E : \mathcal{B}(\mathbb{R}) \to \mathcal{L}(\mathfrak{H})$ is a resolution of identity on $\mathcal{B}(\mathbb{R})$.*

Proof We have to show 1°–5° on p. 177.

1° Since

$$|m_{x,y}(\emptyset)| = |\varphi_\emptyset(x, y)| = 0 \quad \text{for all} \quad x, y \in \mathfrak{H}$$

by (7.23), it holds true that $E(\emptyset) = 0$. Setting $n = 0$ in (7.22), we have

$$\langle x, y \rangle = \frac{1}{\sqrt{2\pi}} m_{x,y}(\mathbb{T}) = \langle E(\mathbb{T})x, y \rangle \quad \text{for all} \quad x, y \in \mathfrak{H}.$$

Hence $E(\mathbb{T}) = I$.

[13] We owe this to Lax [6] Chap. 31, especially Section 31.7.

2° $E(S)$ is symmetric by Theorem 7.11. Applying 3° (to be shown below) to the special case $S = T$, we obtain

$$E(S) = E(S)^2.$$

So $E(S)$ is a projection by Theorem 7.6.

3° Taking account of the relation between $m_{x,y}$ and $\langle E(\cdot)x, y \rangle$, we have

$$\langle U^n x, y \rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} dm_{x,y}(\theta) = \int_{\mathbb{T}} e^{-in\theta} \langle E(d\theta)x, y \rangle. \tag{7.25}$$

Substituting n by $n + k$, we also have

$$\langle U^{n+k} x, y \rangle = \int_{\mathbb{T}} e^{-in\theta} e^{-ik\theta} \langle E(d\theta)x, y \rangle. \tag{7.26}$$

Replacing x in (7.25) by $U^k x$, we further obtain

$$\langle U^{n+k} x, y \rangle = \int_{\mathbb{T}} e^{-in\theta} \langle E(d\theta) U^k x, y \rangle. \tag{7.27}$$

The comparison of (7.26) and (7.27) gives[14]

$$e^{-ik\theta} \langle E(d\theta)x, y \rangle = \langle E(d\theta) U^k x, y \rangle. \tag{7.28}$$

χ_S denotes the characteristic function of $S \in \mathcal{B}(\mathbb{R})$. Then it follows from (7.25) ($n = 0$) that

$$\begin{aligned}
\int_{\mathbb{T}} \chi_S(\theta) e^{-ik\theta} \langle E(d\theta)x, y \rangle = \langle E(S) U^k x, y \rangle. \\
= \langle U^k x, E(S) y \rangle \quad (\text{by } E(S) = E(S)^*) \\
\underset{(7.25)}{=} \int_{\mathbb{T}} e^{-ik\theta} \langle E(d\theta)x, E(S) y \rangle.
\end{aligned} \tag{7.29}$$

Since the first and the fourth terms in (7.29) are equal,

$$\chi_S(\theta) \langle E(d\theta)x, y \rangle = \langle E(d\theta)x, E(S) y \rangle.$$

[14] If the Fourier transforms of two Radon measures on $\mathbb{T}$ are equal, then the two measures are coincide.

Integrating both sides on $T \in \mathcal{B}(\mathbb{T})$,[15] we have

$$\int_{\mathbb{T}} \chi_T(\theta)\chi_S(\theta)\langle\, E(d\theta)x,\, y\,\rangle = \langle\, E(T)x,\, E(S)y\,\rangle.$$

The left-hand side is equal to $\langle\, E(S \cap T)x,\, y\,\rangle$, and so

$$\langle\, E(S \cap T)x,\, y\,\rangle = \langle\, E(T)x,\, E(S)y\,\rangle = \langle\, E(S)E(T)x,\, y\,\rangle.$$

This proves $E(S \cap T) = E(S)E(T)$.

4° Assume $S \cap T = \emptyset$. Then

$$m_{x,y}(S \cup T) = \langle\, E(S \cup T)x,\, y\,\rangle. \tag{7.30}$$

On the other hand,

$$\begin{aligned} m_{x,y}(S \cup T) &= m_{x,y}(S) + m_{x,y}(T) = \langle\, E(S)x,\, y\,\rangle + \langle\, E(T)x,\, y\,\rangle \\ &= \langle\, (E(S) + E(T))x,\, y\,\rangle. \end{aligned} \tag{7.31}$$

Comparing (7.30) and (7.31), we obtain

$$E(S \cup T) = E(S) + E(T).$$

5° It is easily observed that $m_{x,y}(\cdot) = \langle\, E(\cdot)x,\, y\,\rangle$ is a complex measure by its construction. □

In addition to 1°–5°, $E(\cdot)$ has some more properties.

6° $E(S)U = UE(S)$.

In fact,

$$\langle\, U^n(Ux),\, y\,\rangle = \langle\, U^n x,\, U^{-1}y\,\rangle$$

since U is unitary. Hence

$$m_{Ux,y} = m_{x,U^{-1}y}. \tag{7.32}$$

Both sides of (7.32) can be rewritten as

$$\begin{aligned} m_{Ux,y}(S) &= \langle\, E(S)Ux,\, y\,\rangle, \\ m_{x,U^{-1}y}(S) &= \langle\, E(S)x,\, U^{-1}y\,\rangle = \langle\, UE(S)x,\, y\,\rangle. \end{aligned} \tag{7.33}$$

[15] $\chi_T\chi_S = \chi_{S\cap T}$.

The relations (7.32) and (7.33) imply $E(S)U = UE(S)$.

Needless to say, all the properties corresponding to Theorem 7.8 are satisfied. That is, the following statements hold good for $S, T \in \mathcal{B}(\mathbb{T})$ and $x \in \mathfrak{H}$.

7° $m_{x,x}(S) = \langle E(S)x, x \rangle = \|E(S)x\|^2$.

8° $m_{x,x}(\cdot)$ is a positive measure on $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ and

$$m_{x,x}(\mathbb{T}) = \|x\|^2.$$

9° $E(S)E(T) = E(T)E(S)$.

10° If $S \cap T = \emptyset$, the images of $E(S)$ and $E(T)$ are orthogonal.

$\langle U^n x, y \rangle$ can be expressed in terms of $E(\cdot)$ in the form

$$\langle U^n x, y \rangle = \int_{\mathbb{T}} e^{-in\theta} \langle E(d\theta)x, y \rangle = \Big\langle \underbrace{\int_{\mathbb{T}} e^{-in\theta} E(d\theta)\ x}_{\substack{\text{integration with}\\ \text{respect to the}\\ \text{vector measure} E(\cdot)x}}, y \Big\rangle.$$

We may write this relation symbolically as

$$U^n = \int_{\mathbb{T}} e^{-in\theta} dE.$$

In the special case of $n = 1$, we have

$$U = \int_{\mathbb{T}} e^{-i\theta} dE.$$

This is the **spectral representation** of U.

Remark 7.6 The above result is based upon the observation that the numerical sequence $\{a_n = \langle U^n x, x\rangle\}$ defined by the corresponding composition $U^n (n \in \mathbb{Z})$ of a unitary operator U is positive semi-definite. By Herglotz's theorem, a_n's can be represented as the Fourier coefficients of a positive Radon measure on $\mathbb{T}$.

We can generalize this result to the family $\{U_n | \ n \in \mathbb{Z}\}$ of unitary operators with parameter n which satisfies

$$U_n U_m = U_{n+m} \quad (n, m \in \mathbb{Z}), \qquad T_0 = I.$$

(The family $\{U^n | \ n \in \mathbb{Z}\}$ of compositions is its special case.) In this case, too, a quite similar representation is possible. That is, there exists uniquely some bounded symmetric operator $E(S)$, for each $S \in \mathcal{B}(\mathbb{T})$, which satisfies

$$U_n = \int_{\mathbb{T}} e^{-in\theta} dE(\theta), \qquad n \in \mathbb{Z}.$$

In the next section, we consider the representation problem of the group of unitary operators where the parameter space is $\mathbb{R}$ rather than $\mathbb{Z}$. It may be naturally conjectured that Bochner's theorem plays a key role instead of Herglotz's theorem.

7.6 One-Parameter Group of Unitary Operators and its Spectral Representation: Stone's Theorem

Definition 7.6 A family $\{T_t | t \in \mathbb{R}\}$ of bounded linear operators on a Hilbert space $\mathfrak{H}$ is called a **one-parameter group** of linear operators[16] if

$$T_t T_s = T_{t+s} \quad (t, s \in \mathbb{R}), \quad T_0 = I.$$

A one-parameter group $\{T_t\}$ is said to be of **class-C_0** if

$$s\text{-}\lim_{t \to t_0} T_t x = T_{t_0} x \quad \text{for all} \quad x \in \mathfrak{H}$$

for each $t_0 \in \mathbb{R}$.[17]

We now consider the special case of a one-parameter group consisting of unitary operators.[18]

Lemma 7.1 *Let $\{U_t | t \in \mathbb{R}\}$ be a one-parameter group of class-C_0 consisting of unitary operators. A function $\varphi : \mathbb{R} \to \mathbb{C}$ is defined by*

$$\varphi(t) = \langle U_t x, x \rangle.$$

Then φ enjoys the following properties:

(i) *φ is continuous.*
(ii) $\varphi(-t) = \overline{\varphi(t)}$.
(iii) $\displaystyle\sum_{i,j=1}^{n} \varphi(t_i - t_j) z_i \bar{z}_j \geqq 0$

for any finite $t_1, t_2, \cdots, t_n \in \mathbb{R}$ and $z_1, z_2, \cdots, z_n \in \mathbb{C}$.

Proof The continuity of φ is clear. We also note that

$$U_{-t} = U_t^{-1} = U_t^*, \tag{7.34}$$

[16] If t is restricted to $[0, \infty)$, $\{T_t \mid t \geqq 0\}$ is called a **semi-group**.

[17] A one-parameter group $\{T_t\}$ of class-C_0 is also said to be **strongly continuous**.

[18] This theory was established by Stone [8]. I benefited much from Lax [6] Chap. 35, as well as Dunford [3].

since $\{U_t | t \in \mathbb{R}\}$ is a one-parameter group. By (7.32),

$$\varphi(-t) = \langle U_{-t}x, x \rangle = \langle U_t^* x, x \rangle = \overline{\langle x, U_t^* x \rangle} = \overline{\langle U_t x, x \rangle} = \overline{\varphi(t)}.$$

Finally, the positive semi-definiteness follows from a simple computation:

$$\sum \varphi(t_i - t_j) z_i \bar{z}_j = \sum \langle U_{t_i - t_j} x, x \rangle z_i \bar{z}_j = \sum \langle U_{t_j}^{-1} U_{t_i} x, x \rangle z_i \bar{z}_j$$

$$= \sum \langle U_{t_i} x, U_{t_j} x \rangle z_i \bar{z}_j = \left\langle \sum z_i U_{t_i} x, \sum z_j U_{t_j} x \right\rangle = \left\| \sum z_i U_{t_i} x \right\|^2 \geqq 0$$

for $t_1, t_2, \cdots, t_n \in \mathbb{R}$ and $z_1, z_2, \cdots, z_n \in \mathbb{C}$. □

We apply Bochner's theorem (Theorem 6.11, p. 150) to the function $\varphi(t)$ to obtain its representation as the Fourier transform of a positive Radon measure $m(\cdot\,;\ x)$ on $\mathbb{R}$. That is, there exists a measure $m(\cdot\,;\ x) \in \mathfrak{M}_+(\mathbb{R})$ such that

$$\varphi(t) = \langle U_t x, x \rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-it\lambda} dm(\lambda;\ x).$$

For fixed $x, y \in \mathfrak{H}$, $\langle U_t x, y \rangle$ is expressed as a linear combination of

$$\langle U_t(x \pm y), x \pm y \rangle, \quad \langle U_t(x \pm iy), x \pm iy \rangle.$$

Hence there exists some (complex) measure $m(\cdot\,;\ x, y)$ which satisfies

$$\langle U_t x, y \rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-it\lambda} dm(\lambda;\ x, y). \tag{7.35}$$

Theorem 7.13 *Let $\{U_t \mid t \in \mathbb{R}\}$ be a one-parameter group of class-C_0 consisting of unitary operators. Then there exists a Borel complex measure $m(\cdot\,;\ x, y)$ for each $x, y \in \mathfrak{H}$ such that*

$$\langle U_t x, y \rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-it\lambda} dm(\lambda;\ x, y). \tag{7.36}$$

Such a measure $m(\cdot\,;\ x, y)$ is unique.

The properties of $m(\cdot\,;\ x, y)$ are summarized as follows. We can prove it in a similar way to Theorem 7.10.

Theorem 7.14 *$m(\cdot\,;\ x, y)$ has the following properties:*

(i) *$(x, y) \mapsto m(\cdot\,;\ x, y)$ is a sesquilinear functional.*
(ii) *$(x, y) \mapsto m(\cdot\,;\ x, y)$ is skew-symmetric.*
(iii) *$\frac{1}{\sqrt{2\pi}} |m(S;\ x, y)| \leqq \|x\| \cdot \|y\|$ for any $S \in \mathcal{B}(\mathbb{R})$.*

It is obvious that $m(\cdot\,;\ x, x) = m(\cdot\,;\ x)$.

Thus if we define a function $\Phi_S : \mathfrak{H} \times \mathfrak{H} \to \mathbb{C}$ $(S \in \mathcal{B}(\mathbb{R}))$ by

$$\Phi_S : (x, y) \mapsto \frac{1}{\sqrt{2\pi}} m(S;\ x, y),$$

it is a skew-symmetric sesquilinear functional with

$$\|\Phi_S(x, y)\| \leqq \|x\| \cdot \|y\|.$$

Consequently, by Corollary 7.1, there exists, for each $S \in \mathcal{B}(\mathbb{R})$, a bounded symmetric operator $E(S) \in \mathcal{L}(\mathfrak{H})$ which satisfies

$$\Phi_S(x, y) = \frac{1}{\sqrt{2\pi}} m(S;\ x, y) = \langle E(S)x, y \rangle, \quad x, y \in \mathfrak{H}. \tag{7.37}$$

$E(S)$ is uniquely determined and $\|E(S)\| = \|\Phi_S\|$.

By Theorem 7.13 and (7.37), (7.36) can be reexpressed eventually in the form

$$\langle U_t x, y \rangle = \int_{\mathbb{R}} e^{-it\lambda} \langle E(d\lambda)x, y \rangle.$$

This is the final form of Stone's theorem.

Theorem 7.15 (Stone) *Let $\{U_t | t \in \mathbb{R}\}$ be a one-parameter group of class-C_0 consisting of unitary operators. Then there exists, for each $S \in \mathcal{B}(\mathbb{R})$, a unique bounded symmetric operator $E(S) \in \mathcal{L}(\mathfrak{H})$ which satisfies*

$$\langle U_t x, y \rangle = \int_{\mathbb{R}} e^{-it\lambda} \langle E(d\lambda)x, y \rangle.$$

The proof of Stone's theorem explained here is based upon Bochner's theorem. However, it is remarkable that Bochner's theorem can be deduced from Stone's theorem, conversely. That is, these two theorems are mathematically equivalent.[19]

Let $\varphi \in \mathfrak{C}^b(\mathbb{R}, \mathbb{C})$ be positive semi-definite. We denote by $\mathfrak{H}$ the set of functions of $\mathbb{R}$ into $\mathbb{C}$, the values of which are zero except some finite points. $\mathfrak{H}$ is a vector space with usual operations and an inner product $\langle \cdot, \cdot \rangle$[20]:

$$\begin{aligned}
(x + y)(t) &= x(t) + y(t), \\
(\alpha x)(t) &= \alpha \cdot x(t), \\
\langle x, y \rangle &= \sum_{t,s \in \mathbb{R}} \varphi(t - s) x(t) \overline{y(s)}.
\end{aligned}$$

[19] We owe this to Yosida [9] pp. 346–347.

[20] It is not necessarily true that $\langle x, x \rangle = 0 \Rightarrow x = 0$. $\langle x, x \rangle \geqq 0$ follows from the positive semi-definiteness of φ.

$\mathfrak{H}$ is a pre-Hilbert space under these operations. We define a subspace[21] $\mathfrak{N}$ of $\mathfrak{H}$ by

$$\mathfrak{N} = \{x \in \mathfrak{H} | \langle x, x \rangle = 0\}.$$

Let ξ_x, ξ_y be elements of the quotient space $\mathfrak{H}/\mathfrak{N}$ which contain x and y, respectively. $\mathfrak{H}/\mathfrak{N}$ becomes a pre-Hilbert space under the inner product

$$\langle \xi_x, \xi_y \rangle = \langle x, y \rangle.$$

The completion $\tilde{\mathfrak{H}}$ of $\mathfrak{H}/\mathfrak{N}$ is a Hilbert space.

It is clear that the operator $U_\tau : \mathfrak{H} \to \mathfrak{H}$ $(\tau \in \mathbb{R})$ defined by

$$(U_\tau x)(t) = x(t - \tau), \quad x \in \mathfrak{H}$$

satisfies

$$\langle U_\tau x, U_\tau y \rangle = \langle x, y \rangle, \quad x, y \in \mathfrak{H},$$
$$U_\tau U_\sigma = U_{\tau+\sigma}, \quad \tau, \sigma \in \mathbb{R}, \quad U_0 = I.$$

Hence if we define another operator $\tilde{U}_\tau; \tilde{\mathfrak{H}} \to \tilde{\mathfrak{H}}(\tau \in \mathbb{R})$ by

$$(\tilde{U}_\tau \xi_x)(t) = (U_\tau x)(t),$$

$\{\tilde{U}_\tau\}$ is a one-parameter group of unitary operators on $\tilde{\mathfrak{H}}$; and it is of class-C_0 by the continuity of φ.

Therefore, by Stone's theorem, there exists, for each $E \in \mathcal{B}(\mathbb{R})$, a symmetric operator $E(S)$ such that

$$\langle \tilde{U}_\tau \xi_x, \xi_y \rangle = \langle U_\tau x, y \rangle = \int_{\mathbb{R}} e^{-i\tau\lambda} \langle E(d\lambda)x, y \rangle.$$

If we define a complex measure $m(\cdot\,;\ x, y)$ by

$$m(S;\ x, y) = \sqrt{2\pi} \langle E(S)x, y \rangle,$$

$m(\cdot\,;\ x, x)$ is real by Theorem 7.14.

[21] For any two elements x and y of $\mathfrak{N}$, we must have $x + y \in \mathfrak{N}$. It can be shown as follows. By the positive semi-definiteness of φ, $\varphi(-u) = \overline{\varphi(u)}$ (Lemma 6.3, p. 131). And we may apply Schwarz's inequality to $\langle \cdot, \cdot \rangle$. Since $\langle x, x \rangle = \langle y, y \rangle = 0$, we have

$$\langle x + y, x + y \rangle = 2\mathcal{R}e\langle x, y \rangle.$$

Hence by Schwarz's inequality,

$$|\langle x + y, x + y \rangle| \leqq 2\langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2} = 0.$$

$x + y \in \mathfrak{N}$ immediately follows.

Specifying $x_0 \in \mathfrak{H}$ by

$$x_0(t) = \begin{cases} 1 & \text{if} \quad t = \tau, \\ 0 & \text{if} \quad t \neq \tau, \end{cases}$$

we obtain

$$\langle U_\tau x_0, x_0 \rangle = \sum_{t,s \in \mathbb{R}} \varphi(t-s) U_\tau x_0(t) \overline{x_0(s)} = \sum_{t,s \in \mathbb{R}} \varphi(t-s) x_0(t-\tau) \overline{x_0(s)}$$

by definition of $\langle \cdot, \cdot \rangle$. Since $x_0(t-\tau) \neq 0$ and $x_0(s) \neq 0$ only when $t = 2\tau$ and $s = \tau$, it follows that

$$\langle U_\tau x_0, x_0 \rangle = \varphi(\tau).$$

Consequently,

$$\varphi(\tau) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\tau\lambda} dm(\lambda;\ x_0, x_0).$$

This proves Bochner's theorem.

References

1. Arveson, W.: An Invitation to C^*-Algebra. Springer, Berlin (1976)
2. Dixmier, J.: C^*-Algebras. North-Holland, Amsterdam (1977)
3. Dunford, N.: On one-parameter groups of linear transformations. Ann. Math. **39**, 305–356 (1938)
4. Evans, L.: Partial Differential Equations. American Mathematical Society, Providence (1998)
5. Kato, T.: Iso Kaiseki (Functional Analysis). Kyoritsu Shuppan, Tokyo (1957) (Originally published in Japanese)
6. Lax, P.D.: Functional Analysis. Wiley, New York (2002)
7. Maruyama, T.: Kansu Kaisekigaku (Functional Analysis). Keio Tsushin, Tokyo (1980) (Originally published in Japanese)
8. Stone, M.H.: On one-parameter unitary groups in Hilbert space. Ann. Math. **33**, 643–648 (1932)
9. Yosida, K.: Functional Analysis, 3rd edn. Springer, New York (1971)

Chapter 8
Fourier Analysis of Periodic Weakly Stationary Processes

During the decade around 1930, the world economy was thrown into a serious depression that nobody had previously experienced.

A number of economic theories were proposed in order to analyze and control the violence of business fluctuations.

We enumerate, say, the interaction theory of accelerator and multiplier by P.A. Samuelson and J.R. Hicks, the nonlinear oscillation model by N. Kaldor and R. Goodwin, the analysis based upon the differential equation with delay by R. Frisch, and so on. Furthermore, we should remember that J.M. Keynes' innovative macroeconomic theory paved the way leading to dynamic theory of economic fluctuations.

From a mathematical point of view, the work of E. Slutsky [29], a Ukrainian mathematician, is particularly remarkable.[1] It was published in *Econometrica*, 1937. He tried to explain more or less regular fluctuations of macro-economic movements based upon the overlapping effects of random shocks. However, frankly speaking, his analysis was rather experimental and devoid of mathematical rigor.[2]

In this chapter and the next, we investigate periodic and almost periodic behaviors of stationary stochastic processes and give a systematic exposition of the mathematical skeleton of the Slutsky problem from the viewpoint of classical Fourier analysis. The Bochner–Herglotz theory concerning Fourier representation of positive semi-definite functions provides the key analytical tool.[3]

---

[1] R. Frisch [6] also deserves special attention.

[2] See Maruyama [22] for the outline of the theories of business fluctuations. See also Samuelson [24, 25], Hicks [9], Kaldor [11], R. Frisch [6], Keynes [18], Slutsky [28–30].

[3] Kawata [16, 17], Maruyama [20] and Wold [31] are classical works on Fourier analysis of stationary stochastic processes, which provided me with all the basic mathematical background. Among more recent literature, I wish to mention Brémaud [1]. Granger and Newbold [7] Chap. 2, Hamilton [8] Chap. 3 and Sargent [26] Chap. XI are textbooks written from the standpoint of economics.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_8

8.1 Stochastic Processes of Second Order

Let $(\Omega, \mathcal{E}, P)$ be a probability space, and T a subset of $\mathbb{R}$. T is usually interpreted as the space of time and, in this chapter, T is assumed to be either $\mathbb{R}$ or $\mathbb{Z}$. A function $X : T \times \Omega \to \mathbb{C}$ is said to be a stochastic process if the function $\omega \mapsto X(t, \omega)$ is $(\mathcal{E}, \mathcal{B}(\mathbb{C}))$-measurable for any fixed $t \in T$.

The trajectory of $X(t, \omega)$ for a fixed ω, that is, the function $t \mapsto X(t, \omega)$, is called the **sample function** of this stochastic process.

Let $\mathcal{T}$ be the set of all the finite tuples of elements of T, that is

$$\mathcal{T} = \{\mathbf{t} = (t_1, t_2, \cdots, t_n) \mid t_j \in T,\ \ j = 1, 2, \cdots, n,\ \ n \in \mathbb{N}\}.$$

$X_{\mathbf{t}}(\omega)$ denotes the vector

$$X_{\mathbf{t}}(\omega) = (X(t_1, \omega), X(t_2, \omega), \cdots, X(t_n, \omega)), \quad \mathbf{t} \in \mathcal{T}.$$

The set function $\nu_{X_{\mathbf{t}}} : \mathcal{B}(\mathbb{C}^n) \to \mathbb{R}$ defined by

$$\nu_{X_{\mathbf{t}}}(E) = P\{\omega \in \Omega | X_{\mathbf{t}}(\omega) \in E\}, \quad E \in \mathcal{B}(\mathbb{C}^n)$$

is a measure on $\mathcal{B}(\mathbb{C}^n)$, called the distribution of $X_{\mathbf{t}}(\omega)$. The set of all the distributions $\{\nu_{X_{\mathbf{t}}} \,|\, \mathbf{t} \in \mathcal{T}\}$ is called the system of finite dimensional distributions determined by $X(t, \omega)$.

$\{\nu_{X_{\mathbf{t}}} \,|\, \mathbf{t} \in \mathcal{T}\}$ is determined by a given stochastic process $X(t, \omega)$. Conversely, assume now that a probability measure $\nu_{\mathbf{t}}$ on $(\mathbb{C}^n, \mathcal{B}(\mathbb{C}^n))$ is given for each $\mathbf{t} = (t_1, t_2, \cdots, t_n) \in \mathcal{T}$. The set of all of them is $\{\nu_{\mathbf{t}} \,|\, \mathbf{t} \in \mathcal{T}\}$. Does there exist a stochastic process, the system of finite dimensional distributions of which is $\{\nu_{\mathbf{t}} \,|\, \mathbf{t} \in \mathcal{T}\}$? The positive answer to this question is given by the famous theorem due to A.N. Kolmogorov.[4]

Kolmogorov's Theorem *Assume that a family $\{\nu_{\mathbf{t}} \,|\, \mathbf{t} \in \mathcal{T}\}$ of probability measures satisfies the following, (i) and (ii). Then there exists a stochastic process $X : T \times \Omega \to \mathbb{C}$ (or $\mathbb{R}$), the system of finite dimensional distributions of which is $\{\nu_{\mathbf{t}} \,|\, \mathbf{t} \in \mathcal{T}\}$.*

(i) *If $\mathbf{t}_1 = (t_1, t_2, \cdots, t_n)$ and $\mathbf{t}_2 = (t_1, t_2, \cdots, t_n, t_{n+1})$,*

$$\nu_{\mathbf{t}_1}(E) = \nu_{\mathbf{t}_2}(E \times \mathbb{C}) \quad \textit{for} \quad E \in \mathcal{B}(\mathbb{C}^n).$$

(ii) *Let π be a permutation of n letters, $\mathbf{t} = (t_1, t_2, \cdots, t_n) \in \mathcal{T}$, $\mathbf{s} = (s_1, s_2, \cdots s_n) = \pi(\mathbf{t})$. Then*

[4] See Kawata [17] pp. 5–7, for the proof. Loève [19] pp. 92–94 and Dudley [5] pp. 346–349 are also suggestive.

$$\nu_{\mathbf{t}}(E) = \nu_{\mathbf{s}}(\pi(E)) \quad \textit{for} \quad E \in \mathcal{B}(\mathbb{C}^n),$$

where $\pi(E) = \{(x_{s_1}, \cdots, x_{s_n}) \mid (x_{t_1}, \cdots, x_{t_n}) \in E\}$.

From now on, the space T of time is assumed to be $\mathbb{R}$, unless otherwise mentioned.

A stochastic process $X(t, \omega)$ is said to be of the ***p*-th order** if

$$\mathbb{E}|X(t, \omega)|^p = \int_\Omega |X(t, \omega)|^p dP(\omega) < \infty,$$

where $\mathbb{E}$ is the expectation operator. In this chapter, the special case of $p = 2$ is our main concern.

Let $X : T \times \Omega \to \mathbb{C}$ be a stochastic process of the second order. The following two functions play a crucial role:

$$R : T \times T \to \mathbb{C}, \quad \rho : T \times T \to \mathbb{C}.$$

$R(s, t)$ is defined by

$$R(s, t) = \mathbb{E}X(s, \omega)\overline{X(t, \omega)}, \quad s, t \in T$$

and is called the **correlation function** of X. $\rho(s, t)$ is defined by

$$\rho(s, t) = \mathbb{E}[X(s, \omega) - \mathbb{E}X(s, \omega)][\overline{X(t, \omega) - \mathbb{E}X(t, \omega)}], \quad s, t \in T$$

and is called the **covariance function** of X.

We list several basic properties of covariance function $\rho(s, t)$ of a stochastic process $X(t, \omega)$:

1° $\rho(s, t) = R(s, t) - \mathbb{E}X(s, \omega)\mathbb{E}\overline{X(t, \omega)}$.

Proof

$$\begin{aligned}
\rho(s, t) &= \mathbb{E}[X(s, \omega)\overline{X(t, \omega)} - X(s, \omega)\mathbb{E}\overline{X(t, \omega)} - \overline{X(t, \omega)}\mathbb{E}X(s, \omega) \\
&\qquad + \mathbb{E}X(s, \omega) \cdot \mathbb{E}\overline{X(t, \omega)}] \\
&= \mathbb{E}[X(s, \omega)\overline{X(t, \omega)}] - \mathbb{E}X(s, \omega) \cdot \mathbb{E}\overline{X(t, \omega)} \\
&\qquad\qquad - \mathbb{E}\overline{X(t, \omega)} \cdot \mathbb{E}X(s, \omega) + \mathbb{E}X(s, \omega) \cdot \mathbb{E}\overline{X(t, \omega)} \\
&= R(s, t) - \mathbb{E}X(s, \omega) \cdot \mathbb{E}\overline{X(t, \omega)}.
\end{aligned}$$

□

2° $\rho(s, t) = \overline{\rho(t, s)}$.
3° $\rho(t, t) = \mathbb{E}|X(t, \omega) - \mathbb{E}X(t, \omega)|^2 \geqq 0$.
$\rho(t, t)$ is nothing other than the **variance** of $X(t, \omega)$.

4° $|\rho(s,t)|^2 \leqq \rho(s,s)\cdot\rho(t,t)$.

Proof By Schwarz's inequality,

$$|\rho(s,t)|^2 \leqq \mathbb{E}|X(s,\omega)-\mathbb{E}X(s,\omega)|^2 \times \mathbb{E}|X(t,\omega)-\mathbb{E}X(t,\omega)|^2 = \rho(s,s)\cdot\rho(t,t).$$

□

5° The $(n \times n)$-matrix

$$M = (\rho(t_i,t_j))_{1\leqq i,j\leqq n}$$

defined for any $t_i, t_j \in \mathbb{R}$ $(i, j = 1, 2, \cdots, n)$ and $n \in \mathbb{N}$ is an Hermite matrix which determines an Hermitian form of positive semi-definite.

Proof It is obvious that M is Hermite by **2°**. M determines an Hermitian form of positive semi-definite, since

$$\begin{aligned}\sum_{i,j=1}^{n}\rho(t_i,t_j)z_i\bar{z}_j &= \mathbb{E}\left\{\sum_{i,j=1}^{n}[X(t_i,\omega)-\mathbb{E}X(t_i,\omega)]\times[\overline{X(t_j,\omega)-\mathbb{E}X(t_j,\omega)}]z_i\bar{z}_j\right\}\\ &= \mathbb{E}\left|\sum_{i=1}^{n}[X(t_i,\omega)-\mathbb{E}X(t_i,\omega)]z_i\right|^2 \geqq 0.\end{aligned}$$

□

Let $X : T\times\Omega \to \mathbb{C}(T=\mathbb{R})$ be a stochastic process of the second order. When we regard X as a function of ω, fixing $t \in T$, we sometimes write it as $X_t(\omega)$.

Definition 8.1 If a function $A : \mathbb{R} \to \mathfrak{L}^2(\Omega, \mathbb{C})$ defined by

$$A : t \mapsto X(t,\omega)$$

is continuous at a point $t_0 \in \mathbb{R}$; i.e.

$$\mathbb{E}|X(t,\omega)-X(t_0,\omega)|^2 \to 0 \quad \text{as} \quad t \to t_0,$$

we say that $X(t,\omega)$ is **strongly continuous** at t_0. If $X(t,\omega)$ is strongly continuous at every $t \in \mathbb{R}$, $X(t,\omega)$ is said to be strongly continuous on $\mathbb{R}$.

In other words, $X(t,\omega)$ is strongly continuous at t_0 if there exists some $\delta > 0$, for each $\varepsilon > 0$, such that

$$\mathbb{E}|X(t,\omega)-X(t_0,\omega)|^2 < \varepsilon$$

if $|t-t_0| < \delta$. Even if $X(t,\omega)$ is strongly continuous on $\mathbb{R}$, the magnitude of δ depends upon ε and t, in general. If δ depends upon ε but is independent of t, we say that $X(t,\omega)$ is **uniformly strongly continuous**.

1° If $X(t,\omega)$ is strongly continuous, the function $t \mapsto \mathbb{E}X(t,\omega)$ is continuous.

Proof At any fixed $t_0 \in T$, we have

$$|\,\mathbb{E}X(t,\omega) - \mathbb{E}X(t_0,\omega)\,|^2 = |\,\mathbb{E}[X(t,\omega) - X(t_0,\omega)]\cdot 1\,|^2$$
$$\leqq \mathbb{E}\,|\,X(t,\omega) - X(t_0,\omega)\,|^2 \cdot \mathbb{E}(1)^2 \quad \text{(Schwarz's inequality)}$$
$$= \mathbb{E}\,|\,X(t,\omega) - X(t_0,\omega)\,|^2 \to 0 \quad \text{as} \quad t \to t_0.$$

□

2° $R(s,t)$ is continuous if and only if both $\mathbb{E}X(t,\omega)$ and $\rho(s,t)$ are continuous.

Proof Suppose that $R(s,t)$ is continuous. Making use of the proof of **1°** again, we have

$$|\,\mathbb{E}X(t,\omega) - \mathbb{E}X(t_0,\omega)\,|^2 \leqq \mathbb{E}\,|\,X(t,\omega) - X(t_0,\omega)\,|^2 \quad \text{(see the proof of 1°)}$$
$$= \mathbb{E}\,|\,X(t,\omega)\,|^2 + \mathbb{E}\,|\,X(t_0,\omega)\,|^2 - 2\mathcal{R}e\,\mathbb{E}X(t,\omega)\overline{X(t_0,\omega)}$$
$$= R(t,t) + R(t_0,t_0) - 2\mathcal{R}e\,R(t,t_0). \tag{8.1}$$

The formula (8.1) converges to 0 as $t \to t_0$, since $R(s,t)$ is continuous. Hence $t \mapsto \mathbb{E}X(t,\omega)$ is continuous.

$\rho(s,t)$ is computed as

$$\rho(s,t) = R(s,t) - \mathbb{E}X(s,\omega)\mathbb{E}\overline{X(t,\omega)}. \tag{8.2}$$

The second term on the right-hand side of (8.2) is continuous as shown above. So $\rho(s,t)$ is also continuous.

Conversely, suppose that both $\mathbb{E}X(t,\omega)$ and $\rho(s,t)$ are continuous. The continuity of $R(s,t)$ follows from (8.2). □

Theorem 8.1 (strong continuity) *The following statements are equivalent:*

(i) $X(t,\omega)$ *is strongly continuous.*
(ii) $R(s,t)$ *is continuous.*
(iii) $R(s,s)$ *is continuous in* s*, and the function* $s \mapsto R(s,t)$ *is continuous for each fixed* $t \in T$.

Proof (i)⇒(ii): For $s_0, t_0 \in \mathbb{R}$, we have

$$R(s,t) - R(s_0,t_0) = \mathbb{E}[X(s,\omega)\overline{X(t,\omega)} - X(s_0,\omega)\overline{X(t_0,\omega)}]$$
$$= \mathbb{E}\{[X(s,\omega) - X(s_0,\omega)]\overline{X(t,\omega)}\}$$
$$+\,\mathbb{E}\{X(s_0,\omega)[\overline{X(t,\omega)} - \overline{X(t_0,\omega)}]\}. \tag{8.3}$$

By Schwarz's inequality,

$$(\mathbb{E}\,|\,X(s,\omega)-X(s_0,\omega)\,|\,\cdot\,|\,\overline{X(t,\omega)}\,|\,)^2 \leqq \mathbb{E}\,|\,X(s,\omega)-X(s_0,\omega)\,|^{\,2}\cdot\mathbb{E}\,|\,\overline{X(t,\omega)}\,|^{\,2}. \tag{8.4}$$

$$(\mathbb{E}\,|\,X(s_0,\omega)\,|\,\cdot\,|\,\overline{X(t,\omega)}-\overline{X(t_0,\omega)}\,|\,)^2 \leqq \mathbb{E}\,|\,X(s_0,\omega)\,|^{\,2}\cdot\mathbb{E}\,|\,\overline{X(t,\omega)}-\overline{X(t_0,\omega)}\,|^{\,2}. \tag{8.5}$$

We obtain (ii) by (i), (8.3), (8.4) and (8.5).

(ii)⇒(iii): Obvious.

(iii)⇒(i): We have already observed in (8.1) that

$$\mathbb{E}\,|\,X(t,\omega)-X(t_0,\omega)\,|^{\,2}\to 0 \quad \text{as} \quad t\to t_0.$$

□

We introduce another concept of continuity which is weaker than strong continuity.

Given a stochastic process $X(t,\omega)$ of the second order, we define a subspace $\mathfrak{H}(X)$ of $\mathfrak{L}^2(\Omega,\mathbb{C})$ by

$$\mathfrak{H}(X)=\overline{\text{span}}\{X(t,\omega)\,|\,t\in\mathbb{R}\},$$

which is called the linear subspace determined by X. It is clear that $\mathfrak{H}(X)$ is a Hilbert space (under the inner product of $\mathfrak{L}^2$).

Definition 8.2 $X(t,\omega)$ is $\mathfrak{H}$-**weakly continuous** at $t_0\in\mathbb{R}$ if

$$\mathbb{E}Y(\omega)\overline{X(t,\omega)}\to\mathbb{E}Y(\omega)\overline{X(t_0,\omega)} \quad \text{as} \quad t\to t_0$$

for any $Y\in\mathfrak{H}(X)$. $X(t,\omega)$ is said to be $\mathfrak{H}$-**weakly continuous on** $\mathbb{R}$ or just $\mathfrak{H}$-weakly continuous if it is $\mathfrak{H}$-weakly continuous at every $t_0\in\mathbb{R}$.

If $X(t,\omega)$ is strongly continuous, it is $\mathfrak{H}$-weakly continuous. In fact, it immediately follows from the evaluation (Schwarz's inequality) that

$$|\mathbb{E}Y(\omega)\overline{X(t,\omega)}-\mathbb{E}Y(\omega)\overline{X(t_0,\omega)}|\leqq\|Y\|_2\cdot\|X(t,\omega)-X(t_0,\omega)\|_2$$

for any $Y\in\mathfrak{L}^2$.

Theorem 8.2 (separability of $\mathfrak{H}(X)$) *If $X(t,\omega)$ is $\mathfrak{H}$-weakly continuous, $\mathfrak{H}(X)$ is a separable Hilbert space.*

Proof Let D be the set of all random variables of the form

$$\sum_{j=1}^{n} r_j X(t_j,\omega), \quad n\in\mathbb{N}, \quad r_j : \text{rational complex},[5]\; t_j\in\mathbb{Q}.$$

[5] That is, both the real and imaginary parts are rationals.

Then D is a countable set in $\mathfrak{H}(X)$. If we define

$$\mathfrak{M} = \overline{\operatorname{span}} D,$$

$\mathfrak{M}$ is a separable closed subspace of $\mathfrak{H}(X)$. So our target is to show that $\mathfrak{M} = \mathfrak{H}(X)$.

Since $\mathfrak{H}(X)$ is a Hilbert space, it can be expressed as

$$\mathfrak{H}(X) = \mathfrak{M} \oplus \mathfrak{M}^{\perp},$$

where $\mathfrak{M}^{\perp}$ is the orthogonal complement of $\mathfrak{M}$. Hence any $Y \in \mathfrak{H}(X)$ can be uniquely represented in the form

$$Y(\omega) = V(\omega) + W(\omega), \quad V \in \mathfrak{M}, \quad W \in \mathfrak{M}^{\perp}.$$

Then we can show $W = 0$ as an element of $\mathfrak{L}^2$. Let t_0 be an element of $\mathbb{R}$ and $\{t_n\}$ a sequence in $\mathbb{Q}$ converging to t_0. The $\mathfrak{H}$-weak continuity of $X(t, \omega)$ implies

$$\mathbb{E} W(\omega)\overline{X(t_n, \omega)} \to \mathbb{E} W(\omega)\overline{X(t_0, \omega)} \quad \text{as} \quad n \to \infty. \tag{8.6}$$

However,

$$\mathbb{E} W(\omega)\overline{X(t_n, \omega)} = 0 \quad \text{for all } n, \tag{8.7}$$

since $X(t_n, \omega) \in D \subset \mathfrak{M}$ and $W \in \mathfrak{M}^{\perp}$. Hence it is obvious by (8.6) and (8.7) that

$$\mathbb{E} W(\omega)\overline{X(t_0, \omega)} = 0.$$

Therefore it follows that

$$\mathbb{E} W(\omega)\left[\overline{\sum_{j=1}^{n} c_j X(t_j, \omega)}\right] = 0$$

for any element $\sum_{j=1}^{n} c_j X(t_j, \omega)$, $(c_j \in \mathbb{C},\ t_j \in \mathbb{R},\ n \in \mathbb{N})$ of $\operatorname{span}\{X(t, \omega) \mid t \in \mathbb{R}\}$.
According to the extension by continuity, we obtain

$$\mathbb{E} W(\omega)\overline{Z(\omega)} = 0 \quad \text{for all} \quad Z \in \mathfrak{H}(X),$$

from which $W = 0$ (in $\mathfrak{L}^2$) follows.

By the uniqueness of orthogonal decomposition, we obtain

$$Y(\omega) = V(\omega),$$

from which $Y \in \mathfrak{M}$ follows. Since this holds good for any $Y \in \mathfrak{H}(X)$, we obtain $\mathfrak{H}(X) \subset \mathfrak{M}$. Thus we prove that $\mathfrak{M} = \mathfrak{H}(X)$. □

If $X(t, \omega)$ is strongly continuous, it is $\mathfrak{H}$-weakly continuous, and so $\mathfrak{H}(X)$ is separable by Theorem 8.2.

8.2 Weakly Stationary Stochastic Processes

Throughout this section, $T = \mathbb{R}$, $\mathbb{Z}$ or $\mathbb{N}$ and $(\Omega, \mathcal{E}, P)$ is a probability space. $X(t, \omega) : T \times \Omega \to \mathbb{C}$ is a stochastic process. In the case of $T = \mathbb{Z}$ or $\mathbb{N}$, we sometimes write $X_t(\omega)$ instead of $X(t, \omega)$.

Definition 8.3 A stochastic process $X(t, \omega)$ is **strongly stationary** if the distribution of $(X(t_1 + t, \omega), X(t_2 + t, \omega), \cdots, X(t_n + t, \omega))$ is independent of t for any $\mathbf{t} = (t_1, t_2, \cdots, t_n) \in \mathcal{T}$.

We define $\Phi_{t_1,t_2,\cdots,t_n}(E)$ for each $E \in \mathcal{B}(\mathbb{C}^n)$ by

$$\Phi_{t_1,t_2,\cdots,t_n}(E) = P\{\omega \in \Omega | (X(t_1, \omega), \cdots, X(t_n, \omega)) \in E\}.$$

Then $X(t, \omega)$ is strongly stationary if and only if

$$\Phi_{t_1,t_2,\cdots,t_n} = \Phi_{t_1+t,t_2+t,\cdots,t_n+t}, \quad n \in \mathbb{N}, \quad t_1, t_2, \cdots, t_n \in T.$$

Definition 8.4 $X(t, \omega)$ is said to be **weakly stationary** if the following conditions are satisfied:

(i) The absolute moment of the second order is finite:

$$\mathbb{E}|X(t, \omega)|^2 < \infty \quad \text{for each} \quad t \in T.$$

(ii) The expectation is constant throughout time:

$$\mathbb{E}X(t, \omega) = m(t) = m \quad \text{constant for all} \quad t \in T.$$

(iii) The covariance depends only upon the difference $u = s - t$ of times:

$$\rho(s, t) = \rho(s + h, t + h) \quad \text{for any} \quad s, t, h \in T.$$

In this case, the convariance $\rho(s, t)$ is a function of $s - t$. Hence we write it as $\rho(u)$, $u = s - t$ instead of $\rho(s, t)$.[6]

[6]We should use distinct notations for one-variable function $\rho(u)$ and two variable function $\rho(s, t)$. However, we use the same notation because there seems no possibility of confusion. $\rho(u)$ is also called the **autocovariance** and $\rho(u)/\rho(0)$ is called the **autocorrelation** of the process.

Theorem 8.3 *If a stochastic process $X(t, \omega)$ is strongly stationary and*

$$\mathbb{E}|X(t_0, \omega)|^2 < \infty$$

for some $t_0 \in T$, then $X(t, \omega)$ is weakly stationary.

Proof $X(t, \omega)$ is a stochastic process of the second order, since

$$\mathbb{E}|X(t, \omega)|^2 = \mathbb{E}|X(t_0, \omega)|^2 < \infty \quad \text{for all} \quad t \in T$$

by the strong stationarity. Furthermore,

$$m(t) = \int_{\mathbb{C}} \xi \, d\Phi_t(\xi),$$
$$\rho(s, t) = \int_{\mathbb{C}} \int_{\mathbb{C}} (\xi - m(s))\overline{(\eta - m(t))} d\Phi_{s,t}(\xi, \eta).$$

Hence $X(t, \omega)$ is weakly stationary. □

Remark 8.1 The strong stationarity does not imply the weak one because the condition $\mathbb{E}|X(t, \omega)|^2 < \infty$ is not required in the definition of the strong one.

Example 8.1 Let $X_j : \Omega \to \mathbb{C}$ $(j = 1, 2, \cdots, n)$ be mutually orthogonal random variables in $\mathfrak{L}^2(\Omega, \mathbb{C})$ with mean 0, i.e.

$$\mathbb{E}X_j(\omega) = 0, \quad j = 1, 2, \cdots, n,$$
$$\mathbb{E}X_j(\omega)\overline{X_k(\omega)} = 0 \quad \text{if} \quad j \neq k. \quad \text{(orthogonality)}$$

We define $X(t, \omega)$ by

$$X(t, \omega) = \sum_{j=1}^{n} e^{i\lambda_j t} X_j(\omega), \quad t \in \mathbb{R},$$

where $\lambda_j \in \mathbb{R}$ $(j = 1, 2, \cdots, n)$. Then $X(t, \omega)$ is weakly stationary with

$$\mathbb{E}X(t, \omega) = 0 \quad \text{for all} \quad t \in \mathbb{R}, \tag{8.8}$$

$$\rho(u) = \sum_{j=1}^{n} e^{i\lambda_j u} \mathbb{E}|X_j(\omega)|^2. \tag{8.9}$$

In fact, (8.8) is obvious. $\rho(s, t)$ is computed as

$$\rho(s,t) = \mathbb{E}X(s,\omega)\overline{X(t,\omega)} = \mathbb{E}\sum_{j=1}^{n}\sum_{k=1}^{n} e^{i\lambda_j s}e^{-i\lambda_k t}X_j(\omega)\overline{X_k(\omega)}$$
$$= \sum_{j=1}^{n} e^{i\lambda_j(s-t)}\mathbb{E}|X_j(\omega)|^2.$$

So we obtain (8.9) by setting $u = s - t$.

Example 8.2 Let $X_j, Y_j : \Omega \to \mathbb{R}$ $(j = 1, 2, \cdots, n)$ be orthogonal random variables with

$$\mathbb{E}X_j(\omega) = \mathbb{E}Y_j(\omega) = 0, \quad j = 1, 2, \cdots, n,$$
$$\mathbb{E}|X_j(\omega)|^2 = \mathbb{E}|Y_j(\omega)|^2 = 1, \quad j = 1, 2, \cdots, n.$$

We define $X(t,\omega)$ by

$$X(t,\omega) = \sum_{j=1}^{n} a_j[X_j(\omega)\cos\lambda_j t + Y_j(\omega)\sin\lambda_j t],$$

where $\alpha_j, \lambda_j \in \mathbb{R}$ $(j = 1, 2, \cdots, n)$.

Then it is clear that

$$\mathbb{E}X(t,\omega) = 0 \quad \text{for all} \quad t \in \mathbb{R}.$$

Since X_j's and Y_j's are random variables with mean 0 and variance 1, the covariance $\rho(s,t)$ is computed as

$$\rho(s,t) = \mathbb{E}X(s,\omega)\overline{X(t,\omega)}$$
$$= \mathbb{E}\Bigg[\sum_{j=1}^{n} a_j^2(X_j(\omega)^2\cos\lambda_j s\cos\lambda_j t + Y_j(\omega)^2\sin\lambda_j s\sin\lambda_j t)\Bigg]$$
$$= \sum_{j=1}^{n} a_j^2\cos\lambda_j(s-t).$$

Hence we have, setting $u = s - t$, that

$$\rho(u) = \sum_{j=1}^{n} a_j^2\cos\lambda_j u.$$

This proves that $X(t,\omega)$ is weakly stationary.

Example 8.3 (moving average) Let $X_n : \Omega \to \mathbb{C}$ $(n \in \mathbb{Z})$ be orthogonal random variables in $\mathfrak{L}^2$ (i.e. $\mathbb{E}X_j(\omega)\overline{X_k(\omega)} = 0$ for $j \neq k$) with

$$\mathbb{E}X_n = 0, \quad \mathbb{E}|X_n(\omega)|^2 = \sigma^2 \quad \text{(independent of } n\text{)}$$

for each $n \in \mathbb{Z}$. Define $Y_n(\omega)(n \in \mathbb{Z})$ by

$$Y_n(\omega) = \sum_{k=-\infty}^{\infty} c_{k-n} X_k(\omega), \quad n \in \mathbb{Z}, \tag{8.10}$$

where $\{c_\nu\}_{\nu \in \mathbb{Z}} \in l_2(\mathbb{C})$. Then we can prove that the right-hand side of (8.10) strongly converges in $\mathfrak{L}^2$ and $\{Y_n(\omega)\}$ is weakly stationary with

$$\mathbb{E}Y_n(\omega) = 0, \quad n \in \mathbb{Z}, \tag{8.11}$$

and

$$\rho(u) = \sigma^2 \sum_{\nu=-\infty}^{\infty} c_{u+\nu}\bar{c}_\nu.$$

We start by proving $\mathfrak{L}^2$-convergence of the right-hand side of (8.10). Fix $n \in \mathbb{Z}$. Suppose first that $p, p' \in \mathbb{Z}$ satisfy $0 \leqq p < p'$. We have, by a simple computation,

$$\begin{aligned} \mathbb{E}\left|\sum_{k=p}^{p'} c_{k-n} X_k(\omega)\right|^2 &= \sum_{j,k=p}^{p'} c_{j-n}\overline{c_{k-n}}\mathbb{E}X_j(\omega)\overline{X_k(\omega)} \\ &= \sum_{k=p}^{p'} |c_{k-n}|^2\sigma^2 \qquad \text{(by orthogonality)} \\ &= \sigma^2 \sum_{k=p}^{p'} |c_{k-n}|^2 = \sigma^2 \sum_{\nu=p-n}^{p'-n} |c_\nu|^2. \end{aligned} \tag{8.12}$$

When p and p' get larger (n being fixed), $p - n$ and $p' - n$ become positive and their absolute values grow indefinitely. Therefore (8.12) $\to 0$ as $p, p' \to \infty$, since $\{c_\nu\} \in l_2(\mathbb{C})$. (See Fig. 8.1.)

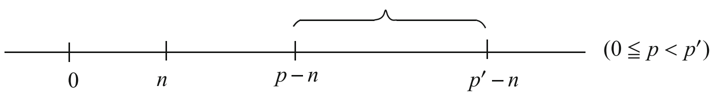


Fig. 8.1 Calculation of (8.12) – 1

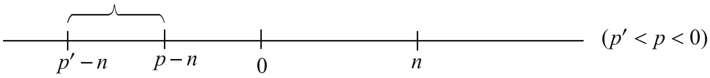


Fig. 8.2 Calculation of (8.12) – 2

Similarly, in the case of $p' < p < 0$, we have (Fig. 8.2)

$$\mathbb{E}\left|\sum_{k=p'}^{p} c_{k-n}X_k(\omega)\right|^2 = \sigma^2 \sum_{\nu=p'-n}^{p-n} |c_\nu|^2 \quad \to 0 \quad \text{as} \quad p, p' \to -\infty.$$

We consider partial sums of (8.10):

$$U(\omega) = \sum_{k=p}^{q} c_{k-n}X_k(\omega), \qquad V(\omega) = \sum_{k=p'}^{q'} c_{k-n}X_k(\omega)$$

with[7] $p' < p < 0, 0 \leqq q < q'$

and evaluate the distance between them based upon preliminary computations shown above (Fig. 8.3):

$$\begin{aligned}
\|U(\omega) - V(\omega)\|_2^2 &= \left\| \sum_{\nu=p'-n}^{p-n} c_\nu X_{n+\nu}(\omega) + \sum_{\nu=q-n}^{q'-n} c_\nu X_{n+\nu}(\omega)\right\|_2^2 \\
&= \left\| \sum_{\nu=p'-n}^{p-n} c_\nu X_{n+\nu}(\omega)\right\|_2^2 + \left\| \sum_{\nu=q-n}^{q'-n} c_\nu X_{n+\nu}(\omega)\right\|_2^2 \quad \text{(by orthogonality)} \\
&= \sigma^2 \left\{ \sum_{\nu=p'-n}^{p-n} |c_\nu|^2 + \sum_{\nu=q-n}^{q'-n} |c_\nu|^2 \right\} \\
&\to 0 \quad \text{as} \quad p, p' \to -\infty, \ q, q' \to \infty.
\end{aligned}$$

This proves that the sequence of partial sums

$$\sum_{k=p}^{q} c_{k-n}X_k(\omega) \quad p < 0, \ q \geqq 0, \ p \to -\infty, \ q \to \infty$$

[7] In this evaluation, we are examining the situation where the upper limits (q, q') of the sums tend to ∞, and the lower limits (p, p') tend to $-\infty$. So, without loss of generality, we may assume $p', p < 0, \ 0 \leqq q, q'$. There are various cases other than $p' < p$ and $q < q'$. However, we can treat them in the same manner.

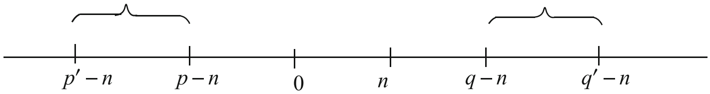


Fig. 8.3 Calculation of $\|U(\omega) - V(\omega)\|_2^2$

of the right-hand side of (8.10) is Cauchy in $\mathfrak{L}^2$. Hence the right-hand side of (8.10) strongly converges in $\mathfrak{L}^2$. Its limit is, as the above argument shows, determined uniquely and independent of the choice of $p \to -\infty$ and $q \to \infty$. So $Y_n(\omega)$ is defined without any ambiguity.

The zero mean condition (8.11) is verified as follows. By definition[8]

$$\mathbb{E}Y_n(\omega) = \mathbb{E}\left[\underset{p\to\infty}{\text{l.i.m.}} \sum_{k=-p}^{p} c_{k-n} X_k(\omega)\right] \quad (p > 0).$$

Since Ω is a probability space, $\mathfrak{L}^2$-convergence implies $\mathfrak{L}^1$-convergence. Hence we obtain

$$\mathbb{E}Y_n(\omega) = \lim_{p\to\infty} \mathbb{E} \sum_{k=-p}^{p} c_{k-n} X_k(\omega) = 0.$$

Finally, we evaluate the covariance:

$$\rho(m,n) = \mathbb{E}Y_m(\omega)\overline{Y_n(\omega)} = \sigma^2 \lim_{p\to\infty} \mathbb{E} \sum_{k=-p}^{p} c_{k-m}\overline{c_{k-n}} = \sigma^2 \sum_{\nu=-\infty}^{\infty} c_{n-m+\nu}\bar{c}_\nu.$$

Writing $u = n - m$, we obtain

$$\rho(u) = \sigma^2 \sum_{\nu=-\infty}^{\infty} c_{u+\nu}\bar{c}_\nu.$$

Thus we conclude that $\{Y_n(\omega)\}$ is a weakly stationary process.

The weakly stationary stochastic process defined by (8.10) is called a **moving average process** of $\{X_n(\omega)\}$ or a **linear stochastic process** generated by $\{X_n(\omega)\}$.

[8]It is known that the right-hand side converges strongly. Hence we can specify the upper and lower limits of $\sum$ as p and $-p$. l.i.m. denotes the limit in $\mathfrak{L}^2$.

Definition 8.5 A stochastic process $X_n : \Omega \to \mathbb{C}$ $(n \in \mathbb{Z})$ is called a **white noise** if it satisfies:

(i) $\mathbb{E}X_n(\omega) = 0 \quad$ for all $n \in \mathbb{Z}$,
(ii) $\mathbb{E}\,|\,X_n(\omega)|^2 = \sigma^2 \quad$ for all $n \in \mathbb{Z}$, and
(iii) $\mathbb{E}X_n(\omega)\overline{X_m(\omega)} = 0 \quad$ for all $n \neq m$.

The orthogonality expressed in (iii) is also called the condition of **serially uncorrelatedness**.

In terms of this terminology, Example 8.3 says that **a moving average process of a white noise is weakly stationary**.

The next results concern the relations between the strong continuity of a weakly stationary process and the continuity of its covariance.

Theorem 8.4 *Let $X : \mathbb{R} \times \Omega \to \mathbb{C}$ be a weakly stationary stochastic process with*

$$\mathbb{E}X(t, \omega) = 0 \quad \textit{for all} \quad t \in \mathbb{R}.$$

Then the following statements are equivalent:

(i) *$X(t, \omega)$ is strongly continuous at $t = 0$.*
(ii) *$X(t, \omega)$ is uniformly strongly continuous on $\mathbb{R}$.*
(iii) *$\rho(u)$ is continuous at $u = 0$.*
(iv) *$\rho(u)$ is uniformly continuous on $\mathbb{R}$.*

Proof (i)⇔(ii): Assume that $X(t, \omega)$ is strongly continuous at $t = 0$. The uniform continuity follows from

$$\begin{aligned}
\mathbb{E}|X(t+h, \omega) - X(t, \omega)|^2 &= \mathbb{E}|X(t+h, \omega)|^2 + \mathbb{E}|X(t, \omega)|^2 - 2\mathcal{R}e\mathbb{E}X(t+h, \omega)\overline{X(t, \omega)} \\
&= \mathbb{E}|X(h, \omega)|^2 + \mathbb{E}|X(0, \omega)|^2 - 2\mathcal{R}e\mathbb{E}X(h, \omega)\overline{X(0, \omega)} \\
&\qquad\qquad \text{(by weak stationarity)} \\
&= \mathbb{E}|X(h, \omega) - X(0, \omega)|^2. \qquad (8.13)
\end{aligned}$$

The converse assertion is obvious.

(ii)⇒(iv): By a simple calculation, we have

$$\begin{aligned}
|\rho(u+h) - \rho(u)|^2 &= \left|\mathbb{E}\left[X(u+h, \omega)\overline{X(0, \omega)} - X(u, \omega)\overline{X(0, \omega)}\right]\right|^2 \\
&\leqq \mathbb{E}|X(u+h, \omega) - X(u, \omega)|^2 \cdot \mathbb{E}|X(0, \omega)|^2. \qquad (8.14)
\end{aligned}$$

(iv) follows from (8.14) and (ii).

(iv)⇒(iii): Obvious.

(iii)⇒(ii): The right-hand side of (8.13) is rewritten as

$$\begin{aligned}
\mathbb{E}|X(t+h,\omega)-X(t,\omega)|^2 &= \mathbb{E}|X(h,\omega)-X(0,\omega)|^2 \qquad \text{(by (8.13))}\\
&= 2\mathbb{E}|X(0,\omega)|^2 - 2\mathcal{R}e\mathbb{E}X(h,\omega)\overline{X(0,\omega)}\\
&\qquad\qquad \text{(by weak stationarity)}\\
&= 2[\rho(0)-\mathcal{R}e\rho(h)]. \qquad (8.15)
\end{aligned}$$

(ii) follows from (8.15) and (iii). □

Definition 8.6 If $X(t,\omega)$ is $(\mathcal{L}\otimes\mathcal{E},\mathcal{B}(\mathbb{C}))$-measurable, $X(t,\omega)$ is called a **measurable stochastic process**, where $\mathcal{L}$ is the Lebesgue σ-field on $\mathbb{R}$.

The next theorem shows the strong continuity of a measurable weakly stationary stochastic process.

Theorem 8.5 (Crum)[9] *The covariance function of a measurable weakly stationary process $X(t,\omega)$ is continuous. Hence $X(t,\omega)$ is strongly continuous.*

Proof We may assume that $\mathbb{E}X(t,\omega)=0$ without loss of generality. It is sufficient to show the continuity of $\rho(u)$ at $u=0$, by Theorem 8.4.

To start with, we verify

$$e^{-t^2}X(t,\omega)\in\mathfrak{L}^2(\mathbb{R},\mathbb{C}) \quad \text{a.e.}$$

In fact, it immediately follows from

$$\mathbb{E}\int_{\mathbb{R}} e^{-2t^2}|X(t,\omega)|^2dt = \int_{\mathbb{R}} e^{-2t^2}\mathbb{E}|X(t,\omega)|^2dt = \rho(0)\int_{\mathbb{R}} e^{-2t^2}dt < \infty.$$

Define a couple of functions, $Y(\omega)$ and $Z(u,\omega)$, as follows:

$$Y(\omega) = \int_{\mathbb{R}} e^{-2t^2}|X(t,\omega)|^2dt,$$

$$Z(u,\omega) = \int_{\mathbb{R}} |e^{-t^2}X(t,\omega) - e^{-(t+u)^2}X(t+u,\omega)|^2dt. \qquad (8.16)$$

(Of course, they are defined for almost every ω.)

By the continuity of the shift operator (Theorem 5.1, p. 101), the function

$$u\mapsto e^{-(t+u)^2}X(t+u,\omega) \qquad (\mathbb{R}\to\mathfrak{L}^2)$$

is continuous. This implies that

[9] Crum [3].

$$Z(u,\omega) \to 0 \quad \text{as} \quad u \to 0 \quad \text{a.e.} \tag{8.17}$$

By the definition (8.16) of Z,[10]

$$\begin{aligned}
0 &\leqq Z(u,\omega) \\
&\leqq 2\left[\int_{\mathbb{R}} e^{-2t^2}|X(t,\omega)|^2 dt + \int_{\mathbb{R}} e^{-2(t+u)^2}|X(t+u,\omega)|^2 dt\right] \\
&= 4Y(\omega).
\end{aligned}$$

Since $\mathbb{E}Y(\omega) < \infty$, (8.17) and the dominated convergence theorem imply

$$\lim_{u\to 0} \mathbb{E}Z(u,\omega) = 0. \tag{8.18}$$

On the other hand,

$$\begin{aligned}
Z(u,\omega) &= \int_{\mathbb{R}} [e^{-2t^2}|X(t,\omega)|^2 + e^{-2(t+u)^2}|X(t+u,\omega)|^2 \\
&\qquad\qquad - 2\mathcal{R}e\, e^{-t^2-(t+u)^2} X(t+u,\omega)\overline{X(t,\omega)}]dt \\
&= 2\int_{\mathbb{R}} [e^{-2t^2}|X(t,\omega)|^2 - e^{-t^2-(t+u)^2}\mathcal{R}e X(t+u,\omega)\overline{X(t,\omega)}]dt \\
&\qquad\qquad \text{(by changing variables).}
\end{aligned}$$

Hence

$$\begin{aligned}
\mathbb{E}Z(u,\omega) &= 2\int_{\mathbb{R}} [e^{-2t^2}\rho(0) - e^{-t^2-(t+u)^2}\mathcal{R}e\,\rho(u)]dt \\
&= 2\rho(0)\int_{\mathbb{R}} e^{-2t^2}dt - 2e^{-u^2}\mathcal{R}e\,\rho(u)\int_{\mathbb{R}} e^{-2t^2-2tu}dt \\
&= 2[\rho(0) - \mathcal{R}e\rho(u)]e^{-u^2}\int_{\mathbb{R}} e^{-2t^2-2tu}dt + 2\rho(0)\int_{\mathbb{R}} e^{-2t^2}(1-e^{-2tu-u^2})dt.
\end{aligned}$$

Writing, for the sake of simplicity,

[10] It is well-known that

$$|\alpha - \beta|^2 \leqq 2(|\alpha|^2 + |\beta|^2)$$

for any $\alpha, \beta \in \mathbb{C}$, in general. Specifying α and β as

$$\alpha = e^{-t^2}X(t,\omega), \quad \beta = e^{-(t+u)^2}X(t+u,\omega),$$

we obtain the desired result.

$$I_1 = \int_{\mathbb{R}} e^{-2t^2-2tu}dt, \quad I_2 = \int_{\mathbb{R}} e^{-2t^2}(1-e^{-2tu-u^2})dt,$$

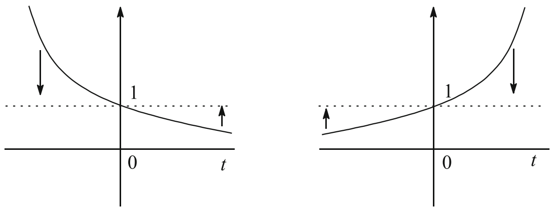
we obtain

$$\mathbb{E}Z(u,\omega) = 2[\rho(0) - \mathcal{R}e\,\rho(u)]e^{-u^2}I_1 + 2\rho(0)I_2. \tag{8.19}$$

By the monotone convergence theorem, it holds good that[11]

[11] The graph of the function $t \mapsto e^{-2tu}$ is depicted as follows according to the sign of u.

(a) (b)



(a) In the case of $u > 0$, e^{-2tu} monotonically converges to 1 at each t as $u \to 0$. To say more in detail, it is monotonically increasing in the region $t > 0$, and decreasing in the region $t < 0$. Since we may assume $0 < u \leqq 1$, without loss of generality, $e^{-2t^2-2tu} \leqq e^{-2t^2-2t}$ in the region $t < 0$. Hence applying the dominated convergence theorem and the monotone convergence theorem to the first and the second terms, respectively, of the right-hand side of the integral

$$\int_{\mathbb{R}} e^{-2t^2-2tu}dt = \int_{-\infty}^{0} e^{-2t^2}e^{-2tu}dt + \int_{0}^{\infty} e^{-2t^2}e^{-2tu}dt,$$

we obtain

$$\int_{\mathbb{R}} e^{-2t^2-2tu}dt \to \int_{\mathbb{R}} e^{-2t^2}dt \tag{†}$$

as $u \to 0$ $(u > 0)$.

(b) A similar argument applies to the case of $u < 0$, $u \to 0$.

In general, assume that a numerical sequence $\{u_n\}$ (u_n may either be positive or negative) converges to 0. Suppose

$$\int_{\mathbb{R}} e^{-2t^2}e^{-2tu_n}dt \nrightarrow \int_{\mathbb{R}} e^{-2t^2}dt.$$

Then for sufficiently small $\varepsilon > 0$, there exists a subsequence $\{u_{n'}\}$ of $\{u_n\}$ such that

$$\left| \int_{\mathbb{R}} e^{-2t^2-2tu_{n'}}dt - \int_{\mathbb{R}} e^{-2t^2}dt \right| \geqq \varepsilon \quad \text{for all } n'.$$

If we choose a further subsequence $\{u_{n''}\}$ of $\{u_{n'}\}$ of the type (a) or (b), then we must have

$$\left| \int_{\mathbb{R}} e^{-2t^2-2tu_{n''}}dt - \int_{\mathbb{R}} e^{-2t^2}dt \right| \geqq \varepsilon \quad \text{for all } n''.$$

A contradiction occurs.

$$\lim_{u\to 0} I_1 = \int_{\mathbb{R}} e^{-2t^2} dt, \quad \lim_{u\to 0} I_2 = 0. \tag{8.20}$$

Hence it follows that

$$\rho(0) - \mathcal{R}e\,\rho(u) \to 0 \quad \text{as} \quad u \to 0$$

by (8.18), (8.19) and (8.20).That is,

$$\mathcal{R}e\,\rho(u) \to \rho(0) \quad \text{as} \quad u \to 0. \tag{8.21}$$

Taking account of $|\rho(u)| \leqq \rho(0)$ (cf. 4° on p. 196), we have

$$(\mathcal{R}e\,\rho(u))^2 + (\mathcal{I}m\,\rho(u))^2 \leqq \rho(0)^2. \tag{8.22}$$

By (8.21) and (8.22),

$$\mathcal{I}m\,\rho(u) \to 0 \quad \text{as} \quad u \to 0.$$

Thus we obtain the desired result:

$$\rho(u) \to \rho(0) \quad \text{as} \quad u \to 0.$$

□

8.3 Periodicity of Weakly Stationary Stochastic Process

Let $X : T \times \Omega \to \mathbb{C}$ be a weakly stationary stochastic process. Then it is easy to see that its covariance function $\rho(u)$ is positive semi-definite. In fact,

$$\begin{aligned}\sum_{i,j=1}^{n} \rho(t_i - t_j)\lambda_i\overline{\lambda_j} &= \sum_{i,j=1}^{n} \mathbb{E}X(t_i,\omega)\overline{X(t_j,\omega)}\lambda_i\overline{\lambda_j} \\ &= \mathbb{E}\Big|\sum_{i,j=1}^{n} X(t_j,\omega)\lambda_j\Big|^2 \geqq 0\end{aligned}$$

for any $t_j \in T$ and $\lambda_j \in \mathbb{C}$ $(j = 1, 2, \cdots, n)$.

In the case of $T = \mathbb{Z}$, the following result is immediately obtained by Herglotz's theorem 6.4 (p. 133).

Theorem 8.6 (Spectral representation of $\rho : T = \mathbb{Z}$**)** *If $X : \mathbb{Z} \times \Omega \to \mathbb{C}$ is a weakly stationary process, its covariance function $\rho(u)$ can be expressed as the Fourier transform of certain positive Radon measure ν on $\mathbb{T}$:*

$$\rho(n) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\theta} d\nu(\theta).$$

Such a measure ν is determined uniquely.

In the case of $T = \mathbb{R}$, the corresponding result holds good.

Theorem 8.6′ (Spectral representation of $\rho : T = \mathbb{R}$) *If $X : \mathbb{R} \times \Omega \to \mathbb{C}$ is a measurable and weakly stationary process, its covariance function $\rho(u)$ can be expressed as the Fourier transform of certain positive Radon measure ν on $\mathbb{R}$:*

$$\rho(u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iut} d\nu(t).$$

Such a measure ν is determined uniquely.

Since $X(t, \omega)$ is measurable, $\rho(u)$ is continuous by Theorem 8.5. We have already confirmed the positive semi-definiteness of $\rho(u)$. Hence Theorem 8.6′ follows from Bochner's Theorem 6.11 (p. 150).

ν is not necessarily a probability measure because $\rho(0) = 1/\sqrt{2\pi}$ is not necessarily satisfied.

The Radon measure ν appearing Theorem 8.6 or Theorem 8.6′ is called the **spectral measure** of $X(t, \omega)$.

The function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(\alpha) = \nu((-\infty, \alpha)), \quad \alpha \in \mathbb{R}$$

is called the **spectral distribution function** of $X(t, \omega)$. If ν is absolutely continuous with respect to the Lebesgue measure dt, the Radon–Nikodým derivative $p(t) \in \mathfrak{L}^1(\mathbb{R}, \mathbb{R})$, $p(t) \geqq 0$ is called the **spectral density function** of $X(t, \omega)$; i.e.

$$\nu(E) = \int_E p(t) dt.$$

The spectral density function, if it exists, is unique. We will discuss later on the condition which assures the existence of a spectral density function of a stochastic process.

Example 8.4 We recapitulate Example 8.1 on p. 201. That is, consider a stochastic process

$$X(t, \omega) = \sum_{j=1}^{n} e^{i\lambda_j t} X_j(\omega), \quad t \in \mathbb{R} \qquad (\lambda_j \in \mathbb{R},\ j = 1, 2, \cdots, n),$$

where $X_j : \Omega \to \mathbb{C}$ $(j = 1, 2, \cdots, n)$ is a random variable with mean 0 and $X_j \perp X_k$ $(j \neq k)$. Then $X(t, \omega)$ is weakly stationary and

$$\mathbb{E}X(t,\omega) = 0 \quad \text{for all} \quad t \in \mathbb{R},$$

$$\rho(u) = \sum_{j=1}^{n} e^{i\lambda_j u}\mathbb{E}|X_j(\omega)|^2$$

as we have already verified. We may assume $\lambda_j \in [-\pi, \pi)$, without loss of generality. $\mathbb{E}|X_j(\omega)|^2$ is the variance of $X_j(\omega)$ and is denoted by σ_j^2. The spectral measure of $X(t,\omega)$ is a Radon measure on $\mathbb{T}$ which assigns a mass

$$\sqrt{2\pi}\mathbb{E}|X_j(\omega)|^2 = \sqrt{2\pi}\sigma_j^2$$

to each $t = -\lambda_j$; i.e.

$$\frac{1}{\sqrt{2\pi}}\nu = \sum_{j=1}^{n} \sigma_j^2 \delta_{-\lambda_j},$$

where δ_{λ_j} is a Dirac measure concentrating at λ_j. There is no spectral density function. Furthermore, the spectral distribution function is given by

$$\frac{1}{\sqrt{2\pi}}F(\alpha) = \begin{cases} 0 \quad \text{for} \quad \alpha < -\lambda_n, \\ \displaystyle\sum_{j=k}^{n} \sigma_j^2 \quad \text{for} \quad -\lambda_k \leqq \alpha < -\lambda_{k-1} \quad (k = 2, 3, \cdots, n), \\ \displaystyle\sum_{j=1}^{n} \sigma_j^2 \quad \text{for} \quad -\lambda_1 \leqq \alpha, \end{cases}$$

where $\lambda_1 < \lambda_2 < \cdots < \lambda_n$.

Example 8.5 Look at Example 8.2 on p. 202. That is, consider a stochastic process

$$X(t,\omega) = \sum_{j=1}^{n} a_j[X_j(\omega)\cos\lambda_j t + Y_j(\omega)\sin\lambda_j t],$$

where $X_j, Y_j : \Omega \to \mathbb{R}$ $(j = 1, 2, \cdots, n)$ are orthogonal random variables with mean 0 and variance 1. $a_j, \lambda_j \in \mathbb{R}$ $(j = 1, 2, \cdots, n)$. Then $X(t,\omega)$ is weakly stationary and

$$\mathbb{E}X(t,\omega) = 0 \quad \text{for all} \quad t \in \mathbb{R},$$

$$\rho(u) = \sum_{j=1}^{n} a_j^2 \cos\lambda_j u. \tag{8.23}$$

We may assume again that $\lambda_j \in \mathbb{T}$, without loss of generality. The spectral measure of $X(t, \omega)$ is, by (8.23), the Radon measure on $\mathbb{T}$ which assigns a mass $\sqrt{2\pi} a_j^2/2$ to each $t = \pm\lambda_j$; i.e.

$$\nu = \sqrt{2\pi} \sum_{j=1}^{n} \frac{a_j^2}{2} \delta_{\lambda_j} + \sqrt{2\pi} \sum_{j=1}^{n} \frac{a_j^2}{2} \delta_{-\lambda_j} = \sqrt{2\pi} \sum_{j=1}^{n} \frac{a_j^2}{2} (\delta_{\lambda_j} + \delta_{-\lambda_j}).$$

There is no spectral density function. The spectral distribution function of $X(t, \omega)$ is a step function with a jump $\sqrt{2\pi} a_j^2/2$ at each $t = \pm\lambda_j$.

The spectral measure of a moving average process (Example 8.3) will be discussed later on.

Here arises a natural question.[12] When a positive Radon measure ν on T is given, does there exist a weakly stationary stochastic process the spectral measure of which is exactly ν? The following proposition answers this question positively.

Theorem 8.7 *A positive Radon measure ν on $\mathbb{T}$ is assumed to be given. Then there exist a probability space $(\Omega, \mathcal{E}, P)$ and a stochastic process $X_n : \Omega \to \mathbb{C} (n \in \mathbb{Z})$ which satisfies the following conditions:*

(i) *$X_n(\omega)$ is a weakly stationary process.*
(ii) *ν is the spectral measure of $X_n(\omega)$.*

The next result is a continuous-time ($T = \mathbb{R}$) version of the above.

Theorem 8.7′ *A positive Radon measure ν on $\mathbb{R}$ is assumed to be given. Then there exist a probability space $(\Omega, \mathcal{E}, P)$ and a stochastic process $X : \mathbb{R} \times \Omega \to \mathbb{C}$ which satisfies the following conditions:*

(i) *$X(t, \omega)$ is a measurable and weakly stationary process.*[13]
(ii) *ν is the spectral measure of $X(t, \omega)$.*

Proof Let ν be a positive Radon measure on $\mathbb{R}$. If we define a set function θ on $\mathcal{B}(\mathbb{R})$ by

$$\theta(E) = \nu(E)/\nu(\mathbb{R}), \quad E \in \mathcal{B}(\mathbb{R}),$$

then θ is a Radon probability measure on $\mathbb{R}$. As is well-known in probability theory, there exist a certain probability space $(\Omega, \mathcal{E}, P)$ and a couple of (real) random variables $Y(\omega)$, $Z(\omega)$ which are defined on Ω and satisfy the following conditions[14]:

[12] Kawata [16], pp. 150–151.

[13] $X(t, \omega)$ is strongly continuous by Theorem 8.5.

[14] Let $\Phi_1, \Phi_2, \cdots$ be a sequence of Borel probability measures. Then there exist a certain probability space $(\Omega, \mathcal{E}, P)$ and mutually independent random variables, the distributions of which are $\Phi_1, \Phi_2, \cdots$ (Itô [10], p. 68).

a. $Y(\omega)$ and $Z(\omega)$ are independent.
b. The distribution of $Z(\omega)$ is θ.
c. $\mathbb{E}Y(\omega) = 0, \quad \mathbb{E}Y(\omega)^2 = \frac{1}{\sqrt{2\pi}}\nu(\mathbb{R})$.

Define a stochastic process $X(t,\omega)$ by

$$X(t,\omega) = Y(\omega)e^{-iZ(\omega)t}.$$

Then it satisfies[15]

$$\begin{aligned}
\mathbb{E}X(t,\omega) &= \mathbb{E}Y(\omega)\mathbb{E}e^{-iZ(\omega)t} = 0, \\
\rho(t+u,t) &= \mathbb{E}X(t+u,\omega)\overline{X(t,\omega)} = \mathbb{E}[Y(\omega)^2 e^{-iZ(\omega)u}] \\
&= \frac{1}{\sqrt{2\pi}}\nu(\mathbb{R})\int_{\mathbb{R}} e^{-i\lambda u}\,d\theta(\lambda) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-i\lambda u}\,d\nu(\lambda).
\end{aligned}$$

The covariance $\rho(t+u,t)$ of $X(t,\omega)$ depends only upon u, and it is represented as the Fourier transform of ν.

Since the Fourier transform of ν is uniformly continuous,[16] $X(t,\omega)$ is strongly continuous. The measurability of $X(t,\omega)$ is also clear by its definition. □

We now investigate the periodic behaviors of weakly stationary processes.

Definition 8.7 Let $X : T \times \Omega \to \mathbb{C}$ be a weakly stationary process with the covariance function $\rho(u)$. $X(t,\omega)$ is called a **periodic weakly stationary process** with period τ or, in short, τ-periodic if the $\rho(u)$ is periodic with period τ; i.e. $\rho(u+\tau) = \rho(u)$.

Theorem 8.8 is a characterization of periodic processes in the discrete-time case. Theorem 8.8′ is its continuous-time version.

Theorem 8.8 *Let $X : \mathbb{Z} \times \Omega \to \mathbb{C}$ be a weakly stationary process with the spectral measure ν. Then the following three statements are equivalent:*

---

[15] Let $X_1, X_2, \cdots$ be independent real random variables and $g_1, g_2, \cdots : \mathbb{R} \to \mathbb{C}$ Borel measurable. Then $Y_1 = g_1(X_1), Y_2 = g_2(X_2), \cdots$ are independent random variables (Itô [10], p. 66). Based upon this result, $Y(\omega)$ and $e^{-iZ(\omega)t}$ in the text are independent.

[16] If we denote by $\hat{\nu}$ the Fourier transform of ν, we have

$$\begin{aligned}
|\hat{\nu}(t+h) - \hat{\nu}(t)| &= \frac{1}{\sqrt{2\pi}}\left|\int_{\mathbb{R}} (e^{-i(t+h)x} - e^{-itx})\,d\nu(x)\right| \\
&\leqq \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} |e^{-ihx} - 1|\,d\nu(x) \\
&\to 0 \quad \text{as} \quad h \to 0,
\end{aligned}$$

by the dominated convergence theorem.

(i) $X_n(\omega)$ *is* τ*-periodic.*
(ii) $X_{n+\tau}(\omega) - X_n(\omega) = 0 \quad a.e.(\omega) \quad$ *for all* $n \in \mathbb{Z}$.
(iii) *If* $E \in \mathcal{B}(\mathbb{T})$ *and* $E \cap \{2k\pi/\tau | k \in \mathbb{Z}\} = \emptyset$*, then* $\nu(E) = 0$.

Theorem 8.8′ *Let* $X : \mathbb{R} \times \Omega \to \mathbb{C}$ *be a measurable and weakly stationary process with the spectral measure* ν*. Then the following three statements are equivalent:*

(i) $X(t, \omega)$ *is* τ*-periodic.*
(ii) $X(t + \tau, \omega) - X(t, \omega) = 0.\ a.e.(\omega) \quad$ *for all* $t \in \mathbb{R}$.
(iii) *If* $E \in \mathcal{B}(\mathbb{R})$ *and* $E \cap \{2k\pi/\tau | k \in \mathbb{Z}\} = \emptyset$*, then* $\nu(E) = 0$.

Proof (i)⇒(ii): We assume that $\rho(\cdot)$ is a periodic function with period T, and show

$$\mathbb{E}|X(t + T, \omega) - X(t, \omega)|^2 = 0,$$

which is equivalent to (ii). It is easily verified by a direct computation:

$$\begin{aligned}\mathbb{E}|X(t + T, \omega) - X(t, \omega)|^2 &= \mathbb{E}|X(t + T, \omega)|^2 + \mathbb{E}|X(t, \omega)|^2 - 2\mathcal{R}e\mathbb{E}X(t+T, \omega)\overline{X(t, \omega)} \\ &= 2\rho(0) - 2\mathcal{R}e\rho(T) \\ &= 2(\rho(0) - \mathcal{R}e\rho(T)) \\ &= 0 \quad \text{(by (i))}.\end{aligned}$$

(ii)⇒(i): (i) follows from (ii) because

$$\begin{aligned}|\rho(u + T) - \rho(u)|^2 &= |\mathbb{E}[X(u + T, \omega)\overline{X(0, \omega)} - X(u, \omega)\overline{X(0, \omega)}]|^2 \\ &\leqq \mathbb{E}|X(u + T, \omega) - X(u, \omega)|^2 \cdot \mathbb{E}|X(0, \omega)|^2 \\ &\quad \text{(Schwarz's inequality)} \\ &= 0 \quad \text{(by (ii))}.\end{aligned}$$

(i)⇒(iii): If $\rho(\cdot)$ is a periodic function with period T, then

$$0 = 2\rho(0) - \rho(T) - \rho(-T) = 2\int_{\mathbb{R}} (1 - \cos tT)d\nu(t).$$

Since $1 - \cos tT \geqq 0$, we must have $\nu(E) = 0$ ($E \in \mathcal{B}(\mathbb{R})$) if E contains no point t such that $1 - \cos tT = 0$; i.e. $E \cap \{2k\pi/T | k \in \mathbb{Z}\} = \emptyset$.

(iii)⇒(i): Assume (iii). Let ν be the spectral measure of $X(t, \omega)$; i.e.

$$\rho(u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iut} d\nu(t).$$

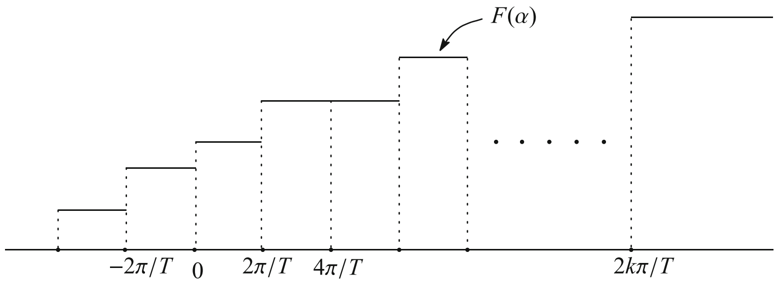


Fig. 8.4 Spectral distribution function

Writing $a_k = \nu(\{2\pi k/T\})$ $(k \in \mathbb{Z})$, we obtain

$$\rho(u) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} a_k e^{-iu\cdot 2\pi k/T}$$

by (iii). This is clearly a T-periodic function. □

Remark 8.2 The spectral distribution function of $X(t, \omega)$ is defined by

$$F(\alpha) = \nu((-\infty, \alpha]), \quad \alpha \in \mathbb{R}.$$

Then (iii) means that $F(\alpha)$ is a step function with possible discontinuities at $\{2k\pi/T | k \in \mathbb{Z}\}$. See Fig. 8.4.

If $X(t, \omega)$ is a T-periodic weakly stationary process, the spectral measure concentrates on a countable set in $\mathbb{T}$ or $\mathbb{R}$, informally called the energy set of $X(t, \omega)$, such that the distance of any adjacent two points is some multiple of $2\pi/T$. We have to keep in mind that the periodic weakly stationary process can not have a spectral density function. Sargent [26] (Chap. XI) talks about the spectral density function in this case. It is, according to Sargent, zero on $(2(k-1)\pi/T, 2k\pi/T)$ and has a "spike" at a jump-point of $F(\alpha)$. This exposition is something like a parable and not correct mathematically. We need Schwartz's distribution theory to grasp the situation exactly.

8.4 Orthogonal Measures

Our crucial objects in this chapter include the spectral representation of a weakly stationary process and the absolute continuity of spectral measures with respect to the Lebesgue measure. For the analysis of these problems, the concept of orthogonal measures is indispensable. We now state its definition and basic properties.

$\mathcal{B}(\mathbb{R})$ denotes the Borel σ-field on $\mathbb{R}$, m is the Lebesgue measure on $\mathbb{R}$, and $\mathcal{B}^*(\mathbb{R}) = \{S \in \mathcal{B}(\mathbb{R}) | m(S) < \infty\}$. $(\Omega, \mathcal{E}, P)$ is a probability space.

Definition 8.8 A function $\xi : \mathcal{B}(\mathbb{R}) \times \Omega \to \mathbb{C}$ is called an **orthogonal measure** (or $\mathfrak{L}^2$-orthogonal measure) if the following three conditions are satisfied:

(i) The function $\omega \mapsto \xi(S, \omega)$ is in $\mathfrak{L}^2(\Omega, \mathbb{C})$ for every $S \in \mathcal{B}^*(\mathbb{R})$.

(ii) If $S_1, S_2, \cdots \in \mathcal{B}^*(\mathbb{R})$ are mutually disjoint and $\bigcup_{n=1}^{\infty} S_n \in \mathcal{B}^*(\mathbb{R})$, then

$$\xi\left(\bigcup_{n=1}^{\infty} S_n, \omega\right) = \sum_{n=1}^{\infty} \xi(S_n, \omega) \text{ in } \mathfrak{L}^2(\Omega, \mathbb{C}).^{17} \quad (\sigma\text{-additivity})$$

(iii) If $S, S' \in \mathcal{B}^*(\mathbb{R})$ and $S \cap S' = \emptyset$, then

$$\mathbb{E}\xi(S, \omega)\overline{\xi(S', \omega)} = 0 \quad \text{(orthogonality)}.$$

If (i), (ii) and (iii) are satisfied on $\mathcal{B}(\mathbb{R})$ instead of $\mathcal{B}^*(\mathbb{R})$, $\xi(S, \omega)$ is called a **finite orthogonal measure**.

Let ξ be an orthogonal measure. The function $\omega \mapsto \xi(\mathbb{R}, \omega)$ may not be in $\mathfrak{L}^2(\Omega, \mathbb{C})$ in general. However, it is in $\mathfrak{L}^2(\Omega, \mathbb{C})$ if ξ is a finite orthogonal measure.

The concept of orthogonal measures is similarly defined in the case of $T = \mathbb{T}$. In this case, any orthogonal measure is automatically finite.

Theorem 8.9 *Let $\xi : \mathcal{B}(\mathbb{R}) \times \Omega \to \mathbb{C}$ be a finite orthogonal measure. The set function $\nu_\xi : \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ defined by*

$$\nu_\xi(S) = \|\xi(S, \omega)\|_2^2, \quad S \in \mathcal{B}(\mathbb{R})$$

is a finite measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proof It is obvious that $\nu_\xi(S) \geqq 0$ for all $S \in \mathcal{B}(\mathbb{R})$. Furthermore, if $S_n \in \mathcal{B}(\mathbb{R})$ $(n = 1, 2, \cdots)$ are mutually disjoint, then

$$\nu_\xi\left(\bigcup_{n=1}^{\infty} S_n\right) = \left\|\xi\left(\bigcup_{n=1}^{\infty} S_n, \omega\right)\right\|_2^2 \underset{\text{(ii)}}{=} \lim_{n\to\infty} \left\|\sum_{j=1}^{n} \xi(S_j, \omega)\right\|_2^2$$

$$\underset{\text{(iii)}}{=} \lim_{n\to\infty} \sum_{j=1}^{n} \|\xi(S_j, \omega)\|_2^2 = \sum_{n=1}^{\infty} \nu_\xi(S_n).$$

This proves the σ-additivity of $\nu_\xi(\cdot)$. $\nu_\xi(\mathbb{R}) < \infty$ is clear by (i). $\square$

---

[17] Hence if the right-hand side is a finite sum,

$$\xi\left(\bigcup_{j=1}^{p} S_j, \omega\right) = \sum_{j=1}^{p} \xi(S_j, \omega) \text{ a.e.}$$

Theorem 8.9 does not hold for a general orthogonal measure. If ν_ξ satisfies

$$\nu_\xi(S) = cm(S) \quad \text{for all} \quad S \in \mathcal{B}^*(\mathbb{R}),$$

ξ is called a **Khintchine orthogonal measure**.

We now define the integration by means of a finite orthogonal measure.

Let $\varphi : \mathbb{R} \to \mathbb{C}$ be a simple function of the form

$$\varphi(x) = \sum_{j=1}^{p} c_j \chi_{S_j}(x), \quad c_j \in \mathbb{C}, \quad S_j \in \mathcal{B}(\mathbb{R}), \tag{8.24}$$

$$S_i \cap S_j = \emptyset \ (i \neq j), \quad m(S_j) < \infty,$$

(χ_{S_j} is the characteristic function of S_j). m is the Lebesgue measure on $\mathbb{R}$ and it is also denoted by $dx, dy, \cdots$ in order to specify integrating variables. We define the integral of φ by

$$\int_{\mathbb{R}} \varphi(x)\xi(dx, \omega) = \sum_{j=1}^{p} c_j \xi(S_j, \omega). \tag{8.25}$$

Of course, the representation (8.24) of φ is not unique. However, the integral is well-defined in the sense that the value of (8.25) is uniquely determined independently of the representation (8.24). This can be proved in the same manner as for the ordinary Lebesgue integrals.

The integral of a linear combination $\sum_{k=1}^{n} \alpha_k \varphi_k(x)$ ($\alpha_k \in \mathbb{C}$) of simple functions is defined by

$$\int_{\mathbb{R}} \sum_{k=1}^{n} \alpha_k \varphi_k(x)\xi(dx, \omega) = \sum_{k=1}^{n} \alpha_k \int_{\mathbb{R}} \varphi_k(x)\xi(dx, \omega).$$

Defining a measure ν_ξ as in Theorem 8.9, we obtain

$$\left\| \int_{\mathbb{R}} \varphi(x)\xi(dx, \omega) \right\|_2^2 = \int_{\mathbb{R}} |\varphi(x)|^2 \nu_\xi(dx). \tag{8.26}$$

In fact, it is verified by

$$\text{left-hand side of (8.26)} = \sum_{k=1}^{n} |c_k|^2 \nu_\xi(S_k) = \text{right-hand side of (8.26)}.$$

The set of measurable functions $f : \mathbb{R} \to \mathbb{C}$ which is square-integrable with respect to ν_ξ is denoted by $\mathfrak{L}^2_{\nu_\xi}(\mathbb{R}, \mathbb{C})$.

For any $f \in \mathfrak{L}^2_{\nu_\xi}(\mathbb{R}, \mathbb{C})$, there exists a sequence $\{\varphi_n\}$ of simple functions in $\mathfrak{L}^2_{\nu_\xi}(\mathbb{R}, \mathbb{C})$ such that

$$\|f - \varphi_n\|^2_{\nu_\xi,2} = \int_{\mathbb{R}} |f(x) - \varphi_n(x)|^2 d\nu_\xi \to 0 \quad \text{as} \quad n \to \infty. \tag{8.27}$$

By (8.26) and (8.27),

$$\left\| \int_{\mathbb{R}} \varphi_m(x)\xi(dx, \omega) - \int_{\mathbb{R}} \varphi_n(x)\xi(dx, \omega) \right\|_2^2$$
$$= \int_{\mathbb{R}} |\varphi_m(x) - \varphi_n(x)|^2 d\nu_\xi \to 0 \quad \text{as} \quad n, m \to \infty.$$

That is, $\{\varphi_n\}$ is Cauchy in $\mathfrak{L}^2_{\nu_\xi}$. Hence

$$\int_{\mathbb{R}} \varphi_n(x)\xi(dx, \omega)$$

converges to some element of $\mathfrak{L}^2_{\nu_\xi}$. We call this function the integral of f with respect to $\xi(S, \omega)$ and write it as

$$\int_{\mathbb{R}} f(x)\xi(dx, \omega). \tag{8.28}$$

It can be shown as usual that the limit (8.28) is determined independently of the choice of $\{\varphi_n\}$ converging to f.

The integration by an orthogonal measure has basic properties as follows:

1° For $f_1, f_2, \cdots, f_p \in \mathfrak{L}^2_{\nu_\xi}(\mathbb{R}, \mathbb{C}), \alpha_1, \alpha_2, \cdots, \alpha_p \in \mathbb{C}$,

$$\int_{\mathbb{R}} \sum_{k=1}^{p} \alpha_k f_k(x)\xi(dx, \omega) = \sum_{k=1}^{p} \alpha_k \int_{\mathbb{R}} f_k(x)\xi(dx, \omega).$$

2° Let $\{f_n\}$ be a sequence which converges to $f \in \mathfrak{L}^2_{\nu_\xi}(\mathbb{R}, \mathbb{C})$ in $\mathfrak{L}^2_{\nu_\xi}$. Then

$$\int_{\mathbb{R}} f_n(x)\xi(dx, \omega) \to \int_{\mathbb{R}} f(x)\xi(dx, \omega) \quad \text{as} \quad n \to \infty$$

in $\mathfrak{L}^2(\Omega, \mathbb{C})$.

The two formulas in Theorem 8.10 will be effectively made use of in later discussions. 2° stated above can be proved by them.

Theorem 8.10 (D-K formulas)

$$\mathbb{E}\int_{\mathbb{R}} f(x)\xi(dx,\omega)\int_{\mathbb{R}} \overline{g(x)\xi(dx,\omega)} = \int_{\mathbb{R}} f(x)\overline{g(x)}d\nu_\xi(x)$$
$$\textit{for any}\quad f,g \in \mathfrak{L}^2_{\nu_\xi}(\mathbb{R},\mathbb{C}). \tag{8.29}$$

In particular,

$$\left\| \int_{\mathbb{R}} f(x)\xi(dx,\omega)\right\|_2^2 = \int_{\mathbb{R}} |f(x)|^2 d\nu_\xi(x)$$
$$\textit{for } f \in \mathfrak{L}^2_{\nu_\xi}(\mathbb{R},\mathbb{C}). \tag{8.30}$$

For the sake of later reference, we call (8.29) and (8.30) the **the Doob–Kawata formulas** (D–K formulas).

Proof We prove (8.30) first. If f is a simple function in $\mathfrak{L}^2_{\nu_\xi}$, (8.30) is clear. If f is a general element of $\mathfrak{L}^2_{\nu_\xi}$, there exists a sequence $\{\varphi_n\}$ of simple functions in $\mathfrak{L}^2_{\nu_\xi}$ such that

$$\|f-\varphi_n\|^2_{\nu_\xi,2} \to 0 \quad \text{as} \quad n\to\infty$$

($\|\cdot\|_{\nu_\xi,2}$ is the $\mathfrak{L}^2$-norm with respect to ν_ξ). It follows that

$$\int_{\mathbb{R}} |\varphi_n(x)|^2 d\nu_\xi(x) \to \int_{\mathbb{R}} |f(x)|^2 d\nu_\xi(x). \tag{8.31}$$

On the other hand,

$$\left\| \int_{\mathbb{R}} \varphi_n(x)\xi(dx,\omega)\right\|_2^2 \to \left\| \int_{\mathbb{R}} f(x)\xi(dx,\omega)\right\|_2^2, \tag{8.32}$$

by the definition (8.28) of integration. Since the left-hand side of (8.31) and that of (8.32) are equal, (8.30) follows.

Next we turn to (8.29). By (8.30), we obtain

$$\left\| \int_{\mathbb{R}} (f(x)-g(x))\xi(dx,\omega)\right\|_2^2 = \int_{\mathbb{R}} |f(x)-g(x)|^2 d\nu_\xi. \tag{8.33}$$

Both sides of (8.33) are evaluated as follows:

$$\begin{aligned}\text{left-hand side of (8.33)} &= \left\| \int_{\mathbb{R}} f(x)\xi(dx,\omega) \right\|_2^2 + \left\| \int_{\mathbb{R}} g(x)\xi(dx,\omega) \right\|_2^2 \\ &\quad - 2\mathcal{R}e\mathbb{E} \int_{\mathbb{R}} f(x)\xi(dx,\omega) \int_{\mathbb{R}} \overline{g(x)\xi(dx,\omega)}. \\ \text{right-hand side of (8.33)} &= \int_{\mathbb{R}} |f(x)|^2 d\nu_\xi(x) + \int_{\mathbb{R}} |g(x)|^2 d\nu_\xi(x) \\ &\quad - 2\mathcal{R}e \int_{\mathbb{R}} f(x)\overline{g(x)} d\nu_\xi(x).\end{aligned}$$

Hence by (8.30) and (8.33), we obtain

$$\mathcal{R}e\mathbb{E} \int_{\mathbb{R}} f(x)\xi(dx,\omega) \cdot \int_{\mathbb{R}} \overline{g(x)\xi(dx,\omega)} = \mathcal{R}e \int_{\mathbb{R}} f(x)\overline{g(x)} d\nu_\xi(x). \tag{8.34}$$

Replacing $f(x) - g(x)$ in (8.33) by $if(x) - g(x)$, (8.34) holds good after changing $f(x)$ to $if(x)$; i.e.

$$\mathcal{I}m\mathbb{E} \int_{\mathbb{R}} f(x)\xi(dx,\omega) \cdot \int_{\mathbb{R}} \overline{g(x)\xi(dx,\omega)} = \mathcal{I}m \int_{\mathbb{R}} f(x)\overline{g(x)} d\nu_\xi(x). \tag{8.35}$$

The formula (8.29) follows from (8.34) and (8.35). □

8.5 Spectral Representation of Weakly Stationary Stochastic Processes: Cramér–Kolmogorov Theorem

In Sect. 8.3, we discussed the problem of representing the convariance function of a weakly stationary stochastic process by the Fourier transform of some positive Radon measure. In this section, we proceed to the Fourier analytic representation of a weakly stationary process itself.

Theorem 8.11 (Cramér–Kolmogorov)[18] *Let $X : \mathbb{R} \times \Omega \to \mathbb{C}$ be a measurable and weakly stationary process with $\mathbb{E}X(t,\omega) = a$ (for all $t \in \mathbb{R}$). Then there exists an orthogonal measure $\xi : \mathcal{B}(\mathbb{R}) \times \Omega \to \mathbb{C}$ which satisfies*

$$X(t,\omega) = a + \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\lambda t}\xi(d\lambda,\omega). \tag{8.36}$$

[18] According to Itô [10] p. 255, Kolmogorov's important article was published in *C.R. Acad. Sci. URSS*, **26** (1940), 115–118. However, I have never read it, very regrettably. That is why I dropped it from the reference list.

Conversely, the stochastic process $X(t, \omega)$ represented by the above formula in terms of an orthogonal measure $\xi(S, \omega)$ is weakly stationary.

ξ is uniquely determined corresponding to X.

The same result holds good for the case $T = \mathbb{Z}$. In this case, $\mathcal{B}(\mathbb{R})$ should be replaced by $\mathcal{B}(\mathbb{T})$ and the scope of integration in (8.36) should be $\mathbb{T}$ instead of $\mathbb{R}$.

Remark 8.3

1° Since we are assuming the measurability of $X(t, \omega)$, $X(t, \omega)$ is strongly continuous by Crum's theorem (Theorem 8.5).

2° Let M be a complete metric space and D a dense subset in M. If a function $f : D \to D$ is an isometric (and so uniformly continuous) surjection, f is, of course, a bijection. By extension by continuity,[19] f can be uniquely extended to an isometric function of M into M. We denote the extended function by the same notation f. $f : M \to M$ is an isometric injection. It can be proved that f is also a surjection.

Let y be any point of M. Since $\bar{D} = M$, there exists a sequence $\{y_n\}$ in D which converges to y:

$$y_n \to y \quad \text{as} \quad n \to \infty, \qquad y_n \in D.$$

Since $f : D \to D$ is a bijection, there exists a unique $x_n \in D$, for each n, such that $f(x_n) = y_n$. Since $\{y_n\}$ is convergent and f is isometry, $\{x_n\}$ is Cauchy in D. Hence, by the completeness of M, $\{x_n\}$ converges to some $x \in M$:

$$x_n \to x \quad \text{as} \quad n \to \infty.$$

$f(x_n) \to f(x)$ and $f(x_n) = y_n \to y$ imply $y = f(x)$. This proves that $f : M \to M$ is surjective.

We make use of this reasoning in the proof.

Proof of Theorem 8.11 We may assume $a = 0$ without loss of generality. Otherwise we apply our reasoning below to $X'(t, \omega) = X(t, \omega) - a$. If we define

$$\mathfrak{M} = \operatorname{span}\{X(t, \omega) | t \in \mathbb{R}\}, \qquad \mathfrak{H}(X) = \bar{\mathfrak{M}},$$

an element of $\mathfrak{M}$ has the form

$$\sum_{i=1}^{n} a_i X(t_i, \omega), \qquad a_i \in \mathbb{C}.$$

$\mathfrak{H}(X)$ is a closed subspace of $\mathfrak{L}^2(\Omega, \mathbb{C})$, and is itself a Hilbert space.

---

[19] See Maruyama [21] pp. 74–76.

We now proceed to define an operator $U_\tau (\tau \in \mathbb{R})$ on $\mathfrak{H}(X)$. To start with, we define

$$U_\tau X(t, \omega) = X(t + \tau, \omega).$$

If $W(\omega) \in \mathfrak{M}$ is expressed in the form

$$W(\omega) = \sum_{i=1}^{n} a_i X(t_i, \omega), \tag{8.37}$$

$U_\tau W$ is defined by

$$U_\tau W(\omega) = \sum_{i=1}^{n} a_i X(t_i + \tau, \omega).$$

We have to check that $U_\tau W$ is well-defined in the sense that the value defined above is independent of the expression (8.37) of $W(\omega)$. Let

$$W(\omega) = \sum_{j=1}^{m} b_j X(s_j, \omega)$$

be another expression of $W(\omega)$. We have to show

$$\sum_{i=1}^{n} a_i X(t_i + \tau, \omega) = \sum_{j=1}^{m} b_j X(s_j + \tau, \omega). \tag{8.38}$$

Taking account of $a = 0$, we have ($\langle \cdot, \cdot \rangle$ is the inner product of $\mathfrak{L}^2$, ρ is the covariance)

$$\begin{aligned} \left\langle \sum_i \alpha_i X(t_i + \tau, \omega), \sum_j \beta_j X(s_j + \tau, \omega) \right\rangle &= \sum_{i,j} \alpha_i \bar{\beta}_j \rho(t_i - s_j) \\ &= \left\langle \sum_i \alpha_i X(t_i, \omega), \sum_j \beta_j X(s_j, \omega) \right\rangle, \quad \alpha_i, \beta_j \in \mathbb{C}, \end{aligned} \tag{8.39}$$

from which

$$\left\| \sum_i \alpha_i X(t_i + \tau, \omega) \right\|_2^2 = \left\| \sum_i \alpha_i X(t_i, \omega) \right\|_2^2 \tag{8.39$'$}$$

follows. As a special case of (8.39′), it follows that

$$\left\| \sum_{i=1}^{n} a_i X(t_i+\tau, \omega) - \sum_{j=1}^{m} b_j X(s_j+\tau, \omega) \right\|_2^2 = \left\| \sum_{i=1}^{n} a_i X(t_i, \omega) - \sum_{j=1}^{n} b_j X(s_j, \omega) \right\|_2^2.$$

Since the right-hand side is 0, (8.38) is verified.

Thus the operator U_τ is defined on $\mathfrak{M}$, and its images are also contained in $\mathfrak{M}$.

$$U_\tau : \mathfrak{M} \to \mathfrak{M}.$$

We list elementary properties of U_τ:

1° (linearity) $U_\tau(\alpha V + \beta W) = \alpha U_\tau V + \beta U_\tau W, \quad \alpha, \beta \in \mathbb{C}; \; V, W \in \mathfrak{M}.$

2° (group) $U_\tau U_\theta W = U_{\tau+\theta} W, \quad W \in \mathfrak{M}. \quad U_0 = I$ (identity).

3° (isometry) $\langle U_\tau V, U_\tau W \rangle = \langle V, W \rangle, \quad V, W \in \mathfrak{M}$. In particular, $\|U_\tau V\|_2^2 = \|V\|_2^2$ (by (8.39), (8.39′)).

4° (surjection) Any element

$$\sum_{i=1}^{n} a_i X(t_i, \omega)$$

of $\mathfrak{M}$ is a value of U_τ since

$$U_\tau\left(\sum_{i=1}^{n} a_i X(t_i - \tau, \omega) \right) = \sum_{i=1}^{n} a_i X(t_i, \omega).$$

U_τ can be uniquely extended to an isometric linear operator on $\mathfrak{H}(X)$. This extension is also denoted by the same notation U_τ. By Remark 8.3, 2° (p. 222),

$$U_\tau : \mathfrak{H}(X) \to \mathfrak{H}(X), \quad \tau \in \mathbb{R}$$

is a surjection. That is, U_τ is a unitary operator on $\mathfrak{H}(X)$ and $\{U_\tau | \tau \in \mathbb{R}\}$ is a one-parameter group.

We next show the strong continuity of this one-parameter group; i.e. for $\tau_0 \in \mathbb{R}$

$$s\text{-}\lim_{\tau \to \tau_0} U_\tau V = U_{\tau_0} V \quad \text{for all} \quad V \in \mathfrak{H}(X).$$

For any $V \in \mathfrak{H}(X)$ and $\varepsilon > 0$, there exists some $W \in \mathfrak{M}$ such that

$$\|V - W\|_2 < \frac{\varepsilon}{3}. \tag{8.40}$$

If $W(\omega)$ has the form

$$W(\omega) = \sum_{i=1}^{n} a_i X(t_i, \omega),$$

we have the evaluation

$$\|U_\tau W - U_{\tau_0} W\|_2 \leqq \sum_{i=1}^{n} |a_i| \cdot \|X(t_i + \tau, \omega) - X(t_i + \tau_0, \omega)\|_2.$$

By the strong continuity of X (Remark 8.3, 1° on p. 222),

$$\|U_\tau W - U_{\tau_0} W\|_2 < \frac{\varepsilon}{3} \tag{8.41}$$

provided that τ and τ_0 are sufficiently near.

$$\|U_\tau V - U_{\tau_0} V\|_2 \leqq \|U_\tau V - U_\tau W\|_2 + \|U_\tau W - U_{\tau_0} W\|_2 + \|U_{\tau_0} W - U_{\tau_0} V\|_2.$$

The first and third terms of the right-hand side are less than $\varepsilon/3$ by the isometry of U_τ and U_{τ_0} together with (8.40). And the second term is also less than $\varepsilon/3$ by (8.41), provided that $|\tau - \tau_0|$ is very small.

Hence

$$\|U_\tau V - U_{\tau_0} V\|_2 < \varepsilon$$

provided that $|\tau - \tau_0|$ is sufficiently small. This holds good for every $V \in \mathfrak{H}(X)$.

Thus it is confirmed that $\{U_\tau\}$ is a strongly continuous one-parameter group of unitary operators on $\mathfrak{H}(X)$. According to Stone's theorem (Theorems 7.13, 7.15, pp. 188–189), it has a representation

$$U_\tau = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\tau\lambda} E(d\lambda), \tag{8.42}$$

where $E(\cdot)$ is a resolution of the identity on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, the values of which are bounded symmetric operators. The more detailed expression of (8.42) is

$$\langle U_\tau V, W \rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-it\lambda} \langle E(d\lambda) V, W \rangle, \quad V, W \in \mathfrak{H}(X).$$

Define a function $\xi : \mathcal{B}(\mathbb{R}) \times \Omega \to \mathbb{C}$ by

$$\xi(S, \omega) = E(S) X(0, \omega), \quad S \in \mathcal{B}(\mathbb{R}).$$

We also define $\nu_\xi(S)$ $(S \in \mathcal{B}(\mathbb{R}))$ by[20]

$$\begin{aligned}\nu_\xi(S) &= \langle E(S)X(0,\omega),\ X(0,\omega)\rangle_{\mathfrak{L}^2(\Omega,\mathbb{C})} \\ &= \langle E^2(S)X(0,\omega),\ X(0,\omega)\rangle \qquad (E(S) \text{is a projection}) \\ &= \langle E(S)X(0,\omega),\ E(S)X(0,\omega)\rangle \qquad (E(S) \text{is symmetric}) \\ &= \|\xi(S,\omega)\|_2^2 \quad (\text{by the definition of } \xi), \quad S \in \mathcal{B}(\mathbb{R}),\end{aligned} \tag{8.43}$$

where $\langle\cdot,\cdot\rangle$ is the inner product of $\mathfrak{L}^2(\Omega,\mathbb{C})$.

Then ξ is an $\mathfrak{L}^2$-orthogonal measure. In fact, the set function $S \mapsto \xi(S,\omega)$ is σ-additive for any fixed $\omega \in \Omega$. The function $\omega \mapsto \xi(S,\omega)$ is square-integrable for any fixed $S \in \mathcal{B}(\mathbb{R})$. It remains to show the orthogonality. For $S_1, S_2 \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}\langle \xi(S_1,\omega), \xi(S_2,\omega)\rangle &= \langle E(S_1)X(0,\omega), E(S_2)X(0,\omega)\rangle \\ &= \langle E(S_2)E(S_1)X(0,\omega), X(0,\omega)\rangle \quad (E(\cdot) \text{ is symmetric}) \\ &= \langle E(S_1\cap S_2)X(0,\omega), X(0,\omega)\rangle \quad (3^\circ \text{on p. } 177) \\ &= \nu_\xi(S_1\cap S_2) \quad (\text{by } (8.43)).\end{aligned}$$

Hence if $S_1 \cap S_2 = \emptyset$,

$$\mathbb{E}\xi(S_1,\omega)\overline{\xi(S_2,\omega)} = \nu_\xi(S_1\cap S_2) = 0.$$

Finally, $X(t,\omega)$ can be represented as[21]

$$X(t,\omega) = U_t X(0,\omega) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-i\lambda t}E(d\lambda)X(0,\omega) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-i\lambda t}\xi(d\lambda,\omega).$$

This completes the proof. □

The relation

$$\nu(S) = \nu_\xi(S) \underset{(8.43)}{=} \|\xi(S,\omega)\|_2^2$$

[20] $\nu_\xi(S) = \|\xi(S,\omega)\|_2^2$ is a special case of Theorem 7.8(i) (p. 177).

[21] The last equality is verified as follows. We consider a simple function $\varphi(\lambda) = \sum_{i=1}^n \alpha_i \chi_{S_i}(\lambda)$ $(S_i \cap S_j = \emptyset$ if $i \neq j)$.

$$\int_{\mathbb{R}} \varphi(\lambda)dE(\lambda)X(0,\omega) = \sum_{i=1}^n \alpha_i E(S_i)X(0,\omega) = \sum_{i=1}^n \alpha_i \xi(S_i,\omega) = \int_{\mathbb{R}} \varphi(\lambda)\xi(d\lambda,\omega).$$

The integration of $e^{-i\lambda t}$ is a limit of such integrations of simple functions.

holds good for the covariance $\rho(u)$ and the spectral measure ν. In fact,

$$\begin{aligned}\rho(u) &= \langle X(u,\omega), X(0,\omega)\rangle \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-i\lambda u}\langle E(d\lambda)X(0,\omega), X(0,\omega)\rangle \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-i\lambda u} d\nu_\xi(\lambda).\end{aligned}$$

Consequently, $\nu = \nu_\xi$ by the uniqueness of the spectral measure.

8.6 Spectral Density Functions

As we saw in previous sections, *any periodic weakly stationary process can not have spectral density functions*.

Hence a problem arises: under what conditions does a weakly stationary process have a spectral density function?[22]

Let $\xi : \mathcal{B}(\mathbb{T}) \times \Omega \to \mathbb{C}$ be an orthogonal measure with $\mathbb{E}\xi(S,\omega) = 0$ for any $S \in \mathcal{B}(\mathbb{T})$. There exist a probability space $(\Omega', \mathcal{E}', P')$ and a weakly stationary process $Y_n(\omega') : \Omega' \to \mathbb{C}(n \in \mathbb{Z})$, the spectral measure of which is the Lebesgue measure m on $\mathbb{T}$ (by Theorem 8.7). We denote by $\mathbb{E}_\omega$ (resp. $\mathbb{E}_{\omega'}$) the expectation operator on $(\Omega, \mathcal{E}, P)$ (resp. $(\Omega', \mathcal{E}', P')$). The Cramér–Kolmogorov theorem assures the representation

$$Y_n(\omega') = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{T}} e^{-i\lambda n}\eta(d\lambda, \omega')$$

for some orthogonal measure $\eta : \mathcal{B}(\mathbb{T}) \times \Omega' \to \mathbb{C}$. (We assume that $\mathbb{E}_{\omega'}Y_n(\omega') = 0$.) η satisfies:

(a) $\mathbb{E}_{\omega'}\eta(S,\omega') = 0$ for any $S \in \mathcal{B}(\mathbb{T})$, and
(b) $\mathbb{E}_{\omega'}|\eta(S,\omega')|^2 = m(S)$ for any $S \in \mathcal{B}(\mathbb{T})$.

In order to express some relations between ξ and η, we need the "adjunction" method[23] as follows. $\mathbb{E}_{(\omega,\omega')}$ denotes the expectation operator on the product probability space $(\Omega \times \Omega', \mathcal{E} \otimes \mathcal{E}', P \otimes P')$. $1(\omega)$ (resp. $1(\omega')$) denotes the constant function which is identically 1 on Ω (resp. on Ω'). Then the following conditions hold good:

[22] This problem was studied by Doob [4] Chap. X, §8, Chap. XI, §8 and Kawata [17] pp. 69–73. I try to clarify the subtle details embedded in their works. cf. Maruyama [23].

[23] See Doob [4] Chap. II, §2.

(i) $\mathbb{E}_{(\omega,\omega')}\eta(S,\omega')1(\omega) = \mathbb{E}_{\omega'}\eta(S,\omega') = 0$ for any $S \in \mathcal{B}(\mathbb{T})$.
(ii) $\mathbb{E}_{(\omega,\omega')}|\eta(S,\omega')1(\omega)|^2 = \nu_\eta(S) = m(S)$ for any $S \in \mathcal{B}(\mathbb{T})$.
(iii) $\mathbb{E}_{(\omega,\omega')}\xi(S,\omega)1(\omega') \cdot \overline{\eta(S',\omega')1(\omega)} = \mathbb{E}_\omega\xi(S,\omega)\mathbb{E}_{\omega'}\overline{\eta(S',\omega')} = 0$ for any S and $S' \in \mathcal{B}(\mathbb{T})$.

Theorem 8.12 *Let $X_n(\omega)$ $(n \in \mathbb{Z})$ be a weakly stationary process with the spectral measure ν. Assume also that $\mathbb{E}X_n(\omega) = 0$. Then the following two statements are equivalent:*

(i) *$X_n(\omega)$ has the spectral density function.*
(ii) *$X_n(\omega)$ is a linear stochastic process; that is, there exist a sequence $\{c_n\}$ of complex numbers such that*

$$\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty$$

and a stochastic process $Z_n(\omega)(n \in \mathbb{Z})$ with

$$\begin{aligned} &\mathbb{E}Z_n(\omega) = 0, \quad \mathbb{E}|Z_n(\omega)|^2 = 1, \\ &\mathbb{E}Z_n(\omega)\overline{Z_m(\omega)} = 0 \quad \text{if} \quad n \neq m \end{aligned} \tag{8.44}$$

which satisfy

$$X_n(\omega) = \sum_{k=-\infty}^{\infty} c_{k-n} Z_k(\omega) \quad a.e.$$

(The convergence on the right-hand side is in $\mathfrak{L}^2(\Omega, \mathbb{C})$.)

Proof (i)⇒(ii): Let $p(\lambda) \geqq 0$ be the Radon–Nikodým derivative of ν. We also define

$$\alpha(\lambda) = \sqrt{p(\lambda)}.$$

Since $\alpha(\lambda) \in \mathfrak{L}^2(\mathbb{T}, \mathbb{C})$ (actually real-valued), it can be expanded by the Fourier series[24]:

$$\alpha(\lambda) \sim \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \alpha_k e^{ik\lambda}, \tag{8.45}$$

$$\sum_{k=-\infty}^{\infty} |\alpha_k|^2 < \infty,$$

where α_k is the Fourier coefficient corresponding to $(1/\sqrt{2\pi})e^{ik\lambda}(k \in \mathbb{Z})$.

[24] The convergence of the series in (8.45) is in $\mathfrak{L}^2(\mathbb{T}, \mathbb{C})$. However, the series is, actually, convergent a.e. and is equal to $\alpha(\lambda)$ according to Carleson's theorem. cf. Carleson [2].

By the Cramér–Kolmogorov theorem, $X_n(\omega)$ is represented as the Fourier transform of some orthogonal measure $\xi(S, \omega)$. According to the argument preceeding the theorem, there exist a probability space $(\Omega', \mathcal{E}', P')$ and an orthogonal measure $\eta(S, \omega') : \mathcal{B}(\mathbb{T}) \times \Omega' \to \mathbb{C}$ which satisfies (i), (ii) and (iii) mentioned above (p. 228).

Divide $\mathbb{T}$ into $\mathbb{T}_+$ and $\mathbb{T}_0$, defined as $\mathbb{T}_+ = \{t \in \mathbb{T} | \alpha(t) > 0\}$, $\mathbb{T}_0 = \{t \in \mathbb{T} | \alpha(t) = 0\}$.

We define a couple of functions, $\alpha_1(\lambda)$ and $\alpha_2(\lambda)$, by

$$\alpha_1(\lambda) = \begin{cases} \frac{1}{\alpha(\lambda)} & \text{on } \mathbb{T}_+, \\ 0 & \text{on } \mathbb{T}_0, \end{cases}$$

$$\alpha_2(\lambda) = \begin{cases} 0 & \text{on } \mathbb{T}_+, \\ 1 & \text{on } \mathbb{T}_0. \end{cases}$$

Furthermore, define a function $\gamma' : \mathcal{B}(\mathbb{T}) \times \Omega \times \Omega' \to \mathbb{C}$ by[25]

$$\begin{aligned} \gamma'(S, \omega, \omega') &= \int_S \alpha_1(\lambda)\xi(d\lambda, \omega)1(\omega') + \int_S \alpha_2(\lambda)\eta(d\lambda, \omega')1(\omega) \\ &= \int_S \alpha_1(\lambda)\xi(d\lambda, \omega) + \int_S \alpha_2(\lambda)\eta(d\lambda, \omega'). \end{aligned} \tag{8.46}$$

For the definition (8.46) to be possible, we have to check that $\alpha_1(\cdot) \in \mathfrak{L}^2_{\nu_\xi}$ and $\alpha_2(\cdot) \in \mathfrak{L}^2_{\nu_\eta}$, where

$$\nu_\xi(S) = \mathbb{E}_\omega |\xi(S, \omega)|^2 \quad \text{and} \quad \nu_\eta(S) = \mathbb{E}_{\omega'} |\eta(S, \omega')|^2, \quad S \in \mathcal{B}(\mathbb{T}),$$

respectively. But there is no difficulty in checking it as shown in the following:

$$\begin{aligned} \int_S |\alpha_1(\lambda)|^2 d\nu_\xi &= \int_{S \cap \mathbb{T}_+} \frac{1}{\alpha(\lambda)^2} p(\lambda) dm(\lambda) \\ &= \int_{S \cap \mathbb{T}_+} dm(\lambda) = m(S \cap \mathbb{T}_+) < \infty, \end{aligned} \tag{8.47}$$

$$\int_S |\alpha_2(\lambda)|^2 d\nu_\eta = \int_{S \cap \mathbb{T}_0} dm(\lambda) = m(S \cap \mathbb{T}_0) < \infty.$$

---

[25] In the case $\mathbb{T}_0 = \emptyset$ (and so $\alpha(\lambda)$ never vanishes), the discussion becomes much easier, since it is enough to define $\gamma(S, \omega)$ simply by

$$\gamma(S, \omega) = \int_S \frac{1}{\alpha(\lambda)} \xi(d\lambda, \omega)$$

for any $S \in \mathcal{B}(\mathbb{T})$. Clearly, $\nu_\gamma(S) = m(S)$.

$\gamma'(S, \omega, \omega')$ is an orthogonal measure. For instance, the orthogonality is proved as follows. Let S and $S' \in \mathcal{B}(\mathbb{T})$ be disjoint. Then we obtain

$$\begin{aligned}
\mathbb{E}_{(\omega,\omega')}\gamma'(S,\omega,\omega')\overline{\gamma'(S',\omega,\omega')} \\
&= \mathbb{E}_{\omega}\int_S \alpha_1(\lambda)\xi(d\lambda,\omega)\int_{S'}\alpha_1(\lambda)\overline{\xi(d\lambda,\omega)} \\
&+ \mathbb{E}_{(\omega,\omega')}\int_S \alpha_1(\lambda)\xi(d\lambda,\omega)\int_{S'}\alpha_2(\lambda)\overline{\eta(d\lambda,\omega')} \\
&+ \mathbb{E}_{(\omega,\omega')}\int_S \alpha_2(\lambda)\eta(d\lambda,\omega')\int_{S'}\alpha_1(\lambda)\overline{\xi(d\lambda,\omega)} \\
&+ \mathbb{E}_{\omega'}\int_S \alpha_2(\lambda)\eta(d\lambda,\omega')\int_{S'}\alpha_2(\lambda)\overline{\eta(d\lambda,\omega')}.
\end{aligned} \tag{8.48}$$

The second and third terms of (8.48) are zero because of the orthogonality of ξ and η in the sense of (iii) (p. 228). The first and the fourth terms are also zero because $S \cap S' = \emptyset$.

By a computation similar to (8.47) and (8.48), we have[26]

$$\mathbb{E}_{(\omega,\omega')}|\gamma'(S,\omega,\omega')|^2 = m(S), \quad S \in \mathcal{B}(\mathbb{T}).$$

Finally, a function $\gamma : \mathcal{B}(\mathbb{T}) \times \Omega \to \mathbb{C}$ is defined by

$$\gamma(S,\omega) = \mathbb{E}_{\omega'}\gamma'(S,\omega,\omega'). \tag{8.49}$$

26

$$\begin{aligned}
\mathbb{E}_{(\omega,\omega')}|\gamma'(S,\omega,\omega')|^2 \\
&= \mathbb{E}_{\omega}\left|\int_S \alpha_1(\lambda)\xi(d\lambda,\omega)\right|^2 \\
&+ \mathbb{E}_{\omega'}\left|\int_S \alpha_2(\lambda)\eta(d\lambda,\omega')\right|^2 \\
&+ 2\mathcal{R}e\mathbb{E}_{(\omega,\omega')}\int_S \alpha_1(\lambda)\xi(d\lambda,\omega)1(\omega')\int_S \overline{\alpha_2(\lambda)\eta(d\lambda,\omega')1(\omega)} \\
&= \int_S |\alpha_1(\lambda)|^2\nu_\xi(d\lambda) + \int_S |\alpha_2(\lambda)|^2\nu_\eta(d\lambda) \\
&+ 2\mathcal{R}e\mathbb{E}_{(\omega,\omega')}\int_{S\cap\mathbb{T}_+}\alpha_1(\lambda)\xi(d\lambda,\omega)\int_{S\cap\mathbb{T}_0}\overline{\alpha_2(\lambda)\eta(d\lambda,\omega')} \\
&= m(S\cap\mathbb{T}_+) + m(S\cap\mathbb{T}_0) + 0.
\end{aligned}$$

Taking account of the properties of γ', γ is shown to be an orthogonal measure with $\nu_\gamma(S) = m(S \cap \mathbb{T}_+)$.

The Cramér–Kolmogorov representation theorem gives[27]

$$\begin{aligned} X_n(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\lambda} \xi(d\lambda, \omega) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\lambda} \alpha(\lambda) \gamma(d\lambda, \omega). \end{aligned} \tag{8.50}$$

Since the Fourier series (8.45) converges to $\alpha(\lambda)$ in $\mathfrak{L}^2$, it follows that

$$X_n(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \int_{\mathbb{T}} \alpha_k e^{-i(n-k)\lambda} \gamma(d\lambda, \omega) \text{ a.e.} \tag{8.51}$$

In fact, (8.51) is verified by the computation:

$$\begin{aligned} &\mathbb{E}_\omega \left| X_n(\omega) - \frac{1}{2\pi} \sum_{k=-p}^{p} \int_{\mathbb{T}} \alpha_k e^{i(k-n)\lambda} \gamma(d\lambda, \omega) \right|^2 \\ &= \mathbb{E}_\omega \left| X_n(\omega) - \frac{1}{2\pi} \int_{\mathbb{T}} \sum_{k=-p}^{p} \alpha_k e^{i(k-n)\lambda} \gamma(d\lambda, \omega) \right|^2 \\ &= \mathbb{E}_\omega \left| \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\lambda} \left[\alpha(\lambda) - \frac{1}{\sqrt{2\pi}} \sum_{k=-p}^{p} \alpha_k e^{ik\lambda} \right] \gamma(d\lambda, \omega) \right|^2 \quad \text{(by (8.50))} \end{aligned}$$

27

$$\begin{aligned} &\int_{\mathbb{T}} e^{-in\lambda} \alpha(\lambda) \gamma(d\lambda, \omega) \\ &= \int_{\mathbb{T}_+} e^{-in\lambda} \alpha(\lambda) \cdot \frac{1}{\alpha(\lambda)} \xi(d\lambda, \omega) + \mathbb{E}_{\omega'} \int_{\mathbb{T}_0} e^{-in\lambda} \alpha(\lambda) \cdot \alpha_2(\lambda) \eta(d\lambda, \omega') \\ &= \int_{\mathbb{T}_+} e^{-in\lambda} \xi(d\lambda, \omega) = \int_{\mathbb{T}} e^{-in\lambda} \xi(d\lambda, \omega). \end{aligned}$$

The final equality is justified by

$$\begin{aligned} \mathbb{E}_\omega \left| \int_{\mathbb{T}_0} e^{-in\lambda} \xi(d\lambda, \omega) \right|^2 &= \int_{\mathbb{T}_0} \nu_\xi(d\lambda) \quad \text{(by D-K formula)} \\ &= \int_{\mathbb{T}_0} p(\lambda) dm = 0 \quad (p(\lambda) = 0 \text{ on } \mathbb{T}_0). \end{aligned}$$

$$= \frac{1}{2\pi} \int_{\mathbb{T}} \left| \alpha(\lambda) - \frac{1}{\sqrt{2\pi}} \sum_{k=-p}^{p} \alpha_k e^{ik\lambda} \right|^2 d\nu_\gamma(\lambda) \qquad \text{(by D-K formula)}$$

$$= \frac{1}{2\pi} \int_{\mathbb{T}_+} \left| \alpha(\lambda) - \frac{1}{\sqrt{2\pi}} \sum_{k=-p}^{p} \alpha_k e^{ik\lambda} \right|^2 dm(\lambda) \qquad (\text{by } \nu_\gamma(S) = m(S \cap \mathbb{T}_+))$$

$$\leqq \frac{1}{2\pi} \int_{\mathbb{T}} \left| \alpha(\lambda) - \frac{1}{\sqrt{2\pi}} \sum_{k=-p}^{p} \alpha_k e^{ik\lambda} \right|^2 dm(\lambda)$$

$$\to 0 \quad \text{as} \quad p \to \infty \qquad \text{(by (8.45))}.$$

If we define

$$Z_j(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{ij\lambda} \gamma(d\lambda, \omega),$$

and $c_k = (1/\sqrt{2\pi})\alpha_k$, then

$$X_n(\omega) = \sum_{k=-\infty}^{\infty} c_k Z_{k-n}(\omega) = \sum_{j=-\infty}^{\infty} c_{n+j} Z_j(\omega) \quad (\text{in } \mathfrak{L}^2).$$

It is easy to check that $Z_n(\omega)$ satisfies the condition (8.44). $\mathbb{E}Z_j(\omega) = 0$ is clear from the definition (8.49) of γ ($\mathbb{E} = \mathbb{E}_\omega$). The variance $= 1$ and the orthogonality come from

$$\begin{aligned}
&2\pi \mathbb{E} Z_n(\omega)\overline{Z_m(\omega)} \\
&= \mathbb{E} \int_{\mathbb{T}} e^{in\lambda} \gamma(d\lambda, \omega) \int_{\mathbb{T}} e^{-im\lambda} \overline{\gamma(d\lambda, \omega)} \\
&= \int_{\mathbb{T}} e^{i(n-m)\lambda} d\nu_\gamma = \int_{\mathbb{T}_+} e^{i(n-m)\lambda} dm(\lambda) \\
&= \int_{\mathbb{T}} e^{i(n-m)\lambda} dm(\lambda) - \int_{\mathbb{T}_0} e^{i(n-m)\lambda} d\nu_\gamma = \int_{\mathbb{T}} e^{i(n-m)\lambda} dm(\lambda) \\
&= \delta_{n,m} \times 2\pi.
\end{aligned}$$

The second equality is justified by D-K formula.

(ii)⇒(i): Conversely, assume that $X_n(\omega)$ is a moving average of a stochastic process $Z_n(\omega)$ which satisfies (8.44). Since $Z_n(\omega)$ is weakly stationary, there exists an orthogonal measure $\xi(S, \omega)$ on $\mathcal{B}(\mathbb{T}) \times \Omega$ such that

$$Z_n(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} e^{-in\lambda} \xi(d\lambda, \omega)$$

by the Cramér–Kolmogorov theorem. If we define $\nu_\xi(S) = \|\xi(S,\omega)\|_2^2$ as usual, ν_ξ is the spectral measure of $Z_n(\omega)$, which has the spectral density function $1/\sqrt{2\pi}$ (constant function).[28]

Consequently, we obtain, for some $\{c_n\} \in l_2(\mathbb{C})$, that

$$\begin{aligned} X_n(\omega) &= \underset{N\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \sum_{k=-N}^{N} c_{k-n} \int_{\mathbb{T}} e^{-ik\lambda} \xi(d\lambda,\omega) \\ &= \underset{N\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} \sum_{k=-N}^{N} c_{k-n} e^{-ik\lambda} \xi(d\lambda,\omega). \end{aligned} \tag{8.52}$$

Since $\{c_n\} \in l_2(\mathbb{C})$, c_n is the Fourier coefficient of some $C(\lambda) \in \mathfrak{L}^2(\mathbb{T},\mathbb{C})$, and[29]

$$\left\| \frac{1}{\sqrt{2\pi}} \sum_{k=-p}^{q} c_k e^{-ik\lambda} - C(\lambda) \right\|_2 \to 0 \quad \text{as} \quad p, q \to \infty.$$

Hence (writing $k - n = j$),

$$\sum_{k=-N}^{N} c_{k-n} e^{-ik\lambda} = e^{-in\lambda} \sum_{j=-N-n}^{N-n} c_j e^{-ij\lambda}$$

tends to $\sqrt{2\pi} e^{-in\lambda} C(\lambda)$ in $\mathfrak{L}^2$ as $N \to \infty$ for fixed n; i.e.

$$\left\| \sum_{k=-N}^{N} c_{k-n} e^{-ik\lambda} - \sqrt{2\pi} e^{-in\lambda} C(\lambda) \right\|_2 \to 0 \quad \text{as} \quad N \to \infty. \tag{8.53}$$

---

[28] Since $Z_n(\omega)$ is a white noise, the covariance is given by

$$\mathbb{E} Z_{n+u}(\omega)\overline{Z_n(\omega)} = \begin{cases} 1 & \text{if} \quad u = 0, \\ 0 & \text{if} \quad u \neq 0. \end{cases}$$

[29] The convergence of $(1/\sqrt{2\pi}) \sum_{k=-p}^{q} c_k e^{-ik\lambda}$ to $C(\lambda)$ also holds good "almost everywhere" thanks to the Carleson theorem. Hence c_k is the Fourier coefficient of $C(\lambda)$ corresponding to $(1/\sqrt{2\pi}) e^{-ik\lambda}$.

Taking account of the fact that the density function of ν_ξ is $1/\sqrt{2\pi}$ as remarked above, we have

$$
\begin{aligned}
&\mathbb{E}\Big|X_n(\omega) - \int_{\mathbb{T}} C(\lambda)e^{-in\lambda}\xi(d\lambda,\omega)\Big|^2 \\
&= \mathbb{E}\Big|\underset{N\to\infty}{\mathrm{l.i.m.}}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{T}}\sum_{k=-N}^{N} c_{k-n}e^{-ik\lambda}\xi(d\lambda,\omega) - \int_{\mathbb{T}} C(\lambda)e^{-in\lambda}\xi(d\lambda,\omega)\Big|^2 \quad \text{(by (8.52))} \\
&= \lim_{N\to\infty}\frac{1}{2\pi}\mathbb{E}\Big|\int_{\mathbb{T}}\Big[\sum_{j=-N-n}^{N-n} c_j e^{-ij\lambda} - \sqrt{2\pi}C(\lambda)\Big]e^{-in\lambda}\xi(d\lambda,\omega)\Big|^2 \\
&= \lim_{N\to\infty}\frac{1}{2\pi}\int_{\mathbb{T}}\Big|\sum_{j=-N-n}^{N-n} c_j e^{-ij\lambda} - \sqrt{2\pi}C(\lambda)\Big|^2 d\nu_\xi \quad \text{(by D-K formula)} \\
&= \lim_{N\to\infty}\int_{\mathbb{T}}\Big|\frac{1}{\sqrt{2\pi}}\sum_{j=-N-n}^{N-n} c_j e^{-ij\lambda} - C(\lambda)\Big|^2 \frac{1}{\sqrt{2\pi}}dm(\lambda) \\
&= 0 \quad \text{(by (8.53))}.
\end{aligned}
$$

Hence we obtain

$$
X_n(\omega) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{T}}\sqrt{2\pi}C(\lambda)e^{-in\lambda}\xi(d\lambda,\omega), \quad \text{a.e.}
$$

If we define $\theta(S,\omega)$ by

$$
\theta(S,\omega) = \int_S \sqrt{2\pi}C(\lambda)\xi(d\lambda,\omega), \quad S \in \mathcal{B}(\mathbb{T}),
$$

then $\theta(S,\omega)$ is an orthogonal measure[30] and

$$
X_n(\omega) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{T}} e^{-in\lambda}\theta(d\lambda,\omega).
$$

[30] The orthogonality, for instance, can be verified as follows. If S and $S' \in \mathcal{B}(\mathbb{T})$ are disjoint,

$$
\begin{aligned}
\mathbb{E}\theta(S,\omega)\overline{\theta(S',\omega)} &= \mathbb{E}\int_{\mathbb{T}} C(\lambda)\chi_S(\lambda)\xi(d\lambda,\omega)\int_{\mathbb{T}}\overline{C(\lambda)\chi_{S'}(\lambda)\xi(d\lambda,\omega)} \\
&= \int_{\mathbb{T}} |C(\lambda)|^2\chi_S(\lambda)\chi_{S'}(\lambda)d\nu_\xi = 0 \quad \text{(by D-K formula)}.
\end{aligned}
$$

This is the spectral representation of $X_n(\omega)$ in terms of $\theta(S, \omega)$. Consequently, the spectral measure ν of $X_n(\omega)$ is given by

$$\begin{aligned}\nu(S) &= \mathbb{E}|\theta(S, \omega)|^2 = 2\pi \int_S |C(\lambda)|^2 d\nu_\xi \\ &= \sqrt{2\pi} \int_S |C(\lambda)|^2 dm(\lambda),\end{aligned}$$

which is, of course, absolutely continuous with respect to m. □

We prepare a lemma to be used in the proof of Theorem 8.13. We denote by $\mathcal{F}_2$ (resp. $\mathcal{F}_2^{-1}$) the Fourier transform (resp. inverse Fourier transform) in the sense of Plancherel.[31]

Lemma 8.1

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(z)g(z)e^{-iuz} dm(z) = \frac{1}{\sqrt{2\pi}} \langle \mathcal{F}_2^{-1} f(\lambda - u), \mathcal{F}_2^{-1} \bar{g}(\lambda) \rangle$$

for any f and $g \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$.

(We have to note $f \cdot g \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$. $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathfrak{L}^2(\mathbb{R}, \mathbb{C})$.)

Proof By definition of the inverse Fourier transform in the sense of Plancherel, we have

$$\begin{aligned}\mathcal{F}_2^{-1}(f(z)e^{-izu})(\lambda) &= \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} f(z)e^{-izu}e^{iz\lambda} dm(z) \\ &= \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} f(z)e^{i(\lambda-u)z} dm(z) \qquad (8.54) \\ &= \mathcal{F}_2^{-1} f(\lambda - u).\end{aligned}$$

Taking account of $f \cdot g \in \mathfrak{L}^1$, we obtain

$$\begin{aligned}&\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(z) \cdot g(z)e^{-izu} dm(z) \\ &\qquad = \frac{1}{\sqrt{2\pi}} \langle f(z)e^{-izu}, \bar{g}(z) \rangle\end{aligned}$$

[31] We can also establish

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(\lambda)g(\lambda)e^{-iz\lambda} dm(\lambda) = \frac{1}{\sqrt{2\pi}} (\mathcal{F}_2 f * \mathcal{F}_2 g)(z).$$

(cf. Kawata [15] pp. 282–283.)

$$= \frac{1}{\sqrt{2\pi}} \langle \mathcal{F}_2^{-1}(f(z)e^{-izu})(\lambda), \mathcal{F}_2^{-1}\bar{g}(\lambda)\rangle$$

(by Parseval equality)

$$= \frac{1}{\sqrt{2\pi}} \langle \mathcal{F}_2^{-1} f(\lambda - u), \mathcal{F}_2^{-1}\bar{g}(\lambda)\rangle \quad \text{(by (8.54))}.$$

□

Theorem 8.13 *Let $X(t,\omega)(t \in \mathbb{R})$ be a measurable and weakly stationary process with the spectral measure ν. Assume also that $\mathbb{E}X(t,\omega) = 0$. Then the following two statements are equivalent:*

(i) *$X(t,\omega)$ has the spectral density function.*
(ii) *There exists a measurable and weakly stationary process $X' : \mathbb{R} \times \Omega \to \mathbb{C}$ of the form*

$$X'(t,\omega) = \int_{\mathbb{R}} w(\lambda - t)\gamma(d\lambda, \omega) \ a.e.\ (\omega),$$

the spectral measure of which is ν, where $w(\cdot) \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$ and the orthogonal measure $\gamma(S,\omega) : \mathcal{B}(\mathbb{R}) \times \Omega \to \mathbb{C}$ satisfies

$$\nu_\gamma(S) = \mathbb{E}|\gamma(S,\omega)|^2 = m(S \cap \mathbb{R}_+), \ \ S \in \mathcal{B}^*(\mathbb{R}).$$

(The definition of $\mathbb{R}_+$ is given below.)

Proof By the Cramér–Kolmogorov theorem, there exists an orthogonal measure $\xi(S,\omega)$ on $\mathcal{B}(\mathbb{R}) \times \Omega$ such that

$$X(t,\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\lambda t}\xi(d\lambda, \omega).$$

(i)⇒(ii): Let $p(\lambda) \geqq 0$ be the Radon–Nikodým derivative of ν. If we define

$$\alpha(\lambda) = \sqrt{p(\lambda)},$$

then $\alpha \in \mathfrak{L}^2(\mathbb{R}, \mathbb{R})$.

The covariance $\rho(u)$ of $X(t,\omega)$ can be written as

$$\begin{aligned}\rho(u) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iuz} d\nu(z) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iuz}\alpha(z) \cdot \alpha(z) dm(z) \\ &= \frac{1}{\sqrt{2\pi}} \langle \mathcal{F}_2^{-1}\alpha(\lambda - u),\ \mathcal{F}_2^{-1}\overline{\alpha}(\lambda)\rangle.\end{aligned} \tag{8.55}$$

The third equality is assured by Lemma 8.1.

Divide $\mathbb{R}$ into $\mathbb{R}_+$ and $\mathbb{R}_0$ defined as $\mathbb{R}_+ = \{t \in \mathbb{R} | \alpha(t) > 0\}$, $\mathbb{R}_0 = \{t \in \mathbb{R} | \alpha(t) = 0\}$.

As in Theorem 8.12, we introduce the orthogonal measure $\eta(S, \omega) : \mathcal{B}(\mathbb{R}) \times \Omega \to \mathbb{C}$, two functions $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$, and another orthogonal measure $\gamma(S, \omega)$. We can construct η and γ so as to be

$$\nu_\eta(S) = m(S), \quad \nu_\gamma = m(S \cap \mathbb{R}_+) \quad \text{for} \quad S \in \mathcal{B}(\mathbb{R}).$$

Note that the values of ν_η and ν_γ admit ∞. Since $\alpha(\cdot)$ is in $\mathfrak{L}^2$ with respect to the Lebesgue measure, it is also in $\mathfrak{L}^2$ with respect to ν_γ.

If we define a weakly stationary process $X'(t, \omega)$ by

$$X'(t, \omega) = \frac{1}{\sqrt[4]{2\pi}} \int_{\mathbb{R}} \alpha(\lambda - t)\gamma(d\lambda, \omega),$$

then the covariance $\rho'(u)$ of $X'(t, \omega)$ is[32]

$$\begin{aligned} \rho'(u) &= \mathbb{E} X'(t+u, \omega)\overline{X'(t, \omega)} \\ &= \frac{1}{\sqrt{2\pi}} \mathbb{E} \int_{\mathbb{R}} \alpha(\lambda - (t+u))\gamma(d\lambda, \omega) \int_{\mathbb{R}} \overline{\alpha(\lambda - t)\gamma(d\lambda, \omega)} \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \alpha(\lambda - (t+u))\overline{\alpha(\lambda - t)d}\nu_\gamma(\lambda) \qquad \text{(by D-K formula)} \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathcal{F}_2^{-1}\alpha(\lambda - u)\overline{\mathcal{F}_2^{-1}\alpha(\lambda)}dm(\lambda). \end{aligned} \tag{8.56}$$

[32] The third line of (8.56)

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}} \Big\{ \int_{\mathbb{R}_+} \alpha(\lambda - (t+u))\alpha(\lambda - t)dm(\lambda) + \int_{\mathbb{R}_0} \underbrace{\alpha(\lambda - (t+u))\alpha(\lambda - t)}_{(\dagger)} d\nu_\gamma \Big\} \\ &= \frac{1}{\sqrt{2\pi}} \Big\{ \int_{\mathbb{R}} (\dagger)dm(\lambda) - \int_{\mathbb{R}_0} (\dagger)dm(\lambda) + \int_{\mathbb{R}_0} (\dagger)d\nu_\gamma \Big\} \\ &= \frac{1}{\sqrt{2\pi}} \Big\{ \int_{\mathbb{R}} (\dagger)dm(\lambda) - \int_{\substack{\lambda\in\mathbb{R}_0 \\ \lambda - t\in\mathbb{R}_+}} (\dagger)dm(\lambda) - \int_{\substack{\lambda\in\mathbb{R}_0 \\ \lambda - t\in\mathbb{R}_0}} (\dagger)dm(\lambda) \\ &\qquad + \int_{\substack{\lambda\in\mathbb{R}_0 \\ \lambda - t\in\mathbb{R}_+}} (\dagger)d\nu_\gamma(\lambda) + \int_{\substack{\lambda\in\mathbb{R}_0 \\ \lambda - t\in\mathbb{R}_0}} (\dagger)d\nu_\gamma(\lambda) \Big\} \\ &= \frac{1}{\sqrt{2\pi}} \Big\{ \int_{\mathbb{R}} (\dagger)dm(\lambda) - \int_{(\mathbb{R}_0 - t)\cap\mathbb{R}_+} \alpha(\lambda' - u)\alpha(\lambda')dm(\lambda') \\ &\qquad + \int_{(\mathbb{R}_0 - t)\cap\mathbb{R}_+} \alpha(\lambda' - u)\alpha(\lambda')d\nu_\gamma(\lambda') \Big\}. \end{aligned}$$

The last two terms cancel out. So we obtain (8.56).

By (8.55) and (8.56), we get $\rho = \rho'$. That is, the spectral density function of $X'(t,\omega)$ is $p(\cdot)$.

(ii)⇒(i): Conversely, assume (ii). The covariance $\rho(u)$ of

$$X(t,\omega) = \int_{\mathbb{R}} w(\lambda - t)\gamma(d\lambda,\omega)$$

is given by

$$\rho(u) = \int_{-\infty}^{\infty} \mathcal{F}_2^{-1} w(\lambda - u)\overline{\mathcal{F}_2^{-1} w(\lambda)}dm(\lambda)$$

as in the calculation of $\rho'(\cdot)$ above.

Then we get, again by Lemma 8.1,

$$\begin{aligned}
\rho(u) &= \int_{\mathbb{R}} \mathcal{F}_2^{-1} w(\lambda - u)\overline{\mathcal{F}_2^{-1} w(\lambda)}dm(\lambda) \\
&= \langle \mathcal{F}_2^{-1} w(\lambda - u), \mathcal{F}_2^{-1} w(\lambda)\rangle \\
&= \int_{\mathbb{R}} |w(z)|^2 e^{-iuz} dm(z) \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \sqrt{2\pi}|w(z)|^2 e^{-iuz} dm(z).
\end{aligned}$$

Finally, if we define $p(z) = \sqrt{2\pi}|w(z)|^2$, $p(z)$ is the spectral density function of $X(t,\omega)$. □

8.7 A Note on Slutsky's Work

So far, we have investigated the periodic behaviors of weakly stationary stochastic processes.[33] In particular, most of this chapter is devoted to a characterization of the periodicity by means of the concept of spectral measures.

As we have already noted, it is E. Slutsky who initiated research on the periodic behaviors of moving average processes generated by white noises. He discovered a special case of moving average which asymptotically approaches to some sine or cosine curve. However, his work is rather experimental and is, regrettably, devoid of any general principle which characterizes the periodicity. This shortcoming seems to accrue from the lack of Fourier analytic viewpoint.

[33]This section was added according to Professor S. Kusuoka's suggestion. It is not included in the Japanese edition of this monograph.

In this section, we make a digression to a brief exposition of what Slutsky did in his pioneering work.

Let $V(n)$ be a function of discrete time $n \in \mathbb{Z}$. The operator B defined by

$$BV(n) = V(n-1), \quad n \in \mathbb{Z} \tag{8.57}$$

is called the **backward operator** or **lag operator**.[34] B^k denotes the repeated applications of B k-times; i.e.

$$B^k V(n) = V(n-k), \quad n \in \mathbb{Z}. \tag{8.58}$$

The backward operator is sometimes used in a polynomial form:

$$\kappa(B) = a_0 + a_1 B + \cdots + a_p B^p, \tag{8.59}$$

i.e.

$$\kappa(B)V(n) = a_0 V(n) + a_1 V(n-1) + \cdots + a_p V(n-p).$$

Let $(\Omega, \mathcal{E}, P)$ be a probability space and $X_n : \Omega \to \mathbb{R} (n \in \mathbb{Z})$ a white noise. Slutsky [28] constructs the following stochastic process based upon X_n. First take two-period moving sums:

$$\begin{aligned} X_n^{(1)} &= X_n + X_{n-1} = (1+B)X_n, \\ X_n^{(2)} &= X_n^{(1)} + X_{n-1}^{(1)} = (1+B)^2 X_n, \\ &\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots \\ X_n^{(\mu)} &= X_n^{(\mu-1)} + X_{n-1}^{(\mu-1)} = (1+B)^\mu X_n. \end{aligned} \tag{8.60}$$

Then take the second differences consecutively:

$$\begin{aligned} \Delta X_n^{(\mu)} &= X_n^{(\mu)} - X_{n-1}^{(\mu)} = (1-B)X_n^{(\mu)}, \\ \Delta^2 X_n^{(\mu)} &= \Delta X_n^{(\mu)} - \Delta X_{n-1}^{(\mu)} = (1-B)^2 X_n^{(\mu)}, \\ &\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots \\ \Delta^\nu X_n^{(\mu)} &= \Delta^{\nu-1} X_n^{(\mu)} - \Delta^{\nu-1} X_{n-1}^{(\mu)} = (1-B)^\nu X_n^{(\mu)}. \end{aligned} \tag{8.61}$$

μ and ν are some natural numbers.

---

[34] For backward operators, see Granger and Newbold [7] Chap. 1.

Slutsky proposes a weakly stationary stochastic process

$$Y_n^{\mu,\nu}(\omega) = \Delta^\nu X_n^{(\mu)}(\omega) = (1-B)^\nu (1+B)^\mu X_n(\omega), \tag{8.62}$$

where μ and ν are parametric natural numbers. This is a candidate, the behavior of which approaches asymptotically to a simple periodic curve, like sine/cosine.

Before we proceed to Slutsky's work, let us evaluate the second difference of a simple periodic function, say

$$y_n = A\cos\alpha n \quad (A > 0), \tag{8.63}$$

for the sake of comparison:

$$\begin{aligned}
\Delta^2 y_n &= y_n - 2y_{n-1} + y_{n-2} \\
&= A\{\cos\alpha n - 2\cos\alpha(n-1) + \cos\alpha(n-2)\} \\
&= A\left\{2\cos\frac{\alpha n + \alpha n - 2\alpha}{2}\cos\frac{\alpha n - \alpha n + 2\alpha}{2} - 2\cos\alpha(n-1)\right\} \qquad (8.64) \\
&= 2A\cos\alpha(n-1)(\cos\alpha - 1) \\
&= 2A(\cos\alpha - 1)y_{n-1}.
\end{aligned}$$

Since $|\cos\alpha| \leqq 1$, we obtain $-a \equiv 2A(\cos\alpha - 1) \leqq 0$ and (8.64) can be written as

$$\Delta^2 y_n = -a y_{n-1}. \tag{8.65}$$

Let $Z_n : \Omega \to \mathbb{R}$ be a weakly stationary stochastic process with $\mathbb{E}Z_n = 0$ and $\mathbb{E}|Z_n|^2 = \sigma^2$ (constant independent of n). If Z_n is exactly equal to y_n, the second difference of Z_n is given by $\Delta^2 Z_n = -aZ_{n-1}$. However, in general, this equality is not satisfied. The gap G_n between them, i.e. $\Delta^2 Z_n = -aZ_{n-1} + G_n$, is interpreted as a measure of deviation of Z_n from the periodic curve $y_n = A\cos\alpha n$ in the spirit of Slutsky.

The value of a which minimizes

$$\mathbb{E}(G_n^2) = \mathbb{E}(\Delta^2 Z_n + aZ_{n-1})^2 \tag{8.66}$$

is given by

$$\begin{aligned}
a &= -\frac{\mathbb{E}\Delta^2 Z_n \cdot Z_{n-1}}{EZ_{n-1}^2} = -\frac{E\Delta^2 Z_n \cdot Z_{n-1}}{\sigma^2} \\
&= -\frac{1}{\sigma^2}\{\mathbb{E}(Z_n - 2Z_{n-1} + Z_{n-2})Z_{n-1}\}
\end{aligned}$$

$$= -\frac{1}{\sigma^2}\{\mathbb{E}Z_n Z_{n-1} - 2\mathbb{E}Z_{n-1}^2 + \mathbb{E}Z_{n-1}Z_{n-2}\} \tag{8.67}$$
$$= -\frac{1}{\sigma^2}\{2\rho(1) - 2\sigma^2\}$$
$$= 2\left(1 - \frac{\rho(1)}{\sigma^2}\right).$$

$\rho(1)$ is the covariance of Z_n; i.e. $\rho(1) = \mathbb{E}Z_m Z_{m-1}$ for any $m \in \mathbb{Z}$. Substituting (8.67) into (8.66), we obtain

$$\mathbb{E}(G_n^2) = \mathbb{E}\left(\Delta^2 Z_n + 2\left(1 - \frac{\rho(1)}{\sigma^2}\right)Z_{n-1}\right)^2. \tag{8.68}$$

Since $\mathbb{E}G_n = 0$, $\mathbb{E}G_n^2$ is the variance of G_n. Hence, for any $t > 0$,

$$P\{\omega \in \Omega | \ |G_n| \geqq t\} \leqq \frac{\mathbb{E}(G_n^2)}{t^2} \tag{8.69}$$

as is well-known (Tchebichev's inequality).[35]

As we have already seen, Slutsky proposes a weakly stationary stochastic process

$$Y_n^{\mu,\nu}(\omega) = \Delta^\nu X_n^{(\mu)}(\omega) = (1-B)^\nu (1+B)^\mu X_n(\omega), \tag{8.62}$$

where μ and $\nu \in \mathbb{N} \cup \{0\}$ are parameters. This is a candidate, the behavior of which approaches asymptotically to a simple periodic curve like sine/cosine curves, as the parameters, μ and ν, go to infinity in a specific manner.

Slutsky obtained a couple of theorems. The first theorem elucidates a set of conditions under which a sequence of stochastic processes converges (in some sense) to a simple periodic function. The second theorem proves the convergence of $Y_n^{\mu,\nu}$ defined by (8.62) to a periodic function by means of the first theorem.

[I] *Suppose that* $\{Z_n^\mu(\omega)\}$ *is a sequence of (real-valued) stochastic processes which satisfy the conditions a–d:*

a. $\mathbb{E}Z_n^\mu = 0$ *for all* $n \in \mathbb{Z}$ *and* $\mu \in \mathbb{N} \cup \{0\}$.
b. $\mathbb{E}|\mathbb{Z}_n^\mu|^2 = \sigma_\mu^2$ *(constant independent of* n*).*
c. The autocorrelation r_m^μ *between* Z_n^μ *and* Z_{n-m}^μ *is denoted by*

$$r_m^\mu = \frac{\mathbb{E}Z_n^\mu Z_{n-m}^\mu}{\sigma_\mu^2} \text{ (independent of } n).$$

In particular, $|r_1^\mu| \leqq C < 1$ *for sufficiently large* μ.

[35] Loève [19] p. 11.

d. The correlation coefficient between $\Delta^2 Z_n^\mu$ and Z_{n-1}^μ, *that is*

$$\frac{\mathbb{E}\Delta^2 Z_n^\mu \cdot Z_{n-1}^\mu}{\sqrt{\mathbb{E}(\Delta^2 Z_n^\mu)^2 \mathbb{E}(Z_{n-1}^\mu)^2}} = \frac{\mathbb{E}\Delta^2 Z_n^\mu \cdot Z_{n-1}^\mu}{\sigma_\mu^2}$$

tends to -1 *as* $\mu \to \infty$.

Then there exists some sinusoid $A \cos 2\pi n/T$ *for which the sequence* $\{G_n^\mu\}$ *of gaps (defined by* $G_n^\mu = \Delta^2 Z_n^\mu + a^\mu Z_{n-1}^\mu$ *as above) converges to zero in probability; that is, for any* $\varepsilon, \eta > 0$, *there exists some* μ_0 *such that*

$$P\left\{\omega \in \Omega \,\middle|\, \left|\frac{G_n^\mu}{\sigma_\mu}\right| \geqq \varepsilon\right\} \leqq \eta \quad \text{for all} \ \ \mu \geqq \mu_0.$$

The stochastic process $Y_n^{\mu,\nu}$ defined above has two parameters, μ and ν. However, if we fix the ratio ν/μ at a constant, μ (or ν) may be regarded as the only parameter associated with the process. Thus we denote the process simply by Y_n^μ instead of $Y_n^{\mu,\nu}$.

Slutsky proves that the sequence $\{Y_n^\mu\}_{\mu\in\mathbb{N}}$ of stochastic processes satisfies all the requirements *a*–*d* of Theorem [I].

For instance, it follows that

$$|r_1^\mu| \leqq C < 1 \quad \text{for all} \ \ \mu$$

in this case, since

$$r_1^\mu = \frac{\mathbb{E}Y_n^\mu Y_{n-1}^\mu}{\sigma_\mu^2} = \frac{\mu - \nu}{\mu + \nu + 1} < \frac{1 - \nu/\mu}{\nu/\mu + 1}$$

and $\nu/\mu =$ constant. Thus the condition *c* is affirmed.

The autocorrelation between $\Delta^2 Y_n^\mu$ and Y_{n-1}^μ is evaluated as

$$-\sqrt{\frac{(2\nu+1)(\mu+\nu+2)}{(2\nu+3)(\mu+\nu+1)}}.$$

The value tends to -1 as $\mu \to \infty$, ν/μ being fixed. The condition *d* is thus confirmed.

Theorem [II] immediately follows.

[II] *The conclusion of Theorem [I] holds good for the sequence* $\{Y_n^\mu\}_{\mu\in\mathbb{N}}$ *defined above instead of* $\{Z_n^\mu\}_{\mu\in\mathbb{N}}$.

We shall skip the proofs and more rigorous discussions because Slutsky's work was done outside the world of Fourier analysis.[36]

[36] Sargent [26] Chap. 11 and Shinkai [27] Chaps.7–8 are very suggestive.

References

1. Brémaud, P.: Fourier Analysis and Stochastic Processes. Springer, Cham (2014)
2. Carleson, L.: On the convergence and growth of partial sums of fourier series. Acta Math. **116**, 135–157 (1966)
3. Crum, M.M.: On positive definite functions. Proc. London Math. Soc. **6**, 548–560 (1956)
4. Doob, J.L.: Stochastic Processes. Wiley, New York (1953)
5. Dudley, R.M.: Real Analysis and Probability. Wadsworth and Brooks, Pacific Grove (1988)
6. Frisch, R.: Propagation problems and inpulse problems in dynamic economics. In: Economic Essays in Honor of Gustav Cassel. Allen and Unwin, London (1933)
7. Granger, C.W.J., Newbold., P.: Forecasting Economic Time Series, 2nd edn. Academic, Orlando (1986)
8. Hamilton, J.D.: Time Series Analysis. Princeton University Press, Princeton (1994)
9. Hicks, J.R.: A Contribution to the Theory of the Trade Cycle. Oxford University Press, London (1950)
10. Itô, K.: Kakuritsu-ron (Probability Theory). Iwanami Shoten, Tokyo (1953) (Originally published in Japanese)
11. Kaldor, N.: A model of the trade cycle. Econ. J. **50**, 78–92 (1940)
12. Kawata, T.: Ohyo Sugaku Gairon (Elements of Applied Mathematics), I, II. Iwanami Shoten, Tokyo (1950, 1952) (Originally published in Japanese)
13. Kawata, T.: Fourier Henkan to Laplace Henkan (Fourier and Laplace Transforms). Iwanami Shoten, Tokyo (1957) (Originally published in Japanese)
14. Kawata, T.: On the Fourier series of a stationary stochastic process, I, II. Z. Wahrsch. Verw. Gebiete **6**, 224–245 (1966); **13**, 25–38 (1969)
15. Kawata, T.: Fourier Kaiseki (Fourier Analysis). Sangyo Tosho, Tokyo (1975) (Originally published in Japanese)
16. Kawata, T.: Fourier Kaiseki to Tokei (Fourier Analysis and Statistics). Kyoritsu Shuppan, Tokyo (1985) (Originally published in Japanese)
17. Kawata, T.: Teijyo Kakuritsu Katei (Stationary Stochastic Processes). Kyoritsu Shuppan, Tokyo (1985) (Originally published in Japanese)
18. Keynes, J.M.: The General Theory of Employment, Interest and Money. Macmillan, London (1936)
19. Loève, M.: Probability Theory, 3rd edn. van Nostrand, Princeton (1963)
20. Maruyama, G.: Harmonic Analysis of Stationary Stochastic Processes. Mem. Fac. Sci. Kyushu Univ. Ser. A, **4**, 45–106 (1949)
21. Maruyama, T.: Suri-keizaigaku no Hoho (Methods in Mathematical Economics). Sobunsha, Tokyo (1995) (Originally published in Japanese)
22. Maruyama, T.: Shinko Keizai Genron (Principles of Economics), 3rd edn. Iwanami Shoten, Tokyo (2013) (Originally published in Japanese)
23. Maruyama, T.: Fourier Analysis of Periodic Weakly Stationary Processes. A Note on Slutsky's Observation. Adv. Math. Eco. **20**, 151–180 (2016)
24. Samuelson, P.A.: Interaction between the multiplier analysis and the principle of acceleration. Rev. Econ. Stud. **21**, 75–78 (1939)
25. Samuelson, P.A.: A synthesis of the principle of acceleration and the multiplier. J. Polit. Econ. **47**, 786–797 (1939)
26. Sargent, T.J.: Macroeconomic Theory. Academic, New York (1979)
27. Shinkai, Y.: Keizai-hendo no Riron (Theory of Economic Fluctuations). Iwanami Shoten, Tokyo (1967) (Originally published in Japanese)
28. Slutsky, E.: Alcune proposizioni sulla teoria delle funzioni aleatorie. Giorn. Inst. Ital. degli Attuari **8**, 193–199 (1937)

29. Slutsky, E.: The summation of random causes as the source of cyclic processes. Econometrica **5**, 105–146 (1937)
30. Slutsky, E.: Sur les fonctions aléatoires presque-periodiques et sur la décomposition des fonctions aléatoires stationnaires en composantes. Actualités Sci. Ind. **738**, 38–55 (1938)
31. Wold, H.: A Study in the Analysis of Stationary Time Series, 2nd edn. Almquist and Wicksell, Uppsala (1953)

Chapter 9
Almost Periodic Functions and Weakly Stationary Stochastic Processes

It is a basic idea for the classical theory of Fourier series to express periodic functions as compositions of harmonic waves. This idea can be successfully extended to nonperiodic functions by means of Fourier transforms. However, we will be confronted with a lot of obstacles when we consider $\mathfrak{L}^p$-function spaces in the case $p > 2$.

H. Bohr developed an investigation to mitigate this difficulty in the 1920s. S. Bochner as well as J. von Neumann succeeded Bohr's approach, and their endeavors yielded an abundant harvest in the form of the theory of almost periodic functions.

In this chapter, we are mainly interested in its connection with the theory of weakly stationary stochastic processes. It was observed in the last chapter that the periodicity of a weakly stationary process is characterized by some regular discreteness of the spectral measure. On the other hand, the almost periodicity of a weakly stationary process is characterized by a nonregular discreteness of the spectral measure.[1]

9.1 Almost Periodic Functions

Definition 9.1 Let $f : \mathbb{R} \to \mathbb{C}$ be a function and ε a positive real number. $\tau \in \mathbb{R}$ is called an ε-**almost period** of f if

$$\sup_{x \in \mathbb{R}} |f(x-\tau) - f(x)| < \varepsilon.$$

[1] I am much indebted to Dunford and Schwartz [7] pp. 281–285, Katznelson [8] pp. 191–200, Kawata [9] II, pp. 96–103, 149–152, [10], [11] pp. 78–86, Loomis [12], Rudin [16, 17] for the contents of this chapter. The classical works cited above are Bohr [4, 5], von Neumann [19], Bochner [2] and Bochner and von Neumann [3].

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_9

f is said to be (uniformly) **almost periodic** in the sense of Bohr [5] if

(i) f is continuous, and
(ii) there exists $\Lambda = \Lambda(\varepsilon, f) \in \mathbb{R}$ for each $\varepsilon > 0$ such that any interval, the length of which is $\Lambda(\varepsilon, f)$, contains an ε-almost period of f.

The set $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$ of all the almost periodic functions of $\mathbb{R}$ into $\mathbb{C}$ forms a closed subalgebra of $\mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$, as will be shown later on.

Example 9.1 $\tau = 0$ is an ε-almost period for any $\varepsilon > 0$.

If f is a τ-periodic function, the period τ of f is an ε-almost period for any $\varepsilon > 0$.

Suppose that f is uniformly continuous. Then for any $\varepsilon > 0$, τ is an ε-almost period if $|\tau|$ is sufficiently small.

Example 9.2 Any periodic continuous function $f : \mathbb{R} \to \mathbb{C}$ is almost periodic.

$f(x) = \cos x + \cos \pi x$ is not periodic.[2] However, $f(x)$ is almost periodic because, as we show later on (Theorem 9.2), a sum of almost periodic functions is almost periodic.

If f is almost periodic, $|f|$, $\bar{f}$, af, $f(\lambda x)$ are also almost periodic, where $a \in \mathbb{C}$, $\lambda \in \mathbb{R}$.

The following two lemmata show the boundedness and the uniform continuity of an almost periodic function.

Lemma 9.1 *Suppose that $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$ and $\varepsilon > 0$. Let I be any closed interval of length $\Lambda(\varepsilon, f)$. Then $f(\mathbb{R})$ is contained in the ε-neighborhood of $f(I)$. Therefore f is bounded.*

Proof Choose any $x \in \mathbb{R}$. Let τ be an ε-almost period contained in $x - I$. Of course $x - \tau \in I$, and

$$|f(x) - f(x - \tau)| < \varepsilon. \tag{9.1}$$

Hence $f(\mathbb{R})$ is contained in the ε-neighborhood of $f(I)$; i.e.

$$\sup_{x \in \mathbb{R}} |f(x)| \leqq \sup_{x \in I} |f(x)| + \varepsilon.$$

So f is bounded. □

Corollary 9.1 *If $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, then $f^2 \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$.*

[2]In fact, we note that $f(0) = 2$ and $f(x) \neq 2$ if $x \neq 0$. If $f(x) = 2$, we must have $\cos x = \cos \pi x = 1$. It follows that both $x = 2k\pi$ $(k \in \mathbb{Z})$ and $\pi x = 2k\pi$ $(k \in \mathbb{Z})$ would have to hold good at the same time. However, it is impossible.

Proof For any $x, \tau \in \mathbb{R}$,

$$f^2(x-\tau) - f^2(x) = (f(x-\tau) + f(x))(f(x-\tau) - f(x)).$$

Let ε be any positive number and τ an $\varepsilon/(2\|f\|_\infty)$-almost period. Then it follows that

$$|f^2(x-\tau) - f^2(x)| \leqq 2\|f\|_\infty \cdot \frac{\varepsilon}{2\|f\|_\infty} = \varepsilon,$$

which shows that τ is an ε-almost period of f^2. □

Lemma 9.2 *Any almost periodic function is uniformly continuous.*

Proof Assume $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$. For $\varepsilon > 0$, we define $\Lambda = \Lambda(\varepsilon/3, f)$. f is uniformly continuous on the interval $[0, \Lambda]$. Hence it follows that, for a sufficiently small $\delta(\varepsilon) > 0$,

$$\sup_{x \in [0,\Lambda]} |f(x+\eta) - f(x)| < \frac{\varepsilon}{3} \quad \text{if} \quad |\eta| < \delta(\varepsilon). \tag{9.2}$$

Let $y \in \mathbb{R}$ and τ an $\varepsilon/3$-almost period of f contained in the interval $[y - \Lambda, y]$. Then

$$|f(y+\eta)-f(y)| \leqq |f(y+\eta)-f(y-\tau+\eta)|+|f(y-\tau+\eta)-f(y-\tau)|+|f(y-\tau)-f(y)|. \tag{9.3}$$

The second term of the right-hand side of (9.3) is less than $\varepsilon/3$ if $|\eta| < \delta(\varepsilon)$. (Note that $y - \tau \in [0, \Lambda]$.) The first and the third terms on the right-hand side are less than $\varepsilon/3$, respectively, since τ is an $\varepsilon/3$-almost period. Hence we have, by (9.3).

$$|f(y+\eta) - f(y)| < \varepsilon \quad \text{if} \quad |\eta| < \delta(\varepsilon).$$

□

9.2 $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$ as a Closed Subalgebra of $\mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$

In the preceding section, we showed that an almost periodic function is bounded and uniformly continuous. We now proceed to prove that $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$ is a closed subalgebra of $\mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$.

For $f \in \mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$, we write

$$W_0(f) = \{f_y(x) = f(x-y) | y \in \mathbb{R}\}. \tag{9.4}$$

Theorem 9.1 (Bochner) *The following two statements are equivalent for* $f \in \mathfrak{L}^2(\mathbb{R}, \mathbb{C})$*:*

(i) *f is an almost periodic function.*
(ii) *$W_0(f)$ is relatively compact in $\mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$.*

Proof (i)⇒(ii): Since $\mathfrak{L}^\infty$ is a complete metric space, it is sufficient to verify that $W_0(f)$ is totally bounded in order to show its relative compactness. Let $\Lambda = \Lambda(\varepsilon/2, f)$ for $\varepsilon > 0$. We make a decomposition $\eta_1 < \eta_2 < \cdots < \eta_J$ of $[0, \Lambda]$. By the uniform continuity of f (Theorem 9.2), we obtain for any $y_0 \in [0, \Lambda]$

$$\inf_{1 \leqq j \leqq J} \| f_{y_0} - f_{\eta_j} \|_\infty < \frac{\varepsilon}{2} \tag{9.5}$$

provided that the decomposition is sufficiently fine.[3] For any $y \in \mathbb{R}$, let τ be an $\varepsilon/2$-almost period of f contained in $[y - \Lambda, y]$. $y_0 \equiv y - \tau \in [0, \Lambda]$ and

$$\| f_y - f_{y_0} \|_\infty = \| f_y - f_{y-\tau} \|_\infty < \frac{\varepsilon}{2}. \tag{9.6}$$

It follows from (9.5) and (9.6) that

$$\inf_{1 \leqq j \leqq J} \| f_y - f_{\eta_j} \|_\infty < \varepsilon \quad \text{for any} \quad y \in \mathbb{R}. \tag{9.7}$$

This shows that f_{η_j} $(1 \leqq j \leqq J)$ is an ε-net of $W_0(f)$. Hence $W_0(f)$ is totally bounded.

(ii)⇒(i): Conversely, assume that $W_0(f)$ is totally bounded. Then, for any $\varepsilon > 0$, there exist some $y_j \in \mathbb{R}$ $(1 \leqq j \leqq J)$ such that

$$W_0(f) \subset \bigcup_{j=1}^{J} B_\varepsilon(f_{y_j}). \tag{9.8}$$

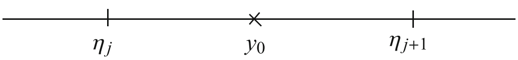
[3] For any $\varepsilon > 0$, there exists some $\delta > 0$ such that

$$|f(u) - f(v)| < \frac{\varepsilon}{2} \quad \text{if} \quad |u - v| < \delta.$$

We make a decomposition $\eta_j (1 \leqq j \leqq J)$ of $[0, \Lambda]$ so that distances of any two adjacent points are less than δ. $y_0 \in [0, \Lambda]$ is contained in a small interval $[\eta_j, \eta_{j+1}]$. Hence

$$|f(x - y_0) - f(x - \eta_j)| < \frac{\varepsilon}{2}.$$

(Note that $|(x - y_0) - (x - \eta_j)| = |y_0 - \eta_j| < \delta$.)



This holds good for any $x \in \mathbb{R}$, and so (9.5) follows.

We may assume that

$$B_\varepsilon(f_{y_j}) \cap W_0(f) \neq \emptyset \quad \text{for all } j, \tag{9.9}$$

without loss of generality. Define

$$\Lambda = 2 \operatorname*{Max}_{1 \leqq j \leqq J} |y_j|. \tag{9.10}$$

We now show that any interval I of length Λ contains an ε-almost period of f. Let y be the midpoint of I; i.e.

$$I = \left[y - \frac{\Lambda}{2}, y + \frac{\Lambda}{2} \right]. \tag{9.11}$$

By (9.8), there exists some $j_0 \in \{1, 2, \cdots, J\}$ such that $\| f_y - f_{y_{j_0}} \|_\infty < \varepsilon$.[4] Let $\tau = y - y_{j_0}$. Then it is clear that $\tau \in I$, since $|y_{j_0}| \leqq \Lambda/2$ by (9.10). (cf. Fig. 9.1.)

On the other hand,

$$\| f_\tau - f \|_\infty = \| f_{\underbrace{\tau + y_{j_0}}_{=y}} - f_{y_{j_0}} \|_\infty < \varepsilon. \tag{9.12}$$

Thus I contains an ε-almost period τ of f.

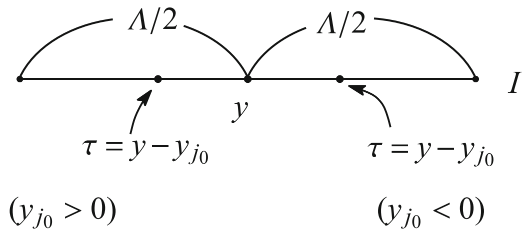
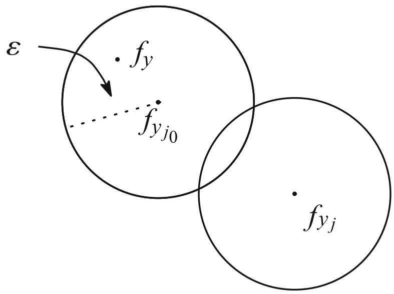


Fig. 9.1 Relation between y and y_{j_0}

4



Finally, we have to show that f is continuous. We can actually prove a stronger result,

$$\lim_{\eta\to 0}\|f_\eta - f\|_\infty = 0 \qquad \text{(uniform continuity).} \tag{9.13}$$

Let f_{y_j} $(1 \leqq j \leqq J)$ be an ε-net of $W_0(f)$ which satisfies (9.8). If we define E_j by

$$E_j = \{\tau \,|\, f_\tau \in B_\varepsilon(f_{y_j})\}, \quad 1 \leqq j \leqq J, \tag{9.14}$$

we have

$$\bigcup_{j=1}^{J} E_j = \mathbb{R}.$$

Hence there exists at least one E_j, say E_{j_0}, which has a positive measure. It is well-known that $E_{j_0} - E_{j_0}$ is a neighborhood of 0.[5] $y = y' - y''$ $(y', y'' \in E_{j_0})$ satisfies

$$\begin{aligned}|f_y(x) - f(x)| &\leqq |f(x-(y'-y'')) - f(x+y''-y_{j_0})| + |f(x+y''-y_{j_0}) - f(x)|\\ &= |f(x+y''-y') - f(x+y''-y_{j_0})|\\ &\qquad + |f(x+y''-y'') - f(x+y''-y_{j_0})|\\ &< \varepsilon + \varepsilon = 2\varepsilon,\end{aligned}$$

from which (9.13) follows. □

Definition 9.2 Let $f \in \mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$. The **translation closed convex hull** is the set

$$W(f) = \overline{\text{co}} \bigcup_{|a|\leqq 1} W_0(af).^{6}$$

Remark 9.1

1° $W(f)$ is the set of all the $\|\cdot\|_\infty$-limit of functions of the form

$$\sum a_k f_{x_k}, \quad x_k \in \mathbb{R}, \quad \sum |a_k| \leqq 1. \tag{9.15}$$

To prove this, we first note that an element of

$$\text{co} \bigcup_{|a|\leqq 1} W_0(af)$$

[5] This delicate result is called **Steinhaus' theorem**. See Dudley [6] p. 80 and Stromberg [18] pp. 297–298.

[6] $\overline{\text{co}}A$ is the closed convex hull of a set A. co$A$ is the convex hull of A.

can be expressed in the form

$$\sum_{k=1}^{p} \alpha_k \cdot a_k f_{x_k}, \tag{9.16}$$

$$|a_k| \leqq 1, \ x_k \in \mathbb{R}, \ \alpha_k \geqq 0, \ \sum_{k=1}^{p} \alpha_k = 1.$$

Writing $\beta_k = \alpha_k \cdot a_k$, we can rewrite (9.16) as

$$\sum_{k=1}^{p} \beta_k f_{x_k}, \quad \sum_{k=1}^{p} |\beta_k| \leqq 1.$$

Hence any element of $W(f)$ is a $\|\cdot\|_\infty$-limit of functions of the form (9.15).

Conversely, consider a function of the form (9.15). Writing

$$a_k = |a_k| e^{i\theta_k}$$

(without loss of generality, we may assume $\sum |a_k| \neq 0$), we have

$$\sum_{k=1}^{p} a_k f_{x_k} = \sum_{k=1}^{p} |a_k| e^{i\theta_k} f_{x_k} = \sum_{k=1}^{p} \frac{|a_k| e^{i\theta_k}}{\sum_{j=1}^{p} |a_j|} \left(\sum_{j=1}^{p} |a_j| \right) f_{x_k}.$$

Furthermore, if we define $\alpha_k = |a_k| / \sum |a_j|$, α_k's are nonnegative real numbers, the sum of which is equal to 1. On the other hand, $a_k' = e^{i\theta_k}(\sum |a_j|)$ is a complex number with $|a_k'| \leqq 1$. Hence any function of the form (9.15) can be expressed in the form (9.16). It follows that any $\|\cdot\|_\infty$-limit of functions of the form (9.15) is contained in $W(f)$.

2° If f is uniformly continuous, in particular, $W(f)$ is equal to the $\|\cdot\|_\infty$-closure of the set of functions of the form

$$\varphi * f, \quad \varphi \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C}), \quad \|\varphi\|_1 \leqq 1.$$

In order to see this, it is sufficient to show that any function of the form

$$\sum_{k=1}^{K} a_k f_{x_k}, \quad x_k \in \mathbb{R}, \quad \sum_{k=1}^{K} |a_k| \leqq 1$$

can be $\|\cdot\|_\infty$-approximated by a convolution $\varphi * f$, $\varphi \in \mathfrak{L}^1$, $\|\varphi\|_1 \leqq 1$. Since f is uniformly continuous, there exists some $\delta > 0$, for each $\varepsilon > 0$, such that

$$|f(u) - f(v)| < \varepsilon \quad \text{if} \quad |u - v| < \delta,$$

where it is possible to choose $\delta > 0$ so that $[x_k - \delta, x_k + \delta]$ $(k = 1, 2, \cdots, K)$ are mutually disjoint. Define a function $\varphi(y)$ by

$$\varphi(y) = \begin{cases} \dfrac{a_k}{2\delta} & \text{for} \quad y \in [x_k - \delta, x_k + \delta], \\ 0 & \text{for} \quad y \notin \displaystyle\bigcup_{k=1}^{K} [x_k - \delta, x_k + \delta]. \end{cases}$$

Since $\sum |a_k| \leqq 1$, it is obvious that $\|\varphi\|_1 \leqq 1$. We also have

$$\begin{aligned} &\left| \int_{x_k-\delta}^{x_k+\delta} f(x-y)\varphi(y)dy - a_k f(x - x_k) \right| \\ &\qquad = \left| \frac{1}{2\delta} \int_{x_k-\delta}^{x_k+\delta} f(x-y)a_k dy - \frac{1}{2\delta} \int_{x_k-\delta}^{x_k+\delta} a_k f(x-x_k)dy \right| \\ &\qquad \leqq \frac{|a_k|}{2\delta} \int_{x_k-\delta}^{x_k+\delta} |f(x-y) - f(x-x_k)|dy \leqq \frac{|a_k|}{2\delta} \cdot 2\delta \cdot \varepsilon = \varepsilon |a_k|, \end{aligned}$$

where $|(x - y) - (x - x_k)| = |x_k - y| < \delta$. Summing up for all k, we obtain

$$\left| \int_{\mathbb{R}} f(x-y)\varphi(y)dy - \sum_{k=1}^{K} a_k f(x - x_k) \right| \leqq \varepsilon \sum_{k=1}^{K} |a_k| \leqq \varepsilon.$$

3° For $f \in \mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$,

$$W(e^{i\xi x} f) = \{e^{i\xi x} g | g \in W(f)\}.$$

To prove this, we first note that any element of $\operatorname{co} \bigcup_{|a| \leqq 1} W_0(ae^{i\xi x} f)$ can be expressed in the form

$$\begin{aligned} \sum_k a_k e^{i\xi(x-x_k)} f(x - x_k) &= e^{i\xi x} \sum_k \underbrace{a_k e^{-i\xi x_k}}_{=b_k} f(x - x_k) \\ &= e^{i\xi x} \sum_k b_k f(x - x_k) \in e^{i\xi x} \operatorname{co} \bigcup_{|b| \leqq 1} W_0(bf), \end{aligned}$$

where $x_k \in \mathbb{R}$, $\sum |a_k| \leqq 1$, according to 1°. Hence

$$\operatorname{co} \bigcup_{|a| \leqq 1} W_0(ae^{i\xi x} f) \subset e^{i\xi x} \operatorname{co} \bigcup_{|b| \leqq 1} W_0(bf). \tag{9.17}$$

Conversely, if $g \in \operatorname{co} \bigcup_{|b| \leqq 1} W_0(bf)$,

$$\begin{aligned} e^{i\xi x} g(x) &= \sum_k b_k f(x - x_k) \cdot e^{i\xi x} = \sum_k b_k f(x - x_k) e^{i\xi(x-x_k)} \cdot e^{i\xi x_k} \\ &= \sum_k \underbrace{b_k e^{i\xi x_k}}_{=a_k} f(x - x_k) e^{i\xi(x-x_k)} = \sum_k a_k e^{i\xi(x-x_k)} f(x - x_k), \end{aligned}$$

where $x_k \in \mathbb{R}$, $\sum |a_k| \leqq 1$. Since $\sum |b_k| \leqq 1$,

$$\operatorname{co} \bigcup_{|a| \leqq 1} W_0(ae^{i\xi x} f) \supset e^{i\xi x} \operatorname{co} \bigcup_{|b| \leqq 1} W_0(bf). \tag{9.18}$$

By (9.17) and (9.18), both sides are equal. Taking closures of them, we have

$$W(e^{i\xi x} f) = e^{i\xi x} W(f).$$

3° is thus confirmed.

Lemma 9.3 *For $f \in \mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$, the following three statements are equivalent:*

(i) *$W(f)$ is compact with respect to $\| \cdot \|_\infty$.*
(ii) *$W_0(f)$ is relatively compact with respect to $\| \cdot \|_\infty$.*
(iii) *f is almost periodic.*

Proof (ii)⇔(iii) is known by Theorem 9.1 and (i)⇒(ii) is obvious. So it remains to show (ii)⇒(i).

Let $h \in W(f)$. For any $\varepsilon > 0$, there exists $g \in \operatorname{co} \bigcup_{|a| \leqq 1} W_0(af)$ such that

$$\|h - g\|_\infty < \varepsilon. \tag{9.19}$$

We express g as

$$g(x) = \sum_k a_k f_{x_k}(x), \quad \sum_k |a_k| \leqq 1 \tag{9.20}$$

(cf. Remark 9.1, 1°).

Since $W_0(f)$ is relatively compact, there exists an ε-net $\{f_{y_1}, f_{y_2}, \cdots, f_{y_J}\}$ of $W_0(f)$ consisting of finite points. Each f_{x_k} in (9.20) is contained in an ε-neighborhood of some f_{y_j} $(1 \leqq j \leqq J)$. Let $j(k)$ be any one of such j's. Then we have

$$\begin{aligned}\left|\sum_k a_k f_{x_k}(x) - \sum_k a_k f_{y_{j(k)}}(x)\right| &= \left|\sum_k a_k\{f_{x_k}(x) - f_{y_{j(k)}}(x)\}\right| \\ &\leqq \sum_k |a_k| \cdot |f_{x_k}(x) - f_{y_{j(k)}}| = \varepsilon.\end{aligned} \tag{9.21}$$

Collecting k's such that $j(k) = j$, for each j we write

$$\sharp(j) = \{k | j(k) = j\}.$$

Then we obtain

$$\sum_k a_k f_{y_{j(k)}}(x) = \sum_{j=1}^{J} \left\{ \sum_{k \in \sharp(j)} a_k f_{y_j}(x) \right\} = \sum_{j=1}^{J} \underbrace{\left(\sum_{k \in \sharp(j)} a_k \right)}_{=b_j} f_{y_j}(x) = \sum_{j=1}^{J} b_j f_{y_j}(x), \tag{9.22}$$

where $\sum |b_j| \leqq 1$. It follows from (9.21) and (9.22) that

$$\left\| g(x) - \sum_{j=1}^{J} b_j f_{y_j}(x) \right\|_\infty < \varepsilon. \tag{9.23}$$

If we introduce the norm $\|z\|_1 = \sum_{j=1}^{J} |z_j|$ in $\mathbb{R}^J$, $(b_1, b_2, \cdots, b_J)$ is contained in its unit ball. Let $c_1, c_2, \cdots, c_N$ be an $\varepsilon J^{-1} \|f\|_\infty^{-1}$-net of the unit ball. Then

$$\left\| \sum_j b_j f_{y_j}(x) - \sum_j b'_j f_{y_j}(x) \right\|_\infty < \varepsilon \tag{9.24}$$

for some $b'_j \in \{c_1, c_2, \cdots, c_N\}$.[7] By (9.19), (9.23) and (9.24),

[7]Select some $c_n \in \{c_1, c_2, \cdots, c_N\}$ such that $\|(b_1, b_2, \cdots, b_J) - c_n\|_1 < \varepsilon J^{-1} \|f\|_\infty^{-1}$ and express this c_n as

$$c_n = (c_{n,1}, c_{n,2}, \cdots, c_{n,J}) = (b'_1, b'_2, \cdots, b'_J).$$

Then the following evaluation follows.

$$\left\| \sum_j b_j f_{y_j}(x) - \sum_j c_{n,j} f_{y_j}(x) \right\|_\infty \leqq \left\| \sum_j (b_j - c_{n,j}) f_{y_j}(x) \right\|_\infty = \left\| \sum_j (b_j - b'_j) f_{y_j}(x) \right\|_\infty < \varepsilon.$$

$$\left\| h - \sum_j b'_j f_{y_j} \right\|_\infty < 3\varepsilon.$$

Thus

$$\left\{ \sum_j \alpha_j f_{y_j} \middle| \begin{array}{c} j = 1, 2, \cdots, J, \\ \alpha_j \in \{c_1, c_2, \cdots, c_N\} \end{array} \right\}$$

is a 3ε-net of $W(f)$ consisting of $J \times N$ elements.

Hence $W(f)$ is totally bounded. It is also closed (complete) and so compact. □

Based upon the preparations made above, we can prove the following fundamental theorem.

Theorem 9.2 $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$ *is a closed subalgebra of* $\mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$.

Proof It is clear that

$$W(f+g) \subset W(f) + W(g)$$

for any $f, g \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$. Since $W(f)$ and $W(g)$ are compact by Lemma 9.3, $W(f+g)$ is relatively compact. $W(f+g)$ being closed, however, it is compact. Again by Lemma 9.3, $f + g \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$.

Since $f^2, g^2, (f+g)^2 \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, it is easy to see

$$fg = \frac{1}{2}\{(f+g)^2 - f^2 - g^2\} \in \mathfrak{AP}(\mathbb{R}, \mathbb{C}).$$

Hence $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$ is a subalgebra of $\mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$.

Finally, we have to show that $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$ is closed. For any $f \in \overline{\mathfrak{AP}(\mathbb{R}, \mathbb{C})}$ and $\varepsilon > 0$, there exists some $g \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$ such that

$$\|f - g\|_\infty < \frac{\varepsilon}{3}.$$

Let τ be a $\varepsilon/3$-almost period of g. Then we obtain

$$\|f_\tau - f\|_\infty \leqq \|f_\tau - g_\tau\|_\infty + \|g_\tau - g\|_\infty + \|g - f\|_\infty < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

That is, τ is an ε-almost period of f. So any interval of the length $\Lambda(\varepsilon/3, g)$ contains an ε-almost period of f. Thus f is an almost periodic function. □

9.3 Spectrum of Almost Periodic Functions

We mean by a **trigonometric polynomial** a function of the form

$$f(x) = \sum_{j=1}^{n} a_j e^{i\xi_j x}, \quad \xi_j \in \mathbb{R}.$$

$\xi_j (j = 1, 2, \cdots, n)$ is called a **frequency** of f.

Note that ξ_j may be any real number in this definition, while ξ_j is assumed to be an integer in the ordinary usage of the term "trigonometric polynomial".

Remark 9.2 Each of $e^{i\xi_j x}$ is periodic, and so almost periodic. Hence, by Theorem 9.2, trigonometric polynomials and their uniform limits are almost periodic.

If a measure $\mu \in \mathfrak{M}(\mathbb{R})$ is discrete, i.e.

$$\mu = \sum a_j \delta_{\xi_j}, \quad \sum |a_j| < \infty$$

(δ_{ξ_j} is a Dirac measure concentrating at ξ_j), then its Fourier transform

$$\hat{\mu}(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-itx} d\mu(t) = \frac{1}{\sqrt{2\pi}} \sum a_j e^{-i\xi_j x}$$

is almost periodic.

Definition 9.3 Let $f \in \mathfrak{L}^{\infty}(\mathbb{R}, \mathbb{C})$. The **norm spectrum** $\sigma(f)$ of f is the set of $\xi \in \mathbb{R}$ such that

$$ae^{i\xi x} \in W(f)$$

for $a \neq 0$ with sufficiently small $|a|$.

Remark 9.3

1° $\sigma(f)$ may be empty. For instance, $W(f) \subset \mathfrak{C}_{\infty}(\mathbb{R}, \mathbb{C})$ for any $f \in \mathfrak{C}_{\infty}(\mathbb{R}, \mathbb{C})$. Hence $ae^{i\xi x} \notin W(f)$ for any $a \neq 0$. This implies $\sigma(f) = \emptyset$.

2° For $f \in \mathfrak{L}^{\infty}(\mathbb{R}, \mathbb{C})$,

$$\sigma(e^{i\xi x} f) = \xi + \sigma(f).$$

It is verified as follows:

$$\begin{aligned} \theta \in \sigma(e^{i\xi x} f) \quad &\Leftrightarrow \quad \text{For } a \text{ with sufficiently small } |a|, \\ &\qquad\qquad ae^{i\theta x} \in W(e^{i\xi x} f) \\ &\Leftrightarrow \quad \text{For some } g \in W(f), \\ &\qquad\qquad ae^{i\theta x} = e^{i\xi x} g \quad \text{(Remark 9.1, 3°)} \end{aligned}$$

$$\begin{aligned} \Leftrightarrow \quad & \text{For some } g \in W(f), \\ & \qquad ae^{i(\theta-\xi)x} = g \\ \Leftrightarrow \quad & \theta - \xi \in \sigma(f). \end{aligned}$$

The Fourier transform $\hat{f}$ of $f \in \mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$ which appears in the following discussions should be understood as the one in the sense of distributions. (cf. Appendix C.)

Lemma 9.4 *For any* $f \in \mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$,

$$\sigma(f) \subset \text{supp } \hat{f}.$$

*(*supp $\hat{f}$ *is the support of the distribution* $\hat{f}$.*)*[8]

Proof Since $\widehat{f_y} = e^{-i\xi y}\hat{f}$,[9] we obtain

$$\text{supp } \widehat{f_y} = \text{supp } \hat{f}.$$

Hence

$$\text{supp } \hat{g} \subset \text{supp } \hat{f}$$

for every $g \in W(f)$.[10]

---

[8]cf. Appendix C, p. 397.

[9]In fact, we have

$$\begin{aligned} \widehat{f_y}(\varphi) = f_y(\hat{\varphi}) &= \frac{1}{\sqrt{2\pi}} \int \int \varphi(\xi) e^{-i\xi x} d\xi f(x-y) dx \\ &= \frac{1}{\sqrt{2\pi}} \int \int \varphi(\xi) e^{-i\xi(y+z)} d\xi f(z) dz \\ &= f(\widehat{\varphi \cdot e^{-i\xi y}}) = \hat{f}(\varphi \cdot e^{-i\xi y}) = e^{-i\xi y} \hat{f}(\varphi) \end{aligned}$$

for any $\varphi \in \mathfrak{S}$ by the definition of the Fourier transform of distributions. Hence $\widehat{f_y} = e^{-i\xi y}\hat{f}$. In the course of computations, $f(\cdot)$ and $f_y(\cdot)$ are tempered distributions defined by the functions f and f_y. cf. Appendix C, Sect. C.3, p. 389.

[10]We shall give a brief proof. Since $g \in W(f)$, there exists a sequence $\{g_n\}$ of functions of the form

$$\sum a_k f_{x_k}, \;\; x_k \in \mathbb{R}, \;\; \sum |a_k| \leqq 1$$

such that $\|g_n - g\|_\infty \to 0$ (as $n \to \infty$), by Remark 9.1, 1°. Hence it is easy to see that the sequence of tempered distributions defined by g_n's simply converges to the tempered distribution defined by g. Consequently, $\hat{g_n}$ also simply converges to $\hat{g}$ in $\mathfrak{S}'$ (cf. 2° on p. 86). If the support of a function $\varphi \in \mathfrak{S}$ is contained outside supp$\hat{f}$, $\hat{g}_n(\varphi) = 0$ since supp$\hat{f}_{x_k}$ = supp$\hat{f}$. Therefore $\hat{g}(\varphi) = \lim_{n\to\infty} \hat{g}_n(\varphi) = 0$. We conclude supp$\hat{g} \subset$ supp$\hat{f}$.

Assume that $\sigma(f) \neq \emptyset$ and $\xi \in \sigma(f)$. If $g(x) = ae^{i\xi x} \in W(f)$,

$$\hat{g} = \sqrt{2\pi} a \delta_\xi.$$[11]

Hence we have

$$\operatorname{supp} \hat{g} = \{\xi\}.$$

Consequently,

$$\xi \in \operatorname{supp} \hat{f} \quad \text{for} \quad \xi \in \sigma(f).$$

This implies $\sigma(f) \subset \operatorname{supp} \hat{f}$. $\square$

Definition 9.4 Let $F \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$.

$$\Omega(F, \delta) = \sup_{|y|<\delta} \|F(x+y) - F(x)\|_1$$

is called the $\mathfrak{L}^1$-**modulus of continuity** of F.

We list below several elementary properties of $\Omega(F, \delta)$, which are derived from the definition:

1° $\Omega(\eta F(\eta x), \delta) = \Omega(F, \eta\delta)$. In fact,

$$\text{left-hand side} = \sup_{|y|<\delta} \|\eta F(\eta(x+y)) - \eta F(\eta x)\|_1 = \sup_{|y|<\delta} \|F(z+\eta y) - F(z)\|_1$$

(changing variables: $z = \eta x$)

$$= \sup_{|w|<\delta\eta} \|F(z+w) - F(z)\|_1 = \text{right-hand side}.$$

2° Given $\delta > 0$,[12]

$$\lim_{\eta \to 0} \Omega(\eta F(\eta x), \delta) = 0.$$

[11]The Fourier transform of $e^{i\xi x}$ can be evaluated as in Example 4.9 on p. 85. For any $\varphi \in \mathfrak{S}$, we have

$$\widehat{e^{i\xi x}}(\varphi) = \int_{-\infty}^{\infty} e^{i\xi x} \hat{\varphi}(x) dx = \sqrt{2\pi} \varphi(\xi).$$

Hence

$$\widehat{e^{i\xi x}} = \sqrt{2\pi} \delta_\xi$$

is derived.

[12]Since $F \in \mathfrak{L}^1$, the mapping $y \mapsto \tau_y F$ $(\mathbb{R} \to \mathfrak{L}^1)$ is uniformly continuous by Theorem 5.1 (p. 101). It is, of course, continuous at $y = 0$. Hence there exists some $\theta > 0$, for each $\varepsilon > 0$, such that $\|\tau_y F - F\|_1 < \varepsilon$ provided that $|y| < \theta$. Combining this result and 1°, we obtain 2°.

3° If we define g_η by

$$g_\eta = \eta F(\eta x) * h$$

for $h \in \mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$, then

$$|g_\eta(x) - g_\eta(y)| \leqq \Omega(F, \eta|x-y|) \cdot \|h\|_\infty.$$

4° g_η is defined in the same manner as in 3°. If $\lim_{\eta\to 0} g_\eta(x)$ exists at each point x, then

$$\lim_{\eta\to 0} g_\eta(x) = c(\text{constant}) \quad \text{for all} \quad x \in \mathbb{R}.$$

If $x \neq y$, we obtain

$$|g_\eta(x) - g_\eta(y)| \leqq \Omega(\eta F(\eta x), |x-y|) \cdot \|h\|_\infty$$

by 1° and 2°. The right-hand side converges to 0 as $\eta \to 0$, according to 2°. Hence 4° follows.

5° Assume that $\|F\|_1 \leqq 1$ and $\int_{\mathbb{R}} F(x)dx \neq 0$. If $f \in \mathfrak{L}^\infty(\mathbb{R}, \mathbb{C})$ is uniformly continuous,

$$g_\eta = \eta F(\eta x) * f \in W(f)$$

by Remark 9.1, 2°. If the uniform limit of g_η exists, the limit is contained in $W(f)$. In the case $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, a limiting point (with respect to the uniform convergence) exists, since $W(f)$ is compact (Lemma 9.3).

Lemma 9.5 *Assume that* $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, $F \in \mathfrak{L}^1(\mathbb{R}, \mathbb{R})$, $F(x) \geqq 0$ *and* $\int_{\mathbb{R}} F(x)dx = 1$. *Then the limit*

$$\lim_{\eta\to 0} \eta F(\eta x) * f \quad (\textit{uniform convergence})$$

exists and

$$\lim_{\eta\to 0} \eta F(\eta x) * f = 0 \quad \Leftrightarrow \quad 0 \notin \sigma(f).$$

Proof Since f is uniformly continuous, $g_\eta = \eta F(\eta x) * f \in W(f)$ by 5° just stated. $W(f)$ being compact, there exists some sequence $\eta_n \to 0$ such that $\{g_{\eta_n}\}$ has a limit and the limit is constant by 4° above.

1° Suppose that

$$g_{\eta_n}(x) \to 0 \quad (\text{uniform convergence}). \tag{9.25}$$

Then $0 \in W(f)$. It is also easy to see that

$$\lim_{n\to\infty} \eta_n F(\eta_n x) * h = 0 \quad \text{(uniform convergence)} \tag{9.26}$$

for any $h \in W(f)$.[13]

The proof proceeds as follows. Computing the left-hand side for $h(x) = f_\xi(x) = f(x-\xi) \in W_0(f)$, we obtain

$$\begin{aligned}
\eta_n F(\eta_n x) * f_\xi &= \int_{\mathbb{R}} \eta_n F(\eta_n(x-y)) f_\xi(y)dy = \int_{\mathbb{R}} \eta_n F(\eta_n(x-y)) f(y-\xi)dy \\
&= \int_{\mathbb{R}} \eta_n F(\eta_n(x-z-\xi)) f(z)dz \quad \text{(changing variables: } y-\xi = z\text{)} \\
&= \int_{\mathbb{R}} \eta_n F(\eta_n(x-\xi-z)) f(z)dz \to 0 \qquad \text{as} \quad n\to\infty.
\end{aligned}$$

Although this relation is valid for fixed $x-\xi$, we also obtain, by (9.25), that

$$\lim_{n\to\infty} \eta_n F(\eta_n x) * f_\xi = 0 \quad \text{(uniform convergence).} \tag{9.27}$$

Therefore the same relation as (9.27) holds good even if we replace f_ξ by a function of the form

$$\sum_k a_k f_{\xi_k}, \quad \xi_k \in \mathbb{R}, \quad \sum_k |a_k| \leqq 1 \tag{9.28}$$

($\sum_k$ is a finite sum). $h \in W(f)$ is a uniform limit of a sequence of functions of the form (9.28). Hence we conclude that (9.26) is true.

In particular, if h is a constant c contained in $W(f)$,[14] we obtain

$$\eta F(\eta x) * c = c, \tag{9.29}$$

from which it follows that $c = 0$ by (9.26). Hence if (9.25) is satisfied, the only constant contained in $W(f)$ is 0. By the definition of $\sigma(f)$, we have $0 \notin \sigma(f)$.

---

[13] Obviously (9.26) holds good if we use f instead of h.

[14] It can be verified by

$$\eta F(\eta x) * c = \int_{\mathbb{R}} \eta F(\eta(x-y))cdy = c\int_{\mathbb{R}} \eta F(\eta(x-y))dy = c\int_{\mathbb{R}} F(u)du = c$$

(changing variables: $\eta(x-y) = u$).

Suppose that there exists a limiting point of $\{g_\eta(x)\}$ other than 0; i.e.

$$\lim_{n\to\infty} \theta_n F(\theta_n x) * f = \alpha \neq 0 \quad \text{(uniform convergence)} \tag{9.30}$$

for some sequence $\theta_n \to 0$. If this is the case, $\alpha \in W(f)$.[15] But this is impossible because the only constant contained in $W(f)$ is 0.

Thus, if (9.25) is satisfied,

$$\lim_{\eta\to 0} \eta F(\eta x) * f = 0 \quad \text{(uniform convergence)}$$

and $0 \notin \sigma(f)$. This is the first case.

2° Suppose next that a limiting point of g_η is not 0; that is, (9.29) holds good for some sequence $\theta_n \to 0$. In this case, we obtain

$$\lim_{n\to\infty} \theta_n F(\theta_n x) * (f - \alpha) = 0.$$

Hence it follows that

$$\lim_{\eta\to 0} \eta F(\eta x) * f = \alpha \quad \text{(uniform convergence)}.$$

Here we have the alternative case ; i.e. g_η uniformly converges to some constant other than 0, and $0 \in \sigma(f)$. □

As is illuminated by the computation in (9.29),

$$\lim_{\eta\to 0} \eta F(\eta x) * f$$

is determined independently of F provided that $F \in \mathfrak{L}^1(\mathbb{R}, \mathbb{R}), F \geqq 0, \int_{\mathbb{R}} F(x)dx = 1$.

Definition 9.5 Let $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$. The number $M(f)$ which satisfies

$$0 \notin \sigma(f - M(f))$$

is called the **mean value** of f.

The concept of a norm spectrum can be characterized in terms of the mean value.

[15] By definition of $W(f)$, we have $0 \in \sigma(f)$.

Theorem 9.3 *For* $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$,

$$M(f) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(y)dy.$$

Proof It holds good that

$$M(f) = \lim_{\eta\to 0} \eta F(\eta x) * f$$

by Lemma 9.5. As is remarked above, the value of $M(f)$ is determined independently of the choice of F provided that $F \in \mathfrak{L}^1(\mathbb{R}, \mathbb{C})$, $F \geqq 0$, $\int_{\mathbb{R}} F(x)dx = 1$.

Let us specify F as

$$F(x) = \begin{cases} \frac{1}{2} & \text{if} \quad |x| < 1, \\ 0 & \text{if} \quad |x| \geqq 1 \end{cases}$$

and T as $T = 1/\eta$. Then

$$\begin{aligned} \eta F(\eta x) * f &= \int_{\mathbb{R}} \eta F(\eta(x-y)) f(y)dy \\ &= \int_{x-\frac{1}{\eta}}^{x+\frac{1}{\eta}} \eta \underbrace{F(\eta(x-y))}_{=1/2} f(y)dy \qquad \text{(cf. Fig. 9.2)} \\ &= \int_{x-T}^{x+T} \frac{1}{T} F\left(\frac{1}{T}(x-y)\right) f(y)dy \\ &= \int_{-T}^{T} \frac{1}{2T} f(y)dy + \int_{T}^{x+T} \frac{1}{T} F\left(\frac{1}{T}(x-y)\right) f(y)dy \\ &= \frac{1}{2T} \int_{-T}^{T} f(y)dy + \frac{1}{2T} \int_{T}^{x+T} f(y)dy \\ &= I_1 + I_2. \end{aligned} \tag{9.31}$$

$$|I_2| \leqq \frac{1}{2T} \int_{T}^{x+T} |f(y)|dy \leqq \frac{|x|}{2T} \|f\|_\infty \to 0 \quad \text{as} \quad T \to \infty. \tag{9.32}$$

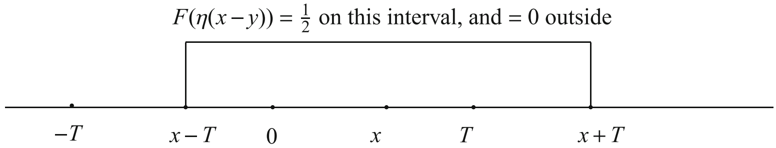


Fig. 9.2 $F(\eta(x-y))$

By (9.31) and (9.32), we obtain

$$M(f) = \lim_{T\to\infty} I_1 = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(y)dy.$$

□

Theorem 9.4 *The following three statements are equivalent for $f \in \mathfrak{AP}(\mathbb{R},\mathbb{C})$:*

(i) $\xi \in \sigma(f)$.
(ii) $0 \in \sigma(fe^{-i\xi x})$.
(iii) $M(fe^{-i\xi x}) \neq 0$.

This result is clear from Remark 9.3, 2° and Lemma 9.5.

Corollary 9.2[16]

(i) *If we define*

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{i\xi x} d\mu(\xi)$$

for $\mu \in \mathfrak{M}(\mathbb{R})$, then μ is the Fourier transform of f. (Hence any $\mu \in \mathfrak{M}(\mathbb{R})$ is the Fourier transform of some function in $\mathfrak{L}^\infty$.)

(ii) *Let $f \in \mathfrak{AP}(\mathbb{R},\mathbb{C})$. If $\mu \in \mathfrak{M}(\mathbb{R})$ satisfies $\mu = \hat{f}$ and $\mu\{0\} \neq 0$, then $0 \in \sigma(f)$.*[17]

Proof (i) Let $\mu \in \mathfrak{M}(\mathbb{R})$. If we define $f \in \mathfrak{L}^\infty$ by

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{i\xi x} d\mu(\xi),$$

[16] Katznelson [8] 2nd edn., pp. 159–160. There is a difference in the explanation of this result between the second edition and the third one. The latter seems to contain a slip.
[17] $\hat{f}$ is the Fourier transform of f in the sense of distribution.

then we obtain[18]

$$
\begin{aligned}
\hat{f}(\varphi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{\varphi}(x) e^{i\xi x} d\mu(\xi) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left\{ \int_{\mathbb{R}} \hat{\varphi}(x) e^{i\xi x} dx \right\} d\mu(\xi) \\
&= \int_{\mathbb{R}} \varphi(\xi) d\mu(\xi) \quad \text{for any } \varphi \in \mathfrak{S} \quad \text{(inverse transform).}
\end{aligned}
$$

Thus $\hat{f} = \mu$ follows.

(ii) We specify F as Fejér's summation kernel[19] $K_\eta(x) = \eta K(\eta x)$, and define

$$
g_\eta = \eta K(\eta x) * f. \tag{9.33}
$$

Then $\hat{g}_\eta$ is evaluated as[20]

$$
\hat{g}_\eta = \sqrt{2\pi} \hat{K}\left(\frac{\xi}{\eta}\right) \cdot \hat{f} = \sqrt{2\pi} \hat{K}\left(\frac{\xi}{\eta}\right) \cdot \mu. \tag{9.34}
$$

Regarding $\hat{g}_\eta$ as an element of $\mathfrak{S}'$, we obtain

$$
w^*\text{-} \lim_{\eta \to 0} \hat{g}_\eta = \sqrt{2\pi} \mu\{0\} \delta_0. \tag{9.35}
$$

Hence g_η w^*-converges to the inverse Fourier transform of $\sqrt{2\pi}\mu\{0\}\delta_0$, i.e. $\mu\{0\}$ thanks to 2° on p. 86. That is,

$$
w^*\text{-} \lim_{\eta \to 0} g_\eta = \mu\{0\} \neq 0. \tag{9.36}
$$

g_η has a uniform limit by Lemma 9.5. The limit is nothing other than (9.36). Hence, again by Lemma 9.5, we have $0 \in \sigma(f)$. □

Thus if $\hat{f} \in \mathfrak{M}(\mathbb{R})$, $M(f)$ can be characterized as

$$
\mu\{0\} = \hat{f}(\{0\}) = M(f).
$$

[18] $\hat{f}$ is the Fourier transform of the tempered distribution defined by f.

[19] See Chap. 5, Sect. 5.4. $K(x) = \dfrac{1}{2\pi}\left(\sin\dfrac{x}{2} \Big/ \dfrac{x}{2}\right)^2$.

[20] $\hat{f}$, $\hat{g}_\eta$ are the Fourier transforms of distributions.

Similarly, we have

$$\hat{f}(\{\xi\}) = M(fe^{-i\xi x}). \tag{9.37}$$

The relation (9.37) is the representation of the discrete part of the measure $\hat{f}$. It will be shown later that the continuous part of $\hat{f}$ actually does not exist (cf. Corollary 9.3).

It is easy to show the following formulas concerning the average $M(\cdot)$ on $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$:

1° $M(f + g) = M(f) + M(g)$.
2° $M(af) = aM(f), \quad a \in \mathbb{C}$.
3° $M(f_y) = M(f)$.

Lemma 9.6 *Let $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$. Then $M(f) > 0$ if $f \geqq 0$ and f is not identically* 0.

Proof We may assume $f(0) > 0$ by 3°. Hence

$$f(x) > \alpha \quad \text{on} \quad (-\alpha, \alpha)$$

for sufficiently small $\alpha > 0$. Let $\Lambda = \Lambda(\alpha/2, f)$. Then any interval of the length Λ contains an $\alpha/2$-almost period τ of f. Without loss of generality, we may assume $\Lambda \geqq 2\alpha$. Since

$$f(x) > \frac{\alpha}{2} \quad \text{on} \quad (\tau - \alpha, \tau + \alpha),$$

we have

$$\int_I f(x)dx \geqq \frac{\alpha}{2} \cdot 2\alpha = \alpha^2$$

for an interval I of the length Λ. Consequently, we obtain, by Theorem 9.3,

$$M(f) = \lim_{n\to\infty} \frac{1}{2n\Lambda} \int_{-n\Lambda}^{n\Lambda} f(x)dx \geqq \frac{\alpha^2}{\Lambda}.$$

□

Corollary 9.3 *If $\mu \in \mathfrak{M}(\mathbb{R})$ and $\hat{\mu} \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, then μ is discrete.*

Proof μ can be expressed in the form $\mu = \mu_d + \mu_c$, where μ_d is the discrete part and μ_c is the continuous part. If $\hat{\mu} \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, then $\hat{\mu_c} \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, since $\hat{\mu}_d \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$ (cf. Remark 9.2). So we can conclude $|\hat{\mu}_c|^2 \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$.

By Wiener's theorem 6.15 (p. 160), we obtain

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |\hat{\mu}_c(x)|^2 dx = 0.$$

Therefore $\mu_c = 0$ by Lemma 9.6. □

9.4 Fourier Series of Almost Periodic Functions

Let $\langle \cdot, \cdot \rangle_M$ be an operation on $\mathfrak{AP}(\mathbb{R}, \mathbb{C}) \times \mathfrak{AP}(\mathbb{R}, \mathbb{C})$ defined by

$$\langle f, g \rangle_M = M(f\bar{g}), \quad f, g \in \mathfrak{AP}(\mathbb{R}, \mathbb{C}). \tag{9.38}$$

$\langle \cdot, \cdot \rangle_M$ is an inner product on $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$, and so $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$ becomes a pre-Hilbert space with respect to this operation.

The family $\{e^{i\xi x} | \xi \in \mathbb{R}\}$ forms an orthonormal family with respect to $\langle \cdot, \cdot \rangle_M$. In fact, it can immediately be verified by

$$\langle e^{i\xi x}, e^{i\eta x} \rangle_M = \lim_{T \to 0} \frac{1}{2T} \int_{-T}^{T} e^{i(\xi - \eta)x} dx = \begin{cases} 1 & \text{if} \quad \xi = \eta, \\ 0 & \text{if} \quad \xi \neq \eta. \end{cases} \tag{9.39}$$

For $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, the numbers defined by

$$\langle f, e^{i\xi x} \rangle_M = M(fe^{-i\xi x}) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f(x) e^{-i\xi x} dx \tag{9.40}$$

are called the **Fourier coefficients** of f with respect to $\{e^{i\xi x} | \xi \in \mathbb{R}\}$, and are denoted by $\hat{f}\{\xi\}$.[21] For any finite numbers $\xi_1, \xi_2, \cdots, \xi_n \in \mathbb{R}$, we obtain

$$\begin{aligned} &\langle f - \sum_{j=1}^{n} \hat{f}\{\xi_j\} e^{i\xi_j x},\ f - \sum_{j=1}^{n} \hat{f}\{\xi_j\} e^{i\xi_j x} \rangle_M \\ &\qquad = \langle f,\ f \rangle_M - \sum_{j=1}^{n} \overline{\hat{f}\{\xi_j\}} \langle f,\ e^{i\xi_j x} \rangle_M \\ &\qquad\qquad - \sum_{j=1}^{n} \hat{f}\{\xi_j\} \langle e^{i\xi_j x},\ f \rangle_M + \sum_{j=1}^{n} |\hat{f}\{\xi_j\}|^2 \\ &\qquad = M(|f|^2) - \sum_{j=1}^{n} |\hat{f}\{\xi_j\}|^2 \geqq 0. \end{aligned}$$

Hence it follows that

[21] We use such a notation because we wish to interpret the Fourier transform of f as something like a measure and its value at ξ as the mass at ξ.

$$\sum_{j=1}^{n} |\hat{f}\{\xi_j\}|^2 \leqq \langle f, \ f\rangle_M = M(|f|^2), \tag{9.41}$$

which is obviously a familiar inequality corresponding to Bessel's inequality. It implies that there are at most countable ξ's such that $\hat{f}\{\xi\} \neq 0$. Some exposition may be required to clarify this point.

Assume that there are infinitely many ξ's such that $|\hat{f}\{\xi\}| \geqq 1$. Choose any countable elements $\xi_1, \xi_2, \cdots$ among them. Then the sum on the left-hand side of (9.41) is greater than any $n \in \mathbb{N}$ and exceeds $M(|f|^2)$ if n is very large. This contradicts (9.41). Furthermore, there are only finite ξ's such that $1/(n+1) \leqq |\hat{f}\{\xi\}| < 1/n$ $(n = 1, 2, \cdots)$. Suppose that there are infinitely many ξ's (say, $\xi_1, \xi_2, \cdots$) which satisfy this inequality for some n. If we sum up $(n+1)^2$ terms $|\hat{f}(\xi_j)|^2$, the sum is greater than 1. Hence the left-hand side of (9.41) grows indefinitely. A contradiction to (9.41) occurs again. Since

$$\xi \in \sigma(f) \Leftrightarrow M(fe^{-i\xi x}) = \hat{f}\{\xi\} \neq 0,$$

$\sigma(f)$ is a countable set.

Each $\xi_n (n \in \mathbb{Z})$ contained in $\sigma(f)$ is called a **Fourier index** of f, and the formal series

$$\sum_{n=-\infty}^{\infty} c_n e^{i\xi_n x}$$

is called the **Fourier series** of an almost periodic function f, where $c_n = M(fe^{-i\xi_n x})$ and $\xi_n \in \sigma(f)$.

Theorem 9.5 *For $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, $\sigma(f)$ is a countable set.*

For a pair of f and $g \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, we define the **mean convolution** $(f \underset{M}{*} g)(x)$ by

$$(f \underset{M}{*} g)(x) = M_y(f(x-y)g(y)) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(x-y)g(y)dy. \tag{9.42}$$

$M_y(f(x-y)g(y))$ is well-defined because $f(x-y)g(y)$ is an almost periodic function in y for each fixed x.

Lemma 9.7 *$f \underset{M}{*} g \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$ for $f, g \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$. If $M(|g|) \leqq 1$, then $f \underset{M}{*} g \in W(f)$.*

Proof Let $\Lambda = \Lambda(\varepsilon/\|g\|_\infty, f)$. Then any interval of length Λ contains an $\varepsilon/\|g\|_\infty$-almost period τ. It is not difficult to show that

$$\sup_x |(f \underset{M}{*} g)(x - \tau) - (f \underset{M}{*} g)(x)| \leqq \varepsilon. \tag{9.43}$$

In fact, (9.43) follows from

$$\begin{aligned}
|(f \underset{M}{*} g)(x - \tau) - (f \underset{M}{*} g)(x)| &= \left| \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(x - \tau - y)g(y)dy \right.\\
&\qquad \left. - \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(x - y)g(y)dy \right| \\
&\leqq \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |f(x-\tau - y) - f(x-y)||g(y)|dy \\
&\leqq \lim_{T\to\infty} \frac{1}{2T} \cdot 2T \cdot \frac{\varepsilon}{\|g\|_\infty} \cdot \|g\|_\infty \leqq \varepsilon.
\end{aligned}$$

That is, any interval of length Λ contains an ε-almost period of $f \underset{M}{*} g$. Hence $f \underset{M}{*} g$ is almost periodic.

Assume next that $M(|g|) < 1$; i.e.

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |g(y)|dy < 1.$$

Then

$$\frac{1}{2T} \int_{-T}^{T} |g(y)|dy < 1$$

for sufficiently large T. If we define a function φ by

$$\varphi(y) = \begin{cases} \frac{1}{2T} g(y) & \text{on} \quad [-T, T], \\ 0 & \text{on} \quad [-T, T]^c, \end{cases}$$

then $\|\varphi\|_1 < 1$ and $\varphi * f \in W(f)$ (cf. p. 251, 2°) That is,

$$\frac{1}{2T} \int_{-T}^{T} f(x - y)g(y)dy \in W(f) \tag{9.44}$$

for sufficiently large T. The left-hand side of (9.44) is convergent for each x by (9.42). $f \underset{M}{*} g$ is the uniform limit of the left-hand side of (9.44) which is in $W(f)$. Since $W(f)$ is $\|\cdot\|_\infty$-compact, $f \underset{M}{*} g \in W(f)$.

Assume that $M(|g|) = 1$. For any $\varepsilon > 0$,

$$\frac{1}{2T}\int_{-T}^{T}|g(y)|dy < 1+\varepsilon, \quad \text{i.e.} \quad \frac{1}{2T}\int_{-T}^{T}\frac{|g(y)|}{1+\varepsilon}dy < 1$$

holds good for sufficiently large $T > 0$. If we define a function φ_ε by

$$\varphi_\varepsilon(y) = \begin{cases} \frac{1}{2T(1+\varepsilon)}g(y) & \text{on} \quad [-T, T], \\ 0 & \text{on} \quad [-T, T]^c, \end{cases}$$

then $\|\varphi_\varepsilon\|_1 < 1$. Applying the argument above, we obtain $f \underset{M}{*} g/(1+\varepsilon) \in W(f)$. Since $f \underset{M}{*} g/(1+\varepsilon)$ uniformly converges to $f \underset{M}{*} g$ as $\varepsilon \downarrow 0$, we obtain $f \underset{M}{*} g \in W(f)$. □

The operation of mean convolution enjoys all the basic properties of ordinary convolution. We now list a few of them.

1° $(\widehat{f \underset{M}{*} g})\{\xi\} = \hat{f}\{\xi\} \cdot \hat{g}\{\xi\}$ for $f, g \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$.

This relation follows from a direct computation:

$$\begin{aligned} (\widehat{f \underset{M}{*} g})\{\xi\} &= M_x(M_y(f(x-y)g(y))e^{-i\xi x}) \\ &= M_x M_y(f(x-y)e^{-i\xi(x-y)}g(y)e^{-i\xi y}) \\ &= \hat{f}\{\xi\}\hat{g}\{\xi\}. \end{aligned}$$

2° $f \underset{M}{*} e^{i\xi x} = \hat{f}\{\xi\}e^{i\xi x}$ for $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$.

3° If we define a function $g(x)$ by

$$g(x) = \sum_j \hat{g}\{\xi_j\}e^{i\xi_j x} \qquad \text{(finite sum)},$$

then $f \underset{M}{*} g (f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C}))$ is given by

$$f \underset{M}{*} g = \sum_j \hat{g}\{\xi_j\}\hat{f}\{\xi_j\}e^{i\xi_j x}.$$

Let $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$. If we define a function f^* by

$$f^*(x) = \overline{f(-x)}, \tag{9.45}$$

it is easy to see

$$\hat{f^*}\{\xi\} = \overline{\hat{f}\{\xi\}}. \tag{9.46}$$

Hence we obtain

$$(\widehat{f \underset{M}{*} f^*})\{\xi\} = |\hat{f}\{\xi\}|^2. \tag{9.47}$$

In the case of $\|f\|_\infty \leqq 1$, we obtain $f \underset{M}{*} f^* \in W(f)$.

Lemma 9.8 *Suppose that* $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$.

(i) $h = f \underset{M}{*} f^*$ *is positive semi-definite.*
(ii) h *can be represented as the Fourier transform of a positive Radon measure on* $\mathbb{R}$.
(iii) $\hat{h}$ *is a positive Radon measure on* $\mathbb{R}$.

Proof (i) For $x_j \in \mathbb{R}, z_j \in \mathbb{C}$ $(j = 1, 2, \cdots, p)$, we obtain

$$\begin{aligned}
\sum_{i,j=1}^{p} h(x_i - x_j) z_i \bar{z}_j &= \sum_{i,j} \left[\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(x_i - x_j - y) f^*(y) dy \right] z_i \bar{z}_j \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \sum_{i,j} f(x_i - x_j - y) \overline{f(-y)} dy z_i \bar{z}_j \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-x_j-T}^{-x_j+T} \sum_{i,j} f(x_i + u) \overline{f(x_j + u)} z_i \bar{z}_j du \\
&\qquad \text{(changing variables: } u = -x_j - y) \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-x_j-T}^{-x_j+T} \left| \sum_j f(x_j + u) \cdot z_j \right|^2 du \\
&\geqq 0.
\end{aligned}$$

(ii) Clear from Bochner's theorem (Theorem 6.11, p. 150).
(iii) If $h = \hat{\mu}$ for some $\mu \in \mathfrak{M}_+(\mathbb{R})$, we have

$$\begin{aligned}
\hat{h}(\varphi) = h(\hat{\varphi}) = \hat{\mu}(\hat{\varphi}) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{\varphi}(\xi) \int_{\mathbb{R}} e^{-i\xi x} d\mu(x) d\xi \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \hat{\varphi}(\xi) e^{i(-x)\xi} d\xi \right] d\mu(x) = \int_{\mathbb{R}} \varphi(-x) d\mu(x) \\
&= \int_{\mathbb{R}} \varphi(x) d\mu(x) = \mu(\varphi) \quad \text{for any } \varphi \in \mathfrak{S}.
\end{aligned}$$

It follows that $\hat{h} = \mu$, where $h(\cdot)$, $\mu(\cdot)$ are distributions defined by h and μ, the Fourier transforms of which are $\hat{h}$, $\hat{\mu}$, respectively. □

It is obvious that

$$h = \hat{\mu} \Leftrightarrow \hat{h} = \mu \tag{9.48}$$

by (iii).

Since the Radon measure $\hat{h}$ is discrete by Corollary 9.3, it can be expressed as

$$\hat{h} = \sum \hat{h}\{\xi\}\delta_\xi . \tag{9.49}$$

Consequently,

$$h(x) = \sum \hat{h}\{\xi\}e^{i\xi x} \underset{(9.47)}{=} \sum |\hat{f}\{\xi\}|^2 e^{i\xi x}. \tag{9.50}$$

If $x = 0$, in particular, we have

$$M(|f|^2) = h(0) = \sum |\hat{f}\{\xi\}|^2. \tag{9.51}$$

This is **Parseval's equality for almost periodic functions**.

The Fourier indices are at most countable in any one of (9.49), (9.50) and (9.51). Hence the sums $\sum$ appearing in them are countable sums.

The orthonormal system $\{e^{i\xi x} | \xi \in \mathbb{R}\}$ in $\mathfrak{AP}(\mathbb{R}, \mathbb{C})$ is complete by (9.51) and Theorem 1.6 (p. 16). Thus

$$f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C}), \ \ f \neq 0 \Rightarrow \sigma(f) \neq \emptyset.$$

Lemma 9.9 *For any finite real numbers $\xi_1, \xi_2, \cdots, \xi_n$ and $\varepsilon > 0$, there exists a trigonometric polynomial P such that:*

(i) $P(0) \geqq 0$,
(ii) $M(P) = 1$, *and*
(iii) $\hat{P}\{\xi_j\} > 1 - \varepsilon$ *for* $j = 1, 2, \cdots, n$.

Proof Assume first that $\xi_1, \xi_2, \cdots, \xi_n$ are all integers. Choose $m \in \mathbb{Z}$ so that

$$\frac{1}{\varepsilon} \operatorname*{Max}_{1 \leqq j \leqq n} |\xi_j| < m.$$

If we adopt Fejér's kernel

$$K_m(x) = \sum_{k=-m}^{m} \left(1 - \frac{|k|}{m+1}\right) e^{ikx},$$

as $P(x)$, then $P(x) = K_m(x)$ satisfies (i)–(iii).

We proceed to the general case. Note that there exist real numbers $\lambda_1, \lambda_2, \cdots, \lambda_q$ for $\xi_1, \xi_2, \cdots, \xi_n$ such that:

(a) $\sum_{h=1}^{q} \theta_h \lambda_h = 0, \theta_h \in \mathbb{Q} \Rightarrow \theta_h = 0$ for all h, and
(b) there exist $A_{j,h} \in \mathbb{Z}$ which satisfy

$$\xi_j = \sum_{h=1}^{q} A_{j,h} \lambda_h, \ \ j = 1, 2, \cdots, n$$

for each $j = 1, 2, \cdots, n$.

We can verify this fact by induction in n.[22]

[22]In the case $n = 1$, $P(x) = e^{i\xi_1 x}$ works. Next assume that there exist $\lambda_1, \lambda_2, \cdots, \lambda_q$ which satisfy (a) and (b) for $\xi_1, \xi_2, \cdots, \xi_{n-1}$. Consider $\xi_1, \xi_2, \cdots, \xi_n$. We may assume that

$$\xi_n \notin \sum_{h=1}^{q} \lambda_h \mathbb{Z}$$

for any $\lambda_1, \lambda_2, \cdots, \lambda_q$ which satisfy (a) and (b). (Otherwise ξ_n is also expressible in the form (b) by using $\lambda_1, \lambda_2, \cdots \lambda_q$.) Define λ by

$$\lambda = \xi_n - \sum_{h=1}^{q} \lambda_h z_h \quad \text{for} \quad z_h \in \mathbb{Z}.$$

Assume that

$$\sum_{h=1}^{q} \theta_h \lambda_h + \theta \lambda = 0, \quad \theta_h, \theta \in \mathbb{Q}.$$

If $\theta = 0$, then $\theta_h = 0$ $(h = 1, 2, \cdots, q)$ by (a). Hence we may concentrate at the case $\theta \neq 0$. Since

$$\sum_{h=1}^{q} \theta_h \lambda_h + \theta \lambda = \sum_{h=1}^{q} \theta_h \lambda_h + \theta \left(\xi_n - \sum_{h=1}^{q} \lambda_h z_h \right) = \sum_{h=1}^{q} (\theta_h - \theta z_h) \lambda_h + \theta \xi_n = 0,$$

it is obvious that

$$\xi_n = -\sum_{h=1}^{q} \frac{\theta_h - \theta z_h}{\theta} \lambda_h.$$

Expressing $\theta_h / \theta = v_h / u_h$ $(u_h, v_h \in \mathbb{Z}, h = 1, 2, \cdots, q)$, we denote by u^* the least common multiple of u_h's. If we write $\theta_h / \theta = v_h^* / u^*$, we obtain

$$\xi_n = -\sum_{h=1}^{q} \left(\frac{v_h^*}{u^*} - z_h \right) \lambda_h = -\sum_{h=1}^{q} (v_h^* - u^* z_h) \frac{\lambda_h}{u^*}.$$

Furthermore, if we write $\lambda_h^* = \lambda_h / u^*$ $(h = 1, 2, \cdots, q)$, ξ_j's can be reexpressed as

$$\xi_j = \sum_{h=1}^{q} u^* A_{j,h} \lambda_h^*, \ \ j = 1, 2, \cdots, n-1,$$

Choose $\varepsilon_1 > 0$ and $m \in \mathbb{N}$ so that

$$(1-\varepsilon_1)^q > 1-\varepsilon, \quad \varepsilon_1 > \frac{\operatorname*{Max}_{j,k} |A_{j,k}|}{m}.$$

Define a function $P(x)$ by

$$P(x) = \prod_{h=1}^{q} K_m(\lambda_h x). \tag{9.52}$$

Then it is easy to see that $P(x)$ satisfies (i).

In order to prove (ii), we first rewrite $P(x)$ as

$$P(x) = \sum \left(1 - \frac{|k_1|}{m+1}\right) \cdots \left(1 - \frac{|k_q|}{m+1}\right) e^{i(k_1\lambda_1 + \cdots + k_q\lambda_q)x}, \tag{9.53}$$

where the summation $\sum$ on the right-hand side of (9.53) runs over all $(k_1, k_2, \cdots, k_q)$ such that $|k_1| \leqq m, \cdots, |k_q| \leqq m$. (By (a), the expression (9.53) is unique.) It follows that

$$\hat{P}(0) = \text{the constant of the right-hand of (9.53)} = M(P) = 1$$

from (9.53). Furthermore,

$$\hat{P}\{\xi_j\} = [\text{the constant of the right-hand of (9.53)}] \times \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{i(\sum_{h=1}^{q}(k_h - A_{j,h})\lambda_h)x} dx.$$

Here remain only terms such as $k_h = A_{j,h}$ in the above integration. Hence

$$\hat{P}\{\xi_j\} = \hat{P}\left\{\sum_{h=1}^{q} A_{j,h}\lambda_h\right\} = \prod_{h=1}^{q}\left(1 - \frac{|A_{j,h}|}{m+1}\right)$$

$$\xi_n = -\sum_{h=1}^{q}(v_h^* - u^* z_h)\lambda_h^*.$$

Thus $\lambda_1^*, \lambda_2^*, \cdots, \lambda_q^*$ satisfy (a) and (b) for $\xi_1, \xi_2, \cdots, \xi_{n-1}$ and also

$$\xi_n \in \sum_{h=1}^{q} \lambda_h^* \mathbb{Z}.$$

This contradicts our assumption. Hence θ must be zero.

$$\geqq (1 - \varepsilon_1)^q > 1 - \varepsilon.$$

This proves (iii). □

Theorem 9.6 *For any $f \in \mathfrak{AP}(\mathbb{R}, \mathbb{C})$, there exists a sequence of trigonometric polynomials in $W(f)$ which uniformly converges to f.*

Proof $\sigma(f)$ is a countable set by Theorem 9.5. So we may write $\sigma(f) = \{\xi_1, \xi_2, \cdots\}$. Let P_n be a trigonometric polynomial which satisfies (i)–(iii) of Lemma 9.9 for $\xi_1, \xi_2, \cdots, \xi_n$ and $\varepsilon = 1/n$. If we define $T_n = f \underset{M}{*} P_n$, then $T_n \in W(f)$ by Lemma 9.7. It also follows that

$$\lim_{n\to\infty} \hat{T}_n\{\xi_j\} = \hat{f}\{\xi_j\} = 0 \quad \text{for all} \quad \xi_j \in \sigma(f)$$

by Lemma 9.9 (iii). If $\xi \notin \sigma(f)$, then

$$\hat{T}_n\{\xi\} = \hat{f}\{\xi\} = 0 \quad \text{for all } n.$$

On the other hand, let g be a $\|\cdot\|_\infty$-limiting point of T_n in the compact set $W(f)$. If a subsequence $\{T_{n'}\}$ of $\{T_n\}$ uniformly converges to g, we have

$$\lim_{n'\to\infty} \hat{T}_{n'}\{\xi\} = \hat{g}\{\xi\}.$$

Hence

$$\hat{g}\{\xi\} = \hat{f}(\xi) \quad \text{for all } \xi.$$

By the remark concerning the uniqueness of Fourier coefficients, we must have $g = f$. Consequently, $\{T_n\}$ uniformly converges to f. □

9.5 Almost Periodic Weakly Stationary Stochastic Processes

Let $(\Omega, \mathcal{E}, P)$ be a probability space. T is either $\mathbb{Z}$ or $\mathbb{R}$, interpreted as time space. For the sake of simplicity, we assume $\mathbb{E}X(t, \omega) = 0$ $(t \in \mathbb{R})$.

Let $X : T \times \Omega \to \mathbb{C}$ be a weakly stationary process with the covariance function $\rho(u)$. $X(t, \omega)$ is called an **almost periodic weakly stationary process** if $\rho(u)$ is almost periodic.

The following theorems play roles similar to those of Theorems 8.8 and 8.8′ (pp. 214–215).

Theorem 9.7 *Let $X : \mathbb{Z} \times \Omega \to \mathbb{C}$ be a weakly stationary process with the spectral measure ν. Then the following three statements are equivalent:*

(i) *$X_n(\omega)$ is an almost periodic process.*

(ii) *There exists $\Gamma = \Gamma(\varepsilon, X) \in \mathbb{R}$ for each $\varepsilon > 0$ such that any interval, the length of which is $\Gamma(\varepsilon, X)$, contains some $\tau \in \mathbb{Z}$ which satisfies*

$$\sup_{n\in\mathbb{Z}} \mathbb{E}|X_{n+\tau}(\omega) - X_n(\omega)|^2 < \varepsilon.$$

(iii) *ν is discrete.*

Theorem 9.7′ *Let $X : \mathbb{R} \times \Omega \to \mathbb{C}$ be a measurable and weakly stationary process with the spectral measure ν. Then the following three statements are equivalent:*

(i) *$X(t, \omega)$ is an almost periodic process.*
(ii) *There exists $\Gamma = \Gamma(\varepsilon, X) \in \mathbb{R}$ for each $\varepsilon > 0$ such that any interval, the length of which is $\Gamma(\varepsilon, X)$, contains some $\tau \in \mathbb{R}$ which satisfies*

$$\sup_{t\in\mathbb{R}} \mathbb{E}|X(t+\tau, \omega) - X(t, \omega)|^2 < \varepsilon.$$

(iii) *ν is discrete.*

We prove only Theorem 9.7′. (Theorem 9.7 can be proved in an entirely similar manner.)

Proof of Theorem 9.7′ [23] (i)⇒(ii): Assume that $\rho(\mu)$ is almost periodic. Then we have

$$\begin{aligned}\sup_{t\in\mathbb{R}} \mathbb{E}|X(t+\tau, \omega) - X(t, \omega)|^2 &= 2[\rho(0) - \mathcal{R}e\,\rho(\tau)] \\ &= 2\mathcal{R}e[\rho(0) - \rho(\tau)] \leqq 2|\rho(0) - \rho(\tau)| \\ &\leqq 2\sup_{u\in\mathbb{R}} |\rho(u+\tau) - \rho(u)|.\end{aligned} \tag{9.54}$$

Since $\rho(\cdot)$ is almost periodic, there exists a number $\Lambda(\varepsilon/2, \rho) > 0$, for each $\varepsilon > 0$, such that any interval the length of which is $\Lambda(\varepsilon/2, \rho)$ contains an $\varepsilon/2$-almost period, τ. Since

$$\begin{aligned}&\left\{\tau \in \mathbb{R} \,\middle|\, \sup_{u\in\mathbb{R}} |\rho(u+\tau) - \rho(u)| < \frac{\varepsilon}{2}\right\} \\ &\qquad \subset \left\{\tau \in \mathbb{R} \,\middle|\, \sup_{t\in\mathbb{R}} \mathbb{E}|X(t+\tau, \omega) - X(t, \omega)|^2 < \varepsilon\right\}\end{aligned}$$

by (9.54), we see that (ii) is satisfied by setting $\Gamma(\varepsilon, X) = \Lambda(\varepsilon/2, \rho)$.

(ii)⇒(i): Assume (ii). By a simple calculation, we obtain the inequality

$$|\rho(u+\tau) - \rho(u)|^2 = |\mathbb{E}[X(u+\tau, \omega)\overline{X(0, \omega)} - X(u, \omega)\overline{X(0, \omega)}]|^2$$

[23] The proof of (i)⇔(ii) is due to Kawata [11] pp. 80–82.

$$\leqq \mathbb{E}|X(u+\tau,\omega) - X(u,\omega)|^2 \mathbb{E}|X(0,\omega)|^2$$
$$\text{(by Schwarz's inequality).}$$

It follows that

$$|\rho(u+\tau) - \rho(u)| \leqq [\mathbb{E}|X(t+\tau,\omega) - X(u,\omega)|^2]^{1/2}\rho(0)^{1/2}.$$

Hence we have

$$\left\{\tau \in \mathbb{R} \,\middle|\, \sup_{u\in\mathbb{R}} \mathbb{E}|X(u+\tau,\omega) - X(u,\omega)|^2 < \frac{\varepsilon^2}{\rho(0)}\right\}$$
$$\subset \left\{\tau \in \mathbb{R} \,\middle|\, \sup_{u\in\mathbb{R}} |\rho(u+\tau) - \rho(u)| < \varepsilon\right\}.$$

Thus (i) holds good by setting $\Lambda(\varepsilon,\rho) = \Gamma(\varepsilon^2/\rho(0), X)$.

(i)⇒(iii): The covariance function ρ is represented by the Fourier transform of some positive Radon measure ν on $\mathbb{R}$. Since $\rho(u) = \hat{\nu}(u)$ is almost periodic, ν must be discrete by Corollary 9.3.

(iii)⇒(i): Assume that the spectral measure ν of X is discrete, say

$$\nu = \sum_{n=1}^{\infty} a_n \delta_{\xi_n} \quad (\delta_{\xi_n} : \text{Dirac measure}).$$

Then the covariance can be expressed as

$$\rho(u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\lambda u} d\nu(\lambda) = \frac{1}{\sqrt{2\pi}} \sum_{n=1}^{\infty} a_n e^{-i\xi_n u}. \tag{9.55}$$

Since $a_n \geqq 0$, the series (9.55) is absolutely and hence uniformly convergent. Thus $\rho(u)$ is the uniform limit of trigonometric polynomials,[24] and hence almost periodic. □

A (measurable) weakly stationary process $X(t,\omega)$ is τ-periodic if and only if its spectral measure ν concentrates on a countable set in $\mathbb{T}$ or $\mathbb{R}$ such that the distance of any adjacent two points is some multiple of $2\pi/\tau$ (cf. Theorem 8.8′ (p. 215)).

$X(t,\omega)$ is almost periodic if and only if ν is discrete (cf. Theorem 9.7′).

[24] A function of the form

$$f(x) = \sum_{j=1}^{n} a_j e^{-i\xi_j u} \quad (\xi_j \in \mathbb{R})$$

is called a trigonometric polynomial. Any trigonometric polynomial is almost periodic.

Therefore if $X(t,\omega)$ is periodic or almost periodic, the spectral measure ν can never be absolutely continuous with respect to the Lebesgue measure.

Now under what conditions does the spectral measure have the density function? The answers were given in Theorem 8.12 (p. 228) and Theorem 8.13 (p. 236). ν is absolutely continuous with respect to the Lebesgue measure if and only if $X(t,\omega)$ is a moving average process of white noise.

So we can conclude that *any moving average process can be neither periodic nor almost periodic.*

However, we know that the set $\mathfrak{M}_+^d(\mathbb{R}) = \left\{ \sum_{j=1}^{n} \alpha_j \delta_{x_j} \middle| \alpha_j \in \mathbb{R}, x_j \in \mathbb{R}, n \in \mathbb{N} \right\}$ of all the discrete positive Radon measures is w^*-dense in $\mathfrak{M}_+(\mathbb{R})$, the set of all the positive Radon measures (cf. Billingsley [1] p. 237, Malliavin [13] Chap. II, §6, Maruyama [14] Chap. 8,§2).

The theorem below immediately follows from this observation.

Theorem 9.8 *Let $X(t,\omega)$ be a measurable and weakly stationary process on $\mathbb{R}\times\Omega$ with the spectral measure ν. Then there exists a sequence of almost periodic weakly stationary processes $X^k(t,\omega)$ with the spectral measure ν^k such that ν^k converges to ν in the w^*-topology.*

Proof Since $\mathfrak{M}_+^d(\mathbb{R})$ is w^*-dense in $\mathfrak{M}_+(\mathbb{R})$ (metrizable), there exists a sequence $\nu^k \in \mathfrak{M}_+^d(\mathbb{R})$ which w^*-converges to ν. By Theorem 8.7′ (p. 213), there is a measurable and weakly stationary process $X^k(t,\omega)$, the spectral measure of which is exactly equal to ν^k. $X^k(t,\omega)$ is almost periodic because $\nu^k \in \mathfrak{M}_+^d(\mathbb{R})$. □

A similar result also holds good for the discrete time case. (In this case $\mathbb{R}$ should be replaced by $\mathbb{T}$ as usual.)

Theorem 9.8 tells us that *any weakly stationary process $X(t,\omega)$ can be approximated by a sequence $\{X^k(t,\omega)\}$ of almost periodic weakly stationary processes* in the sense that the sequence of the spectral measures of $X^k(t,\omega)$ converges to the spectral measure of $X(t,\omega)$ in the w^*-topology.[25]

References

1. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
2. Bochner, S.: Beiträge zur Theorie der fastperiodischen Funktionen, I, II. Math. Ann. **96**, 119–147, 383–409 (1927)
3. Bochner, S., von Neumann, J.: Almost periodic functions in a group, II. Trans. Amer. Math. Soc. **37**, 21–50 (1935)
4. Bohr, H.: Zur Theorie der fastperiodischen Funktionen. I–III. Acta Math. **45**, 29–127 (1925); **46**, 101–214 (1925); **47**, 237–281 (1926)
5. Bohr, H.: Fastperiodische Funktionen. Springer, Berlin (1932)

[25] This observation is based upon Maruyama [15].

6. Dudley, R.M.: Real Analysis and Probability. Wadsworth and Brooks, Pacific Grove (1988)
7. Dunford, N., Schwartz, J.T.: Linear Operators, Part 1, Interscience, New York (1958)
8. Katznelson, Y.: An Introduction to Harmonic Analysis, 3rd edn. Cambridge University Press, Cambridge (2004)
9. Kawata, T.: Ohyo Sugaku Gairon (Elements of Applied Mathematics). I, II. Iwanami Shoten, Tokyo (1950, 1952) (Originally published in Japanese)
10. Kawata, T.: On the Fourier series of a stationary stochastic process, I, II. Z. Wahrsch. Verw. Gebiete, **6**, 224–245 (1966); **13**, 25–38 (1969)
11. Kawata, T.: Teijo Kakuritsu Katei (Stationary Stochastic Processes). Kyoritsu Shuppan, Tokyo (1985) (Originally published in Japanese)
12. Loomis, L.: The spectral characterization of a class of almost periodic functions. Ann. Math. **72**, 362–368 (1960)
13. Malliavin, P.: Integration and Probability. Springer, New York (1995)
14. Maruyama, T.: Sekibun to Kansu-kaiseki (Integration and Functional Analysis). Springer, Tokyo (2006) (Originally published in Japanese)
15. Maruyama, T.: Fourier analysis of periodic weakly stationary processes. A note on Slutsky's observation. Adv. Math. Econ. **20**, 151–180 (2016)
16. Rudin, W.: Weak almost periodic functions and Fourier–Stieltjes transforms. Duke Math. J. **26**, 215–220 (1959)
17. Rudin, W.: Fourier Analysis on Groups. Interscience, New York (1962)
18. Stromberg, K.R.: An Introduction to Classical Real Analysis. American Mathematical Society, Providence (1981)
19. von Neumann, J.: Almost periodic functions in a group, I. Trans. Amer. Math. Soc. **36**, 445–492 (1934)

Chapter 10
Fredholm Operators

A bounded linear operator acting between Banach spaces is called a Fredholm operator if the dimension of its kernel and the codimension of its image are both finite. An equation defined by a Fredholm operator sometimes enjoys a nice property which reduces the difficulties associated with infinite dimension to the finite dimensional problem. The object of this chapter is to study the basic elements of Fredholm operators, which will be made use of in the next chapter in the context of bifurcation theory.

It is not easy to extend the spectral theory of linear operators between finite dimensional spaces to infinite dimensional case. Several special kinds of operators are known for which detailed spectral theory can be constructed even in an infinite dimensional context; compact operators and Fredholm operators are typical and remarkable examples. In this book, we do not intend to go into details of spectral theory itself.[1]

10.1 Direct Sums and Projections

All the normed linear spaces in this chapter are assumed to be real.

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two subspaces of a vector space $\mathfrak{X}$. If every $x \in \mathfrak{X}$ can be expressed uniquely as $x = m_1 + m_2$ for some $m_1 \in \mathfrak{M}_1$ and $m_2 \in \mathfrak{M}_2$, $\mathfrak{X}$ is said to be a **direct sum** of $\mathfrak{M}_1$ and $\mathfrak{M}_2$, and this situation is denoted by $\mathfrak{X} = \mathfrak{M}_1 \oplus \mathfrak{M}_2$.

Suppose that $\mathfrak{X} = \mathfrak{M}_1 \oplus \mathfrak{M}_2$. If $x \in \mathfrak{X}$ is expressed uniquely as

$$x = m_1 + m_2; \quad m_1 \in \mathfrak{M}_1, \ m_2 \in \mathfrak{M}_2,$$

[1] We are indebted to Kuroda [3] Chap. 11 and Zeidler [6] Chaps. 3 and 5. See also Kato [2] Chap. 6 for the spectral theory of compact operators.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_10

the operators

$$P_1 : x \mapsto m_1 \quad (\mathfrak{X} \to \mathfrak{M}_1),$$
$$P_2 : x \mapsto m_2 \quad (\mathfrak{X} \to \mathfrak{M}_2)$$

are called the **projection** to $\mathfrak{M}_1$ (along $\mathfrak{M}_2$) and the projection to $\mathfrak{M}_2$ (along $\mathfrak{M}_1$), respectively. Both P_1 and P_2 are linear operators and have basic properties as follows:

1° $P_1^2 = P_1, \quad P_2^2 = P_2$.
2° $P_1(\mathfrak{X}) = \mathfrak{M}_1, \quad P_2(\mathfrak{X}) = \mathfrak{M}_2$.
3° $\mathfrak{M}_2 = (I - P_1)(\mathfrak{X}) = \mathrm{Ker} P_1$,
$\mathfrak{M}_1 = (I - P_2)(\mathfrak{X}) = \mathrm{Ker} P_2$.

Theorem 10.1 *Suppose that $\mathfrak{X}$ is a linear space and a linear operator $P : \mathfrak{X} \to \mathfrak{X}$ satisfies $P^2 = P$. Then $\mathfrak{X} = \mathfrak{M}_1 \oplus \mathfrak{M}_2$, where*

$$\mathfrak{M}_1 = P(\mathfrak{X}), \quad \mathfrak{M}_2 = (I - P)(\mathfrak{X}).$$

Proof It is clear that $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are linear subspaces of $\mathfrak{X}$. For each $x \in \mathfrak{X}$, we define m_1 and m_2 as

$$m_1 = Px, \quad m_2 = (I - P)x.$$

Then, obviously,

$$x = m_1 + m_2\ ; \quad m_1 \in \mathfrak{M}_1,\ m_2 \in \mathfrak{M}_2.$$

Such an expression is unique. If another expression

$$x = m_1' + m_2'\ ; \quad m_1' \in \mathfrak{M}_1,\ m_2' \in \mathfrak{M}_2$$

is possible, there exist $x', x'' \in \mathfrak{M}$ which satisfy

$$m_1' = Px', \quad m_2' = (I - P)x''$$

by the definitions of $\mathfrak{M}_1$ and $\mathfrak{M}_2$. Since $P^2 = P$, we have

$$Pm_1' = m_1', \quad Pm_2' = 0.$$

Hence it holds good that

$$m_1 = Px = Pm_1' + Pm_2' = Pm_1' = m_1'.$$

Similarly, it is easy to show that $m_2 = m_2'$. □

Theorem 10.2 *Let $\mathfrak{X}$ be a linear space and $\mathfrak{M}$ its subspace.*

(i) *There exists a subspace $\mathfrak{N}$ of $\mathfrak{X}$ such that $\mathfrak{X} = \mathfrak{M} \oplus \mathfrak{N}$.*
(ii) *If $\mathfrak{X} = \mathfrak{M} \oplus \mathfrak{N}$, then*

$$\mathrm{codim}\mathfrak{M} = \dim\mathfrak{N}.$$

(iii) *If $\mathfrak{X} = \mathfrak{M} \oplus \mathfrak{N}$, then*

$$\dim\mathfrak{X} = \dim\mathfrak{M} + \dim\mathfrak{N} = \dim\mathfrak{M} + \mathrm{codim}\mathfrak{M}.$$

(Hence in the case of $\dim\mathfrak{X} < \infty$, the equality $\mathrm{codim}\mathfrak{M} = \dim\mathfrak{X} - \dim\mathfrak{M}$ holds good.)

Proof (i) Let $\mathcal{P}$ be the set of all the linear operators $P : D(P) \to \mathfrak{M}$ such that the domain $D(P)$ of P includes $\mathfrak{M}$ and $P(x) = x$ for all $x \in \mathfrak{M}$. Of course, $\mathcal{P} \neq \emptyset$.[2] Define a partial ordering $\leq$ on $\mathcal{P}$ by

$$P_1 \leq P_2 \Leftrightarrow P_2 \text{ is an extension of } P_1.$$

Then any chain in $\mathcal{P}$ has an upper bound in $\mathcal{P}$. Hence, by Zorn's lemma, $\mathcal{P}$ has a maximal element P_0 with respect to $\leq$. It can be shown that P_0 is defined on the whole space $\mathfrak{X}$. Assume, on the contrary, that $D(P_0) \subsetneq \mathfrak{X}$. Define a subspace $\mathfrak{N}$ by $\mathfrak{N} = D(P_0) + \mathrm{span}\{x_0\}$, where x_0 is any element of $\mathfrak{X} \setminus D(P_0)$. Then the operator $P : \mathfrak{N} \to \mathfrak{M}$ defined by

$$P(x + \alpha x_0) = P_0(x), \quad x \in D(P_0), \quad \alpha \in \mathbb{R}$$

is an element of $\mathcal{P}$ and is an extension of P_0. This contradicts the maximality of P_0.

Furthermore, $P_0^2 = P_0$ can be verified easily. In fact, $P_0(P_0x) = P_0x$, since $P_0x \in \mathfrak{M}$ for any $x \in \mathfrak{X}$. Hence (i) follows by Theorem 10.1.

(ii) Assume that $\mathfrak{X} = \mathfrak{M} \oplus \mathfrak{N}$. Define a linear operator $T : \mathfrak{M} \to \mathfrak{X}/\mathfrak{N}$ by

$$Tx = \xi_x, \quad x \in \mathfrak{M},$$

where ξ_x is the residue class $x + \mathfrak{N}$ of x modulo $\mathfrak{N}$. Since T is clearly a bijection, we obtain

$$\dim\mathfrak{M} = \dim\mathfrak{X}/\mathfrak{N}.$$

[2]Let $P : \mathfrak{M} \to \mathfrak{M}$ the identity mapping on $\mathfrak{M}$. This P is clearly contained in $\mathcal{P}$.

(iii) It is enough to show only $\dim\mathfrak{X} = \dim\mathfrak{M} + \dim\mathfrak{N}$. If $\dim\mathfrak{M} = \infty$ or $\dim\mathfrak{N} = \infty$, it is obvious that $\dim\mathfrak{X} = \infty$. So there is nothing to prove in this case. If $\dim\mathfrak{M} < \infty$ and $\dim\mathfrak{N} < \infty$, the union of bases of $\mathfrak{M}$ and $\mathfrak{N}$ forms a basis of $\mathfrak{X}$. The desired result immediately follows. □

We now introduce a topology to $\mathfrak{X}$. Suppose that $\mathfrak{X}$ is a normed linear space and $\mathfrak{X} = \mathfrak{M}_1 \oplus \mathfrak{M}_2$ for some pair of subspaces, $\mathfrak{M}_1$ and $\mathfrak{M}_2$. Let P_1 and P_2 be the projections to $\mathfrak{M}_1$ and $\mathfrak{M}_2$, respectively. Then it is almost obvious that the following three statements are equivalent:

(i) P_1 is continuous.
(ii) P_2 is continuous.
(iii) Both P_1 and P_2 are continuous.

If any one of (i)–(iii) holds good, $\mathfrak{X}$ is said to be a **topological direct sum** of $\mathfrak{M}_1$ and $\mathfrak{M}_2$. $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are called **topological complements** of each other.

Theorem 10.3 *Let $\mathfrak{X}$ be a normed linear space, with a subspace $\mathfrak{M}$.*

(i) *$\mathfrak{M}$ has a topological complement if and only if there exists a continuous linear operator $P : \mathfrak{X} \to \mathfrak{X}$ such that*

$$a.\ P^2 = P, \qquad b.\ P(\mathfrak{X}) = \mathfrak{M}.$$

(ii) *If $\mathfrak{M}$ has a topological complement, $\mathfrak{M}$ is closed.*

Theorem 10.4 *Let $\mathfrak{X}$ be a Banach space such that $\mathfrak{X} = \mathfrak{M}_1 \oplus \mathfrak{M}_2$ for some subspaces $\mathfrak{M}_1, \mathfrak{M}_2$. Then the following two statements are equivalent:*

(i) *$\mathfrak{X}$ is a topological direct sum of $\mathfrak{M}_1$ and $\mathfrak{M}_2$.*
(ii) *$\mathfrak{M}_1$ and $\mathfrak{M}_2$ are closed.*

Proof Since (i)⇒(ii) is obvious, we have only to prove (ii)⇒(i).

Let P_1 and P_2 be the projections to $\mathfrak{M}_1$ and to $\mathfrak{M}_2$, respectively. We have to show their continuity. By the remark stated above, it is sufficient to show the continuity of either of P_1 and P_2, say P_1. Furthermore, P_1 is continuous if its graph $G(P_1) = \{(x, m_1) \in \mathfrak{X} \times \mathfrak{M}_1 \mid m_1 = P_1 x\}$ is closed in $\mathfrak{X} \times \mathfrak{M}_1$.[3] Let $\{x_n\}$ be a sequence in $\mathfrak{X}$ such that

$$m_{1n} = P_1 x_n, \quad m_{2n} = P_2 x_n, \quad n = 1, 2, \cdots$$

Then we have, of course, $(x_n, m_{1n}) \in G(P_1)$. If

$$(x_n, m_{1n}) \to (x^*, m_1^*) \quad \text{as} \quad n \to \infty,$$

[3] This is so by the closed graph theorem (Dunford and Schwartz [1] pp. 57–58, Maruyama [4] pp. 145–146).

then $m_{2n} \to m_2^* \equiv x^* - m_1^*$ (as $n \to \infty$). By (ii), $m_1^* \in \mathfrak{M}_1$ and $m_2^* \in \mathfrak{M}_2$. Hence

$$m_1^* = P_1 x^*, \quad m_2^* = P_2 x^*$$

and so $(x^*, m_1^*) \in G(P_1)$. Therefore $G(P_1)$ is closed. □

Theorem 10.5 provides several cases in which a Banach space can be a topological direct sum. We need a new concept of a biorthogonal system to prove it.

Definition 10.1 Suppose that $\mathfrak{X}$ is a normed linear space and that $x_1, x_2, \cdots, x_n \in \mathfrak{X}$ and $\Lambda_1, \Lambda_2, \cdots, \Lambda_n \in \mathfrak{X}'$.[4] $\{x_j, \Lambda_j\}_{j=1,2,\cdots,n}$ is called a **biorthogonal system** of $\mathfrak{X}$ if it satisfies

$$\langle \Lambda_i, x_j \rangle = \delta_{ij} \quad \text{for all} \quad i, j = 1, 2, \cdots, n.$$

Lemma 10.1 *Let $\mathfrak{X}$ be a normed linear space. If $\Lambda_1, \Lambda_2, \cdots, \Lambda_{k+1} \in \mathfrak{X}'$ satisfy $\mathrm{Ker}\Lambda_1 \cap \cdots \cap \mathrm{Ker}\Lambda_k \subset \mathrm{Ker}\Lambda_{k+1}$, then Λ_{k+1} can be expressed as a linear combination of $\Lambda_1, \Lambda_2, \cdots, \Lambda_k$.*

Proof We prove this lemma by induction.

Suppose first that $k = 1$, and $\mathrm{Ker}\Lambda_1 \subset \mathrm{Ker}\Lambda_2$. Without loss of generality, we may assume that $\Lambda_1 \neq 0$. Since $\mathfrak{X}/\mathrm{Ker}\Lambda_1 \cong \mathbb{R}$, we obtain

$$\mathfrak{X} = \mathrm{Ker}\Lambda_1 \oplus \mathrm{span}\{x_0\}, \quad x_0 \notin \mathrm{Ker}\Lambda_1 \quad \text{(topological direct sum)}$$

by Theorem 10.5 (iii). Any $x \in \mathfrak{X}$ can be expressed as

$$x = u + \alpha x_0, \quad u \in \mathrm{Ker}\Lambda_1, \quad \alpha \in \mathbb{R}.$$

Taking account of $\mathrm{Ker}\Lambda_1 \subset \mathrm{Ker}\Lambda_2$, we obtain

$$\Lambda_1(x) = \alpha\Lambda_1(x_0), \quad \Lambda_2(x) = \alpha\Lambda_2(x_0).$$

Since $\Lambda_1(x_0) \neq 0$, we can define $\gamma = \Lambda_2(x_0)/\Lambda_1(x_0)$. Thus $\Lambda_2(x) = \gamma\Lambda_1(x)$.

Next we assume that the claim of the lemma holds good if the number of given elements of $\mathfrak{X}'$ is k. Consider the case of $k + 1 (k \geqq 2)$. We write

$$\tilde{\Lambda}_j = \Lambda_j|_{\mathrm{Ker}\Lambda_1}, \quad j = 2, 3, \cdots, k+1.$$

If we regard $\mathrm{Ker}\Lambda_1$ as a normed linear space,

$$\tilde{\Lambda}_j \in (\mathrm{Ker}\Lambda_1)', \quad j = 2, 3, \cdots, k+1.$$

[4] $\mathfrak{X}'$ is the dual space of $\mathfrak{X}$. $\langle \Lambda_i, x_j \rangle$ means the value of Λ_j at x_j; i.e. $\Lambda_j(x_j)$. δ_{ij} is Kronecker's delta.

By assumption

$$\mathrm{Ker}\tilde{\Lambda}_2 \cap \mathrm{Ker}\tilde{\Lambda}_3 \cap \cdots \cap \mathrm{Ker}\tilde{\Lambda}_k \subset \mathrm{Ker}\tilde{\Lambda}_{k+1}.$$

Hence, by assumption of induction, we can express $\tilde{\Lambda}_{k+1}$ as $\tilde{\Lambda}_{k+1} = \gamma_2\tilde{\Lambda}_2 + \cdots + \gamma_k\tilde{\Lambda}_k$ $(\gamma_j \in \mathbb{R})$. If we define $\Phi \in \mathfrak{X}'$ by

$$\Phi = \Lambda_{k+1} - (\gamma_2\Lambda_2 + \cdots + \gamma_k\Lambda_k) \in \mathfrak{X}',$$

Φ satisfies

$$\Phi = \tilde{\Lambda}_{k+1} - (\gamma_2\tilde{\Lambda}_2 + \cdots + \gamma_k\tilde{\Lambda}_k) = 0 \quad \text{on} \quad \mathrm{Ker}\Lambda_1.$$

That is $\mathrm{Ker}\Lambda_1 \subset \mathrm{Ker}\Phi$. Consequently, we can express Φ as $\Phi = \gamma_1\Lambda_1$, following the case of $k = 1$. Finally, we obtain

$$\Lambda_{k+1} = \gamma_1\Lambda_1 + \gamma_2\Lambda_2 + \cdots + \gamma_k\Lambda_k.$$

□

Lemma 10.2 *Let $\mathfrak{X}$ be a normed linear space.*

(i) *If $x_1, \cdots, x_n \in \mathfrak{X}$ are linearly independent, there exist some $\Lambda_1, \cdots, \Lambda_n \in \mathfrak{X}'$ such that $\{x_j, \Lambda_j\}_{j=1,2,\cdots,n}$ is a biorthogonal system.*
(ii) *If $\Lambda_1, \cdots, \Lambda_n \in \mathfrak{X}'$ are linearly independent, there exist some $x_1, \cdots, x_n \in \mathfrak{X}$ such that $\{x_j, \Lambda_j\}_{j=1,2,\cdots,n}$ is a biorthogonal system.*

Proof (i) Let $\mathfrak{M}_j$ be a subspace of $\mathfrak{X}$ spanned by $x_1, \cdots, x_{j-1}, x_{j+1}, \cdots, x_n$, excluding x_j. Of course, $x_j \notin \mathfrak{M}_j$. By the Hahn–Banach theorem, there exists a $\Lambda_j \in \mathfrak{X}'$ which satisfies $\Lambda_j(x_j) = 1$ and vanishes on $\mathfrak{M}_j$. Combining $x_1, x_2, \cdots, x_n$ with $\Lambda_1, \Lambda_2, \cdots, \Lambda_n$ thus chosen, we obtain a biorthogonal system.

(ii) By Lemma 10.1, $\bigcap_{i \neq j} \mathrm{Ker}\Lambda_i \setminus \mathrm{Ker}\Lambda_j \neq \emptyset$ for every j. Hence there exists some $x_j \in \mathfrak{X}$ which satisfies

$$\Lambda_j(x_j) = 1, \quad \Lambda_i(x_j) = 0 \quad \text{for} \quad i \neq j.$$

Thus $\{x_j, \Lambda_j\}_{j=1,2,\cdots,n}$ is a biorthogonal system. □

Theorem 10.5 *Let $\mathfrak{X}$ be a Banach space, with a subspace $\mathfrak{M}$. Then $\mathfrak{M}$ has a topological complement if either one of the following conditions is satisfied:*

(i) *$\mathfrak{X}$ is a Hilbert space, with a closed subspace $\mathfrak{M}$.*
(ii) $\dim\mathfrak{M} < \infty$.
(iii) *$\mathfrak{M}$ is closed, and* $\mathrm{codim}\mathfrak{M} < \infty$.

Proof (i) is explained in almost all the textbooks on functional analysis,[5] so we may omit it here.

(ii) Let $\{x_1, x_2, \cdots, x_n\}$ be the basis of $\mathfrak{M}$. By Lemma 10.1, there exist $\Lambda_1, \Lambda_2, \cdots, \Lambda_n \in \mathfrak{X}'$ such that $\{x_j, \Lambda_j\}_{j=1,\cdots,n}$ is a biorthogonal system. Define a linear operator $P : \mathfrak{X} \to \mathfrak{X}$ by

$$Px = \sum_{j=1}^{n} \Lambda_j(x)x_j; \quad x \in \mathfrak{X}.$$

Then P is continuous and satisfies $P^2 = P$, since

$$\begin{aligned} P^2x &= \sum_{k=1}^{n} \Lambda_k(Px)x_k = \sum_{k=1}^{n} \Lambda_k\left(\sum_{j=1}^{n} \Lambda_j(x)x_j\right)x_k \\ &= \sum_{k=1}^{n}\left[\sum_{j=1}^{n} \Lambda_j(x)\Lambda_k(x_j)\right]x_k = \sum_{k=1}^{n} \Lambda_k(x)x_k = Px; \quad x \in \mathfrak{X}. \end{aligned}$$

We note that $Px \in \mathfrak{M}$ for any $x \in \mathfrak{X}$, since Px is a linear combination of $x_1, x_2, \cdots, x_n$. Hence $P(\mathfrak{X}) \subset \mathfrak{M}$. Conversely, it is also clear that $P(\mathfrak{X}) \supset \mathfrak{M}$, since $Px_j = x_j (j = 1, 2, \cdots, n)$. So we obtain $P(\mathfrak{X}) = \mathfrak{M}$. By Theorem 10.3, $\mathfrak{M}$ has a topological complement.

(iii) Let

$$\xi_{y_k} \equiv y_k + \mathfrak{M}; \quad k = 1, 2, \cdots, p$$

be a basis of $\mathfrak{X}/\mathfrak{M}$. (Note that $\dim(\mathfrak{X}/\mathfrak{M}) < \infty$.) For any $x \in \mathfrak{X}$, there exists some k such that $x \in y_k + \mathfrak{M}$. Hence there exists some $z \in \mathfrak{M}$ such that

$$x = y_k + z. \tag{10.1}$$

If we define $\mathfrak{N} = \operatorname{span}\{y_1, y_2, \cdots, y_p\}$, it is clear that $\mathfrak{X} = \mathfrak{M} + \mathfrak{N}$. We now show that the expression (10.1) is unique.

The only important point to be checked for proving the uniqueness is $\mathfrak{M} \cap \mathfrak{N} = \{0\}$. In fact, if

$$0 \neq \sum_{k=1}^{p} \alpha_k y_k \in \mathfrak{M} \cap \mathfrak{N},$$

[5] See Maruyama [4] p. 190, Yosida [5] p. 82 for instance.

then we have

$$0 = \xi_{\sum_{k=1}^{p} \alpha_k y_k} = \sum_{k=1}^{p} \alpha_k \xi_{y_k} \quad \text{(0 on the left-hand side means } \xi_0\text{)},$$

which is a contradiction to the assumption that ξ_{y_k} $(k = 1, 2, \cdots, p)$ is a basis of $\mathfrak{X}/\mathfrak{M}$.

Suppose that $x \in \mathfrak{X}$ can be expressed in two ways,

$$x = z + y = z' + y', \quad z, z' \in \mathfrak{M}, \quad y, y' \in \mathfrak{N}.$$

Then we have obviously

$$z - z' = y' - y \in \mathfrak{M} \cap \mathfrak{N}.$$

This implies that $z = z'$ and $y = y'$. Hence $\mathfrak{X} = \mathfrak{M} \oplus \mathfrak{N}$. By Theorem 10.4, it is a topological direct sum. □

If a subspace $\mathfrak{M}$ satisfies (iii), then any subspace $\mathfrak{N}$ containing $\mathfrak{M}$ also satisfies (iii).

Theorem 10.6 *Let $\mathfrak{X}$ be a Banach space and $\mathfrak{M}$ a closed subspace with codim $\mathfrak{M} < \infty$. Then any subspace $\mathfrak{N}$ containing $\mathfrak{M}$ is closed and codim $\mathfrak{N} < \infty$.*

Proof The canonical mappings, $\pi_1 : \mathfrak{X} \to \mathfrak{X}/\mathfrak{M}$ and $\pi_2 : \mathfrak{N} \to \mathfrak{N}/\mathfrak{M}$, are continuous in the quotient topologies for $\mathfrak{X}/\mathfrak{M}$ and $\mathfrak{N}/\mathfrak{M}$. Since dim $\mathfrak{X}/\mathfrak{M} < \infty$, it is obvious that dim $\mathfrak{N}/\mathfrak{M} < \infty$. Hence $\mathfrak{N}/\mathfrak{M}$ is a closed subspace of $\mathfrak{X}/\mathfrak{M}$. It follows that $\mathfrak{N} = \pi_1^{-1}(\mathfrak{N}/\mathfrak{M})$ is closed.

Suppose that $\xi_{x_1}, \xi_{x_2}, \cdots, \xi_{x_n} \in \mathfrak{X}/\mathfrak{N}$ are linearly independent, where $x_j \in \mathfrak{X}$ and $\xi_{x_j} = x_j + \mathfrak{N}(j = 1, 2, \cdots, n)$. Since

$$\sum_{j=1}^{n} \alpha_j x_j \equiv 0 \pmod{\mathfrak{N}} \Rightarrow \quad \alpha_j = 0 \quad \text{for all } j,$$

we also obtain

$$\sum_{j=1}^{n} \alpha_j x_j \equiv 0 \pmod{\mathfrak{M}} \Rightarrow \quad \alpha_j = 0 \quad \text{for all } j.$$

That is, $x_1 + \mathfrak{M}, \ x_2 + \mathfrak{M}, \ \cdots, x_n + \mathfrak{M}$ are linearly independent in $\mathfrak{X}/\mathfrak{M}$. Consequently, we have

$$\dim \mathfrak{X}/\mathfrak{N} \leqq \dim \mathfrak{X}/\mathfrak{M}.$$

□

Lemma 10.3 *Let $\mathfrak{X}$ be a normed linear space, with a closed subspace $\mathfrak{M}$. Then* $\dim \mathfrak{M}^{\perp} < \infty$ *if and only if there exists a subspace $\mathfrak{N}$ which satisfies $\mathfrak{X} = \mathfrak{M} \oplus \mathfrak{N}$ and* $\dim \mathfrak{N} < \infty$. *In this case,* $\dim \mathfrak{M}^{\perp} = \dim \mathfrak{N}$.[6]

Proof (Sufficiency) Assume that $\mathfrak{X} = \mathfrak{M} \oplus \mathfrak{N}$ and $\dim \mathfrak{N} = n < \infty$. Let $v_1, v_2, \cdots, v_n$ be a basis of $\mathfrak{N}$. Then there exist $\Lambda_j \in \mathfrak{M}^{\perp} (j = 1, 2, \cdots, n)$ such that $\{v_j, \Lambda_j\}_{j=1,2,\cdots,n}$ is a biorthogonal system. We can prove this in a way similar to Lemma 10.2 (i). That is, if we define

$$\mathfrak{N}_j = \mathrm{span}\{v_1, \cdots, v_{j-1}, v_{j+1}, \cdots, v_n\}, \quad j = 1, 2, \cdots, n$$

and $\mathfrak{M}_j = \mathfrak{M} \oplus \mathfrak{N}_j$, then $\mathfrak{M}_j$ is a closed subspace by Theorem 10.3. Of course $v_j \notin \mathfrak{M}_j$. By the Hahn–Banach theorem, there exists $\Lambda_j \in \mathfrak{X}'$, which satisfies

$$\Lambda_j(v_j) = 1, \quad \Lambda_j(v) = 0 \quad \text{on } \mathfrak{M}_j.$$

$\Lambda_1, \Lambda_2, \cdots, \Lambda_n$ thus defined clearly fulfill the desired conditions.

$\Lambda_1, \Lambda_2, \cdots, \Lambda_n$ are linearly independent.[7] We have only to verify that any $\Lambda \in \mathfrak{M}^{\perp}$ is a linear combination of $\Lambda_1, \Lambda_2, \cdots, \Lambda_n$ in order to conclude that $\mathfrak{M}^{\perp} = \mathrm{span}\{\Lambda_1, \Lambda_2, \cdots, \Lambda_n\}$ and $\dim \mathfrak{M}^{\perp} = n$. Let $\Lambda \in \mathfrak{M}^{\perp}$. Defining

$$\Phi = \sum_{j=1}^{n} \Lambda(v_j)\Lambda_j,$$

we obtain

$$\Phi(v_k) = \sum_{j=1}^{n} \Lambda(v_j)\delta_{jk} = \Lambda(v_k).$$

On the other hand, $x \in \mathfrak{X}$ can be expressed as

$$x = u + \sum_{j=1}^{n} \alpha_j v_j, \quad u \in \mathfrak{M}.$$

[6] $\mathfrak{M}^{\perp}$ denotes the **annihilator**, i.e.

$$\mathfrak{M}^{\perp} = \{\Lambda \in \mathfrak{X}' | \Lambda x = 0 \quad \text{for all } x \in \mathfrak{M}\}.$$

[7] Set $\alpha_1\Lambda_1 + \cdots + \alpha_n\Lambda_n = 0$. Evaluate the left-hand side at v_j. By the definition of a biorthogonal system, it follows that $\alpha_j = 0$.

Since $\Lambda, \Phi \in \mathfrak{M}^\perp$, we have

$$\Lambda(x) = \Phi(x) \quad \text{for all} \quad x \in \mathfrak{X}$$

by $\Lambda(v_j) = \Phi(v_j)$. That is, $\Lambda = \Phi$. Hence Λ is a linear combination of Λ_j's.

(Necessity) Assume that $\dim \mathfrak{M}^\perp = n < \infty$ and $\Lambda_1, \Lambda_2, \cdots, \Lambda_n$ form a basis of $\mathfrak{M}^\perp$. By Lemma 10.2, there exists a biorthogonal system of the form $\{u_j, \Lambda_j\}_{j=1,2,\cdots,n}$.

Let $\mathfrak{N} = \text{span}\{v_1, v_2, \cdots, v_n\}$. We show that $\mathfrak{X} = \mathfrak{M} \oplus \mathfrak{N}$ (topological direct sum). It is easy to verify $\mathfrak{M} \cap \mathfrak{N} = \{0\}$. Any $x \in \mathfrak{M} \cap \mathfrak{N}$ can be expressed as

$$x = \sum_{j=1}^{m} \alpha_j v_j, \quad \alpha_j \in \mathbb{R},$$

since $x \in \mathfrak{N}$. On the other hand,

$$\Lambda_j(x) = \sum_{k=1}^{n} \alpha_k \Lambda(v_k) = \alpha_j = 0$$

by $x \in \mathfrak{M}$. Hence $x = 0$.

For any $x \in \mathfrak{X}$, we define

$$w = \sum_{j=1}^{n} \Lambda_j(x) v_j \in \mathfrak{N}, \quad u = x - w.$$

We have only to show that $u \in \mathfrak{M}$. On the contrary, suppose that $u \notin \mathfrak{M}$. By the Hahn–Banach theorem, there exists $\Lambda \in \mathfrak{M}^\perp$, which satisfies $\Lambda(u) \neq 0$. Λ can be expressed in the form

$$\Lambda = \sum_{j=1}^{n} \beta_j \Lambda_j, \quad \beta_j \in \mathbb{R}.$$

Since

$$\Lambda_j(w) = \sum_{k=1}^{n} \Lambda_k(x) \Lambda_j(v_k) = \sum_{k=1}^{n} \Lambda_k(x) \delta_{jk} = \Lambda_j(x),$$

we obtain $\Lambda(w) = \Lambda(x)$. Hence $\Lambda(u) = 0$, which contradicts $\Lambda(u) \neq 0$. □

Lemma 10.4 *Suppose that $\mathfrak{X}$ is a normed linear space. If $\mathfrak{M}$ is a closed subspace of $\mathfrak{X}$ and $\mathfrak{N}$ is a finite dimensional subspace of $\mathfrak{X}$ such that $\mathfrak{M} \cap \mathfrak{N} = \{0\}$, then $\mathfrak{M} \oplus \mathfrak{N}$ is a closed subspace.*

Proof Let $\{x_n\}$ be a sequence in $\mathfrak{M} \oplus \mathfrak{N}$ which converges to x^*. We have to show that $x^* \in \mathfrak{M} \oplus \mathfrak{N}$. Each x_n can be expressed as

$$x_n = u_n + v_n \ ; \quad u_n \in \mathfrak{M}, v_n \in \mathfrak{N}.$$

Then we observe that the sequence $\{\|v_n\|\}$ is bounded. If it is unbounded, there exists some subsequence $\{v_{n'}\}$ of $\{v_n\}$ such that $\|v_{n'}\| \to \infty$ (as $n' \to \infty$). Define

$$z_{n'} = \frac{v_{n'}}{\|v_{n'}\|}.$$

Then we obtain

$$\left\| \frac{u_{n'}}{\|v_{n'}\|} + z_{n'} \right\| = \frac{1}{\|v_{n'}\|} \|u_{n'} + v_{n'}\| = \frac{1}{\|v_{n'}\|} \|x_{n'}\| \to 0 \quad \text{as} \quad n' \to \infty.$$

However, $\{z_{n'}\}$ is a sequence in the unit sphere $S_{\mathfrak{N}}$ of $\mathfrak{N}$, and $u_{n'}/\|v_{n'}\| \in \mathfrak{M}$. So the above result implies

$$\inf_{z \in S_{\mathfrak{N}}} \rho(z, \mathfrak{M}) = 0,$$

where $\rho(z, \mathfrak{M})$ is the distance between z and $\mathfrak{M}$. Since $S_{\mathfrak{N}}$ is compact by $\dim \mathfrak{N} < \infty$, there exists some $z_0 \in S_{\mathfrak{N}}$ such that

$$\rho(z_0, \mathfrak{M}) = 0.$$

On the other hand, we are assuming that $\mathfrak{M}$ is closed and $\mathfrak{M} \cap \mathfrak{N} = \{0\}$. Hence we obtain

$$\rho(z_0, \mathfrak{M}) > 0$$

since $z_0 \in S_{\mathfrak{N}}$ (and so $z_0 \neq 0$). Contradiction.

Thus we have proved that $\{\|v_n\|\}$ is bounded. Hence there exists a subsequence $\{v_{n'}\}$ (different subsequence from $\{v_{n'}\}$ used above) of $\{v_n\}$ such that

$$v_{n'} \to v^* \in \mathfrak{N}.$$

Since

$$u_{n'} = x_{n'} - v_{n'}$$

and $x_{n'} \to x^*$, $v_{n'} \to v^*$, we obtain

$$u_{n'} \to x^* - v^* \equiv u^* \quad \text{as} \quad n' \to \infty.$$

$u^* \in \mathfrak{M}$ because $\mathfrak{M}$ is closed. We conclude that $x^* \in \mathfrak{M} \oplus \mathfrak{N}$. $\square$

10.2 Fredholm Operators: Definitions and Examples

Definition 10.2 Let $\mathfrak{X}$ and $\mathfrak{Y}$ be normed linear spaces. A bounded linear operator $T : \mathfrak{X} \to \mathfrak{Y}$ is called a **Fredholm operator** if it satisfies

$$\dim \mathrm{Ker} T < \infty \quad \text{and} \quad \mathrm{codim} T(\mathfrak{X}) < \infty.$$

If T is a Fredholm operator,

$$\kappa(T) = \dim \mathrm{Ker} T - \mathrm{codim} T(\mathfrak{X})$$

is called the **index** of T.

Example 10.1 If $\mathfrak{X} = \mathbb{R}^n$ and $\mathfrak{Y} = \mathbb{R}^m$, any linear operator $T : \mathbb{R}^n \to \mathbb{R}^m$ is Fredholm. Since $\mathfrak{X}/\mathrm{Ker} T \cong T(\mathfrak{X})$,

$$\mathrm{codim}\, \mathrm{Ker} T = \dim T(\mathfrak{X}).$$

Consequently, we obtain

$$\kappa(T) = \dim \mathrm{Ker} T - \mathrm{codim} T(\mathfrak{X}) = (n - \mathrm{codim}\, \mathrm{Ker} T) - (m - \dim T(\mathfrak{X})) = n - m.$$

Example 10.2 Let $\mathfrak{X} = \mathfrak{C}^1([a, b], \mathbb{R})$ and $\mathfrak{Y} = \mathfrak{C}([a, b], \mathbb{R})$.[8] A linear operator T is given by

$$T : f \mapsto f', \quad f \in \mathfrak{C}^1.$$

Since $\mathrm{Ker} T$ is the space of constant functions, $\dim \mathrm{Ker} T = 1$. Furthermore, $\mathrm{codim} T(\mathfrak{X}) = 0$, since $T(\mathfrak{X}) = \mathfrak{Y}$. Hence T is Fredholm and $\kappa(T) = 1$.

Example 10.3 Let $\lambda \in \mathbb{R}\backslash\{0\}$ and a continuous function $K(x, y) : [a, b] \times [a, b] \to \mathbb{R}$ be given. Define an operator $T : \mathfrak{L}^2([a, b], \mathbb{R}) \to \mathfrak{L}^2([a, b], \mathbb{R})$ by

$$Tf(x) = \int_a^b K(x, y) f(y) dy - \lambda f(x).$$

($\mathfrak{X} = \mathfrak{Y} = \mathfrak{L}^2([a, b], \mathbb{R})$.) As is well-known, the operator

$$S : f \mapsto \int_a^b K(x, y) f(y) dy$$

[8] The topologies of $\mathfrak{C}^1$ and $\mathfrak{C}$ are defined by the $\mathfrak{C}^1$-norm and the uniform convergence norm, respectively.

is a compact operator on $\mathfrak{L}^2([a, b], \mathbb{R})$. Since $T = S - \lambda I$, $\dim \mathrm{Ker} T = \mathrm{codim} T(\mathfrak{X})$.[9] Hence T is Fredhom and $\kappa(T) = 0$.

Theorem 10.7 *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. If $T : \mathfrak{X} \to \mathfrak{Y}$ is a bounded linear operator and* $\mathrm{codim} T(\mathfrak{X}) < \infty$*, then $T(\mathfrak{X})$ is closed.*

Proof By Theorem 10.2, there exists some subspace $\mathfrak{Z}$ such that[10]

$$\mathfrak{Y} = T(\mathfrak{X}) \oplus \mathfrak{Z}.$$

Since $\dim \mathfrak{Z} < \infty$, $\mathfrak{Z}$ is a closed subspace. Define an operator $S : \mathfrak{X} \times \mathfrak{Z} \to \mathfrak{Y}$ by

$$S(x, z) = Tx + z; \quad x \in \mathfrak{X}, \ z \in \mathfrak{Z}.$$

Then S is a bounded linear operator of the Banach space $\mathfrak{X} \times \mathfrak{Z}$ onto another Banach space $\mathfrak{Y}$. Hence, by the open mapping theorem,[11] S is an open mapping. If we define an operator

$$\tilde{S} : \mathfrak{X} \times \mathfrak{Z}/\mathrm{Ker} S \to \mathfrak{Y}$$

by

$$\tilde{S} : (x, z) + \mathrm{Ker} S \mapsto Tx + z,$$

$\mathfrak{X} \times \mathfrak{Z}/\mathrm{Ker} S$ and $\mathfrak{Y}$ are isomorphic (as Banach spaces) through $\tilde{S}$. $T(\mathfrak{X})$ coincides with the image of the Banach space $\mathfrak{X} \times \{0\}/\mathrm{Ker} S$ by $\tilde{S}$. Thus we can conclude that $T(\mathfrak{X})$ is complete (and so closed). □

Theorem 10.8 *Let $\mathfrak{X}$ be a Banach space, and $K : \mathfrak{X} \to \mathfrak{X}$ a compact operator.*

(i) $\dim \mathrm{Ker}(I + K) < \infty$.
(ii) $\dim \mathrm{Ker}(I + K') < \infty$.[12]
(iii) *The image of $I + K$ is closed.*

---

[9] Let $A : \mathfrak{X} \to \mathfrak{X}$ be a compact operator and $\lambda \neq 0$. All the following four numbers are finite and equal. (A' is the dual operator of A.)

$$\alpha = \dim \mathrm{Ker}(A - \lambda I), \quad \beta = \mathrm{codim}(A - \lambda I)(\mathfrak{X}),$$
$$\alpha' = \dim \mathrm{Ker}(A' - \lambda L), \quad \beta' = \mathrm{codim}(A' - \lambda I)(\mathfrak{X}').$$

[10] We can not have recourse to Theorem 5.1 because the closedness of $T(\mathfrak{X})$ has not been shown yet.

[11] See Maruyama [4] pp. 139–142, Yosida [5] pp. 75–77.

[12] K' is the dual operator of K.

Proof (i) Let S be the unit ball of $\mathrm{Ker}(I+K)$. Since $I=-K$ on $\mathrm{Ker}(I+K)$, we have $S=-K(S)$. So S is relatively compact because K is a compact operator. S being closed, it is compact. Hence $\dim \mathrm{Ker}(I+K)<\infty$.

(ii) Similarly, (ii) follows as in (i), since $(I+K)'=I+K'$ and K' is a compact operator.

(iii) According to (i) and Theorem 10.5, there exists a closed subspace $\mathfrak{M}$ such that $\mathfrak{X}=\mathfrak{M}\oplus\mathrm{Ker}(I+K)$ (topological direct sum). It should be observed that there exists a number $\theta>0$ such that

$$\|m\| \leqq \theta\|(I+K)m\| \quad \text{for all} \quad m\in\mathfrak{M}. \tag{10.2}$$

If there is no such θ, there exists a sequence $\{m_n\}$ of $\mathfrak{M}$, which satisfies[13]

$$m_n\in\mathfrak{M}, \quad \|m_n\|=1, \quad (I+K)m_n\to 0.$$

Since K is a compact operator. $\{Km_n\}$ has a convergent subsequence $\{Km_{n'}\}$. Let y^* be its limit. Then $m_{n'}\to -y^*$, since

$$(I+K)m_{n'}=m_{n'}+Km_{n'}\to 0, \quad Km_{n'}\to y^*.$$

That is, $y^*\in\mathrm{Ker}(I+K)$. $-y^*\in\mathfrak{M}$ because $\mathfrak{M}$ is closed. Hence we obtain

$$y^*\in\mathfrak{M}\cap\mathrm{Ker}(I+K).$$

It follows that $y^*=0$. However, it is a contradiction to

$$m_{n'}\to -y^*, \quad \|m_{n'}\|=1.$$

Finally, we show that $(I+K)(\mathfrak{X})$ is closed. Assume that

$$(I+K)x_n=u_n\to u^*, \quad x_n\in\mathfrak{X}.$$

Since $\mathfrak{X}=\mathfrak{M}\oplus\mathrm{Ker}(I+K)$, we can express u_n as

$$u_n=(I+K)m_n, \quad m_n\in\mathfrak{M}. \tag{10.3}$$

[13]The reason is as follows. For any $\varepsilon>0$, there exists some $v_\varepsilon\in\mathfrak{M}$ such that $\|v_\varepsilon\|>\varepsilon^{-1}\|(I+K)v_\varepsilon\|$. So defining $m_\varepsilon=v_\varepsilon/\|v_\varepsilon\|$, we have

$$\|(I+K)m_\varepsilon\|<\varepsilon.$$

By (10.2) and (10.3), we have

$$\|m_n - m_k\| \leqq \theta \|u_n - u_k\|.$$

Hence $\{m_n\}$ is Cauchy, and so it has a limit m^*. By the continuity of $I + K$,

$$u^* = \lim_{n\to\infty} u_n = \lim_{n\to\infty} (I + K)m_n = (I + K)m^*,$$

which implies $u^* \in (I + K)(\mathfrak{X})$. □

10.3 Parametrix

To start with, we prepare a useful lemma, which will be made use of frequently.

Lemma 10.5 *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces, and $T : \mathfrak{X} \to \mathfrak{Y}$ a Fredholm operator. Assume that $\mathfrak{M} \subset \mathfrak{X}$ and $\mathfrak{Z} \subset \mathfrak{Y}$ are subspaces which satisfy*

$$\mathfrak{X} = \mathfrak{V} \oplus \mathfrak{M}, \ \ \mathfrak{Y} = \mathfrak{Z} \oplus \mathfrak{R} \quad \textit{(topological direct sum)},$$

where $\mathfrak{V} = \mathrm{Ker} T$ and $\mathfrak{R} = T(\mathfrak{X})$. Then the restriction $T|_{\mathfrak{M}}$ of T to $\mathfrak{M}$ is an isomorphism between two normed spaces, $\mathfrak{M}$ and $\mathfrak{R}$.

Proof It is obvious that $T|_{\mathfrak{M}} : \mathfrak{M} \to \mathfrak{R}$ is linear. Any $x \in \mathfrak{X}$ is uniquely expressed as $x = v + m$ ($v \in \mathfrak{V}, m \in \mathfrak{M}$). Hence $Tx = Tv + Tm = 0 + Tm$. Thus we obtain $\mathfrak{R} = T|_{\mathfrak{M}}(\mathfrak{M})$. $T|_{\mathfrak{M}}$ is injective, since $\{m \in \mathfrak{M} \mid T|_{\mathfrak{M}} m = 0\} = \mathfrak{V} \cap \mathfrak{M} = \{0\}$. This proves that $T|_{\mathfrak{M}}$ is a bijection. Therefore it has the inverse $T|_{\mathfrak{M}}^{-1} : \mathfrak{R} \to \mathfrak{M}$. $\mathfrak{M}$ and $\mathfrak{R}$ are closed by Theorem 10.4. Hence, by the open mapping theorem,[14] $T|_{\mathfrak{M}}$ is an open mapping. We conclude that $T|_{\mathfrak{M}}$ is an isomorphism between $\mathfrak{M}$ and $\mathfrak{R}$. □

Theorem 10.9 (parametrix) *Let $\mathfrak{X}, \mathfrak{Y}$ be Banach spaces and $T : \mathfrak{X} \to \mathfrak{Y}$ a bounded linear operator. Then the following statements are equivalent:*

(i) *T is a Fredholm operator.*
(ii) *There exist bounded linear operators*

$$P_l, P_r : \mathfrak{Y} \to \mathfrak{X}$$

as well as compact linear operators

$$C_l : \mathfrak{X} \to \mathfrak{X}, \quad C_r : \mathfrak{Y} \to \mathfrak{Y}$$

[14] See footnote 11, p. 291.

such that

$$P_l T = I + C_l, \tag{10.4}$$

$$T P_r = I + C_r. \tag{10.5}$$

Proof (i)⇒(ii): If T is Fredholm, there exist subspaces $\mathfrak{M} \subset \mathfrak{X}$ and $\mathfrak{Z} \subset \mathfrak{Y}$ such that

$$\mathfrak{X} = \mathfrak{V} \oplus \mathfrak{W}, \quad \mathfrak{Y} = \mathfrak{Z} \oplus \mathfrak{R} \quad \text{(topological direct sum)},$$

where $\mathfrak{V} \equiv \mathrm{Ker} T$ and $\mathfrak{R} \equiv T(\mathfrak{X})$. (Note that $\mathfrak{R}$ is closed by Theorem 10.7.) Let

$$\pi_{\mathfrak{V}} : \mathfrak{X} \to \mathfrak{V}, \quad \pi_{\mathfrak{Z}} : \mathfrak{Y} \to \mathfrak{Z}$$

be canonical projections corresponding to the above direct sums. The restriction $T|_{\mathfrak{M}} : \mathfrak{M} \to \mathfrak{R}$ of T to $\mathfrak{M}$ is an isomorphism by Lemma 10.5. Let $S : \mathfrak{Y} \to \mathfrak{M}$ be an operator of $\mathfrak{Y}$ into $\mathfrak{M}$, which maps each element $y = z + r$ $(z \in \mathfrak{Z}, r \in \mathfrak{R})$ of $\mathfrak{Y}$ to $T|_{\mathfrak{M}}^{-1}(r)$; i.e.

$$S : (z + r) \mapsto T|_{\mathfrak{M}}^{-1}(r).$$

Then $ST : \mathfrak{X} \to \mathfrak{X}$ and $TS : \mathfrak{Y} \to \mathfrak{Y}$ can be expressed as

$$ST = I - \pi_{\mathfrak{V}},$$

$$TS = I - \pi_{\mathfrak{Z}}.$$

Since T is Fredholm,

$$\dim \pi_{\mathfrak{V}}(\mathfrak{X}) < \infty, \quad \dim \pi_{\mathfrak{Z}}(\mathfrak{Y}) < \infty.$$

Hence both $\pi_{\mathfrak{V}}$ and $\pi_{\mathfrak{Z}}$ are compact operators. Defining

$$P_l = P_r = S,$$

$$C_l = -\pi_{\mathfrak{V}}, \quad C_r = -\pi_{\mathfrak{Z}},$$

we obtain (ii).

(ii)⇒(i): We next assume (ii). By (10.4),

$$\mathrm{Ker} T \subset \mathrm{Ker}(I + C_l).$$

Theorem 10.7 tells us that $\dim \mathrm{Ker}(I + C_l) < \infty$. Hence

$$\dim \mathrm{Ker} T \leqq \dim \mathrm{Ker}(I + C_l) < \infty.$$

On the other hand, we have

$$(I + C_r)(\mathfrak{Y}) \subset T(\mathfrak{X}) = \mathfrak{R}$$

by (10.5). $(I + C_r)(\mathfrak{Y})$ is closed (by Theorem 10.7) and $\operatorname{codim}(I + C_r)(\mathfrak{Y}) < \infty$ (see footnote 9 on p. 291). We obtain $\operatorname{codim}\mathfrak{R} < \infty$ by Theorem 10.6.

We conclude that T is Fredholm. □

The operators P_l and P_r, the existence of which has just been confirmed by the above lemma, are called the **left parametrix** and the **right parametrix**, respectively.

Remark 10.1

1° The operator $S : \mathfrak{Y} \to \mathfrak{X}$ obtained above is Fredholm and $\kappa(S) = -\kappa(T)$. (It is clear that $\operatorname{Ker} S = \mathfrak{Z}$ and $S(\mathfrak{Y}) = \mathfrak{M}$.)

2° As is observed from the proof above, (i)⇒(ii) can be reformulated in a more exact form.

"Let T be a Fredholm operator. Then there exist a bounded linear operator $S : \mathfrak{Y} \to \mathfrak{X}$ and a couple of compact operators $C_l : \mathfrak{X} \to \mathfrak{X}$ and $C_r : \mathfrak{Y} \to \mathfrak{Y}$ which satisfy

$$ST = I + C_l, \tag{10.4'}$$

$$TS = I + C_r. \tag{10.5'}$$

Furthermore:

(a) S is Fredholm,
(b) $\dim C_l(\mathfrak{X}) < \infty, \quad \dim C_r(\mathfrak{Y}) < \infty$,
(c) $\kappa(S) = -\kappa(T)$."

10.4 Product of Fredholm Operators

Theorem 10.10 *Let $\mathfrak{X}, \mathfrak{Y}$ and $\mathfrak{Z}$ be Banach spaces, and let $T : \mathfrak{X} \to \mathfrak{Y}$ and $S : \mathfrak{Y} \to \mathfrak{Z}$ be Fredholm operators. Then the product $ST = S \circ T : \mathfrak{X} \to \mathfrak{Z}$ is also Fredholm and*

$$\kappa(ST) = \kappa(S) + \kappa(T).$$

Proof Since T and S are Fredholm, $\dim \operatorname{Ker} T < \infty$ and $\operatorname{codim} S(\mathfrak{Y}) < \infty$, which implies that $S(\mathfrak{Y})$ is closed (by Theorem 10.7). Then there exist subspaces $\mathfrak{W} \subset \mathfrak{X}$ and $\mathfrak{U} \subset \mathfrak{Z}$ such that

$$\mathfrak{X} = \mathrm{Ker} T \oplus \mathfrak{W}, \tag{10.6}$$

$$\mathfrak{Z} = \mathfrak{U} \oplus S(\mathfrak{Y}) \quad \text{(topological direct sum)} \tag{10.7}$$

by Theorem 10.5. $\mathfrak{Y}$ can be expressed as a direct sum in terms either of T or S. If we define

$$\mathfrak{Y}_1 = T(\mathfrak{X}) \cap \mathrm{Ker} S, \tag{10.8}$$

$\dim \mathfrak{Y}_1 < \infty$. Since $T(\mathfrak{X})$ and $\mathrm{Ker} S$ are closed subspaces, they can be expressed as

$$T(\mathfrak{X}) = \mathfrak{Y}_1 \oplus \mathfrak{Y}_2, \tag{10.9}$$

$$\mathrm{Ker} S = \mathfrak{Y}_1 \oplus \mathfrak{Y}_3 \tag{10.10}$$

(topological direct sums) by Theorem 10.5. Suppose that $y \in T(\mathfrak{X}) \cap \mathfrak{Y}_3$. Then $y \in T(\mathfrak{X}) \cap \mathrm{Ker} S = \mathfrak{Y}_1$ by (10.8). By $\mathfrak{Y}_1 \cap \mathfrak{Y}_3 = \{0\}$, we obtain $y = 0$. That is, $T(\mathfrak{X}) \cap \mathfrak{Y}_3 = \{0\}$. Hence we can consider $T(\mathfrak{X}) \oplus \mathfrak{Y}_3$. Since $T(\mathfrak{X})$ is closed and $\mathrm{codim} T(\mathfrak{X}) < \infty$, $T(\mathfrak{X} \oplus \mathfrak{Y}_3)$ is closed and $\mathrm{codim}(T(\mathfrak{X}) \oplus \mathfrak{Y}_3) < \infty$ by Theorem 10.6. Applying Theorem 10.5 repeatedly, we obtain

$$\mathfrak{Y} = (T(\mathfrak{X}) \oplus \mathfrak{Y}_3) \oplus \mathfrak{Y}_4 = T(\mathfrak{X}) \oplus (\mathfrak{Y}_3 \oplus \mathfrak{Y}_4) \tag{10.11}$$

for some subspace $\mathfrak{Y}_4$. By (10.9) and (10.10),

$$\mathfrak{Y} = \mathfrak{Y}_1 \oplus \mathfrak{Y}_2 \oplus \mathfrak{Y}_3 \oplus \mathfrak{Y}_4 = \mathfrak{Y}_2 \oplus \mathfrak{Y}_4 \oplus \mathrm{Ker} S. \tag{10.12}$$

Since

$$\mathrm{codim} T(\mathfrak{X}) = \dim \mathfrak{Y}_3 + \dim \mathfrak{Y}_4$$

by (10.11), we obtain

$$\kappa(T) = \dim \mathrm{Ker} T - (\dim \mathfrak{Y}_3 + \dim \mathfrak{Y}_4). \tag{10.13}$$

We also have

$$\dim \mathrm{Ker} S = \dim \mathfrak{Y}_1 + \dim \mathfrak{Y}_3$$

by (10.10). Hence

$$\kappa(S) = \dim \mathfrak{Y}_1 + \dim \mathfrak{Y}_3 - \dim \mathfrak{U} \tag{10.14}$$

by (10.7).

Noting that $T|_{\mathfrak{W}} : \mathfrak{W} \to T(\mathfrak{X})$ is an isomorphism by Lemma 10.5, we have, from (10.9), that

$$\mathfrak{W} = T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_1) \oplus T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_2). \tag{10.15}$$

Combining this with (10.6), we have

$$\mathfrak{X} = \mathrm{Ker}T \oplus T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_1) \oplus T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_2). \tag{10.16}$$

It is clear that

$$\mathrm{Ker}T \oplus T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_1) \subset \mathrm{Ker}(ST). \tag{10.17}$$

ST is injective on $T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_2)$. In fact, S is injective on $\mathfrak{Y}_2$, since $T(\mathfrak{X}) = (T(\mathfrak{X}) \cap \mathrm{Ker}S) \oplus \mathfrak{Y}_2$ by (10.8) and (10.9). And T maps $T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_2)$ to $\mathfrak{Y}_2$ one-to-one (Lemma 10.5). Consequently,

$$\mathrm{Ker}(ST) = \mathrm{Ker}T \oplus T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_1), \tag{10.18}$$

$$ST(\mathfrak{X}) = ST(T|_{\mathfrak{W}}^{-1}(\mathfrak{Y}_2)) = S\mathfrak{Y}_2. \tag{10.19}$$

It follows from (10.18) that

$$\dim \mathrm{Ker}(ST) = \dim \mathrm{Ker}T + \dim \mathfrak{Y}_1. \tag{10.20}$$

According to (10.12), S maps $\mathfrak{Y}_2 \oplus \mathfrak{Y}_4$ onto $S(\mathfrak{Y})$ one-to-one (Lemma 10.5). Since $\mathfrak{Y}_2$ is a closed subspace and $\dim \mathfrak{Y}_4 < \infty$, $\mathfrak{Y}_2 \oplus \mathfrak{Y}_4$ is a closed subspace by Lemma 10.4. Applying Lemma 10.5 again, we verify that $S|_{\mathfrak{Y}_2 \oplus \mathfrak{Y}_4}$ is an isomorphism between $\mathfrak{Y}_2 \oplus \mathfrak{Y}_4$ and $S(\mathfrak{Y})$. Hence $ST(\mathfrak{X})$ is closed by (10.19).

Since

$$\mathfrak{Z} = \mathfrak{U} \oplus S(\mathfrak{Y}) = \mathfrak{U} \oplus S(\mathfrak{Y}_2) \oplus S(\mathfrak{Y}_4),$$

we obtain

$$\mathrm{codim} ST(\mathfrak{X}) = \dim \mathfrak{U} + \dim \mathfrak{Y}_4 \tag{10.21}$$

by (10.19). By (10.20) and (10.21),

$$\kappa(ST) = \dim \mathrm{Ker}T + \dim \mathfrak{Y}_1 - (\dim \mathfrak{U} + \dim \mathfrak{Y}_4).$$

Taking account of (10.13) and (10.14), we conclude

$$\kappa(ST) = \kappa(S) + \kappa(T).$$

□

10.5 Stability of Indices

Theorem 10.11 *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces, and $T : \mathfrak{X} \to \mathfrak{Y}$ a Fredholm operator. Then there exists some $\varepsilon > 0$ such that $\kappa(U) = \kappa(T)$ for any Fredholm operator $U : \mathfrak{X} \to \mathfrak{Y}$ with $\|U - T\| < \varepsilon$.*

Proof By Remark 10.1, there exist a bounded linear operator $S : \mathfrak{Y} \to \mathfrak{X}$ and a pair of compact operators, $C_l : \mathfrak{X} \to \mathfrak{X}$ and $C_r : \mathfrak{Y} \to \mathfrak{Y}$ which satisfy (10.4′), (10.5′) as well as (a), (b) and (c) there. Let $\varepsilon = \|S\|^{-1}$.[15]

If $V : \mathfrak{X} \to \mathfrak{Y}$ is a bounded linear operator with $\|V\| < \varepsilon$, then $SV : \mathfrak{X} \to \mathfrak{X}$ satisfies $\|SV\| \leqq \|S\| \cdot \|V\| < 1$ and $VS : \mathfrak{Y} \to \mathfrak{Y}$ also satisfies $\|VS\| < 1$. Hence $I + SV$ and $I + VS$ have bounded inverses $(I + SV)^{-1}$ and $(I + VS)^{-1}$.[16]

If we define $W = T + V$, (10.4′) and (10.5′) become

$$SW = I + SV + C_l,$$
$$WS = I + VS + C_r.$$

Multiplying $(I + SV)^{-1}$ by the left of the first formula, and multiplying $(I + VS)^{-1}$ by the right of the second, we obtain

$$(I + SV)^{-1}SW = I + (I + SV)^{-1}C_l,$$
$$WS(I + VS)^{-1} = I + C_r(I + VS)^{-1}.$$

Writing $(I + SV)^{-1}S = P_l$, $S(I + VS)^{-1} = P_r$, $(I + SV)^{-1}C_l = C_l'$ and $C_r(I + VS)^{-1} = C_r'$, we can rewrite the above relations as

$$P_l W = I + C_l', \quad W Pr = I + C_r'.$$

Since $\dim C_l'(\mathfrak{X}) < \infty$ and $\dim C_r'(\mathfrak{Y}) < \infty$ (Remark 10.1, 2°), C_l' and C_r' are compact operators. Hence, by Theorem 10.9, W is a Fredholm operator.

We next evaluate the index. Since $\kappa(S) = -\kappa(T)$ and $\kappa((I + SV)^{-1}) = 0$, we obtain

$$\kappa(P_l W) = \kappa((I + SV)^{-1}SW) = \kappa((I + SV)^{-1}) + \kappa(S) + \kappa(W) = -\kappa(T) + \kappa(W)$$

by Theorem 10.10. On the other hand, $\dim C_l'(\mathfrak{X}) < \infty$ implies $\kappa(I + C_l') = 0$. Therefore $\kappa(E) = \kappa(T)$. □

[15] In the case of $\|S\| = 0$, we may consider $\varepsilon = \infty$. But such a case occurs when $\dim \mathfrak{X} < \infty$ and $\dim \mathfrak{Y} < \infty$.

[16] Maruyama [4] pp. 153–154.

Theorem 10.12 *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. If $T : \mathfrak{X} \to \mathfrak{Y}$ is a Fredholm operator and $K : \mathfrak{X} \to \mathfrak{Y}$ is a compact operator, then $T + K$ is Fredholm and $\kappa(T + K) = \kappa(T)$.*

Proof There exist parametrices $P_l, P_r : \mathfrak{Y} \to \mathfrak{X}$ and compact operators $C_l : \mathfrak{X} \to \mathfrak{X}$, $C_r : \mathfrak{Y} \to \mathfrak{Y}$ such that

$$P_l T = I + C_l, \quad T P_r = I + C_r.$$

Consequently,

$$P_l(T + K) = I + C_l + P_l K,$$
$$(T + K)P_r = I + C_r + K P_r.$$

Since K is a compact operator, so are $P_l K$ and $K P_r$. So we observe that P_l and P_r are parametrices of $T + K$. That is, $T + K$ is Fredholm.

Define a continuous mapping $\varphi : [0, 1] \to \mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ (space of bounded linear operators) by

$$\varphi(t) = T + tK.$$

Since tK is a compact operator, $T + tK$ is Fredholm for every $t \in [0, 1]$. $\kappa(T + tK)$ is locally constant on $[0, 1]$ by Theorem 10.11. Hence, by connectedness of $[0, 1]$,

$$\kappa(T + tK) = \text{constant} \quad \text{for all} \quad t \in [0, 1].$$

□

References

1. Dunford, N., Schwartz, J.T.: Linear Operators, Part 1. Interscience, New York (1958)
2. Kato, T.: Iso Kaiseki (Functional Analysis). Kyoritsu Shuppan, Tokyo (1957) (Originally published in Japanese)
3. Kuroda, N.: Kansu Kaiseki (Functional Analysis). Kyoritsu Shuppan, Tokyo (1980) (Originally published in Japanese)
4. Maruyama, T.: Suri-keizaigaku no Hoho (Methods in Mathematical Economics). Sobunsha, Tokyo (1995) (Originally published in Japanese)
5. Yosida, K.: Functional Analysis. 3rd edn. Springer, New York (1971)
6. Zeidler, E.: Applied Functional Analysis. Springer, New York (1995)

Chapter 11
Hopf Bifurcation Theorem

The Hopf bifurcation theorem provides an effective criterion for finding periodic solutions for ordinary differential equations. Although various proofs of this classical theorem are known, there seems to be no easy way to arrive at the goal. Among them, the idea of Ambrosetti and Prodi [1] is particularly noteworthy.

They start by formulating the Hopf theorem in an abstract fashion and then try to deduce the classical result from it. Let $F(\omega, \mu, \cdot)$ be a smooth function of a Banach space $\mathfrak{X}$ into another one $\mathfrak{Y}$ with a pair (ω, μ) of real parameters. In order to find some bifurcation point (ω^*, μ^*) of the equation $F(\omega, \mu, x) = 0$, the derivative $D_x F(\omega^*, \mu^*, 0)$ of F with respect to $x \in \mathfrak{X}$ at $(\omega^*, \mu^*, 0)$ plays a crucial role. As will be stated in Theorem 11.1 exactly, the condition that both the dimension of the kernel of $D_x F(\omega^*, \mu^*, 0)$ and the codimension of the image of $D_x F(\omega^*, \mu^*, 0)$ are 2, together with another condition, assures that there occurs a bifurcation phenomenon at (ω^*, μ^*).

Based upon this result, Ambrosetti and Prodi successfully paved the way to deducing the classical Hopf theorem and elucidate its mathematical structure.[1]

Let $f(\mu, x)$ be a given function of $\mathbb{R} \times \mathbb{R}^n$ into $\mathbb{R}^n$. We consider the ordinary differential equation

$$F(\omega, \mu, x(\cdot)) \equiv \omega \frac{dx(\cdot)}{dt} - f(\mu, x(\cdot)) = 0 \tag{11.1}$$

with a pair (ω, μ) of parameters. Here Ambrosetti and Prodi adopt some suitable function space $\mathfrak{C}^r$(resp. $\mathfrak{C}^{r-1}$) as $\mathfrak{X}$ (resp. $\mathfrak{Y}$). In order to apply the abstract theorem mentioned above to this concrete problem, we have to calculate the dimension of the kernel of the operator

[1]See Hopf [8] for the original formulation. Crandall and Rabinowitz [4] is extremely suggestive.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8_11

$$D_x F(\omega^*, \mu^*, 0) : x \mapsto \omega^* \frac{dx}{dt} - D_x f(\mu^*, 0)x \tag{11.2}$$

of $\mathfrak{X}$ into $\mathfrak{Y}$ and the codimension of its image. Ambrosetti and Prodi's idea to overcome this problem is very simple in a sense. They first expand the function $x(t)$ in the uniformly convergent Fourier series

$$x(t) = \sum_{k=-\infty}^{\infty} u_k e^{ikt}. \tag{11.3}$$

If the derivative $dx/dt = \dot{x}(\cdot)$ can be expanded in the form

$$\dot{x}(t) = \sum_{k=-\infty}^{\infty} iku_k e^{ikt}, \tag{11.4}$$

we obtain

$$D_x F(\omega^*, \mu^*, 0)x = \sum_{k=-\infty}^{\infty} [ik\omega^* I - D_x f(\mu^*, 0)]u_k e^{ikt} \tag{11.5}$$

by substituting (11.3) and (11.4) into (11.2), where I is the $(n \times n)$-identity matrix. Hence the kernel of $D_x F(\omega^*, \mu^*, 0)$ consists of all the $x(\cdot)$ such that $[ik\omega^* I - D_x f(\mu^*, 0)]u_k = 0$ for all $k \in \mathbb{Z}$ (the set of all the integers). And when we wish to determine the image of $D_x F(\omega^*, \mu^*, 0)$, we have only to examine the equation

$$[ik\omega^* I - D_x f(\mu^*, 0)]u_k = v_k \quad \text{for all} \quad k \in \mathbb{Z}, \tag{11.6}$$

where the v_k's are the Fourier coefficients of $y \in \mathfrak{Y}$. The image of $D_x F(\omega^*, \mu^*, 0)$ is the set of all the elements y of $\mathfrak{Y}$, for which the equation (11.6) is solvable.

The main purpose of the present chapter is to establish the Hopf theorem in the framework of some Sobolev space instead of $\mathfrak{C}^r$. This approach seems to enable us to simplify the technical details in the course of the proof to some extent. The basic result (cf. p. 37) due to Carleson [2] and Hunt [9] plays a crucial role in our theory.

Incidentally, we have to note that we require the condition that $x(\cdot)$ is of the class $\mathfrak{C}^r$, $r \geqq 3$ and $y(\cdot)$ is of the class $\mathfrak{C}^{r-1}$ in order to justify the expression (11.4) if we remain in the space $\mathfrak{C}^r$. As will be shown later, we have, for any $\varepsilon > 0$, that

$$\begin{aligned}\left\| \sum_{k=-\infty}^{\infty} iku_k e^{ikt} \right\| &\leqq \sum_{k=-\infty}^{\infty} \|ku_k\| \\ &\leqq \sum_{|k|<N} \|ku_k\| + \varepsilon \sum_{|k| \geqq N} \frac{1}{|k|^{r-1}}\end{aligned}$$

for sufficiently large N. Hence if we assume $r \geqq 3$, the left-hand side is uniformly convergent and the expression (11.4) is valid. Thus although Ambrosetti and Prodi assume only $r = 1$, we have to impose some more restrictions on the smoothness of $x(\cdot)$ and $y(\cdot)$, and some more subtle and careful account of the magnitudes of the Fourier coefficients seems to be required.

The other object of this chapter is to fortify these analytical details and to complete the Ambrosetti and Prodi theory from the standpoint of the classical Fourier analysis.

Furthermore, we apply this theory to laying a solid mathematical foundation for N. Kaldor's theory of business fluctuations.[2]

11.1 Ljapunov–Schmidt Reduction Method

Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. Consider an equation

$$f(\lambda, x) = 0$$

defined by a function $f \in \mathfrak{C}^p(\mathbb{R} \times \mathfrak{X}, \mathfrak{Y})$ $(p \geqq 1)$. Although this equation is defined in the framework of infinite dimensional spaces in general, it can be solved by reducing it to the case of finite variables and finite equations if certain conditions are satisfied. This is the so-called Ljapunov–Schmidt reduction method.

Suppose that a function $f \in \mathfrak{C}^p(\mathbb{R} \times \mathfrak{X}, \mathfrak{Y})$ $(p \geqq 1)$ satisfies the following four conditions. We use several notations:

$$T = D_x f(\lambda^*, 0), \quad \mathfrak{V} = \mathrm{Ker} T, \quad \mathfrak{R} = T(\mathfrak{X})$$

where $\lambda^* \in \mathbb{R}$.

(i) $f(\lambda^*, 0) = 0$.
(ii) T is not invertible.
(iii) $\mathfrak{V}$ has a topological complement in $\mathfrak{X}$:

$$\mathfrak{X} = \mathfrak{V} \oplus \mathfrak{W} \quad \text{(topological direct sum)}.$$

(iv) $\mathfrak{R}$ is closed, and has a topological complement $\mathfrak{Z}$:

$$\mathfrak{Y} = \mathfrak{Z} \oplus \mathfrak{R} \quad \text{(topological direct sum)}.$$

Remark 11.1 If T is Fredholm, (iii) and (iv) are satisfied. (cf. Theorem 10.4 on p. 282, Theorem 10.7 on p. 291.)

[2]This chapter is based upon Ambrosetti and Prodi [1] and Maruyama [13, 14]. Masuda [16] Part I, Chap. 6 is quite helpful as a good exposition of bifurcation theory in general.

Let $P : \mathfrak{Y} \to \mathfrak{Z}$ and $Q : \mathfrak{Y} \to \mathfrak{R}$ be the projections of $\mathfrak{Y}$ into $\mathfrak{Z}$ and $\mathfrak{R}$, respectively. If we express each $x \in \mathfrak{X}$ as

$$x = v + w; \quad v \in \mathfrak{V}, \ w \in \mathfrak{W},$$

$f(\lambda, x) = 0$ is equivalent to

$$Pf(\lambda, v + w) = 0, \tag{11.7a}$$

$$Qf(\lambda, v + w) = 0. \tag{11.7b}$$

Defining

$$\varphi(\lambda, x) = f(\lambda, x) - Tx,$$

we obtain

$$f(\lambda, x) = Tw + \varphi(\lambda, v + w),$$

since $Tv = 0$, where $x = v + w$ $(v \in \mathfrak{V}, w \in \mathfrak{W})$.[3] The equality (11.7b) can be rewritten as

$$Qf(\lambda, v + w) = QTw + Q\varphi(\lambda, v + w) = Tw + Q\varphi(\lambda, v + w) = 0, \tag{11.8}$$

since $Tw \in \mathfrak{R}$. Furthermore, if we define

$$\Phi(\lambda, v, w) = Tw + Q\varphi(\lambda, v + w),$$

then $\Phi \in \mathfrak{C}^p(\mathbb{R} \times \mathfrak{V} \times \mathfrak{W}, \mathfrak{R})$ and $D_w\Phi(\lambda^*, 0, 0) : \mathfrak{W} \to \mathfrak{R}$ is given by

$$D_w\Phi(\lambda^*, 0, 0) : w \mapsto Tw + QD_x\varphi(\lambda^*, 0)w.$$

Since $\varphi(\lambda, x) = f(\lambda, x) - Tx$ by definition, it follows that

$$D_x\varphi(\lambda^*, 0) = D_x f(\lambda^*, 0) - T = 0.$$

Hence

$$D_w\Phi(\lambda^*, 0, 0) = T|_{\mathfrak{W}}.$$

We should note that $T|_{\mathfrak{W}} : \mathfrak{W} \to \mathfrak{R}$ is an isomorphism (cf. Lemma 10.5, p 293).

[3] In the case of $f(\lambda^*, 0)$, in particular, $f(\lambda^*, x) = f(\lambda^*, x) - f(\lambda^*, 0) = D_x f(\lambda^*, 0)x + \varphi(\lambda^*, x)$. Thus, $\varphi(\lambda^*, x)$ is a residue term of the linear approximation of $f(\lambda^*, x)$.

We now apply the implicit function theorem[4] to the equation (11.7b) ⇔ (11.8):

$$\Phi(\lambda, v, w) = Tw + Q\varphi(\lambda, v + w) = 0.$$

Of course, $\Phi(\lambda^*, 0, 0) = 0$. There exist a neighborhood Γ_{λ^*} of λ^*, a neighborhood $\Gamma_{0,\mathfrak{V}}$ of $0 \in \mathfrak{V}$, a neighborhood $\Gamma_{0,\mathfrak{W}}$ of $0 \in \mathfrak{W}$ and a unique function $\gamma \in \mathfrak{C}^p(\Gamma_{\lambda^*} \times \Gamma_{0,\mathfrak{V}}, \Gamma_{0,\mathfrak{W}})$ such that:

a. $\Phi(\lambda, v, \gamma(\lambda, v)) = 0 \quad \text{for all} \quad (\lambda, v) \in \Gamma_{\lambda^*} \times \Gamma_{0,\mathfrak{V}}$,
b. $\gamma(\lambda^*, 0) = 0$.

Remark 11.2

1° If $f(\lambda, 0) = 0$ for all $\lambda \in \mathbb{R}$,

$$\gamma(\lambda, 0) = 0 \quad \text{for all} \quad \lambda \in \Gamma_{\lambda^*}.$$

2° $D_v\gamma(\lambda^*, 0) = 0$.

Substituting $w = \gamma(\lambda, v)$ to (11.7a), we obtain

$$P(f(\lambda, v + \gamma(\lambda, v))) = 0 \quad \text{for all} \quad (\lambda, v) \in \Gamma_{\lambda^*} \times \Gamma_{0,\mathfrak{V}}. \tag{11.9}$$

This is called the **bifurcation equation**. Combining with $w = \gamma(\lambda, v)$, it is locally (that is, in $\Gamma_{\lambda^*} \times \Gamma_{0,\mathfrak{V}} \times \Gamma_{0,\mathfrak{W}}$) equivalent to $f(\lambda, x) = 0$. When $D_x f(\lambda^*, 0) = T$ is Fredholm in particular, the number of variables and the number of equations are finite. So the problem is reduced to the finite dimensional case.

11.2 Abstract Hopf Bifurcation Theorem

Let $\mathfrak{X}$ and $\mathfrak{Y}$ be a pair of real Banach spaces. $F(\omega, \mu, x)$ is assumed to be a function of the class $\mathfrak{C}^2(\mathbb{R}^2 \times \mathfrak{X}, \mathfrak{Y})$ which satisfies

$$F(\omega, \mu, 0) = 0 \quad \text{for all} \quad (\omega, \mu) \in \mathbb{R}^2.$$

A point $(\omega^*, \mu^*) \in \mathbb{R}^2$ is called a *bifurcation point* of F if $(\omega^*, \mu^*, 0)$ is in the closure of the set

$$S = \{(\omega, \mu, x) \in \mathbb{R}^2 \times \mathfrak{X} \mid x \neq 0, F(\omega, \mu, x) = 0\}.$$

[4] See Maruyama [12], pp. 281–283 for the implicit function theorem in infinite dimensional spaces.

We shall use several notations for the sake of simplicity:

$$T = D_x F(\omega^*, \mu^*, 0),$$
$$\mathfrak{V} = \operatorname{Ker} T, \quad \mathfrak{R} = T(\mathfrak{X}),$$
$$M = D^2_{x,\mu} F(\omega^*, \mu^*, 0),$$
$$N = D^2_{x,\omega} F(\omega^*, \mu^*, 0).$$

T is the derivative of F with respect to x at $(\omega^*, \mu^*, 0)$. It is a bounded linear operator of $\mathfrak{X}$ into $\mathfrak{Y}$. $\mathfrak{V}$ and $\mathfrak{R}$ are the kernel and the image of T, respectively. M (resp. N) is the second derivative of F with respect to (x, μ) (resp. (x, ω)) at $(\omega^*, \mu^*, 0)$. Since $D^2_{x,\mu} F$ is the bounded linear operator of $\mathbb{R}$ into $\mathfrak{L}(\mathfrak{X}, \mathfrak{Y})$ (the Banach space of all the bounded linear operators of $\mathfrak{X}$ into $\mathfrak{Y}$), it can be identified with an element of $\mathfrak{L}(\mathfrak{X}, \mathfrak{Y})$. The same is true for $D^2_{x,\omega} F$.

The following theorem is an abstract version of the Hopf bifurcation theorem due to Ambrosetti and Prodi [1] (pp. 136–139).

Theorem 11.1 *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be real Banach spaces. Assume that $F(\omega, \mu, x)$ is a function of the class $\mathfrak{C}^2(\mathbb{R}^2 \times \mathfrak{X}, \mathfrak{Y})$ which satisfies the following two conditions:*

1. $\dim \mathfrak{V} = 2$. $\mathfrak{R}$ *is closed and* $\operatorname{codim} \mathfrak{R} = 2$.
   We represent $\mathfrak{X}$ and $\mathfrak{Y}$ in the forms of topological direct sums:

$$\mathfrak{X} = \mathfrak{V} \oplus \mathfrak{W}, \quad \mathfrak{Y} = \mathfrak{Z} \oplus \mathfrak{R}.$$

   We denote by P the projection of $\mathfrak{Y}$ into $\mathfrak{Z}$, and by Q the projection of $\mathfrak{Y}$ into $\mathfrak{R}$.
2. *There exists some point $v* \in \mathfrak{V}$ such that PMv^* and PNv^* are linearly independent.*[5]
   Then (ω^*, μ^*) is a bifurcation point of F.

The condition that $\dim\mathfrak{V} = 2$ and $\operatorname{codim}\mathfrak{R} = 2$ implies that $T : \mathfrak{X} \to \mathfrak{Y}$ is a Fredholm operator with index zero.

The proof of this theorem is based upon the Ljapunov–Schmidt reduction method, which is explained in the preceding section.

Proof Expressing

$$x = v + w\ ; \quad v \in \mathfrak{V}, \quad w \in \mathfrak{W},$$

we can rewrite $F(\omega, \mu, x)$ in the form

$$F(\omega, \mu, x) = Tx + \varphi(\omega, \mu, x) = Tw + \varphi(\omega, \mu, v + w). \tag{11.10}$$

[5]Let $\mathfrak{L}(\mathfrak{X}, \mathfrak{Y})$ be the space of bounded linear operators of $\mathfrak{X}$ into $\mathfrak{Y}$. M and N are bounded linear operators of $\mathbb{R}$ into $\mathfrak{L}(\mathfrak{X}, \mathfrak{Y})$. Hence each of M and N can be regarded as a point of $\mathfrak{L}(\mathfrak{X}, \mathfrak{Y})$. Mv^* and Nv^* are points of $\mathfrak{Y}$. Thus we can let P act on them.

This is the definition of the function $\varphi : \mathbb{R}^2 \times \mathfrak{V} \times \mathfrak{W} \to \mathfrak{Y}$. $QF(\omega, \mu, x) = 0$ can also be rewritten as

$$\begin{aligned} QF(\omega, \mu, v + w) &= QTw + Q\varphi(\omega, \mu, v + w) \\ &= Tw + Q\varphi(\omega, \mu, v + w) \\ &= 0. \end{aligned}$$

If we define

$$\Phi(\omega, \mu, v, w) = Tw + Q\varphi(\omega, \mu, v + w),$$

then $\Phi \in \mathfrak{C}^2(\mathbb{R}^2 \times \mathfrak{V} \times \mathfrak{W}, \mathfrak{Y})$ and

$$D_w\Phi(\omega^*, \mu^*, 0, 0)w = Tw + QD_x\varphi(\omega^*, \mu^*, 0)w. \tag{11.11}$$

By the definition (11.10) of φ,

$$D_x\varphi(\omega^*, \mu^*, 0) = D_xF(\omega^*, \mu^*, 0) - T = 0. \tag{11.12}$$

The equations (11.11) and (11.12) imply

$$D_w\Phi(\omega^*, \mu^*, 0, 0) = T|_{\mathfrak{W}},$$

which gives an isomorphism between $\mathfrak{W}$ and $\mathfrak{R}$.[6]

Being based upon the above observation, we apply the implicit function theorem to the equation $\Phi = 0$. Then there exist a neighborhood G of $(\omega^*, \mu^*, 0)$ and a function $\gamma \in \mathfrak{C}^2(G, \mathfrak{W})$ such that

$$\begin{aligned} &\Phi(\omega, \mu, v, \gamma(\omega, \mu, v)) = 0 \quad \text{for all} \quad (\omega, \mu, v) \in G, \\ &\gamma(\omega^*, \mu^*, 0) = 0. \end{aligned} \tag{11.13}$$

We also obtain

$$D_v\gamma(\omega^*, \mu^*, 0) = -[D_w\Phi(\omega^*, \mu^*, 0, 0)]^{-1} \circ D_v\Phi(\omega^*, \mu^*, 0, 0) = 0$$

(cf. Remark 11.2, 2°).

Thus the bifurcation equation (corresponding to (11.9))

$$P(F(\omega, \mu, v + \gamma(\omega, \mu, v))) = 0 \tag{11.14}$$

holds good.

[6] cf. Lemma 10.5 on p. 293.

Define a function $h : \mathbb{R}^3 \to \mathfrak{Z}$ by

$$h(\omega, \mu, s) = PF(\omega, \mu, sv^* + \gamma(\omega, \mu, sv^*)).$$

h is of class $\mathfrak{C}^2$ and satisfies $h(\omega, \mu, 0) = 0$. The partial derivative of h with respect to s is given by

$$D_s h(\omega, \mu, s) = PD_x F(\omega, \mu, sv^* + \gamma(\omega, \mu, sv^*)) \cdot [v^* + D_v \gamma(\omega, \mu, sv^*) v^*]. \tag{11.15}$$

We denote it by $\chi(\omega, \mu, s)$; i.e.

$$\chi(\omega, \mu, s) = D_s h(\omega, \mu, s).$$

χ is of class $\mathfrak{C}^1$ and satisfies

$$\chi(\omega^*, \mu^*, 0) = D_s h(\omega^*, \mu^*, 0) = 0$$

by (11.13) and (11.15).

We now try to find derivatives of χ. First of all, we have

$$\begin{aligned} D_\mu \chi(\omega^*, \mu^*, 0) &= PD^2_{x,\mu} F(\omega^*, \mu^*, 0)[v^* + D_v \gamma(\omega^*, \mu^*, 0) v^*] \\ &\qquad + \underbrace{PD_x F(\omega^*, \mu^*, 0)}_{=0}[D^2_{v,\mu} \gamma(\omega^*, \mu^*, 0) v^*] \\ &= PD^2_{x,\mu} F(\omega^*, \mu^*, 0) v^* = PMv^*. \end{aligned}$$

Similarly,

$$D_\omega \chi(\omega^*, \mu^*, 0) = PNv^*.$$

$PMv^*, PNv^* \in \mathfrak{Z}$ and $\dim \mathfrak{Z} = 2$. So we may regard PMv^* and PNv^* as two elements of $\mathbb{R}^2$. By assumption,

$$\det D_{(\omega,\mu)} \chi(\omega^*, \mu^*, 0) \neq 0.$$

Applying the implicit function theorem to the equation $\chi(\omega, \mu, s) = 0$, we can solve this equation with respect to (ω, μ) locally in the neighborhood of $(\omega^*, \mu^*, 0)$. That is, there exist functions, $\omega(s)$ and $\mu(s)$, which are defined in a neighborhood of $s = 0$, and satisfy

$$\begin{aligned} &\chi(\omega(s), \mu(s), s) = 0, \\ &\omega(0) = \omega^*, \quad \mu(0) = \mu^*. \end{aligned}$$

By this relation, the bifurcation equation (11.14) becomes

$$PF(\omega(s), \mu(s), sv^* + \gamma(\omega(s), \mu(s), sv^*)) = 0.$$

We observe

$$(\omega(s), \mu(s)) \to (\omega^*, \mu^*) \quad \text{as} \quad s \to 0.$$

If we define

$$x_s = sv^* + \gamma(\omega(s), \mu(s), sv^*),$$

then $x_s \neq 0$ $(s \neq 0)$ and $x_s \to 0$ as $s \to 0$.

Thus we conclude that (ω^*, μ^*) is a bifurcation point of F. □

11.3 Classical Hopf Bifurcation for Ordinary Differential Equations

We now turn to the classical bifurcation phenomena of periodic solutions for an ordinary differential equation. Let $f(\mu, x)$ be a function of the class $\mathfrak{C}^2(\mathbb{R}\times\mathbb{R}^n, \mathbb{R}^n)$, and consider the differential equation

$$\frac{dx}{ds} = f(\mu, x). \tag{11.16}$$

Changing the time variable s by the relation

$$t = \omega s \quad (\omega \neq 0),$$

we rewrite the equation (11.16) as

$$\frac{dx}{dt} = \frac{1}{\omega} f(\mu, x),$$

$$\text{i.e.} \quad \omega \frac{dx}{dt} = f(\mu, x). \tag{11.16$'$}$$

This is an ordinary differential equation with two real parameters, ω and μ. For the sake of simplicity, we assume that the function $f(\mu, x)$ satisfies

$$f(\mu, 0) = 0 \quad \text{for all} \quad \mu \in \mathbb{R}.$$

We denote by $\mathfrak{W}^{1,2}_{2\pi}(\mathbb{R}, \mathbb{R}^n)$ the set of all the 2π-periodic and absolutely continuous functions $x : \mathbb{R} \to \mathbb{R}^n$ such that $\dot{x}|_{[0,2\pi]} \in \mathfrak{L}^2([0, 2\pi], \mathbb{R}^n)$, where $\dot{x}|_{[0,2\pi]}$ denotes the restriction of $\dot{x} = dx/dt$ to the interval $[0, 2\pi]$; i.e.

$$\mathfrak{W}^{1,2}_{2\pi} = \{x : \mathbb{R} \to \mathbb{R}^n \mid x \text{ is } 2\pi\text{-periodic, absolutely continuous and} \\ \dot{x}(\cdot)|_{[0,2\pi]} \in \mathfrak{L}^2([0, 2\pi], \mathbb{R}^n)\}.$$

$\mathfrak{W}^{1,2}_{2\pi}$ is a Banach space under the norm

$$\|x\|_{\mathfrak{W}^{1,2}_{2\pi}} = \Big(\int_0^{2\pi} \|x(t)\|^2 dt\Big)^{1/2} + \Big(\int_0^{2\pi} \|\dot{x}(t)\|^2 dt\Big)^{1/2}.$$

We also denote by $\mathfrak{L}^2_{2\pi}(\mathbb{R}, \mathbb{R}^n)$ the set of all the 2π-periodic measurable functions $y : \mathbb{R} \to \mathbb{R}^n$ such that $y|_{[0,2\pi]} \in \mathfrak{L}^2([0, 2\pi], \mathbb{R}^n)$; i.e.

$$\mathfrak{L}^2_{2\pi} = \{y : \mathbb{R} \to \mathbb{R}^n \mid y \text{ is } 2\pi\text{-periodic and } y|_{[0,2\pi]} \in \mathfrak{L}^2([0, 2\pi], \mathbb{R}^n)\}.$$

$\mathfrak{L}^2_{2\pi}$ is a Banach space under the norm

$$\|y\|_{\mathfrak{L}^2_{2\pi}} = \Big(\int_0^{2\pi} \|y(t)\|^2 dt\Big)^{1/2}.$$

In this section, we adopt $\mathfrak{W}^{1,2}_{2\pi}$ as $\mathfrak{X}$ and $\mathfrak{L}^2_{2\pi}$ as $\mathfrak{Y}$, respectively; i.e.

$$\mathfrak{X} = \mathfrak{W}^{1,2}_{2\pi}, \quad \mathfrak{Y} = \mathfrak{L}^2_{2\pi}.$$

Define the function $F : \mathbb{R}^2 \times \mathfrak{X} \to \mathfrak{Y}$ by

$$F(\omega, \mu, x) = \omega\frac{dx}{dt} - f(\mu, x).$$

Then we can prove that F is a function of the class $\mathfrak{C}^2(\mathbb{R}^2 \times \mathfrak{X}, \mathfrak{Y})$ provided that the following assumptions are satisfied.

Assumption 1

(i) *There exists some constants α and $\beta \in \mathbb{R}$ such that*

$$\|f(\mu, x)\| \leqq \alpha + \beta\|x\| \quad \textit{for all} \quad x \in \mathbb{R}^n.$$

(ii) *There exists some constant ρ such that*

$$\|D_x f(\mu, x)\|, \quad \|D^2 f(\mu, x)\| \leqq \rho \quad \textit{for all} \quad x \in \mathbb{R}^n.$$

The proof of the fact $F(\cdot) \in \mathfrak{C}^2(\mathbb{R}^2 \times \mathfrak{X}, \mathfrak{Y})$ *will be given in the next section.*

It is obvious that

$$F(\omega, \mu, 0) = 0 \quad \text{for all} \quad (\omega, \mu) \in \mathbb{R}^2.$$

$(\omega^*, \mu^*) \in \mathbb{R}^2$ is called a *bifurcation point* of F if there exists a sequence (ω_n, μ_n, x_n) in $\mathbb{R}^2 \times \mathfrak{X}$ such that

$$\begin{cases} F(\omega_n, \mu_n, x_n) = 0, \\ (\omega_n, \mu_n) \to (\omega^*, \mu^*) \quad \text{as} \quad n \to \infty, \quad \text{and} \\ x_n \neq 0, \quad x_n \to 0 \quad \text{as} \quad n \to \infty. \end{cases}$$

Each x_n is a nontrivial (not identically zero) periodic solution with period 2π of the equation

$$\omega_n \frac{dx}{dt} - f(\mu_n, x) = 0.$$

Hence, changing the time-variable to s again, we obtain a periodic solution

$$X_n(s) = x_n(\omega_n s)$$

with period $\tau_n = 2\pi/\omega_n$ for the original equation (11.16). Consequently, we have that

$$\tau_n \to \tau^* = \frac{2\pi}{\omega^*}, \quad \| X_n \|_{\mathfrak{W}^{1,2}_{2\pi/\omega_n}} \to 0 \quad \text{as} \quad n \to \infty,$$

provided that $\omega^* \neq 0$. The target of our investigations is to find a bifurcation point of F according to the principle of Theorem 11.1. We have to note that the derivative

$$D_x F(\omega, \mu, 0) : \ x \mapsto \omega \dot{x} - D_x f(\mu, 0)x \tag{11.17}$$

is to play the most important role in the course of our discussions. ($\dot{x}$ means dx/dt.) Of course, $D_x F(\omega, \mu, 0)$ is a bounded linear operator of $\mathfrak{X}$ into $\mathfrak{Y}$. If we write

$$A_\mu = D_x f(\mu, 0),$$

A_μ is an $(n \times n)$-matrix and (11.17) can be rewritten in the form

$$D_x F(\omega, \mu, 0)x = \omega \dot{x} - A_\mu x.$$

Here we need a couple of assumptions to be imposed upon the matrix A_μ at some (ω^*, μ^*).

Assumption 2 *A_{μ^*} is regular, and $\pm i\omega^*(\omega^* > 0)$ are simple eigenvalues of A_{μ^*}.*

Assumption 3 *None of $\pm ik\omega^*$ $(k \neq \pm 1)$ is an eigenvalue of A_{μ^*}.*

11.4 Smoothness of *F*

In this section, we examine the differentiability of the so-called Nemyckii operator.[7]

Let us define the operator Φ on $\mathfrak{W}^{1,2}_{2\pi}$ by

$$\Phi(x(\cdot)) = f(\mu, x(\cdot)), \quad x \in \mathfrak{W}^{1,2}_{2\pi}$$

for any fixed μ. If Assumption 1(i) is satisfied, then $\Phi(x(\cdot))$ is in $\mathfrak{L}^2_{2\pi}$.

Lemma 11.1 *Under Assumption 1(i), the operator $\Phi : \mathfrak{W}^{1,2}_{2\pi} \to \mathfrak{L}^2_{2\pi}$ is continuous.*

Proof Assume that $x_n(\cdot) \to x_0(\cdot)$ (as $n \to \infty$) in $\mathfrak{W}^{1,2}_{2\pi}$. It obviously implies that $x_n(\cdot) \to x_0(\cdot)$ (as $n \to \infty$) in $\mathfrak{L}^2_{2\pi}$.

Then there exists some subsequence $\{x_{n'}(\cdot)\}$ and a function $\varphi(\cdot) \in \mathfrak{L}^2([0, 2\pi], \mathbb{R})$ such that

$$x_{n'}(t) \to x_0(t) \quad \text{a.e. on} \quad [0, 2\pi], \text{ and}$$

$$\|x_{n'}(t)\| \leqq \varphi(t) \quad \text{a.e. on} \quad [0, 2\pi].$$

Since f is $\mathfrak{C}^2$ by assumption, we must have

$$\Phi(x_{n'}(\cdot)) = f(\mu, x_{n'}(\cdot)) \to \Phi(x_0(\cdot)) = f(\mu, x_0(\cdot)) \quad \text{a.e.} \tag{11.18}$$

Taking account of Assumption 1(i), we obtain

$$\|f(\mu, x_{n'}(\cdot))\| \leqq \alpha + \beta\|x_{n'}(\cdot)\| \leqq \alpha + \beta\varphi(t) \quad \text{a.e.} \tag{11.19}$$

Now (11.18) and (11.19) imply that

$$\begin{aligned}\|\Phi(x_{n'}(\cdot)) - \Phi(x_0(\cdot))\|^2_{\mathfrak{L}^2_{2\pi}} &= \int_0^{2\pi} \|f(\mu, x_{n'}(t)) - f(\mu, x_0(t))\|^2 dt \\ &\to 0 \quad \text{as} \quad n' \to \infty\end{aligned}$$

by the dominated convergence theorem.

[7] Related topics are discussed in Ambrosetti-Prodi [1] pp. 17–21.

If follows that $\Phi(x_n(\cdot)) \to \Phi(x_0(\cdot))$ in $\mathfrak{L}^2_{2\pi}$. (If not, a contradiction would occur to the above argument.) □

Lemma 11.2 *If Assumption 1 is satisfied, then the operator* $\Phi : \mathfrak{W}^{1,2}_{2\pi} \to \mathfrak{L}^2_{2\pi}$ *is twice continuously differentiable.*

Proof Let us begin by evaluating the first variation of Φ. For any $x, z \in \mathfrak{W}^{1,2}_{2\pi}$, we have

$$\begin{aligned}
&\frac{1}{\lambda}[\Phi(x(\cdot)+\lambda z(\cdot)) - \Phi(x(\cdot))](t)\\
&\qquad = \frac{1}{\lambda}[f(\mu, x(t)+\lambda z(t)) - f(\mu, x(t))]\\
&\qquad = \frac{1}{\lambda}[D_x f(\mu, x(t))\lambda z(t) + o(\lambda z(t))]\\
&\qquad \to D_x f(\mu, x(t))z(t) \quad \text{as} \quad \lambda \to 0,
\end{aligned} \tag{11.20}$$

where we must have

$$D_x f(\mu, x(\cdot))z(\cdot) \in \mathfrak{L}^2_{2\pi} \quad \text{for any} \quad x(\cdot), z(\cdot) \in \mathfrak{W}^{1,2}_{2\pi} \tag{11.21}$$

by Assumption 1 and $z(\cdot) \in \mathfrak{W}^{1,2}_{2\pi}$.

We can also confirm that, for any $\varepsilon > 0$,

$$\begin{aligned}
&\left\| \frac{1}{\lambda}[\Phi(x(\cdot)+\lambda z(\cdot)) - \Phi(x(\cdot))](t) - D_x f(\mu, x(t))z(t) \right\|^2\\
&\qquad = \left\| D_x f(\mu, x(t))z(t) + \frac{o(\lambda z(t))}{\lambda} - D_x f(\mu, x(t))z(t) \right\|^2\\
&\qquad \leqq \varepsilon^2 \cdot \|z(t)\|^2 \qquad \text{for sufficiently small } \lambda.
\end{aligned}$$

Hence we obtain, by (11.20) and the dominated convergence theorem, that

$$\frac{1}{\lambda}[\Phi(x(\cdot)+\lambda z(\cdot)) - \Phi(x(\cdot))] \to D_x f(\mu, x(\cdot))z(\cdot) \qquad \text{in } \mathfrak{L}^2_{2\pi} \text{ as } \lambda \to 0.$$

Therefore Φ has the first variation of the form (11.21).

It is easy to check that the mapping

$$z(\cdot) \mapsto D_x f(\mu, x(\cdot))z(\cdot)$$

is a bounded linear operator of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}^2_{2\pi}$. Hence Φ is Gâteaux-differentiable.

Furthermore, Φ turns out to be Fréchet-differentiable[8] since

$$x(\cdot) \mapsto D_x f(\mu, x(\cdot))$$

is a continuous mapping of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}(\mathfrak{W}^{1,2}_{2\pi}, \mathfrak{L}^2_{2\pi})$ (the space of bounded linear operators of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}^2_{2\pi}$).[9]

Finally, we can prove that Φ is twice continuously differentiable by a similar method as above. So we omit the details. □

Writing $T = D_x F(\omega^*, \mu^*, 0)$, we have

$$Tx = 0 \quad \text{if and only if} \quad \omega^* \dot{x} - A_{\mu^*} x = 0.$$

In order to apply Theorem 11.1 to our classical problem in Sect. 11.3, we have to start by confirming that (a) the dimension of the kernel of T is 2, and (b) the codimension of the image of T is also 2. For brevity, we have to show that

$$\begin{aligned} &\dim \ \operatorname{Ker} T = 2, \quad \text{and} \\ &\operatorname{codim} T(\mathfrak{X}) = 2. \end{aligned}$$

Henceforth, we denote Ker T by $\mathfrak{V}$ and $T(\mathfrak{X})$ by $\mathfrak{R}$.

Remark 11.3 Actually, we can confirm a priori that T is a Fredholm operator with index zero when $\omega^* = 0$. Hence we do not have to check both $\dim \mathfrak{V} = 2$ and $\operatorname{codim} \mathfrak{R} = 2$, because either one follows from the other automatically.

The operator $T : x \mapsto \omega^* \dot{x} - A_{\mu^*} x \quad (\mathfrak{W}^{1,2}_{2\pi} \to \mathfrak{L}^2_{2\pi})$ can be rewritten as

$$\begin{aligned} Tx &= \omega^* \dot{x} - A_{\mu^*} x \\ &= \omega^* \dot{x} + x - x - A_{\mu^*} x \\ &= \omega^* \dot{x} + x - (I + A_{\mu^*}) x. \end{aligned}$$

Since the mapping $x \mapsto \omega^* \dot{x} + x$ $(\omega^* \neq 0)$ is an isomorphism between $\mathfrak{W}^{1,2}_{2\pi}$ and $\mathfrak{L}^2_{2\pi}$, it is a Fredholm operator with index zero. We also know that the inclusion mapping of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}^2_{2\pi}$ is a compact operator.[10] Consequently, the mapping $x \mapsto (I + A_{\mu^*})x$ is also a compact operator of $\mathfrak{W}^{1,2}_{2\pi}$ into $\mathfrak{L}^2_{2\pi}$. (Note that A_{μ^*} is a bounded operator of $\mathfrak{L}^2_{2\pi}$ into itself.) Thus we confirm that the operator T can be expressed

[8]Let $\mathfrak{V}$ and $\mathfrak{W}$ be a couple of Banach spaces. Assume that a function φ of an open subset U of $\mathfrak{V}$ into $\mathfrak{W}$ is Gâteaux-differentiable in a neighborhood V of $x \in U$. We denote by $\delta\varphi(v)$ the Gâteaux-derivative of φ at v. If the function $v \mapsto \delta\varphi(v)(V \to \mathfrak{L}(\mathfrak{V}, \mathfrak{W}))$ is continuous, then φ is Fréchet-differentiable. cf. Maruyama [12] pp. 236–237.

[9]The continuity of the mapping can be proved in the same manner as in the proof of Lemma 11.1. Assumption 1(i) is used again for the dominated convergence argument.

[10]This is a special case of the Rellich–Kondrachov Compactness Theorem. Evans [5] pp. 272–274.

as a sum of a Fredholm operator with index zero and a compact operator. It follows from Theorem 10.12 (p. 299) that T is also a Fredholm operator with index zero.

However, we prefer an elementary way to prove both $\dim\mathfrak{V} = 2$ and $\operatorname{codim}\mathfrak{R} = 2$ without having recourse to the Fredholm operator theory discussed above.

I appreciate Professor S. Kusuoka's suggestion on this point.

11.5 dim $\mathfrak{V} = 2$

Expanding $x \in \mathfrak{X}$ in the Fourier series, we obtain

$$x(t) = \sum_{k=-\infty}^{\infty} u_k e^{ikt}, \quad u_k \in \mathbb{C}^n, \tag{11.22}$$

where u_k is a vector, the j-th coordinate of which is given by

$$\frac{1}{2\pi}\int_0^{2\pi} x_j(t)e^{-ikt}dt, \quad j = 1, 2, \cdots, n.$$

We must have the relation

$$u_{-k} = \bar{u}_k \quad \text{(conjugate)}, \quad k \in \mathbb{Z},$$

since $x(t)$ is a real vector. Furthermore, we have to keep in mind that the series (11.22) is uniformly convergent by Theorem 2.5 (p. 37).[11]

The k-th Fourier coefficient (times $1/\sqrt{2\pi}$, rigorously speaking) of $\dot{x}(\cdot)$is given by iku_k. Since $x(\cdot) \in \mathfrak{W}^{1,2}_{2\pi}$, it is clear that $\dot{x}(\cdot) \in \mathfrak{L}^2_{2\pi}$. Hence we obtain

$$\dot{x}(t) = \sum_{k=-\infty}^{\infty} iku_k e^{ikt} \quad \text{a.e.}, \tag{11.23}$$

that is, the Fourier series of $\dot{x}(\cdot)$ given by the right-hand side of (11.23) converges a.e. and equal to $\dot{x}(\cdot)$. This result is justified by Carleson's theorem (p. 37). It follows that

$$\omega^*\dot{x} - A_{\mu^*}x = \sum_{k=-\infty}^{\infty} [ik\omega^* I - A_{\mu^*}]u_k e^{ikt} = 0 \quad \text{a.e.}$$

[11] If a 2π-periodic function $\varphi : \mathbb{R} \to \mathbb{R}$ is absolutely continuous and its derivative φ' belongs to $\mathfrak{L}^2([0, 2\pi], \mathbb{R})$, then the Fourier series of φ uniformly converges to φ on $\mathbb{R}$. The k-th Fourier coefficient of φ' is given by $ik\hat{\varphi}(k)$, where $\hat{\varphi}(k)$ is the k-th Fourier coefficient of φ.

By the uniqueness of the Fourier coefficients (with respect to the complete orthonormal system $(1/\sqrt{2\pi})e^{ikt}$; $k = 0, \pm 1, \pm 2, \cdots$), we must have

$$[ik\omega^* I - A_{\mu^*}]u_k = 0 \quad \text{for all} \quad k \in \mathbb{Z}. \tag{11.24}$$

It is enough to find all $x \in \mathfrak{X}$, the Fourier coefficients of which satisfy (11.24), in order to determine $\mathfrak{V}$.

By the regularity of A_{μ^*} in Assumptions 2 and 3,

$$ik\omega^* I - A_{\mu^*}, \quad k \neq \pm 1$$

are all invertible. Therefore we have simply

$$u_k = 0 \quad \text{for} \quad k \neq \pm 1.$$

Thus all the coefficients in (11.22) besides $k = \pm 1$ must be zero. Consequently, all we have to do is to find functions whose Fourier coefficients $u_{\pm 1}$ satisfy

$$[\pm i\omega^* I - A_{\mu^*}]u_{\pm 1} = 0. \tag{11.25}$$

Since $\pm i\omega^*$ are simple eigenvalues of A_{μ^*} by Assumption 2, there exists $\xi \in \mathbb{C}^n$ $(\xi \neq 0)$ such that[12]

$$\text{Ker}\{i\omega^* I - A_{\mu^*}\} = \text{span}\{\xi\}. \tag{11.26}$$

On the other hand,

$$\text{Ker}\{-i\omega^* I - A_{\mu^*}\} = \text{span}\{\bar{\xi}\}.$$

The Fourier coefficients $u_{\pm 1}$ which satisfy (11.25) can be expressed as

$$u_{+1} = a\xi, \quad u_{-1} = b\bar{\xi} \quad (a, b \in \mathbb{C}).$$

Then any real solution $x(t)$ of (11.23) is of the form:

$$x(t) = a\xi e^{it} + \bar{a}\bar{\xi}e^{-it}.$$

If we set $a = \alpha + i\beta$, $\xi = \gamma + i\delta$ $(\alpha, \beta \in \mathbb{R};\ \gamma, \delta \in \mathbb{R}^n)$, it follows from simple calculations that

$$\begin{aligned} x(t) &= 2Re[a\xi e^{it}] \\ &= 2[\alpha(\gamma \cos t - \delta \sin t) - \beta(\gamma \sin t + \delta \cos t)]. \end{aligned}$$

[12] span$\{\xi\}$ denotes the subspace of $\mathbb{C}^n$ spanned by ξ.

Writing $p(t) = \gamma \cos t - \delta \sin t$ and $q(t) = \gamma \sin t + \delta \cos t$, we have

$$x(t) = 2\alpha p(t) - 2\beta q(t),$$

where α and β are any real numbers. It can easily be checked that $p(t)$ and $q(t)$ are linearly independent.[13]

Thus we have shown that $p(\cdot)$ and $q(\cdot)$ form a base of $\mathfrak{V}$. And so $\dim\mathfrak{V} = 2$.

11.6 codim $\mathfrak{R} = 2$

We now proceed to examining the dimension of the quotient space $\mathfrak{Y}/\mathfrak{R}$ of $\mathfrak{Y}$ modulo $\mathfrak{R}$.

Writing

$$Tx = y, \quad x \in \mathfrak{X}, \quad y \in \mathfrak{Y},$$

we again expand x and y in the Fourier series. Let u_k (resp. v_k) be the Fourier coefficients (times $1/\sqrt{2\pi}$) of x (resp. y). Then we must have

$$\sum_{k=-\infty}^{\infty} [ik\omega^* I - A_{\mu^*}]u_k e^{ikt} = \sum_{k=-\infty}^{\infty} v_k e^{ikt}. \tag{11.27}$$

Since $x \in \mathfrak{W}^{1,2}_{2\pi}$, $y \in \mathfrak{L}^2_{2\pi}$ and both of them are 2π-periodic, the right-hand side of (11.27) converges a.e. again by Carleson's theorem (p. 37). By the uniqueness of the Fourier coefficients, we must have

$$[ik\omega^* I - A_{\mu^*}]u_k = v_k \quad \text{for all} \quad k \in \mathbb{Z}. \tag{11.28}$$

---

[13] Set $\mu p(t)+\nu q(t) = \mu(\gamma \cos t-\delta \sin t)+\nu(\gamma \sin t+\delta \cos t) = (\mu\gamma+\nu\delta) \cos t+(\nu\gamma-\mu\delta) \sin t = 0$. Then we have

$$\begin{cases} \mu\gamma + \nu\delta = 0, \\ \nu\gamma - \mu\delta = 0. \end{cases}$$

It follows that

$$\begin{cases} \mu\nu\gamma + \nu^2\delta = 0, \\ \mu\nu\gamma - \mu^2\delta = 0. \end{cases}$$

Hence $(\nu^2 + \mu^2)\delta = 0$. If $\mu \neq 0$ or $\nu \neq 0$, δ must be zero. And so $\xi = \gamma$, that is $\xi = \bar{\xi}$ (real vector). Thus we get a contradiction.

By the regularity of A_{μ^*} in Assumptions 2 and 3, u_k $(k \neq \pm 1)$ in (11.28) can be solved uniquely in the form:

$$\begin{aligned} u_0 &= -A_{\mu^*}^{-1} v_0, \\ u_k &= [ik\omega^* I - A_{\mu^*}]^{-1} v_k, \quad k \neq 0, \ \pm 1. \end{aligned} \tag{11.29}$$

Since the norm of $(1/ik\omega^*)A_{\mu^*}$ is less than 1 for sufficiently large $|k|$'s (say $|k| \geqq k_0$), we have

$$\begin{aligned} [ik\omega^* I - A_{\mu^*}]^{-1} &= \frac{1}{ik\omega^*}\left[I - \frac{1}{ik\omega^*}A_{\mu^*}\right]^{-1} \\ &= \frac{1}{ik\omega^*}\left[I + \frac{1}{ik\omega^*}A_{\mu^*} + \frac{1}{(ik\omega^*)^2}A_{\mu^*}^2 + \cdots\right] \\ &= \frac{1}{ik\omega^*}I + O\left(\frac{1}{k^2}\right) \quad \text{for} \quad |k| \geqq k_0. \end{aligned} \tag{11.30}$$

It follows, from (11.29) and (11.30), that

$$u_k = [ik\omega^* I - A_{\mu^*}]^{-1} v_k = \frac{1}{ik\omega^*} v_k + O\Big(\frac{1}{k^2}\Big) v_k \quad \text{for} \quad |k| \geqq k_0.$$

Consequently,

$$\begin{aligned} \sum_{k \neq 0, \pm 1} u_k e^{ikt} = &\sum_{\substack{|k| < k_0 \\ k \neq 0, \pm 1}} [ik\omega^* I - A_{\mu^*}]^{-1} v_k e^{ikt} \\ &+ \sum_{|k| \geqq k_0} \frac{v_k}{ik\omega^*} e^{ikt} + \sum_{|k| \geqq k_0} O\Big(\frac{1}{k^2}\Big) v_k e^{ikt}. \end{aligned} \tag{11.31}$$

We claim that the second term of the right-hand side of (11.31) is of the class $\mathfrak{W}_{2\pi}^{1,2}$. In order to prove it, we define the function $\theta(t)$ by

$$\theta(t) = \sum_{|k| \geqq k_0} v_k e^{ikt}. \tag{11.32}$$

Then $\theta(t)$ is of the class $\mathfrak{L}_{2\pi}^2$. Let $\Theta(t)$ be the indefinite integral of $\theta(t)$, that is

$$\Theta(t) = \int_0^t \theta(\tau) d\tau. \tag{11.33}$$

Clearly, $\Theta(t)$ is of the class $\mathfrak{W}^{1,2}_{2\pi}$ and $\dot{\Theta}(t) = \theta(t)$ a.e. Expanding $\theta(t)$ in the Fourier series, we confirm that the coefficient $\hat{\theta}(0)$ corresponding to $k = 0$ is zero. Hence we obtain by Remark 3.3 (p. 57)[14] that

$$\hat{\Theta}(k) = \frac{\hat{\theta}(k)}{ik} \quad \text{for all} \quad k \neq 0,$$

$$\text{i.e.} \quad \frac{1}{\omega^*}\Theta(t) = \sum_{|k| \geqq k_0} \frac{v_k}{ik\omega^*} e^{ikt}$$

$$= \text{the second term of (11.31).}$$

This is of the class $\mathfrak{W}^{1,2}_{2\pi}$ by (11.33).

We can also prove that the third term of the right-hand side of (11.31) is of the class $\mathfrak{W}^{1,2}_{2\pi}$ by a similar argument as above. In this case we define the function $\theta(t) \in \mathfrak{L}^2_{2\pi}$ by

$$\theta(t) = \sum_{|k| \geqq k_0} ikO\left(\frac{1}{k^2}\right) v_k \cdot e^{ikt}$$

instead of (11.32). $\Theta(t)$ is the indefinite integral of $\theta(t)$ as in (11.33).

The first term of the right-hand side of (11.31) is obviously of the class $\mathfrak{W}^{1,2}_{2\pi}$. This finishes the proof of the claim that the right-hand side of (11.31) is of the class $\mathfrak{W}^{1,2}_{2\pi}$.

Thus denoting the Fourier coefficients of any $y \in \mathfrak{Y}$ by v_k's, the function

$$\sum_{k \neq \pm 1} u_k e^{ikt} = -A^{-1}_{\mu^*} v_0 + \sum_{k \neq 0, \pm 1} [ik\omega^* I - A_{\mu^*}]^{-1} v_k e^{ikt}$$

is of the class $\mathfrak{W}^{1,2}_{2\pi}$ and u_k's defined here satisfy the relation (11.28).

[14] Let $f : \mathbb{R} \to \mathbb{R}$ (we may replace $\mathbb{R}$ by $\mathbb{R}^n$) be a 2π-periodic function which is integrable on $[-\pi.\pi]$. Furthermore, we assume $\hat{f}(0) = 0$ ($\hat{f}(0)$ is the Fourier coefficient corresponding to $k = 0$). If we define

$$F(t) = f(0) + \int_0^t f(\tau) d\tau,$$

F is a 2π-periodic continuous function and

$$\hat{F}(k) = \frac{1}{ik} \hat{f}(k), \quad k \neq 0.$$

We shall now go over to $k = \pm 1$. The equation

$$[\pm i\omega^* I - A_{\mu^*}]u_{\pm 1} = v_{\pm 1} \tag{11.34}$$

does or does not have a solution. Since

$$\mathrm{codim}[\pm i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = 1$$

by Assumption 2, there must exist some $\varphi \in \mathbb{C}^n$ $(\varphi \neq 0)$ such that

$$\mathbb{C}^n/[i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = \mathrm{span}\{\varphi + [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)\}.$$

On the other hand, we also have

$$\mathbb{C}^n/[-i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = \mathrm{span}\{\bar{\varphi} + [-i\omega^* I - A_{\mu^*}](\mathbb{C}^n)\}.$$

An element of $\mathfrak{Y}$ is not contained in $\mathfrak{R}$ if and only if its Fourier coefficients $v_{\pm 1}$ (corresponding to $k = \pm 1$) do not admit the existence of $u_{\pm 1}$ which satisfy (11.34). Such vector v_1(resp. v_{-1}) is contained in a class of span $\{\varphi + [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)\}$ (resp. $\bar{\varphi} + [-i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$). Therefore any element of $\mathfrak{Y}/\mathfrak{R}$ can be expressed as

$$a\varphi e^{it} + b\bar{\varphi}e^{-it} + \mathfrak{R}; \quad a, b \in \mathbb{C}.$$

In order to find a real solution, we should set $b = \bar{a}$. If we write $\varphi = \gamma + i\delta,\ a = \alpha + i\beta$ $(\alpha, \beta \in \mathbb{R};\ \gamma, \delta \in \mathbb{R}^n)$, we obtain (in the same manner as the calculations on pp. 316–317)

$$\begin{aligned} a\varphi e^{it} + b\bar{\varphi}e^{-it} &= 2[\alpha(\gamma \cos t - \delta \sin t) - \beta(\gamma \sin t + \delta \cos t)] \\ &= 2\alpha p(t) - 2\beta q(t), \end{aligned}$$

where $p(t) = \gamma \cos t - \delta \sin t$ and $q(t) = \gamma \sin t + \delta \cos t$. Thus we have confirmed that any element of $\mathfrak{Y}/\mathfrak{R}$ can be expressed as a linear combination of the two linearly independent elements, $p(t) + \mathfrak{R}$ and $q(t) + \mathfrak{R}$. Hence codim $\mathfrak{R} = 2$.

Remark 11.4 We can specify $\varphi = \xi$ (see (11.26) for the definition of ξ). Therefore we will henceforth choose ξ as φ.

Remark 11.5 We can represent $\mathfrak{X} = \mathfrak{W}^{1,2}_{2\pi}$ and $\mathfrak{Y} = \mathfrak{L}^2_{2\pi}$ in the form of topological direct sums:

$$\mathfrak{X} = \mathfrak{V} \oplus \mathfrak{W}, \quad \mathfrak{Y} = \mathfrak{Z} \oplus \mathfrak{R},$$

by choosing suitable subspaces $\mathfrak{W} \subset \mathfrak{X}$ and $\mathfrak{Z} \subset \mathfrak{Y}$, respectively. Of course, $\mathfrak{V} = \mathrm{Ker}T$ and $\mathfrak{R} = T(\mathfrak{X})$ as stated on p. 303. We also denote by P the projection of $\mathfrak{Y}$ into $\mathfrak{Z}$ corresponding to the direct product defined above.

11.7 Linear Independence of PMv^* and PNv^* (1)

According to Assumption 2, $\pm i\omega^*$ are simple eigenvalues of A_{μ^*}. Hence $\mathbb{C}^n$ can be expressed as a direct sum as (cf. Fig. 11.1)

$$\mathbb{C}^n = \mathrm{Ker}[\pm i\omega^* I - A_{\mu^*}] \oplus [\pm i\omega^* I - A_{\mu^*}](\mathbb{C}^n). \tag{11.35}$$

We shall now concentrate on the case $+i\omega^*$. (The case $-i\omega^*$ can be discussed similarly.)

Let $\eta \in \mathbb{C}^n$ $(\eta \neq 0)$ be any vector which is orthogonal to $[i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$. Define a function $g : \mathbb{R} \times \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n \times \mathbb{C}$ by

$$g(\mu, \lambda, \theta) = \begin{pmatrix} (\lambda I - A_\mu)(\xi + \theta) \\ \langle \eta, \theta \rangle \end{pmatrix}.$$

($\langle \cdot, \cdot \rangle$ denotes the inner product, that is $\langle \eta, \theta \rangle = \sum_{j=1}^n \eta_j \bar{\theta}_j$.) We should note again that the vector ξ is defined by (11.26) on p. 316. Then the function g is of the class $\mathfrak{C}^1$ and satisfies

$$g(\mu^*, i\omega^*, 0) = 0.$$

We are now going to solve the equation $g(\mu, \lambda, \theta) = 0$ locally with respect to (λ, θ) in terms of μ in some neighborhood of $(\mu^*, i\omega^*, 0)$. The derivative of g with respect to (λ, θ) is given by

$$D_{(\lambda,\theta)} g(\mu^*, i\omega^*, 0)(\lambda, \theta) = \begin{pmatrix} \lambda\xi + (i\omega^* I - A_{\mu^*})\theta \\ \langle \eta, \theta \rangle \end{pmatrix}.$$

Here $D_{(\lambda,\theta)} g(\mu^*, i\omega^*, 0)$ ($(n+1)\times(n+1)$ matrix) is regular by (11.35).[15] Applying the implicit function theorem, we obtain the following lemma.

Lemma 11.3 *There exist a couple of functions, $\lambda(\mu)$ and $\theta(\mu)$ of the $\mathfrak{C}^1$-class which are defined in some neighborhood of μ^* and satisfy*

[15]For any $(\alpha_0, \beta_0) \in \mathbb{C}^n \times \mathbb{C}$, there exist some $\lambda_0 \in \mathbb{C}$ and $\gamma_0 \in (i\omega^* I - A_{\mu^*})(\mathbb{C}^n)$ such that $\alpha_0 = \lambda_0 \xi + \gamma_0$. Such λ_0 and γ_0 are unique. Let $(\alpha_0, \beta_0) = (0, 0)$. Then we must have $\lambda_0 = 0$ and $\gamma_0 = 0$. The equation

$$\begin{pmatrix} (i\omega^* I - A_{\mu^*})\theta \\ \langle \eta, \theta \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

has a unique solution $\theta = 0 (\in \mathbb{C}^n)$ because $\mathrm{Ker}[i\omega^* I - A_{\mu^*}] \cap \mathrm{Ker}\langle \eta, \cdot \rangle = \mathrm{Ker}[i\omega^* I - A_{\mu^*}] \cap [i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = \{0\}$. Thus we conclude that $D_{(\lambda,\theta)} g(\mu^*, i\omega^*, 0)$ is injective.

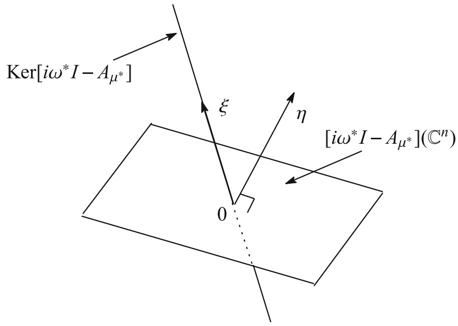


Fig. 11.1 Orthogonal decomposition of $\mathbb{C}^n$

$$\begin{pmatrix} (\lambda(\mu)I - A_\mu)(\xi + \theta(\mu)) \\ \langle \eta, \theta(\mu) \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{11.36}$$

and

$$\lambda(\mu^*) = i\omega^*, \quad \theta(\mu^*) = 0. \tag{11.37}$$

Denoting $\xi + \theta(\mu)$ in (11.36) by $\xi(\mu)$, we can rewrite the relations (11.36) and (11.37) as follows:

$$A_\mu \xi(\mu) = \lambda(\mu)\xi(\mu), \tag{11.36$'$}$$

$$\lambda(\mu^*) = i\omega^*, \ \xi(\mu^*) = \xi. \tag{11.37$'$}$$

11.8 Linear Independence of PMv^* and PNv^* (2)

Since $\mathrm{codim}[i\omega^* I - A_{\mu^*}](\mathbb{C}^n) = 1$ by Assumption 2, there exists a nonzero vector $\eta \in \mathbb{C}^n$ such that

$$\langle \eta, \kappa \rangle = 0 \quad \text{for all} \quad \kappa \in [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$$

as we have seen above. Taking account of the fact[16] that $\xi \notin [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$, we can choose η so that

$$\langle \eta, \xi \rangle = 1.$$

The function

$$\Pi : \ \kappa \mapsto \langle \eta, \kappa \rangle \xi, \ \kappa \in \mathbb{C}^n$$

[16] $i\omega^*$ is a simple eigenvalue, again by Assumption 2.

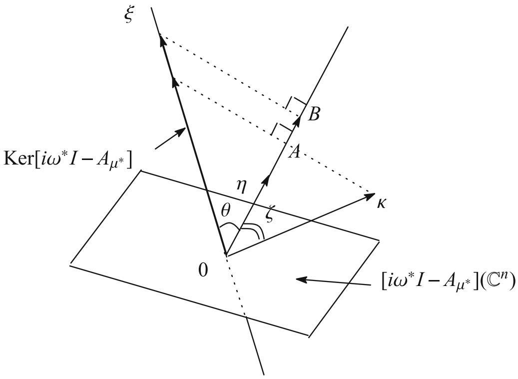


Fig. 11.2 Spectral projection

is called the *spectral projection*[17] associated with ξ. We can similarly define the spectral projection $\overline{\Pi}$ associated with $\bar{\xi}$ by using $\bar{\eta}$ instead of η.

Writing

$$A'_{\mu^*} = \frac{d}{d\mu} A_\mu \Big|_{\mu=\mu^*}$$

for the sake of simplicity, we get the following result.

Lemma 11.4 $\Pi A'_{\mu^*}\xi = \lambda'(\mu^*)\xi, \quad \overline{\Pi} A'_{\mu^*}\bar{\xi} = \overline{\lambda'(\mu^*)}\bar{\xi}.$

[17]Look at Fig. 11.2. For the sake of an intuitive exposition, the vectors η, ξ and κ are treated as real vectors. By $\langle \eta, \xi \rangle = \| \eta \| \cdot \| \xi \| \cos\theta = 1$, it follows that $\| \eta \| = 1 / \| \xi \| \cos\theta$. Hence

$$\Pi(\kappa) = \xi \langle \eta, \kappa \rangle = (\| \kappa \| \cos\zeta / \| \xi \| \cos\theta) \cdot \xi.$$

Since $\| \kappa \| \cos\zeta = 0A$ (the length of the segment) and $\| \xi \| \cos\theta = 0B$,

$$\Pi(\kappa) = \frac{0A}{0B} \cdot \xi.$$

Here ζ is the angle between η and κ, and θ is the one between ξ and η. κ can be represented uniquely as $\kappa = \alpha\eta + \beta z$ for some $\alpha, \beta \in \mathbb{C}$ and $z \in [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$. On the other hand, η can be represented uniquely in the form $\eta = a\xi + bz'$ for some $a, b \in \mathbb{C}$ and $z' \in [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$. Since $\| \eta \|^2 = \langle a\xi + bz', \eta \rangle = a\langle \xi, \eta \rangle + b\langle z', \eta \rangle = a$, it follows that $\eta = \| \eta \|^2 \xi + bz'$. Hence we have

$$\kappa = \alpha\eta + \beta z = \alpha \| \eta \|^2 \xi + (\alpha b z' + \beta z).$$

Furthermore, $\Pi(\kappa) = \langle \eta, \kappa \rangle \xi = \langle \eta, \alpha \| \eta \|^2 \xi + (\alpha b z' + \beta z) \rangle \xi = \alpha \| \eta \|^2 \xi$. ($\langle \eta, \alpha b z' + \beta z \rangle = 0$ because $\alpha b z' + \beta z \in [i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$.) Thus we obtain

$$\kappa = \Pi(\kappa) + (\alpha b z' + \beta z).$$

This is the direct sum of $\mathbb{C}^n$ corresponding to $\mathrm{span}\{\xi\}$ and $[i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$.

Proof It is enough to prove only the first part. It is straightforward that

$$A_\mu \xi = A_\mu(\xi - \xi(\mu)) + A_\mu \xi(\mu) = A_\mu(\xi - \xi(\mu)) + \lambda(\mu)\xi(\mu).$$

Define

$$\xi' = \frac{d}{d\mu}\xi(\mu)\Big|_{\mu=\mu^*}.$$

Then it follows that

$$\begin{aligned} A'_{\mu^*}\xi &= A'_{\mu^*}(\xi - \xi(\mu^*)) - A_{\mu^*}\xi' + \lambda'(\mu^*)\xi(\mu^*) + \lambda(\mu^*)\xi' \\ &= \lambda'(\mu^*)\xi(\mu^*) + (\lambda(\mu^*)I - A_{\mu^*})\xi' \\ &= \lambda'(\mu^*)\xi + (i\omega^* I - A_{\mu^*})\xi'. \end{aligned}$$

Taking account of the fact $\Pi[(i\omega^* I - A_{\mu^*})\xi'] = 0$, we must have

$$\Pi A'_{\mu^*}\xi = \lambda'(\mu^*)\xi.$$

□

We have finished the preparation for the spectral projection. We shall now go back to our task to evaluate PNv^* and PMv^*. Recall that $M = D^2_{x,\mu}F(\omega^*, \mu^*, 0)$ and $N = D^2_{x,\omega}F(\omega^*, \mu^*, 0)$.

If we specify $v^* \in \mathfrak{V}$ as

$$v^* = \xi e^{it} + \bar{\xi} e^{-it},$$

it follows that

$$\begin{aligned} D_x F(\omega, \mu, 0)v^* &= \omega \dot{v}^* - A_\mu v^* \\ &= i\omega(\xi e^{it} - \bar{\xi} e^{-it}) - A_\mu(\xi e^{it} + \bar{\xi} e^{-it}). \end{aligned} \tag{11.38}$$

Evaluation of PNv^*. We obtain, by (11.38), that

$$\underbrace{D^2_{x\omega}F(\omega^*, \mu^*, 0)}_{N} v^* = i\xi e^{it} - i\bar{\xi} e^{-it}. \tag{11.39}$$

In general, the projection Py of $y \in \mathfrak{Y}$ into $\mathfrak{Z}$ can be calculated as

$$Py = \Pi(v_1)e^{it} + \Pi(v_{-1})e^{-it},$$

where $v_{\pm 1}$ are the Fourier coefficients of y corresponding to $k = \pm 1$.[18] Hence, by (11.39), PNv^* is evaluated as

$$PNv^* = i\langle \eta, \xi \rangle \xi e^{it} - i\langle \bar{\eta}, \bar{\xi} \rangle \bar{\xi} e^{-it} = i\xi e^{it} - i\bar{\xi} e^{-it}. \tag{11.40}$$

Evaluation of PMv^*. Dividing $\lambda(\mu)$ (obtained by Lemma 11.3) into real and imaginary parts, we write

$$\lambda(\mu) = \alpha(\mu) + i\beta(\mu).$$

We also write

$$\lambda'(\mu) = \alpha'(\mu) + i\beta'(\mu).$$

It follows from (11.38) that

$$\underbrace{D^2_{x\mu} F(\omega, \mu, 0)}_{M} v^* = -A'_{\mu^*}(\xi e^{it} + \bar{\xi} e^{-it}).$$

Therefore we have by Lemma 11.4 that

$$\begin{aligned} PMv^* &= -\Pi A'_{\mu^*} \xi e^{it} - \overline{\Pi} A_{\mu^*} \bar{\xi} e^{-it} \\ &= -\lambda'(\mu^*) \xi e^{it} - \overline{\lambda'(\mu^*)} \bar{\xi} e^{-it} \\ &= -\alpha'(\mu^*)(\xi e^{it} + \bar{\xi} e^{-it}) - \beta'(\mu^*)(\xi e^{it} - \bar{\xi} e^{-it}). \end{aligned} \tag{11.41}$$

Comparing (11.40) and (11.41), we get a simple fact that PNv^* and PMv^* are linearly independent if and only if $\alpha'(\mu^*) \neq 0$.

Thus all the requirements in Theorem 11.1 are fulfilled if we make an additional assumption that $\alpha'(\mu^*) \neq 0$.

Theorem 11.2 *Let $f(\mu, x) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a function of the class $\mathfrak{C}^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$ which satisfies $f(\mu, 0) = 0$ for all $\mu \in \mathbb{R}$. Suppose that Assumptions 1–3 as well as the condition $\alpha'(\mu^*) \neq 0$ are satisfied. Then (ω^*, μ^*) is a bifurcation point of $F(\omega, \mu, x) = \omega dx/dt - f(\mu, x)$.*

Remark 11.6 (Due to S. Kusuoka) The additional condition $\alpha'(\mu^*) \neq 0$ can be expressed in an alternative equivalent form.

[18] Express each of the Fourier coefficients of y by the direct sum corresponding to span$\{\xi\}$ and $[i\omega^* I - A_{\mu^*}](\mathbb{C}^n)$, and delete all the terms which do not contribute to the former.

We denote by $\mathfrak{V}_0$ (resp. $\mathfrak{W}_0$) the kernel (resp. the image) of $\omega^{*^2} I + A^2_{\mu^*}$; i.e.

$$\mathfrak{V}_0 = \mathrm{Ker}\{\omega^{*^2} I + A^2_{\mu^*}\},$$
$$\mathfrak{W}_0 = [\omega^{*^2} I + A^2_{\mu^*}](\mathbb{R}^n).$$

We have, of course, that

$$\mathbb{R}^n = \mathfrak{V}_0 \oplus \mathfrak{W}_0. \tag{11.42}$$

Let $\pi_0 : \mathbb{R}^n \to \mathfrak{V}_0$ be the projection of $\mathbb{R}^n$ into $\mathfrak{V}_0$ corresponding to the above direct product (11.42). Then we can prove that

$$\alpha'(\mu^*) = 0 \quad \text{if and only if} \quad \mathrm{tr} A'_{\mu^*}\pi_0 \neq 0,$$

where tr is the trace of a matrix.

Proof Let $\mathfrak{V}_0^{\mathbb{C}} = \mathbb{C} \otimes \mathfrak{V}_0$ and $\pi_0^{\mathbb{C}} : \mathbb{C}^n \to \mathfrak{V}_0^{\mathbb{C}}$ be the complexifications of $\mathfrak{V}_0$ and π_0, respectively. Recall the formula (11.26) in Sect. 11.5; i.e.

$$\mathrm{Ker}\{i\omega^* I - A_{\mu^*}\} = \mathrm{span}\{\xi\}.$$

We write

$$\xi = a + ib; \quad a, b \in \mathbb{R}^n.$$

It can easily be checked that

$$A_{\mu^*} a = -\omega^* b \quad \text{and} \quad A_{\mu^*} b = \omega^* a.$$

Consequently, we also have that a and b are contained in $\mathfrak{V}_0$. They must be linearly independent. Hence there exist some $\tilde{a}$ and $\tilde{b} \in \mathbb{R}^n$ which satisfy the following conditions:

$$\langle \tilde{a}, a \rangle = 1, \quad \langle \tilde{a}, b \rangle = 0,$$
$$\langle \tilde{b}, a \rangle = 0, \quad \langle \tilde{b}, b \rangle = 1, \text{ and}$$
$$\langle \tilde{a}, w \rangle = \langle \tilde{b}, w \rangle = 0 \quad \text{for all } w \in \mathfrak{W}_0.$$

We then define the vector η by

$$\eta = \frac{1}{2}(\tilde{a} + i\tilde{b}).$$

This η clearly satisfies the condition in Sect. 11.7 (p. 321) that $\eta \neq 0$ and η is orthogonal to $[i\omega^* I - A_{\mu^*}](\mathbb{C}^*)$.

Since $\Pi A'_{\mu^*}\xi = \lambda'(\mu^*)\xi$ by Lemma 11.4, it follows that

$$\lambda'(\mu^*) \equiv \langle \eta, A'_{\mu^*}\xi \rangle = \frac{1}{2}\langle \tilde{a} + i\tilde{b}, A'_{\mu^*}\xi \rangle.$$

Finally, we obtain the desired result by

$$\alpha'(\mu^*) = Re\lambda'(\mu^*) = \frac{1}{2}\{\langle \tilde{a}, A_{\mu^*}a \rangle + \langle \tilde{b}, A'_{\mu^*}b \rangle\} = \frac{1}{2}\mathrm{tr}A'_{\mu^*}\pi_0.$$

This proves the desired result. □

Example 11.1 Consider the differential equation of the Van der Pol type

$$\ddot{x} - (\mu - 3x^2)\dot{x} + x = 0, \tag{11.43}$$

where μ is a real parameter.[19] The equation (11.43) is equivalent to

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -x + (\mu - 3x^2)y. \end{cases} \tag{11.44}$$

Let $u = (x, y)$. Define $f(\mu, u) = (y, -x + (\mu - 3x^2)y)$, and set $\omega^* = 1$. Then (11.44) is reduced to the form

$$\omega\dot{u} = f(\mu, u).$$

If we define $F(\omega, \mu, u) = \omega\dot{u} - f(\mu, u)$, then $(\omega^*, \mu^*) = (1, 0)$ is a bifurcation point of F.

$D_u f(\mu, 0) = A_\mu$ is given by

$$A_\mu = \begin{pmatrix} 0 & 1 \\ -1 & \mu \end{pmatrix}.$$

Hence the eigenvalues $\lambda(\mu)$ of A_μ are

$$\lambda(\mu) = \frac{1}{2}(\mu \pm \sqrt{\mu^2 - 4}).$$

In the case of $\mu = \mu^* = 0$, we have

$$\lambda(0) = \pm i.$$

[19] $\dot{x}$ and $\ddot{x}$ denote the first and second derivatives of x with respect to t, respectively.

There are clearly simple eigenvalues of $A_{\mu^*} = A_0$ and there is no other eigenvalue. A_0 is regular. Writing $\lambda(\mu) = \alpha(\mu) + i\beta(\mu)$, we finally obtain $\alpha'(0) = 1/2 \neq 0$. Consequently, we can conclude that $(\omega^*, \mu^*) = (1, 0)$ is a bifurcation point of F.

11.9 Hopf Bifurcation in $\mathfrak{C}^r$

In the preceding sections, we examined the Hopf bifurcation phenomena in the framework of the Sobolev space $\mathfrak{X} = \mathfrak{W}^{1,2}_{2\pi}$. However, we have to note that some technical modifications are required when we consider the same problem in the alternative space consisting of periodic smooth functions.

Here we specify a couple of function spaces, $\mathfrak{X}$ and $\mathfrak{Y}$, as

$$\mathfrak{X} = \{x \in \mathfrak{C}^r(\mathbb{R}, \mathbb{R}^n) \mid x(t + 2\pi) = x(t) \text{ for all } t\},$$

$$\mathfrak{Y} = \{x \in \mathfrak{C}^{r-1}(\mathbb{R},\mathbb{R}^n) \mid y(t + 2\pi) = y(t) \text{ for all } t\},$$

where $r \geqq 3$.

Furthermore, the function $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be of the class $\mathfrak{C}^{r-1}(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$.

1° The first point to be examined is the formula (11.23) in Sect. 11.5. In the alternative setting, we have to proceed somewhat more carefully as follows.

Since x is of the class $\mathfrak{C}^r \quad (r \geqq 3)$, we must have

$$u_k = o\left(\frac{1}{|k|^r}\right) \quad \text{as} \quad |k| \to \infty.$$[20]

Therefore, for any $\varepsilon > 0$, there exists some $N \in \mathbb{N}$ such that $|u_k| \leqq \varepsilon \cdot (1/|k|^r) \quad (|k| \geqq N)$. By differentiating the right-hand side of (11.22) of Sect. 11.5 termwise, we obtain

$$\begin{aligned}
\left\| \sum_{k=-\infty}^{\infty} iku_k e^{ikt} \right\| &\leqq \sum_{k=-\infty}^{\infty} \|ku_k\| \\
&\leqq \sum_{|k|<N} \|ku_k\| + \varepsilon \sum_{|k|\geqq N} |k| \cdot \frac{1}{|k|^r} \\
&= \sum_{|k|<N} \|ku_k\| + \varepsilon \sum_{|k|\geqq N} \frac{1}{|k|^{r-1}}.
\end{aligned}$$

[20] See Theorem 3.6 (p. 59).

Thus, taking account of the condition $r \geqq 3$, the series obtained by the termwise differentiation of the right-hand side of (11.22) is uniformly convergent. Hence we obtain[21]

$$\dot{x}(t) = \sum_{k=-\infty}^{\infty} iku_k e^{ikt}.$$

2° The second modification is required in Sect. 11.6. The argument which succeeds the formula (11.31) should be replaced by the following one.

We claim that the second term of the right-hand side of (11.31) in Sect. 11.6 is of the class $\mathfrak{C}^r$. In order to prove it, we define the function $\theta(t)$ by

$$\theta(t) = \sum_{|k| \geqq k_0} v_k e^{ikt}.$$

Then $\theta(t)$ is of the class $\mathfrak{C}^{r-1}$. Let $\Theta(t)$ be the indefinite integral of $\theta(t)$, that is

$$\Theta(t) = \int_0^t \theta(\tau) d\tau. \tag{11.45}$$

Clearly, $\Theta(t)$ is continuously differentiable and $\dot{\Theta}(t) = \theta(t)$. Expanding $\theta(t)$ in the Fourier series, we confirm that the coefficient $\hat{\theta}(0)$ corresponding to $k = 0$ is zero. Hence we obtain by Remark 3.3 (p. 57) again[22] that

$$\hat{\Theta}(k) = \frac{\hat{\theta}(k)}{ik} \quad \text{for all} \quad k \neq 0,$$

$$\text{i.e.} \quad \frac{1}{\omega^*}\Theta(t) = \sum_{|k| \geqq k_0} \frac{v_k}{ik\omega^*} e^{ikt}$$

$$= \text{the second term of (11.31)}.$$

This is of the class $\mathfrak{C}^r$ by (11.45).

The third term of the right-hand side of (11.31) in Sect. 11.6 is of the class $\mathfrak{C}^{r-1}$, and converges at every t. Differentiating termwise formally, we get

$$\sum_{|k| \geqq k_0} O\left(\frac{1}{k^2}\right) v_k \cdot ike^{ikt}. \tag{11.46}$$

---

[21] Assume that $\varphi_n : [a, b] \to \mathbb{R}$ is of class $\mathfrak{C}^1$ $(n = 1, 2, \cdots)$ and $S(t) = \sum_{n=1}^{\infty} \varphi_n(t)$ is convergent. If $\sum_{n=1}^{\infty} \varphi_n'(t)$ converges uniformly, $S(t)$ is differentiable and $S'(t) = \sum_{n=1}^{\infty} \varphi_n'(t)$. cf. Takagi [21], pp. 158–159, Stromberg [20] pp. 214–215. We assumed $r \geqq 3$ to use this theorem.

[22] See footnote 11 on p. 315.

The series (11.31) is uniformly convergent, since $v_k = o(1/k^{r-1})$ $(r \geqq 3)$. Hence the third term of (11.31) in Sect. 11.6 is differentiable and (11.46) deduced above is exactly its derivative. Thus the third term is of the class $\mathfrak{C}^r$.

The first term of the right-hand side of (11.31) is obviously of the class $\mathfrak{C}^r$. This finishes the proof of the claim that the right-hand side of (11.46) is of the class $\mathfrak{C}^r$.

No change is required at all in the remaining part of the proof.

Theorem 11.3 *Let $f(\mu, x) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a function of the class $\mathfrak{C}^{r-1}(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$ $(r \geqq 3)$ which satisfies $f(\mu, 0) = 0$ for all $\mu \in \mathbb{R}$. Suppose that Assumptions 2, 3 and $\alpha'(\mu^*) \neq 0$ are satisfied. Then (ω^*, μ^*) is a bifurcation point of $F(\omega, \mu, x) = \omega dx/dt - f(\mu, x)$.*

11.10 Kaldorian Business Fluctuations

The phenomena of quasi-regular business fluctuations have been attracting the interest of numerous economists from both of theoretical and empirical viewpoints. In particular, the theories to describe and explain economic fluctuations have deepened very much since the "Great Depression", through the 1930s. I have also to add Frisch's famous paper [6] and Keynes'book [11] as indispensable contributions providing solid foundations for further developments since then.

Among the various economic theories on business fluctuations, the Hicks [7]–Samuelson [17, 18] model and the Kaldor [10] model should be regarded as the most typical and influential. The former tries to explain the business cycles as a result of interactions between the multiplier and the accelerator, whereas the latter is based upon the so-called "profit principle" of investment which is embodied in nonlinear investment functions. In the present chapter, we concentrate upon the Kaldorian theory of business fluctuations. Yasui [23] tried to clarify the mathematical structure of the theory and successfully reduced the fundamental dynamics due to Kaldor to the Liénard differential equation, which was quite familiar in the field of mathematical analysis. Yasui also investigated the periodic behavior of the solutions, having recourse to the theory of nonlinear oscillations which was then in the course of vigorous development.[23]

However, from the mathematical viewpoint, some more rigorous reasoning seems to be required in order to establish the existence of periodic solutions.

In this section, we present an approach to the existence problem of periodic solutions of the Kaldor–Yasui equation. The proof is based upon the Hopf bifurcation theorem.

According to the profit principle due to N. Kaldor, we assume that the gross investment I depends upon the (real) national income Y and the capital stock K. The relation between I and (Y, K) is expressed by the equation

$$I = \psi(Y, K),$$

[23]For an introductory exposition of business cycle theory, see Maruyama [15], Chap. 18.

which satisfies the conditions

$$D_Y\psi > 0 \quad \text{and} \quad D_K\psi < 0,$$

where $D_Y\psi$ and $D_K\psi$ are the partial derivatives of ψ with respect to Y and K, respectively. From now on, we specify the function ψ as

$$\psi(Y, K) = F(Y) - \mu K \quad (\mu > 0), \tag{11.47}$$

for the sake of simplicity. The function F is assumed to be twice continuously differentiable. If we denote by $\delta > 0$ the rate of depreciation of capital, the condition

$$I = \delta K \tag{11.48}$$

describes the situation of zero net investment. The saving function $S(Y)$ is simply given by

$$S(Y) = sY \quad (0 < s < 1).$$

In the state of zero net investment, we must have, by (11.47) and (11.48), that

$$F(Y) - \mu K = \delta K,$$

that is,

$$K = \frac{1}{\mu + \delta} F(Y).$$

It follows that

$$I = \frac{\delta}{\mu + \delta} F(Y), \tag{11.49}$$

since $I = \delta K$. The equation (11.49) expresses the relation between Y and I, under which there is no variation of the capital stock K. In Fig. 11.3, the relation is depicted as the curve RR' with upward slope.

The equilibrium $S = I$ on RR' is attained at the level of Y such that

$$sY = \frac{\delta}{\mu + \delta} F(Y).$$

We denote this level of Y by Y_0. The corresponding level K_0 of the capital stock and that $I = 0$ of the (gross) investment are determined by (11.48) and (11.49).

The difference between $I = \psi(Y, K)$ and $I_0 = \psi(Y_0, K_0)$ is calculated as

$$I - I_0 = F(Y) - F(Y_0) - \mu(K - K_0). \tag{11.50}$$

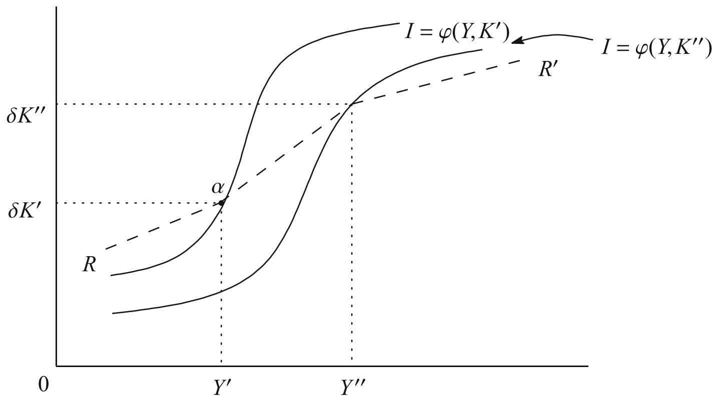


Fig. 11.3 Koldor's investment function

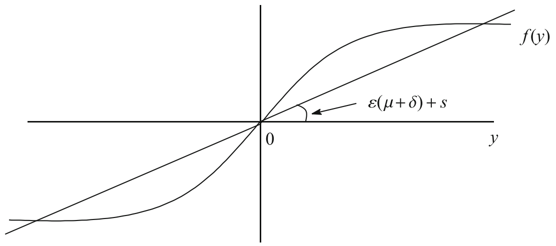


Fig. 11.4 Relation of y and i

Changing the variables as $i = I - I_0$, $y = Y - Y_0$, $k = K - K_0$ and $f(y) = F(Y) - F(Y_0)$, we can rewrite the equation (11.50) in the form

$$i = f(y) - \mu k. \tag{11.51}$$

Similarly, the relation (11.49) is reexpressed as (cf. Fig. 11.4)

$$i = \frac{\delta}{\mu + \delta} f(y).$$

Assume further that the rate of change in y is proportional to the difference between the investment and saving, say

$$\varepsilon \dot{y} = i - sy, \quad \varepsilon > 0, \tag{11.52}$$

where $\dot{y}$ means the derivative of y with respect to time.

Substituting (11.51) into (11.52), we have

$$\varepsilon \dot{y} = f(y) - \mu k - sy, \tag{11.53}$$

$$\text{i.e.} \quad \varepsilon\dot{y} - f(y) + \mu k + sy = 0. \tag{11.53$'$}$$

According to (11.51) again, the net investment is given by

$$\dot{k} = i - \delta k = f(y) - \mu k - \delta k. \tag{11.54}$$

Differentiating (11.53′) with respect to time t, we obtain

$$\begin{aligned}
\varepsilon\ddot{y} - f'(y)\dot{y} + \mu\dot{k} + s\dot{y} &= \varepsilon\ddot{y} - f'(y)\dot{y} + \mu[f(y) - \mu k - \delta k] + s\dot{y} \quad \text{(by (11.54))} \\
&= \varepsilon\ddot{y} + [s - f'(y)]\dot{y} + \mu[\varepsilon\dot{y} + sy - \delta k] \quad \text{(by (11.53))} \\
&= \varepsilon\ddot{y} + [s + \mu\varepsilon - f'(y)]\dot{y} + \mu(sy - \delta k) \\
&= \varepsilon\ddot{y} + [s + \mu\varepsilon - f'(y)]\dot{y} + \mu sy + \delta\varepsilon\dot{y} \\
&\quad - \delta f(y) + \delta sy \quad \text{(by (11.53))} \\
&= \varepsilon\ddot{y} + [\varepsilon(\mu + \delta) + s - f'(y)]\dot{y} + s(\mu + \delta)y - \delta f(y) \\
&= 0. \qquad (11.55)
\end{aligned}$$

Dividing both sides of (11.55) by ε, we have

$$\ddot{y} + \frac{1}{\varepsilon}\varphi(y)\dot{y} + g(y) = 0, \tag{11.56}$$

where

$$\varphi(y) = \varepsilon(\mu + \delta) + s - f'(y), \tag{11.57}$$

and

$$g(y) = \frac{\delta}{\varepsilon}\left[\frac{s(\mu + \delta)}{\delta}y - f(y)\right].$$

The differential equation (11.56) is clearly of the Liénard type. This is the fundamental dynamic equation of the Kaldor–Yasui theory of business cycles.

It is a duty for economic theorists to look for some striking sufficient conditions which guarantees the existence of periodic solutions for the differential equation (11.56). We now present a couple of alternative approaches leading to this target. See Chang and Smyth [3] and Shinasi [19] as related contributions.

The Kaldor–Yasui equation (11.56) can be rewritten as

$$\begin{pmatrix} \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} z \\ -\frac{1}{\varepsilon}[\varepsilon(\mu + \delta) + s - f'(y)]z - \frac{\delta}{\varepsilon}\left[\frac{s(\mu + \delta)}{\delta}y - f(y)\right] \end{pmatrix}. \tag{11.58}$$

If we denote ${}^t(y, z)$ by u and the right-hand side of (11.58) by $h(\mu, u)$, regarding μ as a parameter, (11.58) can be shortened as

$$\dot{u} = h(\mu, u). \tag{11.59}$$

Consider an equation

$$K(\omega, \mu, u) = \omega\dot{u} - h(\mu, u) = 0, \tag{11.60}$$

where ω is another parameter. Writing $A_\mu = D_u h(\mu, 0)$, we obtain

$$A_\mu = \begin{pmatrix} 0 & 1 \\ -\dfrac{s(\mu+\delta)}{\varepsilon} + \dfrac{\delta}{\varepsilon} f'(0) & -\dfrac{1}{\varepsilon}\big[\varepsilon(\mu+\delta) + s - f'(0)\big] \end{pmatrix}.$$

We specify the value μ^* of μ so that[24]

$$-\frac{s(\mu+\delta)}{\varepsilon} + \frac{\delta}{\varepsilon} f'(0) = -1, \tag{11.61}$$

and

$$-\frac{1}{\varepsilon}[\varepsilon(\mu+\delta) + s - f'(0)] = 0.$$

The value of ω^* is specified as $\omega^* = 1$. Then $\pm i\omega^* = \pm i$ are simple eigenvalues of A_{μ^*}. If we write the eigenvalue of A_μ by $\lambda(\mu) = \alpha(\mu) \pm i\beta(\mu)$, we can easily confirm that $\alpha'(\mu^*) \neq 0$. We thus obtain the following theorem by Theorem 11.3.

Theorem 11.4 $(\omega^*, \mu^*) = (1, \varepsilon(\delta^2 + 1)/(s - \delta\varepsilon))$ *is a bifurcation point of* $K(\omega, \mu, u)$ *defined by* (11.60).

This implies the existence of a nontrivial periodic solution for (11.58) $\Leftrightarrow$ (11.59) (with period near 2π) corresponding to μ near to μ^*.

[24] We obtain $s = -\varepsilon(\mu+\delta) + f'(0)$. It follows from this relation and (11.61) that

$$\varepsilon = s(\mu+\delta) - \delta f'(0) = s(\mu+\delta) - \delta(s + \varepsilon(\mu+\delta)),$$

which gives

$$\mu^* = \frac{\varepsilon}{s - \delta\varepsilon}(\delta^2 + 1).$$

11.11 Ljapunov's Center Theorem

It is well-known that trajectories of planar linear ordinary differential equations can be classified by eigenvalues of coefficient matrix. In particular, the equilibrium point is a center if all the eigenvalues are purely imaginary.

However, the situation becomes quite different in the case of nonlinear systems as illustrated by the following example.

Example 11.2 Consider a nonlinear ordinary differential equation

$$\begin{aligned} \dot{x} &= -y - x(x^2 + y^2), \\ \dot{y} &= x - y(x^2 + y^2). \end{aligned} \tag{11.62}$$

The matrix which gives a linear approximation of the right-hand side of (11.62) around the origin 0 is

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The eigenvalues of A are $\pm i$. However, trajectories of the equation (11.62) do not show the behaviors moving around a center. In fact, we obtain

$$\frac{d}{dt}\left[\frac{1}{2}(x^2 + y^2)\right] = x\dot{x} + y\dot{y} = -(x^2 + y^2)^2$$

for any solution $(x(t), y(t))$ of (11.62). Hence it follows that

$$x^2 + y^2 = \frac{1}{2t + C} \quad (C \text{ is constant}).$$

The trajectory is not a closed orbit. Thus the origin can not be a center.

Under what conditions does a center occur for a nonlinear differential equation? We next study this problem by making use of Hopf's bifurcation theorem.

Let Ω be an open set in $\mathbb{R}^n$ and Ω' be an open subset of Ω. Suppose that a function $f : \Omega \to \mathbb{R}^n$ is continuous and has partial derivatives. A continuous differentiable function $u : \Omega' \to \mathbb{R}$ is called a **first integral** of the differential equation

$$\dot{x} = f(x) \tag{11.63}$$

if

$$u(\varphi(t)) = \text{constant} \tag{11.64}$$

for any solution $\varphi(t)$ of (11.63) such that its trajectory is contained in Ω'. The condition (11.64) means that the value of $u(\varphi(t))$ depends upon the choice of $\varphi(t)$ but is independent of t.

Lemma 11.5 *A function u is a first integral of* (11.63) *if and only if*

$$\langle Du(x), f(x)\rangle = 0 \quad \textit{for all} \quad x \in \Omega'. \tag{11.65}$$

Proof Suppose that u is a first integral of (11.63). Let $\varphi(t, \xi)$ be a solution of (11.63) with the initial value $\xi \in \Omega'$; i.e. $\varphi(0, \xi) = \xi$. By definition of the first integral,

$$\frac{d}{dt} u(\varphi(t, \xi))|_{t=0} = \langle Du(\xi), f(\xi)\rangle = 0.$$

This relation holds good for any $\xi \in \Omega'$. Hence (11.65) follows.

Suppose conversely that (11.65) holds good. Let $\varphi(t)$ be a solution of (11.63), the trajectory of which is contained in Ω'. If we define a function $v(t)$ by

$$v(t) = u(\varphi(t)),$$

we obtain

$$\frac{d}{dt} v(t) = \langle Du(\varphi(t)), f(\varphi(t))\rangle = 0$$

by (11.65). Consequently, $v(t) = u(\varphi(t))$ =constant. □

Example 11.3 Let a function $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be differentiable. Consider a differential equation called the **Hamiltonian system**

$$\begin{aligned} \dot{x} &= -D_y H(x, y), \\ \dot{y} &= D_x H(x, y). \end{aligned} \tag{11.66}$$

Then H itself is a first integral of (11.66).

Lemma 11.6 *Let $f \in \mathfrak{C}^1(\mathbb{R}^n, \mathbb{R}^n)$. Suppose that $u \in \mathfrak{C}^1(\mathbb{R}^n, \mathbb{R})$ is a first integral of the differential equation.*

$$\dot{x} = f(x). \tag{11.67}$$

If x is a solution of

$$\dot{x} = f(x) + \mu Du(x), \qquad \mu \in \mathbb{R} \tag{11.68}$$

with period T, then x is also a solution of (11.67) *with period T.*

Proof There is nothing to prove in the case $\mu = 0$. So we may assume that $\mu \neq 0$. Let x be a T-periodic solutions of (11.68). If we define

$$v(t) = u(x(t)),$$

then we obtain

$$\begin{aligned}\dot{v}(t) &= \frac{d}{dt}u(x(t)) = \langle Du(x(t)), \dot{x}(t)\rangle \\ &= \langle Du(x(t)), f(x(t)) + \mu Du(x(t))\rangle \\ &\qquad \text{(by (11.68))} \\ &= \langle Du(x(t)), f(x(t))\rangle + \mu\|Du(x(t))\|^2.\end{aligned}$$

Since u is a first integral of (11.67), $\langle Du(x(t)), f(x(t))\rangle = 0$. Hence

$$\dot{v}(t) = \mu\|Du(x(t))\|^2. \tag{11.69}$$

If $\mu > 0$, we obtain

$$\dot{v} \geqq 0. \tag{11.70}$$

So $v(t)$ is nondecreasing. On the other hand,

$$v(0) = u(x(0)) = u(x(T)) = v(T) \tag{11.71}$$

by the periodicity of x. Hence $\dot{v}(t) = 0$ by (11.70) and (11.71). It follows from (11.69) that

$$\mu Du(x(t)) = 0.$$

We obtain the same result in the case of $\mu < 0$. Thus we conclude that x is a T-periodic solution of (11.67). □

Theorem 11.5 (Ljapunov) *Let $f \in \mathfrak{C}^2(\mathbb{R}^n, \mathbb{R}^n)$ satisfy $f(0) = 0$. Assume that $A = Df(0)$ satisfies the following two conditions:*

(i) *A is regular and $\pm i\omega^*(\omega^* > 0)$ are simple eigenvalues of A.*
(ii) *$ik\omega^*$ is not an eigenvalue of A for all $k \in \mathbb{Z}$ other than $k \neq \pm 1$.*

Assume further that $u \in \mathfrak{C}^3(\mathbb{R}^n, \mathbb{R})$ is a first integral of (11.63) and $D^2u(0) \equiv B$ is regular.

Then $(\omega^*, \mu^*) = (\omega^*, 0)$ is a bifurcation point of

$$\omega\dot{x} = f(x) + \mu Du(x)$$

in the space

$$\mathfrak{X} = \{x \in \mathfrak{C}^r(\mathbb{R}, \mathbb{R}^n) \mid x(t + 2\pi) = x(t) \quad \textit{for all} \;\; t\}, \;\; r \geqq 3.$$

Proof The following reasoning depends upon Theorem 11.3:

1° Define a function $\varphi \in \mathfrak{C}^2$ by

$$\varphi(\mu, x) = f(x) + \mu Du(x). \tag{11.72}$$

φ plays a role corresponding to f in Theorem 11.3.

2° Since $\langle Du(\xi), f(x)\rangle = 0$ $(x \in \mathbb{R}^n)$ by Lemma 11.5, we obtain

$$\langle D^2u(\xi)y, f(\xi)\rangle + \langle Du(\xi), Df(\xi)y\rangle = 0 \quad \text{for all} \quad \xi, y \in \mathbb{R}^n. \tag{11.73}$$

If $\xi = 0$, in particular,

$$\langle D^2u(0)y, f(0)\rangle + \langle Du(0), Df(0)y\rangle = 0 \quad \text{for all} \quad y \in \mathbb{R}^n. \tag{11.74}$$

The first term of (11.74) is zero, since $f(0) = 0$. Hence

$$\langle Du(0), Ay\rangle = \langle {}^tA Du(0), y\rangle = 0 \quad \text{for all} \quad y \in \mathbb{R}^n.$$

By the regularity of A, we must have $Du(0) = 0$. It follows from the definition (11.72) that

$$\varphi(\mu, 0) = f(0) + \mu Du(0) = 0.$$

3° Writing $A_\mu = D_x\varphi(\mu, 0)$, we have

$$A_\mu = D_x f(0) + \mu D^2u(0) = A + \mu B.$$

The following two conditions are satisfied for $\mu^* = 0$ by (i) and (ii):

(a) A_{μ^*} is regular, and $\pm i\omega^*$ are simple eigenvalues of A_{μ^*}.
(b) $ik\omega^*$ is not an eigenvalue of A_{μ^*} for all $k \in \mathbb{Z}$ other than $k = \pm 1$.

4° We are now going to verify the final condition of the Hopf theorem. Differentiating (11.73) with respect to ξ, we obtain at $\xi = 0$ that

$$\begin{aligned} &\langle D^3u(0)(y, z), f(0)\rangle + \langle Df(0)z, D^2u(0)y\rangle + \langle D^2u(0)z, Df(0)y\rangle \\ &\qquad\qquad + \langle Du(0), D^2f(0)(y, z)\rangle = 0 \quad \text{for all} \quad y, z \in \mathbb{R}^n. \end{aligned}$$

The first and the fourth terms on the left-hand side disappear, since $f(0) = 0$ and $Du(0) = 0$. Hence

$$\langle Az, By\rangle + \langle Ay, Bz\rangle = 0 \quad \text{for all} \quad y, z \in \mathbb{R}^n.$$

Since B is a symmetric matrix,

$$\langle z, {}^tABy\rangle + \langle z, BAy\rangle = \langle z, ({}^tAB + BA)y\rangle = 0 \quad \text{for all} \quad y, z \in \mathbb{R}^n$$

and so

$$ {}^tAB + BA = 0. \tag{11.75}$$

Applying a suitable transformation ($\pm i\omega^*$ are simple eigenvalues of A), we obtain

$$A = \begin{pmatrix} S & 0 \\ 0 & R \end{pmatrix},$$

where

$$S = \begin{pmatrix} 0 & -\omega^* \\ \omega^* & 0 \end{pmatrix}$$

and $\pm i\omega^*$ are not contained in the spectrum of R. (S is a (2×2)-matrix, and R is a $(n-2) \times (n-2)$-matrix.) On the other hand, let us decompose B as

$$B = \begin{pmatrix} U & W \\ {}^tW & V \end{pmatrix},$$

where U(resp. V) is a (2×2)-symmetric matrix (resp. $(n-2) \times (n-2)$-symmetric matrix) and W is a $2 \times (n-2)$-matrix. It follows from (11.75) that

$$SU = US, \tag{11.76}$$

$$SW = WR. \tag{11.77}$$

For instance, (11.76) can be verified as follows. Looking at the multiplication of (2×2)-submatrices in the upper left zones in the calculation

$${}^tAB + BA = \left(\begin{array}{cc|c} 0 & \omega^* & 0 \\ -\omega^* & 0 & \\ \hline 0 & & {}^tR \end{array}\right) \left(\begin{array}{c|c} U & W \\ \hline {}^tW & V \end{array}\right) + \left(\begin{array}{c|c} U & W \\ \hline {}^tW & V \end{array}\right) \left(\begin{array}{cc|c} 0 & -\omega^* & 0 \\ \omega^* & 0 & \\ \hline 0 & & R \end{array}\right) = 0,$$

we can easily observe that

$$-SU + US = 0,$$

which is equivalent to (11.76).

Write the symmetric matrix U as

$$U = \begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

By (11.76), we must have $b = 0$, since $\omega^* \neq 0$. Hence

$$U = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

Furthermore, the matrix W is zero. In order to see this, we decompose W as

$$W = \begin{pmatrix} X \\ Y \end{pmatrix},$$

where $X, Y \in \mathbb{R}^{n-2}$. By (11.77), we have

$$XR + \omega^* Y = 0, \tag{11.78}$$

$$YR - \omega^* X = 0. \tag{11.79}$$

It follows from (11.79) that

$$X = \frac{1}{\omega^*} YR. \tag{11.80}$$

Substituting (11.80) into (11.78), we have

$$\begin{aligned} \frac{1}{\omega^*} YR^2 + \omega^* Y &= \frac{1}{\omega^*} Y[R^2 + {\omega^*}^2 I] \\ &= \frac{1}{\omega^*} Y(R + i\omega^* I)(R - i\omega^* I) \qquad (11.81) \\ &= 0. \end{aligned}$$

$Y = 0$ follows from (11.81), since $\pm i\omega^*$ are not eigenvalues of R. Hence $X = 0$ by (11.80). Thus we obtain $W = 0$. That is, B has the following form:

$$B = \left(\begin{array}{cc|c} a & 0 & \mathbf{0} \\ 0 & a & \\ \hline \mathbf{0} & & V \end{array}\right).$$

$a \neq 0$ by the regularity of B.

Consequently, $A + \mu B$ is expressed as

$$A + \mu B = \left(\begin{array}{cc|c} \mu a & -\omega^* & \mathbf{0} \\ \omega^* & \mu a & \\ \hline \mathbf{0} & & R + \mu V \end{array}\right).$$

Find eigenvalues $\lambda(\mu)$ of $A + \mu B$ with parameter μ, which are equal to $\pm i\omega^*$ when $\mu = 0$. Then

$$\lambda(\mu) = \mu a \pm i\omega^* \quad (a \neq 0).$$

We have verified all the conditions of the Hopf theorem. □

The above theorem tells us that the differential equation

$$\omega_n \dot{x} = f(x) + \mu_n Du(x), \quad n = 1, 2, \cdots$$

has a nontrivial solution $x_n(t)$ with period 2π and $x_n \to 0$ (in $\mathfrak{C}^r$) as $\omega_n \to \omega^*$ and $\mu_n \to 0$.

Specifying $\omega^* = 1$, we obtain a solution x_n with period $T_n = 2\pi/\omega_n$ of

$$\dot{x} = f(x) + \mu_n Du(x), \quad n = 1, 2, \cdots \tag{11.82}$$

such that $x_n \to 0$ as $\mu_n \to 0$. According to Lemma 11.6, however, x_n is a solution not only of (11.82) but also of

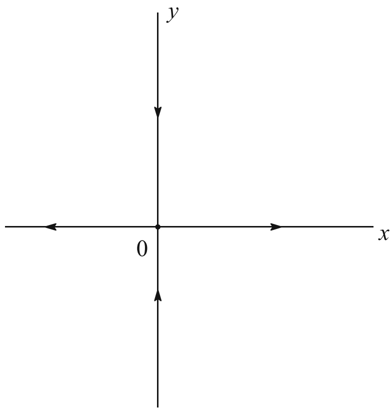
$$\dot{x} = f(x).$$

That is, there exists a sequence of nontrivial solutions with period $2\pi/\omega_n$ such that $x_n \to 0$ as $n \to \infty$.

Remark 11.7

1° The domain of f, which determines the differential equation, is assumed to be $\mathbb{R}^n$ in Theorem 11.5. It is clear that $\mathbb{R}^n$ can be substituted by any open set Ω containing 0. In this case, the domain of the first integral may be Ω. That is, what we require is "$u(\varphi(t)) =$ constant for any solution $\varphi(t)$ of (11.67) such that its trajectory is contained in Ω".

Fig. 11.5 Trajectories of Volterra–Lotka equation (1)



2° In Theorem 11.5, the origin 0 is considered as an equilibrium point of (11.67) and $A = Df(0)$. If an equilibrium point x_0 is nonzero, then the problem can be reduced to the above case by changing the variable $v = x - x_0$.

Example of Application (Volterra–Lotka Differential Equation) Consider a planar system of ordinary differential equations

$$\begin{aligned} \dot{x} &= x(a - by), \\ \dot{y} &= y(cx - d), \end{aligned} \tag{11.83}$$

where a, b, c and d are positive constants ($\dot{x}$ and $\dot{y}$ denote the derivatives with respect to the time variable t). $x(t)$ and $y(t)$ may be interpreted, for instance, as the quantity of corn at time t and the population at time t. The life of a man is supported by corn. (11.83) means that the growth ratios, $\dot{x}/x$ and $\dot{y}/y$, are equal to $(a - by)$ and $(cx - d)$, respectively. Expressing the right-hand side of (11.83) as a vector $f(x, y)$, we can rewrite (11.83) in the form

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = f(x, y).$$

The equilibrium points of (11.83) are $(0, 0)$ and $(d/c, a/b) \equiv (x_0, y_0)$. The trajectories with initial values on axes are depicted in Fig. 11.5.

Adapting to the interpretation as above, we restrict our attention to $x \geqq 0$, $y \geqq 0$. When $x = 0$ and $y \geqq 0$, $x(a - by)|_{x=0} = 0$. When $x \geqq 0$ and $y = 0$, $y(cx - d)|_{y=0} = 0$. Hence a nonnegative solution $(x(t), y(t))$ of (11.83) exists for any nonnegative initial value (α_0, β_0).[25] We denote by Ω the interior of the first orthant;

[25] See Yamaguti [22], p. 25.

i.e. $\Omega = \{(x, y) \in \mathbb{R}^2 | x > 0, y > 0\}$. Any solution with initial value in Ω remains in Ω.

Evaluating $Df(x_0, y_0)$ at the equilibrium point (x_0, y_0), we obtain

$$Df(x_0, y_0) = \begin{pmatrix} 0 & -bd/c \\ ca/b & 0 \end{pmatrix}.$$

The eigenvalues are $\pm i\sqrt{ad} \equiv i\omega_0$. Both of them are simple.

We next look for a first integral. By a simple manipulation of (11.83), we obtain

$$acx - bdy = c\dot{x} + b\dot{y}, \tag{11.84}$$

$$d\frac{\dot{x}}{x} - (acx - bdy) + a\frac{\dot{y}}{y} = 0. \tag{11.85}$$

Substituting (11.85) into (11.84), we further obtain

$$c\dot{x} - d\frac{\dot{x}}{x} + b\dot{y} - a\frac{\dot{y}}{y} = 0. \tag{11.86}$$

Any solution of (11.83) must satisfy (11.86). Integrating (11.86) with respect to a solution $(x(t), y(t))$ which remains in Ω, we get

$$cx(t) - d\log x(t) + by(t) - a\log y(t) = \text{const}.$$

Define a function $u : \Omega \to \mathbb{R}$ by

$$u(x, y) = cx - d\log x + by - a\log y.$$

Then it satisfies

$$\frac{d}{dt}u(x(t), y(t)) = 0.$$

That is, $u(x, y)$ is a first integral of (11.83).

Finally,

$$D^2u(x_0, y_0) = \begin{pmatrix} d/x_0^2 & 0 \\ 0 & a/y_0^2 \end{pmatrix}$$

is regular.

Hence, according to Theorem 11.5, (x_0, y_0) must be a center. (cf. Fig. 11.6.)

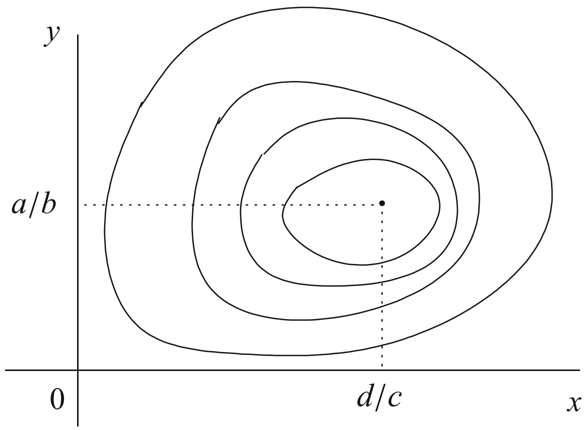


Fig. 11.6 Trajectories of Volterra–Lotka equation (2)

References

1. Ambrosetti, A., Prodi, G.: A Primer of Nonlinear Analysis. Cambridge University Press, Cambridge (1993)
2. Carleson, L.: On the convergence and growth of partial sums of Fourier series. Acta Math. **116**, 135–157 (1966)
3. Chang, W.W., Smyth, D.J.: The existence and persistence of cycles in a non-linear model of Kaldor's 1940 model re-examined. Rev. Econ. Stud. **38**, 37–44 (1971)
4. Crandall, M.G., Rabinowitz, P.H.: The Hopf bifurcation theorem. MRC Technical Summary Report, No. 1604, University of Wisconsin Math. Research Center (1976)
5. Evans, L.: Partial Differential Equations. American Mathematical Society, Providence (1998)
6. Frisch, R.: Propagation problems and inpulse problems in dynamic economics. In: Economic Essays in Honor of Gustav Cassel. Allen and Unwin, London (1933)
7. Hicks, J.R.: A Contribution to the Theory of the Trade Cycle. Oxford University Press, London (1950)
8. Hopf, E.: Abzweigung einer periodishcen Lösung von einer stationären Lösung eines Differentialsystems. Ber. Math. Phys. Sächsische Akademie der Wissenschafters, Leipzig **94**, 1–22 (1942)
9. Hunt, R.A.: On the convergence of Fourier series. In: Proceedings of the Conference on Orthogonal Expanisions and their Continuous Analogues, Southern Illinois University Press, Carbondale, pp. 234–255 (1968)
10. Kaldor, N.: A model of the trade cycle. Econ. J. **50**, 78–92 (1940)
11. Keynes, J.M.: The General Theory of Employment, Interest and Money. Macmillan, London (1936)
12. Maruyama, T.: Suri-keizaigaku no Hoho (Methods in Mathematical Economics). Sobunsha, Tokyo (1995) (Originally published in Japanese)
13. Maruyama, T.: Existence of periodic solutions for Kaldorian business fluctuations. Contemp. Math. **514**, 189–197 (2010)
14. Maruyama, T.: On the Fourier analysis approach to the Hopf bifurcation theorem. Adv. Math. Econ. **15**, 41–65 (2011)
15. Maruyama, T.: Shinko Keizai Genron (Principles of Economics). Iwanami Shoten, Tokyo (2013) (Originally published in Japanese)
16. Masuda, K. (ed.): Ohyo-kaiseki Handbook (Handbook of Applied Analysis). Springer, Tokyo (2010) (Originally published in Japanese)

17. Samuelson, P.A.: Interaction between the multiplier analysis and the principle of acceleration. Rev. Econ. Stud. **21**, 75–78 (1939)
18. Samuelson, P.A.: A synthesis of the principle of acceleration and the multiplier. J. Polit. Econ. **47**, 786–797 (1939)
19. Shinasi, G.J.: A nonlinear dynamic model of short run fluctuations. Rev. Econ. Stud. **48**, 649–656 (1981)
20. Stromberg, K.R.: An Introduction to Classical Real Analysis. American Mathematical Society, Providence (1981)
21. Takagi, T.: Kaiseki Gairon (Treatise on Analysis), 3rd edn. Iwanami Shoten, Tokyo (1961) (Originally published in Japanese)
22. Yamaguti, M: Hisenkei Gensho no Sugaku (Mathematics of Nonlinear Phenomena). Asakura Shoten, Tokyo (1972) (Originally published in Japanese)
23. Yasui, T.: Jireishindo to Keiki Junkan. In:Kinko-bunseki no Kihon Mondai. (Self-oscillations and trade cycles. In: Fundamental Problems in Equilibrium Analysis.) Iwanami Shoten, Tokyo (1965) (Originally published in Japanese)

Appendix A
Exponential Function $e^{i\theta}$

The imaginary exponential function $e^{i\theta}$ incessantly appears in Fourier analysis. Everybody knows that $e^{i\theta}$ is a complex number, the real and the imaginary parts of which are $\cos\theta$ and $\sin\theta$, respectively. If we take it for granted, the trajectory of $e^{i\theta}$ is exactly the unit circle in the complex plane $\mathbb{C}$. However, it is not so easy for us to give a coherent exposition of this concept. We would like to provide an idea on how to look at this concept systematically in order to promote more rigorous reasoning.[1]

A.1 Complex Exponential Function

We define a function e^z of a complex variable z by

$$e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n. \tag{A.1}$$

Since the convergence radius of the power series on the right-hand side of (A.1) is infinite, (A.1) is absolutely convergent. The derivative is

$$\frac{d}{dz} e^z = e^z. \tag{A.2}$$

It is easy to see that

$$e^{z+z'} = e^z \cdot e^{z'} \tag{A.3}$$

[1] I am indebted to Cartan [1] and Takagi [5] very much. See van der Waerden [6] for basic knowledge on group theory.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8

for $z, z' \in \mathbb{C}$. In fact, it is verified by[2]

$$\begin{aligned} e^z \cdot e^{z'} &= \sum_{n=0}^{\infty} \frac{1}{n!} z^n \cdot \sum_{n=0}^{\infty} \frac{1}{n!} z'^n \\ &= \sum_{n=0}^{\infty} \sum_{p=0}^{n} \frac{1}{p!(n-p)!} z^p z'^{(n-p)} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (z+z')^n \\ &= e^{z+z'}. \end{aligned}$$

Since $e^z \cdot e^{-z} = 1$ by (A.3), we observe that $e^z \neq 0$ for any $z \in \mathbb{C}$.

---

[2]In general, let $\sum_{n=0}^{\infty} u_n$ and $\sum_{n=0}^{\infty} v_n$ be absolutely convergent series. Define

$$w_n = \sum_{p=0}^{n} u_p v_{n-p}.$$

Then $\sum_{n=0}^{\infty} w_n$ is also absolutely convergent and

$$\sum_{n=0}^{\infty} w_n = \left(\sum_{n=0}^{\infty} u_n \right) \left(\sum_{n=0}^{\infty} v_n \right). \tag{†}$$

We can prove this relation as follows. Defining

$$\alpha_p = \sum_{n=p}^{\infty} |u_n|, \quad \beta_q = \sum_{n=q}^{\infty} |v_n|,$$

we obtain

$$\sum_{n=0}^{\infty} |w_n| \leqq \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} |u_p| \, |v_q| = \alpha_0 \cdot \beta_0.$$

For $m \geqq 2n$, the absolute value of

$$\sum_{k=0}^{m} w_k - \left(\sum_{k=0}^{n} u_k \right) \left(\sum_{k=0}^{n} v_k \right)$$

is bounded by some finite sum of $|u_p| \cdot |v_q|$ (either p or q is greater than n). This sum is, in turn, bounded by $\alpha_0 \beta_{n+1} + \beta_0 \alpha_{n+1}$. Hence it converges to 0 as $n \to \infty$. Thus we verify (†).

If we write $z \in \mathbb{C}$ as $z = x + iy$ $(x, y \in \mathbb{R})$, then

$$e^z = e^x \cdot e^{iy}.$$

The study of e^z is eventually reduced to examining the properties of the real exponential function e^x and of the imaginary exponential function e^{iy}.

It does not seem necessary here to give any expositions concerning the properties of e^x and the logarithmic function as its inverse.[3] So we exclusively concentrate on the imaginary exponential function e^{iy} $(y \in \mathbb{R})$.

A.2 Imaginary Exponential Function

By definition of $e^{i\theta}$ given by (A.1) in the preceding section, $e^{-i\theta} = \overline{e^{i\theta}} (\theta \in \mathbb{R})$. Hence

$$e^{i\theta} \cdot e^{-i\theta} = |e^{i\theta}|^2 = 1,$$

and so

$$|e^{i\theta}| = 1 \quad \text{for all} \quad \theta \in \mathbb{R}. \tag{A.4}$$

The set $U = \{z \in \mathbb{C} \,|\, |z| = 1\}$ forms a unit circle of the complex plane $\mathbb{C}$, and is a group with respect to the multiplication. A function $\varphi : \mathbb{R} \to U$ defined by

$$\varphi(\theta) = e^{i\theta}, \quad \theta \in \mathbb{R} \tag{A.5}$$

is a (group) homomorphism of an additive group $\mathbb{R}$ into the multiplicative group U.

In order to examine φ closely, we divide $e^{i\theta}$ into the real part and the imaginary part. We define $\cos\theta$ as the real part of $e^{i\theta}$ and $\sin\theta$ as its imaginary part; i.e.

$$e^{i\theta} = \cos\theta + i\sin\theta. \tag{A.6}$$

Then $\cos\theta$ and $\sin\theta$ are given by the power series

$$\begin{aligned} \cos\theta &= 1 - \frac{1}{2}\theta^2 + \cdots + \frac{(-1)^n}{(2n)!}\theta^{2n} + \cdots, \\ \sin\theta &= \theta - \frac{1}{3!}\theta^3 + \cdots + \frac{(-1)^n}{(2n+1)!}\theta^{2n+1} + \cdots, \end{aligned} \tag{A.7}$$

[3] cf. Stromberg [4] pp. 134–136, Takagi [5] pp. 189–193.

the convergence radii of which are both infinite.[4] The derivatives are evaluated as

$$\frac{d}{d\theta}\cos\theta = -\sin\theta, \quad \frac{d}{d\theta}\sin\theta = \cos\theta. \tag{A.8}$$

Since $\cos 0 = 1$ and $\cos\theta$ is continuous, there exists some $\theta_0 > 0$ such that

$$\cos\theta > 0 \quad \text{for all} \quad \theta \in [0, \theta_0]. \tag{A.9}$$

Hence, by (A.8), $\sin\theta$ is strictly increasing on $[0, \theta_0]$ and $\sin\theta_0$ is positive, since $\sin 0 = 0$.

We now show that there exists some $\theta > 0$ such that $\cos\theta = 0$. Suppose that $\cos\theta > 0$ on $[\theta_0, \theta_1]$. We know that

$$\cos\theta_1 - \cos\theta_0 = -\int_{\theta_0}^{\theta_1} \sin\theta d\theta. \tag{A.10}$$

$\cos\theta > 0$ on the interval of integration. So $\sin\theta$ is increasing there and $\sin\theta > \sin\theta_0 \equiv a$. Hence

$$\int_{\theta_0}^{\theta_1} \sin\theta d\theta > a(\theta_1 - \theta_0).$$

---

[4]Consider a series $\sum_{n=1}^{\infty} a_n(x)$ of functions $a_n : \mathbb{R} \to \mathbb{R}$.

1° If there exists a sequence $\{c_n\}$ of positive numbers such that

$$|a_n(x)| \leqq c_n \ ; \ n = 0, 1, 2, \cdots, \text{ and } \sum_{n=1}^{\infty} c_n < \infty,$$

on some interval I, then $\sum_{n=1}^{\infty} a_n(x)$ uniformly converges on I.

2° If each $a_n(x)$ is continuously differentiable, $\sum_{n=1}^{\infty} a_n(x)$ converges, and $\sum_{n=1}^{\infty} a_n'(x)$ is uniformly convergent, then $\sum_{n=1}^{\infty} a_n(x)$ is also continuously differentiable and

$$\frac{d}{dx}\sum_{n=1}^{\infty} a_n(x) = \sum_{n=1}^{\infty} \frac{d}{dx} a_n(x).$$

(cf. Stromberg [4] pp. 213–215, Takagi [5] pp. 155–159.)

By (A.10) and $\cos\theta_1 > 0$, we obtain

$$\theta_1 - \theta_0 < \frac{1}{a}\cos\theta_0.$$

Thus we must conclude that there exists some point $\theta \in [\theta_0, \theta_0 + (1/a)\cos\theta_0]$ such that $\cos\theta = 0$.

We define $\pi/2$ as the minimum positive number which satisfies $\cos\theta = 0$. This is the definition of π.

On the interval $[0, \pi/2]$, $\cos\theta$ is monotonically decreasing from 1 to 0, and $\sin\theta$ is increasing from 0 to 1. That is, φ is a bijection of $[0, \pi/2]$ onto

$$U_1 = \{(u, v) \in U \mid u \geqq 0, v \geqq 0\}.$$

The inverse of φ is also continuous, since $[0, \pi/2]$ is compact. Thus $[0, \pi/2]$ is homeomorphic to U_1 through φ.

Similarly, φ gives a homeomorphic relation between each of the following interval and the corresponding quarter of the unit circle:

$$\left[\frac{\pi}{2}, \pi\right] \longleftrightarrow U_2 = \{(u, z) \in U \mid u \leqq 0, v \geqq 0\},$$
$$\left[\pi, \frac{3}{2}\pi\right] \longleftrightarrow U_3 = \{(u, z) \in U \mid u \leqq 0, v \leqq 0\},$$
$$\left[\frac{3}{2}\pi, 2\pi\right) \longleftrightarrow U_4 = \{(u, z) \in U \mid u \geqq 0, v < 0\}.$$

For instance, suppose θ moves on $[\pi/2, \pi]$. Write $\theta = \pi/2 + \xi$. When θ moves from $\pi/2$ to π, ξ moves from 0 to $\pi/2$. It is easy to see that[5]

$$\begin{aligned} e^{i\theta} &= e^{i(\pi/2+\xi)} = e^{i\cdot\pi/2} \cdot e^{i\xi} \\ &= \left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right)e^{i\xi} = -\sin\xi + i\cos\xi. \end{aligned}$$

When θ moves on $[\pi/2, \pi]$ and so ξ moves on $[0, \pi/2]$, the real part of $e^{i\theta}$ decreases from 0 to -1, and the imaginary part also decreases from 1 to 0. Therefore the trajectory of $\varphi(\theta) = e^{i\theta}$ ($\theta \in [\pi/2, \pi]$) is the quarter U_2 of the unit circle. We can apply similar reasonings to other pairs.

Hence $[0, 2\pi)$ and U are homeomorphic to each other through φ.

Of course, U is also homeomorphic to $[-\pi, \pi)$ instead of $[0, 2\pi)$.

[5]The second equality is justified because the function φ defined by (A.5) is a homomorphism.

Since the unit element of U is 1, the kernel $\mathrm{Ker}\varphi = \{\theta \in \mathbb{R} | \varphi(\theta) = 1\}$ of the homomorphism φ is the set of multiples of 2π; i.e.

$$\mathrm{Ker}\varphi = 2\pi\mathbb{Z} = \{2\pi n | n \in \mathbb{Z}\}. \tag{A.11}$$

A.3 Torus $\mathbb{R}/2\pi\mathbb{Z}$

$2\pi\mathbb{Z} = \{2\pi n | n \in \mathbb{Z}\}$ is a subgroup of the additive group $\mathbb{R}$. The factor group $\mathbb{R}/2\pi\mathbb{Z}$ of $\mathbb{R}$ modulo $2\pi\mathbb{Z}$ is called the (one dimensional) **torus**, which is denoted by $\mathbb{T}$. $\mathbb{T}$ and the unit circle U is isomorphic through φ defined in the preceding section: i.e.

$$\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z} \cong U. \tag{A.12}$$

This is justified by the isomorphism theorem of groups and (A.11) in the preceding section.

$\mathbb{T}$ endowed with the quotient topology is a Hausdorff topological space. Let $\xi : \mathbb{R} \to \mathbb{T}$ be a canonical mapping; i.e. a function which maps each $\theta \in \mathbb{R}$ to its residue class $\theta + 2\pi\mathbb{Z}$. By the definition of the quotient topology, ξ is continuous. Since $\mathbb{T}$ is the image of $[0, 2\pi]$ by ξ, $\mathbb{T}$ is compact in the quotient topology.

If we define another function $\tilde{\varphi} : \mathbb{T} \to U$ by

$$\tilde{\varphi}(\theta + 2\pi\mathbb{Z}) = \varphi(\theta), \quad \theta \in \mathbb{R}, \tag{A.13}$$

$\tilde{\varphi}$ is a continuous bijection.[6]

[6] The continuity of $\tilde{\varphi}$ can be proved as follows. Let Σ be an open neighborhood of $\tilde{\varphi}(\theta_0 + 2\pi\mathbb{Z}) = \varphi(\theta_0)$. By the continuity of φ, there exists an open set $\Theta (\subset \mathbb{R})$ containing θ_0 such that

$$\varphi(\theta) \in \Sigma \quad \text{for all} \quad \theta \in \Theta.$$

Define a set Θ' by

$$\Theta' = \{\theta + 2\pi z \,|\, \theta \in \Theta, z \in \mathbb{Z}\} = \bigcup_{z \in \mathbb{Z}} (\Theta + 2\pi z).$$

Θ' is an open set in $\mathbb{R}$ which contains θ_0 and satisfies

$$\varphi(\theta) \in \Sigma \quad \text{for all} \quad \theta \in \Theta'.$$

Let Γ be the set of residue classes containing elements of Θ'; i.e.

$$\Gamma = \{\theta + 2\pi\mathbb{Z} \,|\, \theta \in \Theta'\}.$$

Consequently, $\tilde{\varphi}$ is a homeomorphism between $\mathbb{T}$ and U. (Note that U is compact.)

We conclude that the torus $\mathbb{T}$ and the unit circle U are isomorphic as groups and homeomorphic as topological spaces.

$\tilde{\varphi}^{-1}(u) \in \mathbb{T}$ is uniquely determined for each $u \in U$; that is, a unique real number modulo $2\pi\mathbb{Z}$ is determined. We call this real number the **argument** of u and denote it by $arg\ u$.

For any complex number $z \neq 0$, we define its argument by

$$arg\, z = arg\left(\frac{z}{|z|}\right). \tag{A.14}$$

(Since $z/|z| \in U$, the right-hand side has already been defined above.) We must note that $arg\, z$ is defined at most modulo $2\pi\mathbb{Z}$. Making use of this notation, z can be expressed as

$$z = |z|e^{i\, arg\, z}. \tag{A.15}$$

Given a complex number $t \in \mathbb{C}$ ($t \neq 0$), the complex number z which satisfies $e^z = t$ is given by

$$\log|t| + i\, arg\, t. \tag{A.16}$$

This complex number is the definition of $\log t$, i.e. $\log t = \log|t| + i\, arg\, t$. Since $arg\, t$ is determined only modulo $2\pi\mathbb{Z}$, $\log t$ is also defined only modulo $2\pi i \cdot \mathbb{Z}$.

Let D be a region (open connected set) in $\mathbb{C} \setminus \{0\}$. A continuous function $f : D \to \mathbb{C}$ which satisfies $e^{f(t)} = t$ ($t \in D$) (that is, $f(t)$ is a value of $\log t$) is called a **branch** of $\log t$.

If $f(t)$ is a branch of $\log t$, then its derivative $f'(t)$ exists and $f'(t) = 1/t$.

Proof For $h \neq 0$ with sufficiently small $|h|$ (and so $t + h \in D$), we have

$$\frac{f(t+h) - f(t)}{h} = \frac{f(t+h) - f(t)}{e^{f(t+h)} - e^{f(t)}}.$$

As $h \to 0$, this number approaches to the inverse of the limit of $(e^{z'} - e^z)/(z' - z)$ as $z' \to z = f(t)$. Hence $f'(t)$ is the inverse of the derivative of e^z at $z = f(t)$, that is, $e^{-f(t)} = 1/t$. □

Then $\theta_0 + 2\pi\mathbb{Z} \in \Gamma$ and $\xi^{-1}(\Gamma) = \Theta'$. Hence, by the definition of the quotient topology, $\Gamma (\subset \mathbb{T})$ is open in $\mathbb{T}$. It also holds good that

$$\tilde{\varphi}(\theta + 2\pi\mathbb{Z}) = \varphi(\theta) \in \Sigma \quad \text{for all} \quad \theta \in \Theta' \quad (\text{i.e. for all } \theta + 2\pi\mathbb{Z} \in \Gamma).$$

This proves the continuity of $\tilde{\varphi}$.

A.4 A Homomorphism of $\mathbb{R}$ into U

We define a function $\psi_x : \mathbb{R} \to U$ for a fixed $x \in \mathbb{R}$ by

$$\psi_x(\theta) = e^{ix\theta}.$$

ψ_x is a homomorphism between the two groups. How about the converse? That is, is any homomorphism $\eta : \mathbb{R} \to U$ represented in the form $\eta(\theta) = e^{ix\theta}$ for some $x \in \mathbb{R}$?[7]

Since η is continuous and $\eta(0) = 1$, there exists some $\delta > 0$ such that

$$\int_0^{\delta} \eta(t)dt = \alpha \neq 0. \tag{A.17}$$

Furthermore,

$$\begin{aligned} \alpha\eta(\theta) = \eta(\theta) \int_0^{\delta} \eta(t)dt &= \int_0^{\delta} \eta(\theta + t)dt \\ &= \int_{\theta}^{\theta+\delta} \eta(t)dt \quad \text{(changing variables)} \end{aligned} \tag{A.18}$$

because η is a homomorphism. Although we assume only the continuity of η, η actually turns out to be differentiable by (A.18).[8]

$$\begin{aligned} \alpha\eta'(\theta) &= \int_0^{\delta} \eta'(\theta + t)dt = \eta(\theta + \delta) - \eta(\theta) \\ &= \eta(\theta)\eta(\delta) - \eta(\theta) = \eta(\theta)(\eta(\delta) - 1). \end{aligned}$$

Hence

$$\eta'(\theta) = \eta(\theta) \cdot \frac{\eta(\delta) - 1}{\alpha}.$$

Since $(\eta(\delta) - 1)/\alpha$ on the right-hand side is exactly equal to $\eta'(0)$,[9] we have

$$\eta'(\theta) = x\eta(\theta), \quad x = \eta'(0). \tag{A.19}$$

[7] We owe this to Rudin [3] pp. 12–13.

[8] Maruyama [2] pp. 381–382.

[9]

$$\alpha\eta'(0) = \int_o^{\delta} \eta'(t)dt = \eta(\delta) - \eta(0) = \eta(\delta) - 1.$$

Solving this differential equation with the initial value $\eta(0) = 1$, we obtain

$$\eta(\theta) = e^{ix\theta} \tag{A.20}$$

by the differentiation formula of a branch of $\log t$.[10]

Thus we conclude that any continuous homomorphism $\psi : \mathbb{R} \to U$ can be expressed in the form $\psi(\theta) = e^{i\theta x}$ for a uniquely determined $x \in \mathbb{R}$.

A.5 Functions and σ-Fields on the Torus

We denote by $\mathfrak{F}(\mathbb{T}, \mathbb{C})$ the set of complex-valued functions on $\mathbb{T}$ and by $\mathfrak{F}_{2\pi}(\mathbb{R}, \mathbb{C})$ the set of complex-valued functions on $\mathbb{R}$ which are 2π-periodic. Then there is a one-to-one correspondence between these two sets. Define an operator $T : \mathfrak{F}(\mathbb{T}, \mathbb{C}) \to \mathfrak{F}_{2\pi}(\mathbb{R}, \mathbb{C})$ by

$$(Tf)(\theta) = f(\xi(\theta)) = f(\theta + 2\pi\mathbb{Z}), \quad f \in \mathfrak{F}(\mathbb{T}, \mathbb{C}).$$

The kernel Ker$T$ of T is $\{0\}$. Hence $\mathfrak{F}(\mathbb{T}, \mathbb{C})$ and $\mathfrak{F}_{2\pi}(\mathbb{R}, \mathbb{C})$ are isomorphic as linear spaces. (ξ is the canonical mapping which maps each θ to its residue class $\theta + 2\pi\mathbb{Z}$.)

We denote by $\mathfrak{C}(\mathbb{T}, \mathbb{C})$ the space of complex-valued continuous functions on $\mathbb{T}$, and by $\mathfrak{C}^b(\mathbb{R}, \mathbb{C})$ the space of bounded complex-valued continuous functions on $\mathbb{R}$. Both of them are Banach spaces with the uniform convergence norm. If we define a linear operator $T : \mathfrak{C}(\mathbb{T}, \mathbb{C}) \to \mathfrak{C}^b(\mathbb{R}, \mathbb{C})$ by

$$(Tf)(\theta) = f(\xi(\theta)), \quad f \in \mathfrak{C}(\mathbb{T}, \mathbb{C}),$$

Tf is a continuous function with period 2π.

So if we denote by $\mathfrak{C}^b_{2\pi}(\mathbb{R}, \mathbb{C})$ the space of bounded complex-valued continuous functions with period 2π, $\mathfrak{C}(\mathbb{T}, \mathbb{C})$ is isomorphic to $\mathfrak{C}^b_{2\pi}(\mathbb{R}, \mathbb{C})$ as Banach spaces.

It is also clear that $\mathfrak{C}(\mathbb{T}, \mathbb{C})$ is also isomorphic to $\{f \in \mathfrak{C}([-\pi, \pi), \mathbb{C}) \mid \lim_{\theta \uparrow \pi} f(\theta) = f(-\pi)\}$.

Since $\mathbb{T}$ and U are homeomorphic, the measurable spaces $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ and $(U, \mathcal{B}(U))$ are isomorphic, where $\mathcal{B}(\cdot)$ denotes the Borel σ-field. Although $[-\pi, \pi)$ and U are not homeomorphic, $(U, \mathcal{B}(U))$ and $([-\pi, \pi), \mathcal{B}([-\pi, \pi)))$ are isomorphic. So we can consider $[-\pi, \pi)$ instead of U or $\mathbb{T}$ as a measurable space.

[10] $\log \eta(\theta) = x\theta +$ constant. But the constant must be 0 since $\eta(0) = 1$. On the other hand, $\log \eta(\theta) = \log|\eta(\theta)| + i\, arg\eta(\theta) = i\, arg\eta(\theta)$. $i\, arg\eta(\theta) = \theta x$. If we express x as $x = a + ib$, we obtain $i\, arg\eta(\theta) = ib(\theta)$. Hence $\eta(\theta) = e^{i\, arg\eta(\theta)} = e^{ib\theta}$. This proves (A.20).

References

1. Cartan, H.: Théorie élémentaires des fonctions analytiques d'une ou plusieurs variables complexes. Hermann, Paris (1961) (English edn.) Elementary Theory of Analytic Functions of One or Several Complex Variables. Addison Wesley, Reading (1963)
2. Maruyama, T.: Suri-keizaigaku no Hoho (Methods in Mathematical Economics). Sobunsha, Tokyo (1995) (Originally published in Japanese)
3. Rudin, W.: Fourier Analysis on Groups. Interscience, New York (1962)
4. Stromberg, K.R.: An Introduction to Classical Real Analysis. American Mathematical Society, Providence (1981)
5. Takagi, T.: Kaiseki Gairon (Treatise on Analysis), 3rd edn. Iwanami Shoten, Tokyo (1961) (Originally published in Japanese)
6. van der Waerden, B.L.: Algebra, nouvelle édn. Unger, New York (1970)

Appendix B
Topics from Functional Analysis

We have studied the theory of Fourier transforms of distributions, and then applied it to the theory of almost periodic functions. Spaces like $\mathfrak{D}(\Omega)$, $\mathfrak{D}(\Omega)'$, $\mathfrak{S}(\Omega)$ and $\mathfrak{S}(\Omega)'$, which play basic roles in the theory in question, are not normed spaces. So more abstract theories of locally convex topological vector spaces are indispensable. However, it seems to be a digression for us to describe the entire landscape of the theory in this book. We ask our readers to read through some reliable textbooks on functional analysis for systematic knowledge.[1] In this appendix, we are going to give a supplementary exposition of a couple of topics for the sake of readers' convenience; on the concept of inductive limit and on the dual space of a locally convex topological vector space.

The topology of $\mathfrak{D}(\Omega)$ which appears in Appendix C is determined by the inductive limit and its dual $\mathfrak{D}(\Omega)'$ is the space of distributions. Hence the basic theories explained in this appendix are more or less required in order to study the distributions at a rigorous level.

B.1 Inductive Limit Topology

A distribution is a continuous linear functional on a topological vector space consisting of certain smooth functions. The topology of its domain is defined by the inductive limit. So we start by explaining this concept.

[1]For general theories of locally convex spaces, Bourbaki [1], and Grothendieck [2] are most reliable works. We are much indebted to Schwartz [6] Chap. III and Schwartz [5] Chap. IV as well as Appendix, Part II for the theories of inductive limit and barreled spaces, which are discussed in relation to the theory of distributions.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8

Let $\mathfrak{X}$ be a vector space and $\{\mathfrak{X}_\alpha\}_{\alpha\in A}$ a family of subspaces of $\mathfrak{X}$ such that $\mathfrak{X} = \bigcup_{\alpha\in A} \mathfrak{X}_\alpha$. We suppose that each $\mathfrak{X}_\alpha$ is a locally convex Hausdorff topological vector (or linear) space (abbreviated as LCHTVS) and its topology is denoted by $\mathcal{T}_\alpha$. We also denote by $\mathcal{B}_0$ the set of all the balanced convex subsets U such that

$$U \bigcap \mathfrak{X}_\alpha \in \mathcal{N}(0, \mathcal{T}_\alpha)$$

for all $\alpha \in A$, where $\mathcal{N}(0, \mathcal{T}_\alpha)$ is a perfect family of neighborhoods of $0 \in \mathfrak{X}_\alpha$ with respect to $\mathcal{T}_\alpha$.

In fact, such a neighborhood U of 0 certainly exists. Let V_α be a balanced convex neighborhood of $0 \in \mathfrak{X}_\alpha$. Then $U \equiv \mathrm{co}\bigcup_\alpha V_\alpha$ certainly satisfies the required properties.[2]

If we define

$$\mathcal{B} = \{x + U | x \in \mathfrak{X},\ U \in \mathcal{B}_0\},$$

a topology $\mathcal{T}$ of $\mathfrak{X}$, the base of which is $\mathcal{B}$, is determined. By the construction of $\mathcal{B}_0$, $\mathfrak{X}$ endowed with the topology $\mathcal{T}$ is a locally convex topological vector space (not necessarily Hausdorff). This topology $\mathcal{T}$ is called the **inductive limit topology**. The relative topology of $\mathfrak{X}_\alpha$ induced by $\mathcal{T}$ is weaker than $\mathcal{T}_\alpha$. The inductive limit topology can be characterized in several ways. Consider a topology $\mathcal{T}'$ on $\mathfrak{X}$ which satisfies the following conditions:

(a) $\mathfrak{X}$ endowed with $\mathcal{T}'$ is a locally convex topological vector space.
(b) The relative topology $\mathcal{T}'_\alpha$ of $\mathfrak{X}_\alpha$ induced by $\mathcal{T}'$ is weaker than $\mathcal{T}_\alpha$.
(c) The embedding mapping $I_\alpha : \mathfrak{X}_\alpha \to \mathfrak{X} (\alpha \in A)$ is continuous with respect to $\mathcal{T}'$.

Then the inductive limit is characterized as follows.

Theorem B.1

(i) *The inductive limit topology* $\mathcal{T}$ *is the strongest one which satisfies (a) and (b).*
(ii) $\mathcal{T}$ *is the strongest one which satisfies (a) and (c).*

Proof (i) Let $\mathcal{T}'$ be a topology of $\mathfrak{X}$ which satisfies (a) and (b). If $U \in \mathcal{N}(0, \mathcal{T}')$ is a balanced convex set,

$$U \cap \mathfrak{X}_\alpha \in \mathcal{N}(0, \mathcal{T}'_\alpha) \quad \text{for all} \quad \alpha \in A.$$

Since $\mathcal{T}'_\alpha$ is weaker than $\mathcal{T}_\alpha$, we also have

[2] Let $\mathfrak{X}$ be a vector space and M its subset. M is **balanced** if $\alpha x \in M$ for every $x \in M$ and $\alpha \in \mathbb{C}$ with $|\alpha| \leqq 1$. M is said to be **absorbing** if there exists some $\alpha > 0$ for every $x \in \mathfrak{X}$ such that $(1/\alpha)x \in M$. (cf. Maruyama [3] pp. 246–248, Schwartz [5] p. 7.) co$M$ is the convex hull of M.

$$U \cap \mathfrak{X}_\alpha \in \mathcal{N}(0, \mathcal{T}_\alpha) \quad \text{for all} \quad \alpha \in A.$$

Hence $U \in \mathcal{N}(0, \mathcal{T})$. This proves that $\mathcal{T}'$ is weaker than $\mathcal{T}$.

(ii) Let $\mathcal{T}'$ be a topology of $\mathfrak{X}$ which satisfies (a) and (c). If $U \in \mathcal{N}(0, \mathcal{T}')$ is a balanced convex set,

$$U \cap \mathfrak{X}_\alpha \in \mathcal{N}(0, \mathcal{T}_\alpha) \quad \text{for all} \quad \alpha \in A,$$

by (c). Hence $U \in \mathcal{N}(0, \mathcal{T})$. □

The following several theorems show that seminorms, linear functionals and linear operators defined on $(\mathfrak{X}, \mathcal{T})$ are continuous if and only if they are continuous on $(\mathfrak{X}_\alpha, \mathcal{T}_\alpha)(\alpha \in A)$.

Theorem B.2 *Let $(\mathfrak{X}_\alpha, \mathcal{T}_\alpha)$, $\alpha \in A$ be LCHTVS's. Suppose that the vector space $\mathfrak{X} = \bigcup_{\alpha \in A} \mathfrak{X}_\alpha$ is endowed with the inductive limit topology $\mathcal{T}$. The following two statements are equivalent for any seminorm p on $\mathfrak{X}$:*

(i) *p is continuous with respect to $\mathcal{T}$.*
(ii) *p is continuous on $\mathfrak{X}_\alpha$ for all $\alpha \in A$.*

Proof (i)⇒(ii): If p is continuous on $(\mathfrak{X}, \mathcal{T})$, p is, of course, continuous on $(\mathfrak{X}_\alpha, \mathcal{T}'_\alpha)$ (where $\mathcal{T}'_\alpha$ is the relative topology of $\mathfrak{X}_\alpha$ induced by $\mathcal{T}$). Since $\mathcal{T}_\alpha$ is stronger than $\mathcal{T}'_\alpha$, p is continuous on $(\mathfrak{X}_\alpha, \mathcal{T}_\alpha)$.

(ii)⇒(i): If we assume (ii), the set

$$\{x \in \mathfrak{X} | p(x) \leqq 1\}$$

is balanced and convex, and satisfies

$$\{x \in \mathfrak{X} | p(x) \leqq 1\} \cap \mathfrak{X}_\alpha \in \mathcal{N}(0, \mathcal{T}_\alpha) \quad \text{for all} \quad \alpha \in A.$$

Hence, by the definition of the topology of $\mathcal{T}$, we have

$$\{x \in \mathfrak{X} | p(x) \leqq 1\} \in \mathcal{N}(0, \mathcal{T}).$$

Thus p is continuous on $(\mathfrak{X}, \mathcal{T})$.[3] □

Theorem B.3 $(\mathfrak{X}_\alpha, \mathcal{T}_\alpha)$ *and* $(\mathfrak{X}, \mathcal{T})$ *are the same as in Theorem B.2. The following two statements are equivalent for a linear functional $\Lambda : \mathfrak{X} \to \mathbb{C}$:*

[3] $\mathcal{N}(0)$ is the perfect family of neighborhoods of 0. The following five statements are equivalent for a seminorm $p(\cdot)$ on a topological vector space $\mathfrak{X}$. (i) p is continuous. (ii) $\{x \in \mathfrak{X} | p(x) < 1\}$ is open. (iii) $\{x \in \mathfrak{X} | p(x) < 1\} \in \mathcal{N}(0)$. (iv) $\{x \in \mathfrak{X} | p(x) \leqq 1\} \in \mathcal{N}(0)$. (v)$p$ is continuous at 0 (Schwartz [5] p. 10).

(i) *Λ is continuous with respect to $\mathcal{T}$.*
(ii) *Λ is continuous on $\mathfrak{X}_\alpha$ for all α.*

Proof Λ is continuous on $(\mathfrak{X}, \mathcal{T})$.

$\iff$ $|\Lambda(\cdot)|$ is continuous seminorm on $(\mathfrak{X}, \mathcal{T})$.

$\iff$ $|\Lambda(\cdot)|$ is continuous seminorm on $(\mathfrak{X}_\alpha, \mathcal{T}_\alpha)$ for all α.

(Theorem B.2)

$\iff$ Λ is continuous on $(\mathfrak{X}_\alpha, \mathcal{T}_\alpha)$ for all α. □

Corollary B.1 *$(\mathfrak{X}_\alpha, \mathcal{T}_\alpha)$ and $(\mathfrak{X}, \mathcal{T})$ are the same as in Theorem B.2. Let $\mathfrak{Y}$ be a locally convex topological vector space and $T : \mathfrak{X} \to \mathfrak{Y}$ a linear operator. Then the following two statements are equivalent:*

(i) *T is continuous with respect to $\mathcal{T}$.*
(ii) *T is continuous on $\mathfrak{X}_\alpha$ for all $\alpha \in A$.*

The proof is quite easy. It is enough to check $q \circ T$ for a continuous seminorm q on $\mathfrak{Y}$.

The inductive limit topology $\mathcal{T}$ is called a **strictly inductive limit topology** if the following three conditions are satisfied in the definition of $\mathcal{T}$:

(i) $A = \mathbb{N}$.
(ii) $\mathfrak{X}_n \subset \mathfrak{X}_{n+1}$ and the topology $\mathcal{T}_n$ on $\mathfrak{X}_n$ coincides with the relative topology on $\mathfrak{X}_n$, induced by $\mathcal{T}_{n+1}$.
(iii) $\mathfrak{X}_n$ is a closed subspace of $\mathfrak{X}_{n+1}$.

Remark B.1

1° (i) Let $\mathfrak{M}$ be a subspace of a topological vector space $\mathfrak{X}$. Suppose that U is a convex, open, balanced neighborhood of $0 \in \mathfrak{X}$, and V is a convex, open, balanced neighborhood of $0 \in \mathfrak{M}$ such that $U \cap \mathfrak{M} \subset V$. Then $\mathrm{co}(U \cup V)$ is a convex, open, balanced neighborhood of $0 \in \mathfrak{X}$.

(ii) Let q be a continuous seminorm on $\mathfrak{M}$, and p a continuous seminorm on $\mathfrak{X}$ such that

$$q(x) \leqq p(x) \quad \text{for all} \quad x \in \mathfrak{M}.$$

Then there exists a continuous seminorm $\tilde{q}$ on $\mathfrak{X}$ which satisfies

a. $\tilde{q}(x) = q(x)$ for all $x \in \mathfrak{M}$,
b. $\tilde{q}(x) \leqq p(x)$ for all $x \in \mathfrak{X}$.

2° The following results are generalizations of the facts discussed in Chap. 10.

(i) Let $\mathfrak{X}$ be a vector space. Suppose that $\mathfrak{Y}$ and $\mathfrak{Z}$ are orthogonal complements of each other; i.e. $\mathfrak{X} = \mathfrak{Y} \oplus \mathfrak{Z}$. If $P : \mathfrak{X} \to \mathfrak{Y}$ is a projection, it satisfies $P(\mathfrak{X}) = \mathfrak{Y}$, $P^{-1}(\{0\}) = \mathfrak{Z}$ and $P^2 = P$. Conversely, assume that $P : \mathfrak{X} \to \mathfrak{X}$ is a linear operator which satisfies $P^2 = P$. Then there exist uniquely a pair of subspaces $\mathfrak{Y}, \mathfrak{Z} \subset \mathfrak{X}$ such that $\mathfrak{X} = \mathfrak{Y} \oplus \mathfrak{Z}$ and P is a projection of $\mathfrak{X}$ into $\mathfrak{Y}$.

(ii) Suppose that $\mathfrak{X}$ is a Hausdorff topological vector space, and $\mathfrak{Y}, \mathfrak{Z}$ and P are just as in (i). If we define a linear operator $\varphi : \mathfrak{Y} \times \mathfrak{Z} \to \mathfrak{X}$ by

$$\varphi : (x, y) \mapsto x + y,$$

then φ is a injection. In the case that φ is a homeomorphism, $\mathfrak{Y}$ and $\mathfrak{Z}$ are said to be **topological complements** of each other. $\mathfrak{Y}$ and $\mathfrak{Z}$ are topological complements if and only if P is continuous. And $\mathfrak{Y}$ and $\mathfrak{Z}$ are closed subspaces of $\mathfrak{X}$ if P is continuous. Furthermore, the converse holds good if $\dim\mathfrak{Y} < \infty$ or $\dim\mathfrak{Z} < \infty$.

(iii) Let $\mathfrak{X}$ be a LCHTVS. If $\mathfrak{Y}$ is a subspace of $\mathfrak{X}$ such that either

a. $\dim\mathfrak{Y} < \infty$,
   or
b. $\mathfrak{Y}$ is closed and $\operatorname{codim}\mathfrak{Y} < \infty$,

then $\mathfrak{Y}$ has a topological complement $\mathfrak{Z}$.

Theorem B.4 (Continuous extension of seminorm) *Let $(\mathfrak{X}_n, \mathcal{T}_n)$ be LCHTVS's. $\mathfrak{X} = \bigcup_{n=1}^{\infty} \mathfrak{X}_n$ is endowed with a strictly inductive limit topology $\mathcal{T}$. Any continuous seminorm p_n on $(\mathfrak{X}_n, \mathcal{T}_n)$ can be extended as a continuous seminorm on $(\mathfrak{X}, \mathcal{T})$.*

Proof If we define a set V in $\mathfrak{X}_n$ by

$$V = \{x \in \mathfrak{X}_n | p_n(x) \leqq 1\},$$

there exists a convex, open, balanced neighborhood U of $0 \in \mathfrak{X}_{n+1}$ such that

$$U \cap \mathfrak{X}_n \subset V$$

by the definition of the strictly inductive limit topology. If we denote by q the continuous seminorm on $\mathfrak{X}_{n+1}$ determined by U, it holds good that

$$p_n(x) \leqq q(x) \quad \text{for all} \quad x \in \mathfrak{X}_n.$$

Applying Remark B.1, 1° above, we know that p_n can be extended to a continuous seminorm p_{n+1} on $\mathfrak{X}_{n+1}$. Repeating the same process, we obtain a seminorm p on the entire space $\mathfrak{X}$. The restriction of p to each $\mathfrak{X}_k (k \geqq n)$ is continuous. Hence p is continuous on $\mathfrak{X}$ by Theorem B.2. □

It should be noted that **the relative topology of $\mathfrak{X}_n$ induced by $(\mathfrak{X}, \mathcal{T})$ coincides with $\mathcal{T}_n$**. This is an important implication of Theorem B.4.

Corollary B.2 *A strictly inductive limit topology is Hausdorff.*[4]

Remark B.2 Let $\mathfrak{X}$ be a vector space. Suppose that both $\{(\mathfrak{X}_n, \mathcal{T}_n)\}$ and $\{(\tilde{\mathfrak{X}}_m, \tilde{\mathcal{T}}_m)\}$ are increasing sequences of LCHTVS's such that $\mathfrak{X} = \bigcup_n \mathfrak{X}_n = \bigcup_m \tilde{\mathfrak{X}}_m$. Assume further that:

(a) For any $n \in \mathbb{N}$, there exists some $m \in \mathbb{N}$ such that $\mathfrak{X}_n \subset \tilde{\mathfrak{X}}_m$ and the relative topology of $\mathfrak{X}_n$ induced by $\tilde{\mathcal{T}}_m$ coincides with $\mathcal{T}_n$.
(b) For any $m \in \mathbb{N}$, there exists some $n \in \mathbb{N}$ such that $\tilde{\mathfrak{X}}_m \subset \mathfrak{X}_n$ and the relative topology of $\tilde{\mathfrak{X}}_m$ induced by $\mathcal{T}_n$ coincides with $\tilde{\mathcal{T}}_m$.

Then the strictly inductive limit topology on $\mathfrak{X}$ defined by $\{\mathfrak{X}_n\}$ coincides with that defined by $\{\tilde{\mathfrak{X}}_m\}$.

Theorem B.5 $(\mathfrak{X}_n, \mathcal{T}_n)$ *and* $(\mathfrak{X}, \mathcal{T})$ *are the same as in Theorem B.4. The following two statements are equivalent for a set* $A \subset \mathfrak{X}$*:*

(i) *A is bounded in $(\mathfrak{X}, \mathcal{T})$,*
(ii) *A is bounded in $(\mathfrak{X}_n, \mathcal{T}_n)$ for some n.*

Proof It is sufficient to show (i)⇒(ii) because (ii)⇒(i) is obvious.

Suppose that there is no $\mathfrak{X}_n$ such that $A \subset \mathfrak{X}_n$. Then there exist a subsequence $\{\tilde{\mathfrak{X}}_m\}$ of $\{\mathfrak{X}_m\}$ and a sequence $\{x_1 \neq 0, x_2, \cdots, x_m, \cdots\}$ of A which satisfy

- $x_1 \in \tilde{\mathfrak{X}}_1, \ x_m \in \tilde{\mathfrak{X}}_m \setminus \tilde{\mathfrak{X}}_{m-1}; \quad m = 2, 3, \cdots$
- $\mathfrak{X} = \bigcup_{m=1}^{\infty} \tilde{\mathfrak{X}}_m$.
- $\{\tilde{\mathfrak{X}}_m\}$ satisfies (a) and (b) in the above remark.

Note that the strictly inductive limit topology on $\mathfrak{X}$ determined by $\{\tilde{\mathfrak{X}}_m\}$ coincides with $\mathcal{T}$.

Construct a sequence $\{p_n\}$ of seminorms by induction. First of all, choose a continuous seminorm p_1 on $\tilde{\mathfrak{X}}_1$ such that $p_1(x_1) = 1$. Next assume that a continuous seminorm p_{m-1} on $\tilde{\mathfrak{X}}_{m-1}$ is constructed so that

$$p_{m-1}|_{\tilde{\mathfrak{X}}_{m-2}} = p_{m-2}, \quad p_{m-1}(x_{m-1}) = m - 1.$$

We now define

$$\mathfrak{Z} = \mathrm{span}\{x_m\} = \mathbb{C}x_m (\subset \tilde{\mathfrak{X}}_m),$$
$$\mathfrak{M} = \mathrm{span}(\tilde{\mathfrak{X}}_{m-1} \cup \mathfrak{Z}).$$

[4] It is sufficient to find disjoint neighborhoods of 0 and $x \neq 0$. Since $x \in \mathfrak{X}_n$ for some n, there exists a continuous seminorm p_n such that $p_n(x) = 1$. Extend this p_n, by Theorem B.2, to a continuous seminorm on $\mathfrak{X}$.

Applying Remark B.1, 2°(iii), we observe that $\tilde{\mathfrak{X}}_{m-1}$ and $\mathfrak{Z}$ are topological complements in $\mathfrak{M}$ of each other since $\tilde{\mathfrak{X}}_{m-1}$ is closed and $\dim \mathfrak{M}/\tilde{\mathfrak{X}}_{m-1} = \dim \mathfrak{Z} = 1$. Hence $\tilde{\mathfrak{X}}_{m-1} \times \mathfrak{Z}$ and $\tilde{\mathfrak{X}}_{m-1} \oplus \mathfrak{Z} = \mathfrak{M}$ are homeomorphic.

$x \in \mathfrak{M}$ can be represented as

$$x = y + \lambda x_m; \quad \lambda \in \mathbb{C}, \ y \in \tilde{\mathfrak{X}}_{m-1},$$

or $x = (y, \lambda)$, for brevity. If we define q_m by

$$q_m(x) = p_{m-1}(y) + m|\lambda|, \quad x = (y, \lambda) \in \tilde{\mathfrak{X}}_{m-1} \oplus \mathfrak{Z},$$

q_m is a continuous seminorm on $\tilde{\mathfrak{X}}_{m-1} \oplus \mathfrak{Z} = \mathfrak{M}$ and $q_m(x_m) = m$. We can extend q_m to a continuous seminorm p_m on $\tilde{\mathfrak{X}}_m$, by Remark B.1, 1°(ii).

Repeating this process, we obtain a seminorm p on $\mathfrak{X}$ defined by

$$p(x) = p_m(x) \quad \text{if} \quad x \in \tilde{\mathfrak{X}}_m.$$

Then $p|_{\tilde{\mathfrak{X}}_m} = p_m$ $(m \in \mathbb{N})$. Hence p is continuous on each $\tilde{\mathfrak{X}}_m$. It follows that p is continuous on $\mathfrak{X}$ by Theorem B.2. However, p is not bounded on A. This contradicts the boundedness of A.

Thus we have shown that $A \subset \mathfrak{X}_n$ for some n. □

The next important consequence follows immediately from Theorem B.5.

Corollary B.3 *Let $(\mathfrak{X}_n, \mathcal{T}_n)$ and $(\mathfrak{X}, \mathcal{T})$ be the same as in Theorem B.4. The following two statements are equivalent for any net $\{x_\alpha\}$ in $\mathfrak{X}$:*

(i) *$\{x_\alpha\}$ converges to some $x^* \in \mathfrak{X}$ in the strict inductive limit topology.*
(ii) *$\{x_\alpha\}$ is contained in some $\mathfrak{X}_n$ and it converges to $x^* \in \mathfrak{X}_n$.*

Definition B.1 A closed, convex, balanced and absorbing set in a topological vector space is called a **barrel**. A LCHTVS is called a **barreled space** if every barrel is a neighborhood of the origin 0.

The inductive limit topology of barreled spaces is also barreled. We now show it successively.

Theorem B.6 *A LCHTVS $\mathfrak{X}$ is barreled if and only if every lower semi-continuous (l.s.c.) seminorm is continuous.*

Proof Assume first that $\mathfrak{X}$ is a barreled LCHTVS, and a seminorm p is l.s.c. on $\mathfrak{X}$. If we define

$$A = \{x \in \mathfrak{X} | p(x) \leqq 1\}$$

then A is clearly convex, balanced and absorbing. It is closed, since p is l.s.c. Hence A is a barrel and a neighborhood of 0. Therefore p is continuous. (cf. footnote 3 on p. 359.)

Conversely, assume that every l.s.c. seminorm on $\mathfrak{X}$ is continuous. Let A be a barrel. If we define

$$p(x) = \inf\{\lambda > 0 | x \in \lambda A\},$$

then $p(x)$ is finite, since A is absorbing. It is also easily seen that p is a seminorm.[5]

We next show that $A = \{x \in \mathfrak{X} | p(x) \leqq 1\}$. In case of $x \in A$, it is obvious that $p(x) \leqq 1$. If $p(x) < 1$, we must have $x \in A$, since A is balanced. If $p(x) = 1$,

$$(1 - \varepsilon)x \in A \quad \text{for all} \quad \varepsilon \in (0, 1).$$

So we must have $x \in A$, since A is closed. Furthermore, p is l.s.c. again by the closedness of A. Thus p is continuous, and A is a neighborhood of 0. Hence $\mathfrak{X}$ is barreled. □

Lemma B.1 *Let X be a Baire space*[6] *and a function $f : X \to \mathbb{R}$ be l.s.c. Then for any open set U in X, there exists an open set $V \subset U$ such that*

$$\sup_{x \in V} f(x) = M < \infty.$$

Proof U is a Baire space, since it is an open set in a Baire space.[7] We restrict our attention to U. If we define

$$F_n = \{x \in U | f(x) \leqq n\}; \quad n = 1, 2, \cdots,$$

we observe that $\text{int.} F_n \neq \emptyset$ for some n by the Baire category theorem. If we define $V = \text{int.} F_n$, it satisfies the required property. □

Remark B.3 Let X be a Baire space and $\{f_\alpha : X \to \mathbb{R}\}$ a family of continuous functions such that

$$\sup_{\alpha} |f_\alpha(x)| < \infty$$

---

[5] In general, suppose that K is a balanced, absorbing and convex neighborhood of 0 of a LCHTVS. If we define

$$p_K(x) = \inf\left\{r > 0 \,\middle|\, \frac{x}{r} \in K\right\},$$

$p_K(\cdot)$ is a seminorm. It is called the **Minkowski functional** determined by K (cf. Maruyama [3] pp. 250–251 and Yosida [7] p. 24).

[6] A topological space X is called a **Baire space** if $\overline{X \backslash M} = X$ for any set $M \subset X$ of the first category. (cf. Maruyama [3] pp. 28–33.)

[7] Maruyama [3] Theorem 1.14, p. 30.

for each $x \in X$. Then $\{f_\alpha\}$ is uniformly bounded on some open set V in X. (Consider $f(x) = \sup_\alpha |f_\alpha(x)|$.) This is a slight generalization of Osgood's uniform boundedness principle.

Theorem B.7 *If a LCHTVS $\mathfrak{X}$ is a Baire space, it is barreled.*

Proof Let p be a l.s.c. seminorm on $\mathfrak{X}$. Then, by Lemma B.1, there exists an open set $V \in \mathfrak{X}$ such that

$$\sup_{x \in V} p(x) = M < \infty.$$

Pick up any $x \in V$ and define $W = V - x$. Then W is an open neighborhood of 0. We observe that

$$p(w) \leqq 2M \quad \text{for any} \quad w \in W,$$

since $w = v - x$ for some $v \in V$. Hence

$$\frac{1}{2M} W \subset \{x \in \mathfrak{X} | p(x) \leqq 1\}.$$

It tells us that $\{x \in \mathfrak{X} | p(x) \leqq 1\}$ is a neighborhood of 0. So p is continuous. By Theorem B.6, we conclude that $\mathfrak{X}$ is barreled. □

Corollary B.4 *Any Banach space is barreled. Any Fréchet space is barreled.*

Example B.1 $\mathfrak{C}([0, 1], \mathbb{R})$ endowed with $\mathfrak{L}^1$-norm is not barreled. (Check $U = \{f \in \mathfrak{C}([0, 1], \mathbb{R}) \mid \|f\|_\infty \leqq 1\}$.)

Theorem B.8 *Suppose that $\mathfrak{X}$ is a vector space and $\{\mathfrak{X}_\alpha\}_{\alpha \in A}$ is a family of vector subspaces of $\mathfrak{X}$ such that $\mathfrak{X} = \bigcup_{\alpha \in A} \mathfrak{X}_\alpha$. If every $\mathfrak{X}_\alpha$ is barreled and $\mathfrak{X}$ is endowed with the inductive limit topology, then $\mathfrak{X}$ is also barreled.*

Proof If A is a barrel in $\mathfrak{X}$, $A \cap \mathfrak{X}_\alpha$ is a barrel in $\mathfrak{X}_\alpha$ for all α. Since $\mathfrak{X}_\alpha$ is barreled, $A \cap \mathfrak{X}_\alpha$ is a neighborhood of 0 in $\mathfrak{X}_\alpha$. Consequently, A is a neighborhood of 0 in $\mathfrak{X}$ by the definition of the inductive limit topology. So $\mathfrak{X}$ is barreled. □

By Corollary B.4 and Theorem B.8, it is clear that $\mathfrak{E}(\Omega)$, $\mathfrak{D}_K(\Omega)$ and $\mathfrak{D}(\Omega)$, which play crucial roles in the theory of distributions, are all barreled. (See Corollary C.1 for details.)

Definition B.2 A barreled space is called a **Montel space** if any bounded set is relatively compact.

It will be discussed later on that $\mathfrak{D}(\Omega)$ and $\mathfrak{E}(\Omega)$ are typical Montel spaces. (cf. Theorem C.5, p. 384.)

B.2 Duals of Locally Convex Spaces

Definition B.3 Let $\mathfrak{X}$ be a topological vector space (TVS) and $\mathfrak{Y}$ a locally convex topological vector space. We denote by $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ the set of all the continuous linear operators of $\mathfrak{X}$ into $\mathfrak{Y}$. There are three ordinary topologies on $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$:

1° **Simple topology**: If we define $p_{x,q}$ by

$$p_{x,q}(T) = q(Tx); \quad T \in \mathcal{L}(\mathfrak{X}, \mathfrak{Y}) \tag{B.1}$$

for $x \in \mathfrak{X}$ and a continuous seminorm q on $\mathfrak{Y}$, then $p_{x,q}$ is a seminorm on $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$. The topology on $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ generated by the family of seminorms

$$\{p_{x,q} | x \in \mathfrak{X},\ q \text{ is a continuous seminorm on } \mathfrak{Y}\}$$

is called the **simple topology**. In the special case of $\mathfrak{Y} = \mathbb{C}$, the simple topology is nothing other than the weak*-topology.

2° **Compact topology**: If we define $p_{K,q}$ by

$$p_{K,q}(T) = \sup_{x \in K} q(Tx); \quad T \in \mathfrak{L}(\mathfrak{X}, \mathfrak{Y}) \tag{B.2}$$

for a compact set $K \subset \mathfrak{X}$ and a continuous seminorm q on $\mathfrak{Y}$, $p_{K,q}$ is a seminorm on $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$. The topology on $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ generated by the family of seminorms of the form (B.2) is called the **compact topology**.

3° **Strong topology**: If we define $p_{B,q}$ by

$$p_{B,q}(T) = \sup_{x \in B} q(Tx); \quad T \in \mathcal{L}(\mathfrak{X}, \mathfrak{Y}) \tag{B.3}$$

for a bounded set $B \subset \mathfrak{X}$ and a continuous seminorm q on $\mathfrak{Y}$, $p_{B,q}$ is a seminorm on $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$. The topology on $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ generated by the family of seminorms of the form (B.3) is called the **strong topology**.

Theorem B.9 *Let $\mathfrak{X}$ be a LCHTVS. The following two statements are equivalent:*

(i) *$\mathfrak{X}$ is barreled.*
(ii) *Any w^*-bounded set in $\mathfrak{X}'$ is equi-continuous.*

Proof (i)⇒(ii): Let A be a w^*-bounded set in $\mathfrak{X}'$. The seminorm $x \mapsto |\Lambda x|$ defined for each $\Lambda \in A$ is continuous. Hence if we define

$$p_A(x) \equiv \sup_{\Lambda \in A} |\Lambda x|,$$

$p_A(x)$ is finite (by w^*-boundedness of A) and l.s.c. Since $\mathfrak{X}$ is barreled, p_A is continuous by Theorem B.6. Consequently, for each $\varepsilon > 0$, there exists a neighborhood U of $0 \in \mathfrak{X}$, such that

$$\sup_{x \in U} p_A(x) = \sup_{x \in U} \sup_{\Lambda \in A} |\Lambda x| \leqq \varepsilon.$$

That is, A is equi-continuous.

(ii)⇒(i): Conversely, assume (ii). If B is a barrel in $\mathfrak{X}$,[8]

$$B^\circ = \{\Lambda \in \mathfrak{X}' \mid |\Lambda x| \leqq 1 \text{ for all } x \in B\}$$

is w^*-bounded in $\mathfrak{X}'$. It can be proved as follows. Since B is absorbing, we have, for any $z \in \mathfrak{X}$, $\lambda z \in B$ for some $\lambda \in \mathbb{C}$. Hence

$$\sup_{\Lambda \in B^\circ} |\Lambda z| = \sup_{\Lambda \in B^\circ} \frac{1}{|\lambda|} |\Lambda(\lambda z)| \leqq \frac{1}{|\lambda|}.$$

This holds good for any $z \in \mathfrak{X}$. Thus B° is w^*-bounded. Hence B° is equi-continuous by (ii). Defining $B^{\circ\circ} = \{x \in \mathfrak{X} \mid |\Lambda x| \leqq 1 \text{ for all } \Lambda \in B^\circ\}$, we obtain $B = B^{\circ\circ}$ by the Hahn–Banach theorem. It follows that[9]

$$B = B^{\circ\circ} \in \mathcal{N}(0).$$

That is, $\mathfrak{X}$ is barreled. □

Remark B.4 For any set $A \subset \mathfrak{X}$, in general, $A^{\circ\circ}$ is the smallest closed, convex and balanced set containing A.

The proof is not so hard. $A^{\circ\circ} \supset \overline{\text{co}}A$ is obvious, since $A^{\circ\circ}$ is a closed, convex and balanced set containing A. Suppose that $A^{\circ\circ} \subset \overline{\text{co}}A$ is not true. Then there exists some $x_0 \in A^{\circ\circ} \setminus \overline{\text{co}}A$. By the Hahn–Banach theorem, there exists some $\Lambda \in \mathfrak{X}'$ which satisfies

$$\overline{\Lambda(\overline{\text{co}}A)} \bigcap \{\Lambda x_0\} = \emptyset.$$

The shape of $\overline{\Lambda(\overline{\text{co}}A)}$ is:

- a closed disc with center 0 in $\mathbb{C}$ in the case of complex vector space;
- a closed segment with center 0 in $\mathbb{R}$ in the case of real vector space.

In any case,

$$\overline{\Lambda(\overline{\text{co}}A)} = \{z \in \mathbb{C} \text{ or } \mathbb{R} \mid |z| \leqq \rho\}$$

[8]Let A be a set in $\mathfrak{X}$. The **polar set** A° of A is defined by $A^\circ = \{\Lambda \in \mathfrak{X}' \mid |\Lambda x| \leqq 1 \text{ for all } x \in A\}$. We list several simple properties of A°: (i) $\{0\}^\circ = \mathfrak{X}'$, $\mathfrak{X}^\circ = \{0\} \in \mathfrak{X}'$. (ii) $A \subset B \Rightarrow B^\circ \subset A^\circ$. (iii) $(\lambda A)^\circ = (1/\lambda)A^\circ$ for $\lambda \in \mathbb{C} \setminus \{0\}$. (iv) $(A \cup B)^\circ = A^\circ \cap B^\circ$. (cf. Schwartz [5] pp. 56–57.)

[9]For any $0 < \varepsilon < 1$, there exists some $V \in \mathcal{N}(0)$ such that $|\Lambda x| < \varepsilon < 1$ for all $x \in V$ and all $\Lambda \in B^\circ$ by the equi-continuity of B°. Hence $V \subset B^{\circ\circ}$. Thus we obtain $B^{\circ\circ} \in \mathcal{N}(0)$.

for some $\rho \geqq 0$. Since Λx_0 is located outside this set, $|\Lambda x_0| > \rho$. Defining $\Gamma = \Lambda/\rho$, we have $\Gamma \in A^\circ$, since $|\Gamma x| \leqq 1$ for any $x \in \overline{\mathrm{co}}A$. On the other hand, $|\Gamma x_0| > 1$. This contradicts $x_0 \in A^{\circ\circ}$. Hence it follows that $A^{\circ\circ} \subset \overline{\mathrm{co}}A$.

Theorem B.10 *Suppose that $\mathfrak{X}$ is a barreled space and $\mathfrak{Y}$ is a LCHTVS. Then any bounded set in $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ with respect to the simple topology (s-bounded for brevity) is equi-continuous.*

Proof Let $H \subset \mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ be s-bounded; i.e.

$$H(x) \equiv \{Tx \in \mathfrak{Y} | T \in H\}$$

is bounded in $\mathfrak{Y}$ for any $x \in \mathfrak{X}$. We wish to show that

$$U \equiv \bigcap_{T \in H} T^{-1}(V) \in \mathcal{N}_{\mathfrak{X}}(0)$$

for any $V \in \mathcal{N}_{\mathfrak{Y}}(0)$. ($\mathcal{N}_{\mathfrak{X}}(0)$ and $\mathcal{N}_{\mathfrak{Y}}(0)$ are the perfect family of neighborhoods of 0 in $\mathfrak{X}$ and $\mathfrak{Y}$, respectively.) Without loss of generality, we may assume that V is closed, convex and balanced.

Since $T^{-1}(V)$ is closed, convex and balanced for any $T \in H$, so is

$$U = \bigcap_{T \in H} T^{-1}(V).$$

By the s-boundedness of H, there exists some $\lambda \in \mathbb{R}$, for each $x \in \mathfrak{X}$, which satisfies

$$\lambda H(x) = \{T(\lambda x) \in \mathfrak{Y} | T \in H\} \subset V.$$

That is, $\lambda x \in T^{-1}(V)$. Thus we have verified that U is absorbing. So U is a barrel. Since $\mathfrak{X}$ is barreled, $U \in \mathcal{N}_{\mathfrak{X}}(0)$. □

Remark B.5 Suppose that $\mathfrak{X}$ and $\mathfrak{Y} \neq \{0\}$ are LCHTVS's. If every s-bounded set in $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ is equi-continuous, U is barreled. (The converse of Theorem B.10.) It is proved by applying Theorem B.9 to obtain a contradiction.

We now proceed to the generalized Ascoli–Arzelà theorem and then to the Banach–Steinhaus theorem.

Theorem B.11 (Ascoli–Arzelà) *Let X be a topological space and $\mathfrak{Y}$ a locally convex topological vector space. Suppose that $\mathcal{F}$ is an equi-continuous subset of $\mathfrak{C}(X, \mathfrak{Y})$, the space of continuous functions of X into $\mathfrak{Y}$.*

(i) *The closure $\bar{\mathcal{F}}$ of $\mathcal{F}$ with respect to the pointwise convergence topology is an equi-continuous set in $\mathfrak{C}(X, \mathfrak{Y})$.*
(ii) *The following three topologies on $\mathcal{F}$ (similarly for $\bar{\mathcal{F}}$) coincide:*

a. *topology defined by pointwise convergence on a dense subset of* X : $(\mathcal{T}_a)$.
b. *pointwise convergence topology on* X : $(\mathcal{T}_b)$.
c. *uniform convergence on compacta topology on* X : $(\mathcal{T}_c)$.

(iii) *If*

$$\mathcal{F}(x) = \{f(x) \in \mathfrak{Y} | f \in \mathcal{F}\}$$

is relatively compact in $\mathfrak{Y}$ *for each* $x \in X$*, then* $\mathcal{F}$ *is relatively compact in* $\mathcal{T}_c$.

Proof [10] (i) Let $x \in X$ and $V \in \mathcal{N}_{\mathfrak{Y}}(0)$. There exists some balanced $W \in \mathcal{N}_{\mathfrak{Y}}(0)$ such that

$$W + W + W \subset V.$$

Since $\mathcal{F}$ is equi-continuous on X, there exists some $U \in \mathcal{N}_X(x)$ such that

$$f(z) - f(x) \in W \quad \text{for all} \quad z \in U, \ f \in \mathcal{F}.$$

If $\tilde{f} \in \bar{\mathcal{F}}$, there exists some $f \in \mathcal{F}$, for each $z \in U$, such that

$$\tilde{f}(z) - f(z) \in W, \quad \tilde{f}(x) - f(x) \subset W.$$

Therefore

$$\begin{aligned}\tilde{f}(z) - \tilde{f}(x) &= (\tilde{f}(z) - f(z)) + (f(z) - f(x)) + (f(x) - \tilde{f}(x)) \\ &\in W + W + W \subset V.\end{aligned}$$

Hence $\bar{\mathcal{F}}$ is equi-continuous on X. (And so, $\bar{\mathcal{F}} \subset \mathfrak{C}(X, \mathfrak{Y})$.)

(ii) It is obvious that $\mathcal{T}_a \subset \mathcal{T}_b \subset \mathcal{T}_c$. So it is enough to show $\mathcal{T}_c \subset \mathcal{T}_a$; i.e.

$$\mathcal{N}(f, \mathcal{T}_c) \subset \mathcal{N}(f, \mathcal{T}_a)$$

for each $f \in \mathcal{F}$. Without loss of generality, we may assume $f = 0$.

Let $D = \{z_\alpha\}$ be a dense subset of X. Suppose that $K \subset X$ is compact, and p is a continuous seminorm on $\mathfrak{Y}$. Then

$$q_{K,p}(f) = \sup_{x \in K} p(f(x)); \quad f \in \mathcal{F}$$

is a continuous seminorm on $\mathcal{F}$ with respect to $\mathcal{T}_c$. We now show that

$$S \equiv \{f \in \mathcal{F} | q_{K,p}(f) < 1\}$$

belongs to $\mathcal{N}(0, \mathcal{T}_a)$.

[10] Schwartz [5] pp. 78–80.

Define $V \in \mathcal{N}_{\mathfrak{Y}}(0)$ by

$$V = \left\{ y \in \mathfrak{Y} \middle| p(y) < \frac{1}{4} \right\}.$$

By the equi-continuity of $\mathcal{F}$, there exists some $U_x \in \mathcal{N}_X(x)$ for each $x \in K$ such that

$$f(U_x) - f(x) \subset V \quad \text{for all} \quad f \in \mathcal{F}.$$

Since $\{U_x | x \in K\}$ is an open covering of K, it has a finite subcovering $\{U_{x_1}, \cdots, U_{x_n}\}$. $D \cap U_{x_j} \neq \emptyset$ for $j = 1, 2, \cdots, n$ because $\bar{D} = X$. Pick up

$$z_{\alpha_j} \in D \cap U_{x_j}$$

and we write $z_{\alpha_j} = z_j$ for notational simplicity. It is clear that

$$f(U_{x_j}) - f(z_j) = (f(U_{x_j}) - f(x_j)) + (f(x_j) - f(z_j)) \subset V + V \quad \text{for all} \quad f \in \mathcal{F}.$$

Hence there exists some $z_j \in D$ for each $x \in K$ such that

$$p(f(x) - f(z_j)) < \frac{1}{2} \quad \text{for all} \quad f \in \mathcal{F}.$$

Thus it follows that

$$\sup_{x \in K} \operatorname*{Min}_{1 \leqq j \leqq n} p(f(x) - f(z_j)) \leqq \frac{1}{2}.$$

Defining

$$T = \bigcap_{j=1}^{n} \left\{ f \in \mathcal{F} \middle| q_{z_j, p}(f) < \frac{1}{2} \right\},$$

we obtain $T \in \mathcal{N}(0, \mathcal{T}_a)$. For $f \in T$, $x \in K$, we have

$$p(f(x)) \leqq p(f(x) - f(z_j)) + p(f(z_j))$$

and

$$q_{K,p}(f) = \sup_{x \in K} p(f(x)) < \frac{1}{2} + \frac{1}{2} = 1.$$

This implies $f \in S$, that is, $T \subset S$. We conclude that $S \in \mathcal{N}(0, \mathcal{T}_a)$.

(iii) $\mathcal{F}(x) = \{f(x) \in \mathfrak{Y} | f \in \mathcal{F}\}$ is relatively compact in $\mathfrak{Y}$. So, by Tihonov's theorem,

$$\prod_{x \in X} \mathcal{F}(x)$$

is relatively compact in $\mathfrak{Y}^X$. Thus $\mathcal{F}$ is relatively compact with respect to $\mathcal{T}_c$ by (i) and (ii). □

The generalized version of the Banach–Steinhaus theorem can be proved based upon Theorem B.11.

Theorem B.12 (Banach–Steinhaus) *Let $\mathfrak{X}$ be a barreled space and $\mathfrak{Y}$ a LCHTVS. Suppose that a sequence $\{T_n\}$ in $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ has a pointwise limit, that is $\lim_{n\to\infty} T_n(x) \equiv T(x)$ exists for each $x \in \mathfrak{X}$. Then $T \in \mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ and $\{T_n\}$ converges to T with respect to the compact topology.*

Proof $\{T_n\}$ is bounded in $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ with respect to the simple topology, since $\lim_{n\to\infty} T_n(x)$ exists for each $x \in \mathfrak{X}$. Since $\mathfrak{X}$ is barreled, $\{T_n\}$ is equi-continuous by Theorem B.10. Furthermore, thanks to (i) and (ii) of Theorem B.11, $T \in \mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ and $\{T_n\}$ converges to T with respect to the compact topology. □

We are led to the resonance theorem that the boundedness (with respect to the original topology) and the weak boundedness coincide in a locally convex space.

We make use of a quotient space of a locally convex space. We shall explain a few basic points.

Let $\mathfrak{X}$ be a locally convex topological vector space and $\mathfrak{M}$ a subspace of $\mathfrak{X}$. An element of the quotient space $\mathfrak{X}/\mathfrak{M}$, say a residue class containing x, is denoted by ξ_x.

(i) Let p be a seminorm on $\mathfrak{X}$. If we define $\dot{p} : \mathfrak{X}/\mathfrak{M} \to \mathbb{R}$ by

$$\dot{p}(\xi_x) = \inf_{z \in \xi_x} p(z),$$

then $\dot{p}$ is a seminorm, called the **quotient seminorm** on $\mathfrak{X}/\mathfrak{M}$. The following equality holds good:

$$\{\xi_x \in \mathfrak{X}/\mathfrak{M} | \dot{p}(\xi_x) < 1\} = \{\xi_x \in \mathfrak{X}/\mathfrak{M} | p(x) < 1\}.$$

(ii) If the topology on $\mathfrak{X}$ is generated by a family $\{p_\alpha\}$ of seminorms, the topology on $\mathfrak{X}/\mathfrak{M}$ generated by $\{\dot{p}_\alpha\}$ coincides with the quotient topology.

(iii) The quotient topology is Hausdorff if and only if $\mathfrak{M}$ is a closed subspace.

(iv) If $\mathfrak{X}$ is a Fréchet space and $\mathfrak{M}$ is a closed subspace, then $\mathfrak{X}/\mathfrak{M}$ is also a Fréchet space.

(v) If $\mathfrak{X}$ is a barreled space and $\mathfrak{M}$ is a closed subspace, then $\mathfrak{X}/\mathfrak{M}$ is also a barreled space.

(vi) $(\mathfrak{X}/\mathfrak{M})' \cong \mathfrak{M}^\circ$.

Theorem B.13 (Banach–Mackey Resonance Theorem) *Let $\mathfrak{X}$ be a locally convex topological vector space (not necessarily Hausdorff) and $\mathfrak{M}$ a subset of $\mathfrak{X}$. Then M is bounded if and only if it is weakly bounded.*

Proof **1°** It is well-known that the theorem holds good in the case of a normed space.[11]

2° We next consider a locally convex space $\mathfrak{X}$, the topology of which is defined by a single seminorm p. Since

$$\mathfrak{N} = \{x \in \mathfrak{X} | p(x) = 0\}$$

is a closed subspace of $\mathfrak{X}$, $\mathfrak{X}/\mathfrak{N}$ is a locally convex topological vector space. The quotient seminorm $\dot{p}$ defined by p actually turns out to be a norm. If $z \in \xi_x$,

$$z = x + y \quad \text{for some} \quad y \in \mathfrak{N}.$$

Hence

$$p(z) \leqq p(x) + p(y) = p(x).$$

We also obtain $p(x) \leqq p(z)$ similarly. It follows that $p(x) = p(z)$; that is, p is constant on each $\xi_x \in \mathfrak{X}/\mathfrak{M}$. Hence

$$p(x) = \dot{p}(\xi_x).$$

It follows that

$$\dot{p}(\xi_x) = 0 \Leftrightarrow \xi_x = \mathfrak{N}.$$

This proves that $\dot{p}$ is a norm.

For $\Lambda \in \mathfrak{X}'$, there exists a constant C such that

$$|\Lambda(x)| \leqq C \cdot p(x) \quad \text{for all} \quad x \in \mathfrak{X}.$$

(It is verified in a similar way in the case of a normed space.) Since $p(x) = 0$ for $x \in \mathfrak{N}$,

$$\Lambda(x) = 0 \quad \text{for all} \quad x \in \mathfrak{N}.$$

Corresponding to $\Lambda \in \mathfrak{X}'$, we define (cf. Fig. B.1)

$$\dot{\Lambda}(\xi_x) = \dot{\Lambda}(x + \mathfrak{N}) = \Lambda(x).$$

[11] Yosida [7] p. 69, and Maruyama [3] p. 351.

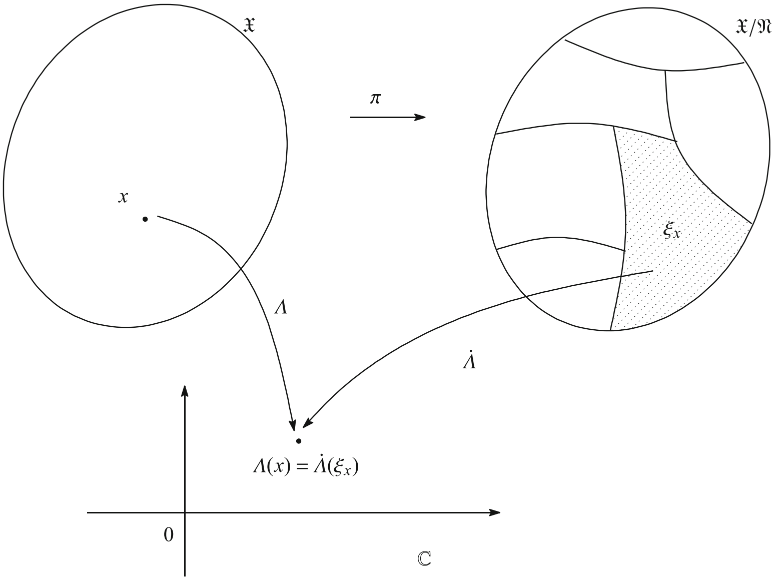


Fig. B.1 Definition of $\dot{\Lambda}$

Then

$$\{\xi_x \in \mathfrak{X}/\mathfrak{N} \mid |\dot{\Lambda}(\xi_x)| < 1\} = \{\xi_x \in \mathfrak{X}/\mathfrak{N} \mid \Lambda(x) < 1\} = \pi(\{x \in \mathfrak{X} \mid \Lambda(x) < 1\})$$
$$(\pi : \mathfrak{X} \to \mathfrak{X}/\mathfrak{N} \text{ is the canonical projection}).$$

This set is open, since π is an open mapping. Hence it is a neighborhood of $0 \in \mathfrak{X}/\mathfrak{N}$. Thus $\dot{\Lambda}$ is continuous and $\dot{\Lambda} \in (\mathfrak{X}/\mathfrak{N})'$.

If we define a linear operator $A : \mathfrak{X}' \to (\mathfrak{X}/\mathfrak{N})'$ by

$$A : \Lambda \mapsto \dot{\Lambda},$$

it is clear that

$$\mathrm{Ker}A = \{0\} \subset \mathfrak{X}'.$$

That is, A is a linear bijection. Hence $\mathfrak{X}'$ and $(\mathfrak{X}/\mathfrak{N})'$ are isomorphic.

So M is weakly bounded in $\mathfrak{X}$ if and only if $\dot{M} \equiv \{\xi_x | x \in M\}$ is weakly bounded in $\mathfrak{X}/\mathfrak{N}$. Since $\mathfrak{X}/\mathfrak{N}$ is a normed space, $\dot{M}$ is strongly bounded in $\mathfrak{X}/\mathfrak{N}$;

$$\text{i.e.} \quad \dot{p}(\xi_x) \leqq B \quad \text{for all} \quad x \in M$$

for some $B < \infty$. This means that

$$p(x) \leqq B \quad \text{for all} \quad x \in M.$$

3° Finally, let $\mathfrak{X}$ be a general locally convex space. Suppose that a family $\{p_\alpha\}$ of seminorms defines the topology on $\mathfrak{X}$. We denote by $\mathfrak{X}_\alpha$ the locally convex space, the topology on which is defined by a single seminorm p_α. The original topology on $\mathfrak{X}$ is, of course, stronger than the topology on $\mathfrak{X}_\alpha$. Hence $\sigma(\mathfrak{X}, \mathfrak{X}')$ is stronger than $\sigma(\mathfrak{X}, \mathfrak{X}'_\alpha)$.

It follows that, if M is weakly bounded in $\mathfrak{X}$, it is also weakly bounded in $\mathfrak{X}_\alpha$. M is bounded in $\mathfrak{X}_\alpha$ by 2°. This holds good for all α, and so

$$\sup_{x \in M} p_\alpha(x) < \infty \quad \text{for all} \quad \alpha.$$

Thus M is bounded on $\mathfrak{X}$. □

Remark B.6 The above theorem does not hold good if $\mathfrak{X}$ is not locally convex. To illuminate the situation by an example, let $\mathfrak{X} = \mathfrak{L}^{1/2}$. Then $\mathfrak{X}' = \{0\}$. It follows that any set in $\mathfrak{X}$ is weakly bounded. However, $\mathfrak{X}$ is not bounded with respect to the original topology.[12]

Corollary B.5 *Any weakly convergent sequence in a locally convex topological vector space is bounded.*

Definition B.4 Let $\mathfrak{X}$ be a locally convex topological vector space. $\mathfrak{X}$ is said to be **semi-reflexive** if the dual space $\mathfrak{X}''$ of the locally convex topological vector space $\mathfrak{X}'$ endowed with the strong topology coincides with $\mathfrak{X}$. $\mathfrak{X}$ is called **reflexive** if $\mathfrak{X}$ is semi-reflexive and the strong topology on $\mathfrak{X}''$ coincides with the topology on $\mathfrak{X}$.

The following two results are most basic to characterize semi-reflexivity and reflexivity.

Theorem B.14 (Mackey) *Let $\mathfrak{X}$ be a LCHTVS. The following two statements are equivalent:*

(i) *$\mathfrak{X}$ is reflexive.*
(ii) *Any bounded set in $\mathfrak{X}$ is weakly relatively compact.*

Theorem B.15 (reflexive vs semi-reflexive) *Let $\mathfrak{X}$ be a LCHTVS. The following two statements are equivalent.*

(i) *$\mathfrak{X}$ is reflexive.*
(ii) *$\mathfrak{X}$ is semi-reflexive and barreled.*

We prepare a couple of lemmata for the proofs.

Lemma B.2 *Let $\mathfrak{X}$ be a LCHTVS and $(\mathfrak{X}')^*$ the algebraic dual space of $\mathfrak{X}'$, endowed with the topology $\sigma((\mathfrak{X}')^*, \mathfrak{X}')$. Then $f \in (\mathfrak{X}')^*$ belongs to $(\mathfrak{X}')'$ (the topology of $\mathfrak{X}'$*

[12] This example is due to Rudin [4] p. 82. See also the revised version, pp. 35–36.

is the strong one) if and only if f belongs to the closure of some bounded set in $\mathfrak{X}$. (Note that we regard $\mathfrak{X}$ as a subspace of $(\mathfrak{X}')^*$.)

Proof For $f \in (\mathfrak{X}')^*$, we obtain

$$f \in \mathfrak{X}'' \Leftrightarrow f \in V^\circ = \{u \in (\mathfrak{X}')^* | \sup_{v \in V} |u(v)| \leqq 1\} \quad \text{for some} \quad V \in \mathcal{N}_{\mathfrak{X}'}(0).$$

A set $V \subset \mathfrak{X}'$ is a neighborhood of 0 (with respect to the strong topology) if and only if there exists a bounded set B in $\mathfrak{X}$ such that $B^\circ \subset V$. Without loss of generality, we may assume that B is also closed, convex and balanced. Thus $f \in (\mathfrak{X}')^*$ belongs to $(\mathfrak{X}')'$ if and only if there is a bounded set B in $\mathfrak{X}$ such that $f \in B^{\circ\circ}$.

If $f \in \bar{B}$, we have $f \in B^{\circ\circ} = \bar{B}$. Hence $f \in \mathfrak{X}''$ follows from the above result. Conversely, if $f \in (\mathfrak{X}')'$, it follows that $f \in B^{\circ\circ}$ for some bounded set B in $\mathfrak{X}$. Since we may admit that B is closed, convex and balanced, we obtain $B^{\circ\circ} = \bar{B}$. Hence $f \in \bar{B}$. $\square$

Lemma B.3 *Suppose that $\mathfrak{X}$ is a locally convex topological vector space and V is a neighborhood of $0 \in \mathfrak{X}$. Then V° is w^*-compact.*

Proof Let V be a neighborhood of $0 \in \mathfrak{X}$. Then V° is equi-continuous.[13] Since V is absorbing,

$$V^\circ(x) = \{\Lambda(x) | \Lambda \in V^\circ\}$$

is a bounded set in $\mathbb{C}$ for all $x \in \mathfrak{X}$. Hence V° is w^*-compact by Theorem B.11, (ii) and (iii). $\square$

Proof of Theorem B.14 (i)$\Rightarrow$(ii): Let $\mathfrak{X}$ be semi-reflexive and B a bounded set in $\mathfrak{X}$. B° is a neighborhood of 0 in $\mathfrak{X}'$ with respect to the strong topology. By Lemma B.3, $B^{\circ\circ}$ is $\sigma(\mathfrak{X}'', \mathfrak{X}')$-compact in $\mathfrak{X}'' = \mathfrak{X}$. Hence $B \subset B^{\circ\circ}$ is weakly relatively compact.

(ii)$\Rightarrow$(i): Assume that any bounded set $B \subset \mathfrak{X}$ is weakly relatively compact. If we regard $\mathfrak{X}$ as a subspace of $(\mathfrak{X}')^*$, the weak topology $\sigma(\mathfrak{X}, \mathfrak{X}')$ coincides with the relative topology of $\mathfrak{X}$ induced by $\sigma((\mathfrak{X}')^*, \mathfrak{X}')$.

We denote by $\overline{B}^{\sigma(\mathfrak{X},\mathfrak{X}')}$ the weak closure of B in $\mathfrak{X}$, and by $\overline{B}^{\sigma((\mathfrak{X}')^*,\mathfrak{X}')}$ the $\sigma((\mathfrak{X}')^*, \mathfrak{X}')$-closure of B in $(\mathfrak{X}')^*$, respectively. Since $\overline{B}^{\sigma(\mathfrak{X},\mathfrak{X}')}$ is weakly compact by assumption, it follows that

$$\overline{B}^{\sigma(\mathfrak{X},\mathfrak{X}')} = \overline{B}^{\sigma((\mathfrak{X}')^*,\mathfrak{X}')}.$$

[13] Let $\mathfrak{X}$ be a LCHTVS and $H \subset \mathfrak{X}'$. The following statements are equivalent:

a. H is equi-continuous.
b. $H \subset V^\circ$ for some $V \in \mathcal{N}_{\mathfrak{X}}(0)$.
c. $H^\circ \in \mathcal{N}_{\mathfrak{X}}(0)$.

If $H = V^\circ$, in particular, V° is equi-continuous.

If $f \in \mathfrak{X}''$, there exists a bounded set $B \subset \mathfrak{X}$ which satisfies

$$f \in \overline{B}^{\sigma((\mathfrak{X}')^*, \mathfrak{X}')} = \overline{B}^{\sigma(\mathfrak{X}, \mathfrak{X}')} \quad (\subset \mathfrak{X})$$

by Lemma B.2. Hence $\mathfrak{X}'' \subset \mathfrak{X}$. The converse inclusion $\mathfrak{X} \subset \mathfrak{X}''$ is obvious. So $\mathfrak{X} = \mathfrak{X}''$; that is $\mathfrak{X}$ is semi-reflexive. □

We now proceed to the proof of Theorem B.15.

Lemma B.4 *Let $\mathfrak{X}$ be a locally convex topological vector space. If V is a barrel in $\mathfrak{X}'$ with respect to the w^*-topology, then V is a neighborhood of* 0 *in $\mathfrak{X}'$ with respect to the strong topology.*

Proof Since V is absorbing, it is easy to see that V° is weakly bounded. V° is also bounded with respect to the original topology by Theorem B.13. Consequently, $V = V^{\circ\circ}$ is a neighborhood of 0 in $\mathfrak{X}'$ with respect to the strong topology. □

Lemma B.5 *If $\mathfrak{X}$ is semi-reflexive, $\mathfrak{X}'$ (endowed with the strong topology) is barreled.*

Proof Let V be a barrel in $\mathfrak{X}'$ (endowed with the strong topology). Since $\mathfrak{X} = \mathfrak{X}''$, $\sigma(\mathfrak{X}', \mathfrak{X}) = \sigma(\mathfrak{X}', \mathfrak{X}'')$. Hence V is a barrel with respect to the w^*-topology. By Lemma B.4, V is a neighborhood of 0 in $\mathfrak{X}'$ with respect to the strong topology. □

Proof of Theorem B.15 If $\mathfrak{X}$ is reflexive, so is $\mathfrak{X}'$ (endowed with the strong topology). Hence $\mathfrak{X}'$ is semi-reflexive and $\mathfrak{X}'' = \mathfrak{X}$ is barreled by Lemma B.3 and Lemma B.4.

Conversely, suppose that $\mathfrak{X}$ is semi-reflexive, and barreled. Compare $\mathfrak{X}$ endowed with the original topology and $(\mathfrak{X}')'$.

Let V be a neighborhood of 0 in $\mathfrak{X}$ with respect to the original topology, and W a neighborhood of 0 in $(\mathfrak{X}')'$. We may assume that both V and W are closed, convex and balanced without loss of generality.

V° is equi-continuous since $V \in \mathcal{N}_{\mathfrak{X}}(0)$ (cf. footnote 13 on p. 375), and

$$V^\circ(x) = \{\Lambda(x) | \Lambda \in V^\circ\}$$

is a bounded set in $\mathbb{C}$ for all $x \in \mathfrak{X}$ because V is absorbing. Hence V° is w^*-relatively compact by Theorem B.11. By $\mathfrak{X} = \mathfrak{X}''$ (semi-reflexivity), $\sigma(\mathfrak{X}', \mathfrak{X}) = \sigma(\mathfrak{X}', \mathfrak{X}'')$. Therefore V° is weakly relatively compact and so weakly bounded. By Theorem B.13, V° is also bounded with respect to the original topology. It follows that

$$V^{\circ\circ} \in \mathcal{N}_{\mathfrak{X}''}(0).$$

Thus we know that the topology on $\mathfrak{X}''$ is stronger than the original topology on $\mathfrak{X}$. (The assumption that $\mathfrak{X}$ is barreled is not used here.)

Suppose next that B is a bounded set in $\mathfrak{X}'$ and define

$$W = \{x \in \mathfrak{X} | p_B(x) \leqq 1\},$$

where

$$p_B(x) = \sup_{\Lambda \in B} |\Lambda(x)|.$$

This means that $W = B^\circ$. W is a closed, convex and balanced set with respect to the original topology. It is easy to see that W is absorbing by the boundedness of B. Hence W is a barrel, and so $W \in \mathcal{N}_{\mathfrak{X}}(0)$. Thus it is verified that the original topology of $\mathfrak{X}$ is stronger than that of $\mathfrak{X}''$. Combining with the previous result, we conclude that $\mathfrak{X}$ is reflexive. □

Remark B.7 **1°** A Montel space is reflexive.
2° If $\mathfrak{X}$ is a reflexive locally convex topological vector space, $\mathfrak{X}'$ (endowed with the strong topology) is also reflexive.

Theorem B.16 (dual of Montel space) *If $\mathfrak{X}$ is a Montel space, $\mathfrak{X}'$ (endowed with the strong topology) is also a Montel space.*

Proof Since $\mathfrak{X}$ is a Montel space, $\mathfrak{X}'$ is reflexive by the above remark. Hence $\mathfrak{X}'$ is barreled by Theorem B.15. There remains to show that any bounded set in $\mathfrak{X}'$ is relatively compact.

Let A be a bounded set in $\mathfrak{X}'$. Since $\mathfrak{X}$ is barreled, A is equi-continuous by Theorem B.10. A is weakly relatively compact by the reflexivity of $\mathfrak{X}'$ and the boundedness of A. According to Theorem B.11, the w^*-topology (= the weak topology) coincides with the compact-topology on A. Hence A is relatively compact with respect to the compact topology.

Any bounded set is relatively compact, since $\mathfrak{X}$ is a Montel space. Therefore the strong topology and the compact topology coincide on $\mathfrak{X}'$. Consequently, A is relatively compact with respect to the strong topology. □

References

1. Bourbaki, N.: Eléments de mathématique; Espaces vectoriels topologiques. Hermann, Paris, Chaps. 1–2 (seconde édition) (1965), Chaps. 3–5 (1964). (English edn.) Elements of Mathematics; Topological Vector Spaces. Springer, Berlin/Heidelberg/New York (1987)
2. Grothendieck, A.: Topological Vector Spaces. Gordon and Breach, New York (1973)
3. Maruyama, T.: Kansu Kaisekigaku (Functional Analysis). Keio Tsushin, Tokyo (1980) (Originally published in Japanese)
4. Rudin, W.: Functional Analysis. McGraw-Hill, New York (1973)

5. Schwartz, L.: Functional Analysis. Courant Institute of Mathematical Sciences, New York University, New York (1964)
6. Schwartz, L.: Théorìe des distributions. Nouvelle édition, entièrement corrigée, refondue et augmentée, Hermann, Paris (1973)
7. Yosida, K.: Functional Analysis, 3rd edn. Springer, New York (1971)

Appendix C
Theory of Distributions

In this appendix, we discuss basic elements of the theory of distributions due to L. Schwartz, which are required to understand the text rigorously. We freely make use of various concepts and results in abstract functional analysis which were supplied in Appendix B.[1]

C.1 The Space $\mathfrak{D}$

A distribution is a continuous linear functional on the space $\mathfrak{D}$ consisting of very smooth functions. We start by constructing the basic space $\mathfrak{D}$.

Let Ω be a nonempty open set in $\mathbb{R}^l$. We denote[2] by $\mathfrak{E}(\Omega)$ the set of all the complex-valued functions defined on Ω which are differentiable infinitely many times. We restrict our attention to the case of $l = 1$ for the sake of simplicity. D^s denotes the differential operator to take the s-th derivative. We define a seminorm $p_{K,m}$ on $\mathfrak{E}(\Omega)$ by

$$p_{K,m}(\varphi) = \sup_{\substack{x \in K \\ s \leqq m}} |D^s \varphi(x)|$$

for each compact set $K \subset \Omega$ and $m \in \mathbb{N} \cup \{0\}$. $\mathfrak{E}(\Omega)$ becomes a LCHTVS endowed with the topology generated by the family $\{p_{K,m} | K \subset \Omega \text{ is compact}, m \in \mathbb{N} \cup \{0\}\}$.

Theorem C.1 (space $\mathfrak{E}(\Omega)$) *$\mathfrak{E}(\Omega)$ is a Fréchet space.*

Proof We can construct a sequence $\{K_1, K_2, \cdots\}$ of compact sets in Ω which satisfies

[1] cf. Schwartz [6] Chaps. I–III, Yosida [8] pp. 46–52, 62–64.

[2] Of course, we may write $\mathfrak{E}^\infty(\Omega, \mathbb{C})$. But we follow the tradition of distribution theory here.

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8

(a) $K_n \subset \text{int}.K_{n+1}$,
(b) $\bigcup_{n=1}^{\infty} K_n = \Omega$.

(For instance, the sequence defined by

$$K_n = \overline{B_n(0)} \bigcap \left\{ x \in \mathbb{R} \,\middle|\, \rho(\Omega^c, x) \geqq \frac{1}{n} \right\} \quad n = 1, 2, \cdots$$

does work.) Then any compact set in Ω is contained in some K_n. Hence the family

$$\{p_{K_n,m} | m, n \in \mathbb{N}\}$$

of countable seminorms determines a topology on $\mathfrak{E}(\Omega)$. $\mathfrak{E}(\Omega)$ endowed with this topology is metrizable.

We next show the completeness. Let $\{\varphi_n\}$ be a Cauchy sequence in $\mathfrak{E}(\Omega)$. Then for any differential operator D^s, $\{D^s\varphi_n\}$ converges to some function $g_s : \Omega \to \mathbb{C}$ uniformly in wide sense. In particular,

$$\varphi_n(x) = D^0\varphi_n(x) \to g_0(x) \; ; \quad x \in \Omega.$$

It follows that $g_s = D^s g_0$. Consequently,

$$g_0 \in \mathfrak{E}(\Omega) \quad \text{and} \quad \varphi_n \to g_0 \quad \text{in } \mathfrak{E}(\Omega).$$

Thus $\mathfrak{E}(\Omega)$ is complete. □

Theorem C.2 (i) *A bounded set in* $\mathfrak{E}(\Omega)$ *is totally bounded.*
(ii) $\mathfrak{E}(\Omega)$ *is not normable.*

Proof (i) Let $\{p_{K_n,m}\}$ be a family of countable seminorms which defines the topology on $\mathfrak{E}(\Omega)$. If A is a bounded set in $\mathfrak{E}(\Omega)$, then

$$\sup_{\varphi \in A} p_{K_n,m}(\varphi) \equiv C_{K_{n,m}} < \infty \quad \text{for each} \quad n, m \in \mathbb{N} \cup \{0\}.$$

For any sequence $\{\varphi_j\}$ in A, $\{D^s\varphi_j\}$ is uniformly bounded and equi-continuous on each K_n and for each differential operator D^s. Then applying the Ascoli–Arzelà theorem (Theorem B.11, p. 368) and Cantor's diagonal method, $\{\varphi_j\}$ has a convergent subsequence in $\mathfrak{E}(\Omega)$. That is, A is sequentially compact in $\mathfrak{E}(\Omega)$, and so totally bounded.

It is recommended that readers prove (ii).[3] □

[3]Let $\mathfrak{X}$ be a Hausdorff topological vector space. The topology is defined by a norm if and only if there exists a convex and bounded neighborhood of 0.

Let K be a compact set in Ω. We define a subspace $\mathfrak{D}_K(\Omega)$ of $\mathfrak{E}(\Omega)$ by

$$\mathfrak{D}_K(\Omega) = \{\varphi \in \mathfrak{E}(\Omega) | \operatorname{supp}\varphi \subset K\}.$$

The relative topology of $\mathfrak{D}_K(\Omega)$ induced by $\mathfrak{E}(\Omega)$ is determined by the family

$$p_{K,m}(\varphi) = \sup_{\substack{x \in K \\ s \leqq m}} |D^s \varphi(x)|; \quad m \in \mathbb{N} \cup \{0\}$$

of seminorms, and $\mathfrak{D}_K(\Omega)$ becomes a LCHTVS.

Theorem C.3 (space $\mathfrak{D}_K$) *$\mathfrak{D}_K(\Omega)$ is a Fréchet space.*

Proof Define a linear functional $\Lambda_x : \mathfrak{E}(\Omega) \to \mathbb{C} (x \in \Omega)$ by

$$\Lambda_x : \varphi \mapsto \varphi(x).$$

Λ_x is continuous by the definition of the topology on $\mathfrak{E}(\Omega)$. Since

$$\mathfrak{D}_K(\Omega) = \bigcap_{x \in \Omega \setminus K} \operatorname{Ker} \Lambda_x,$$

$\mathfrak{D}_K(\Omega)$ is a closed subspace of $\mathfrak{E}(\Omega)$. So $\mathfrak{D}_K(\Omega)$ is a Fréchet space. □

We denote by $\mathfrak{D}(\Omega)$ the set of all elements of $\mathfrak{E}(\Omega)$ with compact support;[4] i.e.

$$\begin{aligned}\mathfrak{D}(\Omega) &= \{\varphi \in \mathfrak{E}(\Omega) | \operatorname{supp}\varphi \text{ is compact}\} \\ &= \bigcup \{\mathfrak{D}_K(\Omega) | \ K \text{ is compact subset in } \Omega\}.\end{aligned}$$

If we choose a sequence $\{K_n\}$ of compact sets in Ω so that it satisfies (a) and (b) on p. 380, the sequence $\{\mathfrak{D}_{K_n}(\Omega)\}$ of Fréchet spaces determines the strictly inductive limit topology on

$$\mathfrak{D}(\Omega) = \bigcup_{n=1}^{\infty} \mathfrak{D}_{K_n}(\Omega).$$

This topology is determined independently of the choice of $\{K_n\}$ (cf. Remark B.2, p. 362).

Elements of $\mathfrak{D}(\Omega)$ are called **test functions**.

[4] We may use a more ordinary notation $\mathfrak{C}_0^{\infty}(\Omega, \mathbb{C})$ instead of $\mathfrak{D}(\Omega)$. Again we follow here the tradition of distribution theory.

The basic properties of $\mathfrak{D}(\Omega)$ are summarized in the next theorem; they are direct consequences of the general theory of inductive limit topology (cf. Appendix B, Sect. B.1).

Theorem C.4 (basic properties of $\mathfrak{D}(\Omega)$)

(i) *$\mathfrak{D}(\Omega)$ is a LCHTVS.*
(ii) *Let K be a compact set in Ω. Then the relative topology of $\mathfrak{D}_K(\Omega)$ induced by $\mathfrak{D}(\Omega)$ coincides with the original topology $\mathfrak{D}_K(\Omega)$.*
(iii) *$B \subset \mathfrak{D}(\Omega)$ is bounded if and only if B is a bounded set in $\mathfrak{D}_K(\Omega)$ for some compact set $K \subset \Omega$.*
(iv) *A net $\{\varphi_\alpha\}$ in $\mathfrak{D}(\Omega)$ converges to $\varphi^* \in \mathfrak{D}(\Omega)$ if and only if there exists some compact set $K \subset \Omega$ (independent of α) such that*

$$\mathrm{supp}\varphi_\alpha \subset K \quad \textit{for all } \alpha$$

and

$$D^s\varphi_\alpha \to D^s\varphi^* \quad \textit{uniformly}$$
$$(\textit{i.e.} \quad \varphi_\alpha \to \varphi^* \quad \textit{in } \mathfrak{D}_K(\Omega))$$

for any differential operator D^s.
(v) *$\mathfrak{D}(\Omega)$ is complete.*

Proof (i)–(iv) are obvious in view of the general theory discussed in Appendix B.1. So it is enough to examine only (v).[5]

Let $\{\varphi_\alpha\}$ be a Cauchy net in $\mathfrak{D}(\Omega)$. The identity mapping

$$I : \mathfrak{D}(\Omega) \to \mathfrak{E}(\Omega)$$

is continuous. Hence $\{\varphi_\alpha\}$ is also a Cauchy net in $\mathfrak{E}(\Omega)$. Since $\mathfrak{E}(\Omega)$ is a Fréchet space by Theorem C.1, there exists some $\varphi^* \in \mathfrak{E}(\Omega)$ such that

$$\varphi_\alpha \to \varphi^* \quad \text{in } \mathfrak{E}(\Omega).$$

We have only to show that the support of φ^* is compact.

Suppose that the support of φ^* is not compact. There exist a sequence $\{K_n\}$ of compact sets, which satisfies (a) and (b) on p. 380, and a sequence $\{x_n\}$ such that

$$x_n \in K_n \setminus K_{n-1}, \quad \varphi^*(x_n) \neq 0 \quad (n = 1, 2, \cdots).$$

[5] We owe this to Choquet [2] Vol.I, pp. 305–306.

If we define

$$V = \left\{ \varphi \in \mathfrak{D}(\Omega) \middle| \varphi \in \mathfrak{D}_{K_n}(\Omega) \Rightarrow \|\varphi\|_\infty < \frac{|\varphi^*(x_n)|}{2} \right\},$$

V is a neighborhood of 0 in $\mathfrak{D}(\Omega)$. Since $\{\varphi_\alpha\}$ is Cauchy, there exists some α_0 such that

$$\varphi_\alpha - \varphi_\beta \in V \quad \text{for all} \quad \alpha, \beta \succ \alpha_0.$$

Since the support of each φ_α is compact,

$$\varphi_{\alpha_0}(x_n) = 0$$

for sufficiently large n. Hence (setting $\beta = \alpha_0$) we have

$$|\varphi_\alpha(x_n)| < \frac{|\varphi^*(x_n)|}{2} \quad \text{for all} \quad \alpha \succ \alpha_0.$$

This, however, contradicts

$$\varphi_\alpha(x) \to \varphi^*(x) \quad \text{pointwise}.$$

□[6]

Remark C.1 If $\{K_n\}$ is a sequence of compact sets in Ω such that

$$K_n \subset \text{int. } K_{n+1}, \quad \bigcup_{n=1}^{\infty} K_n = \Omega,$$

then it is clear that

$$\mathfrak{D}_{K_n}(\Omega) \subset \mathfrak{D}_{K_{n+1}}(\Omega); \quad \bigcup_{n=1}^{\infty} \mathfrak{D}_{K_n}(\Omega) = \mathfrak{D}(\Omega).$$

(i) Each $\mathfrak{D}_{K_n}(\Omega)$ is a closed subspace of $\mathfrak{D}(\Omega)$.
(ii) int.$\mathfrak{D}_{K_n}(\Omega) = \emptyset$ for each n.
(iii) $\mathfrak{D}(\Omega)$ is not metrizable.

The next corollary immediately follows from Theorem C.1, Theorem C.3, Corollary B.4 (p. 365) and Theorem B.8 (p. 365).

Corollary C.1 *$\mathfrak{E}(\Omega)$, $\mathfrak{D}_K(\Omega)$ and $\mathfrak{D}(\Omega)$ are barreled spaces.*

---

[6]It is much easier to show that any Cauchy sequence in $\mathfrak{D}(\Omega)$ converges in $\mathfrak{D}(\Omega)$ itself.

Theorem C.5 *$\mathfrak{D}(\Omega)$ and $\mathfrak{E}(\Omega)$ are Montel spaces.*

Proof Let B a bounded set in $\mathfrak{D}(\Omega)$. By Theorem B.5 (p. 362), B is a bounded set in $\mathfrak{D}_K(\Omega)$ for some compact $K \subset \Omega$. Note that $\mathfrak{D}_K(\Omega)$ is metrizable. Let $\{\varphi_j\}$ be a sequence in B. Then $\{D^s\varphi_j\}$ is uniformly bounded and equi-continuous on K for any differential operator D^s. Hence $\{\varphi_j\}$ has a subsequence convergent in $\mathfrak{D}_K(\Omega)$ by the Ascoli–Arzelà theorem (Theorem B.11, p. 368) and Cantor's diagonal process. That is, B is sequentially compact in $\mathfrak{D}_K(\Omega)$, and so relatively compact in $\mathfrak{D}_K(\Omega)$. Since the identical embedding $I : \mathfrak{D}_K(\Omega) \to \mathfrak{D}(\Omega)$ is continuous (Theorem B.1, p. 358),

$$I(\mathrm{cl}._{\mathfrak{D}_K(\Omega)}B) = \mathrm{cl}._{\mathfrak{D}_K(\Omega)}B$$

is a compact set in $\mathfrak{D}(\Omega)$, which contains B. Hence B is relatively compact.

We can show that $\mathfrak{E}(\Omega)$ is a Montel space in a similar way. □

C.2 Examples of Test Functions and an Approximation Theorem

In this section, some important examples of test functions will be shown. We also discuss approximation of continuous functions by elements of $\mathfrak{D}(\Omega)$, as well as certain "ampleness" of $\mathfrak{D}(\Omega)$.

We prepare a few results, the first of which is a special case of the finite increment formula.

Lemma C.1 *Let I be an open interval in $\mathbb{R}$. Consider a function $f : I \to \mathbb{R}^n$.*

(i) *If f is differentiable on I, then*

$$\|f(y) - f(x)\| \leqq |y - x| \sup_{0 \leqq t \leqq 1} \|Df(x + t(y - x))\|$$

for all $x, y \in I$.

(ii) *If f is continuous on $[x, y]$ and differentiable on (x, y) for $x, y \in I$, then*

$$\|f(y) - f(x)\| \leqq |y - x| \sup_{0 < t < 1} \|Df(x + t(y - x))\|.$$

(iii) *Under the same assumptions as in (ii),*

$$\|f(y) - f(x) - \langle v, y - x\rangle\| \leqq |y - x| \sup_{0 < t < 1} \|Df(x + t(y - x)) - v\|$$

for $v \in \mathbb{R}^n$.

Proof (i) Choose a sufficiently large $M > 0$ so that

$$M > \sup_{0 \leqq t \leqq 1} \|Df(x + t(y - x))\|.$$

(If the right-hand side is ∞, (i) is obvious.) Define a set E by

$$E = \{t \in [0, 1] \mid \|f(x + t(y - x)) - f(x)\| \leqq Mt|x - y|\}.$$

E is closed by the continuity of f, and clearly $0 \in E$. Hence E has the maximum, say s. Suppose that $s < 1$. If $t > s$ and $t - s$ is sufficiently small, then it follows that

$$\begin{aligned}\|f(x + t(y - x)) - f(x)\| &\leqq \|f(x + t(y - x)) - f(x + s(y - x))\| \\ &\quad + \|f(x + s(y - x)) - f(x)\| \\ &\leqq M|(t - s)(y - x)| + Ms|y - x| \\ &= Mt|y - x|.\end{aligned}$$

Hence t is also an element of E. It contradicts the maximality of s on E. So we must have $s = 1$.

(ii) and (iii) are easy.[7] □

Lemma C.2 *Let I be an open interval in $\mathbb{R}$ and $F \subset I$ a closed set. Suppose that a continuous function $f : I \to \mathbb{R}^n$ is identically zero on F and differentiable on $I \setminus F$. Suppose further that $x \in \partial F$ and a sequence $\{y_j\}$ in $I \setminus F$ converges to x. If $Df(y_j) \to 0$ (as $j \to \infty$), then $Df(x)$ exists and is equal to* 0.

The next lemma follows from Lemma C.2.

Lemma C.3 *Let P be a polynomial. Define a function $f : \mathbb{R} \to \mathbb{R}$ by*

$$f(x) = \begin{cases} P\left(\frac{1}{x}\right)\exp\left(-\frac{1}{x}\right) & \text{if} \quad x > 0, \\ 0 & \text{if} \quad x \leqq 0. \end{cases}$$

Then f is differentiable at 0 *and* $Df(0) = 0$.[8]

We now proceed to discuss a typical example of test functions.

Example C.1 Define a function $\varphi : \mathbb{R} \to \mathbb{R}$ by

$$\varphi(x) = \begin{cases} 0 & \text{if} \quad |x| \geqq 1, \\ \exp\left(-\frac{1}{1-x^2}\right) & \text{if} \quad |x| < 1. \end{cases}$$

[7] Consider $x \mapsto f(x) - xv$ for (iii).

[8] We acknowledge Hölmander [3] I, pp. 6–7 for Lemma C.1–C.3.

Then supp$\varphi = [-1, 1]$ and φ is infinitely differentiable at any x with $|x| \gtrless 1$. We have only to check the smoothness of φ at $x = \pm 1$.

If we define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(t) = \begin{cases} \exp\left(-\frac{1}{t}\right) & \text{if} \quad t > 0, \\ 0 & \text{if} \quad t \leqq 0, \end{cases}$$

g is infinitely differentiable at all points (including 0) by repeated application of Lemma C.3. If we also define $h : \mathbb{R} \to \mathbb{R}$ by

$$h(x) = 1 - x^2,$$

it is also infinitely differentiable. Since φ can be represented as

$$\varphi(x) = (g \circ h)(x),$$

φ is infinitely differentiable everywhere. Hence $\varphi \in \mathfrak{D}(\mathbb{R})$.

In the next section, we define distributions as continuous linear functionals on $\mathfrak{D}(\Omega)$. That is, $\mathfrak{D}(\Omega)'$ is the space of distributions. Although detailed discussion is postponed to the next section, we should now observe that any locally integrable function $f : \Omega \to \mathbb{C}$ defines a distribution T_f by

$$T_f(\varphi) = \int_\Omega f(x)\varphi(x)dx, \quad \varphi \in \mathfrak{D}(\Omega).$$

Given two different locally integrable functions f and g, we may ask whether $\mathfrak{D}(\Omega)$ is ample enough to discriminate T_f and T_g. That is, is there any $\varphi \in \mathfrak{D}(\Omega)$ such that $T_f(\varphi) \neq T_g(\varphi)$ for any different locally integrable functions, f and g?

Theorem C.6 *Let $f : \Omega \to \mathbb{C}$ be a locally integrable function. If*

$$\int_\Omega f(x)\varphi(x)dx = 0$$

for any $\varphi \in \mathfrak{D}(\Omega)$, then $f(x) = 0$ a.e. $x \in \Omega$.

Proof We divide our proof in three steps.

1° For any continuous function $\varphi : \Omega \to \mathbb{C}$ with compact support,

$$\int_\Omega f(x)\psi(x)dx = 0.$$

For any $\varepsilon > 0$ and $\delta > 0$, there exists a function $\varphi \in \mathfrak{D}(\Omega)$ such that

$$\text{supp}\varphi \subset B_\delta(\text{supp}\psi), \quad \|\varphi - \psi\|_\infty \leqq \varepsilon.$$

φ satisfies

$$\left|\int_{\Omega}(\varphi(x)-\psi(x))f(x)dx\right| \leqq \varepsilon \int_{B_\delta(\mathrm{supp}\psi)} |f(x)|dx.$$

Since

$$\int_{\Omega} f(x)\varphi(x)dx = 0$$

by assumption, we obtain

$$\left|\int_{\Omega} f(x)\psi(x)dx\right| \leqq \varepsilon \int_{B_\delta(\mathrm{supp}\psi)} |f(x)|dx.$$

Passing to the limit as $\varepsilon \to 0$ (δ being fixed), we verify 1°.

2° For any bounded measurable function $\eta : \Omega \to \mathbb{C}$ with compact support,

$$\int_{\Omega} f(x)\eta(x)dx = 0.$$

We start by proving the existence of a sequence $\{\psi_n : \Omega \to \mathbb{C}\}$ of continuous functions which satisfies:

a. $\psi_n(x) \to \eta(x)$ a.e.,
b. $\mathrm{supp}\psi_n \subset K$ for all n for some compact set $K \subset \Omega$, and
c. $\|\psi_n\|_\infty \leqq \|\eta\|_\infty$.

As is well-known, there exists a sequence $\{\theta_n : \Omega \to \mathbb{C}\}$ of continuous functions which converges to η a.e.[9] Let $K \subset \Omega$ be a compact set such that $\mathrm{supp}\eta \subset \mathrm{int}.K$. Then there exists a continuous function $\alpha : \Omega \to [0, 1]$ such that[10]

$$\alpha(\mathrm{supp}\eta) = \{1\}, \quad \mathrm{supp}\alpha \subset K.$$

Define $\tilde{\theta}_n : \Omega \to \mathbb{C}$ by

$$\tilde{\theta}_n(x) = \alpha(x)\theta_n(x); \quad n = 1, 2, \cdots$$

Each $\tilde{\theta}_n$ is a continuous function with $\mathrm{supp}\tilde{\theta}_n \subset K$ and $\tilde{\theta}_n \to \eta$ a.e. So $\{\tilde{\theta}_n\}$ satisfies a and b.

We express $\tilde{\theta}_n(x)$ as

$$\tilde{\theta}_n(x) = |\tilde{\theta}_n(x)|e^{i \cdot arg\tilde{\theta}_n(x)}$$

[9] Maruyama [5] pp. 232–233.
[10] Maruyama [5] p. 317n.

and define $\beta_n(x)$ and $\psi_n(x)$ by

$$\beta_n(x) = \mathrm{Min}\{\|\eta\|_\infty, |\tilde{\theta}_n(x)|\},$$

$$\psi_n(x) = \beta_n(x)e^{i\cdot arg\tilde{\theta}_n(x)}; \quad n = 1, 2, \cdots$$

Then $\{\psi_n\}$ certainly satisfies all of a, b and c. By 1°, we have

$$\int_\Omega f(x)\psi_n(x)dx = 0; \quad n = 1, 2, \cdots .$$

On the other hand, it follows that

$$\int_\Omega f(x)\eta(x)dx = \int_K \lim_{n\to\infty} f(x)\psi_n(x)dx = \lim_{n\to\infty}\int_\Omega f(x)\psi_n(x)dx = 0$$

by a, b and the dominated convergence theorem.[11] Thus we obtain 2°.

3° $f(x) = 0$ a.e.

Specify a bounded measurable function $\eta : \Omega \to \mathbb{C}$ with compact support as

$$\eta(x) = \begin{cases} 0 & \text{if} \quad |x| > a \quad \text{or} \quad f(x) = 0, \\ e^{-i\cdot argf(x)} & \text{if} \quad |x| \leqq a \quad \text{and} \quad f(x) \neq 0, \end{cases}$$

where $a > 0$. η thus defined is measurable and $|\eta(x)| = 0$ or 1. It is obvious that supp$\eta \subset \{x \in \Omega \mid \|x\| \leqq a\}$. We obtain

$$\int_\Omega f(x)\eta(x)dx = \int_{|x|\leqq a} |f(x)|dx = 0$$

by 2°. Since this holds good for all $a > 0$, $f(x) = 0$ a.e.[12] □

C.3 Distributions: Definition and Examples

We now define the distributions in the sense of L. Schwartz and give several examples.

[11] $|f(x)\psi_n(x)| \leqq \|\eta\|_\infty |f(x)|$, and the right-hand side is integrable.

[12] The proof of Theorem C.5 is due to Schwartz [7] pp. 68–70. The significance of the theorem is discussed clearly in Kolmogorov and Fomin [4] pp. 212–213 (Japanese edn.).

Definition C.1 Each element of $\mathfrak{D}(\Omega)'$, the dual space of $\mathfrak{D}(\Omega)$, is called a **distribution**.

That is, a distribution is a continuous linear functional on $\mathfrak{D}(\Omega)$, the topological structure of which is discussed in Appendix B.

Example C.2 Let $f; \Omega \to \mathbb{C}$ be a locally integrable function. Define an operator $T_f : \mathfrak{D}(\Omega) \to \mathbb{C}$ by

$$T_f(\varphi) = \int_\Omega f(x)\varphi(x)dx.$$

T_f is well-defined, and it is clearly linear. The continuity can be proved as follows. Let $\{\varphi_\alpha\}$ be a net in $\mathfrak{D}(\Omega)$ which converges to $\varphi^* \in \mathfrak{D}(\Omega)$. There exists a compact set $K \subset \Omega$ such that

$$\mathrm{supp}\varphi_\alpha \subset K.$$

(cf. Theorem C.4.) Then it follows that

$$|T_f(\varphi_\alpha) - T_f(\varphi^*)| = \left|\int_K f(x)\{\varphi_\alpha(x) - \varphi^*(x)\}dx\right| \leqq \int_K |f(x)|dx \cdot \sup_x |\varphi_\alpha(x) - \varphi^*(x)|.$$

Since f is locally integrable and $\|\varphi_\alpha - \varphi^*\|_\infty \to 0$, we obtain

$$|T_f(\varphi_\alpha) - T_f(\varphi^*)| \to 0.$$

Hence T_f is continuous.

According to Theorem C.6,

$$f = g \quad \text{a.e.} \quad \Leftrightarrow \quad T_f = T_g$$

for any couple of locally integrable functions, f and g. If we identify two locally integrable functions which are equal a.e., there exists one-to-one correspondence between a locally integrable function f and the distribution T_f which is defined by f. Therefore we sometimes confuse f and T_f consciously and write $f(\varphi)$ instead of $T_f(\varphi)$.

Example C.3 Let D be an arbitrary differential operator and f a locally integrable function. If we define an an operator $T : \mathfrak{D}(\Omega) \to \mathbb{C}$ by

$$T(\varphi) = \int_\Omega f(x)D\varphi(x)dx,$$

T is a distribution.

Example C.4 Define an operator $\delta : \mathfrak{D}(\Omega) \to \mathbb{C}$ by

$$\delta(\varphi) = \varphi(0).$$

δ is a distribution, called the **Dirac distribution**. Similarly, an operator $\delta_{(a)}$ defined by

$$\delta_{(a)}(\varphi) = \varphi(a); \quad a \in \Omega$$

is called the Dirac distribution which assigns a mass 1 at a point a.

Example C.5 Define a space $\mathfrak{K}_K(\Omega)$ of functions by

$$\mathfrak{K}_K(\Omega) = \{\varphi \in \mathfrak{C}(\Omega, \mathbb{C}) | \mathrm{supp}\varphi \subset K\},$$

where $K \subset \Omega$ is compact. That is, $\mathfrak{K}_K(\Omega)$ is the space of all continuous functions, the supports of which are contained in K. $\mathfrak{K}_K(\Omega)$ is a linear space under usual operations. If we define $p_K(\cdot)$ on $\mathfrak{K}_K(\Omega)$ by

$$p_K(\varphi) = \sup_{x \in K} |\varphi(x)|,$$

$p_K(\cdot)$ is a norm. $\mathfrak{K}_K(\Omega)$ is a Banach space under this norm.

Define $\mathfrak{K}(\Omega)$ by

$$\mathfrak{K}(\Omega) = \bigcup \{\mathfrak{K}_K(\Omega) | K \text{ is a compact set in } \Omega\}.$$

Then the strictly inductive limit topology can be defined on $\mathfrak{K}(\Omega)$ by $\{\mathfrak{K}_K(\Omega)\}$ as in $\mathfrak{D}(\Omega)$.

1° $B \subset \mathfrak{K}(\Omega)$ is bounded if and only if B is bounded in $\mathfrak{K}_K(\Omega)$ for some compact set $K \subset \Omega$.

2° A net $\{\varphi_\alpha\}$ in $\mathfrak{K}(\Omega)$ converges to $\varphi^* \in \mathfrak{K}(\Omega)$ if and only if there exists some compact set $K \subset \Omega$ (independent of α) such that

$$\mathrm{supp}\varphi_\alpha \subset K \quad \text{for all } \alpha$$

and

$$\varphi_\alpha \to \varphi^* \quad \text{uniformly.}$$

These results immediately follow from the general theory discussed in Appendix B. Each element of $\mathfrak{K}(\Omega)'$, the dual space of $\mathfrak{K}(\Omega)$, is called a **Radon measure** on Ω.[13]

[13] In Sect. 4.1, a continuous linear functional on the Banach space $\mathfrak{C}_\infty$ (with uniform convergence norm) is called a Radon measure. We give the same name to each element of $\mathfrak{K}(\Omega)'$. However,

A Radon measure is obviously a distribution.

We add here a few remarks concerning the relation between distributions and Radon measures.

Remark C.2 A distribution T is defined by some Radon measure μ (i.e. the restriction of μ to $\mathfrak{D}$ is equal to T) if and only if T is continuous on $\mathfrak{D}_K$ endowed with the relative topology induced by $\mathfrak{K}_K$ for any compact set $K \subset \Omega$. This proposition can be proved as follows. (It is enough to show only the sufficiency since the necessity is obvious.)

(i) Let H be a compact set in $\mathbb{R}$. We denote by $\overline{\mathfrak{D}}_H^{\mathfrak{K}_H}$ the closure of $\mathfrak{D}_H$ in $\mathfrak{K}_H$. Then $T|_{\mathfrak{D}_H}$ can be extended to continuous linear functional $\bar{T}_H$ on $\overline{\mathfrak{D}}_H^{\mathfrak{K}_H}$. ($\overline{\mathfrak{D}}_H^{\mathfrak{K}_H}$ is endowed with the relative topology induced by $\mathfrak{K}_H$.)

(ii) T can be extended to a linear functional $\bar{T}$ on $\bigcup\{\overline{\mathfrak{D}}_H^{\mathfrak{K}_H} | H \text{ is compact}\}$ with $\bar{T} = \bar{T}_H$ for each compact set H.

(iii) $\bigcup\{\overline{\mathfrak{D}}_H^{\mathfrak{K}_H} | H \text{ is compact}\} = \mathfrak{K}$.

(iv) $\bar{T}$ is a Radon measure.

(v) Set $\bar{T} = \mu$. μ is uniquely determined by T.

Thus there is a one-to-one correspondence between some subspace of $\mathfrak{D}'$ and $\mathfrak{K}'$. So a Radon measure can be identified with the corresponding distribution.

A distribution T is said to be real valued if $T(\varphi)$ is real for any real-valued $\varphi \in \mathfrak{D}$. A distribution T is called **positive** if $T(\varphi) \geqq 0$ for any $\varphi \in \mathfrak{D}$ with $\varphi \geqq 0$. **A positive distribution is a Radon measure**.

To see this, we have to show a positive distribution T is continuous on $\mathfrak{D}_K$ endowed with the relative topology induced by $\mathfrak{K}_K$ for each compact K. Let $\{\varphi_\alpha\}$ be a net in $\mathfrak{D}_K$ such that φ_α uniformly converges to 0. Choose $\psi \in \mathfrak{D}$ so that

$$\psi \geqq 0 \quad \text{on } \mathbb{R}\,; \qquad \psi \geqq 1 \quad \text{on } K.$$

Then there exists $\varepsilon_\alpha > 0$ such that

$$|\varphi_\alpha| \leqq \varepsilon_\alpha \psi; \qquad \varepsilon_\alpha \downarrow 0.$$

Write φ_α in the form

---

a rigorous distinction is required. If we restrict $\nu \in \mathfrak{C}'_\infty$ to $\mathfrak{K}$, we obtain $\nu|_{\mathfrak{K}} \in \mathfrak{K}'$. If a net $\{\varphi_\alpha\}$ in $\mathfrak{K}$ converges to $\varphi^* \in \mathfrak{K}$, then there exists some compact set K such that $\text{supp}\varphi_\alpha \subset K$ for all α and $\varphi_\alpha \to \varphi^*$ in $\mathfrak{K}_K$. Hence $\{\varphi_\alpha\}$ uniformly converges to φ^* in $\mathfrak{C}_\infty$. It follows that $\nu(\varphi_\alpha) \to \nu(\varphi^*)$ since $\nu|_{\mathfrak{K}}$ is continuous with respect to the uniform convergence topology. Thus, the concept of a Radon measure as an element of $\mathfrak{K}'$ is more general than that in the sense of Chap. 4. See Bourbaki [1], Choquet [2] Vol. I, Chap. 3 for Radon measures in general. Furthermore, Schwartz [7] Chap. 1.4 should be consulted for a correspondence between distributions and Radon measures.

$$\varphi_\alpha = u_\alpha + i v_\alpha; \quad u_\alpha, v_\alpha \in \mathfrak{D}.$$

Then we obtain

$$-\varepsilon_\alpha \psi \leqq u_\alpha,\ v_\alpha \leqq +\varepsilon_\alpha \psi.$$

Hence it holds good that

$$|T(\varphi_\alpha)| \leqq 2\varepsilon_\alpha T(\psi),$$

since T is a positive distribution.

C.4 Differentiation of Distributions

Let $f : \mathbb{R} \to \mathbb{C}$ be a continuously differentiable function. We denote the distributions defined by f and f' by the same notation.

Evaluating the distribution f', we obtain

$$f'(\varphi) = \int_{\mathbb{R}} f' \cdot \varphi dx = \int_{-\infty}^{+\infty} f' \cdot \varphi dx$$

for $\varphi \in \mathfrak{D}(\mathbb{R})$. By integration by parts, it is equal to

$$f \cdot \varphi \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} f \varphi' dx.$$

Since the support of φ is bounded, the first term is 0. Hence

$$f'(\varphi) = -\int_{\mathbb{R}} f \varphi' dx. \tag{*}$$

This shows how to evaluate values of the distribution defined by the derivative of a continuously differentiable function f. It is easy to generalize this reasoning to the concept of the derivative of a distribution.

For a distribution T on $\mathbb{R}$, we define a new linear functional S on $\mathfrak{D}(\Omega)$ by

$$S\varphi = -T(\varphi'); \quad \varphi \in \mathfrak{D}(\mathbb{R}).$$

Then S is also a distribution. We call it the **generalized derivative** or **distributional derivative** of T, and denote it by T'.

We next evaluate the second distributional derivative T'' of T as follows:

$$T''\varphi = -T'(\varphi') = +T(\varphi'').$$

More generally, the p-th distributional derivative D^pT of T is given by

$$D^pT(\varphi) = (-1)^pT(D^p\varphi).$$

Summing up:

Theorem C.7 *Any distribution $T \in \mathfrak{D}(\mathbb{R})'$ has distributional derivatives of all orders, and the p-th distributional derivative D^pT ($p \in \mathbb{N}$) is given by*

$$D^pT(\varphi) = (-1)^pT(D^p\varphi); \quad \varphi \in \mathfrak{D}(\mathbb{R}).$$

Example C.6 Let $Y(x)$ be the **Heaviside function**; i.e.

$$Y(x) = \begin{cases} 0 & \text{if} \quad x < 0, \\ 1 & \text{if} \quad x > 0. \end{cases}$$

It is not necessary to define a value at $x = 0$. Y can be regarded as a distribution, which is evaluated as

$$Y(\varphi) = \int_0^{+\infty} \varphi(x)dx; \quad \varphi \in \mathfrak{D}(\mathbb{R}).$$

The distributional derivative DY of Y is evaluated as follows:

$$\begin{aligned} DY(\varphi) = Y(D\varphi) &= -\int_{-\infty}^{+\infty} Y(x)\varphi'(x)dx = -\int_0^{+\infty} \varphi'(x)dx = -\varphi(x)\Big|_0^{+\infty} \\ &= -\lim_{a\to+\infty}(\varphi(a) - \varphi(0)) = \varphi(0) = \delta(\varphi)\ ;\ \varphi \in \mathfrak{D}(\mathbb{R}). \end{aligned}$$

Thus $DY = Y'$ is exactly equal to the Dirac distribution δ.

The repeated distributional derivatives of Y are given by

$$Y''(\varphi) = \delta'(\varphi) = -\delta(\varphi') = -\varphi'(0).$$

$$D^{p+1}Y(\varphi) = D^p\delta(\varphi) = (-1)^p\delta(D^p\varphi) = (-1)^pD^p\varphi(0).$$

Example C.7 Suppose that countable points $\{x_\nu\} = \{\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots\}$ are given on $\mathbb{R}$ so that $\lim_{\nu\to\pm\infty} x_\nu = \pm\infty$ and intervals $[x_{\nu-1}, x_\nu]$ are nondegenerate. Let f be "piecewise smooth", that is, infinitely differentiable (in the usual sense) on each $(x_{\nu-1}, x_\nu)$. Assume further that each of the successive derivatives (in the usual sense) of f have discontinuities of the first kind at each x_ν.

$$\sigma_\nu^p = D^pf(x_\nu + 0) - D^pf(x_\nu - 0)$$

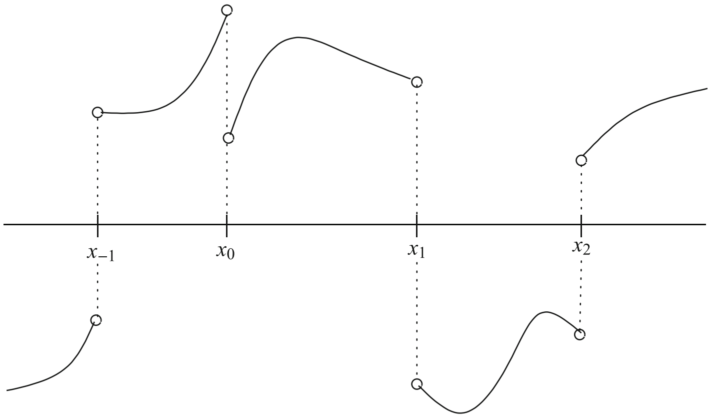


Fig. C.1 Graph of f

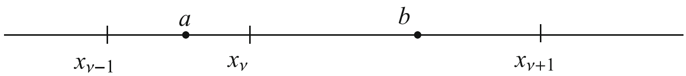


Fig. C.2 Computation of f'

is the "jump" of the p-th derivative of f at x_ν. (The values of f at $x_\nu' s$ are not required to be defined. cf. Fig. C.1.)

We denote by

$$f', f'', \cdots, f^{(p)}, \cdots \tag{*}$$

the successive distributional derivatives of f, and by

$$[f'], [f''], \cdots, [f^{(p)}], \cdots \tag{**}$$

the distributions defined by the (locally integrable) functions which coincide with the successive derivatives (in the usual sense) of f on each $(x_{\nu-1}, x_\nu)$ and are not necessarily defined at x_ν's. We should carefully distinguish (*) and (**). (For instance, if f is the Heaviside function, $f' = \delta$ and $[f'] = 0$.)

We now evaluate the distributional derivatives of such a function f. A couple of points, a and b, are fixed as in Fig. C.2. We try to integrate $f \cdot \varphi'$ ($\in \mathfrak{D}$) on the interval $[a, b]$.

$$\begin{aligned}\int_a^{x_\nu-\varepsilon} f \cdot \varphi' dx &= f \cdot \varphi \Big|_a^{x_\nu-\varepsilon} - \int_a^{x_\nu-\varepsilon} f' \cdot \varphi dx \\ &= f(x_\nu - \varepsilon) \cdot \varphi(x_\nu - \varepsilon) - f(a) \cdot \varphi(a) - \int_a^{x_\nu-\varepsilon} f' \cdot \varphi dx,\end{aligned}$$

where f' in the integral is in the usual sense. Letting $\varepsilon \to 0$, we obtain

$$\int_a^{x_\nu} f \cdot \varphi' dx = f(x_\nu - 0)\varphi(x_\nu) - f(a)\varphi(a) - \int_a^{x_\nu} f' \cdot \varphi dx.$$

Similarly, we have

$$\int_{x_\nu}^b f \cdot \varphi' dx = f(b) \cdot \varphi(b) - f(x_\nu + 0) \cdot \varphi(x_\nu) - \int_{x_\nu}^b f' \cdot \varphi dx.$$

Hence it follows that

$$\begin{aligned}\int_a^b f \cdot \varphi' dx &= f(b) \cdot \varphi(b) - f(a) \cdot \varphi(a) \\ &\quad - [f(x_\nu + 0) - f(x_\nu - 0)]\varphi(x_\nu) - \int_a^b f' \cdot \varphi dx.\end{aligned}$$

If we divide $\mathbb{R}$ into intervals like $[a, b]$, integrate $f \cdot \varphi'$ on each of these small intervals and sum up all, we obtain

$$\int_{-\infty}^{+\infty} f \cdot \varphi' dx = -\sum_\nu \sigma_\nu^0 \varphi(x_\nu) - \int_{-\infty}^{+\infty} f' \cdot \varphi dx,$$

taking account of the support of φ is bounded. Consequently,

$$f'(\varphi) = -f(\varphi') = \sum_\nu \sigma_\nu^0 \varphi(x_\nu) + \int_{-\infty}^{+\infty} f' \cdot \varphi dx = \sum_\nu \sigma_\nu^0 \delta_{(x_\nu)}(\varphi) + [f'](\varphi).$$

Thus we conclude that the distributional derivative of f is given in the form

$$f' = [f'] + \sum_\nu \sigma_\nu^0 \delta_{(x_\nu)}.$$

We should note that the discontinuities of f appear as masses at x_ν's in the distributional derivative.

More generally, the p-th distributional derivative $f^{(p)}$ is given by

$$f^{(p)} = [f^{(p)}] + \sum_\nu \sigma_\nu^{(p-1)} \delta_{(x_\nu)} + \sum_\nu \sigma_\nu^{(p-2)} \delta_{(x_\nu)}^{(1)} + \cdots + \sum_\nu \sigma_\nu^0 \delta_{(x_\nu)}^{(p-1)}.$$

C.5 Topologies on the Space $\mathfrak{D}(\Omega)'$ of Distributions

Simply applying the general theory discussed in Appendix B.2 to $\mathfrak{D}(\Omega)'$, we obtain several basic results.

Theorem C.8 *Endow the space $\mathfrak{D}'$ of distributions with the strong topology.*

(i) *$\mathfrak{D}'$ is complete.*
(ii) *$\mathfrak{D}'$ is a Montel space.*
(iii) *$\mathfrak{D}$ and $\mathfrak{D}'$ are reflexive, and they are the dual spaces of each other.*

Lemma C.4 *If a sequence $\{T_n\}$ of distributions simply converges to a distribution T; i.e.*

$$\lim_{n\to\infty} T_n(\varphi) = T(\varphi) \quad \textit{for any} \quad \varphi \in \mathfrak{D},$$

then $\{D^p T_n\}$ also simply converges to $D^p T$ for any differential operator D^p.

Proof It is immediate, since

$$D^p T_n(\varphi) = (-1)^{|p|} T_n(D^p \varphi) \to (-1)^{|p|} T(D^p \varphi) = D^p T(\varphi)$$

for any $\varphi \in \mathfrak{D}$. □

The next theorem is easily derived from the Banach–Steinhaus theorem (Theorem B.12, p. 371) and Lemma C.4.

Theorem C.9 *Suppose that a sequence $\{T_n\}$ of distributions has a limit*

$$T(\varphi) \equiv \lim_{n\to\infty} T_n(\varphi)$$

for each $\varphi \in \mathfrak{D}$. Then the following results hold good:

(i) *T is a distribution and $\{T_n\}$ converges to T strongly.*
(ii) *For any differential operator D^p,*

$$D^p T = \lim_{n\to\infty} D^p T_n.$$

Corollary C.2 (series of distributions) *Suppose that $T_n (n = 1, 2, \cdots)$ are distributions and the limit*

$$T(\varphi) \equiv \sum_{n=1}^{\infty} T_n(\varphi)$$

exists for each $\varphi \in \mathfrak{D}$.

(i) $T = \sum_n T_n$ *is a distribution.*

(ii) *For any differential operator* D^p,

$$D^p T = \sum_{n=1}^{\infty} D^p T_n.$$

Definition C.2 Let T be a distribution on $\Omega \subset \mathbb{R}$ and U an open subset of Ω. T is said to be **0 in** U (or **vanish in** U) if $T(\varphi) = 0$ for all $\varphi \in \mathfrak{D}(\Omega)$ with supp$\varphi \subset U$.

Theorem C.10 *Let* $\{U_i\}_{i \in I}$ *be a family of open sets in* Ω. *If a distribution* T *vanishes in all* U_i*'s, then* T *vanishes in* $U = \bigcup_{i \in I} U_i$.

Proof Let $\varphi \in \mathfrak{D}$ with supp$\varphi \subset U$. If we add $\Omega \setminus$ suppφ to $\{U_i\}_{i \in I}$, we obtain an open covering of Ω. We denote it by $\{V_j\}_{j \in J}$. Let $\{\beta_j\}_{j \in J}$ be a smooth partition of unity subordinate to $\{V_j\}_{j \in J}$; i.e.

1° $\beta_j : \Omega \to [0, 1]$,
2° supp$\beta_j \subset V_j$,
3° for any compact set $K \subset \Omega$, all but finite β_j, are 0 on K,
4° $\sum_{j \in J} \beta_j(x) = 1$ for all $x \in \Omega$.

Then

$$\varphi = \sum_{j \in J} \beta_j \varphi,$$

and this is actually a finite sum, since suppφ is compact and the property 3° of $\{\beta_j\}$. Hence

$$T(\varphi) = \sum_{j \in J} T(\beta_j \varphi).$$

If suppβ_j is contained in some $U_i (i \in I)$, $T(\beta_j \varphi) = 0$ by assumption. In the case of supp$\beta_j \subset \Omega \setminus$ suppφ, it holds good that $\beta_j \varphi = 0$, and so clearly $T(\beta_j \varphi) = 0$. Therefore $T(\varphi) = 0$. □

According to Theorem C.10, a distribution T is 0 in

$$U \equiv \bigcup \{V \subset \Omega \mid T \text{ vanishes in } V\}.$$

U is the largest open set in which T is 0. $\Omega \setminus U$ is called the support of T.

Definition C.3 The **support** of a distribution $T \in \mathfrak{D}(\Omega)'$ is the smallest closed set F such that T vanishes in $\Omega \setminus F$.

Example C.8 If T is a continuous function,[14] the support of T as a distribution and that as a function coincide.

Example C.9 The support of the Dirac distribution $\delta_{(a)}$ is the singleton $\{a\}$.

We insert here some basic results concerning boundedness in topological vector spaces:

1° Let $\mathfrak{X}$ and $\mathfrak{Y}$ be topological vector spaces and $T : \mathfrak{X} \to \mathfrak{Y}$ a continuous linear operator. Then the image $T(B)$ of a bounded set B in $\mathfrak{X}$ is bounded in $\mathfrak{Y}$.

2° Let $\mathfrak{X}$ be a locally convex topological vector space. The following two statements are equivalent:

a. Any balanced convex set which absorbs every bounded set is a neighborhood of 0 in $\mathfrak{X}$.
b. Any seminorm which is bounded on every bounded set in $\mathfrak{X}$ is continuous.

A locally convex topological vector space is called **bornologic** if it satisfies a ($\Leftrightarrow$ b).

3° A metrizable locally convex topological vector space is bornologic.

Suppose that $A \subset \mathfrak{X}$ is a balanced, convex set which absorbs every bounded set. Since $\mathfrak{X}$ is metrizable, 0 has a countable neighborhood base $\{U_n\}$. We may assume, without loss of generality, that $U_n \subset U_{n-1}$. If A is not a neighborhood of 0, there exists a sequence $\{x_n\}$ in $\mathfrak{X}$ such that $x_n \in (1/n)U_n$, $x_n \notin A$. $\{nx_n\}$ is bounded, since $nx_n \to 0$. Hence

$$nx_n \in \alpha A \text{ for all } n$$

for some $\alpha \geqq 0$. Since $x_n \in (\alpha/n)A$, we have $x_n \in A$ for large n, a contradiction. 4° immediately follows from 3°.

4° $\mathfrak{D}(\Omega)$ and $\mathfrak{E}(\Omega)$ are bornologic.

5° Let $\mathfrak{X}$ and $\mathfrak{Y}$ be locally convex topological vector spaces. $\mathfrak{X}$ is assumed, in addition, to be bornologic. A linear operator $T : \mathfrak{X} \to \mathfrak{Y}$ is continuous, if the image $T(B)$ of any bounded set B in $\mathfrak{X}$ is bounded in $\mathfrak{Y}$.

The next result immediately follows.

Theorem C.11 *A linear functional on $\mathfrak{D}(\Omega)$ (resp. $\mathfrak{E}(\Omega)$) is continuous if and only if the image $T(B)$ of any bounded set B in $\mathfrak{D}(\Omega)$ (resp. $\mathfrak{E}(\Omega)$) is bounded in $\mathbb{C}$.*

Corollary C.3 *The following two statements are equivalent for a linear functional T on $\mathfrak{D}(\Omega)$:*

(i) *T is a distribution.*
(ii) *For any compact set K in Ω, there exist some $C > 0$ and $k \in \mathbb{N}$ such that*

$$|T(\varphi)| \leqq C \sup_{\substack{|p| \leqq k \\ x \in K}} |D^p \varphi(x)| \quad \textit{for all} \quad \varphi \in \mathfrak{D}_K(\Omega).$$

[14] This means that T is a distribution defined by a continuous function, say f. In this case $\mathrm{supp} T = \mathrm{supp} f$.

Proof (i)⇒(ii): If T is a distribution, T is continuous on $\mathfrak{D}_K$ for each compact set K in Ω. (ii) immediately follows.

(ii)⇒(i): If we assume (ii) conversely, T maps any bounded set of $\mathfrak{D}(\Omega)$ into a bounded set. Hence by Theorem C.11, T is continuous. □

Lemma C.5[15]

(i) *The topology of* $\mathfrak{D}(\Omega)$ *is stronger than the relative topology induced by* $\mathfrak{E}(\Omega)$*,*
(ii) $\mathfrak{D}(\Omega)$ *is dense in* $\mathfrak{E}(\Omega)$*.*

Based upon these preparations, we can prove the following important result.
A distribution $T \in \mathfrak{D}(\Omega)'$ **can be extended uniquely to an element of** $\mathfrak{E}(\Omega)'$.

Proof We denote by K the support of T; i.e. $K = \mathrm{supp}T$ (compact). Let ψ be an element of $\mathfrak{D}(\Omega)$, which is identically 1 in a neighborhood of K.[16] (Such a ψ certainly exists.) Define $\tilde{T} : \mathfrak{E}(\Omega) \to \mathbb{C}$ by

[15] We can prove Lemma C.5 in the following way:

(i) Choose a sequence $\{K_n\}$ of compact sets such that $K_n \subset \mathrm{int.}K_{n+1}$ and $\Omega = \bigcup_{n=1}^{\infty} K_n$. Consider a basic neighborhood of 0 in $\mathfrak{E}(\Omega)$, say

$$V_{m,\varepsilon,K} \equiv \{\varphi \in \mathfrak{E} \mid \sup_{p \leqq m, x \in K} |D^p \varphi(x)| < \varepsilon\} \quad (K \text{ is compact}).$$

For $K_n \supset K$, we have

$$V_{m,\varepsilon,K} \bigcap \mathfrak{D}_{K_n}(\Omega) = \{\varphi \in \mathfrak{D}_{K_n}(\Omega) | \sup_{p \leqq m, x \in K} |D^p \varphi(x)| < \varepsilon\}$$
$$\supset \{\varphi \in \mathfrak{D}_{K_n}(\Omega) | \sup_{p \leqq m, x \in K_n} |D^p \varphi(x)| < \varepsilon\}.$$

The last term is a neighborhood of 0 in $\mathfrak{D}_{K_n}(\Omega)$. Hence, by definition of the topology of $\mathfrak{D}(\Omega)$, $V_{m,\varepsilon,K} \bigcap \mathfrak{D}_{K_n}(\Omega)$ is a neighborhood of 0 in $\mathfrak{D}(\Omega)$.

(ii) Let $\varphi_n \in \mathfrak{D}(\Omega)$ satisfy $\varphi_n(K_n) = \{1\}$ and $\varphi_n(\Omega \setminus K_{n+1}) = \{0\}$. For $\varphi \in \mathfrak{E}(\Omega)$, we have $\varphi \cdot \varphi_n \in \mathfrak{D}(\Omega)$. We have to show that $\varphi \cdot \varphi_n \to \varphi$ in $\mathfrak{E}(\Omega)$.

[16] We write $K = \mathrm{supp}T$ (compact). For any $\varepsilon > 0$, there exist finite points $x_1, x_2, \cdots, x_n$ such that

$$K \subset \bigcup_{j=1}^{n} B_\varepsilon(x_j).$$

Let $\{\beta_j\}$ be a smooth partition of unity subordinate to $\{B_\varepsilon(x_j) | j = 1, 2, \cdots n\}$. Let H be a compact neighborhood of K such that $H \subset \bigcup_{j=1}^{n} B_\varepsilon(x_j)$. Define a function $\psi : \Omega \to \mathbb{C}$ by

$$\psi(x) = \sum_{(\mathrm{supp}\beta_j) \cap H \neq \emptyset} \beta_j(x).$$

Then $\psi \in \mathfrak{D}(\Omega)$ and ψ is identically 1 in a neighborhood of K.

$$\tilde{T}(\eta) = T(\psi\eta); \quad \eta \in \mathfrak{E}(\Omega).$$

$\tilde{T}$ is well-defined in the sense that it is determined independently of the choice of ψ.[17]

$\tilde{T}$ is clearly linear. The continuity of $\tilde{T}$ is proved as follows. For any bounded set B in $\mathfrak{E}(\Omega)$, the set $\{\psi\eta|\eta \in B\}$ is bounded in $\mathfrak{D}(\Omega)$. Hence

$$\{\tilde{T}(\eta)|\eta \in B\} = \{T(\psi\eta)|\eta \in B\}$$

is bounded by Theorem C.11. Hence, again by Theorem C.11. $\tilde{T}$ is continuous.

$\tilde{T}$ is an extension of T. In fact, if $\psi \in \mathfrak{D}(\Omega)$, then $\varphi(1-\psi) \in \mathfrak{D}(\Omega)$ and

$$\mathrm{supp}\varphi(1-\psi) \bigcap K = \emptyset.$$

Hence $T(\varphi(1-\psi)) = 0$, which implies

$$\tilde{T}\varphi = T(\psi\varphi) + T(\varphi(1-\psi)) = T\varphi.$$

Thus $\tilde{T}$ is an extension of T.

$\mathfrak{D}(\Omega)$ is dense in $\mathfrak{E}(\Omega)$ (Fréchet space) and $\tilde{T}|_{\mathfrak{D}(\Omega)} = T$ is continuous (and so uniformly continuous). Hence, by the principle of extension by continuity, an extension of T to $\mathfrak{E}(\Omega)$ is uniquely defined. □

The converse assertion is as follows.

An element of $\mathfrak{E}(\Omega)'$ restricted to $\mathfrak{D}(\Omega)$ is a distribution with compact support.

Proof Let $\tilde{T} \in \mathfrak{E}(\Omega)'$. The restriction of $\tilde{T}$ to $\mathfrak{D}(\Omega)$ is denoted by T; i.e.

$$T = \tilde{T}|_{\mathfrak{D}(\Omega)}.$$

It is obvious that T is a linear functional on $\mathfrak{D}(\Omega)$.

Let $\{\varphi_\alpha\}$ be a net in $\mathfrak{D}(\Omega)$ which converges to 0. Then there exists some compact set $K \subset \Omega$ such that

$$\varphi_\alpha \to 0 \quad \text{in } \mathfrak{D}_K(\Omega),$$

[17] If $\theta \in \mathfrak{D}(\Omega)$ is also identically 1 in a neighborhood of K, then $(\psi - \theta)\eta$ is an element of $\mathfrak{D}(\Omega)$, which is identically 0 in a neighborhood of K. Hence it follows that $T(\psi\eta) - T(\theta\eta) = T((\psi-\theta)\eta) = 0$.

and hence

$$\varphi_\alpha \to 0 \quad \text{in } \mathfrak{E}(\Omega).$$

Since $\tilde{T}$ is an element of $\mathfrak{E}(\Omega)'$, we have

$$T(\varphi_\alpha) = \tilde{T}(\varphi_\alpha) \to 0.$$

Hence T is continuous on $\mathfrak{D}(\Omega)$; i.e. $T \in \mathfrak{D}(\Omega)'$.

Finally, we have to show that supp$T$ is compact. If not, there exists a sequence $\{\varphi_n\}$ in $\mathfrak{D}(\Omega)$ such that

$$\begin{cases} \text{supp}\varphi_n \subset \Omega \setminus \overline{B_n(0)}, \\ \quad (\text{i.e.} \quad \varphi_n(x) = 0 \quad \text{if} \quad \|x\| \leqq n), \\ T(\varphi_n) = 1 \end{cases}$$

for each $n \in \mathbb{N}$. Since suppφ_n goes away farther,

$$\varphi_n \to 0 \quad \text{in } \mathfrak{E}(\Omega).$$

Hence

$$T(\varphi_n) = \tilde{T}(\varphi_n) \to 0.$$

However, it contradicts $T(\varphi_n) = 1$ (for all n). supp$T$ must be compact. □

Combining the two propositions above, we can observe that a distribution with compact support corresponds to an element of $\mathfrak{E}(\Omega)'$ one-to-one. If we identify corresponding T and $\tilde{T}$, we have the following result.

Theorem C.12 *The space of distributions with compact support coincides with* $\mathfrak{E}(\Omega)'$.

Remark C.3 A linear functional $\tilde{T}$ on $\mathfrak{E}(\Omega)$ is continuous if and only if there exist a compact set H in Ω, $C > 0$ and $K \in \mathbb{N}$ such that

$$|\tilde{T}(\eta)| \leqq C \cdot \sup_{p \leqq k, x \in H} |D^p \eta(x)| \quad \text{for all} \quad \eta \in \mathfrak{E}(\Omega).$$

Proof We have only to prove necessity, since the converse is easy. $T = \tilde{T}|_{\mathfrak{D}(\Omega)}$ is a distribution if and only if there exist, for any compact set $K \subset \Omega$, some $C' > 0$ and $k' \in \mathbb{N}$ such that

$$|T(\varphi)| \leqq C' \sup_{p \leqq k', x \in K} |D^p \varphi(x)| \quad \text{for all} \quad \varphi \in \mathfrak{D}_K(\Omega).$$

It holds good, in particular, for $H = \text{supp}\tilde{T}$.

Let ψ be an element of $\mathfrak{D}(\Omega)$ which is identically 1 in a neighborhood of H. If $H = \text{supp}\varphi$, $\psi\eta \in \mathfrak{D}_H(\Omega)$ for any $\eta \in \mathfrak{D}(\Omega)$. Hence

$$|\tilde{T}(\eta)| = |T(\psi\eta)| \leqq C' \sup_{p \leqq k', x \in H} |D^p(\psi\eta)(x)| \leqq C'\{C'' \sup_{p \leqq k', x \in H} |D^p\eta(x)|\}.$$

By setting $C = C' \cdot C''$ and $k = k'$, we obtain the desired result. □

References

1. Bourbaki, N.: Eléments de mathématique; Espaces vectoriels topologiques. Hermann, Paris, Chaps.1–2 (seconde édition) (1965), Chaps. 3–5, (1964). (English edn.) Elements of Mathematics; Topological Vector Spaces. Springer, Berlin/Heidelberg/New York (1987)
2. Choquet, G.: Lectures on Analysis, vol. I. Benjamin, London (1969)
3. Hölmander, L.: The Analysis of Linear Partial Differential Operators, I, II. Springer, Berlin/Heidelberg (1983)
4. Kolmogorov, A.N., Fomin, S.V.: (English edn.) Introductory Real analysis, Prentice-Hall, New York (1970) (Japanese edn., translated from 4th edn.) Kansu-kaiseki no Kiso, 4th edn. Iwanami Shoten, Tokyo (1979) (Originally published in Russian) I use the Japanese edition for quotations.
5. Maruyama, T.: Sekibun to Kansu-kaiseki (Integration and Functional Analysis). Springer, Tokyo (2006) (Originally published in Japanese)
6. Schwartz, L.: Méthods mathématique pour les sciences physiques. Hermann, Paris (1965)
7. Schwartz, L.: Théorìe des distributions. Nouvelle édition, entièrement corrigée, refondue et augmentée, Hermann, Paris (1973)
8. Yosida, K.: Functional Analysis, 3rd edn. Springer, New York (1971)

Addendum

My basic policy in this book is "to restrict myself to a rather narrow field of applications familiar to many economists, and to give full details of mathematical materials instead", as I wrote in the preface. However, I would like to add here a brief memorandom of a topic from recent developments in economic theory.

As is well-known, it is Solow [5] who pointed out the significance of the "vintage" structure of a stock of capital. He wrote that any technological innovations realize their effect on output "only to the extent that they are carried into practice either by net capital formation or by the replacement of old-fashioned equipment by the latest models".

Solow's vintage models of the embodied technological innovation exerted a remarkable influence and aroused much controversy in the theory of capital. We can now observe a renaissance of Solow's model in a new framework.

In the so-called "neo-classical" macro-economic theory, the aggregate output (say, gross domestic product, GDP) is produced by an aggregate stock of capital and labor services under certain technological conditions represented by a production function. The gross investment must be equal to the saving at each time. The surviving fraction of a unit of capital depends upon its "vintage", that is, the date when the machine was produced.

Goldman–Kato–Mui [4] presented and analyzed an integral equation

$$k(t) = \int_o^t g(k(t-\theta))\varphi(\theta)d\theta + h(t), \tag{†}$$

which describes a dynamic behavior of the capital stock $k(t)$ per laborer. Here g, φ and h are given functions. The equation (†) may be regarded as a nonlinear version of Volterra's integral equation of the second kind, involving a "nonlinear convolution". Comparing it with the integral equation of the form

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8

$$k(t) = \int_o^t k(t-\theta)\varphi(\theta)d\theta + h(t), \tag{‡}$$

which is called Feller's renewal equation (cf. Feller [3]), we may view the equation (†) as a nonlinear version of it. A similar but distinct equation is discussed on pp. 75–78 of this book.

In the same year, Benhabib–Rustichini [1] studied an optimal growth problem subject to the capital accumulation described by a nonlinear integral equation in the same category.

From a mathematical viewpoint, one of the deepest analyses was provided by Diekmann and Gils [2]. They considered an integral equation of the form

$$x(t) = \int_0^t B(\tau)g(x(t-\tau))d\tau + f(t), \quad t \geqq 0, \tag{$*$}$$

where B is a $(n \times n)$-matrix-valued $\mathfrak{L}^2$-function and $g : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth function. Both B and g are given. By inserting real parameters, they also examined the existence of periodic solutions having recourse to Hopf's bifurcation theorem (cf. Chap. 11).

Integral equations of this type sometimes appear in epidemiology. (See Thieme [6].)

References

1. Benhabib, J., Rustichini, A.: Vintage capital, investment, and growth. J. Econ. Theory **55**, 323–339 (1991)
2. Diekmann, O., van Gils, S.A.: Invariant manifolds for Volterra integral equations of convolution type. J. Differential Equations **54**, 139–180 (1984)
3. Feller, W.: On the integral equation of renewal theory. Ann. Math. Stat. **12**, 243–267 (1941)
4. Goldman, S.M., Kato, T., Mui, V.-L.: Economic growth and general depreciation. J. Development Economics **34**, 397–400 (1991)
5. Solow, R.M.: Investment and technical progress. In: Arrows, K.J., Karlin, S., Suppes, P. (eds.) Mathematical Methods in the Social Sciences 1959. Stanford University Press, Stanford (1960)
6. Thieme, H.R.: Renewal theorems for some mathematical models in epidemiology. J. Integral Equations **8**, 185–216 (1985)

Name Index

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8

Subject Index

T. Maruyama, *Fourier Analysis of Economic Phenomena*, Monographs in Mathematical Economics 2, https://doi.org/10.1007/978-981-13-2730-8