Tepper L. Gill · Woodford W. Zachary

Functional Analysis and the Feynman Operator Calculus

Springer

Functional Analysis and the Feynman Operator Calculus

Tepper L. Gill • Woodford W. Zachary

Functional Analysis and the Feynman Operator Calculus

Springer

Tepper L. Gill
Departments of Electrical
and Computer Engineering
Howard University
Washington, DC, USA

Woodford W. Zachary
Departments of Electrical
and Computer Engineering
Howard University
Washington, DC, USA

ISBN 978-3-319-27593-2 ISBN 978-3-319-27595-6 (eBook)
DOI 10.1007/978-3-319-27595-6

Library of Congress Control Number: 2015957963

Mathematics Subject Classification (2010): 28A35, 35K15, 35L15, 47D06, 46B99, 46F25, 43A25

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

Two approaches to the mathematical foundations of relativistic quantum theory began in the USA. Both evolved from the application of quantum field methods to electron theory in the late 1940s by Feynman, Schwinger, and Tomonaga (see [SC1]).

The first program is well known and was begun in the early 1950s by Professor A.S. Wightman of Princeton University (1922–2013). Following a tradition inspired by Hilbert, the program was called axiomatic field theory. It sought to provide rigorous justification for the complicated and difficult method of renormalization successfully employed by the physics community (see [SW] and [GJ]). Professor Wightman is considered the founding father of modern mathematical physics, but he also strongly influenced a number of other areas in mathematics.

In 1982, Sokal notice some difficulties with the constructive approach to field theory (the concrete version of axiomatic field theory) and conjectured that this approach may not work as expected in four space-time dimensions (see [SO]). His conjecture was later verified by Aizenman and Graham [AG] at Princeton and Fröhlich [FO] at ETS, Zurich. These results have had a damping effect on research in this direction.

In response to the work of Aizenman, Graham, and Fröhlich, a second, less well-known program was initiated by the present authors at

Howard University in 1986. We sought to understand the issues affecting relativistic quantum theory based on a series of problems suggested by Dirac, Dyson, Feynman, Schwinger, and other major architects of quantum field theory. This book is an outgrowth of our investigations into the mathematical issues facing any attempt to develop a reasonable relativistic quantum theory. Our investigations into the physical foundations are the subject of a future project. (However, those with interest in this subject are directed to [GZ5] and [GMK, see Chap. 5] for some partial results in this direction.)

In 1951, Richard Feynman published what became known as the Feynman operator calculus. It served as the basis for his formulation of quantum electrodynamics, for which he shared the Nobel Prize in Physics with Schwinger and Tomonaga. Freeman Dyson introduced this work to the mathematics and physics communities, providing Feynman's theory both the physical and mathematical legitimacy. Dyson also showed that the two competing formulations of quantum electrodynamics were based on different representations of Heisenberg's S-matrix. Using his understanding of both theories, Dyson made fundamental improvements and simplifications. (It is suggested by Schweber [SC1] that Dyson's contribution is also worthy of the Nobel Prize.)

Feynman's basic idea was to first lay out space-time as one would a photographic film. He then imagined the evolution of a physical system appearing as a three-dimensional motion picture on this film; one seeing more and more of the future as more and more of the film comes into view (see [F]). This gives time its natural role in ordering the flow of events as it does in our conscious view of reality. Feynman suggested that time should serve this role in the manipulation of operator-valued variables in quantum field theory, so that operators acting at different times actually commute. He demonstrated that this approach made it possible to write down and compute highly complicated expressions in a fast and effective manner. In one case, he was able to perform a calculation in one night that had previously taken over 6 months (see [SC1]).

Feynman's faith in his operator calculus is expressed at the end of his book on path integrals (with Hibbs [FH]); he states: "Nevertheless, many of the results and formulations of path integrals can be re-expressed by another mathematical system, a kind of ordered operator calculus. In this form many of the results of the preceding chapters find an analogous but more general representation ... involving noncommuting variables." Feynman is referring to [F], quoted above.

To our knowledge, Fujiwara [FW] is the only physicist other then Dyson who takes Feynman's operator calculus seriously in the early literature (1952). Fujiwara agreed with the ideas and results of Feynman with respect to the operator calculus, but was critical of what he called notational ambiguities, and introduced a slightly different approach. "What is wanted, and what I have striven after, is a logical well ordering of the main ideas concerning the operator calculus. The present study is entirely free from ambiguities in Feynman's notation, which might obscure the fundamental concepts of the operator calculus and hamper the rigorous organization of the disentanglement technique." Fujiwara's main idea was that the Feynman program should be implemented using a sheet of unit operators at every point except at time t, where the true operator should be placed. He called the exponential of such an operator an expansional to distinguish it from the normal exponential so that, loosely speaking, disentanglement becomes the process of going from an expansional to an exponential. (Araki [AK] formally investigated Fujiwara's suggestion.) As will be seen, Fujiwara's fundamental insight is the centerfold of our approach to the problem.

In our approach, the motivating research philosophy was that, the correct mathematical foundation for the Feynman operator calculus should in the least:

(1) Provide a transparent generalization and/or extension of current mathematical theories without sacrificing the physically intuitive and computationally useful methods of Feynman

(2) Provide a rigorous foundation for the general theory of path integrals and its relationship to semigroups of operators and partial differential equations

(3) Provide a direct approach to the mathematical study of time-dependent evolution equations in both the finite and infinite-dimensional setting

(4) Provide a better understanding of some of the major mathematical and physical problems affecting the foundations of relativistic quantum theory

This book is devoted to the mathematical development of the first three items. We also briefly discuss a few interesting mathematical points concerning item (4). (However, as noted earlier, a full discussion of (4) is delayed to another venue.)

While no knowledge of quantum electrodynamics is required to understand the material in this book, at a few junctures, some physical intuition and knowledge of elementary quantum mechanics would be helpful. We assume a mathematical background equivalent to that of a third year graduate student, which includes the standard courses in advanced analysis, along with additional preparation in functional analysis and partial differential equations. A course (or self-study) based on the first volume of Reed and Simon [RS1, see Chap. 1] offers a real advantage. An introduction to probability theory or undergraduate background in physics or chemistry would also be valuable. In practice, unless one has acquired a reasonable amount of mathematical maturity, some of the material could be a little heavy going. (Mathematical maturity means losing the fear of learning topics that are new and/or at first appear difficult.) However, in order to make the transition as transparent as possible, for advanced topics we have provided additional motivation and detail in many of the proofs.

We have three objectives. The first two, the Feynman operator calculus and path integrals and their relationship to the foundations of relativistic quantum theory, occupy a major portion of the book. Our third objective, infinite-dimensional analysis, provides the purely mathematical background for the first two. We have also included some closely related material that has independent interest. In these cases, we also indicate and/or direct the interested reader to the Appendix.

The book is organized in a progressive fashion with each chapter building upon the previous ones. Almost all of the material in Chaps. 2, 3, and 6–8 has not previously appeared in book form. In addition, Chap. 5 is developed using a completely new approach to operator theory on Banach spaces, which makes it almost as easy as the Hilbert space theory.

Chapter 1 is given in two parts. Part I introduces some of the background material, which is useful for review and reference. Basic results and definitions from analysis, functional analysis, and Banach space theory are included and should at least receive a glance before proceeding.

Part II is devoted to the presentation of a few advanced topics which are not normally discussed in the first 2 years of a standard graduate program, but are required for later chapters in the book. The reader should at least review this part to identify unfamiliar topics, so one may return when needed.

Chapter 2 is devoted to the foundations for analysis on spaces with an infinite number of variables. Infinite dimensional analysis is intimately related to the Feynman operator calculus and path integrals and cannot be divorced from any complete study of the subject. Faced directly, the first problem encountered is the need for a reasonable version of Lebesgue measure for infinite-dimensional spaces. However, research into the general problem of measure on infinite-dimensional vector spaces has a long and varied past, with participants living in a number of different countries, during times when scientific communication was constrained by war, isolation, and/or national competition. These conditions have allowed quite a bit of misinformation and folklore to grow up around the subject, so that even some experts have a limited view of the subject. Yamasaki was the first to construct a σ-finite version of Lebesgue measure on $\mathbb{R}^\infty$ in 1980 (see [YA1]), and uniqueness has only been proved recently (2007) by Kirtadze and Pantsulaia [KP2, see Chap. 6]. However, due to the nature of their approach, the work of Yamasaki and Kirtadze and Pantsulaia is only known to specialists in the field.

In Sect. 2.1 the Yamasaki version of Lebesgue measure for $\mathbb{R}^\infty$ is constructed in a manner which is very close to the way one learns measure theory in the standard analysis course. In Sect. 2.2, a version of Lebesgue measure is constructed for every Banach space with a Schauder basis (S-basis). In addition, a general approach to probability measures on Banach spaces is developed. The main result in this direction is that every probability measure ν on $\mathfrak{B}[\mathbb{R}]$ with a density induces a corresponding related family of probability measures $\{\nu_{\mathcal{B}}^n\}$ on every Banach space $\mathcal{B}$, with an S-basis, which is absolutely continuous with respect to Lebesgue measure. Under natural conditions, the family converges to a unique measure $\nu_{\mathcal{B}}$. As particular examples, we prove the existence of universal versions of both the Gaussian and Cauchy measures. Section 2.3 is devoted to measurable functions, the Lebesgue integral, and the standard spaces of functions, continuous, L^p, etc. Section 2.4 studies distributions on uniformly convex Banach spaces. Section 2.5 introduces Schwartz space and the Fourier transform on uniformly convex Banach spaces. This allows us to extend the Pontryagin Duality Theorem to uniformly convex Banach spaces in Sect. 2.6. In addition, we provide a direct solution to the diffusion equation on Hilbert space as an interesting application of our universal representation for Gaussian measure. Sections 2.4–2.6 are not required for a basic understanding of the Feynman operator calculus and the theory of path integrals. However, there are natural connections between these subjects. Thus, those with broader concerns and/or interests in other applications will find the study both rewarding and fruitful.

Chapter 3 introduces the Henstock–Kurzweil integral. This is the easiest to learn and best known of those integrals that integrate nonabsolutely integrable functions and extend the Lebesgue integral. Section 3.1 provides a fairly detailed account of the HK-integral and its properties in the one-dimensional case and a brief discussion of the n-dimensional case. Section 3.2 discusses a new class of Banach spaces (KS^p spaces) that are for nonabsolutely integrable functions as the L^p spaces are for Lebesgue integrable functions. These spaces contain the L^p spaces as continuous dense and compact embeddings. Section 3.3 covers some additional classes of Banach spaces associated with nonabsolutely integrable functions which may have future interest. First, we define an important class of spaces $SD^p[\mathbb{R}^n]$, $1 \leqslant p \leqslant \infty$. These spaces contain the test functions of Schwartz [SCH] $\mathcal{D}[\mathbb{R}^n]$, as a dense continuous embedding. In addition, they have the remarkable property that for any multi-index α, $\|D^\alpha \mathbf{u}\|_{SD} = \|\mathbf{u}\|_{SD}$, where D is the distributional derivative. We call them the Jones strong distribution Banach spaces. As an application, we obtain a nice a priori estimate for the nonlinear term of the classical Navier–Stokes initial-value problem. In Sect. 3.4, we introduce a class of spaces in honor of our deceased colleague Woodford W. Zachary. These spaces all extend the class of functions of bounded mean oscillation to include the HK-integrable functions. (Sections 3.3 and 3.4 are not required for the rest of the book.)

Chapter 4 is devoted to a fairly complete account of analysis and operator theory on Hilbert space. The first part introduces the theory of integration of operator-valued functions, and the second part gives a first course in Hilbert space operator theory. The presentation is standard, but an interesting extension of spectral theory is introduced based on the polar decomposition property of closed densely defined linear operators.

Chapter 5 is devoted to operator theory on Banach spaces, with major emphasis on semigroups of operators. Our approach is novel, as it uses the theory of Chap. 4 in a unique manner, showing that the theory on Banach spaces is much closer to the Hilbert space theory then previously known. In the first section we show that, for uniformly convex Banach spaces with a Schauder basis, it is possible to define the adjoint for every closed densely defined linear operator on the space. (This result is extended to a larger class of spaces and operators in the Appendix (Sect. 5.3).) We give a number of examples so that one can see what the adjoint looks like in concrete cases. In the second section, the adjoint is used to give a parallel treatment of semigroups of operators, which is very close to the Hilbert space theory. In the Appendix (Sect. 5.3), in addition to an extension of the adjoint, we extend the spectral theory and provide a complete version of the Schatten classes of compact operators for uniformly convex Banach spaces with a Schauder basis.

Chapter 6 develops infinite tensor product theory for Hilbert and Banach spaces. The Banach space theory is a new subject, which offers a number of advantages for analysis. Our approach generalizes von Neumann's infinite tensor product Hilbert space theory, so we call them spaces of type v. We use infinite tensor products of Hilbert and Banach spaces to construct the mathematical representation for Feynman's physical film. We also introduce the notion of an exchange operator, which will prove important in Chaps. 7 and 8. (Infinite tensor products of Banach spaces are also natural for the constructive study of analysis in infinite-many variables. We have included a few applications and possibilities in the Appendix (Sect. 6.7).)

In Chap. 7, we develop the Feynman operator theory on Hilbert space, as a compromise for the two classes of potential users. Following Fujiwara's idea, we first define what we mean by time-ordering, prove our fundamental theorem on the existence of time-ordered integrals, and extend the basic semigroup theory to the time-ordered setting. This provides, among other results, a time-ordered version of the Hille–Yosida Theorem. We construct time-ordered evolution operators and prove that they have all the expected properties. We define what is meant by the phrase "asymptotic in the sense of Poincaré" for operators. We then develop a general perturbation theory and use it to prove a generalized version of Dyson's second conjecture for quantum electrodynamics, namely, that all theories generated by semigroups are asymptotic in the operator-valued sense of Poincaré. (Dyson conjectured this result for unitary groups.)

In 1955, Hagg [HA, see Chap. 7] investigated the general conditions which a relativistic quantum theory of interacting particles must satisfy in order to be made mathematically rigorous. One of his major conclusions was that the canonical commutation relations need not have unique solutions and that the interaction representation in sharp time does not exist. It has now been experimentally confirmed that there is quantum interference in time (see Chap. 7, [HW]). Thus, Hagg's assumption of sharp time is not physically valid. In this section, we modify Dyson's theory to include an interaction representation which allows time interference of wave packets. Finally, we show that the Fujiwara–Feynman approach to disentanglement can be implemented in a direct manner. This approach also provides a nice extension to the Trotter–Kato perturbation theory. In the last section we develop a general approach to the mathematical foundations for Feynman's sum over paths, which is used in quantum theory.

Chapter 8 provides a few applications of the operator calculus. We first develop a general theory for time-dependent parabolic and hyperbolic evolution equations. We demonstrate that the operator calculus allows us to unify methods and weaken domain requirements.

We then turn to the Feynman path integral. At this time, there is an extensive literature on the development and application of path integral methods in all aspects of physics, chemistry, mathematics, and engineering, and it is impossible to provide a reasonable discussion of these efforts. As a substitute, we provide references to some of the important works on this subject and introduce a number of interesting examples which are not covered in the literature. Our focus is on the mathematical foundations. We first demonstrate that the Kuelbs–Steadman space, $KS^2[\mathbb{R}^3]$, allows us to construct the elementary path integral in exactly the manner suggested by Feynman. Thus, our approach does not encumber physical intuition or computational efficiency. We further show that $KS^2[\mathbb{R}^3]$ is sufficient to provide a rigorous foundation for the Feynman formulation of quantum mechanics.

In order to further extend our theory, we introduce some results due to Maslov and Shishmarev on hypoelliptic pseudodifferential operators that allow us to construct a general class of path integrals generated by Hamiltonians, which are not perturbations of Laplacians (see Shishmarev [SH]). We then use the results of Chap. 7 and our sum over path theory to generalized and extend the well-known Feynman–Kac Theorem. Our final result is independent of the space of continuous functions, so that the question of the existence of measures is more of a desire than a requirement. (The strong continuity of the underlying semigroup ensures us that, whenever a measure exists, our theory can be easily restricted to the space of continuous paths.) In the last section, we provide a proof of the last remaining conjecture of Dyson, concerning the cause for the ultraviolet divergency of quantum electrodynamics.

Although our major focus is functional analysis and the Feynman operator calculus, it is clear from the topics covered that the book has much to offer for those with general research interests in both pure and applied mathematics. The book can be used as a text for advanced courses in analysis, functional analysis, operator theory, mathematical physics, mathematical foundations of quantum theory, or special topic seminars in these or related subjects.

Those with advanced training in quantum theory, who mainly work on Hilbert spaces, could study the first part of Sect. 3.2 and the proof of Theorem 3.25 in Chap. 3. A review of the first two subsections of Chap. 6, Sect. 6.5.1 of Sect. 6.5, and Sect. 6.6 would be sufficient to understand Chap. 7 and the main section on path integrals in Chap. 8.

Acknowledgments

The research for this book was supported by the Howard University research fund and later by the Departments of E and CE, Mathematics, and Physics. For this, we are truly grateful.

We are grateful to a number of students and colleagues at Howard University who have helped with corrections and/or many suggestions for improvement. Particular thanks goes to Dr. Vernice Steadman (now professor at the University of the District of Columbia), Mrs. Marzett Golden, Mr. Timothy Myers, Professor Neal Hindman, and Professor Daniel Williams. (Some of the research for Chap. 2 was jointly developed with Mr. Myers, while some of the research for Chaps. 5 and 6 was jointly developed with Mrs. Golden.) However, we take full responsibility for any errors and would sincerely appreciate being apprised of any other corrections (and/or suggestions).

We are also appreciative of the many corrections, suggestions, and encouragement received from our colleagues outside Howard and especially Paul Chernoff, Lawrence Evans, Jerome Goldstein, Rhonda Hughes, Paul Humke, Gerald Johnson, Michel Lapidus, Michael Reed, David Skoug, and Herbert Winful.

The work in Chap. 2 would not have been possible without the help of Professor Frank Jones of Rice University. The possibility that we might be able to construct a reasonable version of Lebesgue measure on $\mathbb{R}^\infty$ came about after the first author taught analysis using the excellent text by Jones [J]. We further benefited from his many emails and patient phone conversations concerning problems we encountered with our construction. The remarkable Jones functions used in Chap. 3 to construct the class of Jones strong distribution Banach spaces can be found in his book (see page 249). We have also benefited from insightful comments and suggestions from Lance Nelson of Creighton University, USA; Gogi Pantsulaia, I. Vekua Institute of Mathematics, Georgia; and Erik Talvila, Simon Fraser University, Canada.

The refinement of research leading to Chaps. 7 and 8 was carried out while the first named author was supported as a member of the School of Mathematics of the Institute for Advanced Studies, Princeton NJ. The support of Professors Freeman Dyson, Robert Langlands, and Thomas Spencer of IAS and that of Professors Michael Aizenman and Elliott Lieb of Princeton University was most appreciated. Additional support was provided by the Department of Physics of the University of Michigan and the School of Mathematics, University of Minnesota.

Finally, we would like to sincerely thank Ms. Lynn Kinnear, director of the Smith Memorial Library at the Chautauqua Institution, Chautauqua NY. She made it possible for the first author to spend four summers working in a special section of the library, which provided the perfect environment for research and writing.

This book is dedicated to the Reverend LaVerne M. Gill and to the memory of Professors Gerald Chachere, Albert Turner Bharucha-Reid, George R. Sell, and Woodford William Zachary.

Washington, DC, USA

Tepper L. Gill
Woodford W. Zachary

References

[AG] M. Aizenman, R. Graham, Nucl. Phys. B **225**, 261 (1983)

[AK] H. Araki, Expansional in Banach algebras. Ann. scient. Éc. Norm. Sup. Math. **4**, 67–84 (1973)

[F] R.P. Feynman, An operator calculus having applications in quantum electrodynamics. Phys. Rev. **84**, 108–128 (1951)

[FH] R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965)

[FO] J. Fröhlich, Nucl. Phys. B **200**, 281 (1982)

[FW] I. Fujiwara, Operator calculus of quantized operator. Prog. Theor. Phys. **7**, 433–448 (1952)

[GJ] J. Glimm, A. Jaffe, *Quantum Physics. A Functional Integral Point of View* (Springer, New York, 1987)

[GZ5] T.L. Gill, W.W. Zachary, Two mathematically equivalent versions of Maxwell's equations. Found. Phys. **41**, 99–128 (2011)

[GMK] T.L. Gill, T. Morris, S.K. Kurtz, Foundations for proper-time relativistic quantum theory. J. Phys. Conf. Ser. **615**, 012013 (2015)

[J] F. Jones, *Lebesgue Integration on Euclidean Space*, Revised edn. (Jones and Bartlett Publishers, Boston, 2001)

[KP2] A.P. Kirtadze, G.R. Pantsulaia, Lebesgue nonmeasurable sets and the uniqueness of invariant measures in infinite-dimensional vector spaces. Proc. A. Razmadze Math. Inst. **143**, 95–101 (2007)

[RS1] M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic, New York, 1972)

[SC1] S.S. Schweber, *QED and the Men Who Made It* (Princeton University Press, Princeton, 1994)

[SCH] L. Schwartz, *Théorie des Destributions* (Hermann, Paris, 1966)

[SH] I.A. Shishmarev, On the Cauchy problem and T-products for hypoelliptic systems. Math. USSR Izvestiya **20**, 577–609 (1983)

[SO] A.D. Sokal, Ann. Inst. Henri Poincare Sect. A **37**, 13 (1982)

[SW] R.F. Streater, A.S. Wightman, *PCT, Spin and Statistics and All That* (Benjamin, New York, 1964)

[YA1] Y. Yamasaki, Translationally invariant measure on the infinite-dimensional vector space. Publ. Res. Inst. Math. Sci. **16**(3), 693–720 (1980)

Contents

Preliminary Background

This chapter is composed of two parts: Basic Analysis and Intermediate Analysis.

The first part is a review of some of the basic background that is required from the first 2 years of a standard program in mathematics. There are program differences so that some areas may receive more coverage while others receive less. Our purpose is to provide a reference point for the reader and establish notation. In a few important cases, we have provided proofs of major theorems. In other cases, we delayed a proof when a more general result is proven in a later chapter.

In the second part of this chapter, we include some intermediate to advanced material that is required later. In most cases, motivation is given along with additional proof detail and specific references.

Part I: Basic Analysis

The first part of this chapter is devoted to a brief discussion of the circle of ideas required for advanced parts of analysis and the basics of operator theory. Those with a strong background in theoretical chemistry or physics but little or no formal training in analysis will find Reed and Simon (vol.1) to be an excellent copilot (see below).

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_1

General references for this section are Dunford and Schwartz [DS], Jones [J], Reed and Simon [RS], Royden [RO], and Rudin [RU].

1.1. Analysis

1.1.1. Sets. Let X be a nonempty set, let $\emptyset$ be the emptyset, and let $\mathcal{P}(X)$ be the power set of X (i.e., the set of all subsets of X).

Definition 1.1. Let $A, B, A_n \in \mathcal{P}(X), n \in \mathbb{N}$, then

(1) $A^c = \{a \in X : a \notin A\}$, the compliment of A.

(2) $A \backslash B = A \cap B^c$.

(3) (De Morgan's Laws)

$$\left[\bigcup_{k=1}^{\infty} A_k\right]^c = \bigcap_{k=1}^{\infty} A_k^c, \qquad \left[\bigcap_{k=1}^{\infty} A_k\right]^c = \bigcup_{k=1}^{\infty} A_k^c.$$

We define the lim inf and lim sup for sets by:

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \qquad \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Theorem 1.2. *Let $\{A_n\} \subset \mathcal{P}(X), n \in \mathbb{N}$, then the* lim inf and lim sup *satisfy:*

(1)

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n.$$

(2)

$$\limsup_{n\to\infty} A_n = \{a : a \in A_k \text{ for infinitely many } k\}.$$

(3)

$$\liminf_{n\to\infty} A_n = \{a : a \in A_k \text{ for all but finitely many } k\}.$$

(4)

$$(\limsup_{n\to\infty} A_n)^c = \liminf_{n\to\infty} A_n^c.$$

(5) *If $A_n \supset A_{n+1}$, then*

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = \lim_{n\to\infty} A_n = \bigcap_{k=1}^{\infty} A_k.$$

(6) *If* $A_n \subset A_{n+1}$, *then*

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = \lim_{n\to\infty} A_n = \bigcup_{k=1}^{\infty} A_k.$$

Definition 1.3. Let $A, B \subset X$. (We assume they are nonempty.)

(1) The cartesian product, denoted $A \times B$, is defined by

$$A \times B = \{(a, b) : a \in A,\ b \in B\}.$$

In general, $A \times B \neq B \times A$, so that the order matters. If $\{A_k\}$ is a countable collection of subsets of X, we define the cartesian product by:

$$\prod_{k=1}^{\infty} A_k = \{(a_1, a_2, \dots) : a_k \in A_k\}.$$

Definition 1.4. A map $f : A \to B$ (a function, or a transformation), with domain $D(f) \subset A$ and range $R(f) \subset B$ is a subset $f \subset A \times B$ such that, for each $x \in A$, there is one and only one $y \in B$, with $(x, y) \in f$. We write $y = f(x)$ and call $f(A) = \{f(x) : x \in A\} \subset B$, the image of f and, call $f^{-1}(B) = \{x : f(x) \in B\} \subset A$, the inverse image of B. We say that f is one to one or injective, if for all $x_1 \neq x_2 \in A$, we have that $y_1 = f(x_1) \neq y_2 = f(x_2) \in B$. We say that f is onto or surjective if, for each $y \in B$, there is a $x \in A$, with $y = f(x)$.

1.1.2. Topology. We only consider Hausdorff spaces or spaces with the Hausdorff topology (see below). For an elementary introduction to topology, we recommend Mendelson [ME]. Dugundji [DU] is more advanced, but is also worth consulting.

Definition 1.5. Let X be a nonempty set and let τ be a set of subsets of X. We say that τ defines a Hausdorff topology on X, or that X is Hausdorff, if

(1) X and $\emptyset \in \tau$.

(2) If $O_1, \dots, O_n$ is a finite collection of sets in τ, then $\bigcap_{i=1}^{n} O_i \in \tau$.

(3) If Γ is a index set and, for each $\gamma \in \Gamma$, there is a set $O_\gamma \in \tau$, then $\bigcup_{\gamma \in \Gamma} O_\gamma \in \tau$.

(4) If $x, y \in X$ are any two distinct points, there are two disjoint sets $O_1, O_2 \in \tau$ (i.e., $O_1 \cap O_2 = \emptyset$), such that $x \in O_1$ and $y \in O_2$.

We call the collection τ the open sets of the topology for X. A set $N \in \tau$ is called a neighborhood for each point $x \in N$, and the set $\tau_x \subset \tau$ of all neighborhoods for x is called a complete neighborhood basis for x. Thus, any set O, containing x, also contains some neighborhood basis set $N(x) \in \tau_x$.

A set P is said to be closed if P^c is open. It follows that, if Γ is any index set and, for each $\gamma \in \Gamma$, there is a closed set $P_\gamma \in \tau$, then by De Morgan's Law, $\bigcap_{\gamma \in \Gamma} P_\gamma$ is also closed. Thus, we can also define the <u>same</u> topology τ, using closed sets.

Let $M \neq \emptyset$, be a subset of X.

(1) The interior of M, denoted $\mathrm{int}(M)$, is the union of all $O \in \tau$ such that $O \subset M$. If $x \in \mathrm{int}(M)$, we say that x is an interior point of M.

(2) The closure of M, which we denote by $\overline{M}$, is the set of all $x \in X$ such that, for all $N(x) \in \tau_x$, $N(x) \cap M \neq \emptyset$.

(3) We say that M is dense in X if $\overline{M} = X$. If M is also countable, we say that X is separable.

If M and N are any two subsets of X, then $\overline{M \cup N} = \overline{M} \cup \overline{N}$ and, $\overline{M} = M$ if and only if M is closed.

We say that $x_0 \in X$ is a limit point of $M \subset X$, if $x_0 \in \overline{M \backslash \{x_0\}}$ or equivalently, for every $N(x_0) \in \tau_{x_0}$, there is a $y \in N(x_0)$ and $y \notin M$.

Definition 1.6. Let (X_1, τ_1) and (X_2, τ_2) be two Hausdorff spaces. A function f, with $D(f) = X_1$ and $R(f) \subset X_2$, is said to be continuous at a point $x \in X_1$ if, for each neighborhood basis set $N[f(x)] \in \tau_{2,x}$, there is a neighborhood basis set $N(x) \in \tau_{1,x}$ such that $f[N(x)] \subset N[f(x)]$. In terms of inverse images, this says that $f^{-1}\{N[f(x)]\}$ is open in X_1 for each $N[f(x)]$ in X_2. (A little reflection shows that the above definition may be translated to the one we learned in elementary calculus, using ε's and δ's, when $X_1 = X_2 = \mathbb{R}$.) We say that f is continuous on X_1 if it is continuous at each point of X_1.

The topological space (X, τ) is said to be connected if it is not the disjoint union of two open sets. In a connected space X and $\emptyset$ are the only two sets that are both open and closed.

If Γ is a index set, $\{A_\gamma : \gamma \in \Gamma\} \subset X$ is called a cover of $M \subset X$, if $M \subset \bigcup_{\gamma \in \Gamma} A_\gamma$. If each $A_\gamma \in \tau$, we call $\{A_\gamma : \gamma \in \Gamma\}$ an open cover of M. If in addition Γ is finite, we call it a finite open cover of M.

We say that M is compact if, for every open cover $\{A_\gamma : \gamma \in \Gamma\}$, there always exists a finite subset of Γ, $\gamma_1, \ldots, \gamma_n$ such that $M \subset \bigcup_{k=1}^{n} A_{\gamma_k}$.

Definition 1.7. Let (X_1, τ_1) and (X_2, τ_2) be two topological spaces, with $X_1 \cap X_2 = \emptyset$. The coproduct space $(X, \tau) = (X_1, \tau_1) \oplus (X_2, \tau_2)$ is the unique topological space, with the property that each open set $O \subset X$ is of the form $O = O_1 \cup O_2$, where $O_1 \in \tau_1$ and $O_2 \in \tau_2$.

(X, τ) is also known as the disjoint union space or direct sum space. (If (X_1, τ_1) and (X_2, τ_2) are Hausdorff, then it is easy to see that (X, τ) is Hausdorff.)

1.1.3. σ-Algebras.

Definition 1.8. Let $\mathcal{A} \subset \mathcal{P}(X)$ be a collection of subsets of $X \neq \emptyset$. We say that $\mathcal{A}$ is an algebra if the following holds:

(1) $X, \emptyset \in \mathcal{A}$ and,

(2) If $A, B \in \mathcal{A}$ then $A^c, B^c \in \mathcal{A}$ and $A \cup B \in \mathcal{A}$.
It is easy to verify that:

(3) $A \cap B \in \mathcal{A}$ and $A \setminus B \in \mathcal{A}$.

(4) If n is finite and $\{A_k\} \subset \mathcal{A}$, $1 \leq k \leq n$, then

$$\bigcup_{k=1}^{n} A_k \in \mathcal{A}, \qquad \bigcap_{k=1}^{n} A_k \in \mathcal{A}.$$

Definition 1.9. Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra. We say that $\mathcal{A}$ is a σ-algebra if

$$\bigcup_{k=1}^{\infty} A_k \in \mathcal{A},$$

for any countable family of sets $\{A_k\} \in \mathcal{A}$. It is also easy to see that

$$\bigcap_{k=1}^{\infty} A_k \in \mathcal{A},$$

along with

$$\liminf_{n \to \infty} A_n \in \mathcal{A}$$

and

$$\limsup_{n \to \infty} A_n \in \mathcal{A}.$$

Definition 1.10. If Σ is a nonempty class of subsets of X, the smallest σ-algebra $\mathcal{A}$, with $\Sigma \subset \mathcal{A}$ is called the σ-algebra generated by Σ and is written $\mathcal{A}(\Sigma)$.

Remark 1.11. Since $\Sigma \subset \mathcal{P}(X)$, there is at least one σ-algebra containing Σ.

Lemma 1.12. *If J is an index set and for each $\alpha \in J$, $\mathcal{A}_\alpha$ is σ-algebra, then $\mathcal{A} = \bigcap_{\alpha \in J} \mathcal{A}_\alpha$ is a σ-algebra.*

Definition 1.13. If $\mathcal{A}$ is a σ-algebra of subsets of a nonempty set X, we call the couple $(X, \mathcal{A})$ a measurable space.

Definition 1.14. If $\mathcal{A}$ is a σ-algebra of subsets of a nonempty set X, we call a sequence $\{A_k\} \subset \mathcal{A}$ a partition of X if the sequence is disjoint and $\bigcup_{k=1}^{\infty} A_k = X$.

Definition 1.15. If X is a topological space and Σ is the class of open sets of X, then $\mathcal{A}(\Sigma) = \mathfrak{B}(X)$ is called the Borel σ-algebra of X.

1.1.4. Measure Spaces.

Definition 1.16. Let X be a nonempty set. An outer measure ν^* is a function on $\mathcal{P}(X) \to [0, \infty]$, such that

(1) $\nu^*(\emptyset) = 0$.
(2) If $B \subset A$, then $\nu^*(B) \leq \nu^*(A)$.
(3) If $A \subset \bigcup_{k=1}^{\infty} A_k$, then
$$\nu^*(A) \leq \sum_{k=1}^{\infty} \nu^*(A_k).$$

If for each sequence of disjoint sets $\{A_k\} \subset \mathcal{A}$,
$$\nu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \nu(A_k),$$
we say that ν is a measure. We also say that ν is σ-additive and call the triple $(X, \mathcal{A}, \nu)$ a measure space.

Definition 1.17. Let $(X, \mathcal{A})$ be a measurable space and let $\nu(A) \in \mathbb{C}$, the complex numbers, for each $A \in \mathcal{A}$. We say that ν is a complex measure if $\nu(\emptyset) = 0$ and for each disjoint countable union $\bigcup_{k=1}^{\infty} A_k$ of sets in $\mathcal{A}$, we have
$$\nu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \nu(A_k),$$
where the convergence on the right is absolute.

Definition 1.18. Let $(X, \mathcal{A}, \nu)$ a measure space.

(1) We say that ν is a finite measure if $\nu(X) < \infty$.

(2) We say that ν is concentrated on a set $A \in \mathcal{A}$, if $A = U^c$ and U is the largest open set with the property that $\nu(U) = 0$. We also call A the support of ν.

(3) We say that ν is a regular measure if given $A \in \mathcal{A}$, for each $\varepsilon > 0$, there is a open set O and a closed set K such that: $K \subset A \subset O$ and $\nu(O \setminus K) < \varepsilon$.

(4) We say that ν is a σ-finite measure if there is a sequence $\{A_k\} \subset \mathcal{A}$, with

$$X = \bigcup_{k=1}^{\infty} A_k, \text{ and } \nu(A_k) < \infty.$$

(5) We say that ν is a Radon measure, if the set K in (3) can be chosen as compact or the sequence $\{A_k\} \subset \mathcal{A}$ in (4) can be chosen with each A_k is compact.

(6) We say that ν is a complete measure if $A \in \mathcal{A}$, with $B \subset A$ and $\nu(A) = 0$ then $B \in \mathcal{A}$ and $\nu(B) = 0$.

(7) We say that ν is a probability measure if $\nu(X) = 1$.

(8) We say that a complex measure ν is of bounded variation if

$$|\nu|(X) = \sup \sum_{k=1}^{\infty} |\nu(A_k)| < \infty,$$

where the supremum is taken over all partitions of X. We call $|\nu|(X)$ the total variation of ν.

(9) We say that the complex measure ν is a signed measure if both $|\nu| + \nu$ and $|\nu| - \nu$ are real valued. In this case, we define the positive part and the negative part by: $\nu^+ = \frac{1}{2}(|\nu| + \nu)$ and $\nu^- = \frac{1}{2}(|\nu| - \nu)$. We call this the Jordan Decomposition.

Theorem 1.19 (The Hahn Decomposition Theorem)**.** *Let ν be a signed measure on $(X, \mathcal{A})$. Then there exists a partition X_1, X_2 of X such that, for every $A \in \mathcal{A}$:*

$$\nu^+(A) = \nu(A \cap X_1) \text{ and } \nu^-(A) = -\nu(A \cap X_2).$$

Theorem 1.20 (The Jordan Decomposition Theorem)**.** *Let ν be a signed measure on $(X, \mathcal{A})$. If μ_1 and μ_2 are positive measures and $\nu = \mu_1 - \mu_2$, then $\nu^+ \leq \mu_1$ and $\nu^- \leq \mu_2$.*

Thus, the Jordan decomposition $\nu = \nu^+ - \nu^-$, has the above minimal property. If ν is complex, this decomposition becomes $\nu = \nu_1^+ - \nu_1^- + i(\nu_2^+ - \nu_2^-)$, for two positive measures, ν_1 and ν_2.

Definition 1.21. We say that X is an Abelian group if for each pair $x, y \in X$, $x \oplus y \in X$ and

(1) $x \oplus y = y \oplus x$. (The Abelian property.)
(2) For all $x, y, z \in X$ $(x \oplus y) \oplus z = x \oplus (y \oplus z)$.
(3) There is an element $0 \in X$ called the identity and $x \oplus 0 = 0 \oplus x = 0$, for all $x \in X$.
(4) For each $y \in X$, there is a unique element $y^- \in X$, such that $y \oplus y^- = y^- \oplus y = 0$.
(5) We say that Y is a subgroup of X if $Y \subset X$ and for all $y_1, y_2 \in Y$, $y_1 \oplus y_2 \in Y$, satisfying conditions (1)–(4) above.

The real or complex numbers form an Abelian group with addition (or multiplication if we exclude zero). The rational numbers (real or complex) form a subgroup, with the same exception for multiplication.

When X is an Abelian group (with $\oplus = +$) and $(X, \mathcal{A}, \nu)$ is a measure space, we say that $\mathfrak{T}$ is an admissible translation invariance group for $(X, \mathcal{A}, \nu)$ if $\mathfrak{T}$ is a subgroup of X and $\nu(A - t) = \nu(A)$, for all $t \in \mathfrak{T}$. If $\mathfrak{T} = X$, we say that ν is translation invariant on X.

1.1.5. Integral.

Let $(X, \mathcal{A}, \nu)$ a measure space.

Definition 1.22. Let f be a function on X, $f : X \to K$, where $K = \mathbb{R}$ or $\mathbb{C}$.

(1) We say that f is measurable if $f^{-1}(B) \in \mathcal{A}$, for every set $B \in \mathfrak{B}[K]$, the Borel algebra on K. In this case, we say that $f \in \mathcal{M}[X]$ or $\mathcal{M}$, when X is understood.
(2) We say that two functions f and g are equal almost everywhere and write $f(x) = g(x)$, ν-(a.e.), if they have the same domain and $\nu\{x : f(x) \neq g(x)\} = 0$. In general, a property is said to hold ν-(a.e.) on X if the set of points where this property fails has ν-measure zero.

Definition 1.23. A (nonnegative) simple function s is defined on X by

$$s(x) = \sum_{k=1}^{n} a_k \chi_{A_k}(x),$$

where the $a_k \in [0, \infty)$ and the family of measurable sets $\{A_k\}$ form a (finite) partition of X (i.e., $\nu(A_i \cap A_j) = 0$, $i \neq j$ and $\bigcup_{k=1}^{n} A_k = X$).

(By convention, if need be, we can always add a set A_{n+1} to the collection and define $a_{n+1} = 0$ so that the union is always X.)

Lemma 1.24. *If $0 \leq f \in \mathcal{M}$, then there is a sequence of simple functions $\{s_n\}$, with $s_n \leq s_{n+1}$ and $s_n \to f$ (a.e.) at each point of X, as $n \to \infty$.*

Definition 1.25. If $f : X \to [0, \infty]$ is a measurable function and $A \in \mathfrak{B}(X)$, we define the integral of f over A by:

$$\int_A f(x) d\nu = \lim_{n \to \infty} \int_A s_n(x)\, d\nu,$$

where $\{s_n\}$ is any increasing family of simple functions converging to $f(x)$.

Theorem 1.26. *If f, g are nonnegative measurable functions and $0 \leq c < \infty$, we have:*

(1) *$\int_X f(x) d\nu(x)$ is independent of the family of simple functions used;*

(2) *$0 \leq \int_X f(x) d\nu(x) \leq \infty$;*

(3) *$\int_X c f(x) d\nu(x) = c \int_X f(x) d\nu(x)$;*

(4)

$$\int_X [f(x) + g(x)] d\nu(x) = \int_X f(x) d\nu(x) + \int_X g(x) d\nu(x).$$

(5) *If $f \leq g$, then $\int_X f(x) d\nu(x) \leq \int_X g(x) d\nu(x)$.*

Theorem 1.27 (Fatou's Lemma)**.** *Let $\{f_n\} \subset \mathcal{M}$ be a nonnegative family of functions, then:*

$$\int_X \left(\liminf_{n \to \infty} f_n(x) \right) d\nu(x) \leq \liminf_{n \to \infty} \int_X f_n(x) d\nu(x).$$

Theorem 1.28 (Monotone Convergence Theorem)**.** *Let $\{f_n\} \subset \mathcal{M}$ be a nonnegative family of functions, with $f_n \leq f_{n+1}$. Then:*

$$\lim_{n \to \infty} \int_X f_n(x) d\nu(x) = \int_{\mathcal{B}} \left(\lim_{n \to \infty} f_n(x) \right) d\nu(x).$$

Definition 1.29. If $f \in \mathcal{M}$, we define

$$\int_X f(x) d\nu(x) = \int_X f_+(x) d\nu(x) - \int_X f_-(x) d\nu(x),$$

where $f_+(x) = \frac{1}{2}(|f(x)| + f(x))$ and $f_-(x) = \frac{1}{2}(|f(x)| - f(x))$. We say that f is integrable whenever both integrals on the right are finite. The set of all integrable functions is denoted by $\mathcal{L}^1[X, \mathfrak{B}(X), \nu] = \mathcal{L}^1[X]$.

Remark 1.30. As is carefully discussed in elementary analysis, the functions in $\mathcal{L}^1[X]$ are not uniquely defined. Following tradition, we let $L^1[X]$ denote the set of equivalence classes of functions in $\mathcal{L}^1[X]$ that differ by a set of ν-measure zero. By a slight abuse, we will identify an integrable function f as measurable (in $\mathcal{L}^1[X]$) and its equivalence class in $L^1[X]$. The same convention also applies to functions in $L^p[X]$ and will be used later without further comment.

Theorem 1.31 (Dominated Convergence Theorem)**.** *Let* $f_n \in \mathcal{M}[X, \nu]$, $n \in \mathbb{N}$, $g \in L^1(X)$, *with* $g \geq 0$ *and* $|f_n(x)| \leq g(x)$, ν-(a.e.). *If* $\lim_{n\to\infty} f_n(x)$ *exists* ν-(a.e.), *then* $\lim_{n\to\infty} f_n \in L^1[X]$ *and*

$$\lim_{n\to\infty} \int_X f_n(x) d\nu(x) = \int_X \left(\lim_{n\to\infty} f_n(x)\right) d\nu(x).$$

1.2. Functional Analysis

In this section, we include a few basic background results from functional analysis and Banach space theory. Detailed discussions can be found in Dunford and Schwartz [DS], Hille and Phillips [HP], Lax [L1], Reed and Simon [RS], Rudin [RU], or Yosida [YS].

1.2.1. Topological Vector Spaces.

Definition 1.32. A vector space $\mathfrak{X}$ over $\mathbb{C}$ is an Abelian group under addition that is closed under multiplication by elements of $\mathbb{C}$. That is:

(1) For each $x, y \in \mathfrak{X}$, $x + y \in \mathfrak{X}$.

(2) For all $x, y, z \in \mathfrak{X}$, $x+y = y+x$ and $(x+y)+z = x+(y+z)$.

(3) There is a unique element $0 \in \mathfrak{X}$ called zero and $x + 0 = 0 + x = x$ for all $x \in \mathfrak{X}$.

(4) For all $x \in \mathfrak{X}$, there is a unique element $-x \in \mathfrak{X}$ and $x + (-x) = (-x) + x = 0$.

(5) For all $x, y \in \mathfrak{X}$ and $a, b \in \mathbb{C}$, $ax \in \mathfrak{X}$, $1x = x$, $(ab)x = a(bx)$ and $a(x + y) = ax + ay$. We call $b \in \mathbb{C}$ a scalar.

If $\mathfrak{X}$ is a vector space over $\mathbb{C}$, a mapping $\rho(\cdot) : \mathfrak{X} \to [0, \infty)$ is a seminorm on $\mathfrak{X}$ if:

(1) For each $x,\, y \in \mathfrak{X},\ \rho(x) \geq 0$ and $\rho(x+y) \leq \rho(x) + \rho(y)$.

(2) For each $\lambda \in \mathbb{C}$ and each $x \in \mathfrak{X},\ \rho(\lambda x) = |\lambda|\, \rho(x)$.

Definition 1.33. Let V be a subset of $\mathfrak{X}$.

(1) We say that V is a convex subset of $\mathfrak{X}$ if for each $x,\, y \in V,\ \alpha x + (1-\alpha)y \in V$, for all $\alpha \in [0,1]$.

(2) We say that V is an balanced subset of $\mathfrak{X}$ if for each $x \in V$ and $\alpha \in \mathbb{C}$, with $|\alpha| \leq 1,\ \alpha x \in V$.

(3) We say that V is an absolutely convex subset of $\mathfrak{X}$ if it is both convex and balanced.

(4) We say that V is a absorbent subset of $\mathfrak{X}$ if for each $x \in \mathfrak{X},\ \alpha x \in V$, for some $\alpha > 0$. Thus, every point in $x \in \mathfrak{X}$ is in αV for some positive α.

Definition 1.34. A locally convex topological vector space is a vector space with its topology defined by a family of semi-norms $\{\rho_\gamma\}$, where γ is in some index set Γ. Given any $x \in \mathfrak{X}$, a base of ε-neighborhoods about x is a set of the form $V_{\Gamma_0,\varepsilon}(x)$, where Γ_0 is a finite subset of Γ and

$$V_{\Gamma,\varepsilon}(x) = \{\, y \in \mathfrak{X} : \rho_\gamma(x-y) < \varepsilon,\, \gamma \in \Gamma \,\}.$$

Definition 1.35. A locally convex topological vector space $\mathfrak{X}$ is a Fréchet space if it satisfies the following:

(1) $\mathfrak{X}$ is a Hausdorff space.

(2) The neighborhood base about each $x \in \mathfrak{X}$ is induced by a countable number of seminorms (i.e., Γ is a countable set).

(3) $\mathfrak{X}$ is a complete relative to the family of seminorms.

Theorem 1.36. *The vector space $\mathfrak{X}$ is a Fréchet space if and only if:*

(1) *$\mathfrak{X}$ is a locally convex.*

(2) *There is a metric $d : \mathfrak{X} \times \mathfrak{X} \to [0,\infty)$ such that, for all $x, y, z \in \mathfrak{X},\ d(x+z, y+z) = d(x,y)$.*

(3) *$\mathfrak{X}$ is a complete relative to the metric $d(\cdot,\cdot)$.*

Remark 1.37. If the index Γ for the family of semi-norms is countable, then we can define a metric $d(x,y)$ by:

$$d(x,y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\rho_n(x-y)}{1+\rho_n(x-y)}.$$

A sequence $\{x_n\}$ in a metric space $\mathfrak{X}$ converges to a limit $x \in \mathfrak{X}$ if and only if $\lim_{n\to\infty} d(x_n, x) = 0$. In this case, by the triangle inequality

$$d(x_n, x_m) \leq d(x_n, x) + d(x_m, x).$$

We say that a sequence satisfies the Cauchy convergence condition, or is a Cauchy sequence if

$$\lim_{m,n\to\infty} d(x_n, x_m) = 0.$$

A metric space is said to be complete if every Cauchy sequence converges to a point in the space.

1.2.2. Separable Banach Spaces. Hilbert and Banach spaces are discussed further in Chaps. 4 and 5. Let $\mathcal{B}$ be a vector space over $\mathbb{R}$ or $\mathbb{C}$. We say that $\mathcal{B}$ is separable if it contains a countable dense subset.

Definition 1.38. A norm on a vector space $\mathcal{B}$ is a mapping $\|\cdot\|_{\mathcal{B}}$: $\mathcal{B} \to [0, \infty]$, such that

(1) $\|x\|_{\mathcal{B}} = 0$ if and only if $x = 0$.
(2) $\|ax\|_{\mathcal{B}} = |a| \, \|x\|_{\mathcal{B}}$ for all $x \in \mathcal{B}$ and $a \in \mathbb{C}$.
(3) $\|x + y\|_{\mathcal{B}} \leq \|x\|_{\mathcal{B}} + \|y\|_{\mathcal{B}}$, for all $x, y \in \mathcal{B}$.
(4) We say that $\mathcal{B}$ is uniformly convex if, for each $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) > 0$ such that, for all $x, y \in \mathcal{B}$ with

$$\max(\|x\|, \|y\|) \leqslant 1, \quad \|x - y\| \geqslant \varepsilon \Rightarrow \tfrac{1}{2}\|x + y\| \leqslant 1 - \delta.$$

The topology on $\mathcal{B}$ is generated by the metric defined by:

$$d(x, y) = \|x - y\|_{\mathcal{B}},$$

so that $\{x : \|x - y\|_{\mathcal{B}} < r\}$ is an open ball about y of radius r.

The space $\mathcal{B}$ is complete if every Cauchy sequence in the above norm converges to an element in $\mathcal{B}$. A complete normed space is called a Banach space.

Definition 1.39. Let $\mathcal{B}$ be a Banach space and let A be a transformation on $\mathcal{B}$, with domain $D(A)$ (i.e., $A : D(A) \subset \mathcal{B} \to \mathcal{B}$).

(1) We say that A is a linear operator on $\mathcal{B}$, if $A(ax + by) = aAx + bAy$, for all $a, b \in \mathbb{C}$ and all $x, y \in D(A)$.
(2) We say that A is densely defined if $D(A)$ is dense in $\mathcal{B}$.

(3) We say that A is a closed linear operator if and only if the following condition is satisfied: $\{x_n\} \subset D(A), x_n \to x$ and $Ax_n \to z$ always implies that $x \in D(A)$ and $z = Ax$.

(4) We say that A is a bounded linear operator if and only if $D(A) = \mathcal{B}$ and

$$\sup_{\|x\|_{\mathcal{B}} \leqslant 1} \|Ax\|_{\mathcal{B}} < \infty.$$

In this case we define the norm of A, $\|A\|_{\mathcal{B}}$, by the above supremum.

1.2.2.1. *Dual Spaces.*

Definition 1.40. Let $\mathcal{B}$ be a Banach space.

(1) The dual space $\mathcal{B}'$ is the set of all bounded linear operators $x^* : \mathcal{B} \to \mathbb{C}$ (called bounded linear functionals on $\mathcal{B}$). The norm of x^* is defined by:

$$\|x^*\|_{\mathcal{B}'} = \sup_{\|x\|_{\mathcal{B}} \leq 1} |x^*(x)| = \sup_{\|x\|_{\mathcal{B}} \leq 1} |\langle x, x^* \rangle| \,.$$

With this norm $\mathcal{B}'$ is a Banach space. We write $\mathcal{B}'$ as $\mathcal{B}'_s$ and call it the strong dual. The topology is known as the strong topology.

(2) The weak and weak* topology are defined on $\mathcal{B}$ and $\mathcal{B}'$ respectively in the following manner:

- A sequence $\{x_n\} \subset \mathcal{B}$ is said to converge in the weak topology to $x \in \mathcal{B}$ if and only if, for each bounded linear functional $y^* \in \mathcal{B}'$,

$$\lim_{n \to \infty} y^*(x_n) = y^*(x).$$

We also write $w - \lim_{n\to\infty} x_n = x$.

- A sequence $\{x_n^*\} \subset \mathcal{B}'$ is said to converge in the weak* topology to $x^* \in \mathcal{B}'$ if and only if, for each $y \in \mathcal{B}$,

$$\lim_{n \to \infty} x_n^*(y) = x^*(y).$$

We also write $w^* - \lim_{n\to\infty} x_n^* = x^*$.

(3) If $\mathcal{B} = \mathcal{B}''$, we say that $\mathcal{B}$ is reflexive.

(4) A duality map $\mathcal{J} : \mathcal{B} \mapsto \mathcal{B}'$ is a set

$$\mathcal{J}(u) = \left\{ u^* \in \mathcal{B}' \,\middle|\, u^*(u) = \langle u, u^* \rangle = \|u\|^2 = \|u^*\|^2 \right\}, \text{ for all } u \in \mathcal{B}.$$

Remark 1.41. The following remarks are important.

(1) In the definition, we used x^* to represent an element in $\mathcal{B}'$. The notation used varies with the tradition of the particular topical area. To the extent possible, we will try to be consistent within topics studied and the tradition of the field so that the reader will see some correspondence when consulting references for different topics.

(2) It is easy to see that

$$|y^*(x_n) - y^*(x)| \leq \|x_n - x\|_{\mathcal{B}} \|y^*\|_{\mathcal{B}'}$$

for all $y^* \in \mathcal{B}$, so that norm convergence in $\mathcal{B}$ always implies weak convergence. It is also easy to see that

$$|x_n^*(y) - x^*(y)| \leq \|x_n^* - x^*\|_{\mathcal{B}'} \|y\|_{\mathcal{B}},$$

for all $y \in \mathcal{B}$, so that norm convergence in $\mathcal{B}'$ always implies weak* convergence. However (in both cases), the reverse is not true (see Lax [L1, p. 106]).

(3) It is known that every uniformly convex Banach space is reflexive. Furthermore, when $\mathcal{B}$ is uniformly convex, the duality set $\mathcal{J}(u)$, is single valued and uniquely defined by u. However, if $\mathcal{B}$ is not uniformly convex, the duality set $\mathcal{J}(u)$ can have the power of the continuum.

The following examples will help one see what is possible in concrete cases.

(1) If λ_n is Lebesgue measure on $\mathbb{R}^n$, $u \in L^p[\mathbb{R}^n]$, $1 < p < \infty$ and q is such that $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\mathcal{J}(u)(x) = \|u\|_p^{2-p} |u(x)|^{p-2} u(x) = u^* \in L^q\mathbb{R}^n],$$

and

$$\langle u, u^* \rangle = \|u\|_p^{2-p} \int_{\mathbb{R}^n} |u(x)|^p \, d\lambda_n(x) = \|u\|_p^2 = \|u^*\|_q^2 .$$

Thus, it is easy to see that $(L^p[\mathbb{R}^n])'' = L^p[\mathbb{R}^n]$, so that $L^p[\mathbb{R}^n]$ is reflexive for $1 < p < \infty$.

(2) The space $L^1[\mathbb{R}^n]$ is not reflexive, for if $u \in L^1[\mathbb{R}^n]$, then

$$\mathcal{J}(u)(x) = \{ v \in L^\infty[\mathbb{R}^n]| \; : v(x) \in \{\|u\|_1 \operatorname{sign}[u(x)]\}\},$$

where

$$\operatorname{sign}[u(x)] = \begin{cases} 1, & u(x) > 0, \\ -1, & u(x) < 0, \\ [-1, 1], & u(x) = 0. \end{cases}$$

It follows that $\mathcal{J}(u)(x)$ is uncountable for each $u \in L^1[\mathbb{R}^n]$.

The transpose matrix on $\mathbb{R}^n$ or the transpose conjugate matrix on $\mathbb{C}^n$ has its parallel for Banach spaces. In this case, they are known as dual operators. They are also known as adjoint operators, but we will reserve this term for a special class of operators on Banach spaces, discussed in Chap. 5. We will also use adjoint for the same class defined on Hilbert spaces in the next section and explain the distinction.

Definition 1.42. Let $A : D(A) \to \mathcal{B}$ be a closed linear operator on $\mathcal{B}$ with a dense domain $D(A)$. The dual of A, A' is defined on $\mathcal{B}'$ as follows. Its domain $D(A')$ is the set of all $y^* \in \mathcal{B}'$ for which there exists an $x^* \in \mathcal{B}'$ such that

$$\langle Ax, y^* \rangle = \langle x, x^* \rangle ,$$

for all $x \in D(A)$; in this case we define $A'y^* = x^*$.

A proof of the following theorem can be found in [HP] or [YS].

Theorem 1.43. *Let $A : D(A) \to \mathcal{B}$ be a closed linear operator on $\mathcal{B}$ with a dense domain $D(A)$.*

(1) *Then $A' : D(A') \to \mathcal{B}'$ is a closed linear operator on $\mathcal{B}'$ and its domain $D(A')$ is dense in $\mathcal{B}'$.*
(2) *If, in addition, $\|A\|_{\mathcal{B}} < \infty$, then $D(A') = \mathcal{B}'$ and $\|A'\|_{\mathcal{B}'} = \|A\|_{\mathcal{B}}$.*

1.2.2.2. *Hilbert Space.*

Definition 1.44. An inner product on $\mathcal{B} = \mathcal{H}$ is a bilinear mapping $(\cdot, \cdot)_{\mathcal{H}} : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, such that

(1) $(x, x)_{\mathcal{H}} \geq 0$ and $(x, x)_{\mathcal{H}} = 0$ if and only if $x = 0$.
(2) $(ax + by, z)_{\mathcal{H}} = a(x, z)_{\mathcal{H}} + b(y, z)_{\mathcal{H}}$ and $(w, ax + by)_{\mathcal{H}} = a^c(w, x)_{\mathcal{H}} + b^c(w, y)_{\mathcal{H}}$.

If $(\cdot, \cdot)_{\mathcal{H}}$ is a inner product, it induces a norm on $\mathcal{H}$ by

$$\|x - y\|_{\mathcal{H}} = \sqrt{(x - y, x - y)_{\mathcal{H}}}.$$

If $\mathcal{H}$ is complete with this norm, we call it a Hilbert space.

If $(\cdot,\cdot)_{\mathcal{H}}$ is the inner product on the Hilbert space $\mathcal{H}$, then the same Cauchy–Schwarz inequality from $\mathbb{R}^n$ still holds, $|(x,y)_{\mathcal{H}}| \leq \|x\|_{\mathcal{H}}\,\|y\|_{\mathcal{H}}$.

The following polarization identity also holds for a general Hilbert space:

$$(x,y)_{\mathcal{H}} = \frac{1}{4}\left(\|x+y\|^2_{\mathcal{H}} - \|x-y\|^2_{\mathcal{H}}\right),$$

if the field of $\mathcal{H}$ is $\mathbb{R}$ and

$$(x,y)_{\mathcal{H}} = \frac{1}{4}\left\{\left(\|x+y\|^2_{\mathcal{H}} - \|x-y\|^2_{\mathcal{H}}\right) + i\left(\|x+iy\|^2_{\mathcal{H}} - \|x-iy\|^2_{\mathcal{H}}\right)\right\},$$

if the field of $\mathcal{H}$ is $\mathbb{C}$.

Definition 1.45. Let $A : D(A) \to \mathcal{H}$ be a closed linear operator on $\mathcal{H}$ with a dense domain $D(A)$. The adjoint of A, A^* is defined on $\mathcal{H}$ as follows. Its domain $D(A^*)$ is the set of all $y \in \mathcal{H}$ for which there exists an $x \in \mathcal{H}$ such that

$$(Ax,y)_{\mathcal{H}} = (x,A^*y)_{\mathcal{H}}.$$

We will always call A^* the adjoint of A when it is defined on the same space and A', the dual of A when it is defined on the dual space. In Chap. 5, we will see that the adjoint is also possible for a certain class of Banach spaces, which include the uniformly convex ones.

Theorem 1.43 can be slightly modified to show that $D(A^*)$ is dense in $\mathcal{H}$ and, if $\|A\|_{\mathcal{H}} < \infty$, then $D(A^*) = \mathcal{H}$ and $\|A^*\|_{\mathcal{H}} = \|A\|_{\mathcal{H}}$.

Recall that, two functions $f, g \in \mathcal{H}$ are orthogonal, if $(f,g)_{\mathcal{H}} = 0$ and they are orthonormal if in addition, $\|f\|_{\mathcal{H}} = \|g\|_{\mathcal{H}} = 1$. A set $\{\phi_n\} \subset \mathcal{H}$ is an orthonormal basis for $\mathcal{H}$ if they are orthonormal and each $x \in \mathcal{H}$ can be written as $x = \sum_{k=1}^{\infty} a_k\phi_k$, for a unique family of scalars $\{a_n\} \subset \mathbb{C}$.

Definition 1.46. Let A be a linear operator defined on $\mathcal{H}$.

(1) We say that A is a projection operator if $A^2x = Ax$ for all $x \in \mathcal{H}$.

(2) We say that A is the self-adjoint if $D(A) = D(A^*)$ and $Ax = A^*x$, for all $x \in D(A)$.

(3) We say that a bounded linear operator A is the compact, if for every bounded sequence $\{x_n\} \subset \mathcal{H}$, the sequence $\{Ax_n\}$ has a convergent subsequence.

(4) We say that a compact operator A is trace class if, for some orthonormal basis $\{\phi_n\}$ of $\mathcal{H}$, the trace of A, $tr[A]$ is finite,

where

$$tr[A] = \sum_{n=1}^{\infty} (A\phi_n, \phi_n).$$

It is easy to check that the trace (if it exists) is independent of the basis used.

1.2.3. The Hahn–Banach Theorem.

Theorem 1.47. *Let $\mathcal{B}$ be a Banach space over $\mathbb{C}$ and let $p : \mathcal{B} \to \mathbb{R}$ be such that, for all $x, y \in \mathcal{B}$*

$$p(ax + by) \leq |a|\, p(x) + |a|\, p(y), \text{ whenever } |a| + |b| = 1. \tag{1.1}$$

If $\bar{L}$ is a linear functional defined on a subspace $\mathcal{D} \subset \mathcal{B}$, with $\left|\bar{L}(x)\right| \leq p(x)$, for all $x \in \mathcal{D}$, then $\bar{L}$ can be extended to a linear functional L on $\mathcal{B}$ such that $|L(x)| \leq p(x),\ x \in \mathcal{B}$ and $L(x) = \bar{L}(x)$ on $\mathcal{D}$.

Proof. We first assume that the field is $\mathbb{R}$. Suppose that $x \in \mathcal{B}$ but $x \notin \mathcal{D}$. Let $\mathcal{E} = (x, \mathcal{D})$ be the vector space spanned by x and $\mathcal{D}$. If we have an extension L of $\bar{L}$ from $\mathcal{D}$ to $\mathcal{E}$, it must satisfy

$$L\,(ax + by) = \lambda L(x) + \bar{L}(y),\ y \in \mathcal{D}.$$

and from (1.1), $|a| + |b| = 1$ implies that

$$p(ax + by) \leq |a|\, p(x) + |b|\, p(y).$$

Suppose that $y_1, y_2 \in \mathcal{D},\ a, b > 0,\ a + b = 1.$ Then

$$a\bar{L}(y_1) + b\bar{L}(y_2) = \bar{L}(ay_1 + by_2) \leqslant p[a(y_1 - \tfrac{1}{a}x) + b(y_2 + \tfrac{1}{b}x)]$$
$$\leqslant ap(y_1 - \tfrac{1}{a}x) + bp(y_2 + \tfrac{1}{b}x).$$

We see that for all $y_1, y_2 \in \mathcal{D}$ and all $a, b > 0,\ a + b = 1$, we have

$$\tfrac{1}{a}\left[-p[y_1 - ax) + \bar{L}(y_1)\right] \leqslant \tfrac{1}{b}\left[p(y_2 + bx) - \bar{L}(y_2)\right].$$

It now follows that we must be able to find a number c such that for all $a > 0$,

$$\sup_{y \in D} \tfrac{1}{a}\left[-p[y - ax) + \bar{L}(y)\right] \leqslant c \leqslant \inf_{y \in D} \tfrac{1}{a}\left[p(y + ax) - \bar{L}(y)\right].$$

We can define $L(x) = c$. It is easy to check that $L(x) \leq p(x)$, for all $x \in \mathcal{E}$. We now appeal to Zorn's Lemma (see Yosida [YS, p. 2]), to show that $\bar{L}$ can be extended to all of $\mathcal{B}$, when the field is $\mathbb{R}$.

To extend our result to complex linear functionals, let $\bar{L}$ be given on $\mathcal{D}$ and define $L'(x) = \operatorname{Re}\{\bar{L}(x)\}$, so that it is a real linear functional on $\mathcal{D}$. Since

$$L'(ix) = \operatorname{Re}\{i\bar{L}(x)\} = -\operatorname{Im}\{\bar{L}(x)\},$$

we see that $\bar{L}(x) = L'(x) - iL'(ix)$. Furthermore, since L' is real, it has an extension L, to all of $\mathcal{B}$ such that $L(x) \leq p(x)$, for all $x \in \mathcal{B}$. We can now define $F(x) = L(x) - iL(ix)$. It is easy to check that F is a complex linear functional.

Since $|a| = 1$ implies that $p(ax) = p(x)$, we can set $\theta = \operatorname{Arg}\{F(x)\}$. If we now use the fact that $\operatorname{Re}\{F\} = L$, we have

$$|F(x)| = e^{-i\theta}F(x) = F(e^{-i\theta}x) = L(e^{-i\theta}x) \leqslant p(e^{-i\theta}x) = p(x),$$

we are done. □

Theorem 1.48. *Let M be a linear subspace of $\mathcal{B}$. If $x_0 \in \mathcal{B}$, with $0 < c = d(M, x_0)$, then there exists a bounded linear functional $L(\cdot)$ defined on $\mathcal{B}$ such that*

$$L(x_0) = 1,\ \|L\|_{\mathcal{B}} = \frac{1}{c},\ L(x) = 0, \text{ for all } x \in M.$$

Proof. Let $M_1 = (M, x_0)$ be the subspace spanned by M and x_0. Thus, each point $z \in M_1$ is of the form $z = y + \lambda x_0$, where $y \in M$ and $\lambda \in \mathbb{C}$ are uniquely determined by z. Define $F(\cdot)$ on $\mathcal{B}$ by $F(y + \lambda x_0)$ $= \lambda$. Clearly F is a bounded linear functional, and if $\lambda \neq 0$ then

$$\|y + \lambda x_0\|_{\mathcal{B}} = \left\|\frac{y}{\lambda} + x_0\right\|_{\mathcal{B}} |\lambda| \geqslant c\,|\lambda|\,.$$

It follows that $|F(z)| \leqslant \frac{1}{c}\|z\|_{\mathcal{B}}$, so that $\|F\|_{\mathcal{B}'} \leqslant \frac{1}{c}$. If $\{z_n\} \subset M$, $\|x_0 - z_n\| \to c$, then

$$1 = F(x_0 - z_n) \leqslant \|q\|_{\mathcal{B}}\,\|x_0 - z_n|\,\_\mathcal{B} \to c\,\|F\|_{\mathcal{B}}\,.$$

Thus, $\|F\|_{\mathcal{B}} = \frac{1}{c}$. Thus, by Theorem 1.47 with L replacing F finishes the proof. □

The following is a consequence of the last two results.

Theorem 1.49. *If $\mathcal{B}$ is a Banach space, we have*

(1) *For any $x \in \mathcal{B}$, $x \neq 0$, there exists a linear functional $L \in \mathcal{B}'$ such that $\|L\|_{\mathcal{B}'} = 1$ and $L(x) = \|x\|_{\mathcal{B}}$.*

(2) *If $x \neq y$, there exists a linear functional $L \in \mathcal{B}'$ with $L(x) \neq L(y)$.*

(3) *For* $x \in \mathcal{B}$*,*

$$\|x\|_{\mathcal{B}} = \sup_{L \neq 0} \frac{|L(x)|}{\|L\|_{\mathcal{B}'}} = \sup_{\|L\|=1} |L(x)| .$$

(4) *If* M *is a subspace of* $\mathcal{B}$ *and* $x_0 \in \mathcal{B}$*,* $x_0 \notin \overline{M}$*, then there exists a linear functional* $L \in \mathcal{B}'$ *such that* $L(x_0) = 1$ *and* $L(x) = 0$*, for all* $x \in \overline{M}$*.*

1.2.4. The Baire Category Theorem. In this section we introduce Baire's Theorem and some of its consequences. First we need a definition.

Definition 1.50. Let $\mathcal{B}$ be a Banach space. A subset $E \subset \mathcal{B}$ is said to be nowhere dense if its closure has empty interior. A set is said to be meager (or of the first category) in $\mathcal{B}$ if it is a countable union of nowhere dense sets. A set in $\mathcal{B}$ that is not meager (not of the first category) in $\mathcal{B}$ is said to be nonmeager (of the second category) in $\mathcal{B}$.

Theorem 1.51. (Baire's Theorem) *If* $\mathcal{B}$ *is a Banach space, then the intersection of every countable collection of dense open subsets of* $\mathcal{B}$ *is a dense set in* $\mathcal{B}$*.*

Proof. Let $\{U_1, U_2, U_3, \ldots\}$ be any countable collection of dense open subsets of $\mathcal{B}$. If T_0 is any ball in $\mathcal{B}$ of radius 1, choose a ball T_1 of radius $\frac{1}{2}$ such that the closure of T_1, $\overline{T}_1 \subset U_1 \cap T_0$. (Check that this is possible.) Continue this process, so that at step n, we choose a ball T_n of radius $\frac{1}{n}$ such that $\overline{T}_n \subset U_n \cap T_{n-1}$ and define K by:

$$K = \bigcap_{n=1}^{\infty} \overline{T}_n .$$

It is easy to see that the centers of our nested balls form a Cauchy sequence that converges to a point in K, so that K is nonempty. Since $K \subset T_0$ and $K \subset T_n$ for each n, we see that the intersection of T_0 with $\bigcap_{n=1}^{\infty} U_n$ is nonempty. □

The following two lemmas are required for our proof of the Banach–Steinhaus Theorem in the next section. The second lemma is true for an arbitrary index set, but for our use the restriction of the index set to $\mathbb{R}^+$ is sufficient.

Lemma 1.52. *Let $\mathcal{B}$ be a Banach space. Suppose that $\{V_1, V_2, V_3, \ldots\}$ is a countable collection of closed subsets of $\mathcal{B}$ with* $\operatorname{int}(V_n) = \emptyset$. *Then* $V = \bigcap\limits_{n=1}^{\infty} V_n = \emptyset$.

Proof. Since V is meager and $\operatorname{int}(V) \subset V$, it follows that $\operatorname{int}(V)$ is meager. By Baire's Theorem, we see that $\operatorname{int}(V) = \emptyset$. □

Lemma 1.53. *Let $\mathcal{B}$ be a Banach space. Suppose that $\{f_t\}$, $t \in \mathbb{R}^+$ is a pointwise bounded family of continuous real-valued functions on $\mathcal{B}$. Then the family is uniformly bounded on some nonempty open subset of $\mathcal{B}$.*

Proof. Suppose that $|f_t(\varphi)| \le c_\varphi$ for all $t \in \mathbb{R}^+$ and define

$$V_n^t = \{\varphi \in B \mid |f_t(\varphi)| \le n\}.$$

It is clear that V_n^t is closed in $\mathcal{B}$, since f_t is continuous. Therefore the set:

$$V_n = \bigcap_{n=1}^{\infty} \{\varphi \in B \mid |f_t(\varphi)| \le n\},$$

defined for each n, is also closed in $\mathcal{B}$. Since f_t is pointwise bounded, we have that $\mathcal{B} = \bigcup_{n=1}^{\infty} V_n$. If $\operatorname{int}(V_m) = \emptyset$ for all m, then from Lemma 1.52, $\bigcup_{n=1}^{\infty} V_n$ is meager. Since $\mathcal{B}$ is of the second category, this is a contradiction. Therefore, $\operatorname{int}(V_m) \neq \emptyset$ for some m. If we set $M = m$ and $U = \operatorname{int}(V_m)$, it follows that $\{f(t)\}$, $t \in \mathbb{R}^+$ is uniformly bounded on U. □

1.2.5. The Banach–Steinhaus Theorem. The next important result, known in the early literature as the Banach–Steinhaus Theorem, is much better known now as the principle of uniform boundedness.

Theorem 1.54 (Uniform Boundedness Theorem)**.** *Let $\{T(t)\}$ be a family of continuous mappings on the Banach space $\mathcal{B}$ for $t \in \mathbb{R}^+$. If for each $\varphi \in \mathcal{B}$, the family $\{\|T(t)\varphi\|_{\mathcal{B}}\}$ is bounded for all $t \in \mathbb{R}^+$, then the $\{\|T(t)\|_{\mathcal{B}}\}$ is a bounded family.*

Proof. For each $t \in \mathbb{R}^+$ define $f_t : \mathcal{B} \to \mathbb{R}^+$ by $f_t(\varphi) = \|T(t)\varphi\|_{\mathcal{B}}$. Since the norm is continuous, we see that f_t is also continuous. From Lemma 1.53, there is a nonempty open set $V_{n_0} \subset \mathcal{B}$ and $f_t(\varphi) \le n_0$ for all $t \in \mathbb{R}^+$ and all $\varphi \in V_{n_0}$. Without loss of generality, we can assume that $U = \{\varphi \mid \|\varphi\|_{\mathcal{B}} < r\} \subset V_{n_0}$ for some $r > 0$. It follows that, for

$\varphi_0 \in U$, $\|f_t(\varphi + r\phi)\|_{\mathcal{B}} \leqslant n_0$ for all $t \in \mathbb{R}^+$ and all ϕ with $\|\phi\|_{\mathcal{B}} < 1$. This implies that

$$r\|T\|_{\mathcal{B}} = r \sup_{\|\phi\|_{\mathcal{B}} \leqslant 1} \|T(t)\phi\|_{\mathcal{B}} = \sup_{\|\phi\|_{\mathcal{B}} \leqslant 1} \|T(t)\left[\varphi_0 + r\phi\right] - T(t)\varphi_0\|_{\mathcal{B}}$$
$$\leqslant n_0 + \|T(t)\varphi_0\|_{\mathcal{B}} < \infty.$$

□

The next result (the open mapping theorem) is one of the important theorems in functional analysis. It is used to prove the two theorems that follow. The first of the two will be used in the next section, while the second is fundamental for Chaps. 4 and 5.

Theorem 1.55 (Open Mapping Theorem)**.** *Let $\mathcal{B}_1, \mathcal{B}_2$ be two Banach spaces and let A be a continuous linear surjective mapping of $\mathcal{B}_1 \to \mathcal{B}_2$. Then, whenever U is an open set in $\mathcal{B}_1$, $A[U]$ is an open set in $\mathcal{B}_2$.*

Proof. It suffices to show that, for every open ball U about zero in $\mathcal{B}_1$, $A[U]$ contains an open ball about zero in $\mathcal{B}_2$. Hence, fix U and let $\{U_0, U_1, U_2, \ldots\}$ be a sequence of open balls of radius $r/2^n$, $(n = 0,\ 1,\ 2, \ldots)$, where r is chosen so that $U_0 \subset U$. We are done if we can prove that there is an open set W such that:

$$W \subset \overline{A(U_1)} \subset A(\mathbf{U}),$$

where $\overline{A(U_1)}$ is the closure of $A(U_1)$. Since $U_2 - U_2 \subset U_1$, we first need to prove that $W \subset \overline{A(U_1)}$. To do this, note that:

$$\overline{A(U_1)} \supset \overline{A(U_2) - A(U_2)} \supset \overline{A(U_2)} - \overline{A(U_2)}.$$

We will be done with this part of the proof if we show that the interior of $\overline{A(U_2)}$ is nonempty. But

$$A(\mathcal{B}_1) = \bigcup_{m=1}^{\infty} mA(U_2),$$

since U_2 is a ball centered at zero and A is a surjection. Therefore, at least one of the $mA(U_2)$ is of the second category in $\mathcal{B}_2$. But, as the mapping $\varphi \to m\varphi$ is a homeomorphism of $\mathcal{B}_2$ onto $\mathcal{B}_2$, $\overline{A(U_2)}$ is nonmeager in $\mathcal{B}_2$. Therefore, there exists an open set $W \subset \overline{A(U_2)}$.

To prove that $\overline{A(U_1)} \subset A(U)$, let $\varphi_1 \in \overline{A(U_1)}$ be fixed, and observe by the first part that

$$\left(\varphi_1 - \overline{A(U_2)}\right) \bigcap A(U) \neq \emptyset.$$

Thus, there is a $y_1 \in U_1$ with $A(y_1) \in \varphi_1 - \overline{A(U_2)}$. Now, for any $n \geq 1$, $\overline{A(U_n)}$ contains an open neighborhood of zero. Hence, assume that $\varphi_n \in \overline{A(U_n)}$ has been chosen with

$$\left(\varphi_n - \overline{A(U_{n+1})}\right) \bigcap A(U_n) \neq \emptyset.$$

This means there is a $y_n \in U_n$ such that $A(y_n) \in \varphi_n - \overline{A(U_{n+1})}$. Set $y_{n+1} = y_n - A(\varphi_n)$. Then $y_{n+1} \in \overline{A(U_{n+1})}$ and we continue the construction. It is easy to see that the sums $\varphi_1 + \varphi_2 + \varphi_3 + \cdots + \varphi_n$ form a convergent Cauchy sequence which converges to some $\varphi \in \mathcal{B}$, and $\|\varphi\| < r$. It follows that $\varphi \in U$ and, as

$$\sum_{n=1}^{m} A(\varphi_n) = \sum_{n=1}^{m} (y_n - y_{n+1}) = y_1 - y_{m+1},$$

we see that $y_{m+1} \to 0$ since A is continuous. Thus, $y_1 = A(\varphi_1) \in A(U)$. Since φ_1 was arbitrary, we see that $\overline{A(U_1)} \subset A(U)$. □

Theorem 1.56 (Inverse Mapping Theorem)**.** *Let $\mathcal{B}_1, \mathcal{B}_2$ be two Banach spaces and let A be a continuous bijective linear mapping of $\mathcal{B}_1 \to \mathcal{B}_2$. Then $A^{-1} : \mathcal{B}_2 \to \mathcal{B}_1$ is continuous.*

Proof. Since A is continuous, injective, and surjective, it is an open mapping. As A^{-1} exists and, since $A^{-1}\{A[\mathcal{O}]\} = \mathcal{O}$ for all open sets, A^{-1} is continuous. □

Theorem 1.57 (Closed Graph Theorem)**.** *Let $\mathcal{B}$ be a Banach space and let A a closed linear operator on $\mathcal{B}$. If $D(A) = \mathcal{B}$, then $A \in L[\mathcal{B}]$.*

Proof. By definition, $G(A)$ is closed and is a Banach space in the norm $\|(\varphi, A\varphi)\| = \|\varphi\|_{\mathcal{B}} + \|A\varphi\|_{\mathcal{B}}$. Consider the two continuous mappings: $\pi_1 : (\varphi, A\varphi) \to \varphi$, $\pi_2 : (\varphi, A\varphi) \to A\varphi$. Since π_1 is a bijection, π_1^{-1} is continuous so $A = \pi_2 \circ \pi_1^{-1}$ is also continuous. □

Part II: Intermediate Analysis

In this second part of this chapter, we introduce a number of topics that are rarely covered in the first 2 years of the standard graduate programs. These topics will be used at a number of points in the book and are collected here for reference as needed.

S-Basis. In this section, we review a few results that belong to Banach space theory proper. We provide a few proofs, but all the results can be found in Carothers [CA]. Let $\mathcal{B}$ be a separable Banach space.

Definition 1.58. A sequence $(x_n) \in \mathcal{B}$ is called a Schauder basis (S-basis) for $\mathcal{B}$ if $\|x_n\|_{\mathcal{B}} = 1$ for all n and, for each $f \in \mathcal{B}$, there is a unique sequence (a_n) of scalars such that

$$x = \lim_{k \to \infty} \sum_{n=1}^{k} a_n x_n = \sum_{n=1}^{\infty} a_n x_n.$$

All spaces of interest in this book have an S-basis. However, it is known that there are separable Banach spaces without an S-basis (see Carothers [CA] or Diestel [DI]).

Example 1.59. *Let $\mathcal{B} = \ell_p,\ 1 < p < \infty$, where*

$$\ell_p = \left\{ x = (x_1, \cdots) : \sum_{k=1}^{\infty} |x_k|^p < \infty \right\}.$$

The set of vectors $\{e_k\}$, where

$$e_k = \left(0, 0, \ldots, \overset{k}{1}, 0, \ldots \right),$$

form a norm-one S-basis for this space (see [CA]).

If $\Omega = [0,1]$ and $\mathcal{B} = L^p[\Omega],\ 1 < p < \infty$, the family of vectors

$$\{1, \cos(2\pi t), \sin(2\pi t)\cos(4\pi t), \sin(4\pi t), \ldots\}$$

is a norm-one S-basis for $\mathcal{B}$ (see [CA]).

It is easy to see that every Banach space with an S-basis is separable. Let $\mathcal{P}_n x = \sum_{k=1}^{n} a_k x_k$ and define a new norm on $\mathcal{B}$ by

$$|||x|||_{\mathcal{B}} = \sup_n \|\mathcal{P}_n x\|_{\mathcal{B}} = \sup_n \left\| \sum_{k=1}^{n} a_k x_k \right\|_{\mathcal{B}}.$$

Example 1.60. *Let $\Omega = [0,1]$ and $\mathcal{B} = L^p[\Omega]$ over the complex numbers. If $x(t) \in \mathcal{B}$, define*

$$c_k = \int_0^1 e^{-2\pi i k t} x(t)dt, \quad k = 0, \pm 1, \pm 2, \ldots$$

It is easy to see that,

$$|||x|||_{\mathcal{B}} = \sup_n \int_0^1 \left| \sum_{k=-n}^{n} c_k e^{2\pi i t} \right|^2 dt$$

defines a norm on $\mathcal{B}$.

Theorem 1.61. *The norm $|||\cdot|||_{\mathcal{B}}$, is an equivalent norm on $\mathcal{B}$ and*

$$\|x\|_{\mathcal{B}} \le |||x|||_{\mathcal{B}} = \sup_n \|\mathcal{P}_n x\|_{\mathcal{B}}.$$

Proof. Since $\lim_n \mathcal{P}_n x \to x$, it is clear that $|||x|||_{\mathcal{B}}$ is finite for all $x \in \mathcal{B}$. Since the identity map of $(\mathcal{B}, |||\cdot|||_{\mathcal{B}}) \to (\mathcal{B}, ||\|\cdot\||_{\mathcal{B}})$ is continuous, by the Inverse Mapping Theorem (Theorem 1.56), we are done if we can show that this map has a continuous inverse. It suffices to show that $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$ is complete (i.e., a Banach space).

For this, suppose that let (z_k) be a Cauchy sequence in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$. Then $(\mathcal{P}_n z_k)$ is a Cauchy sequence in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$, since $\|\mathcal{P}_n z_i - \mathcal{P}_n z_j\|_{\mathcal{B}} \leqslant |||z_i - z_j|||_{\mathcal{B}}$, for all n (uniformly Cauchy). Thus, if $y_n = \lim_{k\to\infty} \mathcal{P}_n z_k$ then $\lim_{k\to\infty} \|\mathcal{P}_n z_k - y_n\|_{\mathcal{B}} = 0$, uniformly in n.

It now follows using the standard $\frac{\varepsilon}{3}$ argument that (y_n) is a Cauchy sequence in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$. If $y = \lim_{n\to\infty} y_n$ in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$, we are done if we show that $y = \lim_{k\to\infty} z_k$ in $(\mathcal{B}, |||\cdot|||_{\mathcal{B}})$.

Since there is a unique sequence of scalars (a_i) such that $y = \sum_{i=1}^{\infty} a_i x_i$, we see that $y_n = \sum_{i=1}^{n} a_i x_i$, so that $\mathcal{P}_n y = y_n$ and

$$|||z_k - y|||_{\mathcal{B}} = \sup_n \|\mathcal{P}_n z_k - y_n\|_{\mathcal{B}} \to 0 \text{ as } k \to \infty.$$

□

Since, $x \in \mathcal{B}$ implies that $\|\mathcal{P}_n x\|_{\mathcal{B}} < \infty$ for all n, By the Uniform Boundedness Theorem 1.54, we see that $\sup_n \|\mathcal{P}_n\|_{\mathcal{B}} < \infty$.

Definition 1.62. The set $\{\mathcal{P}_n\}$ is called the natural family of projections associated with the S-basis $\{x_n\}$ and $\sup_n \|\mathcal{P}_n\|_{\mathcal{B}} = K$ is called the basis constant of $\{x_n\}$. In terms of the equivalent norm of the last theorem, $K = 1$.

Definition 1.63. Let $\mathcal{B}$ be a Banach space with an S-basis and let x_n^* be the linear functional on $\mathcal{B}$ defined by $x_n^*(x) = a_n$, where $x = \sum_{k=1}^{\infty} a_k x_k$. Since $x_n^*(x_m) = \delta_{mn}$, we say that the sequence of pairs $\{x_n^*, x_n\}$ are biorthogonal. We call the family $\{x_n^*\}$ the coordinate functionals.

Definition 1.64. We define the span of a set of vectors $\{x_n\}$, in a vector space $\mathcal{B}$, written $\mathrm{span}(\{x_n\})$, to be the set of all finite linear combinations of subsets of $\{x_n\}$. When $\mathcal{B}$ is a Banach space, we let $[x_n]$ represent the closed subspace of $\mathcal{B}$ generated by $\mathrm{span}(\{x_n\})$.

A proof of the next result can be found in Carothers (see [CA, pp. 67–71]).

Theorem 1.65. *If $\mathcal{B}$ is a reflexive Banach space and $\{x_n\}$ is an* S-*basis for $\mathcal{B}$, then $\{x_n^*\}$ is an S-basis for $\mathcal{B}'$. Furthermore, the natural embedding $j : \mathcal{B} \to \mathcal{B}''$ defined by $x^{**}(y^*) = y^*(x)$ for all $y^* \in \mathcal{B}'$, is an isometric isomorphism.*

1.3. Distributions and Sobolev Spaces

References for this section are Strichartz [SZ], Yosida [YS], Leoni [GL], Reed and Simon [RS], Rudin [RU1], and Evans [EV]. The purpose of this section is to establish the basic ideas for use in Chaps. 2 and 3. However, neither this section nor the material in Chaps. 2 and 3 is a substitute for a complete introduction to the subject. Those with no background should at least consult Strichartz [SZ].

1.3.1. The Test Functions and Distributions.

Definition 1.66. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ be a multi-index of non-negative integers, with $|\alpha| = \sum_{k=1}^n \alpha_k$. We define the operators D_n^α and $D_{\alpha,n}$ by

$$D_n^\alpha = \prod_{k=1}^n \frac{\partial^{\alpha_k}}{\partial x^{\alpha_k}} \quad D_{\alpha,n} = \prod_{k=1}^n \left(\frac{1}{2\pi i}\frac{\partial}{\partial x_k}\right)^{\alpha_k}$$

Let $\mathbb{C}_c(\mathbb{R}^n)$ be the class of infinitely differentiable functions on $\mathbb{R}^n$ with compact support and impose the natural locally convex topology τ on $\mathbb{C}_c(\mathbb{R}^n)$ to obtain $\mathfrak{D}(\mathbb{R}^n)$. A definition in terms of neighborhoods can be found in Leoni [GL] (see also Yosida [YS] and Reed and Simon [RS]).

Definition 1.67. A sequence $\{f_m\}$ converges to $f \in \mathfrak{D}(\mathbb{R}^n)$ with respect to the compact sequential limit topology if and only if there exists a compact set $K \subset \mathbb{R}^n$, which contains the support of $f_m - f$ for each m and $D_n^\alpha f_m \to D_n^\alpha f$ uniformly on K, for every multi-index $\alpha \in \mathbb{N}^n$.

Let $u \in \mathbb{C}^1(\mathbb{R}^n)$ and suppose that $\phi \in \mathbb{C}_c^\infty(\mathbb{R}^n)$ has its support in a ball B_r, of radius $r > 0$. Integration by parts gives:

$$\int_{\mathbb{R}^n} (\phi u_{y_i}) d\lambda_n = \int_{\partial B_r} (u\phi)\nu_i d\mathbf{S} - \int_{\mathbb{R}^n} (u\phi_{y_i}) d\lambda_n,\ 1 \le i \le n,$$

where $\boldsymbol{\nu}$ is the unit outward normal to B_r. Since ϕ vanishes on the ∂B_r, the above reduces to:

$$\int_{\mathbb{R}^n} (\phi u_{y_i})\ d\lambda_n = -\int_{\mathbb{R}^n} (u\phi_{y_i}) d\lambda_n,\ 1 \le i \le n.$$

In the general case, for any $u \in \mathbb{C}^m[\mathbb{R}^n]$ and any multi-index $\alpha = (\alpha_1, \dots, \alpha_n),\ |\alpha| = \sum_{\alpha=1}^n \alpha_i = m$, we have

$$\int_{\mathbb{R}^n} \phi(D^\alpha u) d\lambda_n = (-1)^m \int_{\mathbb{R}^n} u(D^\alpha \phi) d\lambda_n. \tag{1.2}$$

We now observe that the right-hand side of Eq. (1.2) makes sense, even if $D^\alpha u$ does not exist according to our normal definition. This is the basic idea behind the notion of a distributional derivative. Before giving the formal definition, recall that a function $u \in L^1_{\text{loc}}[\mathbb{R}^n]$ if it is Lebesgue integrable on every compact subset of $\mathbb{R}^n$.

Definition 1.68. If α is a multi-index and $u,\ v \in L^1_{\text{loc}}[\mathbb{R}^n]$, we say that v is the α^{th}-weak (or distributional) partial derivative of u and write $D^\alpha u = v$ provided that

$$\int_{\mathbb{R}^n} u(D^\alpha \phi) d\lambda_n = (-1)^{|\alpha|} \int_{\mathbb{R}^n} \phi v\ d\lambda_n$$

for all functions $\phi \in \mathbb{C}_c^\infty[\mathbb{R}^n]$. Thus, v is in the dual space $\mathcal{D}'[\mathbb{R}^n]$ of $\mathcal{D}[\mathbb{R}^n]$.

The next result is easy.

Lemma 1.69. *If a weak αth-partial derivative exists for u, then it is unique λ_n-(a.e.).*

Definition 1.70. If $m \ge 0$ is fixed and $1 \le p \le \infty$, we define the Sobolev space $W^{m,p}[\mathbb{R}^n]$ to be the set of all locally integrable functions $u:\ \mathbb{R}^n \to \mathbb{R}$ such that, for each multi-index α with $|\alpha| \leqslant m$, $D^\alpha u$ exists in the weak sense and belongs to $L^p[\mathbb{R}^n]$.

Extensions and Decompositions. We need an extension theorem for functions defined on a domain of $\mathbb{R}^n$ and a result which shows that a domain in $\mathbb{R}^n$ can be written as a union of nonoverlapping closed cubes. (Proofs of these results can be found in Evans [EV] and Stein [STE], respectively.)

Let $\mathbb{D}$ be a bounded open connected set of $\mathbb{R}^n$ (a domain) with boundary $\partial\mathbb{D}$ and closure $\overline{\mathbb{D}}$.

Definition 1.71. Let k be a positive integer. We say that $\partial\mathbb{D}$ is of class $\mathbf{C}^k$ if, for every point $\mathbf{x} \in \partial\mathbb{D}$, there is a homeomorphism ϕ of a neighborhood U of $\mathbf{x}$ into $\mathbb{R}^n$ such that both ϕ and ϕ^{-1} have k continuous derivatives with

$$\varphi(\mathbb{D} \cap U) \subset \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n > 0\}$$

and

$$\varphi(\partial\mathbb{D} \cap U) \subset \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n = 0\}.$$

Theorem 1.72. *Let $\mathbb{D}$ be a domain in $\mathbb{R}^n$ with $\partial\mathbb{D}$ of class $\mathbb{C}^1$. Let $\mathbb{U}$ be any bounded open set such that $\overline{\mathbb{D}}$, the closure of $\mathbb{D} \subset\subset \mathbb{U}$ (i.e., the closure of $\mathbb{D}$ is a compact subset of $\mathbb{U}$). Then there is a linear operator $\mathfrak{E}$ mapping functions on $\mathbb{D}$ to functions on $\mathbb{R}^n$ such that:*

1. *The operator $\mathfrak{E}$ maps $W^{1,p}[\mathbb{D}]$ continuously into $W^{1,p}[\mathbb{R}^n]$ for all $1 \le p \le \infty$.*
2. *$\mathfrak{E}(f)\,|_{\mathbb{D}} = f$ (e.g., $\mathfrak{E}(\cdot)$ is an extension operator).*
3. *$\mathfrak{E}(f)(x) = 0$ for $x \in \mathbb{U}^c$ (e.g., $\mathfrak{E}(f)$ has support inside $\mathbb{U}$).*

Theorem 1.73. *Let $\mathbb{D}$ be a domain in $\mathbb{R}^n$. Then $\mathbb{D}$ is the union of a sequence of closed cubes $\{\mathbb{D}_k\}$ whose sides are parallel to the coordinate axes and whose interiors are mutually disjoint.*

Thus, if a function f is defined on a domain in $\mathbb{R}^n$, by Theorem 1.72 it can be extended to the whole space. On the other hand, without loss of generality, by Theorem 1.73, we can assume that the domain is a cube with sides parallel to the coordinate axes.

Definition 1.74. If $\mathbb{D}$ is a domain in $\mathbb{R}^n$, we define $W_0^{m,p}[\mathbb{D}]$ to be the closure of $\mathbb{C}_c^\infty(\mathbb{D})$ in $W^{m,p}[\mathbb{D}]$.

Remark 1.75. Thus, $W_0^{m,p}[\mathbb{D}]$ contains those functions $u \in W^{m,p}[\mathbb{D}]$ such that, for all $|\alpha| \le m-1$, $D^\alpha u = 0$ on the boundary of $\mathbb{D}$, $\partial\mathbb{D}$.

We also note that, when $p = 2$ it is standard to use $H^m(\mathbb{D}) = W^{m,2}(\mathbb{D})$ and $H_0^m(\mathbb{D}) = W_0^{m,2}(\mathbb{D})$.

1.4. Tensor Products

Tensor products of Banach spaces are not a part of the normal graduate program. This section is an introduction to the finite theory that is background for the infinite tensor product theory in Chap. 6. At this point, it is assumed that the reader has at least studied Chap. 4 or is already familiar with the material from some other source.

Since tensor products of Banach spaces have a bad reputation, we should at least comment on this "public relations problem." This reputation is due to questions and studies unrelated to partial differential equations, path integrals, stochastic processes, analysis (proper), and the many possible applications in science and engineering. We approach the subject from a more natural point of view, so that its usefulness for these important and equally interesting areas will be transparent.

1.4.1. Elementary Background. For those with no background in tensor products, we begin with $\mathbb{R}^3$, a space which is well known from calculus (any finite dimension will do). There are a number of ways to patch together two copies of $\mathbb{R}^3$ to obtain a new space. The first is called the direct sum:

$$\mathbb{R}^3 \oplus \mathbb{R}^3 = \{(a_1, a_2, a_3, b_1, b_2, b_3) : (a_1, a_2, a_3), (b_1, b_2, b_3) \in \mathbb{R}^3\}.$$

It is clear that $\mathbb{R}^3 \oplus \mathbb{R}^3$ is isomorphic to $\mathbb{R}^6$. There are also two ways we can define a product on $\mathbb{R}^3$; the first is the dot product

$$\begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \sum_{i=1}^{3} b_i a_i,$$

which takes two vectors and produces a scalar. The other is the tensor product

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix},$$

which takes two vectors and produces a 3×3 matrix. It is easy to see that we can write the resulting matrix as a vector in $\mathbb{R}^9 = \mathbb{R}^{3^2}$. Thus, the tensor product of $\mathbb{R}^3$ with itself, written as $\mathbb{R}^3 \otimes \mathbb{R}^3$, is isomorphic to $\mathbb{R}^9$.

Implicit in our use of the dot product is the assumption that the norm induced is the natural one generated by the dot product on $\mathbb{R}^3$:

$$\|\mathbf{a}-\mathbf{b}\|_{\mathbb{R}^3} = \sqrt{(\mathbf{a}-\mathbf{b})\cdot(\mathbf{a}-\mathbf{b})} = \sqrt{\sum_{i=1}^{3}(a_i-b_i)^2}.$$

However, there are a number of other norms possible on $\mathbb{R}^3$ which are not induced by a dot product. For example:

$$\begin{aligned} \|\mathbf{a}-\mathbf{b}\|_p &= \left[\sum_{i=1}^{3}|a_i-b_i|^p\right]^{1/p}, \quad 1 \leqslant p < \infty,\ p \neq 2, \\ \|\mathbf{a}-\mathbf{b}\|_\infty &= \max_{1\leqslant i\leqslant 3}|a_i-b_i|. \end{aligned} \tag{1.3}$$

We will discuss this later. However, the case $p = 2$ is the standard one because it is the only (unique) one generated by a dot product (even in infinite dimensions). Let $\ell_p\left(\mathbb{R}^3\right)$ represent $\mathbb{R}^3$ with the norm $\|\cdot\|_p$. It is easy to check that $\ell_p\left(\mathbb{R}^3\right)$ is a Banach space for $p \neq 2$ and that $\ell_2\left(\mathbb{R}^3\right)$ is a Hilbert space.

We know that $\mathbb{R}^3 \otimes \mathbb{R}^3 = \mathbb{R}^9$. The basic question is, how do we define the norm, so that $\ell_p(\mathbb{R}^3) \otimes \ell_p(\mathbb{R}^3) = \ell_p(\mathbb{R}^9)$. It is known that, on $\mathbb{R}^n$, all norms are equivalent. That is, for any pair p, q, there exists constants $c_{p,q}, C_{p,q}$, such that, for any vector $\mathbf{a}$,

$$c_{p,q}\|\mathbf{a}\|_q \leqslant \|\mathbf{a}\|_p \leqslant C_{p,q}\|\mathbf{a}\|_q.$$

Thus, we can define the dot product on $\mathbb{R}^9$ and use the norm equivalence to obtain all the others. However, in the infinite-dimensional case $(\ell_p(\mathbb{R}^\infty))$, this is no longer true and each of the norms in Eq. (1.3) generates distinct Banach spaces.

Example 1.76. *Let us see what the norm looks like for $\mathbb{R}^2 \otimes \mathbb{R}^2$. A direct computation shows that*

$$\begin{aligned} \mathbf{a}\otimes\mathbf{b} &= \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix} \\ &= \begin{bmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_1 \end{bmatrix} = \left[a_1b_1, a_1b_2, a_2b_1, a_2b_1\right]. \end{aligned}$$

and

$$\begin{aligned}
&\|\mathbf{a}\|_2\|\mathbf{b}\|_2\\
&= \left[\sum_{i=1}^{2} a_i^2\right]^{1/2}\left[\sum_{k=1}^{2} b_k^2\right]^{1/2} = \left[\left(\sum_{i=1}^{2} a_i^2\right)\left(\sum_{k=1}^{2} b_k^2\right)\right]^{1/2} \qquad (1.4)\\
&= \left[a_1^2b_1^2 + a_1^2b_2^2 + a_2^2b_1^2 + a_2^2b_2^2\right]^{1/2} = \|\mathbf{a}\otimes\mathbf{b}\|_4.
\end{aligned}$$

This is a special property called the crossnorm (i.e., $\|\mathbf{a}\otimes\mathbf{b}\|_4 = \|\mathbf{a}\|_2\|\mathbf{b}\|_2$*).*

1.4.2. General Background. We begin with a few concrete examples and ideas that reveal the landscape. Let $f(x) \in \mathbb{C}(\Omega_1)$, $g(y) \in \mathbb{C}(\Omega_2)$, where Ω_1 and Ω_2 are compact sets in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. If we let $F(x,y) = f(x)g(y)$, then $F(x,y) \in \mathbb{C}(\Omega_1 \times \Omega_2)$. It is clear that:

(1)

$$\frac{\partial^2}{\partial x \partial y}F(x,y) = \left[\frac{d}{dx}f(x)\right]\left[\frac{d}{dy}g(y)\right]$$

and

(2)

$$\int\limits_{\Omega_1\times\Omega_2} F(x,y)dxdy = \int\limits_{\Omega_1} f(x)dx \int\limits_{\Omega_2} g(y)dy.$$

If we let S be the set of all finite sums, $S = \{F_m(x,y)\},\ m \in \mathbb{N}$, where

$$F_m(x,y) = \sum\nolimits_{i=1}^{m} f_i(x)g_i(y), \quad m \in \mathbb{N},$$

it is easy to see that S is dense in $\mathbb{C}(\Omega_1 \times \Omega_2)$. We can now ask the natural question; What norm should we use on S so that the completion of S, $\bar{S} = \mathbb{C}(\Omega_1 \times \Omega_2)$? It is not hard to show that the appropriate norm is

$$\|F_m\|_\infty = \sup_{x\in\Omega_1}\sup_{y\in\Omega_2}\left|\sum_{i=1}^{m} f_i(x)g_i(y)\right|. \qquad (1.5)$$

On the other hand, we could replace Eq. (1.5) with

$$\|F_m\|_p = \left[\int_{\Omega_2}\int_{\Omega_1}\left|\sum_{i=1}^{m} f_i(x)g_i(y)\right|^p dxdy\right]^{1/p},$$

where $1 \le p \le \infty$ and ask the same question. If we use this norm on S, we clearly do not expect to get $\mathbb{C}(\Omega_1 \times \Omega_2)$. The theory of tensor products of Banach spaces is designed to make the above precise and reveal the nature of the resulting space.

1.4.3. Tensor Products of Hilbert Spaces. Tensor products of Hilbert spaces are the easiest Banach spaces to study, because as noted earlier, there exists only one norm that will make the result another Hilbert space.

Let $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ be three Hilbert spaces over $\mathbb{C}$. A mapping $T : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_3$ is said to be bilinear if for all $a, b \in \mathbb{C}$ and all $x \in \mathcal{H}_1,\ y \in \mathcal{H}_2$,

(1)
$$T(ax_1 + bx_2, y) = aT(x_1, y) + bT(x_2, y)$$
and

(2)
$$T(x, ay_1 + by_2) = a^cT(x, y_1) + b^cT(x, y_2).$$

Definition 1.77. The complete tensor product of the Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ is a Hilbert space $\mathcal{H}_3$ and a bilinear mapping $T : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_3$, such that

(1) The closed linear span of all the vectors $T(x, y),\ x \in \mathcal{H}_1,\ y \in \mathcal{H}_2$ is equal to $\mathcal{H}_3$.
(2) The inner product for $\mathcal{H}_3$ satisfies:
$$(T(x_1, y_1), T(x_2, y_2))_3 = (x_1, x_2)_1 (y_1, y_2)_2, \tag{1.6}$$

for all pairs $x_1, x_2 \in \mathcal{H}_1$ and $y_1, y_2 \in \mathcal{H}_2$. We call the pair $(\mathcal{H}_3, T)$ the (complete) tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$. We denote the linear span of the Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ by $\mathcal{H}_1 \otimes \mathcal{H}_2$, replace $\mathcal{H}_3$ with $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$ and $T(x, y)$ by $x \otimes y$, which is the standard representation. If we let $x_1 = x_2,\ y_1 = y_2$ in Eq. (1.6), it now reads $\|x \otimes y\|_3^2 = \|x\|_1^2 \|y\|_2^2$ or $\|x \otimes y\|_3 = \|x\|_1 \|y\|_2$. This is the crossnorm relationship we saw in Eq. (1.4).

The tensor product $x \otimes y$ is a bilinear mapping of $\mathcal{H}_1 \times \mathcal{H}_2$ to $\mathcal{H}_1 \hat{\otimes} \mathcal{H}_2$, we can also view it as a functional in the space $B(\mathcal{H}_1, \mathcal{H}_2, \mathbb{C}) = B(\mathcal{H}_1, \mathcal{H}_2)$, of bilinear mappings on $\mathcal{H}_1 \times \mathcal{H}_2$ to $\mathbb{C}$. That is, from Eq. (1.6),
$$(x_1 \otimes y_1, x_2 \otimes y_2)_3 = (x_1, x_2)_1 (y_1, y_2)_2.$$

We will use this interpretation later in the section to define the tensor product of two Banach spaces.

Theorem 1.78. *If the family* $\{\phi_n\}$ *is a orthonormal basis for* $\mathcal{H}_1$ *and the family* $\{\psi_m\}$ *is a orthonormal basis for* $\mathcal{H}_2$*, the family* $\{\phi_n \otimes \psi_m\}$ *is a orthonormal basis for* $\mathcal{H}_3$*.*

Proof. Since $\|\phi_n \otimes \psi_m\|_3 = \|\phi_n\|_1 \|\psi_m\|_2 = 1$, they are normal. Furthermore,

$$(\phi_n \otimes \psi_m, \phi_i \otimes \psi_j)_3 = (\phi_n, \phi_i)_1 (\psi_m, \psi_j)_2 = \delta_{n,i}\delta_{m,j},$$

so they are also orthogonal.

Thus, we are done if we can show that the closure of the linear span of $\{\phi_n \otimes \psi_m\}$ is all of $\mathcal{H}_3$. Let $f \otimes g \in \mathcal{H}_3$, so that $f \in \mathcal{H}_1$ and $g \in \mathcal{H}_2$. Then there exist unique families of constants $\{a_i\}, \{b_j\}$ such that: $f = \sum_{i=1}^{\infty} a_i \phi_i$ and $g = \sum_{j=1}^{\infty} b_j \psi_j$. It is now easy to show that $f \otimes g = \sum_{i,j}^{\infty} a_i b_j \phi_i \otimes \psi_j$. A little reflection will convince the reader that every vector $u \in \mathcal{H}_3$ can be written as $u = \sum_{i,j}^{\infty} a_{ij} \phi_i \otimes \psi_j$, for some (unique) constants $\{a_{ij}\}$, so that the family $\{\phi_i \otimes \psi_j\}$ is a basis for $\mathcal{H}_3$. □

1.5. Tensor Products of Banach Spaces

If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two Banach spaces, the algebraic tensor product of $\mathcal{B}_1$ and $\mathcal{B}_2$ is denoted by $\mathcal{B}_1 \otimes \mathcal{B}_2$, and every element $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$ may be written as $\phi = \sum_{i=1}^{n} \phi_1^i \otimes \phi_2^i$, where $\{\phi_1^i\} \in \mathcal{B}_1, \{\phi_2^i\} \in \mathcal{B}_2$ and n is some nonnegative integer. We denote by $B(\mathcal{B}_1, \mathcal{B}_2) = B(\mathcal{B}_1, \mathcal{B}_2, \mathbb{C})$, the space of all continuous bilinear functionals on $\mathcal{B}_1 \times \mathcal{B}_2$. If l is a bilinear form on $\mathcal{B}_1 \times \mathcal{B}_2$, it generates a natural linear functional $\hat{l}$ on $\mathcal{B}_1 \otimes \mathcal{B}_2$ defined by evaluation:

$$\left\langle \varphi \otimes \psi, \hat{l} \right\rangle = l(\varphi, \psi),\ (\varphi, \psi) \in \mathcal{B}_1 \times \mathcal{B}_2,\ l \in B\left(\mathcal{B}_1, \mathcal{B}_2\right).$$

Also, $\langle \mathcal{B}_1 \otimes \mathcal{B}_2, \mathcal{B}'_1 \otimes \mathcal{B}'_2 \rangle$ defines a (strong) dual system by:

$$\langle \varphi \otimes \psi, \varphi^* \otimes \psi^* \rangle = \langle \varphi, \varphi^* \rangle \langle \psi, \psi^* \rangle,\ (\varphi, \psi) \in \mathcal{B}_1 \times \mathcal{B}_2,\ (\varphi^*, \psi^*) \in \mathcal{B}'_1 \times \mathcal{B}'_2.$$

It follows that we can consider $\mathcal{B}_1 \otimes \mathcal{B}_2$ as the space of continuous bilinear functionals on $\mathcal{B}'_1 \times \mathcal{B}'_2$ ($\mathcal{B}_1 \otimes \mathcal{B}_2 \subset B\left(\mathcal{B}'_1, \mathcal{B}'_2\right)$), and $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ as the space of continuous bilinear functionals on $\mathcal{B}_1 \times \mathcal{B}_2$, so that $\mathcal{B}'_1 \otimes \mathcal{B}'_2 \subset B\left(\mathcal{B}_1, \mathcal{B}_2\right)$.

For notation consistent with the field, when studying one of the L^p-type spaces (with $1 \le p \le \infty$), we will use $\Delta_p(\cdot)$ in place of $\|\cdot\|_p$. Although there are many norms that may be defined on $\mathcal{B}_1 \otimes \mathcal{B}_2$ such that the completion is a Banach space, we will always use the one that is natural for the spaces under consideration. This means that we will restrict our attention to spaces of direct interest for analysis, applied mathematics, mathematical physics, and probability theory. (Those with interest in the general theory and other approaches should consult the nice books by Defant and Floret [DOF] and Ryan [RA], along with the references therein.)

Let A_i, $i = 1, 2$ be closed linear operators with domains $D_i \subset \mathcal{B}_i$, $A_i : D_i \subset \mathcal{B}_i \to \mathcal{B}_i$, $i = 1, 2$. The mapping $(\phi_1, \phi_2) \to A_1\phi_1 \otimes A_2\phi_2$ is bilinear from $D(A_1) \times D(A_2) \to \mathcal{B}_1 \otimes \mathcal{B}_2$. The corresponding linear mapping of $D(A_1) \otimes D(A_2)$ into $\mathcal{B}_1 \otimes \mathcal{B}_2$ is denoted by $A_1 \otimes A_2$, and is called the tensor product of the operators A_1 and A_2.

Definition 1.79. Let α be a norm (written $\|\cdot\|_\alpha$) on $\mathcal{B}_1 \otimes \mathcal{B}_2$.

(1) We say that α is a crossnorm if for $\phi_1 \in \mathcal{B}_1,\ \phi_2 \in \mathcal{B}_2$, we have that:

$$\alpha(\phi_1 \otimes \phi_2) = \|\phi_1 \otimes \phi_2\|_\alpha = \|\phi_1\|_{\mathcal{B}_1} \|\phi_2\|_{\mathcal{B}_2}. \tag{1.7}$$

(2) The greatest crossnorm γ on $\mathcal{B}_1 \otimes \mathcal{B}_2$ can be defined on the unit ball in $B(\mathcal{B}_1, \mathcal{B}_2)$. For $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i \in \mathcal{B}_1 \otimes \mathcal{B}_2$,

$$\|\phi\|_\gamma = \sup_{l \in B(\mathcal{B}_1, \mathcal{B}_2)} \left\{ |\langle \phi, l \rangle| : \sum_{i=1}^n \phi_1^i \otimes \phi_2^i = \sum_{k=1}^m \psi_1^i \otimes \psi_2^i \right\}.$$

This norm is equivalent to:

$$\|\phi\|_\gamma = \inf \left\{ \sum_{k=1}^m \left\|\psi_1^i\right\|_{\mathcal{B}_1} \left\|\psi_2^i\right\|_{\mathcal{B}_2} : \sum_{i=1}^n \phi_1^i \otimes \phi_2^i = \sum_{k=1}^m \psi_1^i \otimes \psi_2^i \right\}.$$

(3) The least crossnorm λ is the norm induced on $\mathcal{B}_1 \otimes \mathcal{B}_2$ by the topology of bi-equicontinuous convergence in $B(\mathcal{B}'_1, \mathcal{B}'_2)$. That is, for $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$ and $(F_1, F_2) \in \mathcal{B}'_1 \times \mathcal{B}'_2$,

$$\|\phi\|_\lambda = \sup \left\{ |\langle \phi, F_1 \otimes F_2 \rangle| : \|F_1\|_{\mathcal{B}'_1} \leqslant 1,\ \|F_2\|_{\mathcal{B}'_2} \leqslant 1 \right\}. \tag{1.8}$$

Remark 1.80. For the spaces we are interested in, the least crossnorm $\lambda = \Delta_\infty$, while the greatest crossnorm $\gamma = \Delta_1$.

Definition 1.81. Let α be a given crossnorm on $\mathcal{B}_1 \otimes \mathcal{B}_2$. We say that:

(1) The crossnorm α is a reasonable crossnorm if the dual norm α' induced by the dual of $\mathcal{B}_1 \otimes^\alpha \mathcal{B}_2$ is a crossnorm on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$.

(2) The crossnorm α is uniform relative to $\mathcal{B}_1$ and $\mathcal{B}_2$ if, for $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$ and $A_1,\ A_2 \in L[\mathcal{B}_1], L[\mathcal{B}_2]$:

$$\sup_{\|\phi\|_\alpha \leqslant 1} \|(A_1 \otimes A_2)\phi\|_\alpha \leqslant \|A_1\|_{\mathcal{B}_1} \|A_2\|_{\mathcal{B}_2} . \tag{1.9}$$

(3) If conditions (1) and (2) are satisfied, we say that the crossnorm α is a relative tensor norm.

If α is reasonable, then the norm α' on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ induced by $(\mathcal{B}_1 \otimes^\alpha \mathcal{B}_2)'$ is also reasonable. We denote by $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ the completion of $\mathcal{B}_1 \otimes \mathcal{B}_2$ with respect to α, and by $\mathcal{B}'_1 \hat{\otimes}^{\alpha'} \mathcal{B}'_2$ the completion of $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ with respect to α'. In general, $\mathcal{B}'_1 \hat{\otimes}^{\alpha'} \mathcal{B}'_2$ can be identified with a closed subspace of $(\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2)'$ (cf. Schatten [S], in Chap. 6).

Remark 1.82. Our definition of a relative uniform norm depends on the spaces under consideration. This is a restriction of the conventional definition, which is independent of the spaces, and is called a uniform norm. We refer to Defant and Floret [DOF] for a complete discussion of the standard case. They follow Grothendieck and replace the notion of a uniform norm with the condition that α has the metric approximation property. This, coupled with the first condition, leads to the definition of a tensor norm.

We now give some examples of the norms and standard spaces of interest. Let Ω be a compact domain in $\mathbb{R}^n$ and let $\mathbb{C}[\Omega]$ be the set of bounded continuous functions on Ω, let $L^p[\Omega, \mathfrak{B}(\Omega), m] = L^p[\Omega]$ be the space of Lebesgue integrable functions on Ω that have finite L^p norm, where m is a measure on Ω, and let $\mathfrak{B}(\Omega)$ be the Borel σ-algebra generated by the open sets of Ω. For any Banach space $\mathcal{B}$, it is easy to see that $\mathbb{C}[\Omega, \mathcal{B}] = \mathbb{C}[\Omega] \hat{\otimes}^\lambda \mathcal{B}$.

Example 1.83. *The following are elementary and most will be proved later. We present them because they are what one would naturally expect:*

(1) *Let* $\mathcal{B} = \mathbb{C}[\Omega]$ *as above, so that* $\mathbb{C}[\Omega] \hat{\otimes}^\lambda \mathbb{C}[\Omega] = \mathbb{C}\left[\Omega \times \Omega\right]$. *If* $A_1 = d/dx$, $A_2 = d/dy$, *then*

$$A_1 \hat{\otimes}^\lambda A_2 = \partial^2/\partial x \partial y = d/dx \hat{\otimes}^\lambda d/dy \ = (d/dx \hat{\otimes}^\lambda I)(I \hat{\otimes}^\lambda d/dy)$$

(see Ichinose [IC70], in Chap. 6).

(2) *Let* $\mathcal{B}_3 = \mathcal{B}_4 = L^1[\Omega]$, *then* $L^1[\Omega] \hat{\otimes}^\gamma L^1[\Omega] = L^1\left[\Omega \times \Omega\right]$ *(see Dunford and Schatten [DSH], in Chap. 6).*

(3) *Let* $\mathcal{B}_5 = \mathcal{B}_6 = L^p[\Omega]$, *then* $L^p[\Omega]\hat{\otimes}^{\Delta_p} L^p[\Omega] = L^p[\Omega \times \Omega]$ *for* $1 \leqslant p < \infty$ *(see Schatten [S], in Chap. 6), where*

$$\Delta_p \left(\sum\nolimits_{i=1}^{n} \phi_i \otimes \varphi_i\right) \equiv^{def} \left\{ \iint\limits_{\Omega\times\Omega} \left|\sum\nolimits_{i=1}^{n} \phi_i(x) \otimes \varphi_i(y)\right|^p dxdy \right\}^{1/p}. \quad (1.10)$$

Remark 1.84. Note that $\Delta_1 = \gamma$, $\Delta_\infty = \lambda$, so that Δ_p is always a tensor norm relative to $L^p[\Omega]$ $(1 \leq p \leq \infty)$. It is easy to show that $L^\infty[\Omega]\,\hat{\otimes}^\lambda L^\infty[\Omega] \subset L^\infty[\Omega \times \Omega]$ with the inclusion proper (see Dunford and Schatten [DSH], in Chap. 6). [Similar results show that Δ_p is also a tensor norm relative to the various Sobolev spaces $W^{m,p}[\Omega]$ (see Adams [A], in Chap. 6).] Finally, we can allow that Ω be a locally compact group or complete separable metric space with minor adjustments.

1.5.1. Basic Results. In this section, we prove a number of basic results (including some of those mentioned above) about the tensor product of spaces in the Δ_p norm, $1 \leq p \leq \infty$.

Theorem 1.85. *If* $\mathcal{B}_1$ *and* $\mathcal{B}_2$ *are Banach spaces, then both* $\gamma = \Delta_1$ *and* $\lambda = \Delta_\infty$ *provide norms on* $\mathcal{B}_1 \otimes \mathcal{B}_2$, *with* $\Delta_\infty(\phi) \leq \Delta_1(\phi)$ *for all* $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$.

Proof. We prove that Δ_1 is a norm and $\Delta_\infty(\phi) \leq \Delta_1(\phi)$ for all $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$. The proof that Δ_∞ is a norm is left as an exercise.

It is clear that $\Delta_1(a\phi) = |a|\,\Delta_1(\phi)$. To prove the triangle inequality, let $\varepsilon > 0$ be given and choose $\phi = \sum_{i=1}^{n} \phi_1^i \otimes \phi_2^i$, $\psi = \sum_{i=1}^{m} \psi_1^i \otimes \psi_2^i$, so that

$$\sum_{i=1}^{n} \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2} \leqslant \Delta_1(\phi) + \frac{\varepsilon}{2}$$

and

$$\sum_{i=1}^{m} \left\|\psi_1^i\right\|_{\mathcal{B}_1} \left\|\psi_2^i\right\|_{\mathcal{B}_2} \leqslant \Delta_1(\psi) + \frac{\varepsilon}{2}.$$

By definition, this implies that

$$\begin{aligned} \Delta_1(\phi + \psi) &\leqslant \sum_{i=1}^{n} \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2} + \sum_{i=1}^{m} \left\|\psi_1^i\right\|_{\mathcal{B}_1} \left\|\psi_2^i\right\|_{\mathcal{B}_2} \\ &\leqslant \Delta_1(\phi) + \Delta_1(\psi) + \varepsilon. \end{aligned}$$

Since this is true for all $\varepsilon > 0$, $\Delta_1(\phi + \psi) \leq \Delta_1(\phi) + \Delta_1(\psi)$.

Suppose that $\Delta_1(\phi) = 0$. Then, for each $\varepsilon > 0$, there exists a representation $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i$ such that $\sum_{i=1}^n \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2} \leq \varepsilon$. It follows that, for all $(F_1 \otimes F_2) \in \mathcal{B}'_1 \otimes \mathcal{B}'_2$,

$$\left|(F_1 \otimes F_2)\left(\sum_{i=1}^n \phi_1^i \otimes \phi_2^i\right)\right| \leq \left|\sum_{i=1}^n F_1(\phi_1^i)F_2(\phi_2^i)\right| \leqslant \varepsilon \left\|F_1\right\|_{\mathcal{B'}_1} \left\|F_2\right\|_{\mathcal{B'}_2}.$$

Since $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ is fundamental for $\mathcal{B}_1 \otimes \mathcal{B}_2$, we must have $\phi = 0$.

In order to show that $\Delta_1(\phi_1 \otimes \phi_2) = \|\phi_1\|_{\mathcal{B}_1} \|\phi_2\|_{\mathcal{B}_2}$, first note that $\Delta_1(\phi_1 \otimes \phi_2) \leq \|\phi_1\|_{\mathcal{B}_1} \|\phi_2\|_{\mathcal{B}_2}$, so we need to only prove the opposite relation. Let $F_1 \otimes F_2 \in \mathcal{B}'_1 \otimes \mathcal{B}'_2$ satisfy $F_1(\phi_1) = \|\phi_1\|_{\mathcal{B}_1}$ and $F_2(\phi_2) = \|\phi_2\|_{\mathcal{B}_1}$ (duality maps). Since

$$\begin{aligned}
&\left|(F_1 \otimes F_2)\left(\sum_{i=1}^n \phi_1^i \otimes \phi_2^i\right)\right| \\
&\leqslant \sum_{i=1}^n \left|(F_1 \otimes F_2)\left(\phi_1^i \otimes \phi_2^i\right)\right| \\
&= \sum_{i=1}^n \left|F_1(\phi_1^i)F_2(\phi_2^i)\right| \leqslant \sum_{i=1}^n \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2},
\end{aligned} \tag{1.11}$$

we see that $|(F_1 \otimes F_2)\phi| \leqslant \Delta_1(\phi)$, where $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i$. Thus, $(F_1 \otimes F_2)(\phi_1 \otimes \phi_2) = \|\phi_1\|_{\mathcal{B}_1} \|\phi_2\|_{\mathcal{B}_2} \leq \Delta_1(\phi_1 \otimes \phi_2)$.

From Eq. (1.11) we see that $|(F_1 \otimes F_2)\phi| \leqslant \Delta_1(\phi)$ for all $\phi \in \mathcal{B}_1 \otimes \mathcal{B}_2$. It follows from the definition of Δ_∞ that $\Delta_\infty(\phi) \leq \Delta_1(\phi)$. □

Let Ω be a domain in $\mathbb{R}^n$, let μ be a measure on Ω, and let $\mathcal{B}$ be a separable Banach space with a Schauder basis.

Theorem 1.86. *The completion of $L^1[\Omega, \mu] \otimes \mathcal{B}$ with the Δ_1 norm, $L^1[\Omega, \mu]\hat{\otimes}^{\Delta_1} B$, is isometrically isomorphic to $L^1[\Omega, \mu; B]$, the space of Bochner integrable functions on Ω with values in $\mathcal{B}$.*

Proof. Let $J : L^1[\Omega, \mu] \times \mathcal{B} \to L^1[\Omega, \mu; \mathcal{B}]$, via $(\phi_1, \phi_b) \to \phi = \phi_1 \otimes \phi_b$. By linearization, this induces a norm one mapping

$$J : L^1[\Omega, \mu]\hat{\otimes}^{\Delta_1} \mathcal{B} \to L^1[\Omega, \mu; \mathcal{B}].$$

It follows that $\|J\phi\|_1 \le \Delta_1(\phi)$ for all $\phi \in L^1[\Omega,\mu]\hat{\otimes}^{\Delta_1}\mathcal{B}$. However, if $\phi \in L^1[\Omega,\mu;\mathcal{B}]$ is a simple function, then $\phi = \sum_{k=1}^{n}\chi_{A_k}\phi_b^k$ and $J\phi = \sum_{k=1}^{n}\chi_{A_k}\otimes\phi_b^k$. Thus,

$$\Delta_1(\phi) \le \sum_{k=1}^{n}\mu(A_k)\left\|\phi_b^k\right\| = \|J\phi\|_1 .$$

It follows that $\Delta_1(\phi) = \|J\phi\|_1$ for all simple functions. Since the class S of simple functions is dense in $L^1[\Omega,\mu]$, we see that $S\otimes\mathcal{B}$ is dense in $L^1[\Omega,\mu]\hat{\otimes}^{\Delta_1}\mathcal{B}$. Since the norm closure of the class of Bochner integrable functions is all of $L^1[\Omega,\mu;\mathcal{B}]$, we see that J is surjective. Furthermore, since $\Delta_1(\phi) = \|J\phi\|_1$ on $S\otimes\mathcal{B}$, the extension to $L^1[\Omega,\mu]\hat{\otimes}^{\Delta_1}B$ is both injective and isometric. Thus J is an isometry. □

Corollary 1.87. *If Ω_1, Ω_2 are domains in $\mathbb{R}^n$ with measures μ_1, μ_2, then*

$$L^1[\Omega_1,\mu_1]\hat{\otimes}^{\Delta_1}L^1[\Omega_2,\mu_2] = L^1[\Omega_1\times\Omega_2,\mu_1\otimes\mu_2].$$

Definition 1.88. Let Ω_1, Ω_2 be compact domains in $\mathbb{R}^n$, then

$$\mathbb{C}(\Omega_1)\otimes\mathbb{C}(\Omega_2) =: \{\phi(x,y)\in\mathbb{C}(\Omega_1\times\Omega_2) \mid \exists n\in\mathbb{N},$$
$$\left\{\phi_1^k(x)\right\}_{k=1}^{n}\subset\mathbb{C}(\Omega_1),\ \left\{\phi_2^k(y)\right\}_{k=1}^{n}\subset\mathbb{C}(\Omega_2) \text{ and } \phi(x,y) = \sum_{k=1}^{n}\phi_1^k(x)\phi_2^k(y)\right\}.$$

(If $\Omega_1 = \mathbb{R}^n$, $\Omega_2 = \mathbb{R}^m$ for some n, m, use the one point compactification and the result still applies.)

This is why the notation $(\phi_1\otimes\phi_2)(x,y) = \phi_1(x)\phi_2(y)$ is used to denote products of functions of two variables (in this case). By the Weierstrass Approximation Theorem, we see that $\mathbb{C}(\Omega_1)\otimes\mathbb{C}(\Omega_2)$ is dense in $\mathbb{C}(\Omega_1\times\Omega_2)$.

Theorem 1.89. $\mathbb{C}(\Omega_1)\hat{\otimes}^{\Delta_\infty}\mathbb{C}(\Omega_2) = \mathbb{C}(\Omega_1\times\Omega_2)$.

Theorem 1.90. *Let $\mathcal{B}_1$, $\mathcal{B}_2$ be separable Banach spaces with a Schauder basis.*

(1) *The norm α is a reasonable crossnorm on $\mathcal{B}_1\otimes\mathcal{B}_2$ if and only if*

$$\Delta_\infty(\phi) \le \alpha(\phi) \le \Delta_1(\phi) \text{ for all } \phi\in\mathcal{B}_1\otimes\mathcal{B}_2.$$

(2) *If α is a reasonable crossnorm, then the norm α' on $\mathcal{B'}_1\otimes\mathcal{B'}_2$ induced by $(\mathcal{B}_1\otimes^{\alpha}\mathcal{B}_2)'$ is also a reasonable crossnorm.*

Proof. To begin, we let $\mathbb{B}_{\mathcal{B}'_i}$ denote the unit ball in $\mathcal{B}'_i$, $i = 1, 2$.

If α is a reasonable crossnorm on $\mathcal{B}_1 \otimes \mathcal{B}_2$, then for any representation of $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i \in \mathcal{B}_1 \otimes \mathcal{B}_2$ we have

$$\alpha(\phi) \leqslant \sum_{i=1}^n \alpha(\phi_1^i \otimes \phi_2^i) = \sum_{i=1}^n \left\|\phi_1^i\right\|_{\mathcal{B}_1} \left\|\phi_2^i\right\|_{\mathcal{B}_2},$$

so that $\alpha(\phi) \leq \Delta_1(\phi)$. To see that $\Delta_\infty(\phi) \leq \alpha(\phi)$, let $F_1 \otimes F_2$ and $F = \sum_{i=1}^n F_1^i \otimes F_2^i$ be in $\mathcal{B}'_1 \otimes \mathcal{B}'_2$. Then

$$\begin{aligned}
&\Delta_\infty(\phi) \\
&= \sup\left\{|(F_1 \otimes F_2)\phi| : F_1 \otimes F_2 \in \mathbb{B}_{\mathcal{B}'_1} \otimes \mathbb{B}_{\mathcal{B}'_2}\right\} \\
&\leqslant \sup\left\{|(F)\phi| : F = \sum_{i=1}^n F_1^i \otimes F_2^i \in \mathbb{B}_{\mathcal{B}'_1} \otimes \mathbb{B}_{\mathcal{B}'_2}\right\} = \alpha(\phi).
\end{aligned}$$

On the other hand, if α is a norm on $\mathcal{B}_1 \otimes \mathcal{B}_2$ with

$$\Delta_\infty(\phi) \leq \alpha(\phi) \leq \Delta_1(\phi), \text{ for all } \phi \in \mathcal{B}_1 \otimes \mathcal{B}_2,$$

then $\Delta_\infty(\phi \otimes \psi) = \alpha(\phi \otimes \psi) = \Delta_1(\phi \otimes \psi)$, so that α is a crossnorm. To see that α' is a crossnorm on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$, use $\Delta_\infty \leq \alpha \leq \Delta_1$ to get that, for $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i \in \mathcal{B}_1 \otimes \mathcal{B}_2$,

$$\begin{aligned}
&\|F_1\|_{\mathcal{B}'_1} \|F_2\|_{\mathcal{B}'_2} \\
&= \sup\left\{|(F_1 \otimes F_2)\phi| : \phi \in \mathcal{B}_1 \otimes \mathcal{B}_2,\ \Delta_1(\phi) \leqslant 1\right\} \\
&\leqslant \alpha'(F_1 \otimes F_2) = \sup\left\{|(F_1 \otimes F_2)\phi| : \phi \in \mathcal{B}_1 \otimes \mathcal{B}_2,\ \alpha(\phi) \leqslant 1\right\} \\
&\leqslant \sup\left\{\,|(F_1 \otimes F_2)\phi||\ \phi \in \mathcal{B}_1 \otimes \mathcal{B}_2,\ \Delta_\infty(\phi) \leqslant 1\right\} = \|F_1\|_{\mathcal{B}'_1} \|F_2\|_{\mathcal{B}'_2}.
\end{aligned}$$

It follows that α' is a reasonable crossnorm. □

We denote by $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ the completion of $\mathcal{B}_1 \otimes \mathcal{B}_2$ with respect to α, and by $\mathcal{B}'_1 \hat{\otimes}^{\alpha'} \mathcal{B}'_2$ the completion of $\mathcal{B}'_1 \otimes \mathcal{B}'_2$ with respect to α'. In general, $\mathcal{B}'_1 \hat{\otimes}^{\alpha'} \mathcal{B}'_2$ can be identified with a closed subspace of $(\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2)'$ (cf. Schatten [S], in Chap. 6).

Theorem 1.91. *If Ω is a domain in $\mathbb{R}^n$ and μ is a measure on $\mathfrak{B}$, then Δ_p is a reasonable crossnorm on $L^p[\Omega, \mathfrak{B}(\Omega), \mu] \otimes \mathcal{B} =: L^p[\Omega] \otimes \mathcal{B}$, $1 \leq p \leq \infty$, for any separable Banach space $\mathcal{B}$, and $L^p[\Omega] \hat{\otimes}^{\Delta_p} \mathcal{B} = L^p[\Omega; \mathcal{B}]$ for $1 \leq p < \infty$.*

Proof. The proof for $p = 1$ was given in Theorem 6.8, so we need to only consider the case for $1 < p \leq \infty$. Let $J : L^p[\Omega] \otimes \mathcal{B} \to L^p[\Omega; \mathcal{B}]$

be defined by $J[f \otimes \phi] = f(\cdot)\phi$. This is clearly an injective mapping. Let $g = \sum_{k=1}^{n} f_k \otimes \phi_i$ and define

$$\Delta_p[g] = \left[\int_\Omega \|g(\omega)\|_{\mathcal{B}}^p \, d\mu(\omega)\right]^{1/p}.$$

It is clear that

$$\Delta_p[f \otimes \phi] = \left[\int_\Omega \|f(\omega)\phi\|_{\mathcal{B}}^p \, d\mu(\omega)\right]^{1/p} = \|\phi\|_{\mathcal{B}} \, \|f\|_p \,,$$

so that $\Delta_p[\cdot]$ is a crossnorm. To see that $\Delta_p[\cdot] \le \Delta_1[\cdot]$, note that

$$\begin{aligned} &\Delta_p[g] \\ &= \left[\int_\Omega \left\| \sum_{k=1}^{n} f_k(\omega)\phi_k \right\|_{\mathcal{B}}^p d\mu(\omega)\right]^{1/p} \\ &\leqslant \left[\sum_{k=1}^{n} \|\phi_k\|_{\mathcal{B}}^p \int_\Omega |f_k(\omega)|^p \, d\mu(\omega)\right]^{1/p} \\ &\leqslant \sum_{k=1}^{n} \|\phi_k\|_{\mathcal{B}} \, \|f_k\|_p, \end{aligned}$$

so that $\Delta_p[g] \leqslant \Delta_1[g]$. To see that $\Delta_\infty[\cdot] \le \Delta_p[\cdot]$, let $F \otimes \Phi \in [L^p]' \otimes \mathcal{B}'$ be in the respective unit balls (i.e., $\|F\|_{p'} \leqslant 1$, $\|\Phi\|_{B'} \leqslant 1$, $\frac{1}{p'} = 1 - \frac{1}{p}$). Then

$$\begin{aligned} &|\langle F \otimes \Phi, g\rangle| \\ &= \left| \int_\Omega F(\omega) \, \langle \Phi, g(\omega)\rangle \, d\mu(\omega) \right| \\ &\le \sup_{\|\Phi\| \leqslant 1, \ \|F\|_{p'} \leqslant 1} |\langle F \otimes \Phi, g\rangle| = \Delta_\infty[g] \\ &\leqslant \|F\|_{p'} \left[\int_\Omega |\langle \Phi, g(\omega)\rangle|^p \, d\mu(\omega)\right]^{1/p} \\ &\leqslant \sup_{\|\Phi\| \leqslant 1} \left[\int_\Omega |\langle \Phi, g(\omega)\rangle|^p \, d\mu(\omega)\right]^{1/p} = \Delta_p[g]. \end{aligned}$$

Thus, $\Delta_p[\cdot]$ is a reasonable crossnorm on $L^p[\Omega] \otimes \mathcal{B}$ for any p, $1 \le p \le \infty$. If $p < \infty$, then the (equivalence class of) step functions

$$S(\mu) \otimes B = \left\{ \sum_{k=1}^{n} \chi_{A_k} \otimes \phi_k : \ n \in \mathbb{N}, \ \mu(A_k) < \infty, \, \phi_k \in B \right\}$$

is dense in $L^p[\Omega;\mathcal{B}]$. This implies that $L^p[\Omega]\otimes\mathcal{B}$ is dense in $L^p[\Omega;\mathcal{B}]$. It follows that

$$L^p[\Omega]\hat{\otimes}^{\Delta_p}B = L^p[\Omega;B].$$

□

Corollary 1.92. *Let* $(\Omega_1,\mathfrak{B}_1,\mu_1)$ *and* $(\Omega_2,\mathfrak{B}_2,\mu_2)$ *be* σ*-finite measure spaces. Then* $L^p[\Omega_1\times\Omega_2,\ \mathfrak{B}_1\times\mathfrak{B}_2,\ \mu_1\times\mu_2] = L^p[\Omega_1]\hat{\otimes}^{\Delta_p}L^p[\Omega_2]$ *for* $1\le p<\infty$, *where* $\mathfrak{B}_1\times\mathfrak{B}_2$ *is the* σ*-algebra generated by* $\mathfrak{B}_1$ *and* $\mathfrak{B}_2$.

Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be separable Banach spaces, and let $A_1\in L[\mathcal{B}_1]$, $A_2\in L[\mathcal{B}_2]$.

Theorem 1.93. *The operator* $A_1\otimes A_2:\mathcal{B}_1\otimes\mathcal{B}_2\to\mathcal{B}_1\otimes\mathcal{B}_2$ *has an unique extension to both* $\mathcal{B}_1\hat{\otimes}^{\Delta_1}\mathcal{B}_2$ *and* $\mathcal{B}_1\hat{\otimes}^{\Delta_\infty}\mathcal{B}_2$ *as a bounded linear operator and*

$$\left\|A_1\hat{\otimes}A_2\right\| = \|A_1\|\,\|A_2\|\,.$$

Proof. We first prove it for Δ_1. Let $\sum_{i=1}^n \phi_1^i\otimes\phi_2^i$ be a representation for $\phi\in\mathcal{B}_1\hat{\otimes}^{\Delta_1}\mathcal{B}_2$. Then

$$\Delta_1[(A_1\otimes A_2)\,\phi] = \Delta_1\left[\sum_{i=1}^n A_1\phi_1^i\otimes A_2\phi_2^i\right] \leqslant \|A_1\|\,\|A_2\|\sum_{i=1}^n \left\|\phi_1^i\right\|_{\mathcal{B}_1}\left\|\phi_2^i\right\|_{\mathcal{B}_2},$$

so that $\Delta_1[(A_1\otimes A_2)\,\phi]\le\|A_1\|\,\|A_2\|\,\Delta_1[\phi]$. It follows that $\|A_1\otimes A_2\| \le \|A_1\|\,\|A_2\|$. However, from $(A_1\otimes A_2)(\phi_1\otimes\phi_2) = (A_1\phi_1)\otimes(A_2\phi_2)$, we see that (using the crossnorm property of Δ_1)

$$\begin{aligned}
&\|A_1\|\,\|A_2\| \\
&= \sup\frac{\|(A_1\phi_1)\|_{\mathcal{B}_1}\,\|(A_2\phi_2)\|_{\mathcal{B}_2}}{\|\phi_1\|_{\mathcal{B}_1}\,\|\phi_2\|_{\mathcal{B}_2}} \\
&= \sup\frac{\|(A_1\phi_1\otimes A_2\phi_2)\|_{\Delta_1}}{\|\phi_1\|_{\mathcal{B}_1}\,\|\phi_2\|_{\mathcal{B}_2}} \\
&= \sup\frac{\|(A_1\otimes A_2)\,(\phi_1\otimes\phi_2)\|_{\Delta_1}}{\|(\phi_1\otimes\phi_2)\|_{\Delta_1}} \leqslant \|A_1\otimes A_2\|\,.
\end{aligned}$$

It follows that $\|A_1\otimes A_2\| = \|A_1\|\,\|A_2\|$. It is clear that this equality holds for the unique extension $A_1\hat{\otimes}A_2$ of $A_1\otimes A_2$ to all of $\mathcal{B}_1\hat{\otimes}^{\Delta_1}\mathcal{B}_2$.

To prove the result for Δ_∞, let $\phi = \sum_{i=1}^n \phi_1^i \otimes \phi_2^i$. Then

$$\begin{aligned}
&\Delta_\infty[(A_1 \otimes A_2)\,\phi] \\
&= \sup\left\{ \sum_{i=1}^n F_1(A_1\phi_1^i)F_2(A_2\phi_2^i) :\ F_1 \in \mathcal{B}'{}_1,\ F_2 \in \mathcal{B}'{}_2;\ \|F_1\| \leqslant 1,\ \|F_2\| \leqslant 1 \right\} \\
&= \sup\left\{ \sum_{i=1}^n (A_1'F_1)(\phi_1^i)(A_2'F_2)(\phi_2^i) :\ F_1 \in \mathcal{B}'{}_1,\ F_2 \in \mathcal{B}'{}_2;\ \|F_1\| \leqslant 1,\ \|F_2\| \leqslant 1 \right\} \\
&\leqslant \|A_1'\|\,\|A_2'\|\,\Delta_\infty[\phi] = \|A_1\|\,\|A_2\|\,\Delta_\infty[\phi].
\end{aligned}$$

Thus, $A_1 \otimes A_2$ has a bounded extension to $\mathcal{B}_1 \hat{\otimes}^{\Delta_\infty} \mathcal{B}_2$. Now let $\varepsilon > 0$ and choose $\phi_1 \in \mathcal{B}_1,\ \phi_2 \in \mathcal{B}_2$ with $\|\phi_1\|_{\mathcal{B}_1} \leq 1,\ \|\phi_2\|_{\mathcal{B}_2} \leq 1$ and, such that

$$\|A_1\phi_1\|_{\mathcal{B}_1} \geqslant (1-\varepsilon)\,\|A_1\|_{\mathcal{B}_1}; \quad \|A_2\phi_2\|_{\mathcal{B}_2} \geqslant (1-\varepsilon)\,\|A_2\|_{\mathcal{B}_2}.$$

Thus, $\Delta_\infty(\phi_1 \otimes \phi_2) \leq 1$ and

$$\|A_1\phi_1\|_{\mathcal{B}_1}\,\|A_2\phi_2\|_{\mathcal{B}_2} = \Delta_\infty[(A_1 \otimes A_2)\,(\phi_1 \otimes \phi_2)] \geqslant (1-\varepsilon)^2\,\|A_1\|\,\|A_2\|.$$

Since ε is arbitrary, $\|A_1 \otimes A_2\| = \|A_1\|\,\|A_2\|$. It follows that the same is true for the unique extension $A_1 \hat{\otimes} A_2$ of $A_1 \otimes A_2$ to all of $\mathcal{B}_1 \hat{\otimes}^{\Delta_\infty} \mathcal{B}_2$.
□

From Theorem 1.93, we see that Δ_1 and Δ_∞ are uniform for all Banach space couples (tensor norms). The following example shows that, for $1 < p < \infty$, we cannot expect Δ_p to be uniform for all Banach space couples.

Let $L^2[\mathbb{R}]$ and $\ell_1(\mathbb{R})$ have the standard definitions, and let $\mathfrak{F}$ be the Fourier transform on $L^2[\mathbb{R}]$, which is an isometry, and let $\mathbf{I}_1$ be the identity on $\ell_1(\mathbb{R})$. If $\mathcal{B}_1 = L^2(\mathbb{R})$, $\mathcal{B}_2 = \ell_1(\mathbb{R})$ and $\alpha = \Delta_2$, we have

$$\Delta_2(\sum_{m=1}^n \varphi_m \otimes \psi_m) \equiv^{def} \left\{ \int_{-\infty}^{\infty} \left\| \sum_{m=1}^n \varphi_m(x)\psi_m(y) \right\|_{\mathcal{B}_2}^2 dx \right\}^{1/2}.$$

Example 1.94. *Set* $f_n = \sum_{m=1}^n \chi_{[m,m+1)} \otimes e_m$, *where* $\chi_{[m,m+1)}(x)$ *is the characteristic function of the interval* $[m, m+1)$, *and* e_m *is the* m*th unit basis vector of* $\ell_1(\mathbb{R})$. *Then*

$$\Delta_2(f_n) = \left\{ \int_{-\infty}^{\infty} \left\| \sum_{m=1}^n \chi_{[m,m+1)}(x)e_m \right\|_{l_1}^2 dx \right\}^{1/2} = \sqrt{n}.$$

However, if we look at the norm of $(\mathfrak{F} \otimes \mathbf{I}_1)f_n$*, we get:*

$$\|(\mathfrak{F} \otimes \mathbf{I}_1)f_n\|_{\Delta_2}$$

$$= \left\{ \int_{-\infty}^{\infty} \left\| \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ixy} \left[\sum_{m=1}^{n} \chi_{[m,m+1]}(y) e_m \right] dy \right\|_{l_1}^{2} dx \right\}^{1/2}$$

$$= \tfrac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \left\| \tfrac{1}{x} \left[\sum_{m=1}^{n} \{ e^{[-i(m+1)x]} - e^{[-imx]} \} e_m \right] \right\|_{l_1}^{2} dx \right\}^{1/2}$$

$$= \tfrac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \tfrac{1}{x^2} \left[\sum_{m=1}^{n} \left| e^{[-i(m+1)x]} - e^{[-imx]} \right| \right]^{2} dx \right\}^{1/2}$$

$$= \frac{n}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \left| \frac{\exp(-ix) - 1}{x} \right|^{2} dx \right\}^{1/2} = n.$$

It follows that $\mathfrak{F} \otimes \mathbf{I}_1$ *cannot extend to a bounded operator on* $L^2[\mathbb{R}]\hat{\otimes}^{\Delta_2} \ell_1(\mathbb{R})$*. Thus,* Δ_2 *is not uniform with respect to* $L^2[\mathbb{R}]$ *and* $\ell_1(\mathbb{R})$*, so that* Δ_2 *is not a tensor norm. However, it is a relative tensor norm for the right space. To see this in the above case, replace* $\ell_1(\mathbb{R})$ *by* $\ell_2(\mathbb{R})$ *and note that, if* e_m *is the mth unit basis vector of* $\ell_2(\mathbb{R})$*, then* $f_n \in L^2[\mathbb{R}]\hat{\otimes}^{\Delta_2} \ell_2(\mathbb{R})$ *and we have:*

$$\Delta_2(f_n) = \left\{ \int_{-\infty}^{\infty} \left\| \sum_{m=1}^{n} \chi_{[m,m+1)}(x) e_m \right\|_{l_2}^{2} dx \right\}^{1/2} = \sqrt{n}.$$

If we now look at the norm of $(\mathfrak{F} \otimes \mathbf{I}_2)f_n$*, we get (the expected result):*

$$\|(\mathcal{F} \otimes \mathbf{I}_2)f_n\|_{\Delta_2}$$

$$= \left\{ \int_{-\infty}^{\infty} \left\| \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ixy} \left[\sum_{m=1}^{n} \chi_{[m,m+1]}(y) e_m \right] dy \right\|_{l_2}^{2} dx \right\}^{1/2}$$

$$= \tfrac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \left\| \tfrac{1}{x} \left[\sum_{m=1}^{n} \{ e^{[-i(m+1)x]} - e^{[-imx]} \} e_m \right] \right\|_{l_2}^{2} dx \right\}^{1/2}$$

$$= \frac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \frac{1}{x^2} \sum_{m=1}^{n} \left| e^{[-i(m+1)x]} - e^{[-imx]} \right|^2 dx \right\}^{1/2}$$

$$= \frac{\sqrt{n}}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \left| \frac{\exp(-ix) - 1}{x} \right|^2 dx \right\}^{1/2} = \sqrt{n}.$$

It follows that $(\mathfrak{F} \otimes \mathbf{I}_2)$ *can be extended to a bounded operator on* $L^2[\mathbb{R}]\hat{\otimes}^{\Delta_2} \ell_2(\mathbb{R})$.

By a theorem of Kwapień [KP], in Chap. 6, $\mathcal{B}$ is isomorphic to a Hilbert space if and only if $(\mathfrak{F} \otimes \mathbf{I}_{\mathcal{B}})$ is continuous on $\mathbf{L}^2[\mathbb{R}]\hat{\otimes}^{\Delta_2}\mathcal{B}$. The point is that Δ_2 is a relative tensor norm which is not a tensor norm. On the other hand, if α is any tensor norm, $\mathfrak{F} \otimes \mathbf{I}_1$ has an extension to a bounded linear operator on $\mathbf{L}^2[\mathbb{R}]\hat{\otimes}^{\alpha} l_1(\mathbb{Z})$ (see [DOF, p. 147]).

We now show that Δ_p is uniform relative to the tensor product of L^p spaces.

Theorem 1.95. *Let* $(\Omega_1, \mathfrak{B}_1, \mu_1)$ *and* $(\Omega_2, \mathfrak{B}_2, \mu_2)$ *be* σ*-finite measure spaces. Let* $A_1 : L^p[\Omega_1] \to L^p[\Omega_1]$ *and* $A_2 : L^p[\Omega_2] \to L^p[\Omega_2]$*. Then, for* $1 < p < \infty$*, the operator*

$$A_1 \otimes A_2 : L^p[\Omega_1] \otimes L^p[\Omega_2] \to L^p[\Omega_1] \otimes L^p[\Omega_2]$$

has an unique extension to a bounded linear operator

$$A_1 \hat{\otimes} A_2 : L^p[\Omega_1]\hat{\otimes}^{\Delta_p} L^p[\Omega_2] \to L^p[\Omega_1 \times \Omega_2],$$

and $\left\| A_1 \hat{\otimes} A_2 \right\| = \|A_1\| \, \|A_2\|$.

Proof. We first show that $I_1 \otimes A_2$ is bounded as an operator mapping $L^p[\Omega_1] \otimes L^p[\Omega_2] \to L^p[\Omega_1 \times \Omega_2]$.

Let $\{\phi_2^i\}$ be a Schauder basis for $L^p[\Omega_2]$ and, for $1 \le i \le n$, $n \in \mathbb{N}$, let $\psi_2^i = A_2 \phi_2^i$. Then, for all scalars $a_1, \ldots, a_n$, we have

$$\left\| \sum_{i=1}^{n} a_i \psi_2^i \right\|_p \leqslant \|A_2\| \left\| \sum_{i=1}^{n} a_i \phi_2^i \right\|_p .$$

It follows that, for arbitrary functions $a_1(\cdot), \ldots, a_n(\cdot) \in L^p[\Omega_1]$,

$$\int_{\Omega_2} \left| \sum_{i=1}^{n} a_i(x) \psi_2^i(y) \right|^p \mu_2(dy) \leqslant \|A_2\|^p \int_{\Omega_2} \left| \sum_{i=1}^{n} a_i(x) \phi_2^i(y) \right|^p \mu_2(dy).$$

Integrating both sides with respect to μ_1, we see that

$$\begin{aligned}&\int_{\Omega_1}\int_{\Omega_2}\left|\sum_{i=1}^n a_i(x)\psi_2^i(y)\right|^p \mu_2(dy)\mu_1(dx)\\ &\leqslant \|A_2\|^p\int_{\Omega_1}\int_{\Omega_2}\left|\sum_{i=1}^n a_i(x)\phi_2^i(y)\right|^p \mu_2(dy)\mu_1(dx).\end{aligned}\tag{1.12}$$

Since μ_1 and μ_2 are σ-finite, we can use Fubini's Theorem to get

$$\begin{aligned}&\int_{\Omega_1}\int_{\Omega_2}\left|\sum_{i=1}^n a_i(x)\psi_2^i(y)\right|^p \mu_2(dy)\mu_1(dx)\\ &=\int_{\Omega_2}\left\|\sum_{i=1}^n a_i(\cdot)\psi_2^i(y)\right\|_p^p \mu_2(dy).\end{aligned}\tag{1.13}$$

If we set $\Phi(x)=\left|\sum_{i=1}^n a_i(x)\psi_2^i(\cdot)\right|^p,\ \ x\in\Omega_1$, then

$$\begin{aligned}&\int_{\Omega_2}\left\|\sum_{i=1}^n a_i(\cdot)\psi_2^i(y)\right\|_p^p \mu_2(dy)\\ &=\left\|\int_{\Omega_1}\Phi(x)\mu_1(dx)\right\|_p^p\leqslant\int_{\Omega_1}\|\Phi(x)\|_p^p\,\mu_1(dx).\end{aligned}\tag{1.14}$$

If we combine Eqs. (6.8)–(6.10), we get

$$\int_{\Omega_2}\left\|\sum_{i=1}^n a_i(\cdot)\psi_2^i(y)\right\|_p^p \mu_2(dy)\leqslant\|A_2\|^p\int_{\Omega_2}\left\|\sum_{i=1}^n a_i(\cdot)\phi_2^i(y)\right\|_p^p \mu_2(dy).$$

It follows that $\|I_1\otimes A_2\|\leqslant\|A_2\|^p$.

Since

$$\begin{aligned}\left|\sum_{i=1}^n a_i(x)\psi_2^i(y)\right|&=\left|A_2\left[\sum_{i=1}^n a_i(x)\phi_2^i(y)\right]\right|\\ &=\left|\left((I_1\otimes A_2)\left[\sum_{i=1}^n a_i\otimes\phi_2^i\right]\right)(x,y)\right|,\end{aligned}$$

we see that

$$\left\| \left[A_2 \left(\sum_{i=1}^{n} a_i(\cdot)\psi_2^i(\cdot) \right) \right] \right\|_p = \left\| \sum_{i=1}^{n} a_i \otimes A_2\phi_2^i \right\|_{\Delta_p}$$

$$= \left\| \left((I_1 \otimes A_2) \left[\sum_{i=1}^{n} a_i \otimes \phi_2^i \right] \right) \right\|_{\Delta_p} \leqslant \|I_1 \otimes A_2\| \left\| \sum_{i=1}^{n} a_i \otimes \phi_2^i \right\|_{\Delta_p}.$$

Thus we see that $\|I_1 \otimes A_2\| = \|A_2\|$.

The same proof (with minor adjustments) shows that $(A_1 \otimes I_2)$ is also bounded as an operator mapping $L^p[\Omega_1]\otimes L^p[\Omega_2] \to L^p[\Omega_1 \times \Omega_2]$. Since $A_1 \otimes A_2 = (I_1 \otimes A_2)(A_1 \otimes I_2)$, we see that $\|A_1 \otimes A_2\| = \|A_1\| \, \|A_2\|$. □

Thus, we see that Δ_p is always a tensor norm relative to $\mathbf{L}^p[\Omega]$ $(1 \leqslant p \leq \infty)$.

Theorem 1.96. *[Schatten* [S]*, in Chap. 6] The norms λ, γ are tensor norms on $\mathcal{B}_1 \otimes \mathcal{B}_2$ and $\lambda \leqslant \gamma$. Furthermore, if α is any norm with $\lambda \leqslant \alpha \leqslant \gamma$, then α is a reasonable crossnorm which is a relative tensor norm that need not be a tensor norm, and $\gamma' \leqslant \alpha' \leqslant \lambda'$ (i.e., α' is a crossnorm on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$, and $\gamma' = \lambda$, $\lambda' = \gamma$).*

Definition 1.97. A relative tensor norm α is said to be faithful if the natural linear mapping of $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ into $\mathbb{L}_s(\mathcal{B}'_1, \mathcal{B}'_2)$, obtained by extending the identity $\mathbf{I}_1 \otimes \mathbf{I}_2$ on $\mathcal{B}_1 \otimes \mathcal{B}_2 \subset \mathcal{B}_1 \hat{\otimes}^\lambda \mathcal{B}_2$ by continuity to the entire space $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$, is one-to-one.

To say that α is faithful means that, if an element of $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ vanishes on $\mathcal{B}'_1 \otimes \mathcal{B}'_2$, it is the zero function. For all of the above spaces, the relative tensor norm is faithful. Indeed, it has been shown by Gelbaum and Gil de Lamadrid [GG], in Chap. 6, that, if both $\mathcal{B}_1$ and $\mathcal{B}_2$ have Schauder bases and α is a relative tensor norm, then $\mathcal{B}_1 \hat{\otimes}^\alpha \mathcal{B}_2$ has a Schauder basis so that α is faithful. The following result is due to Ichinose [IC70], in Chap. 6.

Theorem 1.98. *Let A_1 and A_2 be closed densely defined linear operators on $\mathcal{B}_1$ and $\mathcal{B}_2$ respectively, and let α be a faithful relative tensor norm. Unless one of the extended spectra $\sigma_e(A_1)$ and $\sigma_e(A_2)$ contains* 0 *while the other contains ∞,*

$$(A_1 \hat{\otimes}^\alpha I_2)(I_1 \hat{\otimes}^\alpha A_2) = (I_1 \hat{\otimes}^\alpha A_2)(A_1 \hat{\otimes}^\alpha I_2) = A_1 \hat{\otimes}^\alpha A_2. \tag{1.15}$$

References

[CA] N.L. Carothers, *A Short Course on Banach Space Theory*. London Mathematical Society Student Texts, vol. 64 (Cambridge University Press, Cambridge, 2005)

[DOF] A. Defant, K. Floret, *Tensor Norms and Operator Ideals*. Mathematics Studies, vol. 176 (North-Holland, New York, 1993)

[DI] J. Diestel, *Sequences and Series in Banach Spaces*. Graduate Texts in Mathematics (Springer, New York, 1984)

[DU] J. Dugundji, *Linear Operators Part I: General Theory*. Wiley Classics Edition (Wiley, New York, 1988)

[DS] N. Dunford, J.T. Schwartz, *Linear Operators Part I: General Theory*. Wiley Classics Edition (Wiley, New York, 1988)

[EV] L.C. Evans, *Partial Differential Equations*. Graduate Studies in Mathematics, vol. 18 (American Mathematical Society, Providence, 1998)

[HP] E. Hille, R.S. Phillips, *Functional Analysis and Semigroups*. Colloquium Publications, vol. 31 (American Mathematical Society, Providence, 1957)

[J] F. Jones, *Lebesgue Integration on Euclidean Space*, revised edn. (Jones and Bartlett, Boston, 2001)

[L1] P.D. Lax, *Functional Analysis*, 2nd edn. (Wiley, New York, 2002)

[GL] G. Leoni, *A First Course in Sobolev Spaces.* Graduate Studies in Mathematics, vol. 105 (American Mathematical Society, Providence, 2009)

[ME] B. Mendelson, *Introduction to Topology* (Dover, New York, 1980)

[RS] M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic, New York, 1972)

[RO] H.L. Royden, *Real Analysis*, 2nd edn. (Macmillan, New York, 1968)

[RU1] W. Rudin, *Functional Analysis* (McGraw-Hill, New York, 1973)

[RU] W. Rudin, *Real and Complex Analysis*, 3rd edn. (McGraw-Hill, New York, 1987)

[RA] R. Ryan, *Introduction to Tensor Products of Banach Spaces* (Springer, New York, 2002)

[STE] E. Stein, *Singular Integrals and Differentiability Properties of Functions.* Princeton Mathematical Series, vol. 30 (Princeton University Press, Princeton, 1970)

[SZ] R.S. Strichartz, *A Guide to Distribution Theory and Fourier Transforms* (World Scientific, River Edge, 1994)

[YS] K. Yosida, *Functional Analysis*, 2nd edn. (Springer, New York, 1968)

Chapter 2

Integration on Infinite-Dimensional Spaces

This chapter is required for the foundations of infinite-dimensional analysis. It is assumed that the reader is conversant with Lebesgue measure on $\mathbb{R}^n$, including the standard limit theorems, inequalities, convolution, Fourier transform theory, and Fubini's theorem. With this in mind, we offer a parallel treatment on infinite-dimensional spaces, with a theorem proof protocol. The proof of any theorem that is the same as on $\mathbb{R}^n$ is omitted. We have also added a few interesting topics, which are discussed more fully below. We do not include any exercises, however, any serious question could lead to a research problem. (This statement applies to all chapters except Chaps. 1 and 4.)

This chapter is not required for the Feynman operator calculus of Chap. 7, or for the general path integral theory of Chap. 8. However, one should continue reading at least to Theorem 2.7, in order to understand the notation in Chap. 3.

In the first section we discuss the basic problem for Lebesgue measure on $\mathbb{R}^\infty$. After some historical background, we introduce a new class of open sets for $\mathbb{R}^\infty$. The induced topology is natural for Lebesgue measure on $\mathbb{R}^\infty$. After constructing our measure λ_∞, on $\mathbb{R}^\infty$, we show

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_2

that its restriction to $\mathbb{R}^n$ is equivalent to n-dimensional Lebesgue measure for each n.

In the second section we construct the corresponding version of Lebesgue measure for every Banach space with an S-basis. In addition, we develop a general theory of probability measure for Banach spaces with an S-basis. The interesting result in this direction is our proof that every probability measure with a density, defined on the Borel sets of $\mathbb{R}$, induces a (closely related) family of probability measures on every Banach space with an S-basis, that is absolutely continuous with respect to Lebesgue measure.

In the third section we discuss integrable functions, L^p spaces, product measures, convolutions, and integral inequalities. In the fourth section, we develop a general theory of distributions on uniformly convex Banach spaces, while in the fifth section we construct the Schwartz space and the Fourier transform on uniformly convex Banach spaces. We then use the transform to extend the Pontryagin duality theory to a new class of nonlocally compact groups. Finally, in the last section, we provide a direct solution of the diffusion equation on Hilbert space, as an application of our theory.

2.1. Lebesgue Measure on $\mathbb{R}^\infty$

Historically, the topology for $\mathbb{R}^\infty$ defines open sets to be the cartesian product of an arbitrary finite number of open sets in $\mathbb{R}$, while the remaining infinite number are copies of $\mathbb{R}$ (cylindrical sets). This approach forces any possible measure to assign a finite value to each open set. The success of Kolmogorov's work on the foundations of probability theory [KO] has since embedded this requirement into the fabric of infinite dimensional analysis. Any attempt to construct a reasonable version of Lebesgue measure using the above approach is impossible.

2.1.1. Background. Historically, the first advance in understanding the problem of measures on infinite dimensional spaces was made (indirectly) in 1933 when Haar [HA] proved:

Theorem 2.1. *If G is a locally compact abelian group, then there exists a nonnegative regular translation invariant measure m (Haar measure) on G (i.e., $m(A + x) = m(A)$ for every $x \in G$ and every Borel set A in $\mathfrak{B}[G]$).*

This result stimulated the interest of von Neumann [VN1], who proved that it is the only σ-finite left-invariant Borel measure on such a group (uniqueness up to a multiplicative constant) and Weil [WE] proved that:

Theorem 2.2. *If G is a (separable) topological group and m is a σ-finite left-translation invariant Borel measure on G, then it is always possible to define an equivalent locally compact topology on G.*

Since $\mathbb{R}^\infty$ is a complete separable nonlocally compact metric or Polish group, this necessarily means that a σ-finite translation invariant measure cannot be defined on $\mathbb{R}^\infty$. In 1946 Oxtoby [OX] began the study of translation-invariant Borel measures on Polish groups. He proved the following result attributed to Ulam:

Theorem 2.3. *There always exists a left-invariant Borel measure on any Polish group which assigns positive finite measure to at least one set and vanishes on singletons. However, a σ-finite measure is possible if and only if the group is locally compact.*

In 1959 Sudakov [SU] explicitly proved that: *If $\mathbb{R}^\infty$ is regarded as a linear topological space, then there does not exist a σ-finite translation-invariant Borel measure for $\mathbb{R}^\infty$.* Since then other studies have been conducted on the subject. For additional information, the (relatively) recent papers by Baker [BA1], [BA2] and by Vershik [V], [V1], [V2] are especially recommended. The papers by Hill [HI] and Ritter and Hewitt [RH] are also worth reading.

2.1.2. An Alternate Approach. On $\mathbb{R}^n$, $n < \infty$, it is useful to think of Lebesgue measure in terms of geometric objects. It is also natural to expect that Lebesgue measure will leave geometric objects invariant under translations and rotations, and to assume that rotational and translational invariance is an intrinsic property of Lebesgue measure. In many applications, rotational and translational invariance plays no role at all, it is the σ-finite nature of Lebesgue measure that is critical. In addition to probability theory (stochastic processes), another major motivation for the study of measure on infinite-dimensional space is its importance for the foundations of statistical mechanics and quantum theory. In physical systems translational invariance is intimately related to the total momentum and rotational invariance is intimately related to the total angular momentum. (In the precise language of groups, the total momentum is the generator of translational invariance and the total angular momentum is the generator of rotational invariance.) In both cases, these quantities must be finite in order for a physical system to be well defined. This is impossible in infinite-dimensional space if we require full rotational and translational invariance.

We take a different approach, using a basic requirement on any mapping on Borel sets, which would serve as an acceptable version of Lebesgue measure on $\mathbb{R}^\infty$. In particular, if $I_0 = \left[-\frac{1}{2}, \frac{1}{2}\right]^{\aleph_0}$, then any definition $\lambda_\infty(\cdot)$, of Lebesgue measure must satisfy $\lambda_\infty\left[I_0\right] = 1$. We make this requirement the centerfold of our approach. Let $\mathfrak{B}[\mathbb{R}^n]$ be the Borel σ-algebra for $\mathbb{R}^n$ and $I = [-\frac{1}{2}, \frac{1}{2}]$.

Definition 2.4. Let $A \in \mathfrak{B}[\mathbb{R}^n]$, $n \in \mathbb{N}$. A set of the form $A_n = A \times I_n$, $I_n = \prod_{i=n+1}^\infty I$ is called a *nth*-order box set.

Definition 2.5. Let $A_n = A \times I_n,\ B_n = B \times I_n$ be *nth*-order box sets in $\mathbb{R}^\infty$. We define:

(1) $A_n \cup B_n = (A \cup B) \times I_n$,
(2) $A_n \cap B_n = (A \cap B) \times I_n$, and
(3) $B_n^c = B^c \times I_n$.

Definition 2.6. Let $\mathbb{R}_I^n = \mathbb{R}^n \times I_n$ and let $\mathfrak{B}[\mathbb{R}_I^n]$ be the Borel σ-algebra for $\mathbb{R}_I^n$, where the topology for $\mathbb{R}_I^n$ is defined via the following class of open sets:

$$\mathfrak{O}_n = \{U \times I_n :,\ U \text{ open in } \mathbb{R}^n\}.$$

For any $A \in \mathfrak{B}[\mathbb{R}^n]$, we define $\lambda_\infty(A)$ on $\mathbb{R}_I^n$ by:

$$\lambda_\infty(A_n) = \lambda_n(A) \times \prod_{i=n+1}^\infty \lambda_1(I) = \lambda_n(A),$$

Theorem 2.7. *$\lambda_\infty(\cdot)$ is a measure on $\mathfrak{B}[\mathbb{R}_I^n]$, equivalent to n-dimen sional Lebesgue measure on $\mathbb{R}^n$.*

Corollary 2.8. *The measure $\lambda_\infty(\cdot)$ is both translationally and rotationally invariant on $(\mathbb{R}_I^n, \mathfrak{B}[\mathbb{R}_I^n])$, for each $n \in \mathbb{N}$.*

Thus, we can construct a theory of Lebesgue measure on $\mathbb{R}_I^n$ that completely parallels that on $\mathbb{R}^n$. If we can extend $\lambda_\infty(\cdot)$ to $\mathfrak{B}[\mathbb{R}_I^\infty]$, we will obtain a measure with the following properties:

(1) $\mathbb{R}_I^n \in \mathfrak{B}(\mathbb{R}_I^\infty)$ for all n, so that $\lambda_\infty(\cdot)$ restricted to $\mathbb{R}_I^n$ is equivalent to $\lambda_n(\cdot)$;

(2) $\mathfrak{B}[\mathbb{R}_I^\infty]$ has a large number of open sets of finite measure, a property not shared by $\mathfrak{B}[\mathbb{R}^\infty]$ (with the standard topology on $\mathbb{R}^\infty$); and

(3) $\lambda_\infty(I_0) = 1$.

Thus, $\lambda_\infty(\cdot)$ would certainly qualify as the natural version of Lebesgue measure on $\mathbb{R}^\infty$.

2.1.2.1. **Definition of $\mathbb{R}_I^\infty$.** Recall that $\mathbb{R}^\infty$ is the set of all sequences $\mathbf{x} = (x_1, x_2, \ldots)$, such that $x_i \in \mathbb{R}$. The standard metric for $\mathbb{R}^\infty$ is defined by:

$$d(\mathbf{x}, \mathbf{y}) = \sum\nolimits_{n=1}^{\infty} \frac{1}{2^n} \frac{|x_n - y_n|}{1 + |x_n - y_n|}.$$

This metric makes $\mathbb{R}^\infty$ a complete translation invariant metric space.

Since $\mathbb{R}_I^n \subset \mathbb{R}_I^{n+1}$, we have an increasing sequence so we define $\hat{\mathbb{R}}_I^\infty$ by:

$$\hat{\mathbb{R}}_I^\infty = \lim_{n\to\infty} \mathbb{R}_I^n = \bigcup_{k=1}^{\infty} \mathbb{R}_I^k.$$

Let $\mathfrak{X}_1 = \hat{\mathbb{R}}_I^\infty$ and let τ_1 be the topology induced by the class of open sets $\mathfrak{O}$:

$$\mathfrak{O} = \bigcup_{n=1}^{\infty} \mathfrak{O}_n = \bigcup_{n=1}^{\infty} \{U \times I_n : U \text{ open in } \mathbb{R}^n\}.$$

Let $\mathfrak{X}_2 = \mathbb{R}^\infty \setminus \hat{\mathbb{R}}_I^\infty$ and let τ_2 be discrete topology on $\mathfrak{X}_2$ induced by the discrete metric such that, for $x, y \in \mathfrak{X}_2,\ x \neq y,\ d_2(x, y) = 1$ and for $x = y,\ d_2(x, y) = 0$.

Definition 2.9. We define $(\mathbb{R}_I^\infty, \tau)$ to be the coproduct $(\mathfrak{X}_1, \tau_1) \oplus (\mathfrak{X}_2, \tau_2)$, of $(\mathfrak{X}_1, \tau_1)$ and $(\mathfrak{X}_2, \tau_2)$, so that every open set in $(\mathbb{R}_I^\infty, \tau)$ is the disjoint union of two open sets $G_1 \bigcup G_2$, with G_1 in $(\mathfrak{X}_1, \tau_1)$ and G_2 in $(\mathfrak{X}_2, \tau_2)$.

It follows that $\mathbb{R}_I^\infty = \mathbb{R}^\infty$ as sets, but not as topological spaces. The following result shows that convergence in the τ-topology always implies convergence via the $\mathbb{R}^\infty$-metric.

Theorem 2.10. *Let $x \in \mathbb{R}_I^\infty$, and $\{y^{(k)}\} \subset \mathbb{R}_I^\infty$. If $\{y^{(k)}\}$ converges to x in the τ-topology, then $y^{(k)}$ converges to x in the $\mathbb{R}^\infty$-metric.*

Proof. Denote $x = (x_1, x_2, \ldots)$, $y^k = (y_1^{(k)}, y_2^{(k)}, \ldots)$. If $x \in \mathfrak{X}_2$, then $\{y^{(k)}\}$ converges to x in the τ_2 topology, and if $0 < \varepsilon < 1$, there is

$N \in \mathbb{N}$ such that $d_2(x, y^k) < \varepsilon$ for all $k \geq N$; hence $x_j = y_j^{(k)}$ for $j \in \mathbb{N}$, and $d(x, y^k) = 0 < \varepsilon$.
If however $x \in \mathfrak{X}_1$, then $\{y^{(k)}\}$ converges to x in the τ_1 topology, and $x \in \mathbb{R}_I^m$ for some $m \in \mathbb{N}$. Choose $N \in \mathbb{N}$ so large that $\frac{1}{N} < \frac{\varepsilon}{2}$ and $N > m$. Now, since the set

$$U_N = \left(x_1 - \tfrac{\epsilon}{4}, x_1 + \tfrac{\epsilon}{4}\right) \times \cdots \times \left(x_N - \tfrac{\epsilon}{4}, x_N + \tfrac{\epsilon}{4}\right) \times \left(\prod_{i=N+1}^{\infty} [-1, 1]\right)$$

is open in τ_1, there is $N_1 \in \mathbb{N}$ for which $k \geq N_1$ implies $y^{(k)} \in U_N$, so by the above equation we have that $y_j^{(k)} \in \left(x_j - \frac{\varepsilon}{4}, x_j + \frac{\varepsilon}{4}\right), j = 1, \ldots, N$, and $y_j^{(k)} \in [-1, 1]$ for $j \geq N + 1$. It follows that

$$\begin{aligned} d(x, y^{(k)}) &= \sum_{j=1}^{N} \frac{|x_j - y_j|}{2^j \left(1 + |x_j - y_j^{(k)}|\right)} + \sum_{j=N+1}^{\infty} \frac{|x_j - y_j|}{2^j \left(1 + |x_j - y_j^{(k)}|\right)} \\ &< \frac{\varepsilon}{2} \sum_{j=1}^{N} \frac{1}{2^j} + \sum_{j=N+1}^{\infty} \frac{1}{2^j} \\ &< \frac{\varepsilon}{2} + \frac{1}{2^N} < \frac{\varepsilon}{2} + \frac{1}{N} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Therefore $y^{(k)}$ converges to x in the $\mathbb{R}^\infty$-metric. □

2.1.2.2. **Definition of $\mathfrak{B}[\mathbb{R}_I^\infty]$.** In a similar manner, if $\mathfrak{B}[\mathbb{R}_I^n]$ is the Borel σ-algebra for $\mathbb{R}_I^n$, then $\mathfrak{B}[\mathbb{R}_I^n] \subset \mathfrak{B}[\mathbb{R}_I^{n+1}]$, so we can define $\hat{\mathfrak{B}}[\mathbb{R}_I^n]$ by:

$$\hat{\mathfrak{B}}[\mathbb{R}_I^n] = \lim_{n \to \infty} \mathfrak{B}[\mathbb{R}_I^n] = \bigcup_{k=1}^{\infty} \mathfrak{B}[\mathbb{R}_I^k].$$

Let $\mathfrak{B}[\mathbb{R}_I^\infty]$ be the smallest σ-algebra containing $\hat{\mathfrak{B}}[\mathbb{R}_I^n] \cup \mathcal{P}(\mathbb{R}^\infty \setminus \bigcup_{k=1}^{\infty} \mathbb{R}_I^k)$, where $\mathcal{P}(\cdot)$ is the power set. It is obvious that the class $\mathfrak{B}[\mathbb{R}_I^\infty]$ coincides with the Borel σ-algebra generated by the τ-topology on $\mathbb{R}_I^\infty$. From our definition of $\mathfrak{B}[\mathbb{R}_I^\infty]$, we see that $\mathfrak{B}[\mathbb{R}^\infty] \subset \mathfrak{B}[\mathbb{R}_I^\infty]$. To see that they are equal, it suffices to show that $\mathbb{R}_I^n \in \mathfrak{B}[\mathbb{R}^\infty]$. If

$$O_i^{(m)} = \mathbb{R}^{i-1} \times \left(-\tfrac{1}{2} - \tfrac{1}{m}, \tfrac{1}{2} + \tfrac{1}{m}\right) \times \prod_{k>i} \mathbb{R},$$

then

$$O_i = \bigcap_{m \in \mathbb{N}} O_i^{(m)} = \mathbb{R}^{i-1} \times \left[-\tfrac{1}{2}, \tfrac{1}{2}\right] \times \prod_{k>i} \mathbb{R}.$$

Finally, we have

$$\mathbb{R}_I^n = \bigcap_{i>n} O_i.$$

2.1.3. The Extension of $\lambda_\infty(\cdot)$ to $\mathbb{R}_I^\infty$. We know that $\lambda_\infty(\cdot)$ is a countably additive measure on $\mathfrak{B}[\mathbb{R}_I^n]$ for each $n \in \mathbb{N}$, but we cannot say the same for $\mathfrak{B}[\mathbb{R}_I^\infty]$. In this section, we provide a (constructive) extension of $\lambda_\infty(\cdot)$ to a countably additive measure on $\mathfrak{B}[\mathbb{R}_I^\infty]$.

Let

$$\begin{aligned}\Delta_0 =&\{K \times I_n \in \mathfrak{B}[\mathbb{R}_I^n] : n \in \mathbb{N},\, K \subset \mathbb{R}^n, \text{compact} \\ &\text{and}\, \lambda_\infty(K \times I_n) < \infty\},\end{aligned}$$

$$\Delta = \left\{ P_N = \bigcup_{i=1}^N K_i : N \in \mathbb{N};\ K_i \in \Delta_0 \,\text{and}\, \lambda_\infty\left(K_i \cap K_j\right) = 0,\ i \neq j \right\}.$$

Definition 2.11. If $P_N \in \Delta$, we define

$$\lambda_\infty(P_N) = \sum_{i=1}^N \lambda_\infty(K_i).$$

Since $P_N \in \mathfrak{B}[\mathbb{R}_I^n]$ for some n, and $\lambda_\infty(\cdot)$ is a measure on $\mathfrak{B}[\mathbb{R}_I^n]$, the next result follows:

Lemma 2.12. *If $P_{N_1},\ P_{N_2} \in \Delta$ then:*

(1) *If $P_{N_1} \subset P_{N_2}$, then $\lambda_\infty(P_{N_1}) \leq \lambda_\infty(P_{N_2})$.*
(2) *If $\lambda_\infty(P_{N_1} \cap P_{N_2}) = 0$, then $\lambda_\infty(P_{N_1} \cup P_{N_2}) = \lambda_\infty(P_{N_2}) + \lambda_\infty(P_{N_2})$.*

Definition 2.13. If $G \subset \mathbb{R}_I^\infty$ is any open set, we define:

$$\lambda_\infty(G) = \lim_{N\to\infty} \sup\left\{\lambda_\infty(P_N) :\ P_N \in \Delta,\ P_N \subset G, \right\}.$$

Theorem 2.14. *If $\mathfrak{O}$ is the class of open sets, we have:*

(1) $\lambda_\infty(\mathbb{R}_I^\infty) = \infty$.
(2) *If $G_1,\ G_2 \in \mathfrak{O},\ G_1 \subset G_2$, then $\lambda_\infty(G_1) \leq \lambda_\infty(G_2)$.*
(3) *If $\{G_k\} \subset \mathfrak{O}$, then*

$$\lambda_\infty(\bigcup_{k=1}^\infty G_k) \leq \sum_{k=1}^\infty \lambda_\infty(G_k).$$

(4) *If the* $\{G_k\} \subset \mathfrak{O}$ *and are disjoint, then*

$$\lambda_\infty(\bigcup_{k=1}^{\infty} G_k) = \sum_{k=1}^{\infty} \lambda_\infty(G_k).$$

Proof. The proof of (1) is standard. To prove (2), observe that

$$\{P_N \in \Delta : \; P_N \subset G_1\} \subset \left\{P'_N \in \Delta : \; P'_N \subset G_2\right\},$$

so that $\lambda_\infty(G_1) \le \lambda_\infty(G_2)$. To prove (3), let $P_N \subset \bigcup_{k=1}^{\infty} G_k, \; P_N \in \Delta$. Since P_N is compact, there is a finite number of the G_k which cover P_N, so that $P_N \subset \bigcup_{k=1}^{L} G_k$ for some $L \in \mathbb{N}$. Now, for each G_k, there is a $P_k \in \Delta, \; P_k \subset G_k$. Furthermore, as P_N is arbitrary, we can assume that $P_N \subset \bigcup_{k=1}^{L} P_k$. Since there is an n such that all $P_k \in \mathfrak{B}(\mathbb{R}_I^n)$ and λ_∞ is a measure on $\mathbb{R}_I^n$, we have that

$$\lambda_\infty(P_N) \le \sum_{k=1}^{L} \lambda_\infty(P_{N_k}),$$

so that

$$\lambda_\infty(P_N) \le \sum_{k=1}^{L} \lambda_\infty(P_{N_k}) \leqslant \sum_{k=1}^{L} \lambda_\infty(G_k) \leqslant \sum_{k=1}^{\infty} \lambda_\infty(G_k).$$

It follows that

$$\lambda_\infty\left(\bigcup_{k=1}^{\infty} G_k\right) \le \sum_{k=1}^{\infty} \lambda_\infty(G_k).$$

If the G_k are disjoint, observe that if $P_N \subset {P'}_M$,

$$\lambda_\infty({P'}_M) \ge \lambda_\infty(P_N) = \sum_{k=1}^{L} \lambda_\infty(P_{N_k}).$$

It follows that

$$\lambda_\infty(\bigcup_{k=1}^{\infty} G_k) \ge \sum_{k=1}^{L} \lambda_\infty(G_k).$$

This is true for all L so that this, combined with (3), gives our result. □

If F is an arbitrary compact set in $\mathfrak{B}[\mathbb{R}_I^\infty]$, we define

$$\lambda_\infty(F) = \inf\left\{\lambda_\infty(G) : \; F \subset G, \; G \text{ open}\right\}. \tag{2.1}$$

Remark 2.15. At this point we see the power of $\mathfrak{B}[\mathbb{R}_I^\infty]$. Unlike $\mathfrak{B}[\mathbb{R}^\infty]$, Eq. (2.1) is well defined for $\mathfrak{B}[\mathbb{R}_I^\infty]$ because it has a sufficient number of open sets of finite measure.

2.1.3.1. *Bounded Outer Measure.*

Definition 2.16. Let A be an arbitrary set in $\mathbb{R}_I^\infty$.

(1) The outer measure of A is defined by:

$$\lambda_\infty^*(A) = \inf\left\{\lambda_\infty(G) : A \subset G,\ G \text{ open}\right\}.$$

We let $\mathfrak{L}_0$ be the class of all A with $\lambda_\infty^*(A) < \infty$.

(2) If $A \in \mathfrak{L}_0$, we define the inner measure of A by

$$\lambda_{\infty,(*)}(A) = \sup\left\{\lambda_\infty(F) : F \subset A,\ F \text{ compact}\right\}.$$

(3) We say that A is a bounded measurable set if $\lambda_\infty^*(A) = \lambda_{\infty,(*)}(A)$, and define the measure of A, $\lambda_\infty(A)$, by $\lambda_\infty(A) = \lambda_\infty^*(A)$.

Theorem 2.17. *Let* A, B *and* $\{A_k\}$ *be arbitrary sets in* $\mathbb{R}_I^\infty$ *with finite outer measure.*

(1) $\lambda_{\infty,(*)}(A) \leq \lambda_\infty^*(A)$.

(2) *If* $A \subset B$ *then* $\lambda_\infty^*(A) \leq \lambda_\infty^*(B)$ *and* $\lambda_{\infty,(*)}(A) \leq \lambda_{\infty,(*)}(B)$.

(3) $\lambda_\infty^*(\bigcup_{k=1}^\infty A_k) \leq \sum_{k=1}^\infty \lambda_\infty^*(A_k)$.

(4) *If the* $\{A_k\}$ *are disjoint,* $\lambda_{\infty,(*)}(\bigcup_{k=1}^\infty A_k) \geq \sum_{k=1}^\infty \lambda_{\infty,(*)}(A_k)$.

Proof. The proofs of (1) and (2) are straightforward. To prove (3), let $\varepsilon > 0$ be given. Then, for each k, there exists an open set G_k such that $A_k \subset G_k$ and $\lambda_\infty(G_k) < \lambda_\infty^*(A_k) + \varepsilon 2^{-k}$. Since $\bigcup_{k=1}^\infty A_k \subset \bigcup_{k=1}^\infty G_k$, we have

$$\begin{aligned}
\lambda_\infty^*\left(\bigcup_{k=1}^\infty A_k\right) &\leqslant \lambda_\infty\left(\bigcup_{k=1}^\infty G_k\right) \\
&\leqslant \sum_{k=1}^\infty \lambda_\infty\left(G_k\right) \\
&\leqslant \sum_{k=1}^\infty \left[\lambda_\infty^*\left(A_k\right) + \tfrac{\varepsilon}{2^k}\right] \\
&= \sum_{k=1}^\infty \lambda_\infty^*\left(A_k\right) + \varepsilon.
\end{aligned}$$

Since ε is arbitrary, we are done.

To prove (4), let $F_1, F_2, \ldots, F_N$ be compact subsets of $A_1, A_2, \ldots, A_N$, respectively. Since the A_k are disjoint,

$$\lambda_{\infty,(*)}\left(\bigcup_{k=1}^{\infty} A_k\right) \geqslant \lambda_\infty\left(\bigcup_{k=1}^{N} F_k\right)$$
$$= \sum_{k=1}^{N} \lambda_\infty(F_k).$$

Thus,

$$\lambda_{\infty,(*)}\left(\bigcup_{k=1}^{\infty} A_k\right) \geq \sum_{k=1}^{N} \lambda_{\infty,(*)}(A_k).$$

Since N is arbitrary, we are done. □

The next two important theorems follow from the last one.

Theorem 2.18 (Regularity and Radon)**.** *If A has finite measure, then for every $\varepsilon > 0$ there exist a compact set F and an open set G such that $F \subset A \subset G$, with $\lambda_\infty(G \setminus F) < \varepsilon$.*

Proof. Let $\varepsilon > 0$ be given. Since A has finite measure, it follows from our definitions of $\lambda_{\infty,(*)}$ and λ_∞^* that there is a compact set $F \subset A$ and an open set $G \supset A$ such that

$$\lambda_\infty(G) < \lambda_\infty^*(A) + \tfrac{\varepsilon}{2} \quad \text{and} \quad \lambda_\infty(F) > \lambda_{\infty,(*)}(A) - \tfrac{\varepsilon}{2}.$$

Since $\lambda_\infty(G) = \lambda_\infty(F) + \lambda_\infty(G \setminus F)$, we have:

$$\lambda_\infty(G \setminus F) = \lambda_\infty(G) - \lambda_\infty(F) < (\lambda_\infty(A) + \tfrac{\varepsilon}{2}) - (\lambda_\infty(A) - \tfrac{\varepsilon}{2}) = \varepsilon.$$

□

Theorem 2.19 (Countable Additivity)**.** *If the family $\{A_k\}$ consists of disjoint sets with bounded measure and $A = \bigcup_{k=1}^{\infty} A_k$, with $\lambda_\infty^*(A) < \infty$, then $\lambda_\infty(A) = \sum_{k=1}^{\infty} \lambda_\infty(A_k)$.*

Proof. Since $\lambda_\infty^*(A) < \infty$, we have:

$$\lambda_\infty^*(A) \leqslant \sum_{k=1}^{\infty} \lambda_\infty^*(A_k) = \sum_{k=1}^{\infty} \lambda_{\infty,(*)}(A_k) \leqslant \lambda_{\infty,(*)}(A) \leqslant \lambda_\infty^*(A).$$

It follows that $\lambda_\infty(A) = \lambda_\infty^*(A) = \lambda_{\infty,(*)}(A)$, so that

$$\lambda_\infty(A) = \lambda_\infty\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \lambda_\infty(A_k).$$

□

The following results also hold for nonfinite measures and justifies our somewhat unorthodox approach.

Theorem 2.20. *Let $\{A_n\}$ be a countable family of sets with $\lambda_\infty(A_n) < \infty$ for all n.*

(1) *Then,*

$$\lambda_\infty(\liminf_{n\to\infty} A_n) \leqslant \liminf_{n\to\infty} \lambda_\infty(A_n),$$

$$\limsup_{n\to\infty} \lambda_\infty(A_n) \leqslant \lambda_\infty(\limsup_{n\to\infty} A_n)$$

and

$$\liminf_{n\to\infty} \lambda_\infty(A_n) \leqslant \limsup_{n\to\infty} \lambda_\infty(A_n).$$

(2) (Borel–Cantelli Lemma) *If*

$$\sum_{n=1}^{\infty} \lambda_\infty(A_n) < \infty,$$

then

$$\lambda_\infty(\limsup_{n\to\infty} A_n) = 0.$$

Proof. To prove (1), let

$$A = \limsup_{n\to\infty} A_n;$$

i.e., if we let

$$B_k = \bigcup_{m=k}^{\infty} A_m,$$

then

$$A = \bigcap_{k=1}^{\infty} B_k.$$

Now,

$$\bigcup_{m=1}^{\infty} A_m \supseteq \bigcup_{m=2}^{\infty} A_m \supseteq \bigcup_{m=3}^{\infty} A_m \supseteq \cdots .$$

It follows that

$$\lambda_\infty(A) = \lim_{k\to\infty} \lambda_\infty\left(\bigcup_{m=k}^{\infty} A_m\right).$$

We also have $\lambda_\infty\left(\bigcup_{m=k}^\infty A_m\right) \geq \lambda_\infty(A_m)$ for all $m \geq k$ and $k \in \mathbb{N}$. Hence $\lambda_\infty(A) \geq \limsup_{n\to\infty} \lambda_\infty(A_n)$. Therefore, it follows that

$$\limsup_{n\to\infty} \lambda_\infty\left(A_n\right) \leqslant \lambda_\infty\left(\limsup_{n\to\infty} A_n\right).$$

The other parts of the inequality follow the same lines.

To prove (2), we use the fact, proven above, that

$$\limsup_{n\to\infty} \lambda_\infty(A_n) \leqslant \lambda_\infty(\limsup_{n\to\infty} A_n).$$

It now follows that

$$\lambda_\infty(A) \leqslant \lambda_\infty(B_k) \leqslant \sum_{m=k}^\infty \lambda_\infty(A_m)$$

for all $k \in \mathbb{N}$. Thus, we must have $\lambda_\infty(\limsup_{n\to\infty} A_n) = 0$. □

2.1.3.2. *Unbounded Outer Measure.*

Definition 2.21. Let A be an arbitrary set in $\mathbb{R}_I^\infty$. We say that A is measurable if, for all bounded measurable sets $M \in \mathcal{L}_0$ (see Definition 2.16), $A \cap M \in \mathcal{L}_0$. In this case, we define $\lambda_\infty(A)$ by:

$$\lambda_\infty(A) = \sup\left\{\lambda_\infty(A \cap M) : M \subset \mathcal{L}_0\right\}.$$

We let $\mathcal{L}_I$ be the class of all measurable sets A.

Theorem 2.22. *Let A and $\{A_k\}$ be arbitrary sets in $\mathcal{L}_I$.*

(1) *If $\lambda_\infty^*(A) < \infty$, then $A \in \mathcal{L}_I$ if and only if $A \in \mathcal{L}_0$. In this case, $\lambda_\infty(A) = \lambda_\infty^*(A)$.*

(2) *$\mathcal{L}_I$ is closed under countable unions, countable intersections, differences, and complements.*

(3)

$$\lambda_\infty\left(\bigcup_{k=1}^\infty A_k\right) \leq \sum_{k=1}^\infty \lambda_\infty(A_k).$$

(4) *If $\{A_k\}$ are disjoint,*

$$\lambda_\infty\left(\bigcup_{k=1}^\infty A_k\right) = \sum_{k=1}^\infty \lambda_\infty(A_k).$$

Proof. The proofs are the same as for the bounded measure case. □

Theorem 2.23. *Let A be a $\mathcal{L}_I$-measurable set. Then there exists a Borel set F and a set N with $\lambda_\infty(N) = 0$ such that $A = F \cup N$.*

We close this section with the following additional result.

Theorem 2.24. *Let $\{A_k\}$ be a family of measurable sets.*

(1) *If $A_k \subset A_{k+1}$ for all k, then $\lambda_\infty(A_k)$ is a increasing function of k and*

$$\lambda_\infty\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{n\to\infty} \lambda_\infty\left(A_n\right).$$

(2) *If $A_{k+1} \subset A_k$ for all k and $\lambda_\infty(A_1) < \infty$, then $\lambda_\infty(A_k)$ is a decreasing function of k and*

$$\lambda_\infty\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim_{n\to\infty} \lambda_\infty\left(A_n\right).$$

Proof. To prove (1), let $A = \lim_{k\to\infty} A_k = \bigcup_{k=1}^{\infty} A_k$ and use the fact that

$$A = A_1 \cup \left[\bigcup_{k=1}^{\infty} \left(A_{k+1} \backslash A_k\right)\right],$$

is a disjoint union, to get

$$\lambda_\infty\left(A\right) = \lambda_\infty\left(A_1\right) + \sum_{k=1}^{\infty}\left[\lambda_\infty\left(A_{k+1}\right) - \lambda_\infty\left(A_k\right)\right] = \lim_{n\to\infty} \lambda_\infty\left(A_n\right).$$

To prove (2), we use a variation on the last approach, with

$$A_1 = A \cup \left[\bigcup_{k=1}^{\infty} \left(A_k \backslash A_{k+1}\right)\right],$$

to get

$$\begin{aligned} \lambda_\infty(A_1) &= \lambda_\infty(A) + \sum_{k=1}^{\infty}\left[\lambda_\infty(A_k) - \lambda_\infty(A_{k+1})\right] \\ &= \lambda_\infty(A) + \lambda_\infty(A_1) - \lim_{n\to\infty} \lambda_\infty(A_n). \end{aligned}$$

□

Thus, our reasonable version of $\lambda_\infty(\cdot)$ is a complete regular countably additive Radon measure on $\mathbb{R}_I^\infty = \mathbb{R}^\infty$. The construction is essentially the same one would use to construct Lebesgue measure on $\mathbb{R}^n$.

2.1.4. Equivalent Definition of $\lambda_\infty(\cdot)$. We now consider an equivalent definition of $\lambda_\infty(\cdot)$, that connects our definition with others and is useful for our proof that $\lambda_\infty(\cdot)$ is σ-finite. The following theorem provides a nice characterization of a measure.

Theorem 2.25. *Let X be a nonempty set and let $\mathcal{A}$ be a σ-algebra over X. A mapping $\mu : \mathcal{A} \to [0, \infty]$ is a measure on $\mathbb{X}$ if and only if*

(1) $\mu(\emptyset) = 0$.

(2) *If $A, B \in \mathcal{A}$, and $\mu(A \cap B) = 0$, then $\mu(A \cup B) = \mu(A) + \mu(B)$.*

(3) *If $\{B_n\} \subset \mathcal{A}$ and $B_n \subset B_{n+1}$, then*

$$\mu\left(\bigcup_{k=1}^{\infty} B_k\right) = \lim_{n\to\infty} \mu\left(B_n\right).$$

Proof. If μ is a measure, it is clear that the conditions are satisfied. Thus, we need to only prove that these conditions are sufficient. Since μ is nonnegative and finitely additive, it suffices to show that it is countably additive. Let $\{A_n\} \subset \mathcal{A}$ be disjoint. From

$$\bigcup_{k=1}^{n} A_k \subset \bigcup_{k=1}^{n+1} A_k, \quad \text{and} \quad \mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu\left(A_k\right),$$

if we let

$$B_n = \bigcup_{k=1}^{n} A_k,$$

we have $B_n \subset B_{n+1}$, so we can apply Theorem 2.24(1) to get

$$\lim_{n\to\infty} \mu\left(B_n\right) = \mu\left(\bigcup_{k=1}^{\infty} B_k\right) = \mu\left(\bigcup_{k=1}^{\infty} A_k\right)$$

and

$$\lim_{n\to\infty} \mu\left(B_n\right) = \lim_{n\to\infty} \mu\left(\bigcup_{k=1}^{n} A_k\right) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu\left(A_k\right) = \sum_{k=1}^{\infty} \mu\left(A_k\right).$$

Combining these two results proves the theorem. □

Definition 2.26. For each $m \in \mathbb{N}$, we define a measure μ_m on $\mathfrak{B}[\mathbb{R}_I^\infty]$ by

$$\mu_m(A) = \lambda_\infty(A \cap \mathbb{R}_I^m), \text{ for each } A \in \mathfrak{B}[\mathbb{R}^\infty],$$

and set

$$\mu(A) = \lim_{m\to\infty} \mu_m(A). \tag{2.2}$$

Theorem 2.27. *The mapping* $\mu : \mathfrak{B}[\mathbb{R}_I^\infty] \to [0, \infty]$ *is a measure.*

Proof. It is clear that the first two conditions of Theorem 2.25 are satisfied, so we need to only check the last one. Let $\{A_k\}$ be pairwise disjoint family of sets in $\mathfrak{B}[\mathbb{R}_I^\infty]$. Since $A_k \cap A_l = \emptyset$ unless $k = l$, we see that the same is true for $A_k \cap \mathbb{R}_I^m$ and $A_l \cap \mathbb{R}_I^m$. Fix $N \in \mathbb{N}$ and use the distributive property for sets and condition (2) of Theorem 2.25 (finite additivity for disjoint sets) to get

$$\lambda_\infty \left[\left(\bigcup_{k=1}^N A_k \right) \cap \mathbb{R}_I^m \right] = \sum_{k=1}^N \lambda_\infty \left(A_k \cap \mathbb{R}_I^m \right).$$

Since $A_k \cap \mathbb{R}_I^m \subset A_k \cap \mathbb{R}_I^{m+1}$, all the terms are increasing, and therefore, by condition (3) of Theorem 2.25

$$\mu \left(\bigcup_{k=1}^N A_k \right) = \lim_{m \to \infty} \mu_m \left(\bigcup_{k=1}^N A_k \right)$$
$$= \sum_{k=1}^N \lim_{m \to \infty} \mu_m \left(A_k \right) = \sum_{k=1}^N \mu \left(A_k \right).$$

Where the last equality follows from Eq. (2.2). If we now let $N \to \infty$, we are done. □

Corollary 2.28. *The completion of* μ *is equal to* $\lambda_\infty(\cdot)$*.*

Proof. It is easy to see that they agree on $\mathfrak{B}[\mathbb{R}_I^n]$, for all n. □

Remark 2.29. The above approach is easier, as it gets us a measure quickly. On the other hand, by making $\mathfrak{B}[\mathbb{R}_I^n]$ explicit, we see from the last section that the study of Lebesgue measure on $\mathbb{R}^\infty$ is almost as easy as the same study on $\mathbb{R}^n$. This approach was discovered independently by Gill and Zachary [GZ], in Chap. 8, in 2008. However, Theorem 2.27 and Corollary 2.28 (using a different approach) were first obtained by Yamasaki [YA1] in 1980. Unaware of Yamasaki's result, Kharazishvili independently obtained the same result in 1984. In 1991 Kirtadze and Pantsulaia [KP1] provided yet another approach leading to the same result (see also Pantsulaia [PA]). Finally, in 2007, Kirtadze and Pantsulaia [KP2] proved that if ν is the completion (Lebesgue extension) of μ, then ν is the unique (up to a multiplicative constant), regular σ-finite measure on $\mathbb{R}^\infty$, with $\nu\{[-\frac{1}{2}, \frac{1}{2}]^{\aleph_o}\} = 1$, having ℓ_1 as its maximal translation invariant group. Since it is known that every σ-finite Borel measure on a Fréchet space is a Radon measure [BO], we see that ν is Radon. On the other hand, $\mathbb{R}_I^\infty$ is not a Fréchet space; however, Theorem 2.18 shows that $\lambda_\infty(\cdot)$ is also a Radon measure.

The next two results follow from our new representation of $\lambda_\infty(\cdot)$. First recall that, $\mathbb{R}_I^\infty = \mathfrak{X}_1 \oplus \mathfrak{X}_2$, where $\mathfrak{X}_1 = \bigcup_{n=1}^\infty \mathbb{R}_I^n$ and $\mathfrak{X}_2 = \mathbb{R}_I^\infty \setminus \mathfrak{X}_1$.

Theorem 2.30. $\lambda_\infty \left[\mathbb{R}^\infty \setminus (\bigcup_{n=1}^\infty \mathbb{R}_I^n)\right] = \lambda_\infty \left[\mathbb{R}^\infty \setminus \mathfrak{X}_1\right] = 0.$

Proof. Let $\hat{\mu}_n$ be the (Lebesgue) extension of μ_n to $\mathcal{L}[\mathbb{R}_I^n]$. From Definition 2.6, we have $\mu_n(A) = \lambda_\infty(A \bigcap \mathbb{R}_I^n)$ and $\mu(A) = \lim_{n\to\infty} \mu_n(A)$, for all $A \in \mathbb{R}^\infty$, so that

$$\lambda_\infty \left[\left(\mathbb{R}^\infty \setminus \bigcup_{k=1}^\infty \mathbb{R}_I^k\right)\right] = \lim_{n\to\infty} \hat{\mu}_n \left[\left(\mathbb{R}^\infty \setminus \bigcup_{k=1}^\infty \mathbb{R}_I^k\right) \bigcap \mathbb{R}_I^n\right] = 0.$$

□

It now follows from our definition of $\mathfrak{X}_1$ that:

Corollary 2.31. $\lambda_\infty(\mathfrak{X}_2) = 0.$

Theorem 2.32. *There exists a family of sets* $\{A_k\} \subset \mathfrak{B}[\mathbb{R}_I^\infty]$ *with* $\lambda_\infty(A_k) < \infty$, *and a set* N *of measure zero such that:*

$$\mathbb{R}_I^\infty = \left(\bigcup_{k=1}^\infty A_k\right) \cup N, \tag{2.3}$$

so that $\lambda_\infty(\cdot)$ *is* σ*-finite.*

Proof. Since $\lambda_\infty(\cdot)$ is regular and concentrated on $\mathfrak{X}_1$, we can set $N = \mathfrak{X}_2$. To show Eq. (2.3) holds, let $\{x_k\}$ be the set of vectors in $\mathbb{R}^\infty$ with rational coordinates and let B_k be the unit cube with center x_k so that $\lambda_\infty(B_k) < \infty$ for all k and

$$\mathbb{R}_I^\infty = \bigcup_{k=1}^\infty B_k.$$

Let $A_k = B_k \setminus \mathfrak{X}_2$. □

Remark 2.33. Theorem 2.32 shows that λ_∞ is a σ-finite measure. (This result is known to hold for finite Radon measures on Fréchet spaces.) Since $\lambda_\infty(\mathfrak{X}_2) = 0$, it is clear that the support of λ_∞ is contained in $\mathfrak{X}_1 = \bigcup_{n=1}^\infty \mathbb{R}_I^n$. This result is a special case of a general result due to Yamasaki [YA].

Theorem 2.34 (Yamasaki)**.** *The support of* λ_∞ *is* ℓ_∞.

2.1.5. Translations. In this section we prove that ℓ_1 is the largest (dense) group of admissible translations for $\mathbb{R}_I^\infty$. This result was first proven by Yamasaki [YA].

Let $\mathfrak{T}_{\lambda_\infty}$ be the largest group of admissible translations for $\mathbb{R}_I^\infty$.

Theorem 2.35. *If* $A \in \mathfrak{B}[\mathbb{R}_I^\infty]$ *then* $\lambda_\infty(A - x) = \lambda_\infty(A)$ *if and only if* $\mathfrak{T}_{\lambda_\infty} = \ell_1$.

Proof. Since $\lambda_\infty(\mathfrak{X}_2) = 0$, it suffices to prove our result for $A \subset \mathfrak{B}[\mathfrak{X}_1]$. If $x = (x_i) \in \ell_1$ is fixed, then there is an N_1 such that, for all $n > N_1$, $|x_n| < \frac{1}{2}$, so that $x \in \mathbb{R}_I^n$, for all $n > N_1$.

Since $\mathfrak{X}_1 = \bigcup_{m=1}^\infty \mathbb{R}_I^m$, there is $N_2 \in \mathbb{N}$ such that, for all $n > N_2$, $A \in \mathfrak{B}[\mathbb{R}_I^n]$. Fix $n > \max\{N_1, N_2\}$ and let $A = A_n \times \prod_{k=n+1}^\infty I$, where $A_n \in \mathfrak{B}[\mathbb{R}_I^n]$ and let $X_n = (x_1, x_2, \ldots, x_n)$. Since λ_∞ is translation invariant on $\mathbb{R}_I^n$,

$$\begin{aligned}\lambda_\infty\left[A - x\right] &= \lambda_n\left[A_n - X_n\right] \cdot \prod_{k=n+1}^\infty \lambda\left\{\left[-\tfrac{1}{2}, \tfrac{1}{2}\right] \cap \left[-\tfrac{1}{2} - x_k, \tfrac{1}{2} - x_k\right]\right\} \\ &= \lambda_n\left[A_n\right] \cdot \prod_{k=n+1}^\infty \lambda\left\{\left[-\tfrac{1}{2}, \tfrac{1}{2}\right] \cap \left[-\tfrac{1}{2} - x_k, \tfrac{1}{2} - x_k\right]\right\} \\ &= \lambda_n\left[A_n\right] \cdot \prod_{k=n+1}^\infty (1 - |x_k|),\end{aligned}$$

Since $x \in \ell_1$ and $0 < (1 - |x_k|) \leq 1$, we have

$$\lim_{n\to\infty} \prod_{k=n+1}^\infty (1 - |x_k|) = 1,$$

so that $\lambda_\infty[A - x] = \lambda_\infty[A]$. It follows that this is true for every $x \in \ell_1$ and all $A \in \mathfrak{B}[\mathfrak{X}_1]$.

Now, suppose that $y \in \mathfrak{T}_{\lambda_\infty}$, so that $\lambda_\infty[A - y] = \lambda_\infty[A]$ for all $A \in \mathfrak{B}[\mathfrak{X}_1]$. Since $\mathfrak{B}[\mathbb{R}_I^n] \subset \mathfrak{B}[\mathfrak{X}_1]$ for all n, we must also have $\lambda_\infty[A - y] = \lambda_\infty[A]$ for all $A \in \mathfrak{B}[\mathbb{R}_I^n]$. In this case, with $Y_n = (y_1, y_2, \ldots, y_n)$,

$$\begin{aligned}\lambda_\infty\left[A - y\right] &= \lambda_n\left[A_n - Y_n\right] \cdot \prod_{k=n+1}^\infty \lambda\left\{\left[-\tfrac{1}{2}, \tfrac{1}{2}\right] \cap \left[-\tfrac{1}{2} - y_k, \tfrac{1}{2} - y_k\right]\right\} \\ &= \lambda_n\left[A_n\right] \cdot \prod_{k=n+1}^\infty (1 - |y_k|).\end{aligned}$$

This last result holds if $A = \left[-\frac{1}{2}, \frac{1}{2}\right]^{\aleph_0}$, so that $A_n = \prod_{k=1}^{n} [-\frac{1}{2}, \frac{1}{2}]$. Thus, $1 = \lim_{n\to\infty} \prod_{k=n+1}^{\infty} (1 - |y_k|)$. It follows that $\sum_{k=1}^{\infty} |y_k| < \infty$, so that $y \in \ell_1$. □

2.2. Measure on Banach Spaces

Introduction. In this section, we explore the advantages of our construction of $\lambda_\infty(\cdot)$ on $\mathbb{R}_I^\infty$ as it relates to a measure theory for separable Banach spaces.

2.2.1. Basis for a Banach Space. In what follows, $\mathcal{B}$ will denote a Banach space with a Schauder basis (see Sect. 1.2 of Chap. 1). Recall that.

Definition 2.36. A sequence $(e_n) \in \mathcal{B}$ is called a Schauder basis (S-basis) for $\mathcal{B}$ if $\|e_n\|_{\mathcal{B}} = 1$ and, for each $x \in \mathcal{B}$, there is a unique sequence (x_n) of scalars such that

$$x = \lim_{k\to\infty} \sum_{n=1}^{k} x_n e_n = \sum_{n=1}^{\infty} x_n e_n.$$

We restrict ourselves to Banach spaces with an S-basis in this book in order to avoid a pathology that never comes up in practice.

It is easy to see from the definition of a Schauder basis that, for any sequence (x_n) of scalars associated with a $x \in \mathcal{B}$, $\lim_{n\to\infty} x_n = 0$. Let $J_k = \left[-\frac{1}{2\ln(k+1)}, \frac{1}{2\ln(k+1)}\right]$ and define

$$J^n = \prod_{k=n+1}^{\infty} J_k, \quad J = \prod_{k=1}^{\infty} J_k.$$

Definition 2.37. Let $\{e_k\}$ be an S-basis for $\mathcal{B}$ and let $x = \sum_{n=1}^{\infty} x_n e_n$. Recall that $\mathcal{P}_n(x) = \sum_{k=1}^{n} x_k e_k$ and define $\mathcal{Q}_n x = (x_1, x_2, \ldots, x_n)$.

(1) We define $\mathcal{B}_J^n$ by:

$$\mathcal{B}_J^n = \{\mathcal{Q}_n(x) : x \in \mathcal{B}\} \times J^n,$$

with norm:

$$\|(x_k)\|_{\mathcal{B}_J^n} = \max_{1\le k\le n} \left\| \sum_{i=1}^{k} x_i e_i \right\|_{\mathcal{B}} = \max_{1\le k\le n} \|\mathcal{P}_n(x)\|_{\mathcal{B}}.$$

(2) Since $\mathcal{B}_J^n \subset \mathcal{B}_J^{n+1}$, we set $\mathcal{B}_J^\infty = \bigcup_{n=1}^\infty \mathcal{B}_J^n$.

We define $\mathcal{B}_J$ by

$$\mathcal{B}_J = \left\{ (x_1, x_2, \ldots) : \sum_{k=1}^\infty x_k e_k \in \mathcal{B} \right\} \subset \mathcal{B}_J^\infty$$

and define a norm on $\mathcal{B}_J$ by

$$\|x\|_{\mathcal{B}_J} = \sup_n \|\mathcal{P}_n(x)\|_{\mathcal{B}} = |||x\,|||_{\mathcal{B}}\,.$$

Let $\mathfrak{B}[\mathcal{B}_J^\infty]$ be the smallest σ-algebra containing $\mathcal{B}_J^\infty$ and define $\mathfrak{B}[\mathcal{B}_J] = \mathfrak{B}[\mathcal{B}_J^\infty] \cap \mathcal{B}_J$. By Theorem 1.61 of Chap. 1, we know that,

$$|||\,x|||_{\mathcal{B}} = \sup_n \left\| \sum_{k=1}^n x_k e_k \right\|_{\mathcal{B}} \tag{2.4}$$

is an equivalent norm on $\mathcal{B}$. The following lemma shows that every Banach space with an S-basis has a natural embedding in $\mathbb{R}_I^\infty$.

Lemma 2.38. *When $\mathcal{B}$ carries the equivalent norm (2.4), the operator*

$$T : (\mathcal{B}, |||\cdot|||_{\mathcal{B}}) \to (\mathcal{B}_J, ||\cdot||_{\mathcal{B}_J}),$$

defined by $T(x) = (x_k)$ is an isometric isomorphism from $\mathcal{B}$ onto $\mathcal{B}_J$.

Definition 2.39. We call $\mathcal{B}_J$ the canonical representation of $\mathcal{B}$ in $\mathbb{R}_I^\infty$.

Definition 2.40. With $\mathfrak{B}[\mathcal{B}_J] = \mathcal{B}_J \cap \mathfrak{B}[\mathcal{B}_J^\infty]$, we define the σ-algebra generated by $\mathcal{B}$, and associated with $\mathfrak{B}[\mathcal{B}_J]$ by:

$$\mathfrak{B}_J[\mathcal{B}] = \left\{ T^{-1}(A) \mid A \in \mathfrak{B}[\mathcal{B}_J] \right\} =: T^{-1} \left\{ \mathfrak{B}\left[\mathcal{B}_J\right] \right\}.$$

Remark 2.41. Since $\lambda_\infty(A_J^n) = 0$, for $A_J^n \in \mathfrak{B}[\mathcal{B}_J^n]$, with A_J^n compact, we see that $\lambda_\infty(\mathcal{B}_J^n) = 0$, $n \in \mathbb{N}$, so that $\lambda_\infty(\mathcal{B}_J) = 0$ for every Banach space with an S-basis. Thus, the restriction of λ_∞ to $\mathcal{B}_J$ will not induce a nontrivial measure on $\mathcal{B}$. For this, we use a variation of a method developed by Yamasaki [YA].

Definition 2.42. Define $\bar{\nu}_k$, $\bar{\mu}_k$ on $B \in \mathfrak{B}[\mathbb{R}]$ by

$$\bar{\nu}_k(B) = \frac{\lambda(B)}{\lambda(J_k)}, \quad \bar{\mu}_k(B) = \frac{\lambda(B \cap J_k)}{\lambda(J_k)}$$

and, for elementary sets $B = \prod_{k=1}^\infty B_k$, $B \in \mathfrak{B}[\mathcal{B}_J^n]$, define $\bar{\nu}_J^n$ by:

$$\bar{\nu}_J^n(B) = \prod_{k=1}^n \bar{\nu}_k(B_k) \times \prod_{k=n+1}^\infty \bar{\mu}_k(B_k)$$

Finally, we define ν_J^n to be the (Lebesgue) extension of $\bar{\nu}_J^n$ to all of $\mathcal{B}_J^n$ and define $\nu_J(B) = \lim_{n\to\infty} \nu_J^n(B)$, for all $B \in \mathfrak{B}[\mathcal{B}_J]$.

Theorem 2.43. *The family of measures $\{\nu_J^n\}$ is increasing and, for $m < n$, $\nu_J^n|_{\mathcal{B}_J^m} = \nu_J^m$. Furthermore, ν_J is σ-finite with its support concentrated in $\mathcal{B}_J^\infty$.*

Proof. If B is an elementary set of $\mathfrak{B}[\mathcal{B}_J^n]$, then

$$\nu_J^n(B) = \prod_{k=1}^{m} \nu_J(B_k) \times \prod_{k=m+1}^{n} \nu_k(B_k) \times \prod_{k=n+1}^{\infty} \mu_k(B_k)$$

and $B_k = J_k$ for $k > n$, so that $\prod_{k=n+1}^{\infty} \mu_k(B_k) = 1$.

Since $n > m$, if we restrict to $\mathcal{B}_J^m$, we see that $\nu_J^n|_{\mathcal{B}_J^m} = \nu_J^m$. From this, we also see that the family ν_J^n is increasing. Since ν_J^n is σ-finite for all n, it follows that ν_J is also. Thus, we only need to prove that ν_J is σ-additive. Suppose that $\{A_k\}$ is a disjoint family of subsets of $\mathfrak{B}[\mathcal{B}_J]$. Then, for each n

$$\nu_J^n\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \nu_J^n(A_k).$$

Since the family ν_J^n is increasing with respect to n, we see that

$$\nu_J\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{n\to\infty} \nu_J^n\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \lim_{n\to\infty} \nu_J^n(A_k) = \sum_{k=1}^{\infty} \nu_J(A_k).$$

□

Definition 2.44. If $\mathcal{B}$ is a Banach space with an S-basis and $A \in \mathfrak{B}_J[\mathcal{B}]$, we define $\lambda_{\mathcal{B}}(A) = \nu_J[T(A)]$.

Remark 2.45. Our choice of the family $\{J_n\}$ ensures that every Banach space with an S-basis can be embedded in $\mathbb{R}_I^\infty$ as a closed subspace of $\mathcal{B}_J^\infty$. It is clear that other families $\{J'_n\}$ will also produce a (slightly) different version of Lebesgue measure on $\mathcal{B}$. Thus, unlike λ_∞ on $\mathbb{R}_I^\infty$, $\lambda_{\mathcal{B}}$ is not unique. When the family $\{J_n\}$ is used, we call $\lambda_{\mathcal{B}}$ the canonical version of Lebesgue measure associated with $\mathcal{B}$. (To avoid confusion, the canonical version is the only one we will use in the book.)

2.3. Integrable Functions

In this section we discuss measurable functions for $\mathbb{R}_I^\infty$ and for separable Banach spaces with an S-basis. We then define the Lebesgue integral and extend the standard convergence and limit theorems to our setting. We will be brief because the proofs are close to the same ones for $\mathbb{R}^n$.

2.3.1. Measurable Functions. Since $\mathcal{B}_J \subset \mathbb{R}_I^\infty$, it suffices to discuss functions on $\mathbb{R}_I^\infty$. Let $x = (x_1, x_2, x_3, \dots) \in \mathbb{R}_I^\infty$, let

$$I_n = \prod_{k=n+1}^{\infty} \left[-\tfrac{1}{2}, \tfrac{1}{2}\right]$$

and let

$$h_n(x) = \bigotimes_{k=n+1}^{\infty} \chi_I(x_k),$$

where χ_I is the characteristic function for the interval $I = \left[-\frac{1}{2}, \frac{1}{2}\right]$.

Definition 2.46. Let $\mathcal{M}^n$ represent the class of measurable functions on $\mathbb{R}^n$. If $x \in \mathbb{R}_I^\infty$ and $f^n \in \mathcal{M}^n$, let $\bar{x} = (x_i)_{i=1}^n$, $\hat{x} = (x_i)_{i=n+1}^\infty$, define $f(x) = f^n(\bar{x}) \otimes h_n(\hat{x})$ and let

$$\mathcal{M}_I^n = \left\{ f(x) : f(x) = f^n(\bar{x}) \otimes h_n(\hat{x}),\ x \in \mathbb{R}_I^\infty \right\}.$$

Definition 2.47. A function $f : \mathbb{R}_I^\infty \to \mathbb{R}$ is said to be measurable and write $f \in \mathcal{M}$, if there is a sequence $\{f_n \in \mathcal{M}_I^n\}$ such that $\lim_{n\to\infty} f_n(x) \to f(x)$ λ_∞-(a.e.).

This definition highlights our requirement that all functions on infinite dimensional space must be constructively defined as finite dimensional limits. The existence of functions satisfying Definition 2.47 is not obvious, so we provide a proof below.

Theorem 2.48 (Existence)**.** *Suppose that $f : \mathbb{R}_I^\infty \to (-\infty, \infty)$ and $f^{-1}(A) \in \mathfrak{B}[\mathbb{R}_I^\infty]$ for all $A \in \mathfrak{B}[\mathbb{R}]$, then there exists a family of functions $\{f_n\}$, $f_n \in \mathcal{M}_I^n$, such that $f_n(x) \to f(x)$, λ_∞-(a.e.).*

Proof. It suffices to prove the result for $f \geq 0$. For $x \in \mathbb{R}_I^\infty$, $x = (x_i)_{i=1}^\infty$ and, for each $n \in \mathbb{N}$ and $k = 0, 1, 2, \dots, n2^n$, define

$$E_{n,k} = \left\{ (x_i)_{i=1}^{n2^n} : \tfrac{k}{2^n} < f(x) \leqslant \tfrac{(k+1)}{2^n} \right\}.$$

With $\bar{x} = (x_i)_{i=1}^{n2^n}$ and $\hat{x} = (x_i)_{i=n2^n+1}^{\infty}$, define $f_n(x)$ by:

$$f_n(x) = \frac{1}{2^n} \sum_{k=0}^{n2^n} k\chi_{E_{n,k}}(\bar{x}) \otimes h_{n'}(\hat{x}),$$

where $n' = n2^n + 1$. It is now easy to see that $f_n(x)$ converges to $f(x)$, λ_∞-(a.e.). $\square$

The following theorem shows that $\mathcal{M}$ inherits the vector lattice structure of measurable functions on $\mathbb{R}^n$, $n \in \mathbb{N}$. The proof is almost identical to the case for $\mathcal{M}^n$, so we omit it. (Recall that for any two functions, f, g, $(f \wedge g)(x) = \min\{f(x), g(x)\}$ and $(f \vee g)(x) = \max\{f(x), g(x)\}$.)

Theorem 2.49. *The space $\mathcal{M}$ is a vector lattice: if $a \in \mathbb{R}$ and let $f, g \in \mathcal{M}$, then*

(1) *af, $|f|$, fg and $f + g$ are in $\mathcal{M}$.*
(2) *$f \wedge g$, $f \vee g \in \mathcal{M}$.*
(3) *If $\{g_n\} \subset \mathcal{M}$, then*

$$\sup_n g_n,\ \inf_n g_n,\ \limsup_{n\to\infty} g_n, \text{ and } \liminf_{n\to\infty} g_n \in \mathcal{M}_I.$$

(4) *If $f \in \mathcal{M}$, and there is a set N with $\lambda_\infty(N) = 0$, such that $f(x) = g(x)$ for $x \in \mathbb{R}_I^\infty \setminus N$, then $g \in \mathcal{M}$.*
(5) *$f = f_+ - f_-$, where*

$$f_+ = \tfrac{|f|+f}{2},\ f_- = \tfrac{|f|-f}{2} \in \mathcal{M}.$$

It is natural to ask if pointwise convergence of measurable functions on compact sets is related to uniform convergence for $\lambda_{\mathcal{B}}(\cdot)$ on $\mathcal{B}$, in the same way as it is for $\lambda_n(\cdot)$ on $\mathbb{R}^n$. The following theorem shows that the answer is yes for both $\lambda_\infty(\cdot)$ on $\mathbb{R}^\infty$ and $\lambda_{\mathcal{B}}(\cdot)$ on $\mathcal{B}$. The proof is almost identical to that on $\mathbb{R}^n$, so we omit it. The important point is that both $\lambda_\infty(\cdot)$ and $\lambda_{\mathcal{B}}(\cdot)$ are inner regular (i.e., every measurable set of finite measure can be approximated by compact sets).

Theorem 2.50 (Egoroff's Theorem)**.** *Let $f_n : \mathcal{B}_J \to \mathbb{R}$ be measurable, let A be a measurable set with $0 < \lambda_{\mathcal{B}}[A] < \infty$, and suppose that $f_n \to f$, $\lambda_{\mathcal{B}}$-(a.e.) on A. Then, for each $\varepsilon > 0$, there exists a compact set B_ε such that*

(1) *$\lambda_{\mathcal{B}}[A \backslash B_\varepsilon] < \varepsilon$ and*
(2) *$f_n \to f$ uniformly on B_ε.*

2.3.2. Integration on $\mathbb{R}_I^\infty$. In this section we provide a constructive theory of integration on $\mathbb{R}_I^\infty$ using the known properties of integration on $\mathbb{R}_I^n$. This approach has the advantage that all the standard theorems for Lebesgue measure apply. (The proofs are the same as for integration on $\mathbb{R}^n$.)

2.3.3. L^1-Theory. Let $L^1[\mathbb{R}_I^n]$ be the class of integrable functions on $\mathbb{R}_I^n$. Since $L^1[\mathbb{R}_I^n] \subset L^1[\mathbb{R}_I^{n+1}]$, we define $L^1[\hat{\mathbb{R}}_I^\infty] = \bigcup_{n=1}^\infty L^1[\mathbb{R}_I^n]$. We say that a measurable function $f \in L^1[\mathbb{R}_I^\infty]$, if there is a Cauchy-sequence $\{f_n\} \subset L^1[\hat{\mathbb{R}}_I^\infty]$, such that $\lim_{n\to\infty}\|f_n - f\|_1 = 0$.

Definition 2.51. If $f \in L^1[\mathbb{R}_I^\infty]$, we define the integral of f by:

$$\int_{\mathbb{R}_I^\infty} f(x)d\lambda_\infty(x) = \lim_{n\to\infty}\int_{\mathbb{R}_I^\infty} f_n(x)d\lambda_\infty(x).$$

Theorem 2.52. *If $f \in L^1[\mathbb{R}_I^\infty]$, then the above integral exists.*

Proof. The proof follows from the fact that the sequence in the definition of f is L^1-Cauchy. □

Let $\mathbb{C}_c(\mathbb{R}_I^n)$ be the class of continuous functions on $\mathbb{R}_I^n$ which vanish outside compact sets. We say that a measurable function $f \in \mathbb{C}_c(\mathbb{R}_I^\infty)$, if there is a Cauchy-sequence $\{f_n\} \subset \bigcup_{n=1}^\infty \mathbb{C}_c(\mathbb{R}_I^n) = \mathbb{C}_c(\hat{\mathbb{R}}_I^\infty)$, such that

$$\lim_{n\to\infty}\|f_n - f\|_\infty = 0.$$

We define $\mathbb{C}_0(\mathbb{R}_I^\infty)$, the functions that vanish at ∞, in the same manner.

Lemma 2.53. *If $f \in C_c(\mathbb{R}_I^\infty)$* or *$C_0(\mathbb{R}_I^\infty)$, then f is continuous in the supremum norm.*

Proof. Let $f(\mathbf{x}) \in C_c(\mathbb{R}_I^\infty)$ and let $\{\mathbf{x}_n \ : n \in \mathbb{N}\}$ be any sequence, with $\mathbf{x}_n \in \mathbb{R}_I^n$ such that $\mathbf{x}_n \to \mathbf{x}$ as $n \to \infty$.

Let $\varepsilon > 0$ is given. Since $f \in C_c(\mathbb{R}_I^\infty)$, there is a sequence of functions $\{g_n\}$, with $g_n \in \mathbb{C}_c(\mathbb{R}_I^n)$, for each n such that, $g_n \to f$ in the supremum norm on $\mathbb{C}_c(\hat{\mathbb{R}}_I^\infty)$.

Since the support of f is contained in a compact set, we can choose K_1 such that, for $k \geq K_1$, $|g_k(\mathbf{x}_n) - f(\mathbf{x}_n)| < \frac{\varepsilon}{3}$. We can also choose K_2 such that, for $k \geq K_2$, $|g_k(\mathbf{x}) - f(\mathbf{x})| < \frac{\varepsilon}{3}$. Let $k = \max\{K_1, K_2\}$

be fixed. Since $g_k \in \mathbb{C}_c(\mathbb{R}_I^k)$, it is continuous. Thus, we can choose N such that, for $n \geq N$, $|g_k(\mathbf{x}_n) - g_k(\mathbf{x})| < \frac{\varepsilon}{3}$. We now have:

$$|f(\mathbf{x}_n) - f(\mathbf{x})| \leq |g_k(\mathbf{x}_n) - f(\mathbf{x}_n)| + |g_k(\mathbf{x}) - g_k(\mathbf{x}_n)| + |g_k(\mathbf{x}) - f(\mathbf{x})| < \varepsilon.$$

The same proof applies to $C_0(\mathbb{R}_I^\infty)$ □

Theorem 2.54. $\mathbb{C}_c(\mathbb{R}_I^\infty)$ *is dense in* $L^1[\mathbb{R}_I^\infty]$.

Proof. We prove this result in the standard manner, by reducing the proof to positive simple functions and then to one characteristic function and finally using the approximation theorem to approximate a measurable set which contains a closed set and is contained in an open set. The details are left to the reader. □

In a similar fashion we can define the L^p spaces, $1 < p < \infty$.

2.3.4. Integration on $\mathcal{B}$. In this section, we discuss integration on a Banach space $\mathcal{B}$ with an S-basis. Even if a reasonable theory of Lebesgue measure exists on $\mathcal{B}$, this is not sufficient to make it a useful mathematical tool. In addition, all the theory developed for finite-dimensional analysis, differential operators, Fourier transforms, etc. are also required. Furthermore, applied researchers need operational control over the convergence properties of these tools. In particular, one must be able to approximate an infinite-dimensional problem as a natural limit of the finite-dimensional case in a manner that lends itself to computational implementation. This implies that a useful approach also has a well-developed theory of convergence for infinite sums and products. We will not be able to address all of these desired qualities in this chapter. However, we begin with a few desired qualities for such an integral:

(1) Since λ_∞ restricted to $\mathfrak{B}[\mathbb{R}_I^n]$ is equivalent to λ_n (and $\lambda_{\mathcal{B}}$ restricted to $\mathfrak{B}[\mathcal{B}_J^n]$ is equivalent to λ_n), we require that the integral restricted to $\mathfrak{B}[\mathbb{R}_I^n]$ or $\mathfrak{B}[\mathcal{B}_J^n]$ be the integral on $\mathbb{R}^n$.
(2) If $f(x) \in \mathcal{M}$ (the measurable functions on $\mathbb{R}_I^\infty$), but $f(x) \notin \mathcal{M}_I^n$, for all n, then the integral, if it exists, must be the limit of the integrals of the sequence of functions $\{f_n(x)\}$, where $f_n(x) \in \mathcal{M}_I^n$ and $f_n(x) \to f(x)$ λ_∞-(a.e.).

When there is no chance for confusion, we will identify $\mathcal{B}$ with its canonical representation $\mathcal{B}_J$ in $\mathbb{R}_I^\infty$ and omit the subscript J.

Definition 2.55. Let $f : \mathcal{B} \to [0, \infty]$ be a measurable function and let $\lambda_{\mathcal{B}}$ be constructed using the family $\{J_k\}$. If $\{s_n\} \subset \mathcal{M}$ is a increasing family of nonnegative simple functions with $s_n \in \mathcal{M}_I^n$, for each n and $\lim_{n\to\infty} s_n(x) = f(x)$, $\lambda_{\mathcal{B}}$-(a.e.), we define the integral of f over $\mathcal{B}$ by:

$$\int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}} = \lim_{n\to\infty} \int_{\mathcal{B}} \left[s_n(x) \prod_{i=1}^{n} \lambda(J_i) \right] d\lambda_{\mathcal{B}}(x).$$

Recall that, if $g \in \mathcal{M}_I^n$, then $g(x) = g^n(\bar{x}) \otimes h_n(\hat{x})$, where $\bar{x} = (x_1, x_2, \ldots, x_n)$, $g^n(\bar{x}) \in \mathcal{M}^n$, $\hat{x} = (x_{n+1}, \ldots)$ and $h_n(\hat{x}) = \prod_{k=n+1}^{\infty} \chi_I(x_k)$. From our definition of the integral, it is easy to see that, for each fixed n

$$\int_{\mathcal{B}} \left[s_n(x) \prod_{i=1}^{n} \lambda(J_i) \right] d\lambda_{\mathcal{B}}(x) = \int_{\mathbb{R}^n} s_n^n(\bar{x}) d\lambda_n(\bar{x}).$$

Since the family of integrals is increasing, the limit always exists.

Theorem 2.56. *If $f : \mathcal{B} \to [0, \infty]$ is a measurable function, then the value of the integral $\int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x)$ is independent of the sequence of functions $\{f_m \in \mathcal{M}_I^m\}$ used to define it.*

Proof. The proof is easy if $\int_{\mathcal{B}} f \, d\lambda_{\mathcal{B}} = \infty$, so assume it is finite. Let $\{f_m\}$ be a nonnegative increasing sequence for which $f_m \in \mathcal{M}_I^m$ and $\lim_{m\to\infty} f_m = f$, $\lambda_{\mathcal{B}}$-(a.e.) and let

$$\beta = \sup \left\{ \int_{\mathcal{B}} \left[s \prod_{k=1}^{n} \lambda(J_k) \right] d\lambda_{\mathcal{B}} \; : \; 0 \le s \le f, \; s \in \mathcal{M}_I^n \right\}. \tag{2.5}$$

Choose $m_1 \in \mathbb{N} : m_1 \ge n$ and $f_{m_1} \ge s$. Since the integrals are increasing, we have

$$\begin{aligned}\int_{\mathcal{B}} \left[s \prod_{k=1}^{n} \lambda(J_k) \right] d\lambda_{\mathcal{B}} &\le \int_{\mathcal{B}} \left[f_{m_1} \prod_{k=1}^{m_1} \lambda(J_k) \right] d\lambda_{\mathcal{B}} \\ &\le \lim_{m\to\infty} \int_{\mathcal{B}} \left[f_m \prod_{k=1}^{m} \lambda(J_k) \right] d\lambda_{\mathcal{B}} = \int_{\mathcal{B}} f \, d\lambda_{\mathcal{B}}.\end{aligned}$$

Since the s is arbitrary, $\beta \le \int_{\mathcal{B}} f d\lambda_{\mathcal{B}}$. Since $\int_{\mathcal{B}} f \, d\lambda_{\mathcal{B}} < \infty$, let $\epsilon > 0$, and choose $n_0 \in \mathbb{N}$ such that

$$\int_{\mathcal{B}} f d\lambda_{\mathcal{B}} - \int_{\mathcal{B}} \left[f_{n_0} \prod_{k=1}^{n_0} \lambda(J_k) \right] d\lambda_{\mathcal{B}} < \varepsilon,$$

so that

$$\int_{\mathcal{B}} f d\lambda_{\mathcal{B}} < \int_{\mathcal{B}} \left[f_{n_0} \prod_{k=1}^{n_0} \lambda(J_k) \right] d\lambda_{\mathcal{B}} + \varepsilon \leqslant \beta + \varepsilon.$$

Since this is true for all $\varepsilon > 0$, $\int_{\mathcal{B}} f d\lambda_{\mathcal{B}} \leq \beta$. It follows that

$$\lim_{m\to\infty} \int_{\mathcal{B}} \left[f_m \prod_{k=1}^{m} \lambda(J_k) \right] d\lambda_{\mathcal{B}}$$
$$= \sup_n \left\{ \int_{\mathcal{B}} \left[s \prod_{k=1}^{n} \lambda(J_k) \right] d\lambda_{\mathcal{B}} : 0 \leqslant s \leqslant f, \quad s \in \mathcal{M}_I^n \right\},$$

so that the value of the integral does not depend on the sequence $\{f_m\}$ chosen. □

Theorem 2.57. *If f, g are nonnegative measurable functions and $0 \leq c < \infty$, we have:*

(1) *If $f \geq 0$, then $0 \leq \int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) \leq \infty$;*

(2) $\int_{\mathcal{B}} cf(x) d\lambda_{\mathcal{B}}(x) = c \int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x)$;

(3)

$$\int_{\mathcal{B}} [f(x) + g(x)] d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) + \int_{\mathcal{B}} g(x) d\lambda_{\mathcal{B}}(x);$$

(4) *If $f \leq g$,* then $\int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) \leq \int_{\mathcal{B}} g(x) d\lambda_{\mathcal{B}}(x)$.

2.3.5. Probability Measures on $\mathcal{B}$. If we replace our Lebesgue measure λ_∞, by the infinite product Gaussian measure μ_∞, on $\mathbb{R}_I^\infty$, we obtain a countably additive probability measure, but $\mu_\infty(\ell_2) = 0$. A discussion of this and related issues can be found in Dunford and Schwartz [DS] (see p. 402). They show that, by using the standard projection method onto finite dimensional subspaces, to construct a probability measure directly on ℓ_2 leads to a rotationally invariant measure, which is no longer countable additive. The resolution of this problem led to the development of the Wiener measure [WSRM]. However, another approach via Fourier transforms (or characteristic functions) does allow one to induce a unique countable additive Gaussian measure on ℓ_2, which is the restriction of μ_∞ (see De Prato [DP]). In this section, we take a general approach to the construction of countable additive probability measures on any Banach space $\mathcal{B}$ with an S-basis.

Let ν be any probability measure on $\mathfrak{B}[\mathbb{R}]$, with density f. For each n, each $x \in \mathcal{B}_J^n$, and each $A \in \mathfrak{B}_J[\mathcal{B}]$, define $f_{\mathcal{B}}^n(x)$ by:

$$f_{\mathcal{B}}^n(x) = \left(\bigotimes_{k=1}^{n} f(x_k)\right) \otimes \left(\bigotimes_{k=n+1}^{\infty} \chi_I(x_k)\right)$$

and define $\bar{\nu}_J^n$ on $\mathfrak{B}_J[\mathcal{B}]$ by:

$$\bar{\nu}_J^n(A) = \int_{T(A)\cap\mathcal{B}_J^n} f_{\mathcal{B}}^n(x) d\lambda_{\mathcal{B}}(x),$$

where T is the isometric isomorphism between $\mathcal{B}$ and $\mathcal{B}_J$. Finally, let ν_J^n to be the completion of $\bar{\nu}_J^n$ on $\mathfrak{B}_J[\mathcal{B}]$. Clearly, each member of the family $\{\nu_J^n\}$ defines a probability measure on $\mathcal{B}$, which is absolutely continuous with respect to $\lambda_{\mathcal{B}}$. Furthermore, the restriction of ν_J^n to $\mathcal{B}_J^m$, $m \leq n$, is a probability measure on $\mathbb{R}^m$, which is absolutely continuous with respect to λ_m.

If the sequence of functions $\{f_{\mathcal{B}}^n(\cdot)\}$ converges to a density $f_{\mathcal{B}}(\cdot)$, then

$$\nu_{\mathcal{B}}(A) = \int_{T(A)} f_{\mathcal{B}}(x) d\lambda_{\mathcal{B}}(x), \ A \in \mathfrak{B}_J[\mathcal{B}],$$

is also a probability measure on $\mathcal{B}$, which is absolutely continuous with respect to $\lambda_{\mathcal{B}}$. Our construction is general, but the existence of a limit density is more delicate. Since our family of measures form an inductive system on the family $\{\mathcal{B}_J^n\}$, we can always apply the Daniell–Kolmogorov Extension Theorem, to obtain a probability measure ν on $\mathcal{B}_J^\infty$ (see [YA]). However, this measure need not be full on $\mathcal{B}_J \subset \mathcal{B}_J^\infty$, $\nu[\mathcal{B}_J] = 1$. Even if it is full on $\mathcal{B}_J$, it still need not be absolutely continuous with respect to Lebesgue measure.

We consider two examples.

Example 2.58. *The standard Gaussian measure on* $\mathbb{R}$ *is defined by:*

$$d\mu(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{|x|^2}{2}\right\} d\lambda_1(x).$$

Let $[\sqrt{2\pi}]y = x$. *Then* $d\lambda_1(x) = [\sqrt{2\pi}]d\lambda_1(y)$, *so that* $d\mu$ *becomes:*

$$d\mu^*(y) = \exp\left\{-\pi\,|y|^2\right\} d\lambda_1(y)$$

We call μ^* *the canonical representation of Gaussian measure* μ.

Let $\nu = \mu^*$ in our construction and let $A \in \mathfrak{B}_J[\mathcal{B}]$. Then

$$\begin{aligned}\nu_{\mathcal{B}}^n(A) &= \int_{T(A)} d\nu_{\mathcal{B}}^n(x) \\ &= \int_{T(A)} e^{-\pi \sum_{k=1}^n |x_k|^2} \otimes \left(\bigotimes_{k=n+1}^{\infty} \chi_I(x_k) \right) d\lambda_B(x) \\ &= \int_{A_n} e^{-\pi \sum_{k=1}^n |x_k|^2} d\lambda_n(\bar{x}),\end{aligned}$$

where $A_n = T(A) \cap \mathbb{R}^n$ and $\bar{x} = (x_1, \dots, x_n)$. It follows that,

$$d\nu_{\mathcal{B}}^n(x) = e^{-\pi \sum_{k=1}^n |x_k|^2} \otimes \left(\bigotimes_{k=n+1}^{\infty} \chi_I(x_k) \right) d\lambda_B(x)$$

or

$$d\nu_{\mathcal{B}}^n(x) = e^{-\pi \sum_{k=1}^n |x_k|^2} d\lambda_n(\bar{x})$$

on $\mathfrak{B}[\mathcal{B}_J^n]$, for all $n \in \mathbb{N}$. However, the sequence

$$\left\{ \exp\left\{ -\pi \sum_{k=1}^n |x_k|^2 \right\} \otimes \left(\bigotimes_{k=n+1}^{\infty} \chi_I(x_k) \right) \right\}$$

converges if and only if $\mathcal{B}$ is related to ℓ_p, $1 \leq p \leq 2$, in the sense that $\sum_{n=1}^{\infty} |x_n|^2 < \infty$, for all $x = (x_1, x_2, \dots) \in \mathcal{B}_J$. In this case we can write: $d\nu_{\mathcal{B}}(x) = \exp\left\{ -\pi \sum_{k=1}^{\infty} |x_k|^2 \right\} d\lambda_{\mathcal{B}}(x)$.

Definition 2.59. If $\nu = \mu^*$ we call $\nu_{\mathcal{B}}(\cdot)$ the universal representation of Gaussian measure on $\mathcal{B}$.

Example 2.60. *If $f(y) = \frac{1}{\pi} \frac{1}{1+y^2}$ is the density for the Cauchy distribution, make the change of variables $y = \pi x$. In this case, $f(x) = \frac{1}{1+[\pi x]^2}$. If we set*

$$f_{\mathcal{B}}^n(x) = \left(\bigotimes_{k=1}^n f(x_k) \right) \otimes \left(\bigotimes_{k=n+1}^{\infty} \chi_I(x_k) \right),$$

It is easy to check that for $n > m$,

$$|f_{\mathcal{B}}^n(x) - f_{\mathcal{B}}^m(x)| \leqslant \left| \bigotimes_{k=1}^m f(x_k) \right| \sum_{k=m+1}^n |\pi x_k|^2 \leqslant \pi^2 \sum_{k=m+1}^n |x_k|^2.$$

It follows that $f_{\mathcal{B}}^n(x)$ converges if and only if $\mathcal{B}$ is related to ℓ_p, $1 \leq p \leq 2$, in the sense that $\sum_{n=1}^{\infty} |x_n|^2 < \infty$, for all $x = (x_1, x_2, \dots) \in \mathcal{B}_J$.

It is not hard to see that, for each probability measure ν defined on $\mathfrak{B}[\mathbb{R}]$, with a density $f(x)$, we can construct a corresponding family of measures $\{\nu_J^n\}$ on $\{\mathfrak{B}[\mathcal{B}_J^n]\}$, which are absolutely continuous with respect to $\lambda_{\mathcal{B}}$.

2.3.5.1. *Limit Theorems.* In this section, we consider the standard limit theorems of analysis. The proofs for $\mathbb{R}_I^\infty$ are the same as for $\mathbb{R}^n$, so we only need proofs for $\mathcal{B}$.

Theorem 2.61. (Fatou's Lemma) *Let $\{f^i\} \subset \mathcal{M}$ be a nonnegative family of functions, then:*

$$\int_{\mathcal{B}} \left(\liminf_{i\to\infty} f^i(x) \right) d\lambda_{\mathcal{B}}(x) \leq \liminf_{i\to\infty} \int_{\mathcal{B}} f^i(x) d\lambda_{\mathcal{B}}(x).$$

Proof. Let $f(x) = \liminf_{i\to\infty} f^i(x)$ and let $\{s_n\}$ is an increasing sequence of integrable simple functions with $s_n(x) \to f(x)$, $\lambda_{\mathcal{B}}$-(a.e.). By definition of the integral on $\mathcal{B}$, if $\varepsilon > 0$ is given, there is $N = N(\varepsilon) \in \mathbb{N}$ such that, for all $n > N$,

$$\begin{aligned} \int_{\mathcal{B}} s_n(x) \left[\prod_{k=1}^n \lambda(J_k) \right] d\lambda_{\mathcal{B}}(x) - \varepsilon &\leqslant \int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) \\ &\leqslant \int_{\mathcal{B}} s_n(x) \left[\prod_{k=1}^n \lambda(J_k) \right] d\lambda_{\mathcal{B}}(x) + \varepsilon. \end{aligned}$$

For each i, let $\{s_m^i\}$ be an increasing sequence of integrable simple functions such that

$$\lim_{m\to\infty} s_m^i(x) = f^i(x),\ \lambda_{\mathcal{B}} - (\text{a.e.}).$$

By Fatou's Lemma for Lebesgue measure on $\mathbb{R}^m$,

$$\int_{\mathcal{B}} s_m(x) \left[\prod_{k=1}^m \lambda(J_k) \right] d\lambda_{\mathcal{B}}(x) \leqslant \liminf_{i\to\infty} \int_{\mathcal{B}} s_m^i(x) \left[\prod_{k=1}^m \lambda(J_k) \right] d\lambda_{\mathcal{B}}(x).$$

It follows that, for $m > N$,

$$\int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) \leqslant \liminf_{i\to\infty} \int_{\mathcal{B}} s_m^i(x) \left[\prod_{k=1}^m \lambda(J_k) \right] d\lambda_{\mathcal{B}}(x) + \varepsilon.$$

Since the left-hand side is independent of m, we see that

$$\int_{\mathcal{B}} \left(\liminf_{i\to\infty} f^i(x) \right) d\lambda_{\mathcal{B}}(x) \leq \liminf_{i\to\infty} \int_{\mathcal{B}} f^i(x) d\lambda_{\mathcal{B}}(x).$$

□

Theorem 2.62 (Monotone Convergence Theorem)**.** *Let* $\{f_n\} \subset \mathcal{M}$ *be a nonnegative family of functions, with* $f_n(x) \leq f_{n+1}(x)$*. Then*

$$\lim_{n\to\infty} \int_{\mathcal{B}} f_n(x) d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} \left(\lim_{n\to\infty} f_n(x) \right) d\lambda_{\mathcal{B}}(x).$$

Proof. Since $f_n(x) \leq f_{n+1}(x)$, the limit exists, which we denote by $f(x)$. Since $f_n(x) \nearrow f(x)$, we can use Fatou's Lemma to see that

$$\begin{aligned} \int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) &= \int_{\mathcal{B}} \left(\lim_{n\to\infty} f_n(x) \right) d\lambda_{\mathcal{B}}(x) \\ &= \int_{\mathcal{B}} \left(\liminf_{n\to\infty} f_n(x) \right) d\lambda_{\mathcal{B}}(x) \\ &\leq \liminf_{n\to\infty} \int_{\mathcal{B}} f_n(x) d\lambda_{\mathcal{B}}(x) \\ &= \lim_{n\to\infty} \int_{\mathcal{B}} f_n(x) d\lambda_{\mathcal{B}}(x). \end{aligned}$$

On the other hand,

$$\begin{aligned} \lim_{n\to\infty} \int_{\mathcal{B}} f_n(x) d\lambda_{\mathcal{B}}(x) \geqslant \int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) \\ \geqslant \sup_n \int_{\mathcal{B}} f_n(x) d\lambda_{\mathcal{B}}(x) \\ = \lim_{n\to\infty} \int_{\mathcal{B}} f_n(x) d\lambda_{\mathcal{B}}(x). \end{aligned}$$

so that

$$\lim_{n\to\infty} \int_{\mathcal{B}} f_n(x) d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} \left(\lim_{n\to\infty} f_n(x) \right) d\lambda_{\mathcal{B}}(x).$$

□

Definition 2.63. If $f \in \mathcal{M}$, we define

$$\int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} f_+(x) d\lambda_{\mathcal{B}}(x) - \int_{\mathcal{B}} f_-(x) d\lambda_{\mathcal{B}}(x),$$

and say that f is integrable whenever both integrals on the right are finite. The set of all integrable functions is denoted by $L^1[\mathcal{B}, \mathfrak{B}[\mathcal{B}], \lambda_{\mathcal{B}}] = L^1[\mathcal{B}]$.

Theorem 2.64 (Dominated Convergence Theorem)**.** *Let* $\{g\}, \{f_n\} \subset \mathcal{M}$*, with* $g \geq 0$*,* $g \in L^1(\mathcal{B})$ *and* $|f_n(x)| \leq g(x)$*,* $\lambda_{\mathcal{B}}$-(a.e.)*. If* $\lim_{n\to\infty} f_n(x)$ *exists* $\lambda_{\mathcal{B}}$-(a.e.)*, then* $\lim_{n\to\infty} f_n \in L^1[\mathcal{B}]$ *and*

$$\lim_{n\to\infty} \int_{\mathcal{B}} f_n(x) d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} \left(\lim_{n\to\infty} f_n(x) \right) d\lambda_{\mathcal{B}}(x).$$

Proof. Consider the nonnegative functions $\{g(x)+f_n(x)\}$. By Fatou's Lemma,

$$\begin{aligned}&\int_{\mathcal{B}} (g(x) + f(x))d\lambda_{\mathcal{B}}(x)\\ &\leqslant \liminf_{n\to\infty} \int_{\mathcal{B}} (g(x) + f_n(x))d\lambda_{\mathcal{B}}(x)\\ &= \int_{\mathcal{B}} g(x)d\lambda_{\mathcal{B}}(x) + \liminf_{n\to\infty} \int_{\mathcal{B}} f_n(x)d\lambda_{\mathcal{B}}(x)\end{aligned}$$

and so

$$\int_{\mathcal{B}} f(x)d\lambda_{\mathcal{B}}(x) \leqslant \liminf_{n\to\infty} \int_{\mathcal{B}} f_n(x)d\lambda_{\mathcal{B}}(x).$$

If we now use the nonnegative functions $\{g(x) - f_n(x)\}$, we get:

$$\begin{aligned}&\int_{\mathcal{B}} (g(x) - f(x))d\lambda_{\mathcal{B}}(x)\\ &\leqslant \liminf_{n\to\infty} \int_{\mathcal{B}} (g(x) - f_n(x))d\lambda_{\mathcal{B}}(x)\\ &= \int_{\mathcal{B}} g(x)d\lambda_{\mathcal{B}}(x) + \liminf_{n\to\infty} \int_{\mathcal{B}} -f_n(x)d\lambda_{\mathcal{B}}(x)\end{aligned}$$

and so

$$\begin{aligned}\int_{\mathcal{B}} -f(x)d\lambda_{\mathcal{B}}(x) &\leqslant \liminf_{n\to\infty} \int_{\mathcal{B}} -f_n(x)d\lambda_{\mathcal{B}}(x)\\ &= -\limsup_{n\to\infty} \int_{\mathcal{B}} f_n(x)d\lambda_{\mathcal{B}}(x)\end{aligned}$$

If we put the two inequalities together, we have

$$\limsup_{n\to\infty} \int_{\mathcal{B}} f_n(x)d\lambda_{\mathcal{B}}(x) \leqslant \int_{\mathcal{B}} f(x)d\lambda_{\mathcal{B}}(x) \leqslant \liminf_{n\to\infty} \int_{\mathcal{B}} f_n(x)d\lambda_{\mathcal{B}}(x).$$

□

2.3.6. L^p **Spaces.** In this section, we will be brief because the results are the same as those for $\mathbb{R}_I^\infty$.

Definition 2.65. Let $\mathcal{B}$ be a Banach space with an S-basis, let $L^p[\hat{\mathcal{B}}] = \bigcup_{k=1}^\infty L^p[\mathcal{B}^k]$, and let $\mathbb{C}_0(\hat{\mathcal{B}}) = \bigcup_{n=1}^\infty \mathbb{C}_0(\mathcal{B}^n)$.

(1) We say that a measurable function $f \in L^p[\mathcal{B}]$ if there exists a Cauchy sequence $\{f_m\} \subset L^p[\hat{\mathcal{B}}]$, such that

$$\lim_{m\to\infty} \int_{\mathcal{B}} |f_m(x) - f(x)|^p d\lambda_{\mathcal{B}}(x) = 0.$$

(2) We say that a measurable function $f \in \mathbb{C}_0(\mathcal{B})$, the space of continuous functions that vanish at infinity, if there exists a Cauchy sequence $\{f_m\} \subset \mathbb{C}_0(\hat{\mathcal{B}})$, such that

$$\lim_{m\to\infty} \sup_{x\in\mathcal{B}} |f_m(x) - f(x)| = 0.$$

Lemma 2.66. *If $f \in C_c(\mathcal{B})$* or *$C_0(\mathcal{B})$, then f is continuous.*

Theorem 2.67. *$\mathbb{C}_c(\mathcal{B})$ is dense in $L^p[\mathcal{B}]$.*

Theorem 2.68 (Lusin's Theorem)**.** *Let $f : \mathcal{B} \to \mathbb{R}$ be measurable and let A be a measurable set with $0 < \lambda_{\mathcal{B}}[A] < \infty$. Then, for each $\varepsilon > 0$, there exists a compact set K_ε such that*

(1) *$\lambda_{\mathcal{B}}[A \setminus K_\varepsilon] < \varepsilon$ and*

(2) *$f|_{K_\varepsilon}$ is continuous.*

Proof. Since f is measurable, let $\{s_n\}$ be a sequence of simple functions, with $s_n(x) \to f(x)$, $\lambda_{\mathcal{B}}$-(a.e.). Since each s_n is bounded on a set of finite measure, $s_n \in L^1[\mathcal{B}]$. By Egoroff's Theorem, given $\varepsilon > 0$, there is a compact set K_ε such that $\lambda_{\mathcal{B}}[A \setminus K_\varepsilon] < \varepsilon$ and $s_n \to f$ uniformly on K_ε so that $f \in L^1[A]$. Since the continuous functions with compact support are dense in $L^1[A]$, we can replace the family $\{s_n\}$ by a sequence of continuous functions $\{f_n\}$, which converge uniformly to f on K_ε. Since a uniformly convergent sequence of continuous functions on a compact set converges to a continuous function, we see that $f|_{K_\varepsilon}$ is continuous. □

We close this section with the Radon–Nikodym Theorem for Banach spaces. The proof follows that for $\mathbb{R}^n$. (The important ingredient is the σ-finite nature of $\lambda_{\mathcal{B}}$.)

Let Ω be a subspace of $\mathcal{B}$ and $(\Omega, \mathfrak{B})$ a measurable space, where and $\mathfrak{B}$ a Borel σ-algebra.

Definition 2.69. If μ, μ' are any two measures on $(\Omega, \mathfrak{B})$:

(1) We say that μ' is singular with respect to μ and write it as $\mu' \perp \mu$ if, for each $\varepsilon > 0$, there exists a set $X \in \mathfrak{B}$ such that $\mu'(X) < \varepsilon$ and $\mu(\Omega \backslash X) < \varepsilon$.

(2) We say that μ' is absolutely continuous with respect to μ and write it as $\mu' \ll \mu$ if, for each set $X \in \mathfrak{B}$ such that, if $\mu(X) = 0$, then $\mu'(X) = 0$.

(3) If $\mu' \ll \mu$ and $\mu \ll \mu'$, we say that μ and μ' are equivalent and write $\mu' \approx \mu$.

Theorem 2.70. (Radon–Nikodym) *If μ is a measure, which is absolutely continuous with respect to $\lambda_{\mathcal{B}}$, then there is a nonnegative measurable function f such that, for each $A \in \mathfrak{B}[\mathcal{B}]$, we have*

$$\mu(A) = \int_A f(x) d\lambda_{\mathcal{B}}(x). \tag{2.6}$$

The function f is essentially unique, that is, if there is a measurable function g satisfying Eq. (2.6), then $f = g$, $\lambda_{\mathcal{B}}$-(a.e.). We write $f = \frac{d\mu}{d\lambda_{\mathcal{B}}}$, $\lambda_{\mathcal{B}}$-(a.e.).

2.3.7. Product Measures and Fubini's Theorem. Let Ω_i, $i = 1, 2$ be subspaces of $\mathcal{B}$, with corresponding σ-algebras $\mathfrak{B}_i = \Omega_i \cap \mathfrak{B}(\mathcal{B})$ and measures $m_i = \lambda_{\mathcal{B}}|_{\mathfrak{B}_i}$, so that $(\Omega_i, \mathfrak{B}_i, m_i)$, $i = 1, 2$ are measure spaces.

Definition 2.71. We let $\mathfrak{B}_1 \times \mathfrak{B}_2$ denote the smallest σ-algebra of subsets of $\Omega_1 \times \Omega_2$ which contains all sets of the form $A_1 \times A_2$, with $A_1 \in \mathfrak{B}_1$ and $A_2 \in \mathfrak{B}_2$.

Recall that $\lambda_{\mathcal{B}}$ is σ-finite, so that the same is true for m_1 and m_2. Proofs of the following theorems may be found in Royden [RO].

Theorem 2.72. *There is a unique measure $m_1 \otimes m_2$ defined on $\mathfrak{B}_1 \times \mathfrak{B}_2$, such that*

$$(m_1 \otimes m_2)(A_1 \times A_2) = m_1(A_1) \cdot m_2(A_2).$$

If the above is satisfied, we call $m_1 \otimes m_2$ the product measure of m_1 and m_2.

We can now define $\mathfrak{B}_1 \times \mathfrak{B}_2$ measurable functions $f(x, y)$ on $\Omega_1 \times \Omega_2$. The value of an integrable function $f(x, y)$ on $\Omega_1 \times \Omega_2$ is denoted by

$$\iint_{\Omega_1 \times \Omega_2} f(x, y) d(m_1 \otimes m_2)(x, y) \ \text{ or } \ \iint_{\Omega_1 \times \Omega_2} f(x, y) dm_1(x) dm_2(y).$$

Theorem 2.73 (The Fubini–Tonelli Theorem)**.** *A measurable function $f(x, y)$ defined on $\Omega_1 \times \Omega_2$ is integrable if and only if one of*

$$\int_{\Omega_2} \left\{ \int_{\Omega_1} |f(x, y)| \, dm_1(x) \right\} dm_2(y) \ \textit{or} \ \int_{\Omega_1} \left\{ \int_{\Omega_2} |f(x, y)| \, dm_2(y) \right\} dm_1(x),$$

is finite, then

$$\iint_{\Omega_1 \times \Omega_2} f(x,y)dm_1(x)dm_2(y) = \int\limits_{\Omega_2} \left\{ \int\limits_{\Omega_1} f(x,y)dm_1(x) \right\} dm_2(y)$$

$$= \int\limits_{\Omega_1} \left\{ \int\limits_{\Omega_2} f(x,y)dm_2(y) \right\} dm_1(x).$$

Furthermore,

(1) *for almost all x the function f_x defined by $f_x(y) = f(x,y)$ is an integrable function on Ω_2;*

(2) *for almost all y the function f_y defined by $f_y(x) = f(x,y)$ is an integrable function on Ω_1;*

(3) $\int_{\Omega_2} f(x,y)dm_2(y)$ *is an integrable function on Ω_1; and*

(4) $\int_{\Omega_1} f(x,y)dm_1(x)$ *is an integrable function on Ω_2.*

We need a slightly different version for later. Let $\mathbb{R}_j = \mathbb{R}$, for each $j \in \mathbb{N}$ and set $\mathbb{R}^\infty = \prod_{j=1}^{\infty} \mathbb{R}_j$. Recall that, a slice through $z^0 = \left(z_1^0, \ldots z_i, z_{i+1}^0, \ldots\right)$ parallel to $\mathbb{R}_i$ in $\prod_{j=1}^{\infty} \mathbb{R}_j$ is the set:

$$S\left(z^0; i\right) = \mathbb{R}_i \times \prod \left\{z_k^0 : \; k \neq i\right\} \subset \prod_{j=1}^{\infty} \mathbb{R}_j.$$

Fix $z^0 \in \mathbb{R}^\infty$, let $\Omega_1 = S\left(z^0; i\right) \cap \mathcal{B}$ and let

$$\Omega_2 = \left[\left\{z_i^0\right\} \times \prod \left\{z_k \in \mathbb{R} : k \neq i\right\}\right] \cap \mathcal{B}.$$

If we apply Theorem 2.73, we obtain the following result.

Corollary 2.74. *With the above notation,*

$$\int_{\Omega_1} \int_{\Omega_2} f(z_i, \mathbf{z}_i) d\lambda_{\Omega_1}(z_i) d\lambda_{\Omega_2}(\mathbf{z}_i) = \int\limits_{\Omega_2} \left\{ \int\limits_{\Omega_1} f(z_i, \mathbf{z}_i) d\lambda_{\Omega_1}(z_i) \right\} d\lambda_{\Omega_2}(\mathbf{z}_i)$$

$$= \int\limits_{\Omega_1} \left\{ \int\limits_{\Omega_2} f(z_i, \mathbf{z}_i) d\lambda_{\Omega_2}(\mathbf{z}_i) \right\} d\lambda_{\Omega_1}(z_i) = \int_{\mathcal{B}} f(\mathbf{z}) d\lambda_{\mathcal{B}}(\mathbf{z})$$

Furthermore,

(1) *for almost all z_i the function f_{z_i} defined by $f_{z_i}(\mathbf{z}_i) = f(z_i, \mathbf{z}_i)$, is an integrable function on Ω_2.*

(2) *for almost all* $\mathbf{z}$ *the function* $f_{\mathbf{z}_i}$ *defined by* $f_{\mathbf{z}_i}(z_i) = f(z_i, \mathbf{z}_i)$, *is an integrable function on* Ω_1.

(3) $\int_{\Omega_2} f(z_i, \mathbf{z}_i) d\lambda_{\Omega_2}(\mathbf{z}_i)$ *is an integrable function on* Ω_1;

(4) $\int_{\Omega_1} f(z_i, \mathbf{z}_i)) d\lambda_{\Omega_1}(z_i)$ *is an integrable function on* Ω_2.

2.3.8. Inequalities and Convolution. In this section we briefly discuss the Holder and Minkowski inequalities for $L^p[\mathcal{B}]$, $1 \le p < \infty$, but they also hold for $L^p[\mathbb{R}^\infty]$. The proofs are identical to those for $L^p[\mathbb{R}^n]$, so they are omitted. We also discuss the convolution for functions in $L^1[\mathcal{B}]$, which will be used in Sect. 2.5.1, in the study of the Fourier transform when $\mathcal{B}$ is a uniformly convex Banach space with an S-basis.

Theorem 2.75. The (General) Hölder Inequality: *Let* $p_k \in (1, \infty)$, $k = 1, \ldots, N$ *be such that,* $\sum_{k=1}^N \frac{1}{p_k} = 1$. *If* $f_k \in L^{p_k}[\mathcal{B}]$ *for each* k, *then* $\prod_{k=1}^N f_k \in L^1[\mathcal{B}]$ *and*

$$\left| \int_{\mathcal{B}} \prod_{k=1}^N f_k(x) d\lambda_{\mathcal{B}}(x) \right| \leqslant \prod_{k=1}^N \|f_k\|_{p_k}.$$

Theorem 2.76 (The Minkowski Inequality)**.** *Let* $1 \le p < \infty$ *and* $f, g \in L^p[\mathcal{B}]$. *Then* $f + g \in L^p[\mathcal{B}]$ *and*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p \, .$$

Definition 2.77. Let $f, g \in \mathcal{M}$, the measurable functions on $\mathcal{B}$. The convolution, when it exists, is defined by

$$(f * g)(x) = \int_{\mathcal{B}} f(x - y) g(y) d\lambda_{\mathcal{B}}(y).$$

Theorem 2.78. *Let* $f, g \in L^1[\mathcal{B}]$, *then* $(f * g)(x)$ *exists* $\lambda_{\mathcal{B}}$-(a.e.) *and*

$$\|f * g\|_1 \leqslant \|f\|_1 \|g\|_1 \, .$$

Proof. Since $f, g \in L^1[\mathcal{B}]$, there are sequences $\{f_n\}$, $\{g_n\}$, $\in L^1[\mathcal{B}^n]$ such that $f_n \to f$ (a.e) and $g_n \to g$, $\lambda_{\mathcal{B}}$-(a.e.). Since each $f_n(x)$ and $g_n(x)$ is in $L^1[\mathcal{B}^n]$, we can apply Fubini's Theorem to get

$$\begin{aligned}
&\int_{\mathcal{B}} (f_n * g_n)(x) d\lambda_{\mathcal{B}}(x) \\
&= \int_{\mathcal{B}} d\lambda_{\mathcal{B}}(x) \left[\int_{\mathcal{B}} f_n(y) g_n(x-y) d\lambda_{\mathcal{B}}(y) \right] \\
&= \int_{\mathcal{B}} d\lambda_{\mathcal{B}}(y) \left[\int_{\mathcal{B}} f_n(y) g_n(x-y) d\lambda_{\mathcal{B}}(x) \right] \\
&= \int_{\mathcal{B}} f_n(y) d\lambda_{\mathcal{B}}(y) \int_{\mathcal{B}} g_n(x) d\lambda_{\mathcal{B}}(x) \\
&= \int_{\mathcal{B}} \left[\int_{\mathcal{B}} f_n(y) g_n(x) d\lambda_{\mathcal{B}}(y) \right] d\lambda_{\mathcal{B}}(x).
\end{aligned}$$

It follows from the last equality that

$$\int_{\mathcal{B}} (f_n * g_n)(x) d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} \left[\int_{\mathcal{B}} f_n(y) g_n(x) d\lambda_{\mathcal{B}}(y) \right] d\lambda_{\mathcal{B}}(x).$$

Set $c = \sup_n \|g_n\|_1 \|f_n\|_1$, $h_n(x) = \max\{|g_n(x)| \|f_n\|_1, |f_n(x)| \|g_n\|_1\}$ and $h(x) = \max\{|g(x)| \|f\|_1, |f(x)| \|g\|_1\}$. We now have that,

$$|(f * g)_n (x)| \leqslant h_n(x) \leqslant h(x), \; \lambda_{\mathcal{B}} - \text{(a.e.)}.$$

Since $h(x) \in L^1[\mathcal{B}]$, the dominated conversion theorem shows that:

$$\int_{B} (f * g)(x) d\lambda_{\mathcal{B}}(x) = \lim_{n\to\infty} \int_{B} (f_n * g_n)(x) d\lambda_{\mathcal{B}}(x)$$

and

$$\|f * g\|_1 \leqslant \|f\|_1 \|g\|_1 .$$

□

2.3.9. Young's Theorem. In this section we establish a version of Young's Theorem for every separable Banach space with an S-basis:

Theorem 2.79. (Young) *Let $p, q, r \in [1, \infty]$ with*

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1.$$

If $f \in L^p[\mathcal{B}]$ and $g \in L^q[\mathcal{B}]$, then the convolution of f and g, $f * g$, exists (a.s.), belongs to $L^r[\mathcal{B}]$ and

$$\|f * g\|_r \leqslant \|f\|_p \|g\|_q .$$

Proof. First, fix n and assume that $f_n \in L^p[\mathcal{B}^n]$ and $g_n \in L^q[\mathcal{B}^n]$ are nonnegative with $\|f_n\|_p = \|g_n\|_q = 1$. Let $\frac{1}{q'} = 1 - \frac{1}{q}$ and $\frac{1}{p'} = 1 - \frac{1}{p}$. Now note that

$$\frac{1}{r} + \frac{1}{q'} + \frac{1}{p'} = \frac{1}{r} + \left(1 - \frac{1}{q}\right) + \left(1 - \frac{1}{p}\right) = 1;$$
$$\left(1 - \frac{p}{r}\right) q' = p\left(\frac{1}{p} - \frac{1}{r}\right) q' = p\left(1 - \frac{1}{q}\right) q' = p;$$
$$\left(1 - \frac{q}{r}\right) p' = q\left(\frac{1}{q} - \frac{1}{r}\right) p' = q\left(1 - \frac{1}{p}\right) p' = q.$$

If we use Holder's inequality (for three functions), we can write $(f_n * g_n)(x)$ as:

$$\begin{aligned}
&(f_n * g_n)(x) \\
&= \int_{\mathcal{B}} \left[f(y)_n^{p/r} g(x-y)_n^{q/r}\right] \left[f(y)_n^{1-p/r} g(x-y)_n^{1-q/r}\right] d\lambda_{\mathcal{B}}(y) \\
&\leqslant \left[\int_{\mathcal{B}} f(y)_n^p g(x-y)_n^q d\lambda_{\mathcal{B}}(y)\right]^{1/r} \left[\int_{\mathcal{B}} f(y)^{(1-p/r)q'} d\lambda_{\mathcal{B}}(y)\right]^{1/q'} \\
&\qquad \left[\int_{\mathcal{B}} g(x-y)_n^{(1-q/r)p'} d\lambda_{\mathcal{B}}(y)\right]^{1/p'}.
\end{aligned}$$

This last inequality shows that

$$(f_n * g_n)(x) \leqslant \left[\int_{\mathcal{B}} f(y)_n^p g(x-y)_n^q d\lambda_{\mathcal{B}}(y)\right]^{1/r},$$

and so

$$(f_n * g_n)^r(x) \leqslant \left[\int_{\mathcal{B}} f(y)_n^p g(x-y)_n^q d\lambda_{\mathcal{B}}(y)\right].$$

Hence

$$(f_n * g_n)^r(x) \leqslant (f_n^p * g_n^q)(x).$$

From Theorem 2.79, we have $\|(f_n * g_n)^r\|_1 \leqslant \|f_n^p\|_1 \|g_n^q\|_1 = 1$. Now using the dominating convergence theorem, we get

$$\|(f * g)^r\|_1 \leqslant \|f^p\|_1 \|g^q\|_1 = 1.$$

In the general case, we can write $F(x) = f(x)/\|f\|_p$ and $G(x) = g(x)/\|g\|_q$. ☐

In closing we note that Beckner [BE] and Brascamp–Lieb [BL] have shown that on $\mathbb{R}^n$ we can write Young's inequality as

$$\|f * g\|_r \leqslant (C_{p,q,r;n})^n \|f\|_p \|g\|_q ,$$

where $C_{p,q,r;n} \leq 1$ is sharp. We conjecture that 1 is the sharp constant for $\mathcal{B}$.

2.4. Distributions on Uniformly Convex Banach Spaces

The foundations analysis is based on differentiation and integration. Early in the history of analysis, it was realized that not all functions are differentiable. Around the same time, developments in physics forcefully suggested that information about the actual physical world always appears in the form of averages, or mean values and not as pointwise defined values of functions, as was assumed in classical physics. The theory of distribution attempts to solve this problem for both analysis and physics, by imbedding classical functions into a larger class of generalized functions. The basic idea is to replace pointwise defined functions by their "mean value" in a certain sense, which is described in this section. At this point, a review of Sect. 1.3 of Chap. 1 is recommended.

In this section, we briefly discuss distributions on uniformly convex Banach spaces $\mathcal{B}$, with an S-basis. We are brief, because proofs of the main results for uniformly convex Banach spaces $\mathcal{B}$ (with an S-basis) are direct adaptations of those for $\mathbb{R}^n$. This section is a prelude to the next on Schwartz spaces and Fourier transforms, which are important for the general theory of path integrals, discussed in Chap. 8.

2.4.1. Preliminaries.

Definition 2.80. Let $\mathbb{N}_0^\infty$ be the set of all multi-index infinite tuples $\alpha = (\alpha_1, \alpha_2, \dots)$, with $\alpha_i \in \mathbb{N}$ and all but a finite number of entries are zero. We define the operators D^α and D_α by:

$$D^\alpha = \prod_{k=1}^{\infty} \frac{\partial^{\alpha_k}}{\partial x_i^{\alpha_k}}, \quad D_\alpha = \prod_{k=1}^{\infty} \left(\frac{1}{2\pi i} \frac{\partial}{\partial x_k} \right)^{\alpha_k}$$

(The products are well defined, since each has only a finite number of terms.)

Definition 2.81. Let $\mathcal{B}$ be a uniformly convex Banach space with an S-basis. In this case, $\mathcal{B}'$ is the dual space of $\mathcal{B}$ and the pairing $\langle x, y \rangle$ is always well defined for $x \in \mathcal{B}$ and $y \in \mathcal{B}'$.

(1) We recall that $L^1[\mathcal{B}]$ is the L^1-norm closure of $L^1[\hat{\mathcal{B}}] = \bigcup_{n=1}^{\infty} L^1[\mathcal{B}^n]$.

(2) We say a function $u \in L^1_{\text{loc}}[\mathcal{B}]$ if, for every compact set $K \subset \mathcal{B}$, $u|_K \in L^1[\mathcal{B}]$.

(3) We say that a sequence of functions $\{f_m\} \subset \mathbb{C}^\infty(\mathcal{B}^n)$ converges to a function $f \in \mathbb{C}^\infty(\mathcal{B}^n)$ if and only if, for all multi-indices α, $D^\alpha f \in \mathbb{C}(\mathcal{B}^n)$ and, for $x \in \mathcal{B}^n$ and all $N \in \mathbb{N}$,

$$\limsup_{m\to\infty}\left[\sup_{\alpha}\sup_{\|x\|\leqslant N}|D^\alpha f(x) - D^\alpha f_m(x)|\right] = 0.$$

(4) We say that a function $f \in \mathbb{C}^\infty(\mathcal{B})$ if and only if there exists a sequence of functions $\{f_m\} \subset \mathbb{C}^\infty(\hat{\mathcal{B}}) = \bigcup_{n=1}^\infty \mathbb{C}^\infty(\mathcal{B}^n)$ such that, for all multi-index infinite tuples $\alpha \in \mathbb{N}_0$, $D^\alpha f_m \in \mathbb{C}^\infty(\hat{\mathcal{B}})$ and, for all $x \in \mathcal{B}$ and all $N \in \mathbb{N}$,

$$\limsup_{m\to\infty}\left[\sup_{\alpha}\sup_{\|x\|\leqslant N}|D^\alpha f(x) - D^\alpha f_m(x)|\right] = 0.$$

Definition 2.82. We say that a measurable function $f \in \mathfrak{D}(\mathcal{B})$ if and only if there exists a sequence of functions $\{f_m\} \subset \mathfrak{D}(\hat{\mathcal{B}}) = \bigcup_{n=1}^\infty \mathfrak{D}(\mathcal{B}^n)$ and a compact set $K \subset \mathcal{B}$, which contains the support of $f - f_m$, for all m, and $D^\alpha f_m \to D^\alpha f$ uniformly on K, for every multi-index $\alpha \in \mathbb{N}_0^\infty$. We call the topology of $\mathfrak{D}(\mathcal{B})$ the compact sequential limit topology.

Definition 2.83. The set of all continuous linear functionals $T \in \mathfrak{D}'(\mathcal{B})$, the dual space of $\mathfrak{D}(\mathcal{B})$, is called the space of distributions on $\mathcal{B}$. A family of distributions $\{T_i\} \subset \mathfrak{D}'(\mathcal{B})$ is said to converge to $T \in \mathfrak{D}'(\mathcal{B})$ if, for every $\phi \in \mathfrak{D}(\mathcal{B})$, the numbers $T_i(\phi)$ converge to $T(\phi)$.

The most important class of distributions are functions. For example, if $f \in L^1_{\text{loc}}[\mathcal{B}]$ and $\phi \in \mathbb{C}_c^\infty(\mathcal{B})$, we can use integration by parts to define a generalized definition of the derivative of f by

$$\int_{\mathcal{B}} Df(x)\phi(x)d\lambda_{\mathcal{B}}(x) = -\int_{\mathcal{B}} f(x)D\phi(x)d\lambda_{\mathcal{B}}(x).$$

In particular, if $f(x) = H(x)$, the Heaveside function on $\mathcal{B}$, $x = (x_1, \ldots)$,

$$H(x) = \begin{cases} 1, & x_i \geq 0,\ \forall i \in \mathbb{N}, \\ 0, & \exists i \in \mathbb{N},\ \ni\ x_i < 0. \end{cases}$$

then

$$\begin{aligned}\int_{\mathbb{R}} DH(x)\phi(x)d\lambda_{\mathcal{B}}(x) &= -\int_{\mathcal{B}} H(x)D\phi(x)d\lambda_{\mathcal{B}}(x) \\ &= \phi(0) = \int_{\mathcal{B}} \delta_{\mathcal{B}}(x)\phi(x)d\lambda_{\mathcal{B}}(x),\end{aligned}$$

so that, in the generalized sense of distributions, $DH(x) = \delta(x)$, the Dirac delta function on $\mathcal{B}$.

Definition 2.84. If α is a multi-index and $u,\ v \in L^1_{\text{loc}}[\mathcal{B}]$, we say that v is the αth-weak (or distributional) partial derivative of u and write $D^\alpha u = v$ provided that

$$\int_{\mathcal{B}} u(D^\alpha \phi)d\lambda_{\mathcal{B}} = (-1)^{|\alpha|}\int_{\mathcal{B}} \phi v\ d\lambda_{\mathcal{B}}$$

for all functions $\phi \in \mathbb{C}_c^\infty(\mathcal{B})$. Thus, v is in the dual space $\mathcal{D}'(\mathcal{B})$ of $\mathcal{D}(\mathcal{B})$.

If $u \in L^1_{\text{loc}}[\mathcal{B}]$ and $\phi \in \mathcal{D}(\mathcal{B})$, then we can define $T_u(\cdot)$ by

$$T_u(\phi) = \int_{\mathcal{B}} u\phi d\lambda_{\mathcal{B}}.$$

It is clearly a linear functional on $\mathcal{D}(\mathcal{B})$. If $\{\phi_n\} \subset \mathcal{D}(\mathcal{B})$ and $\phi_n \to \phi$ in $\mathcal{D}(\mathcal{B})$, with the support of $\phi_n - \phi$ contained in a compact set $K \subset \mathcal{B}$, then

$$\begin{aligned}|T_u(\phi_n) - T_u(\phi)| &= \left|\int_{\mathcal{B}} u(x)[\phi_n(x) - \phi(x)]d\lambda_{\mathcal{B}}(x)\right| \\ &\le \sup_{x\in K}|\phi_n(x) - \phi(x)|\int_K |u(x)|d\lambda_{\mathcal{B}}(x).\end{aligned}$$

Thus, by uniform convergence on K, we see that T is continuous, so that $T \in \mathcal{D}'(\mathcal{B})$. Let

$$\|\phi\|_N = \sup_{x\in\mathcal{B}}\{|D^\alpha\phi(x)| : \alpha \in \mathbb{N}_0^\infty,\ |\alpha| \le N\}.$$

The proof of the following theorem is along the same lines as for $\mathcal{D}(\mathbb{R}^n)$.

Theorem 2.85. *Let $\mathcal{D}'(\mathcal{B})$ be the dual space of $\mathcal{D}(\mathcal{B})$.*

(1) *Every differential operator D^α, $\alpha \in \mathbb{N}_0^\infty$ defines a bounded linear operator on $\mathcal{D}(\mathcal{B})$.*

(2) *If $T \in \mathcal{D}'(\mathcal{B})$ and $\alpha \in \mathbb{N}_0^\infty$, then $D^\alpha T \in \mathcal{D}'(\mathcal{B})$ and*

$$(D^\alpha T)(\phi) = (-1)^{|\alpha|}T(D^\alpha\phi),\ \phi \in \mathcal{D}(\mathcal{B}).$$

(3) *If* $|T(\phi)| \leq C \|\phi\|_N$ *for all* $\phi \in \mathcal{D}(K)$, *for some compact set* $K \subset \mathcal{B}$, *then*

$$|(D^\alpha T)(\phi)| \leq C \|\phi\|_{N+|\alpha|}$$

and $D^\alpha D^\beta T = D^\beta D^\alpha T$.

(4) *If* $g = D^\alpha f$ *exists as a classical derivative and* $g \in L^1_{\mathrm{loc}}[\mathcal{B}]$, *then* $T_g \in \mathcal{D}'(\mathcal{B})$ *and*

$$(-1)^{|\alpha|} \int_{\mathcal{B}} f(x)(D^\alpha \phi) d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} g(x)\phi(x)\ d\lambda_{\mathcal{B}}(x)$$

for all $\phi \in \mathcal{D}(\mathcal{B})$.

(5) *If* $f \in \mathbb{C}^\infty(\mathcal{B})$, $T \in \mathcal{D}'(\mathcal{B})$, *then* $fT \in \mathcal{D}'(\mathcal{B})$, *with* $fT(\phi) = T(f\phi)$ *for all* $\phi \in \mathcal{D}(\mathcal{B})$ *and*

$$D^\alpha(fT) = \sum_{\beta \leq \alpha} c_{\alpha\beta}(D^{\alpha-\beta} f)(D^\beta T).$$

Our decision to be brief and restricted because of our objectives should not be interpreted to mean that the study of distributions on uniformly convex Banach spaces is not an interesting and important subject in its own right. For example, to our knowledge, the concept of a weak solution to an initial value problem for a partial differential equation in infinitely many variables has not been formulated. The interested reader is encouraged to explore the many possible questions and applications.

2.5. The Schwartz Space and Fourier Transform

On $\mathbb{R}^n$ the Fourier transform can be defined in a number of equivalent ways. However, in the infinite dimensional case, there is only one possibility. Let $\mathcal{B}$ be a uniformly convex Banach space with an S-basis.

First, we need to specify our conventions. Recall, that $\mathcal{B}$ is its representation $\mathcal{B}_J \in \mathbb{R}^\infty$. Also recall that any $\alpha \in \mathbb{N}_0^\infty$ only has a finite number of nonzero terms. If $x = (x_1, x_2, \ldots) \in \mathcal{B}$ and $\alpha \in \mathbb{N}_0^\infty, \alpha = (\alpha_1, \alpha_2, \ldots)$, we define x^α by $x^\alpha = \prod_{k=1}^\infty x_k^{\alpha_k}$, a finite product of real or complex numbers.

Definition 2.86. Let $e_x(y) = e^{2\pi i\langle x,y\rangle}$ and let $\alpha \in \mathbb{N}_0^\infty$.

(1) If $x \in \mathcal{B}$, we define

$$D_\alpha e_x(y) = x^\alpha e_x(y),\ \text{where} \quad x^\alpha = \prod_{k=1}^\infty x_k^{\alpha_k}.$$

(2) If $P(\cdot)$ is a polynomial with complex coefficients and $x \in \mathcal{B}$, then

$$P(x) = \sum c_\alpha x^\alpha = \sum c_\alpha \prod_{k=1}^{\infty} x_k^{\alpha_k}.$$

(3) We define $P(D)$ and $P(-D)$ by

$$P(D) = \sum c_\alpha D_\alpha, \quad P(-D) = \sum (-1)^{|\alpha|} c_\alpha D_\alpha,$$

so that $P(D)e_x(y) = P(x)e_x(y)$, for all x.

(4) We define τ_y on $f(x) \in \mathcal{M}$ (measurable functions on $\mathcal{B}$), by $\tau_y f(x) = f(x-y)$.

Definition 2.87. A function $f \in \mathbb{C}^\infty(\mathcal{B})$ is called a Schwartz function, or $f \in \mathcal{S}(\mathcal{B})$, if and only if, for all multi-indices α and β in $\mathbb{N}_0^\infty$, the seminorm $\rho_{\alpha,\beta}(f)$ is finite, where

$$\rho_{\alpha,\beta}(f) = \sup_{x\in\mathcal{B}} \left| x^\alpha D^\beta f(x) \right|. \tag{2.7}$$

Theorem 2.88. $\mathcal{S}(\mathcal{B})$ *(respectively* $\mathcal{S}(\mathcal{B}')$*) is a Fréchet space, which is dense in* $\mathbb{C}_0(\mathcal{B})$*.*

Proof. We prove the result for $\mathcal{S}(\mathcal{B})$. If $\alpha, \beta \in \mathbb{N}_0^\infty$ then the function:

$$d(f,g) = \sum_{\alpha,\beta} \frac{\rho_{(\alpha,\beta)}(f-g)}{2^{|\alpha|+|\beta|}\left[1+\rho_{(\alpha,\beta)}(f-g)\right]}$$

is a translation invariant metric on $\mathcal{S}(\mathcal{B})$. It is easy to see that $\mathcal{S}(\mathcal{B})$ is locally convex and that the family of seminorms separate points, so it suffices to prove that $\mathcal{S}(\mathcal{B})$ is complete. Let $\{f_k\}$ be a Cauchy sequence in $\mathcal{S}(\mathcal{B})$. This means that

$$\lim_{m,n\to\infty} d(f_m, f_n) = \sum_{\alpha,\beta} \frac{\rho_{(\alpha,\beta)}(f_m - f_n)}{2^{|\alpha|+|\beta|}\left[1+\rho_{(\alpha,\beta)}(f_m-f_n)\right]} = 0,$$

if and only if $\lim_{m,n\to\infty} \rho_{(\alpha,\beta)}(f_m - f_n) = 0$, for every pair of multi-indices $(\alpha,\beta) \in \mathbb{N}_0^\infty \times \mathbb{N}_0^\infty$. Thus, the sequence of functions $\left\{x^\alpha D^\beta f_k\right\}$, converge uniformly on $\mathcal{B}$ to a bounded function, say $f_{(\alpha,\beta)}$, for each pair $(\alpha,\beta) \in \mathbb{N}_0^\infty \times \mathbb{N}_0^\infty$. In particular, if $(\alpha,\beta) = (0,0)$, then the sequence of functions $\{f_k\}$, converge uniformly on $\mathcal{B}$ to the bounded function $f_{(0,0)}$. It follows from Eq. (2.7) that

$$f_{(\alpha,\beta)}(x) = x^\alpha D^\beta f_{(0,0)}(x)$$

for all pairs $(\alpha, \beta) \in \mathbb{N}_0^\infty \times \mathbb{N}_0^\infty$, so that $f_{(0,0)} \in \mathcal{S}(\mathcal{B})$ and $\mathcal{S}(\mathcal{B})$ is complete. $\square$

2.5.1. The Transform on Uniformly Convex Banach Spaces. In this section we study the Fourier transform when $\mathcal{B}$ is a uniformly convex Banach space with an S-basis. This is the natural framework if one wants to remain close to the what is known on finite dimensional Euclidean space. In the Appendix (Sect. 6.7) to Chap. 6, we offer a different definition of the Fourier transform, which applies to all Banach spaces with an S-basis.

Definition 2.89. For each $f \in L^1[\mathcal{B}, \mathfrak{B}[\mathcal{B}], \lambda_{\mathcal{B}}]$ we define $\mathfrak{F}$ directly by:

$$[\mathfrak{F}(f)](y) = \hat{f}(y) = \int_{\mathcal{B}} \exp\{-2\pi i \langle x, y \rangle\} f(x) d\lambda_{\mathcal{B}}(x), \tag{2.8}$$

where $x \in \mathcal{B}$ and $y \in \mathcal{B}'$ and $\langle x, y \rangle$ is the pairing between $\mathcal{B}$ and $\mathcal{B}'$.

The next theorem shows that the Fourier transform defined above has all the properties we would expect from our understanding of the same object on $\mathbb{R}^n$.

Theorem 2.90. *If $f \in L^1[\mathcal{B}]$, then*

(1) $|\hat{f}(y)| \leqslant \|f\|_1$ *for all* $y \in \mathcal{B}'$.

(2) $\hat{f}(y)$ *is uniformly continuous in* $y \in \mathcal{B}'$ *and* $\|\hat{f}\|_\infty \leqslant \|f\|_1$.

(3)

$$\int_{\mathcal{B}} f(x) d\lambda_{\mathcal{B}}(x) = \hat{f}(0).$$

(4) $\widehat{(f * g)}(y) = \hat{f}(y)\hat{g}(y)$.

Proof. The proof of (1) is clear. To prove (2), let $\{y_n\}$ be a sequence in $\mathcal{B}'$, $y_n \to y$, as $n \to \infty$. Then

$$\left|\hat{f}(y) - \hat{f}(y_n)\right| = \left|\int_{\mathcal{B}} f(x) \left[\exp\{-2\pi i \langle x, y \rangle\} - \exp\{-2\pi i \langle x, y_n \rangle\}\right] d\lambda_{\mathcal{B}}(x)\right|$$
$$\leqslant \int_{\mathcal{B}} |f(x)| \left|\exp\{-2\pi i \langle x, y - y_n \rangle\} - 1\right| d\lambda_{\mathcal{B}}(x).$$

Let $\varepsilon > 0$, be given and choose $r > 0$ such that $\int_{\|x\|_{\mathcal{B}} > r} |f(x)|\, d\lambda_{\mathcal{B}}(x) \leqslant \frac{\varepsilon}{4}$. Now,

$$\begin{aligned}
&\int_{\|x\|_{\mathcal{B}} > r} |f(x)|\, |\exp\{-2\pi i \langle x, y - y_n\rangle\} - 1|\, d\lambda_{\mathcal{B}}(x) \\
&\leqslant \int_{\|x\|_{\mathcal{B}} > r} |f(x)|\, \{|\exp\{-2\pi i \langle x, y - y_n\rangle\}| + 1\}\, d\lambda_{\mathcal{B}}(x) \\
&= 2\int_{\|x\|_{\mathcal{B}} > r} |f(x)|\, d\lambda_{\mathcal{B}}(x).
\end{aligned}$$

It follows that

$$\begin{aligned}
\left|\hat{f}(y) - \hat{f}(y_n)\right| &\leqslant \int_{\|x\|_{\mathcal{B}} \leqslant r} |f(x)| \left|e^{-2\pi i \langle x, y - y_n\rangle} - 1\right| d\lambda_{\mathcal{B}}(x) \\
&\quad + 2\int_{\|x\|_{\mathcal{B}} > r} |f(x)|\, d\lambda_{\mathcal{B}}(x) \\
&\leqslant 2\pi \|y - y_n\|_{\mathcal{B}'} \int_{\|x\|_{\mathcal{B}} \leqslant r} |f(x)|\, \|x\|_{\mathcal{B}}\, d\lambda_{\mathcal{B}}(x) \\
&\quad + 2\int_{\|x\|_{\mathcal{B}} > r} |f(x)|\, d\lambda_{\mathcal{B}}(x) \\
&\leqslant 2\pi \|y - y_n\|_{\mathcal{B}'} \int_{\|x\|_{\mathcal{B}} \leqslant r} |f(x)|\, \|x\|_{\mathcal{B}}\, d\lambda_{\mathcal{B}}(x) + \frac{\varepsilon}{2}.
\end{aligned}$$

If we set

$$\delta = \varepsilon \left[4\pi \int_{\|x\|_{\mathcal{B}} \leqslant r} |f(x)|\, \|x\|_{\mathcal{B}}\, d\lambda_{\mathcal{B}}(x)\right]^{-1},$$

then, as soon as n is large enough ($n > N$, for some N), $\|y - y_n\|_{\mathcal{B}'} < \delta$, then $\left|\hat{f}(y) - \hat{f}(y_n)\right| < \varepsilon$. Since δ does not depend on N, we see that $\hat{f}$ is uniformly continuous.

The proof of (3) follows from the dominating convergence theorem. To prove (4), we first use Theorem 2.78, to see that $(f * g)(x)$ exists $\lambda_{\mathcal{B}}$-(a.e.) and

$$\|f * g\|_1 \leqslant \|f\|_1 \|g\|_1 .$$

We can now use

$$\begin{aligned}(f * g)\,(x)e^{-2\pi i\langle x,y\rangle} &= \int_{\mathcal{B}} f(z)g(x-z)e^{-2\pi i\langle x,y\rangle}d\lambda_{\mathcal{B}}(z)\\ &= \int_{\mathcal{B}} f(z)e^{-2\pi i\langle z,y\rangle}g(x-z)e^{-2\pi i\langle x-z,y\rangle}d\lambda_{\mathcal{B}}(z)\\ &= \left[\left(fe^{-2\pi i\langle y,x\rangle}\right) * \left(ge^{-2\pi i\langle x,y\rangle}\right)\right](x),\end{aligned}$$

to get our conclusion. □

In the case of measures, we can also define the Fourier transform:

Definition 2.91. Let $\mathfrak{M}[\mathcal{B}]$ be the space of finite measures on $\mathcal{B}$, with $\|\mu_{\mathcal{B}}\| = |\mu_{\mathcal{B}}|\,(B)$ for $\mu_{\mathcal{B}} \in \mathfrak{M}[\mathcal{B}]$ and, where $|\mu_{\mathcal{B}}|\,(B)$ is the total variation. If $\mu_{\mathcal{B}} \in \mathfrak{M}[\mathcal{B}]$, we define the Fourier transform of $\mu_{\mathcal{B}}$, $\hat{\mu}_{\mathcal{B}}$, by

$$\hat{\mu}_{\mathcal{B}}(y) = \int_{\mathcal{B}} e^{-2\pi i\langle x,y\rangle}d\mu_{\mathcal{B}}(x) = \lim_{n\to\infty}\int_{\mathcal{B}^n} e^{-2\pi i\langle x,y\rangle_n}d\mu_{\mathcal{B}}(x),$$

where $\langle x,y\rangle_n$ is the pairing of $\mathcal{B}^n$ with $\mathcal{B}'^n$.

The proof of the next theorem is very close to same result for functions.

Theorem 2.92. *If $\mu \in \mathfrak{M}[\mathcal{B}]$, then*

(1) $|\hat{\mu}(y)| \leqslant \|\mu\|$ *for all* $y \in \mathcal{B}'$.
(2) $\hat{\mu}(y)$ *is uniformly continuous and* $\|\hat{\mu}\|_\infty \leqslant \|\mu\|$.
(3)
$$\int_{\mathcal{B}} d\mu_{\mathcal{B}}(x) = \hat{\mu}(0).$$
(4) $\widehat{(\mu * \nu)}(y) = \hat{\mu}(y)\hat{\nu}(y)$.

Theorem 2.93. *Let $f \in L^1[\mathcal{B}]$, then*

(1) $\mathfrak{F}(\tau_z f)(y) = e_{-z}(y)\mathfrak{F}(f)(y)$,
(2) $\tau_z[\mathfrak{F}(f)(y)] = \mathfrak{F}[e_z f](y)$.
(3) $\mathfrak{F} : \mathcal{S}(\mathcal{B}) \to \mathcal{S}(\mathcal{B}')$, *with*

$$\mathfrak{F}[P(D)f](z) = P(z)\mathfrak{F}[f](z) \quad \textit{and} \quad \mathfrak{F}[Pf](z) = P(-D)\mathfrak{F}[f](z).$$

(4) *The transformation* $\mathfrak{F}$ *is a bijective linear continuous mapping of* $\mathcal{S}(\mathcal{B}) \to \mathcal{S}(\mathcal{B}')$ *and* (inversion):

$$\mathfrak{F}^{-1} : \mathcal{S}(\mathcal{B}') \to \mathcal{S}(\mathcal{B})$$

is also continuous.

(5) $\mathfrak{F} : L^1[\mathcal{B}] \to \mathbb{C}_0[\mathcal{B}']$.

Proof. For the proof of (1), we have:

$$\mathfrak{F}[\tau_z f]\,(y) = \int_{\mathcal{B}} e^{-2\pi i\langle x,y\rangle} f(x-z)d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} e^{-2\pi i\langle u+z,y\rangle} f(u)d\lambda_{\mathcal{B}}(u)$$
$$= e^{-2\pi i\langle z,y\rangle}\int_{\mathcal{B}} e^{-2\pi i\langle u,y\rangle} f(u)d\lambda_{\mathcal{B}}(u) = e^{-2\pi i\langle z,y\rangle}\hat{f}(y) = e_{-z}(y)\mathfrak{F}[f]\,(y)\,.$$

The proof of (2) follows from:

$$\tau_z\left[\mathfrak{F}\,(f)\,(y)\right] = \hat{f}(y-z) = \int_{\mathcal{B}} e^{-2\pi i\langle x,y-z\rangle} f(x)d\lambda_{\mathcal{B}}(x)$$
$$= \int_{\mathcal{B}} e^{2\pi i\langle x,z\rangle}e^{-2\pi i\langle x,y\rangle} f(x)d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} e_z\,(x)\,e^{-2\pi i\langle x,y\rangle} f(x)d\lambda_{\mathcal{B}}(x)$$
$$= \mathfrak{F}[e_z f]\,(y)\,.$$

To prove (3), it is easy to see that if $f(x) \in \mathcal{S}(\mathcal{B})$ then $P(D)f(x) \in \mathcal{S}(\mathcal{B})$ and

$$\{(P\,(D)\,f) * e_z\}\,(x) = \{f * P\,(D)\,e_z\}\,(x)$$
$$= f * P\,(z)\,e_z = P\,(z)\,[f * e_z]\,(x) = P\,(z)\int_B e^{-2\pi i\langle x-y,z\rangle} f(x)d\lambda_B(x)$$
$$= P\,(z)\,e^{2\pi i\langle y,z\rangle}\int_B e^{-2\pi i\langle x,z\rangle} f(x)d\lambda_B(x) = P\,(z)\,e^{2\pi i\langle y,z\rangle}\hat{f}(z).$$

If we set $y = 0$, we get the first part, $\mathfrak{F}[P(D)f](z) = P(z)\mathfrak{F}[f](z)$. For the second part, compute $\mathfrak{F}[x_1 f]$ and iterate, using induction to get the general result.

For (4), let $f(x) \in \mathcal{S}(\mathcal{B})$, $P(x)$ be a polynomial, use Leibniz formula and the Closed Graph Theorem, to see that the transformations:

$$f(x) \to P(x)f(x),\ f(x) \to h(x)f(x),\ \text{and}\ f(x) \to x^\alpha D_\beta f(x)$$

are all continuous linear mappings of $\mathcal{S}(\mathcal{B})$ into $\mathcal{S}(\mathcal{B})$. Let $\hat{\mathcal{S}}(\mathcal{B}')$ be the set of all $\hat{f}(y) = \mathfrak{F}[f](y)$, for $f \in \mathcal{S}(\mathcal{B})$. From (3), we see that

$$\mathfrak{F}[Pf](y) \quad \textit{and} \quad \mathfrak{F}[P(D)f](y)$$

belong to $\hat{\mathcal{S}}(\mathcal{B}')$. It is easy to see that $\mathfrak{F}$ is injective, so we need to show it is surjective.

By the definition of $\mathcal{S}[\mathcal{B}]$, we can find a sequence of functions $\{f_n\}$ with $f_n \in \mathcal{S}(\mathcal{B}^n) \subset \mathcal{S}(\mathcal{B})$ and $f_n \to f$, $\lambda_{\mathcal{B}}$-(a.e.) as $n \to \infty$. For what follows, we recall that if $\nu_{\mathcal{B}}$ is our Gaussian measure on $\mathfrak{B}[\mathcal{B}_J^n]$, $n \in \mathbb{N}$, then (with $(\bar{x}) = (x_i)_{i=1}^n$, understood where appropriate),

$$d\nu_{\mathcal{B}}(x) = \left\{ e^{-\pi \sum_{k=1}^n |x_k|^2} \right\} d\lambda_n(x).$$

Furthermore, we see that:

$$\hat{\nu}_{\mathcal{B}^n}(y) = \int_{\mathbb{R}^n} e^{-2\pi i \langle x,y \rangle_n} e^{-\pi \sum_{k=1}^n |x_k|^2} d\lambda_n(x) = e^{-\pi \sum_{k=1}^n |y_k|^2},$$

where $\langle x, y \rangle_n$ is the inner product on $\mathbb{R}^n$. Fix n and apply Fubini's Theorem to the double integral

$$\begin{aligned}
&\int_{\mathcal{B}'} \int_{\mathcal{B}} f_n(x) e^{-2\pi \langle x,y \rangle_n} d\nu_{\mathcal{B}}(x) d\lambda_{\mathcal{B}'}(y) \\
&= \int_{\mathcal{B}'} \int_{\mathcal{B}} f_n(x) e^{-2\pi \langle x,y \rangle_n} \left\{ e^{-\pi \sum_{k=1}^n |x_k|^2} \otimes \left(\bigotimes_{k=n+1}^{\infty} \chi_I(x_k) \right) \right\} \\
&\times d\lambda_{\mathcal{B}}(x) d\lambda_{\mathcal{B}'}(y) \\
&= \int_{\mathcal{B}'} \hat{f}_n(y) \left\{ e^{-\pi \sum_{k=1}^n |y_k|^2} \otimes \left(\bigotimes_{k=n+1}^{\infty} \chi_I(y_k) \right) \right\} d\lambda_{\mathcal{B}'}(y) \\
&= \int_{\mathbb{R}^n} f_n(x) e^{-\pi \sum_{k=1}^n |x_k|^2} d\lambda_n(x)
\end{aligned}$$

An easy calculation shows that, for each n and each $\alpha > 0$, we have

$$\int_{\mathbb{R}^n} \hat{f}_n(y) e^{-\pi \sum_{k=1}^n \left|\frac{y_k}{\alpha}\right|^2} d\lambda_n(y) = \int_{\mathbb{R}^n} f_n(\tfrac{x}{\alpha}) e^{-\pi \sum_{k=1}^n |x_k|^2} d\lambda_n(x)$$

Let $\alpha \to \infty$ to get

$$\hat{\nu}_{\mathcal{B}^n}(0) \int_{\mathbb{R}^n} \hat{f}_n(y) d\lambda_n(y) = f_n(0) \int_{\mathbb{R}^n} e^{-\pi \sum_{k=1}^n |x_k|^2} d\lambda_n(x).$$

Since $\hat{\nu}_{\mathcal{B}^n}(0) = \int_{\mathbb{R}^n} e^{-\pi \sum_{k=1}^n |x_k|^2} d\lambda_n(x) = 1$, it follows that

$$f_n(0) = \int_{\mathcal{B}'} \hat{f}_n(y) d\lambda_{\mathcal{B}'}(y)$$

for all n. By definition of the integral, we have

$$f(0) = \int_{\mathcal{B}'} \hat{f}(y) d\lambda_{\mathcal{B}'}(y).$$

From here, we see that

$$\begin{aligned} f(x) &= (\tau_{-x}f)(0) = \int_{\mathcal{B}'} \widehat{(\tau_{-x}f)}(y)\, d\lambda_{\mathcal{B}'}(y) \\ &= \int_{\mathcal{B}'} \hat{f}(y)\, e^{2\pi i\langle x,y\rangle} d\lambda_{\mathcal{B}'}(y). \end{aligned}$$

It follows that $\mathfrak{F}$ is surjective, $\hat{\mathcal{S}}(\mathcal{B}') = \mathcal{S}(\mathcal{B}')$ and $\mathfrak{F}^{-1}$ is continuous. To prove (5), first note that $\mathcal{S}(\mathcal{B})$ is dense in $L^1[\mathcal{B}]$, so that every $f \in L^1[\mathcal{B}]$ is the limit of a sequence $\{f_n\}$ in $\mathcal{S}(\mathcal{B})$. Since, for every $f \in \mathcal{S}(\mathcal{B})$, $\hat{f} \in \mathcal{S}(\mathcal{B}') \subset \mathbb{C}_0(\mathcal{B}')$ and $\hat{f}_n \in \mathbb{C}_0(\mathcal{B}')$, we are done. □

Corollary 2.94. *Let $\mathcal{B}$ be a Banach space with an S-basis.*

(1) *If $\mathcal{B}$ is equivalent to ℓ_p, $1 \le p \le 2$ and $\nu_{\mathcal{B}}$ is the universal representation of Gaussian measure on $\mathcal{B}$, then $\hat{\nu}_{\mathcal{B}}(y)$ does not exist. However,*

$$\hat{\nu}^n_{\mathcal{B}}(y) = e^{-\pi\sum_{i=1}^{n}|y_i|^2} \otimes \left(\bigotimes_{i=n+1}^{\infty} \chi_I(y_i) \right),$$

is well defined for every $n \in \mathbb{N}$.

(2) *If $\mathcal{B}$ is equivalent to ℓ_p, $2 < p < \infty$ then the universal representation of Gaussian measure $\nu_{\mathcal{B}}$ does not exist on $\mathcal{B}$. However, the universal representation of Gaussian measure*

$$\hat{\nu}_{\mathcal{B}}(y) = e^{-\pi\sum_{i=1}^{\infty}|y_i|^2},$$

does exist on $\mathcal{B}'$.

Recall that $g^c(x)$ be the complex conjugate of $g(x)$

Theorem 2.95. *The mapping $\mathfrak{F} : \mathcal{S}(\mathcal{B}) \to \mathcal{S}(\mathcal{B}')$ extends to a continuous linear isometry of $U : L^2[\mathcal{B}] \to L^2[\mathcal{B}']$ satisfying the following:*

(1) $\int_{\mathcal{B}} f(x)g^c(x)d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}'} \hat{f}(y)\hat{g}^c(y)d\lambda_{\mathcal{B}'}(y)$.

(2) $\int_{\mathcal{B}} |f(x)|^2\, d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}'} |\hat{f}(y)|^2 d\lambda_{\mathcal{B}'}(y)$.

Proof. From the inversion property, we have that

$$\begin{aligned} &\int_{\mathcal{B}} f(x)g^c(x)d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}} g^c(x)\left\{ \int_{\mathcal{B}'} \hat{f}(y)e^{2\pi i<x,y>} d\lambda_{\mathcal{B}'}(y) \right\} d\lambda_{\mathcal{B}}(x) \\ &= \int_{\mathcal{B}'} \hat{f}(y)\left\{ \int_{\mathcal{B}} g^c(x)e^{2\pi i<x,y>} d\lambda_{\mathcal{B}}(x) \right\} d\lambda_{\mathcal{B}'}(y) \end{aligned}$$

The last term in parenthesis is the complex conjugate of $\hat{g}(y)$, so we have Parseval's formula:

$$\int_{\mathcal{B}} f(x)g^c(x)d\lambda_{\mathcal{B}}(x) = \int_{\mathcal{B}'} \hat{f}(y)\hat{g}^c(y)d\lambda_{\mathcal{B}^*}(y). \tag{2.9}$$

If we set $g = f$, we get our second result. Since $\mathcal{S}(\mathcal{B})$ is dense in $L^2[\mathcal{B}]$ and $\mathcal{S}(\mathcal{B}')$ is dense in $L^2[\mathcal{B}']$, we see from Eq. (2.8) that, relative to the L^2 metric, the mapping $\mathfrak{F}$, $f :\to \hat{f}$ is a linear isometry of $\mathcal{S}(\mathcal{B}) \subset L^2[\mathcal{B}]$ onto $\mathcal{S}(\mathcal{B}') \subset L^2[\mathcal{B}']$ (onto by inversion). It now follows that $\mathfrak{F}$ has a unique continuous extension $U = \bar{\mathfrak{F}}$, $U : L^2[\mathcal{B}] \to L^2[\mathcal{B}']$. □

2.6. Application

In this section we explore two of the natural applications of results from this chapter.

2.6.1. Pontryagin Duality. Let G be a locally compact abelian (LCA) group (c.f., $\mathbb{R}^n$). In this section, we follow the conventions of group theory, so that the functional pairing between G and its dual $\hat{G}$ is $\langle x, x^* \rangle$, where $x \in G$ and $x^* \in \hat{G}$.

Recall the following theorem of Haar (Theorem 2.1).

Theorem 2.96. *If G is an LCA group and $\mathfrak{B}(G)$ is the Borel σ-algebra of subsets of G, then there is a nonnegative regular translation invariant measure m (Haar-measure), which is unique up to multiplication by a constant.*

Definition 2.97. A complex valued function $\alpha : G \to \mathbb{C}$ on an LCA group is called a character on G provided that α is a homomorphism and $|\alpha(g)| = 1$ for all $g \in G$.

The set of all continuous characters of G defines a new group $\hat{G}$, called the dual group of G and $(\alpha_1 + \alpha_2)(g) = \alpha_1(g) \cdot \alpha_2(g)$. If we define a map $\gamma : G \to \hat{\hat{G}}$, by $\gamma_g(\alpha) = \alpha(g)$, then the following theorem was first proven by Pontryagin (see [PO] or Rudin [RU1]):

Theorem 2.98 (Pontryagin Duality Theorem)**.** *If G is an LCA group, then the mapping $\gamma : G \to \hat{\hat{G}}$ is an isomorphism of topological groups.*

Pontryagin Duality identifies those groups that are the character groups of their character groups. If the group is not locally compact there is no Haar measure. However Kaplan [KA1] has shown that the class of topological abelian groups for which the Pontryagin Duality holds is closed under the operation of taking infinite products of groups. This result immediately implies that this class is larger than the class of locally compact abelian groups because the infinite product of locally compact non-compact groups (for example, R^∞ or $\mathcal{B}$) is not locally compact (see also [KA2]).

2.6.2. Uniformly Convex Banach Spaces. We want to show that Pontryagin duality theory can be extended to all uniformly convex Banach spaces with an S-basis. The next theorem is a recasting of Theorem 2.90, (3) and (4).

Theorem 2.99. *If $\mathcal{B}$ is a uniformly convex Banach space with an S-basis, then $\mathcal{B}$ and $\mathcal{B}'$ are duals as character groups (i.e., $\mathcal{B}' = \hat{\mathcal{B}}$).*

Proof. From Theorem 2.90, if $x^* \in \mathcal{B}'$, we have:

$$[\mathfrak{F}(f)](x^*) = \hat{f}(x^*) = \int_{\mathcal{B}} \exp\{-2\pi i \langle x, x^* \rangle\} f(x) d\lambda_{\mathcal{B}}(x), \tag{2.10}$$

where $\langle x, x^* \rangle$ is the pairing between $\mathcal{B}$ and $\mathcal{B}'$. From the Plancherel part of Theorem 2.95,

$$\|\hat{f}\|_2^2 = \|f\|_2^2 .$$

It follows that $\mathcal{B}$ and $\mathcal{B}'$ are duals as character groups and

$$f(x) = \int_{\mathcal{B}'} \exp\{2\pi i \langle x, x^* \rangle\} \hat{f}(x^*) d\lambda_{\mathcal{B}'}(x^*).$$

□

Since $\mathcal{B}_J^n = \mathcal{B}_J \cap \mathbb{R}_I^n$, we can represent $\hat{f}_n$ directly as a mapping from $L^2[\mathcal{B}_J^n, \lambda_{\mathcal{B}}] \to L^2[\mathcal{B'}_J^n, \lambda_{\mathcal{B}'}]$, by

$$[\mathfrak{F}(f_n)](x^*) = \hat{f}_n(x^*) = \int_{\mathcal{B}} \exp\{-2\pi i \langle x, x^* \rangle_n\} f_n(x) d\lambda_{\mathcal{B}}(x),$$

where $\langle x, x^* \rangle_n$ is the restricted pairing of x and x^* to $\mathcal{B}_J^n$ and $\mathcal{B'}_J^n$ respectively.

If we define $y(\cdot)$ mapping $\mathcal{B} \to \mathbb{C}$, by $y(x) = exp\{-2\pi i \langle y, x^* \rangle\}$, then $y(x)$ is a character of $\mathcal{B}$. Furthermore, it is easy to see that $(y_1 + y_2)(x) = y_1(x) \cdot y_2(x)$. We now have the extension of the Pontryagin Duality Theorem to all uniformly convex Banach spaces with an S-basis.

Theorem 2.100. *If $\mathcal{B}$ is a uniformly convex Banach space with an S-basis, then the mapping $\gamma_x : \mathcal{B} \to \hat{\hat{\mathcal{B}}}$, defined by $\gamma_x(y) = y(x)$, is an isomorphism of topological groups.*

In case $\mathcal{B} = \mathcal{H}$, is a Hilbert space, we can replace Eq. (2.10) by

$$\hat{f}(x^*) = \mathfrak{F}[f](x^*) = \int_{\mathcal{H}} \exp\{-2\pi i \langle x, x^* \rangle_{\mathcal{H}}\} f(x) d\lambda_{\mathcal{H}}(x), \qquad (2.11)$$

so that $\mathcal{H}$ is self-dual (as expected), when we identify $\mathcal{H}$ with $\mathcal{H}'$. From Eq. (2.11), we also get the expected result that:

$$\mathfrak{F}\left[\exp\{-\pi \left\| x \right\|_{\mathcal{H}}^2\}\right](x^*) = \exp\{-\pi \left\| x^* \right\|_{\mathcal{H}'}^2\}.$$

2.7. The Diffusion Equation

In order to see how the existence of $\lambda_{\mathcal{B}}$ impacts the theory of partial differential equations, we show how to directly solve the initial value problem for the diffusion equation in infinitely many variables.

First, for the diffusion equation on $\mathbb{R}^n$ we have (see Evans [EV]):

$$u_t(\mathbf{x}, t) - \tfrac{1}{2}\Delta_n u(\mathbf{x}, t) = 0, \; u(\mathbf{x}, 0) = \phi_0(\mathbf{x}) \in \mathbb{C}_0(\mathbb{R}^n).$$

It has a solution

$$u(\mathbf{x}, t) = \tfrac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}^n} \exp\left\{-\tfrac{|\mathbf{x}-\mathbf{y}|^2}{2t}\right\} \phi_0(\mathbf{y}) d\lambda_n(\mathbf{y}).$$

A proof of the following is the same as in [EV] (see Theorem 1, p. 47).

Theorem 2.101. *Let $\phi_0 \in \mathbb{C}_0(\mathbb{R}^n)$ and let $u(\mathbf{x}, t)$ be as defined above. Then*

(1) $u \in \mathbb{C}_0^\infty(\mathbb{R}^n, (0, \infty))$,

(2)
$$u_t(\mathbf{x}, t) - \tfrac{1}{2}\Delta u(\mathbf{x}, t) = 0, \; (\mathbf{x} \in \mathbb{R}^n, \; t > 0)$$
and for each $\mathbf{y} \in \mathbb{R}^n$,

(3)
$$\lim_{(\mathbf{x}, t) \to (\mathbf{x}, 0)} u(\mathbf{x}, t) = \phi_0(\mathbf{x})$$

This is certainly a very nice result for an equation with only continuous initial data.

2.7.1. Direct Representation in Hilbert Space. We now consider our universal representation, which provides less smoothness for the solution, but has the nice property that it does not depend on the dimension of the space. Let $\mathbf{y} = \mathbf{x} - \mathbf{z}\sqrt{2\pi t}$. In this case, we can write the solution $u(\mathbf{x}, t)$ as:

$$u(\mathbf{x}, t) = \int_{\mathbb{R}^n} \phi_0\left(\mathbf{x} - \sqrt{2\pi t}\,\mathbf{z}\right) \exp\left\{-\pi\,|\mathbf{z}|_n^2\right\} d\lambda_n(\mathbf{z}).$$

We know that

$$e^{-\pi\|\mathbf{z}\|_n^2} = e^{-\pi|\mathbf{z}|_n^2} \otimes \left(\bigotimes_{k=n+1}^{\infty} \chi_I(z_k)\right).$$

It follows that

$$\int_{\mathbb{R}_I^n} e^{-\pi\|\mathbf{z}\|_n^2} d\lambda_\infty(\mathbf{z}) = \int_{\mathbb{R}^n} e^{-\pi|\mathbf{z}|_n^2} d\lambda_n(\mathbf{z}) \times \prod_{k=n+1}^{\infty} \int_I d\lambda(z)$$
$$= \int_{\mathbb{R}^n} e^{-\pi|\mathbf{z}|_n^2} d\lambda_n(\mathbf{z}) = 1.$$

We now observe that the equation for $u(\mathbf{x}, t)$ makes sense for any separable Hilbert space $\mathcal{H}$, so that:

$$u(\mathbf{x}, t) = \int_{\mathcal{H}} \phi_0\left(\mathbf{x} - \sqrt{2\pi t}\,\mathbf{z}\right) \exp\left\{-\pi\,\|\mathbf{z}\|_{\mathcal{H}}^2\right\} d\lambda_{\mathcal{H}}(\mathbf{z}). \tag{2.12}$$

Theorem 2.102. *If $u(\mathbf{x}, t)$ is defined by Eq. (2.12), it satisfies:*

$$u_t(\mathbf{x}, t) - \tfrac{1}{2}\Delta u(\mathbf{x}, t) = 0,\ u(\mathbf{x}, 0) = \phi_0(\mathbf{x}) \in \mathbb{C}_0^2(\mathcal{H}) \tag{2.13}$$

with

$$\Delta = \sum_{i=1}^{\infty} \frac{\partial^2}{\partial x_i^2}.$$

Proof. To see that Eq. (2.13) is satisfied, let $y_i = x_i - z_i\sqrt{2\pi t}$, so that

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial y_i}, \quad \frac{\partial}{\partial z_i} = -\sqrt{2\pi t}\frac{\partial}{\partial y_i}, \quad \frac{\partial}{\partial t} = \frac{\partial y_i}{\partial t}\frac{\partial}{\partial y_i} = -\sqrt{\tfrac{\pi}{2t}}\mathbf{z}\cdot\nabla.$$

Now, let:

$$c_1 = \sup_x |\phi_0|\,,\ c_2 = \sup_x \|\nabla\phi_0\|\,,\ c_3 = \sup_{i,x}\left|\frac{\partial^2\phi_0}{\partial x_i^2}\right|,$$
$$c = \max\{c_1, c_2, c_3\}\,.$$

so that

$$\begin{aligned}
\left|\phi_0\left(\mathbf{x}-\mathbf{z}\sqrt{2\pi t}\right)e^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2}\right| &\leqslant ce^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2},\\
\left|\frac{\partial}{\partial t}\phi_0\left(\mathbf{x}-\mathbf{z}\sqrt{2\pi t}\right)e^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2}\right| &\leqslant c\left\|\mathbf{z}\right\|_{\mathcal{H}}e^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2},\\
\left|\frac{\partial}{\partial x_i}\phi_0\left(\mathbf{x}-\mathbf{z}\sqrt{2\pi t}\right)e^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2}\right| &\leqslant ce^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2},\\
\left|\frac{\partial^2}{\partial x_i^2}\phi_0\left(\mathbf{x}-\mathbf{z}\sqrt{2\pi t}\right)e^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2}\right| &\leqslant ce^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2}.
\end{aligned}$$

These bounds allow us to justify differentiating inside the integral. If we let $\Omega_1=\mathbb{R}$, $\Omega_2=\mathcal{H}^i$ and use Fubini's theorem, we have:

$$\begin{aligned}
&\frac{\partial^2 u(\mathbf{x},t)}{\partial x_i^2}\\
&=\int_{\mathcal{H}}\left[\frac{\partial^2}{\partial x_i^2}\varphi_0(\mathbf{x}-\mathbf{z}\sqrt{2\pi t})\right]e^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2}d\lambda_{\mathcal{H}}(\mathbf{z})\\
&=\int_{\mathcal{H}^i}e^{-\pi\|\mathbf{z}\|_i^2}\left\{\int_{-\infty}^{\infty}\left[\frac{\partial^2}{\partial x_i^2}\varphi_0(\mathbf{x}-\mathbf{z}\sqrt{2\pi t})\right]\right.\\
&\qquad\left.\times\, e^{-\pi z_i^2}d\lambda(z_i)\right\}d\lambda_{\mathcal{H}}^i(\mathbf{z})
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\partial u(\mathbf{x},t)}{\partial t}\\
&=\sum_{i=1}^{\infty}\left\{\int_{\mathcal{H}}\left[\frac{\partial y_i}{\partial t}\frac{\partial}{\partial y_i}\varphi_0(\mathbf{x}-\mathbf{z}\sqrt{2\pi t})\right]e^{-\pi\|\mathbf{z}\|_{\mathcal{H}}^2}d\lambda_{\mathcal{H}}(\mathbf{z})\right\}\\
&=-\sqrt{\frac{\pi}{2t}}\sum_{i=1}^{\infty}\left\{\int_{\mathcal{H}^i}e^{-\pi\|\mathbf{z}\|_i^2}\left[\int_{-\infty}^{\infty}z_i\left(\frac{\partial}{\partial x_i}\varphi_0(\mathbf{x}-\mathbf{z}\sqrt{2\pi t})\right)\right.\right.\\
&\qquad\left.\left.\times\, e^{-\pi z_i^2}d\lambda(z_i)\right]d\lambda_{\mathcal{H}}^i(\mathbf{z})\right\}
\end{aligned}$$

Thus, combining terms we have:

$$\begin{aligned}
&\frac{\partial u(\mathbf{x},t)}{\partial t}-\frac{1}{2}\sum_{i=1}^{\infty}\frac{\partial^2 u(\mathbf{x},t)}{\partial x_i^2}\\
&=-\sqrt{\frac{\pi}{2t}}\sum_{i=1}^{\infty}\left\{\int_{\mathcal{H}^i}e^{-\pi\|\mathbf{z}\|_i^2}\left[\int_{-\infty}^{\infty}z_i\left(\frac{\partial}{\partial x_i}\varphi_0(\mathbf{x}-\mathbf{z}\sqrt{2\pi t})\right)\right.\right.\\
&\qquad\left.\left.\times\, e^{-\pi z_i^2}d\lambda(z_i)\right]d\lambda_{\mathcal{H}}^i(\mathbf{z})\right\}
\end{aligned}$$

$$-\frac{1}{2}\sum_{i=1}^{\infty}\left\{\int_{\mathcal{H}^i} e^{-\pi\|\mathbf{z}\|_i^2}\left[\int_{-\infty}^{\infty}\left(\frac{\partial^2}{\partial x_i^2}\varphi_0(\mathbf{x}-\mathbf{z}\sqrt{2\pi t})\right)\right.\right.$$
$$\left.\left.\times\ e^{-\pi z_i^2}\,d\lambda(z_i)\right]d\lambda_{\mathcal{H}}^i(\mathbf{z})\right\}.$$

If we integrate by parts in the inner integral above for each i, with

$$d\alpha = -z_i e^{-\pi z_i^2}\,d\lambda(z_i), \quad \alpha = \frac{1}{2\pi}e^{-\pi z_i^2},$$

$$\begin{aligned}
\beta = \frac{\partial\phi_0}{\partial x_i}, \quad d\beta &= \frac{\partial}{\partial z_i}\frac{\partial\phi_0}{\partial x_i}d\lambda(z_i)\\
&= \frac{\partial y_i}{\partial z_i}\frac{\partial}{\partial y_i}\frac{\partial\phi_0}{\partial x_i}d\lambda(z_i)\\
&= \frac{\partial y_i}{\partial z_i}\frac{\partial}{\partial x_i}\frac{\partial\phi_0}{\partial x_i}d\lambda(z_i)\\
&= \frac{\partial y_i}{\partial z_i}\frac{\partial^2\phi_0}{\partial x_i^2}d\lambda(z_i)\\
&= -\sqrt{2\pi t}\frac{\partial^2\phi_0}{\partial x_i^2}d\lambda(z_i).
\end{aligned}$$

so that

$$\begin{aligned}
&-\sqrt{\frac{\pi}{2t}}\int_{-\infty}^{\infty} z_i\frac{\partial\varphi_0}{\partial x_i}e^{-\pi z_i^2}\,d\lambda(z_i)\\
&= \frac{1}{2\pi}\sqrt{\frac{\pi}{2t}}e^{-\pi z_i^2}\left.\frac{\partial\varphi_0}{\partial x_i}\right|_{-\infty}^{\infty} + \frac{1}{2\pi}\sqrt{\frac{\pi}{2t}}\int_{-\infty}^{\infty}\sqrt{2\pi t}\frac{\partial^2\varphi_0}{\partial x_i^2}e^{-\pi z_i^2}\,d\lambda(z_i)\\
&= \frac{1}{2}\int_{-\infty}^{\infty}\frac{\partial^2\varphi_0}{\partial x_i^2}e^{-\pi z_i^2}\,d\lambda(z_i).
\end{aligned}$$

Using this result, we see that

$$\frac{\partial u(\mathbf{x},t)}{\partial t} - \frac{1}{2}\sum_{i=1}^{\infty}\frac{\partial^2 u(\mathbf{x},t)}{\partial x_i^2} = 0.$$

From Eq. (2.12) and the dominating convergence theorem, it is clear that

$$\lim_{(\mathbf{x},t)\to(\mathbf{x},0)} u(\mathbf{x},t) = \phi_0(\mathbf{x})$$

□

Remark 2.103. It is easy to see that the same approach above also applies to the Ornstein–Uhlenbech equation.

It is more interesting to note that a restricted version of Theorem 2.102 is true for every Banach space with an S-basis. The proof is left to the interested reader. For a more general result, see the Appendix (Sect. 6.7) of Chap. 6.

Theorem 2.104. *Let $\mathcal{B}$ be a Banach space with an S-basis and let $\phi_0 \in \mathbb{C}_0^2[\mathcal{B}]$. If we are given the equation:*

$$u_t(\mathbf{x}, t) = \Delta u(\mathbf{x}, t), \quad u(\mathbf{x}, 0) = \phi_0(\mathbf{x}) \in \mathbb{C}_0^2(B), \tag{2.14}$$

then there exist a family of functions $\{\phi_0^n(\mathbf{x})\} \subset \mathbb{C}_0^2[\mathcal{B}]$ with each $\phi_0^n(\mathbf{x}) \in \mathbb{C}_0^2[\mathcal{B}^n]$, such that

(1) *the functions $\phi_0^n(\mathbf{x}) \to \phi_0(\mathbf{x})$ and*

(2) *the second partials, $\frac{\partial^2 \phi_0^n}{\partial x_i^2} \to \frac{\partial^2 \phi_0}{\partial x_i^2}$.*

Furthermore, for each $n \in \mathbb{N}$, the function

$$\begin{aligned} u^n(\mathbf{x}, t) &= \int_{\mathcal{B}^n} \phi_0^n\left(\mathbf{x} - \sqrt{2\pi t}\mathbf{z}\right) d\nu_{\mathcal{B}}(\mathbf{z}) \\ &= \int_{\mathcal{B}^n} \phi_0^n\left(\mathbf{x} - \sqrt{2\pi t}\mathbf{z}\right) e^{-\pi \sum_{i=1}^n |z_i|^2} \otimes \left(\otimes_{i=n+1}^\infty h(z_i)\right) d\lambda_{\mathcal{B}}(\mathbf{z}) \\ &= \int_{\mathbb{R}^n} \phi_0^n\left(\mathbf{x} - \sqrt{2\pi t}\mathbf{z}\right) e^{-\pi \sum_{i=1}^n |z_i|^2} d\lambda_n(\mathbf{z}) \end{aligned} \tag{2.15}$$

solves the initial value problem:

$$u_t^n(\mathbf{x}, t) - \tfrac{1}{2}\Delta_n u^n(\mathbf{x}, t) = 0, \; u^n(\mathbf{x}, 0) = \phi_0^n(\mathbf{x}).$$

References

[BA1] R. Baker, “Lebesgue measure” on $\mathbb{R}^\infty$. Proc. Am. Math. Soc. **113**, 1023–1029 (1991)

[BA2] R. Baker, “Lebesgue measure” on $\mathbb{R}^\infty$, II. Proc. Am. Math. Soc. **132**, 2577–2591 (2004)

[BE] W. Beckner, Inequalities in Fourier analysis. Ann. Math. **102**, 159–182 (1975)

[BL] H.J. Brascamp, E.H. Lieb, Best constants in Young’s inequality, its converse, and its generalization to more than three functions. Adv. Math. **20**, 151–173 (1976)

[BO] V.I. Bogachev, *Differentiable Measures and the Malliavin Calculus*. Mathematical Surveys and Monographs, vol. 164 (American Mathematical Society, Providence, 2010)

[DP] G. Da Prato, *Kolmogorov Equations for Stochastic PDEs*. Advanced Courses in Mathematics - CRM (Barcelona) (Birkhäuser, Boston, 2004)

[DS] N. Dunford, J.T. Schwartz, *Linear Operators Part I: General Theory*. Wiley Classics Edition (Wiley, New York, 1988)

[EV] L.C. Evans, *Partial Differential Equations*. Graduate Studies in Mathematics, vol. 18 (American Mathematical Society, Providence, 1998)

[HA] A. Haar, Der Massbegriff in der Theorie der kontinuierlichen Gruppe. Ann. Math. **34**, 147–169 (1933)

[HI] D. Hill, σ-Finite invariant measures on infinite product spaces. Trans. Am. Math. Soc. **153**, 347–370 (1971)

[KA1] S. Kaplan, Extensions of Pontjagin duality I: infinite products. Duke Math. J. **15**, 649–659 (1948)

[KA2] S. Kaplan, Extensions of Pontjagin duality II: direct and inverse sequences. Duke Math. J. **17**, 419–435 (1950)

[KP1] A.P. Kirtadze, G.R. Pantsulaia, Invariant measures in the space $\mathbf{R}^N$. Soobshch. Akad. Nauk Gruzii **141**, 273–276 (1991) [in Russian]

[KP2] A.P. Kirtadze, G.R. Pantsulaia, Lebesgue nonmeasurable sets and the uniqueness of invariant measures in infinite-dimensional vector spaces. Proc. A. Razmadze Math. Inst. **143**, 95–101 (2007)

[KK] K. Kodaira, S. Kakutani, A nonseparable translation-invariant extension of Lebesgue measure space. Ann. Math. **52**, 574–579 (1950)

[KO] A.N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Springer, Vienna, 1933)

[GL] G. Leoni, *A First Course in Sobolev Spaces.* Graduate Studies in Mathematics, vol. 105 (American Mathematical Society, Providence, 2009)

[OX] J.C. Oxtoby, Invariant measures in groups which are not locally compact. Trans. Am. Math. Soc. **60**, 215–237 (1946)

[PA] G. Pantsulaia, *Invariant and Quasiinvariant Measures in Infinite-Dimensional Topological Vector Spaces* (Nova Science Publishers, New York, 2007)

[PO] L. Pontryagin, *Topological Groups* (Princeton University Press, Princeton, 1946)

[RH] G.E. Ritter, E. Hewitt, Elliott-Morse measures and Kakutani's dichotomy theorem. Math. Z. **211**, 247–263 (1992)

[RO] H.L. Royden, *Real Analysis*, 2nd edn. (Macmillan Press, New York, 1968)

[RU1] W. Rudin, *Fourier Analysis on Groups* (Wiley, New York, 1990)

[SU] V.N. Sudakov, Linear sets with quasi-invariant measure. Dokl. Akad. Nauk SSSR **127**, 524–525 (1959) [in Russian]

[UM] Y. Umemura, On the infinite dimensional Laplacian operator. J. Math. Kyoto Univ. **4**, 477–492 (1964/1965)

[V] A.M. Vershik, Duality in the theory of measure in linear spaces. Sov. Math. Dokl. **7**, 1210–1214 (1967) [English translation]

[V1] A.M. Vershik, Does there exist the Lebesgue measure in the infinite-dimensional space? Proc. Steklov Inst. Math. **259**, 248–272 (2007)

[V2] A.M. Vershik, The behavior of Laplace transform of the invariant measure on the hyperspace of high dimension. J. Fixed Point Theory Appl. **3**, 317–329 (2008)

[VN1] J. von Neumann, The uniqueness of Haar's measure. Rec. Math. Mat. Sbornik N.S. **1**, 721–734 (1936)

[WE] A. Weil, *L'intégration dans les groupes topologiques et ses applications*. Actualités Scientifiques et Industrielles, vol. 869, Paris (1940)

[WSRM] N. Wiener, A. Siegel, W. Rankin, W.T. Martin, *Differential Space, Quantum Systems, and Prediction* (MIT Press, Cambridge, 1966)

[YA1] Y. Yamasaki, Translationally invariant measure on the infinite-dimensional vector space. Publ. Res. Inst. Math. Sci. **16**(3), 693–720 (1980)

[YA] Y. Yamasaki, *Measures on Infinite-Dimensional Spaces* (World Scientific, Singapore, 1985)

Chapter 3

HK-Integral and HK-Spaces

In this chapter we discuss the various integrals that integrate nonabsolutely integrable functions and extend the Lebesgue integral. Our objective is the Henstock–Kurzweil integral (HK-integral). It is the most well developed and the easiest to learn. We use it for the Feynman operator calculus in Chap. 7 and the theory of path integrals in Chap. 8. A second objective is a new class of separable Banach spaces, KS^p $1 \leq p \leq \infty$, which include the HK-integrable functions and contain the L^p spaces, $1 \leq p \leq \infty$, as continuous dense and compact embeddings. The Hilbert space KS^2 is the natural for Feynman's (path integral) formulation of quantum mechanics, discussed in Chap. 8.

In this chapter we will suppress the notation R_I^n, $n \geq 1$ and assume that I is understood. However, we will always use λ_∞ for our measure. This is to remind the reader that the results have direct extensions to the infinite-dimensional case. The extensions will be accomplished in Chap. 6, as a natural application of infinite tensor product Banach space theory.

Summary. The first section provides a brief introduction to the history of finitely additive measures and its place in analysis. Then we introduce and discuss the most important integrals generated

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_3

by finitely additive measures that integrate nonabsolutely integrable functions. We mainly focus on the HK-integral developed by Kurzweil [KW] and Henstock [HS]. Loosely speaking, one uses a version of the Riemann integral with the interior points chosen first, while the size of the base rectangle around any interior point is determined by an arbitrary positive function defined at that point. A fairly complete discussion of the one-dimensional HK-integral is followed by a discussion of the general theory on $\mathbb{R}^n$ and a proof that the Lebesgue integral is a special case. The general HK-integral can be found in Henstock [HS], Tuo-Yeong [TY1], or Pfeffer (see [PF] and [PF1]).

We then turn to the construction and study of the main class of Banach spaces used in the book, KS^p $1 \leq p \leq \infty$. These spaces are natural for the HK-integrable functions and contain the L^p spaces, $1 \leq p \leq \infty$, as continuous dense and compact embeddings.

We have also added a number of interesting topics that we hope will attract a new generation of researchers to this new field of inquiry. These topics are not required to understand (or use) the Feynman operator calculus or path integrals, and can be omitted by those with limited objectives. In the third section, we construct a very interesting class of separable Banach spaces $SD^p[\mathbb{R}^n] 1 \leq p \leq \infty$ which also contain the nonabsolutely integrable functions. These spaces contain the generalized Sobolev spaces and the test functions $\mathcal{D}[\mathbb{R}^n]$, as dense continuous embeddings. In addition, they have the remarkable property that, for any multi-index α, $\|D^\alpha \mathbf{u}\|_{SD} = \|\mathbf{u}\|_{SD}$, where D is the distributional derivative. For this reason, we call them strong distribution Banach spaces. As an application, we obtain a priori bounds for the important nonlinear term of the classical Navier–Stokes initial-value problem. Finally, we introduce yet another family of spaces, which include the HK-integrable functions and contain the space of functions of bounded mean oscillation as a continuous dense embedding.

Background

The standard analysis course gives the distinct impression that the Riemann integral is of limited value. In addition, one acquires an unconscious but natural bias and unease concerning the use of finitely additive set functions as a basis for integration theory. However, in analysis proper finitely additive measures are not seen as unwanted guests. In some cases finitely additive measures appear naturally and have been advocated in others. As noted by Diestel and Uhl [DU], interest in these measures date back to the early works of Hildebrandt [HI] and Fichtenholtz and Kantorovich [FK]. In this regard, we also mention the important works of Alexandroff [AX], Bochner [BO1], [BO2], Dunford and Schwartz [DS], and Yosida and Hewitt [YH]. In response to the work of Leader on finitely additive measures [LE], Diestel and Uhl stated "... one might be convinced that countable additivity is more of a hinderance than help" (see [DU], p. 32).

In probability theory, Blackwell and Dubins, and Dubins and Prikry (see [BD], [DUK], and [DU]) argue forcefully for the intrinsic advantages in using finite additivity in the basic axioms of probability theory. Their position is also supported by de Finetti's [DFN] penetrating analysis of the foundations of probability theory.

3.1. The HK-Integral

In this section, we develop the elementary HK-integral in one dimension. This will make it easy to obtain a sense of the differences relative to the Lebesgue integral. The case for $n > 1$ will be discussed later. For comparison, we first define the Riemann integral. (A very nice elementary account of the HK-integral may be found in Bartle [BR], while McShane [McS] uses this approach to give a nice elementary account of the Lebesgue integral.)

3.1.1. One-Dimensional Riemann Integral.

Definition 3.1. Let $[a, b] \subset \mathbb{R}$. The set $\mathcal{P} = \{([t_{i-1}, t_i], \tau_i) : 1 \leqslant i \leqslant n\}$ is called a tagged partition, where the $\tau_i \in [t_{i-1}, t_i]$ are called the tags. We call $\mathcal{P}$ a tagged δ partition if, for $1 \leqslant i \leqslant n$, $t_i - t_{i-1} < \delta$.

Definition 3.2. Let a tagged partition $\mathcal{P} = \{([t_{i-1}, t_i], \tau_i) : 1 \leqslant i \leqslant n\}$ be given. If $f : [a, b] \to \mathbb{R}$, we define the Riemann sum by (with $\Delta t_i = t_i - t_{i-1}$)

$$\Re(f, \mathcal{P}) = \sum_{i=1}^{n} f(\tau_i)\Delta t_i.$$

We say that the number $I \in \mathbb{R}$ is the Riemann integral of f on the interval $[a, b]$ if for each $\varepsilon > 0$, there exists a $\delta > 0$ such that, if $\mathcal{P}$ is a tagged δ partition,

$$\left| I - \sum_{i=1}^{n} f(\tau_i)\Delta t_i \right| < \varepsilon.$$

3.1.2. One-Dimensional HK-Integral.

Definition 3.3. Let $[a,b] \subset \mathbb{R}$, let $\delta(t)$ map $[a,b] \to (0,\infty)$, we say that the tagged partition $\mathcal{P} = \{([t_{i-1}, t_i], \tau_i) : 1 \leqslant i \leqslant n\}$ is a HK-δ partition if, for $1 \leqslant i \leqslant n$, $t_i - t_{i-1} < \delta(\tau_i)$.

Remark 3.4. Gordon defines the phrase nearly everywhere (n.e.) to mean "except for a countable set."

Definition 3.5. The function $f : [a,b] \to \mathbb{R}$ is said to have an HK-integral if there is a number I such that, for each $\varepsilon > 0$, there exists a function δ from $[a,b] \to (0,\infty)$ such that, whenever $\mathcal{P}$ is a HK-δ partition, then

$$\left| \sum_{i=1}^{n} \Delta t_i f(\tau_i) - I \right| < \varepsilon.$$

In this case, we also write $I = (HK) \int_a^b f(t) d\lambda_\infty(t)$.

In the next two theorems, we see how the HK process extends the Riemann integral and shows in what sense we can think of the HK-integral as the reverse of the derivative. The first result assumes that $F : [a,b] \to \mathbb{R}$ is differentiable at each point, with $f = F'$, while the second only assumes that $f = F'$ (n.e.) on $[a,b]$.

Theorem 3.6. *Let $F : [a,b] \to \mathbb{R}$ be continuous. If F is differentiable at each point of $[a,b]$ and $F'(t) = f(t)$ on $[a,b]$, then $f(t)$ is HK-integrable on $[a,b]$ and*

$$(HK) \int_a^b f(t) d\lambda_\infty(t) = F(b) - F(a).$$

Proof. Since $f(t) = F'(t)$ for $t \in [a,b]$, given $\varepsilon > 0$ there exists a function $\delta_\varepsilon(t)$ such that, for each $t \in [a,b]$, if $s \in [a,b]$ and $0 < |s-t| < \delta_\varepsilon(t)$, then

$$|[F(t) - F(s)] - (s-t) f(t)| \leq \varepsilon |s-t|.$$

It follows that, if $a \leq s \leq t \leq r \leq b$ and $0 < r - s \leq \delta_\varepsilon(t)$,

$$\begin{aligned} &|F(r) - F(s) - (r-s) f(t)| \\ &\leqslant |F(r) - F(t) - (r-t) f(t)| + |F(t) - F(s) - (t-s) f(t)| \\ &\leqslant \varepsilon(r-t) + \varepsilon(t-s) = \varepsilon(r-s). \end{aligned}$$

Let $\mathcal{P}$ be a δ_ε partition for $[a,b]$. Since $F(b)-F(a)=\sum_{i=1}^{n}[F(t_i)-F(t_{i-1})]$, we have that

$$\begin{aligned}
&\left|F(b)-F(a)-\sum_{i=1}^{n} f(\tau_i)\Delta t_i\right| \\
&= \left|\sum_{i=1}^{n}[F(t_i)-F(t_{i-1})-f(\tau_i)\Delta t_i]\right| \\
&\leqslant \sum_{i=1}^{n}|[F(t_i)-F(t_{i-1})-f(\tau_i)\Delta t_i]| \\
&\leqslant \sum_{i=1}^{n}\varepsilon\Delta t_i = \varepsilon\,(b-a)\,.
\end{aligned}$$

Since ε was arbitrary, we see that the integral exists. □

The following refinement of Theorem 3.6 shows the power of replacing a constant with an arbitrary positive function.

Theorem 3.7. *Let $F:[a,b]\to\mathbb{R}$ be continuous. If F is differentiable (n. e.) on $[a,b]$ and $F'(t)=f(t)$ on (a,b), then $f(t)$ is HK-integrable on $[a,b]$ and*

$$(HK)\int_a^b f(t)d\lambda_\infty(t)=F(b)-F(a).$$

Proof. Since $F(t)$ is continuous, let $\varepsilon>0$ be given, then for each $t\in[a,b]$, there is a $\delta_1(t)>0$, such that

$$|F(t)-F(s)|<\frac{\varepsilon}{6},$$

if $s\in[a,b]\cap(t-\delta_1(t),t+\delta_1(t))$.

Since $F(t)$ is differentiable (n.e), there is a countable set N, such that, for each $t\in(a,b)^N=[(a,b)\cap N]^c$, there exists a $\delta_2(t)>0$ so that

$$\left|F'(t)-\frac{F(u)-F(s)}{u-s}\right|\leqslant\frac{\varepsilon}{3(b-a)}$$

when $t\in[u,s]\subseteq(a,b)^N\cap(t-\delta_2(t),t+\delta_2(t))$.

Define $\delta_3(t)$ by

$$\delta_3(t)=\min_{t\in(a,b)}\left\{\delta_1(t),\delta_2(t),\tfrac{1}{2}(b-t),\tfrac{1}{2}(t-a)\right\}$$

and let

$$\delta(t) = \begin{cases} \delta_3(t), & \text{if } t \in (a,b)^N \\ \frac{\varepsilon}{6(|f(b)|+|f(a)|+1)}, & \text{if } t = a \vee b. \end{cases}$$

If $\mathcal{P}$ is a HK-δ partition then

$$\begin{aligned}
&\left|\sum_{i=1}^{n} \{f(\tau_i)\Delta t_i - (F(t_i) - F(t_{i-1}))\}\right| \\
&\leqslant |f(a)(t_1-a)-(F(t_1)-F(a))| + |f(b)(b-t_{n-1})-(F(b)-F(t_{n-1}))| \\
&+ \left|\sum_{i=2}^{n-1} \{f(\tau_i)\Delta t_i - (F(t_i) - F(t_{i-1}))\}\right| \\
&\leqslant [|f(a)|\,\Delta t_1 + |F(t_1) - F(a)|] + [|f(b)|\,\Delta t_n + |F(b) - F(t_{n-1})|] \\
&+ \sum_{i=2}^{n-1} \frac{\varepsilon \Delta t_i}{3(b-a)} \\
&< \frac{\varepsilon\,[|f(a)| + |f(b)|]}{6\,[|f(a)| + |f(b)|] + 1} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3(b-a)} \sum_{i=2}^{n-1} \Delta t_i < \varepsilon.
\end{aligned}$$

Since ε was arbitrary, $f(t)$ is HK-integrable on $[a, b]$ and

$$(HK)\int_a^b f(t)d\lambda_\infty(t) = F(b) - F(a).$$

□

Remark 3.8. Theorem 3.7 is still true if we only require $f = F'$ (a.e.) but the proof is a little more involved (and we do not need the result). However, neither Theorem 3.6 nor Theorem 3.7 is true for the Lebesgue integral.

Example 3.9. *The standard counter example is the continuous function $F(t)$ on $[0,1]$, with $F'(0) = 0$, $F'(t) = 2t sin(1/t^2) - (2/t)cos(1/t^2)$ for irrational $t \in (0,1]$ and $F'(t) = 0$ for rational t. Since $F'(t)$ exists (n.e) on $[0,1]$, it is easy to see that Theorem 3.7 is satisfied and $F(t) = t^2 sin(1/t^2)$. It follows that*

$$(HK)\int_0^1 \left(2t\sin\frac{1}{t^2} - 2\frac{1}{t}\cos\frac{1}{t^2}\right)dt = \sin 1.$$

For another interesting example, let

$$F'(t) = \begin{cases} \frac{2}{t^3} - \frac{2\cos t}{\sin^4 t} & \text{if } t \in (0, \frac{\pi}{2}] \\ 0 & \text{if } t = 0. \end{cases}$$

Then $f(t) = F'(t)$ *is HK-integrable and*

$$(HK)\int_0^{\frac{\pi}{2}} f(t)dt = \frac{2}{3} - \frac{4}{\pi^2}.$$

The proofs of the following are close enough to the Riemann case for the interested reader to fill in the details.

Theorem 3.10. *Let* $f(t), g(t)$ *be HK-integrable on* $[a,b]$ *and* $c \in \mathbb{C}$*, then*

(1) $f(t) + g(t)$ *is HK-integrable on* $[a,b]$ *and*

$$(HK)\int_a^b (f(t) + g(t))\, d\lambda_\infty(t) = (HK)\int_a^b f(t)d\lambda_\infty(t) + (HK)\int_a^b g(t)d\lambda_\infty(t).$$

(2) $cf(t)$ *is HK-integrable and*

$$(HK)\int_a^b cf(t)d\lambda_\infty(t) = c\left\{(HK)\int_a^b f(t)d\lambda_\infty(t)\right\}.$$

(3) *If* $f(t) \le g(t)$*, then*

$$(HK)\int_a^b f(t)d\lambda_\infty(t) \leqslant (HK)\int_a^b g(t)d\lambda_\infty(t).$$

The next few results are true under the same conditions as for the Lebesgue integral and some with more general conditions which are not true for the Lebesgue integral. The interested reader can consult Henstock [HS] for a proof in the general case.

Theorem 3.11. *Let* $\phi : [a,b] \to \mathbb{R}$ *be differentiable at each* $t \in [a,b]$ *and suppose that* F *is differentiable on* $\phi([a,b])$*. If* $f(u) = F'(u)$ *for (n.e.)* $u \in \phi([a,b])$*, then* f *has a HK-integral and:*

$$HK\int_a^b (f \circ \phi)(t)\phi'(t)d\lambda_\infty(t)$$
$$= (F \circ \phi)(t)\big|_a^b = F(t)\big|_{\phi(a)}^{\phi(b)} = HK\int_{\phi(a)}^{\phi(b)} f(u)d\lambda_\infty(u).$$

Proof. The proof is easy. Use the chain rule to see that

$$(F \circ \phi)'(t) = (f \circ \phi)(t)\phi'(t),$$

for all $t \in [a,b]$. If we apply Theorem 3.6 to both sides of the equation, the result follows. □

Let $D([a,b])$ be the set of all HK-integrable functions on $[a,b]$. The following theorem provides another property of the HK-integral that is not true for Riemann or the Lebesgue integral.

Theorem 3.12 (Hake's Theorem). *The function* $f \in D([a,b])$ *if and only if* $f \in \mathcal{D}([a,s])$ *for every* $s \in (a,b)$ *and* $\lim_{s\to b^-} \int_a^s f(t)d\lambda_\infty(t)$ *exists. In this case,* $\lim_{s\to b^-} \int_a^s f(t)d\lambda_\infty(t) = \int_a^b f(t)d\lambda_\infty(t)$.

Theorem 3.13. *Let* $(f_n)_{n=1}^{\infty}$ *be a sequence of functions in* $D([a,b])$. *The following holds.*

(1) *If* $f_n \to f$ *uniformly* λ_∞-(a.e.), *then* $f \in D([a,b])$ *and*

$$\int_a^b f(t)d\lambda_\infty(t) = \lim_{n\to\infty} \int_a^b f_n(t)d\lambda_\infty(t).$$

(2) (Monotone Convergence) *If* $f_n(t) \leq f_{n+1}(t)$ λ_∞-(a.e.) *on* $[a,b]$ *and* $f(t) = \lim_{n\to\infty} f_n(t)$. *Then* $f \in D([a,b])$ *if and only if*

$$\sup_{n\to\infty} \int_a^b f_n(t)d\lambda_\infty(t) < \infty.$$

In this case,

$$\int_a^b f(t)d\lambda_\infty(t) = \lim_{n\to\infty} \int_a^b f_n(t)d\lambda_\infty(t).$$

(3) (Fatou's Lemma) *If* $f_n(t) \geq 0$ *for all* n *and* $\liminf_{n\to\infty} f_n(t) < \infty$ λ_∞-(a.e.). *Then*

$$\int_a^b \liminf_{n\to\infty} f_n(t)d\lambda_\infty(t) \leqslant \liminf_{n\to\infty} \int_a^b f_n(t)d\lambda_\infty(t).$$

(4) (Dominated Convergence) *If* $\lim\limits_{n\to\infty} f_n(t) = f(t)$, λ_∞-(a.e.) *and there exists functions,* $g, h \in D([a,b])$ *such that*

$$g(t) \leqslant f_n(t) \leqslant h(t), \lambda_\infty - \text{(a.e.)},$$

then $f \in D([a,b])$ *and*

$$\int_a^b f(t)d\lambda_\infty(t) = \lim_{n\to\infty} \int_a^b f_n(t)d\lambda_\infty(t).$$

The next theorem provides additional information about the relationship of the HK-integral relative to the Lebesgue. We will prove (1) for $\mathbb{R}^n$ later in the section.

Theorem 3.14. *Let $f(t) : [a,b] \to \mathbb{R}$.*

(1) *If $f(t)$ is Lebesgue integrable on $[a,b]$, then it is HK-integrable on $[a,b]$ and HK-$\int_a^b f(t)d\lambda_\infty(t) =$L-$\int_a^b f(t)d\lambda_\infty(t)$.*

(2) *If $f(t)$ is HK-integrable and bounded on $[a,b]$, then it is Lebesgue integrable on $[a,b]$.*

(3) *If $f(t)$ is HK-integrable and nonnegative on $[a,b]$, then it is Lebesgue integrable on $[a,b]$.*

(4) *If $f(t)$ is HK-integrable on every measurable subset of $[a,b]$, then it is Lebesgue integrable on $[a,b]$.*

For later use, we note that one can define a norm on the class $D(\mathbb{R})$ of HK-integrable functions. Following Alexiewicz [AL], for $f \in D(\mathbb{R})$, we define $\|f\|_D$ by

$$\|f\|_D = \sup_s \left| \int_{-\infty}^{s} f(r)d\lambda_\infty(r) \right|. \tag{3.1}$$

General Theory. In this section, we discuss the various classical integrals, which integrate nonabsolutely integrable functions, in order to see how the HK-integral is related to other approaches. There are two ways to define an integral. One can provide a descriptive definition or an operational (or constructive) definition. A descriptive definition describes the integral in relationship to its derivative without providing any process for its construction. Complete proofs can be found in Gordon [GOR], Tuo-Yeong [TY], [TY1], and Saks [SK]. The general case can be found in Henstock [HS], or Pfeffer ([PF] and [PF1]).

The oscillation $\omega(F,[a,b])$ of a function F on an interval $[a,b]$ is defined by:

$$\omega\left(F,\ [a,b]\right) = \sup\left\{|F(x) - F(y)| :\ a \leqslant y < x \leqslant b\right\}.$$

Definition 3.15. We define the weak variation, $V(F,E)$, and the strong variation, $V_*(F,E)$, by:

$$V(F,E) = \sup\left\{\sum_{i=1}^{n} |F(b_i) - F(a_i)|\right\},$$

$$V_*(F,E) = \sup\left\{\sum_{i=1}^{n} \omega\left(F,\ [a_i,b_i]\right)\right\},$$

where the supremum is taken over all possible finite collections of nonoverlapping intervals that have end points in E.

(1) We say that F is of bounded variation on E, (BV), if $V(F,E) < \infty$.

(2) We say that F is of restricted bounded variation on E, (BV_*), if $V_*(F,E) < \infty$.

(3) We say that F is absolutely continuous on E, ($\mathbb{AC}$), if for each $\varepsilon > 0$, there exists a $\delta > 0$ such that, for every collection $\{[a_i, b_i], 1 \leqslant i \leqslant n\}$, of nonoverlapping intervals with end points in E and $\sum_{i=1}^{n}(b_i - a_i) < \delta$, then

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| < \varepsilon.$$

(4) We say that F is absolutely continuous on E in the restricted sense, $(\mathbb{AC})_*$, if for each $\varepsilon > 0$, there exists a $\delta > 0$ such that, for every collection $\{[a_i, b_i], 1 \leqslant i \leqslant n\}$, of nonoverlapping intervals with end points in E and $\sum_{i=1}^{n}(b_i - a_i) < \delta$, then

$$\sum_{i=1}^{n} \omega\left(F, [a_i, b_i]\right) < \varepsilon.$$

(5) We say that F is generalized absolutely continuous on E, ($\mathbb{ACG}$), if $F|_E$ is continuous and E is a countable union of sets $\{E_i\}$ such that F is ($\mathbb{AC}$) on each E_i.

(6) We say that F is generalized absolutely continuous in the restricted sense on E, $(\mathbb{ACG})_*$, if $F|_E$ is continuous and E is a countable union of sets $\{E_i\}$ such that F is $(\mathbb{AC})_*$ on each E_i.

We recall that the set of functions of bounded variation on $[a,b]$, $BV([a,b])$ is a Banach space with norm $\|h\|_{BV} = \|h\|_\infty + V(h,[a,b])$.

Let $\mathbb{C}^k(E)$ denote the set of functions on E with k continuous derivatives (we let $\mathbb{C}^0(E) = \mathbb{C}(E)$, the continuous functions on E).

Theorem 3.16. *If E is a subset of $[a,b]$, we have:*

$$\mathbb{C}^1(E) \subset \mathbb{AC}(E) \subset \mathbb{ACG}_*(E) \subset \mathbb{ACG}(E) \subset \mathbb{C}(E).$$

3.1.3. Descriptive Definitions. Let E be a measurable subset of $\mathbb{R}$ and let $\lambda_\infty(E)$ denote the Lebesgue measure of E.

Definition 3.17. Let E be a measurable set and let $c \in \mathbb{R}$.

(1) We say that c is a point of density for E if
$$d_c E = \lim_{h \to 0^+} \frac{\lambda_\infty \left(E \cap (c-h, c+h)\right)}{2h} = 1.$$

(2) We say that c is a point of dispersion for E if
$$d_c E = \lim_{h \to 0^+} \frac{\lambda_\infty \left(E \cap (c-h, c+h)\right)}{2h} = 0.$$

(3) We say that a function $F : [a, b] \to \mathbb{R}$ is approximately continuous at $c \in E \subset [a, b]$, if c is a point of density for E and $F|_E$ is continuous at c.

(4) We say that a function $F : [a, b] \to \mathbb{R}$ is approximately differentiable at $c \in E \subset [a, b]$, if c is a point of density for E and $F|_E$ is differentiable at c. In this case, we write the derivative as $F'_{ap}(c)$.

In the next theorem, we tie down the left end point for convenience. (This theorem provides a descriptive definition of the most well known of the possible integrals.) (For others, see Gordon [GOR] or Saks [SK].)

Theorem 3.18. *Let F be a function defined on $[a, b]$ with $F(a) = 0$, then the following holds.*

(1) *If F is $(\mathbb{AC})$ on $[a, b]$, then F' exists (a.e) and, if F' is Lebesgue integrable, then $\int_a^x F'(y)d\lambda_\infty(y) = F(x)$.*

(2) *If F is $(\mathbb{ACG}_*)$ on $[a, b]$, then F' exists (a.e) and $\int_a^x F'(y)d\lambda_\infty(y) = F(x)$ (as an HK, Perron, or restricted Denjoy integral).*

(3) *If F is $(\mathbb{ACG})$ on $[a, b]$, then F'_{ap} exists (a.e) and $\int_a^x F'_{ap}(y)d\lambda_\infty(y) = F(x)$ (as a wide sense Denjoy or Denjoy-Khintchine integral).*

Note the slight but important difference between (1) and (2). This gives the qualitative distinction between the Lebesgue and the HK (Perron or restricted Denjoy) integral. These latter integrals all differ in their construction, with the HK being the simplest and the restricted Denjoy the most complicated. This explains why the HK-integral has become so popular in recent times.

3.1.3.1. *n-Dimensional HK-Integral.* There are a number of ways to approach the process of recovering a function from its derivative in $\mathbb{R}^n$, which defines the HK-integral on Euclidean spaces. The approach of Lee Tuo-Yeong is perfect for our purpose (see [TY] and [TY1]).

All norms are equivalent on $\mathbb{R}^n$; however, for the HK-integral the maximal norm $\|\mathbf{x}\| = \max_{1 \leqslant k \leqslant n} |x_k|$ is natural. With this norm, the closed ball $B(\mathbf{x}, r)$, is a cube centered at $\mathbf{x}$ with sides parallel to the coordinate axis of length $2r$. (It is a closed interval when $n = 1$.) If the closed interval for side i about x_i is $[a_i, b_i]$, we represent $B(\mathbf{x}, r)$ as $B(\mathbf{x}, r) = (\mathbf{J}, \mathbf{x})$, where $\mathbf{J} = \prod_{i=1}^{n} [a_i, b_i]$. We will call $\mathbf{J}$ a closed interval in $\mathbb{R}^n$.

Definition 3.19. If E is a compact ball in $\mathbb{R}^n$, a partition $\mathcal{P}$ of E is a collection $\{(\mathbf{J}_i, \mathbf{x}_i) : \mathbf{x}_i \in \mathbf{J}_i,\ 1 \leqslant i \leqslant m\}$, where $\mathbf{J}_1, \mathbf{J}_2 \ldots \mathbf{J}_m$ are nonoverlapping closed intervals (i.e., $\lambda_\infty [\mathbf{J}_i \cap \mathbf{J}_j] = 0, i \neq j$) and $\bigcup_{i=1}^{m} \mathbf{J}_i = E$.

Definition 3.20. If δ is a positive function on E, we say that $\mathcal{P}$ is a HK-δ partition for E if for each i, $\mathbf{J}_i \subset B'(\mathbf{x}_i, \delta(\mathbf{x}_i))$. The function δ is called a gauge on E.

The next lemma shows that one can always find a HK-δ partition for any compact set $E \subset \mathbb{R}^n$. The proof is based on the Heine–Borel Theorem from elementary analysis.

Lemma 3.21. Cousin's Lemma *If $\delta(\cdot)$ is a positive function on E, then a HK-δ partition exists for E.*

Definition 3.22. A function $f : E \to \mathbb{R}$ is said to be HK-integrable on E, if there exists a number I such that for any $\varepsilon > 0$ there is a gauge δ and HK-δ partition on E such that

$$\left| \sum_{i=1}^{m} f(\mathbf{x}_i) \lambda_\infty [\mathbf{J}_i] - I \right| < \varepsilon. \tag{3.2}$$

In this case, we write

$$I = (HK) \int_E f(\mathbf{x}) d\lambda_\infty(\mathbf{x}).$$

We now show that the Lebesgue integral is a special case of the HK-integral.

First we need the following, which gives an operational (or constructive) meaning to absolute continuity for functions on $\mathbb{R}^n$.

Lemma 3.23. *Let $f \in L^1[\mathbb{R}^n]$. If $\varepsilon > 0$, then there is a $\delta > 0$ such that, whenever E is a measurable set with $\lambda_\infty[E] < \delta$,*

$$\left| \int_E f(\mathbf{x}) d\lambda_\infty(\mathbf{x}) \right| < \varepsilon.$$

Proof. Let ε be given. Since $f = f_+ - f_-$, it suffices to show that the result is true for $f > 0$. In this case, there is a simple function s such that $0 \leq s \leq f$ and

$$\int_E s(\mathbf{x})d\lambda_\infty(\mathbf{x}) > \int_E f(\mathbf{x})d\lambda_\infty(\mathbf{x}) - \frac{\varepsilon}{2}.$$

Since f has a finite integral, there is a constant C with $s \leq C$ for all $\mathbf{x} \in \mathbb{R}^n$. It follows that

$$\int_E f(\mathbf{x})d\lambda_\infty(\mathbf{x}) = \int_E s(\mathbf{x})d\lambda_\infty(\mathbf{x}) + \int_E (f(\mathbf{x}) - s(\mathbf{x}))\, d\lambda_\infty(\mathbf{x})$$
$$\leqslant \int_E C d\lambda_\infty(\mathbf{x}) + \int_E f(\mathbf{x})d\lambda_\infty(\mathbf{x}) - \int_E s(\mathbf{x})d\lambda_\infty(x) < C\lambda_\infty(E) + \frac{\varepsilon}{2}.$$

Thus, if we set $\delta = \frac{\varepsilon}{2C}$, we are done. □

Theorem 3.24. *If E is a measurable subset of $\mathbb{R}^n$ and $f : E \to \mathbb{R}$ has a finite Lebesgue integral on E. Then it is HK-integrable on E and given $\varepsilon > 0$, for each point $\mathbf{x} \in E$ there is an open set $G(\mathbf{x})$ containing $\mathbf{x}$ such that, whenever $\{B_1, B_2, \dots\}$ is a family of nonoverlapping closed sets contained in E such that $\lambda_\infty(E \setminus \cup_{k=1}^\infty B_k) = 0$ and $\mathbf{x}_1, \mathbf{x}_2, \dots$ satisfying*

$$\mathbf{x}_k \in B_k \subset G(\mathbf{x}_k), \text{ then } \left|\sum\nolimits_{i=1}^n f(\mathbf{x}_i)\lambda_\infty(B_i) - I\right| < \varepsilon.$$

It follows that

$$(HK)\int_E f(\mathbf{x})d\lambda_\infty(\mathbf{x}) = (L)\int_E f(\mathbf{x})d\lambda_\infty(x).$$

Proof. Since the Lebesgue integral is absolutely continuous, there is a $\delta > 0$ such that for all measurable sets $A \subset E$, $\lambda_\infty(A) < \delta$ implies that $\int_A |f(\mathbf{x})|\, d\lambda_\infty(\mathbf{x}) < \frac{\varepsilon}{3}$. Let $\varepsilon' = \frac{\varepsilon}{3(\delta + \lambda_\infty(E))}$ and, for $k = 0, \pm 1, \pm 2, \dots$, let

$$E_k = \left\{\mathbf{x} \,\middle|\, (k-1)\varepsilon' < f(\mathbf{x}) \leqslant k\varepsilon' \right\}.$$

For each k, choose an open set $G_k \supset E_k$ such that

$$\lambda_\infty (G_k \setminus E_k) < \tfrac{1}{3}\frac{\delta}{2^{|k|}(|k|+1)},$$

and, with each $\mathbf{x} \in E_k$, associate $G(\mathbf{x}) = G_k$. Suppose $\{B_i\}$ is a family of nonoverlapping closed sets satisfying our condition. Let $\mathbf{x}_i \in E_{k(i)}$, so that $B_i \subset G_{k(i)}$ and $B_i \setminus E_{k(i)} \subset G_{k(i)} \setminus E_{k(i)}$. Then

$$\left|\sum_i f(\mathbf{x}_i)\lambda_\infty(B_i) - I\right| = \left|\sum_i \int_{B_i} (f(\mathbf{x}_i) - f(\mathbf{x}))\, d\lambda_\infty(\mathbf{x})\right|$$
$$\leqslant \sum_i \int_{B_i} |(f(\mathbf{x}_i) - f(\mathbf{x}))|\, d\lambda_\infty(\mathbf{x})$$
$$\leqslant \sum_i \int_{B_i \cap E_{k(i)}} |(f(\mathbf{x}_i) - f(\mathbf{x}))|\, d\lambda_\infty(\mathbf{x}) + \sum_i \int_{B_i \cap E^c_{k(i)}} |f(\mathbf{x}_i)|\, d\lambda_\infty(\mathbf{x})$$
$$+ \sum_i \int_{B_i \cap E^c_{k(i)}} |f(\mathbf{x})|\, d\lambda_\infty(\mathbf{x}) = P + Q + R.$$

Thus, it suffices to show that each of the above terms is less than $\frac{\varepsilon}{3}$. For $\mathbf{x} \in B_i \cap E_{k(i)}$, both $f(\mathbf{x})$, & $f(\mathbf{x}_i)$ lie in the open interval $([k(i) - 1]\varepsilon', k(i)\varepsilon')$, so that

$$P = \sum_i \int_{B_i \cap E_{k(i)}} |(f(\mathbf{x}_i) - f(\mathbf{x}))|\, d\lambda_\infty(\mathbf{x})$$
$$\leqslant \sum_i \int_{B_i \cap E_{k(i)}} \varepsilon' d\lambda_\infty(\mathbf{x}) = \varepsilon'\lambda_\infty(E) < \frac{\varepsilon}{3}.$$

For Q, put those terms together that have a fixed value k, so that we may write

$$Q = \sum_{k=-\infty}^{\infty} \sum_{k(i)=k} \int_{B_i \cap E^c_{k(i)}} |f(\mathbf{x}_i)|\, d\lambda_\infty(\mathbf{x})$$
$$\leqslant \sum_{k=-\infty}^{\infty} \sum_{k(i)=k} (|k| + 1)\varepsilon'\lambda_\infty(B_i \setminus E_k)$$
$$\leqslant \sum_{k=-\infty}^{\infty} (|k| + 1)\varepsilon'\lambda_\infty(G_k \setminus E_k) < \frac{\varepsilon}{3}.$$

We can also write R as

$$R = \sum_i \int\limits_{B_i \cap E_{k(i)}^c} |f(\mathbf{x})| \, d\lambda_\infty(\mathbf{x})$$

$$= \sum_{k=-\infty}^{\infty} \sum_{k(i)=k} \int\limits_{B_i \cap E_{k(i)}^c} |f(\mathbf{x})| \, d\lambda_\infty(\mathbf{x})$$

$$= \int\limits_A |f(\mathbf{x})| \, d\lambda_\infty(\mathbf{x}),$$

where $A = \bigcup_{k=-\infty}^{\infty} \left[\bigcup_{k(i)=k} (B_i \setminus E_{k(i)}) \right]$. Since, by assumption, the sets B_i are nonoverlapping,

$$\lambda_\infty(A) = \sum_{k=-\infty}^{\infty} \sum_{k(i)=k} \lambda_\infty(B_i \setminus E_{k(i)}) \leqslant \sum_{k=-\infty}^{\infty} \lambda_\infty(B_i \setminus E_k) < \tfrac{1}{3} \sum_{k=-\infty}^{\infty} \frac{\delta}{2^{|k|}} = \delta.$$

By absolute continuity of the integral and the definition of δ, $R < \frac{\varepsilon}{3}$. □

3.2. The HK-Type Banach Spaces

The most important factor preventing the widespread use of the HK-integral has been the lack of a natural Banach space structure for this class of functions (as is the case for the Lebesgue integral). The purpose of this section is to introduce a class of Banach spaces that have the correct properties. We focus on the first class (KS^p Spaces) because they are directly related to the Feynman operator calculus and path integral. The other classes are briefly introduced because of their potential for applications in other parts of analysis.

3.2.1. The Canonical KS^p Spaces. Recall that the HK-integral is equivalent to the (restricted) Denjoy integral. If we replace $\mathbb{R}$ by $\mathbb{R}^n$ in Eq. (3.1), for $f \in D(\mathbb{R}^n)$, we have:

$$\|f\|_D = \sup_{r>0} \left| \int_{\mathbf{B}_r} f(\mathbf{x}) d\lambda_\infty(\mathbf{x}) \right| < \infty, \tag{3.3}$$

where $\mathbf{B}_r$ is any closed cube of diagonal r centered at the origin in $\mathbb{R}^n$ with sides parallel to the coordinate axes. (Note, this defines a norm for both the restricted and wide sense Denjoy integrable functions.)

Now, fix n, and let $\mathbb{Q}^n$ be the set $\{\mathbf{x} = (x_1, x_2 \cdots, x_n) \in \mathbb{R}^n\}$ such that x_i is rational for each i. Since this is a countable dense set in $\mathbb{R}^n$, we can arrange it as $\mathbb{Q}^n = \{\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \cdots\}$. For each l and i, let $\mathbf{B}_l(\mathbf{x}^i)$ be the closed cube centered at $\mathbf{x}^i$, with sides parallel to the coordinate axes and edge $e_l = \frac{1}{2^l\sqrt{n}}, l \in \mathbb{N}$. Now choose the natural order which maps $\mathbb{N} \times \mathbb{N}$ bijectively to $\mathbb{N}$:

$$\{(1,1),(2,1),(1,2),(1,3),(2,2),(3,1),(3,2),(2,3),\ldots\}.$$

Let $\{\mathbf{B}_k,\ k \in \mathbb{N}\}$ be the resulting set of (all) closed cubes $\{\mathbf{B}_l(\mathbf{x}^i) \mid (l,i) \in \mathbb{N} \times \mathbb{N}\}$ centered at a point in $\mathbb{Q}^n$ and let $\mathcal{E}_k(\mathbf{x})$ be the characteristic function of $\mathbf{B}_k$, so that $\mathcal{E}_k(\mathbf{x})$ is in $L^p[\mathbb{R}^n] \cap L^\infty[\mathbb{R}^n]$ for $1 \le p < \infty$. Define $F_k(\,\cdot\,)$ on $L^1[\mathbb{R}^n]$ by

$$F_k(f) = \int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x}). \tag{3.4}$$

It is clear that $F_k(\,\cdot\,)$ is a bounded linear functional on $L^p[\mathbb{R}^n]$ for each k, $\|F_k\| \le 1$ and, if $F_k(f) = 0$ for all k, $f = 0$ so that $\{F_k\}$ is fundamental on $L^p[\mathbb{R}^n]$ for $1 \le p \le \infty$. Set $t_k = 2^{-k}$, so that $\sum_{k=1}^\infty t_k = 1$ and define a measure $d\mu$ on $\mathbb{R}^n \times \mathbb{R}^n$ by:

$$d\mu = \left[\sum_{k=1}^\infty t_k \mathcal{E}_k(\mathbf{x})\mathcal{E}_k(\mathbf{y})\right] d\lambda_\infty(\mathbf{x}) d\lambda_\infty(\mathbf{y}).$$

We first construct our Hilbert space. Define an inner product $(\,\cdot\,)$ on $L^1[\mathbb{R}^n]$ by

$$\begin{aligned}(f,g) &= \int_{\mathbb{R}^n \times \mathbb{R}^n} f(\mathbf{x}) g(\mathbf{y})^c d\mu \\ &= \sum_{k=1}^\infty t_k \left[\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right] \left[\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{y}) g(\mathbf{y}) d\lambda_\infty(\mathbf{y})\right]^c.\end{aligned} \tag{3.5}$$

We call the completion of $L^1[\mathbb{R}^n]$, with the above inner product, the Kuelbs–Steadman space, $KS^2[\mathbb{R}^n]$. Steadman [ST] constructed this space by adapting an approach developed by Kuelbs [KB] for other purposes. Her interest was in showing that $L^1[\mathbb{R}^n]$ can be densely and continuously embedded in a Hilbert space which contains the HK-integrable functions. To see that this is the case, let $f \in D(\mathbb{R}^n)$, then:

$$\begin{aligned}\|f\|_{KS^2}^2 &= \sum_{k=1}^\infty t_k \left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|^2 \\ &\leqslant \sup_k \left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|^2 \leqslant \|f\|_D^2,\end{aligned}$$

so $f \in KS^2[\mathbb{R}^n]$.

Theorem 3.25. *The space $KS^2[\mathbb{R}^n]$ contains $L^p[\mathbb{R}^n]$ (for each p, $1 \leqslant p \leqslant \infty$) as dense subspaces.*

Proof. By construction, we know that $KS^2[\mathbb{R}^n]$ contains $L^1[\mathbb{R}^n]$ densely. Thus, we need to only show that $L^q[\mathbb{R}^n] \subset KS^2[\mathbb{R}^n]$ for $q \neq 1$. Let $f \in L^q[\mathbb{R}^n]$ and $q < \infty$. Since $|\mathcal{E}(\mathbf{x})| = \mathcal{E}(\mathbf{x}) \leqslant 1$ and $|\mathcal{E}(\mathbf{x})|^q \leqslant \mathcal{E}(\mathbf{x})$, we have

$$\begin{aligned}
\|f\|_{KS^2} &= \left[\sum_{k=1}^{\infty} t_k \left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|^{\frac{2q}{q}}\right]^{1/2} \\
&\leqslant \left[\sum_{k=1}^{\infty} t_k \left(\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) \left|f(\mathbf{x})\right|^q d\lambda_\infty(\mathbf{x})\right)^{\frac{2}{q}}\right]^{1/2} \\
&\leqslant \sup_k \left(\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) \left|f(\mathbf{x})\right|^q d\lambda_\infty(\mathbf{x})\right)^{\frac{1}{q}} \leqslant \|f\|_q .
\end{aligned}$$

Hence, $f \in KS^2[\mathbb{R}^n]$. For $q = \infty$, first note that $vol(\mathbf{B}_k)^2 \leq \left[\frac{1}{2\sqrt{n}}\right]^{2n}$, so we have

$$\begin{aligned}
\|f\|_{KS^2} &= \left[\sum_{k=1}^{\infty} t_k \left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|^{2}\right]^{1/2} \\
&\leqslant \left[\left[\sum_{k=1}^{\infty} t_k [vol(\mathbf{B}_k)]^2\right] [ess\sup|f|]^2\right]^{1/2} \leqslant \left[\frac{1}{2\sqrt{n}}\right]^n \|f\|_\infty .
\end{aligned}$$

Thus $f \in KS^2[\mathbb{R}^n]$, and $L^\infty[\mathbb{R}^n] \subset KS^2[\mathbb{R}^n]$. □

The fact that $L^\infty[\mathbb{R}^n] \subset KS^2[\mathbb{R}^n]$, while $KS^2[\mathbb{R}^n]$ is separable makes it clear in a very forceful manner that separability is not an inherited property. Before proceeding, we construct $KS^p[\mathbb{R}^n]$.

To construct $KS^p[\mathbb{R}^n]$ for all p, let $f \in L^p[\mathbb{R}^n]$ and define:

$$\|f\|_{KS^p} = \begin{cases} \left\{\sum\limits_{k=1}^{\infty} t_k \left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|^p\right\}^{1/p}, 1 \leqslant p < \infty, \\ \sup_{k \geqslant 1} \left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|, p = \infty. \end{cases}$$

It is easy to see that $\|\cdot\|_{KS^p}$ defines a norm on $L^p[\mathbb{R}^n]$. If $KS^p[\mathbb{R}^n]$ is the completion of $L^p[\mathbb{R}^n]$ with respect to this norm, we have:

Theorem 3.26. *For each q, $1 \leqslant q \leqslant \infty$, $KS^p[\mathbb{R}^n] \supset L^q[\mathbb{R}^n]$ as dense continuous embeddings.*

Proof. As in the previous theorem, by construction $KS^p[\mathbb{R}^n]$ contains $L^p[\mathbb{R}^n]$ densely, so we need to only show that $KS^p[\mathbb{R}^n] \supset L^q[\mathbb{R}^n]$ for $q \neq p$. First, suppose that $p < \infty$. If $f \in L^q[\mathbb{R}^n]$ and $q < \infty$, we have

$$\begin{aligned}
\|f\|_{KS^p} &= \left[\sum_{k=1}^{\infty} t_k \left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|^{\frac{qp}{q}}\right]^{1/p} \\
&\leqslant \left[\sum_{k=1}^{\infty} t_k \left(\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) \left|f(\mathbf{x})\right|^q d\lambda_\infty(\mathbf{x})\right)^{\frac{p}{q}}\right]^{1/p} \\
&\leqslant \sup_k \left(\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) \left|f(\mathbf{x})\right|^q d\lambda_\infty(\mathbf{x})\right)^{\frac{1}{q}} \leqslant \|f\|_q .
\end{aligned}$$

Hence, $f \in KS^p[\mathbb{R}^n]$. For $q = \infty$, we have

$$\begin{aligned}
\|f\|_{KS^p} &= \left[\sum_{k=1}^{\infty} t_k \left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|^{p}\right]^{1/p} \\
&\leqslant \left[\left[\sum_{k=1}^{\infty} t_k [vol(\mathbf{B}_k)]^p\right] [ess\sup |f|]^p\right]^{1/p} \leqslant M \|f\|_\infty .
\end{aligned}$$

Thus $f \in KS^p[\mathbb{R}^n]$, and $L^\infty[\mathbb{R}^n] \subset KS^p[\mathbb{R}^n]$. The case $p = \infty$ is obvious. □

Theorem 3.27. *For $KS^p[\mathbb{R}^n]$, $1 \le p \le \infty$, we have:*

(1) *If $f_n \to f$ weakly in $L^p[\mathbb{R}^n]$, then $f_n \to f$ strongly in $KS^p[\mathbb{R}^n]$ (i.e., every weakly compact subset of $L^p[\mathbb{R}^n]$ is compact in $KS^p[\mathbb{R}^n]$).*

(2) *If $1 < p < \infty$, then $KS^p[\mathbb{R}^n]$ is uniformly convex.*

(3) *If $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$, then the dual space of $KS^p[\mathbb{R}^n]$ is $KS^q[\mathbb{R}^n]$.*

(4) *$KS^\infty[\mathbb{R}^n] \subset KS^p[\mathbb{R}^n]$, for $1 \le p < \infty$.*

Proof. The proof of (1) follows from the fact that if $\{f_n\}$ is any weakly convergent sequence in $L^p[\mathbb{R}^n]$ with limit f, then

$$\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) \left[f_n(\mathbf{x}) - f(\mathbf{x})\right] d\lambda_\infty(\mathbf{x}) \to 0$$

for each k. It follows that $\{f_n\}$ converges strongly to f in $KS^p[\mathbb{R}^n]$.

The proof of (2) follows from a modification of the proof of the Clarkson inequalities for l^p norms (see [CL]).

In order to prove (3), observe that, for $p \neq 2,\ 1 < p < \infty$, the linear functional L_g, defined by

$$\begin{aligned} &L_g(f) \\ &= \|g\|_{KS^p}^{2-p} \sum_{k=1}^{\infty} t_k \left| \int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) g(\mathbf{x}) d\lambda_\infty(\mathbf{x}) \right|^{p-2} \\ &\cdot \left[\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) g(\mathbf{x}) d\lambda_\infty(\mathbf{x}) \right]^c \int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{y}) f(\mathbf{y}) d\lambda_\infty(\mathbf{y}),\ f \in KS^q, \end{aligned}$$

is a duality map on $KS^q[\mathbb{R}^n]$ for each nonzero $g \in KS^p[\mathbb{R}^n]$. We then use the fact that $KS^p[\mathbb{R}^n]$ is reflexive from (2). To prove (4), note that $f \in KS^\infty[\mathbb{R}^n]$ implies that $\left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|$ is uniformly bounded for all k. It follows that $\left|\int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x})\right|^p$ is uniformly bounded for each p, $1 \le p < \infty$. It is now clear from the definition of $KS^\infty[\mathbb{R}^n]$ that:

$$\left[\sum_{k=1}^{\infty} t_k \left| \int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) f(\mathbf{x}) d\lambda_\infty(\mathbf{x}) \right|^p \right]^{1/p} \leqslant M \|f\|_{KS^\infty} < \infty.$$

□

Remark 3.28. There is flexibility in the choice of the family of positive numbers $\{t_k\}$, $\sum_{k=1}^{\infty} t_k = 1$. This is somewhat akin to the metric used for $\mathbb{R}^\infty$. Recall that for any two points $x,\ y \in \mathbb{R}^\infty$, $d(x,y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|x-y|}{1+|x-y|}$. The family of numbers $\{\frac{1}{2^n}\}$ can be replaced by any other sequence of positive numbers whose sum is one, without affecting the topology.

There is also some ambiguity associated with the choice $\mathbb{Q}^n$ and the order for $\mathbb{N} \times \mathbb{N}$. (We have used simplicity to choose the order for $\mathbb{N} \times \mathbb{N}$.) The important fact is that, for any combination of orders, the properties of $KS^2[\mathbb{R}^n]$ are invariant.

We could replace the family of generating functions $\{\mathcal{E}_k,\ k \in \mathbb{N}\}$ by the Hermite functions on $\mathbb{R}^n$. However, since the functions in $KS^p[\mathbb{R}^n]$ depend on the smoothness properties of the family $\{\mathcal{E}_k,\ k \in \mathbb{N}\}$, we see that the definition of $KS^p[\mathbb{R}^n]$ depends on our choice of the $\mathcal{E}_k$. We favor the present family because the $\mathcal{E}_k$ have compact support and the weakest continuity properties.

We used the ℓ_p sequence space relationship to our norm to prove (1) and (3). Recall that $\ell_1 \subset \ell_p \subset \ell_q \subset \ell_\infty$, for $1 \le p < q \le \infty$ (see

Jones [J], p. 241). However, from (5), we see that the KS^p spaces cannot be viewed as special cases of the ℓ_p spaces.

We note that, since $L^1[\mathbb{R}^n] \subset KS^p[\mathbb{R}^n]$ and $KS^p[\mathbb{R}^n]$ is reflexive for $1 < p < \infty$, the second dual $\left\{L^1[\mathbb{R}^n]\right\}^{**} = \mathfrak{M}[\mathbb{R}^n] \subset KS^p[\mathbb{R}^n]$. Recall that $\mathfrak{M}[\mathbb{R}^n]$ is the space of bounded finitely additive set functions defined on the Borel sets $\mathfrak{B}[\mathbb{R}^n]$. This space contains the Dirac delta measure and free-particle Green's function for the Feynman integral.

3.2.2. The Hilbert Space $GS^2[\mathbb{R}^n]$. We now turn to the construction of a second Hilbert space $GS^2[\mathbb{R}^n]$, which is required for the extension of the Feynman operator calculus to non-reflexive Banach spaces. This space is also the natural compliment to $KS^2[\mathbb{R}^n]$ in a certain general sense.

For every separable Banach space $\mathcal{B}$ that is dense in $KS^2[\mathbb{R}^n]$, we want to show that $GS^2[\mathbb{R}^n] \subset \mathcal{B} \subset KS^2[\mathbb{R}^n]$ as continuous dense embeddings. We begin with a nice (and very useful) result due to Lax [L]. Let the Banach space $\mathcal{B}$ be a dense continuous embedding in a separable Hilbert space $\mathcal{H}$, so that there is an $M > 0$ such that $\|x\|_{\mathcal{H}} \leqslant M\|x\|_{\mathcal{B}}$, for all $x \in \mathcal{B}$. In what follows, we assume that $M = 1$.

Theorem 3.29 (Lax)**.** *Let $A \in L[\mathcal{B}]$. If A is self-adjoint on $\mathcal{H}$ (i.e., $(Ax, y)_{\mathcal{H}} = (x, Ay)_{\mathcal{H}}, \forall x, y \in \mathcal{B}$), then A is bounded on $\mathcal{H}$ and $\|A\|_{\mathcal{H}} \leqslant k\|A\|_{\mathcal{B}}$ for some positive constant k.*

Proof. Let $x \in \mathcal{B}$ and, without loss, we can assume that $k = 1$ and $\|x\|_{\mathcal{H}} = 1$. Since A is self-adjoint,

$$\|Ax\|_{\mathcal{H}}^2 = (Ax, Ax) = \left(x, A^2x\right) \leqslant \|x\|_{\mathcal{H}} \left\|A^2x\right\|_{\mathcal{H}} = \left\|A^2x\right\|_{\mathcal{H}}.$$

Thus, we have $\|Ax\|_{\mathcal{H}}^4 \leqslant \left\|A^4x\right\|_{\mathcal{H}}$, so it is easy to see that $\|Ax\|_{\mathcal{H}}^{2n} \leqslant \left\|A^{2n}x\right\|_{\mathcal{H}}$ for all n. It follows that:

$$\begin{aligned}\|Ax\|_{\mathcal{H}} &\leqslant (\left\|A^{2n}x\right\|_{\mathcal{H}})^{1/2n} \leqslant (\left\|A^{2n}x\right\|_{\mathcal{B}})^{1/2n}\\ &\leqslant (\left\|A^{2n}\right\|_{\mathcal{B}})^{1/2n}(\|x\|_{\mathcal{B}})^{1/2n} \leqslant \|A\|_{\mathcal{B}}(\|x\|_{\mathcal{B}})^{1/2n}.\end{aligned}$$

Letting $n \to \infty$, we get that $\|Ax\|_{\mathcal{H}} \leqslant \|A\|_{\mathcal{B}}$ for x in a dense set of the unit ball of $\mathcal{H}$. It follows that

$$\|A\|_{\mathcal{H}} = \sup_{\|x\|_{\mathcal{H}} \leqslant 1} \|Ax\|_{\mathcal{H}} \leqslant \|A\|_{\mathcal{B}}.$$

□

For our second Hilbert space, fix $\mathcal{B}$ and define $GS^2_{\mathcal{B}}[\mathbb{R}^n]$ by:

$$GS^2_{\mathcal{B}}[\mathbb{R}^n] = \left\{ u \in \mathcal{B} \,\middle|\, \sum_{n=1}^{\infty} t_n^{-1} \left|(u, \mathcal{E}_n)_2\right|^2 < \infty \right\}, \quad \text{with}$$

$$(u, v)_1 = \sum\nolimits_{n=1}^{\infty} t_n^{-1} (u, \mathcal{E}_n)_2 (\mathcal{E}_n, v)_2 .$$

For convenience, let $\mathcal{H}_1 = GS^2_{\mathcal{B}}[\mathbb{R}^n]$ and $KS^2[\mathbb{R}^n] = \mathcal{H}_2$. for $u \in \mathcal{B}$, let $T_{12}u$ be defined by $T_{12}u = \sum_{n=1}^{\infty} t_n (u, \mathcal{E}_n)_2 \mathcal{E}_n$.

Theorem 3.30. *The operator* $\mathbf{T}_{12}$ *is a positive trace class operator on* $\mathcal{B}$ *with a bounded extension to* $\mathcal{H}_2$. *In addition,* $\mathcal{H}_1 \subset \mathcal{B} \subset \mathcal{H}_2$ *(as continuous dense embeddings),* $\left(T_{12}^{1/2}u,\ T_{12}^{1/2}v\right)_1 = (u,\ v)_2$ *and* $\left(T_{12}^{-1/2}u,\ T_{12}^{-1/2}v\right)_2 = (u,\ v)_1$.

Proof. First, since terms of the form $\{u_N = \sum_{k=1}^{N} t_n^{-1} (u, \mathcal{E}_k)_2 \mathcal{E}_k : u \in \mathcal{B}\}$ are dense in $\mathcal{B}$, we see that $\mathcal{H}_1$ is dense in $\mathcal{B}$. It follows that $\mathcal{H}_1$ is also dense in $\mathcal{H}_2$.

For the operator T_{12}, we see that $\mathcal{B} \subset \mathcal{H}_2 \Rightarrow (u, \mathcal{E}_n)_2$ is defined for all $u \in \mathcal{B}$, so that $\mathbf{T}_{12}$ maps $\mathcal{B} \to \mathcal{B}$ and:

$$\|T_{12}u\|_{\mathcal{B}}^2 \le \left[\sum\nolimits_{n=1}^{\infty} t_n^2 \|\mathcal{E}_n\|_{\mathcal{B}}^2\right] \left[\sum\nolimits_{n=1}^{\infty} \left|(u, \mathcal{E}_n)_2\right|^2\right] = M \|u\|_2^2 \le M \|u\|_{\mathcal{B}}^2 .$$

Thus, T_{12} is a bounded operator on $\mathcal{B}$. It is clearly trace class and, since $(T_{12}u,\ u)_2 = \sum_{n=1}^{\infty} t_n \left|(u, \mathcal{E}_n)_2\right|^2 > 0$, it is positive. From here, it's easy to see that T_{12} is self-adjoint on $\mathcal{H}_2$; so, by Theorem 3.24, it has a bounded extension to $\mathcal{H}_2$.

An easy calculation now shows that $\left(T_{12}^{1/2}u,\ T_{12}^{1/2}v\right)_1 = (u,\ v)_2$ and $\left(T_{12}^{-1/2}u,\ T_{12}^{-1/2}v\right)_2 = (u,\ v)_1$.

Thus, we see that, given $\mathcal{B}$ dense in $KS^2[\mathbb{R}^n]$, we can find a Hilbert space $GS^2_{\mathcal{B}}[\mathbb{R}^n]$, with the property that:

$$GS^2_{\mathcal{B}}[\mathbb{R}^n] \hookrightarrow \mathcal{B} \hookrightarrow KS^2[\mathbb{R}^n] \ \text{(as continuous dense embeddings)} \tag{3.6}$$

□

Remark 3.31. We call $GS^2_{\mathcal{B}}[\mathbb{R}^n]$ the Gross–Steadman space for $\mathcal{B}$. Historically, Gross [GR] first proved that every real separable Banach space contains a separable Hilbert space as a dense embedding, and that this space is the support of a Gaussian measure.

3.2.3. Uniqueness. Our construction does not produce a unique rigging for a given Banach space. To see this, let Ω be a bounded open subset of $\mathbb{R}^n$, and let $\mathbb{C}_0(\Omega)$ be the set of continuous functions on Ω, which vanish on the boundary. Define $H_0[\Omega]$ and $H_0^1[\Omega]$ by:

$$H_0[\Omega] = \left\{u \in L^2[\Omega] : \; u = 0 \text{ on } \partial\Omega\right\}$$
$$H_0^1[\Omega] = \{u | \; \nabla u \in H_0[\Omega], \; u = 0 \text{ on } \partial\Omega\}.$$

The following example assumes some background in partial differential equations, but shows that the Banach space $\mathbb{C}_0(\Omega)$ has (at least) two pair of Hilbert spaces satisfying $\mathcal{H}_1 \subset \mathbb{C}_0(\Omega) \subset \mathcal{H}_2$, as dense continuous embeddings. It follows that our construction of a rigging is not unique.

Example 3.32. *The first pair is* $H_0^1[\Omega] \subset \mathbb{C}_0(\Omega) \subset H_0[\Omega]$. *In this case, the norms for our respective Hilbert spaces are generated by the following inner products:*

$$\langle u, \mathbf{J}_2 v\rangle = (u, v)_2 = \int_\Omega u(x)v(x)d\lambda_\infty$$
$$\langle u, \mathbf{J}_1 v\rangle = (u, v)_1 = \int_\Omega \nabla u(x) \cdot \nabla v(x)d\lambda_\infty.$$

We can take $\mathbf{J}_2 = \mathbf{I}_2$ *as the dual operator (for our dual bracket). However, from*

$$\langle u, \mathbf{J}_1 v\rangle = \int_\Omega \nabla u(x) \cdot \nabla v(x)d\lambda_\infty,$$

we must take $\mathbf{J}_1 = [-\Delta]$, *with Dirichlet boundary conditions (see Barbu [B], p. 4). It is not hard to show that the natural operator relating the spaces must be* $\mathbf{T}_{12} = [-\Delta]^{-1}$, *which is clearly positive and self-adjoint.*

With additional effort, one can even show that $\mathbf{T}_{12}$ *is a trace class operator on* $H_0^1[\Omega]$, *with a bounded extension to* $H_0[\Omega]$. *Furthermore, as expected, we have* $(u, v)_1 = (\mathbf{T}_{12}^{-1} u, v)_2$ *and* $(u, v)_2 = (\mathbf{T}_{12} u, v)_1$.

For the second case, we follow a variation on the process used to construct KS^2. *Let* $\mathbb{Q}_\Omega^n = \mathbb{Q}^n \cap \Omega$ *and, for each* k, *let* $e_k(x) = \mathcal{E}_k(x)|_\Omega$. *If we let*

$$F_n(u) = \int_\Omega e_n(x)u(x)d\lambda_\infty,$$

then the set of functionals $\{F_n,\ n \in \mathbb{N}\}$ *is fundamental on* $\mathbb{C}_0(\Omega)$ *(and also* $L_0^1[\Omega]$*). Using* $t_n = \frac{1}{2^n}$*, define an inner product on* $\mathbb{C}_0(\Omega)$ *by:*

$$\begin{aligned}(u, v)_2 &= \sum_{n=1}^{\infty} t_n F_n(u) \bar{F}_n(v) \\ &= \sum\nolimits_{n=1}^{\infty} t_n \int_{\Omega} \int_{\Omega} e_n(x) \bar{e}_n(y) u(x) \bar{v}(y) d\lambda_{\infty}(x) d\lambda_{\infty}(y).\end{aligned}$$

Letting $\mathcal{H}_2$ *be the completion of* $\mathbb{C}_0(\Omega)$ *in the norm generated by the inner product, we obtain our Hilbert space.*

Since

$$\|u\|_2 = \left[\sum_{n=1}^{\infty} t_n \left| \int_{\Omega} e_n(x) u(x) d\lambda_{\infty} \right|^2 \right]^{1/2},$$

just as with KS^2*, it is easy to show that* $\mathcal{H}_2$ *contains all of the* $L^p[\Omega]$ *spaces,* $1 \le p \le \infty$ *as continuous dense and compact embeddings.*

Define operator $\mathbf{T}_{12}$ *on* $\mathbb{C}_0[\Omega]$ *by:*

$$\mathbf{T}_{12} u = \sum_{n=1}^{\infty} t_n \left(u, e_n\right)_2 e_n.$$

Since $\mathbb{C}_0[\Omega] \subset \mathcal{H}_2$*,* $(u, e_n)_2$ *is defined for all* $u \in \mathbb{C}_0[\Omega]$*. Thus,* $\mathbf{T}_{12}$ *maps* $\mathbb{C}_0[\Omega] \to \mathbb{C}_0[\Omega]$ *and:*

$$\|\mathbf{T}_{12} u\|_0^2 \le \left[\sum_{n=1}^{\infty} t_n^2 \right] \left[\sum\nolimits_{n=1}^{\infty} \left|(u, e_n)_2\right|^2 \right] = M \left\|u\right\|_2^2 \le M \left\|u\right\|_0^2.$$

Thus, $\mathbf{T}_{12}$ *is a bounded operator on* $\mathbb{C}_0(\Omega)$*. Define* $\mathcal{H}_1$ *by:*

$$\mathcal{H}_1 = \left\{ u \in \mathbb{C}_0(\Omega) \;\middle|\; \sum_{n=1}^{\infty} \tfrac{1}{t_n} \left|(u, e_n)_2\right|^2 < \infty \right\},$$

$$\left(u, v\right)_1 = \sum_{n=1}^{\infty} \tfrac{1}{t_n} \left(u, e_n\right)_2 \left(e_n, v\right)_2.$$

With the above inner product, $\mathcal{H}_1$ *is a Hilbert space and, since terms of the form* $\{u_N = \sum_{k=1}^{N} \frac{1}{t_n} \left(u, e_k\right)_2 e_k : u \in \mathbb{C}_0[\Omega]\}$ *are dense in* $\mathbb{C}_0(\Omega)$*, we see that* $\mathcal{H}_1$ *is dense in* $\mathbb{C}_0(\Omega)$*. It follows that* $\mathcal{H}_1$ *is also dense in* $\mathcal{H}_2$*. It is easy to see that* $\mathbf{T}_{12}$ *is a positive self-adjoint trace class operator with respect to the* $\mathcal{H}_2$ *inner product so, using Lax's Theorem,* $\mathbf{T}_{12}$ *has a bounded extension to* $\mathcal{H}_2$ *and* $\|\mathbf{T}_{12}\|_2 \le \|\mathbf{T}_{12}\|_0$*. Finally, for* $u, v \in \mathcal{H}_1$*,* $(u, v)_1 = (\mathbf{T}_{12}^{-1} u, v)_2$ *and* $(u, v)_2 = (\mathbf{T}_{12} u, v)_1$*. It follows that* $\mathcal{H}_1$ *is continuously embedded in* $\mathcal{H}_2$*, hence also in* $\mathbb{C}_0(\Omega)$*.*

3.3. Spaces of Sobolev Type

In many applications, it is convenient to formulate problems on one of the standard Sobolev spaces $W^{m,p}[\mathbb{R}^n]$. In this section our main interest is in the Jones family of spaces, SD^p, $1 \leq p \leq \infty$. These spaces contain the Kuelbs–Steadman spaces $KS^p[\mathbb{R}^n]$ as well as the Sobolev spaces $W^{m,p}[\mathbb{R}^n]$. In addition, they contain the space $\mathcal{D}(\mathbb{R}^n)$, the test functions, all as a dense continuous embeddings. (In order to closely parallel the conventional theory, we replace λ_∞ with λ_n where appropriate.)

3.3.1. The Jones Family of Spaces SD^p, $1 \leq p \leq \infty$. The theory of distributions is based on the action of linear functionals on a space of test functions. In the original approach of Schwartz [SC], both the test functions and the linear functionals have a natural topological vector space structure, which is not normable. For those interested in applications, this is an inconvenience, requiring additional study and effort. Thus, in most applied contexts, the restricted class of Banach spaces due to Sobolev has proved useful (see Leoni [GL]). In the last section, we extended the Sobolev spaces to the nonabsolutely integrable case. The purpose of this section is to introduce another class of Banach spaces which contain the nonabsolutely integrable functions, but also contains the Schwartz test function space as a dense and continuous embedding. A related approach to the work in this section which also leads to a Banach space structure is due to Talvila (see [TA1] and [TA2]). We believe that these spaces will prove very important in the future, so we repeat some construction details of the KS^p-spaces.

3.3.2. The Jones Functions. In this section we develop the Jones class of spaces, which also contains each of the spaces $W^{k,p}[\mathbb{R}^n]$ $1 \leq p \leq \infty$.

We begin with the construction of a special class of functions in $\mathbb{C}_c^\infty(\mathbb{R}^n)$, but first we need the remarkable Jone's functions.

Definition 3.33. For $x \in \mathbb{R}$, $0 \leq y < \infty$ and $1 < a < \infty$, define the Jone's functions $g(x,y)$, $h(x)$ by:

$$g(x,y) = \exp\left\{-y^a e^{iax}\right\},$$

$$h(x) = \begin{cases} \displaystyle\int_0^\infty g(x,y)dy, & x \in (-\frac{\pi}{2a}, \frac{\pi}{2a}) \\ 0, & \text{otherwise.} \end{cases}$$

The following properties of g are easy to check:

(1)
$$\frac{\partial g(x,y)}{\partial x} = -iay^a e^{iax} g(x,y),$$

(2)
$$\frac{\partial g(x,y)}{\partial y} = -ay^{a-1} e^{iax} g(x,y),$$

so that
$$iy\frac{\partial g(x,y)}{\partial y} = \frac{\partial g(x,y)}{\partial x}.$$

It is also easy to see that $h(x) \in L^1[-\frac{\pi}{2a}, \frac{\pi}{2a}]$ and,
$$\frac{dh(x)}{dx} = \int_0^\infty \frac{\partial g(x,y)}{\partial x} dy = \int_0^\infty iy \frac{\partial g(x,y)}{\partial y} dy. \tag{3.7}$$

Integration by parts in the last expression above shows that $h'(x) = -ih(x)$, so that $h(x) = h(0)e^{-ix}$ for $x \in (-\frac{\pi}{2a}, \frac{\pi}{2a})$. Since $h(0) = \int_0^\infty \exp\{-y^a\} dy$, an additional integration by parts shows that $h(0) = \Gamma(\frac{1}{a}+1)$. For each $k \in \mathbb{N}$ let $a = a_k = \pi 2^{k-1}$, $h(x) = h_k(x)$, $x \in (-\frac{1}{2^k}, \frac{1}{2^k})$ and set $\varepsilon_k = \frac{1}{2^{k+1}}$.

Let $\mathbb{Q}$ be the set of rational numbers in $\mathbb{R}$ and for each $x^i \in \mathbb{Q}$, define

$$f_k^i(x) = f_k(x - x^i) = \begin{cases} c_k \exp\left\{\frac{\varepsilon_k^2}{|x-x^i|^2 - \varepsilon_k^2}\right\}, & \left|x - x^i\right| < \varepsilon_k, \\ 0, & \left|x - x^i\right| \geqslant \varepsilon_k, \end{cases}$$

where c_k is the standard normalizing constant. It is clear that the support, $\mathrm{spt}(f_k^i) \subset [-\varepsilon_k, \varepsilon_k] = [-\frac{1}{2^{k+1}}, \frac{1}{2^{k+1}}] = I_k^i$.

Now set $\chi_k^i(x) = (f_k^i * h_k)(x)$, so that $\mathrm{spt}(\chi_k^i) \subset [-\frac{1}{2^{k+1}}, \frac{1}{2^{k+1}}]$. For $x \in \mathrm{spt}(\chi_k^i)$, we can also write $\chi_k^i(x) = \chi_k(x - x^i)$ as:

$$\begin{aligned} \chi_k^i(x) &= \int_{I_k^i} f_k\left[\left(x - x^i\right) - z\right] h_k(z) d\lambda_n(z) \\ &= \int_{I_k^i} h_k\left[\left(x - x^i\right) - z\right] f_k(z) d\lambda_n(z) \\ &= e^{-i\left(x - x^i\right)} \int_{I_k^i} e^{iz} f_k(z) d\lambda_n(z). \end{aligned}$$

Thus, if $\alpha_{k,i} = \int_{I_k^i} e^{iz} f_k^i(z) d\lambda_n(z)$, we can now define:

$$\xi_k^i(x) = \alpha_{k,i}^{-1} \chi_k^i(-x) = \begin{cases} \frac{1}{n} e^{i(x-x^i)}, & x \in I_k^i \\ 0, & x \notin I_k^i, \end{cases}$$

so that $\left|\xi_k^i(x)\right| < \frac{1}{n}$.

3.3.3. The Construction. To construct our space on $\mathbb{R}^n$, let $\mathbb{Q}^n$ be the set of all vectors $\mathbf{x}$ in $\mathbb{R}^n$, such that for each j, the component x_j is rational. Since this is a countable dense set in $\mathbb{R}^n$, we can arrange it as $\mathbb{Q}^n = \left\{\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \cdots\right\}$. For each k and i, let $\mathbf{B}_k(\mathbf{x}^i)$ be the closed cube centered at $\mathbf{x}^i$ with edge $\frac{1}{2^k\sqrt{n}}$.

We choose the natural order which maps $\mathbb{N} \times \mathbb{N}$ bijectively to $\mathbb{N}$:

$$\{(1,1),\ (2,1),\ (1,2),\ (1,3),\ (2,2),\ (3,1),\ (3,2),\ (2,3),\ \cdots\}$$

and let $\{\mathbf{B}_m,\ m \in \mathbb{N}\}$ be the set of closed cubes $\mathbf{B}_k(\mathbf{x}^i)$ with $(k,i) \in \mathbb{N} \times \mathbb{N}$ and $\mathbf{x}^{\mathbf{i}} \in \mathbb{Q}^n$. For each $\mathbf{x} \in \mathbf{B}_m,\ \mathbf{x} = (x_1, x_2, \ldots, x_n)$, we define $\mathcal{E}_m(\mathbf{x})$ by :

$$\mathcal{E}_m(\mathbf{x}) = \left(\xi_k^i(x_1), \xi_k^i(x_2) \ldots \xi_k^i(x_n)\right) \ \textit{with}$$
$$|\mathcal{E}_m(\mathbf{x})| < 1,\ \ \mathbf{x} \in \prod\nolimits_{j=1}^n I_k^i \text{ and } \mathcal{E}_m(\mathbf{x}) = 0,\ \ \mathbf{x} \notin \prod\nolimits_{j=1}^n I_k^i.$$

It is easy to see that $\mathcal{E}_m(\mathbf{x})$ is in $L^p[\mathbb{R}^n]^n = \mathbf{L}^p[\mathbb{R}^n]$ for $1 \le p \le \infty$. Define $F_m(\,\cdot\,)$ on $\mathbf{L}^p[\mathbb{R}^n]$ by

$$F_m(f) = \int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\lambda_n(\mathbf{x}).$$

It is clear that $F_m(\,\cdot\,)$ is a bounded linear functional on $\mathbf{L}^p[\mathbb{R}^n]$ for each m with $\|F_m\| \le 1$. Furthermore, if $F_m(f) = 0$ for all m, $f = 0$ so that $\{F_m\}$ is a fundamental sequence of functionals on $\mathbf{L}^p[\mathbb{R}^n]$ for $1 \le p \le \infty$.

Set $t_m = \frac{1}{2^m}$ so that $\sum_{m=1}^{\infty} t_m = 1$ and define an inner product $(\,\cdot\,)$ on $\mathbf{L}^1[\mathbb{R}^n]$ by

$$(f,g) = \sum\nolimits_{m=1}^{\infty} t_m \left[\int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\lambda_n(\mathbf{x})\right] \left[\int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{y}) \cdot g(\mathbf{y}) d\lambda_n(\mathbf{y})\right]^c.$$

The completion of $\mathbf{L}^1[\mathbb{R}^n]$ with the above inner product is a Hilbert space, which we denote as $SD^2[\mathbb{R}^n]$.

Theorem 3.34. *For each p, $1 \leqslant p \leqslant \infty$, we have:*

(1) *The space $SD^2[\mathbb{R}^n] \supset \mathbf{L}^p[\mathbb{R}^n]$ as a continuous, dense, and compact embedding.*

(2) *The space $SD^2[\mathbb{R}^n] \supset \mathfrak{M}[\mathbb{R}^n]$, the space of finitely additive measures on $\mathbb{R}^n$, as a continuous dense and compact embedding.*

Proof. Since $SD^2[\mathbb{R}^n]$ contains $\mathbf{L}^1[\mathbb{R}^n]$ densely, to prove (1), we need to only show that $\mathbf{L}^q[\mathbb{R}^n] \subset SD^2[\mathbb{R}^n]$ for $q \neq 1$. Let $f \in \mathbf{L}^q[\mathbb{R}^n]$ and $q < \infty$. By construction, for every m, $|\mathcal{E}_m(\mathbf{x})| < \frac{1}{\sqrt{n}}$ so that there is a constant $C = C(q)$, with $|\mathcal{E}_m(\mathbf{x})|^q \leqslant C|\mathcal{E}_m(\mathbf{x})|$. It follows that:

$$\begin{aligned}\|f\|_{SD^2} &= \left[\sum\nolimits_{m=1}^{\infty} t_m \left|\int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) f(\mathbf{x}) d\lambda_n(\mathbf{x})\right|^{\frac{2q}{q}}\right]^{1/2} \\ &\leqslant C\left[\sum\nolimits_{m=1}^{\infty} t_m \left(\int_{\mathbb{R}^n} |\mathcal{E}_m(\mathbf{x})|\, |f(\mathbf{x})|^q \, d\lambda_n(\mathbf{x})\right)^{\frac{2}{q}}\right]^{1/2} \\ &\leqslant C \sup_m \left(\int_{\mathbb{R}^n} |\mathcal{E}_m(\mathbf{x})|\, |f(\mathbf{x})|^q \, d\lambda_n(\mathbf{x})\right)^{\frac{1}{q}} \leqslant C \|f\|_q .\end{aligned}$$

Hence, $f \in SD^2[\mathbb{R}^n]$. For $q = \infty$, first note that $vol(\mathbf{B}_m)^2 \leq \left[\frac{1}{2\sqrt{n}}\right]^{2n}$, so we have

$$\begin{aligned}\|f\|_{SD^2} &= \left[\sum\nolimits_{m=1}^{\infty} t_m \left|\int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) f(\mathbf{x}) d\lambda_n(\mathbf{x})\right|^2\right]^{1/2} \\ &\leqslant \left[\left[\sum\nolimits_{m=1}^{\infty} t_m [vol(\mathbf{B}_m)]^2\right] [ess \sup |f|]^2\right]^{1/2} \leqslant \left[\frac{1}{2\sqrt{n}}\right]^n \|f\|_\infty .\end{aligned}$$

Thus $f \in SD^2[\mathbb{R}^n]$, and $\mathbf{L}^\infty[\mathbb{R}^n] \subset SD^2[\mathbb{R}^n]$. To prove compactness, suppose $\{f_j\}$ is any weakly convergent sequence in $\mathbf{L}^p[\mathbb{R}^n]$, $1 \leq p \leq \infty$ with limit f. Since $\mathcal{E}_m \in \mathbf{L}^q$, $1/p + 1/q = 1$,

$$\int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot [f_j(\mathbf{x}) - f(\mathbf{x})] \, d\lambda_n(\mathbf{x}) \to 0$$

for each m. It follows that $\{f_j\}$ converges strongly to f in $SD^2[\mathbb{R}^n]$.

To prove (2), we note that $\mathfrak{M}[\mathbb{R}^n] = \mathbf{L}^1[\mathbb{R}^n]^{**} \subset SD^2[\mathbb{R}^n]$. □

Definition 3.35. We call $SD^2[\mathbb{R}^n]$ the Jones strong distribution Hilbert space on $\mathbb{R}^n$.

In order to justify our definition, let α be a multi-index of nonnegative integers, $\alpha = (\alpha_1,\ \alpha_2,\ \cdots\ \alpha_k)$, with $|\alpha| = \sum_{j=1}^{k} \alpha_j$. If D denotes the standard partial differential operator, let

$$D^{\alpha} = D^{\alpha_1} D^{\alpha_2} \cdots D^{\alpha_k}.$$

Theorem 3.36. *Let $\mathcal{D}[\mathbb{R}^n]$ be $\mathbb{C}_c^{\infty}(\mathbb{R}^n)$ equipped with the standard locally convex topology (test functions).*

(1) *If $\phi_j \to \phi$ in $\mathcal{D}(\mathbb{R}^n)$, then $\phi_j \to \phi$ in the norm topology of $SD^2[\mathbb{R}^n]$, so that $\mathcal{D}(\mathbb{R}^n) \subset SD^2[\mathbb{R}^n]$ as a continuous dense embedding.*

(2) *If $T \in \mathcal{D}'(\mathbb{R}^n)$, then $T \in SD^2[\mathbb{R}^n]'$, so that $\mathcal{D}'(\mathbb{R}^n) \subset SD^2[\mathbb{R}^n]'$ as a continuous dense embedding.*

(3) *For any $f,\ g \in SD^2[\mathbb{R}^n]$ and any multi-index α, $(D^{\alpha} f, g)_{SD} = (-i)^{\alpha} (f, g)_{SD}$.*

Proof. To prove (1), suppose that $\phi_j \to \phi$ in $\mathcal{D}(\mathbb{R}^n)$. By definition, there exists a compact set $K \subset \mathbb{R}^n$, which is the support of $\phi_j - \phi$ and $D^{\alpha}\phi_j$ converges to $D^{\alpha}\phi$ uniformly on K for every multi-index α. Let $\{\mathcal{E}_{K_l}\}$ be the set of all $\mathcal{E}_l$, with support $K_l \subset K$. If α is a multi-index, we have:

$$\begin{aligned}
&\lim_{j\to\infty} \|D^{\alpha}\phi_j - D^{\alpha}\phi\|_{SD} \\
&= \lim_{j\to\infty} \left\{ \sum_{l=1}^{\infty} t_{K_l} \left| \int_{\mathbb{R}^n} \mathcal{E}_{K_l}(\mathbf{x}) \cdot [D^{\alpha}\phi_j(\mathbf{x}) - D^{\alpha}\phi(\mathbf{x})]\, d\lambda_n(\mathbf{x}) \right|^2 \right\}^{1/2} \\
&\leqslant M \limsup_{j\to\infty, \mathbf{x}\in K} |D^{\alpha}\phi_j(\mathbf{x}) - D^{\alpha}\phi(\mathbf{x})| = 0.
\end{aligned}$$

Thus, since α is arbitrary, we see that $\mathcal{D}(\mathbb{R}^n) \subset SD^2[\mathbb{R}^n]$ as a continuous embedding. Since $\mathbb{C}_c^{\infty}[\mathbb{R}^n]$ is dense in $\mathbf{L}^1[\mathbb{R}^n]$, $\mathcal{D}(\mathbb{R}^n)$ is dense in $SD^2[\mathbb{R}^n]$. To prove (2) we note that, as $\mathcal{D}(\mathbb{R}^n)$ is a dense locally convex subspace of $SD^2[\mathbb{R}^n]$, by a corollary of the Hahn-Banach Theorem every continuous linear functional, T defined on $\mathcal{D}(\mathbb{R}^n)$, can be extended to a continuous linear functional on $SD^2(\mathbb{R}^n)$. By the Riesz representation theorem, every continuous linear functional T defined on $SD^2[\mathbb{R}^n]$ is of the form $T(f) = (f,g)_{SD}$, for some $g \in SD^2[\mathbb{R}^n]$. Thus, $T \in SD^2[\mathbb{R}^n]'$ and, by the identification $T \leftrightarrow g$ for each T in $\mathcal{D}'(\mathbb{R}^n)$, we can map $\mathcal{D}'(\mathbb{R}^n)$ into $SD^2[\mathbb{R}^n]$ as a continuous dense embedding.

To prove (3), recall that each $\mathcal{E}_m \in \mathbb{C}_c^\infty(\mathbb{R}^n)$ so that, for any $f \in SD^2[\mathbb{R}^n]$,

$$\int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot D^\alpha f(\mathbf{x}) d\lambda_n(\mathbf{x}) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} D^\alpha \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\lambda_n(\mathbf{x}).$$

An easy calculation shows that:

$$(-1)^{|\alpha|} \int_{\mathbb{R}^n} D^\alpha \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\lambda_n(\mathbf{x}) = (-i)^{|\alpha|} \int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\lambda_n(\mathbf{x}).$$

It now follows that, for any $\mathbf{g} \in SD^2[\mathbb{R}^n]$, $(D^\alpha f, \mathbf{g})_{SD^2} = (-i)^{|\alpha|} (f, \mathbf{g})_{SD^2}$. □

Remark 3.37. We note that it is easy to see that $W^{k,p}[\mathbb{R}^n] \subset SD^2[\mathbb{R}^n]$ as a continuous dense embedding, for all k and all p.

3.3.3.1. *Functions of Bounded Variation.* The objective of this section is to show that every HK-integrable function is in $SD^2[\mathbb{R}^n]$. To do this, we need to discuss a certain class of functions of bounded variation. For functions defined on $\mathbb{R}$, the definition of bounded variation is unique. However, for functions on $\mathbb{R}^n$, $n \geq 2$, there are a number of distinct definitions.

The functions of bounded variation in the sense of Cesari are well known to analysts working in partial differential equations and geometric measure theory (see Leoni [GL]).

Definition 3.38. A function $f \in \mathbf{L}^1[\mathbb{R}^n]$ is said to be of bounded variation in the sense of Cesari or $f \in BV_c[\mathbb{R}^n]$, if $f \in \mathbf{L}^1[\mathbb{R}^n]$ and each i, $1 \leq i \leq n$, there exists a signed Radon measure μ_i, such that

$$\int_{\mathbb{R}^n} f(\mathbf{x}) \frac{\partial \phi(\mathbf{x})}{\partial x_i} d\lambda_n(\mathbf{x}) = -\int_{\mathbb{R}^n} \phi(\mathbf{x}) d\mu_i(\mathbf{x}),$$

for all $\phi \in \mathbb{C}_0^\infty(\mathbb{R}^n)$.

The functions of bounded variation in the sense of Vitali [TY1] are well known to applied mathematicians and engineers with interest in error estimates associated with research in control theory, financial derivatives, high speed networks, robotics, and in the calculation of certain integrals. (See, for example [KAA], [NI], [PT], or [PTR] and references therein.) For the general definition, see Yeong ([TY1], p. 175). We present a definition that is sufficient for continuously differentiable functions.

Definition 3.39. A function f with continuous partials is said to be of bounded variation in the sense of Vitali or $f \in BV_v[\mathbb{R}^n]$ if for all intervals (a_i, b_i), $1 \le i \le n$,

$$V(f) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \left| \frac{\partial^n f(\mathbf{x})}{\partial x_1 \partial x_2 \cdots \partial x_n} \right| d\lambda_n(\mathbf{x}) < \infty.$$

Definition 3.40. We define $BV_{v,0}[\mathbb{R}^n]$ by:

$$BV_{v,0}[\mathbb{R}^n] = \{f(\mathbf{x}) \in BV_v[\mathbb{R}^n] : f(\mathbf{x}) \to 0, \text{ as } x_i \to -\infty\},$$

where x_i is any component of $\mathbf{x}$.

The following two theorems may be found in [TY1]. (See p. 184 and 187, where the first is used to prove the second.) If $[a_i, b_i] \subset \mathbb{R}$, we define $[\mathbf{a}, \mathbf{b}] \in \mathbb{R}^n$ by $[\mathbf{a}, \mathbf{b}] = \prod_{k=1}^n [a_i, b_i]$. (The notation (RS) means Riemann–Stieltjes.)

Theorem 3.41. *Let f be HK-integrable on $[\mathbf{a}, \mathbf{b}]$ and let $g \in BV_{v,0}[\mathbb{R}^n]$, then fg is HK-integrable and*

$$(HK)\int_{[\mathbf{a},\mathbf{b}]} f(\mathbf{x})g(\mathbf{x})d\lambda_n(\mathbf{x}) = (RS)\int_{[\mathbf{a},\mathbf{b}]} \left\{(HK)\int_{[\mathbf{a},\mathbf{x}]} f(\mathbf{y})d\lambda_n(\mathbf{y})\right\} dg(\mathbf{x})..$$

Theorem 3.42. *Let f be HK-integrable on $[\mathbf{a}, \mathbf{b}]$ and let $g \in BV_{v,0}[\mathbb{R}^n]$, then fg is HK-integrable and*

$$\left|(HK)\int_{[\mathbf{a},\mathbf{b}]} f(\mathbf{x})g(\mathbf{x})d\lambda_n(\mathbf{x})\right| \le \|f\|_D \, V_{[\mathbf{a},\mathbf{b}]}(g).$$

Lemma 3.43. *The space $HK[\mathbb{R}^n]$, of all HK-integrable functions is contained in $SD^2[\mathbb{R}^n]$.*

Proof. Since each $\mathcal{E}_m(\mathbf{x})$ is continuous and differentiable, $\mathcal{E}_m(\mathbf{x}) \in BV_{v,0}[\mathbb{R}^n]$, so that for $f \in HK[\mathbb{R}^n]$,

$$\|f\|^2_{\mathbf{SD}^2} = \sum\nolimits_{m=1}^{\infty} t_m \left| \int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\mathbf{x} \right|^2 \leqslant \sup_m \left| \int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\mathbf{x} \right|^2$$
$$\leqslant \|f\|^2_{HK} \left[\sup_m V(\mathcal{E}_m)\right]^2 < \infty.$$

It follows that $f \in SD^2[\mathbb{R}^n]$. □

3.3.4. The General Case, $SD^p[\mathbb{R}^n]$, $1 \le p \le \infty$. To construct $SD^p[\mathbb{R}^n]$ for all p and for $f \in \mathbf{L}^p[\mathbb{R}^n]$, define:

$$\|u\|_{SD^p} = \begin{cases} \left\{ \sum\limits_{|\alpha| \leqslant m} \sum\limits_{k=1}^{\infty} t_k \left| \int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) D^\alpha u(\mathbf{x}) d\lambda_n(\mathbf{x}) \right|^p \right\}^{1/p}, & 1 \leqslant p < \infty \\ \sum\limits_{|\alpha| \leqslant m} \sup\limits_{k \geqslant 1} \left| \int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{x}) D^\alpha u(\mathbf{x}) d\lambda_n(\mathbf{x}) \right|, & p = \infty. \end{cases}$$

It is easy to see that $\|\cdot\|_{SD^p}$ defines a norm on $\mathbf{L}^p[\mathbb{R}^n]$. If $SD^p[\mathbb{R}^n]$ is the completion of $\mathbf{L}^p[\mathbb{R}^n]$ with respect to this norm, we have:

Theorem 3.44. *For each q, $1 \leqslant q \leqslant \infty$, $SD^p[\mathbb{R}^n] \supset \mathbf{L}^q[\mathbb{R}^n]$ as dense continuous embeddings.*

Proof. As in the previous theorem, by construction $SD^p[\mathbb{R}^n]$ contains $\mathbf{L}^p[\mathbb{R}^n]$ densely, so we need to only show that $SD^p[\mathbb{R}^n] \supset \mathbf{L}^q[\mathbb{R}^n]$ for $q \neq p$. First, suppose that $p < \infty$. If $f \in \mathbf{L}^q[\mathbb{R}^n]$ and $q < \infty$, we have

$$\|f\|_{SD^p} = \left[\sum\nolimits_{m=1}^{\infty} t_m \left| \int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\lambda_n(\mathbf{x}) \right|^{\frac{qp}{q}} \right]^{1/p}$$
$$\leqslant \left[\sum\nolimits_{m=1}^{\infty} t_m \left(\int_{\mathbb{R}^n} |\mathcal{E}_m(\mathbf{x})|^q \, |f(\mathbf{x})|^q \, d\lambda_n(\mathbf{x}) \right)^{\frac{p}{q}} \right]^{1/p}$$
$$\leqslant \sup_m \left(\int_{\mathbb{R}^n} |\mathcal{E}_m(\mathbf{x})|^q \, |f(\mathbf{x})|^q \, d\lambda_n(\mathbf{x}) \right)^{\frac{1}{q}} \leqslant \|f\|_q \, .$$

Hence, $f \in SD^p[\mathbb{R}^n]$. For $q = \infty$, we have

$$\|f\|_{SD^p} = \left[\sum\nolimits_{m=1}^{\infty} t_m \left| \int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) \cdot f(\mathbf{x}) d\lambda_n(\mathbf{x}) \right|^p \right]^{1/p}$$
$$\leqslant \left[\left[\sum\nolimits_{m=1}^{\infty} t_m [vol(\mathbf{B}_m)]^p \right] [ess \sup |f|]^p \right]^{1/p} \leqslant M \, \|f\|_\infty \, .$$

Thus $f \in SD^p[\mathbb{R}^n]$, and $L^\infty[\mathbb{R}^n] \subset SD^p[\mathbb{R}^n]$. The case $p = \infty$ is obvious. □

Theorem 3.45. *For $SD^p[\mathbb{R}^n]$, $1 \le p \le \infty$, we have:*

(1) *If $p^{-1} + q^{-1} = 1$, then the dual space of $SD^p[\mathbb{R}^n]$ is $SD^q[\mathbb{R}^n]$.*
(2) *The test function space $\mathcal{D}(\mathbb{R}^n)$ is contained in $SD^p[\mathbb{R}^n]$ as a continuous dense embedding.*
(3) *If K is a weakly compact subset of $L^p[\mathbb{R}^n]$, it is a strongly compact subset of $SD^p[\mathbb{R}^n]$.*
(4) *The space $SD^\infty[\mathbb{R}^n] \subset SD^p[\mathbb{R}^n]$.*

Remark 3.46. Reflection reveals that the spaces $KS^p[\mathbb{R}^n]$ and $SD^p[\mathbb{R}^n]$ may be viewed as different poles for the same basic construction. The spaces $KS^p[\mathbb{R}^n]$ use a base composed of characteristic functions of cubes; while spaces $SD^p[\mathbb{R}^n]$ use the same cubes, but as supports for functions in $\mathbb{C}_c^\infty(\mathbb{R}^n)$. In both cases, the largest cube has volume $\left[\frac{1}{2\sqrt{n}}\right]^n$. It is actually this property that makes it possible for $SD^p[\mathbb{R}^n]$ to contain the test function space $\mathcal{D}(\mathbb{R}^n)$ as a continuous dense embedding.

The above observation leads to the following theorem.

Theorem 3.47. *The test function space $\mathcal{D}(\mathbb{R}^n)$ is contained in $KS^p[\mathbb{R}^n]$, $1 \le p \le \infty$ as a continuous dense embedding.*

Proof. Since $KS^\infty[\mathbb{R}^n] \subset KS^p[\mathbb{R}^n]$, as a continuous dense embedding for all p, it suffices to prove the result for $KS^\infty[\mathbb{R}^n]$.

Suppose that $\phi_j \to \phi$ in $\mathcal{D}(\mathbb{R}^n)$. Thus, there exists a compact set $K \subset \mathbb{R}^n$, which is the support of $\phi_j - \phi$ and $D^\alpha \phi_j$ converges to $D^\alpha \phi$ uniformly on K for every multi-index α. Let $\{\mathcal{E}_{K_l}\}$ be the set of all $\mathcal{E}_l$, with support $K_l \subset K$. If α is a multi-index, we have:

$$\begin{aligned}
&\lim_{j\to\infty} \|D^\alpha \phi_j - D^\alpha \phi\|_{KS^\infty} \\
&= \lim_{j\to\infty} \sup_l \left| \int_{\mathbb{R}^n} \mathcal{E}_l(\mathbf{x}) \cdot [D^\alpha \phi_j(\mathbf{x}) - D^\alpha \phi(\mathbf{x})]\, d\lambda_n(\mathbf{x}) \right| \\
&\leqslant \left[\frac{1}{2\sqrt{n}}\right]^n \lim_{j\to\infty} \sup_{\mathbf{x}\in K} |D^\alpha \phi_j(\mathbf{x}) - D^\alpha \phi(\mathbf{x})| = 0.
\end{aligned}$$

Thus, since α is arbitrary, we see that $\mathcal{D}(\mathbb{R}^n) \subset KS^\infty[\mathbb{R}^n]$ as a continuous embedding. □

3.3.5. Application. Let $\{L^2[\mathbb{R}^3]\}^3$ be the Hilbert space of square integrable functions on $\mathbb{R}^3$, let $\mathbb{H}[\mathbb{R}^3]$ be the completion of the set of functions in $\left\{\mathbf{u} \in \mathbb{C}_0^\infty(\mathbb{R}^3)^3 \mid \nabla \cdot \mathbf{u} = 0\right\}$, which vanish at infinity with respect to the inner product of $\{L^2[\mathbb{R}^3]\}^3$. The classical Navier–Stokes initial-value problem (on $\mathbb{R}^3$ and all $T > 0$) is to find a function $\mathbf{u} : [0, T] \times \mathbb{R}^3 \to \mathbb{R}^3$ and $p : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ such that

$$\begin{aligned} &\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \nabla p = \mathbf{f}(t) \text{ in } (0, T) \times \mathbb{R}^3, \\ &\nabla \cdot \mathbf{u} = 0 \text{ in } (0, T) \times \mathbb{R}^3 \text{ (in the weak sense)}, \\ &\mathbf{u}(0, \mathbf{x}) = \mathbf{u}_0(\mathbf{x}) \text{ in } \mathbb{R}^3. \end{aligned} \tag{3.8}$$

The equations describe the time evolution of the fluid velocity $\mathbf{u}(\mathbf{x}, t)$ and the pressure p of an incompressible viscous homogeneous Newtonian fluid with constant viscosity coefficient ν in terms of a given initial velocity $\mathbf{u}_0(\mathbf{x})$ and given external body forces $\mathbf{f}(\mathbf{x}, t)$.

Let $\mathbb{P}$ be the (Leray) orthogonal projection of $\{L^2[\mathbb{R}^3]\}^3$ onto $\mathbb{H}[\mathbb{R}^3]$ and define the Stokes operator by: $\mathbf{A}\mathbf{u} =: -\mathbb{P}\Delta\mathbf{u}$, for $\mathbf{u} \in D(\mathbf{A}) \subset \mathbb{H}^2[\mathbb{R}^3]$, the domain of $\mathbf{A}$. If we apply $\mathbb{P}$ to Eq. (3.8), with $B(\mathbf{u}, \mathbf{u}) = \mathbb{P}(\mathbf{u} \cdot \nabla)\mathbf{u}$, we can recast Eq. (3.8) into the standard form:

$$\begin{aligned} &\partial_t \mathbf{u} = -\nu\mathbf{A}\mathbf{u} - B(\mathbf{u}, \mathbf{u}) + \mathbb{P}\mathbf{f}(t) \text{ in } (0, T) \times \mathbb{R}^3, \\ &\mathbf{u}(0, \mathbf{x}) = \mathbf{u}_0(\mathbf{x}) \text{ in } \mathbb{R}^3, \end{aligned} \tag{3.9}$$

where the orthogonal complement of $\mathbb{H}[\mathbb{R}^3]$ relative to $\{L^2[\mathbb{R}^3]\}^3$, $\{\mathbf{v} : \mathbf{v} = \nabla q,\ q \in \mathbb{H}^1[\mathbb{R}^3]\}$ is used to eliminate the pressure term (see Galdi [GA] or [SY], [T1], [T2]).

Definition 3.48. We say that a velocity vector field in $\mathbb{R}^3$ is reasonable if for $0 \le t < \infty$, there is a continuous function $m(t) > 0$, depending only on t and a constant M_0, which may depend on $\mathbf{u}_0$ and f, such that

$$0 < m(t) \leqslant \|\mathbf{u}(t)\|_{\mathbb{H}} \le M_0.$$

The above definition formalizes the requirement that the fluid has nonzero but bounded positive definite energy. However, this condition still allows the velocity to approach zero at infinity in a weaker norm.

3.3.6. The Nonlinear Term: A Priori Estimates. The difficulty in proving the existence and uniqueness of global-in-time strong solutions for Eq. (3.9) is directly linked to the problem of getting good a priori estimates for the nonlinear term $B(\mathbf{u}, \mathbf{u})$. Let $\mathbb{H}_{sd}$ be the closure of $\mathbb{H} \cap SD^2[\mathbb{R}^3]$ in the SD^2 norm.

Theorem 3.49. *If* $\mathbf{A}$ *is the Stokes operator and* $\mathbf{u}(\mathbf{x},t) \in \mathbb{H}_{sd} \cap D(\mathbf{A})$ *is a reasonable vector field, then*

$$\langle \nu \mathbf{A}\mathbf{u}, \mathbf{u} \rangle_{\mathbb{H}_{sd}} = 3\nu \left\| \mathbf{u} \right\|_{\mathbb{H}_{sd}}^2 . \tag{3.10}$$

For $\mathbf{u}(\mathbf{x},t) \in \mathbb{H}_{sd} \cap D(\mathbf{A})$ *and* $t \in [0,\infty)$, *there exists a constant* $M = M(\mathbf{u}_0, \mathbf{f})) > 0$, *such that*

$$\left| \langle B(\mathbf{u},\mathbf{u}), \mathbf{u} \rangle_{\mathbb{H}_{sd}} \right| \leq M \left\| \mathbf{u} \right\|_{\mathbb{H}_{sd}}^3 . \tag{3.11}$$

We also have that:

$$max\{ \| B(\mathbf{u},\mathbf{v}) \|_{\mathbb{H}_{sd}} ,\ \| B(\mathbf{v},\mathbf{u}) \|_{\mathbb{H}_{sd}} \} \leqslant M \left\| \mathbf{u} \right\|_{\mathbb{H}_{sd}} \left\| \mathbf{v} \right\|_{\mathbb{H}_{sd}} . \tag{3.12}$$

Proof. From the definition of the inner product, for (3.10) we have

$$\langle \nu \mathbf{A}\mathbf{u}, \mathbf{u} \rangle_{\mathbb{H}_{sd}}$$
$$= \nu \sum\nolimits_{m=1}^{\infty} t_m \left[\int_{\mathbb{R}^3} \mathcal{E}_m(\mathbf{x}) \cdot \mathbf{A}\mathbf{u}(\mathbf{x}) d\lambda_3(\mathbf{x}) \right] \left[\int_{\mathbb{R}^3} \mathcal{E}_m(\mathbf{y}) \cdot \mathbf{u}(\mathbf{y}) d\lambda_3(\mathbf{y}) \right]^c .$$

Using the fact that $\mathbf{u} \in D(\mathbf{A})$, it follows that

$$\int_{\mathbb{R}^3} \mathcal{E}_m(\mathbf{y}) \cdot \partial_{y_j}^2 \mathbf{u}(\mathbf{y}) d\lambda_3(\mathbf{y})$$
$$= \int_{\mathbb{R}^3} \partial_{y_j}^2 \mathcal{E}_m(\mathbf{y}) \cdot \mathbf{u}(\mathbf{y}) d\lambda_3(\mathbf{y}) = (i)^2 \int_{\mathbb{R}^3} \mathcal{E}_m(\mathbf{y}) \cdot \mathbf{u}(\mathbf{y}) d\lambda_3(\mathbf{y}).$$

Using this in the above equation and summing on j, we have ($\mathbf{A} = -\mathbb{P}\Delta$)

$$\int_{\mathbb{R}^3} \mathcal{E}_m(\mathbf{y}) \cdot \mathbf{A}\mathbf{u}(\mathbf{y}) d\lambda_3(\mathbf{y}) = 3 \int_{\mathbb{R}^3} \mathcal{E}_m(\mathbf{y}) \cdot \mathbf{u}(\mathbf{y}) d\lambda_3(\mathbf{y}).$$

It follows that

$$\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle_{\mathbb{H}_{sd}}$$
$$= 3 \sum\nolimits_{m=1}^{\infty} t_m \left[\int_{\mathbb{R}^3} \mathcal{E}_m(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) d\lambda_3(\mathbf{x}) \right] \left[\int_{\mathbb{R}^3} \mathcal{E}_m(\mathbf{y}) \cdot \mathbf{u}(\mathbf{y}) d\lambda_3(\mathbf{y}) \right]^c$$
$$= 3 \left\| \mathbf{u} \right\|_{\mathbb{H}_{sd}}^2 .$$

This proves (3.11). To prove (3.12), let

$$b(\mathbf{u},\mathbf{v},\mathcal{E}_m) = \int_{\mathbb{R}^3} (\mathbf{u}(\mathbf{x}) \cdot \nabla \mathbf{v}(\mathbf{x})) \cdot \mathcal{E}_m(\mathbf{x}) d\lambda_3(\mathbf{x})$$

and define the vector $\mathbf{I}$ by $\mathbf{I} = [1,1,1]^t$. We start with integration by parts and $\nabla \cdot \mathbf{u} = 0$, to get

$$b(\mathbf{u},\mathbf{v},\mathcal{E}_m) = -b(\mathbf{u},\mathcal{E}_m,\mathbf{v}) = -i \int_{\mathbb{R}^3} (\mathbf{u}(\mathbf{x}) \cdot \mathbf{I}) (\mathcal{E}_m(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x})) d\lambda_3(\mathbf{x}).$$

From the above equation, we have $(m \leftrightarrow (k,i))$

$$\begin{aligned}|b(\mathbf{u},\mathbf{v},\mathcal{E}_m)| &\leqslant \sqrt{3}\int_{\mathbb{R}^3}|\mathbf{u}(\mathbf{x})||\mathbf{v}(\mathbf{x})|d\lambda_3(\mathbf{x})\sup_k \|\mathcal{E}_m\|_\infty \\ &\leqslant C_1\|\mathbf{u}\|_{\mathbb{H}}\|\mathbf{v}\|_{\mathbb{H}}.\end{aligned}$$

We also have:

$$\left|\int_{\mathbb{R}^3}\mathbf{w}(\mathbf{x})\cdot\mathcal{E}_m(\mathbf{x})d\lambda_3(\mathbf{x})\right| \leqslant C_2\|\mathbf{w}\|_{\mathbb{H}}.$$

If we combine the last two results, we get that:

$$\begin{aligned}&\left|\langle B(\mathbf{u},\mathbf{v}),\mathbf{w}\rangle_{\mathbb{H}_{sd}}\right| \\ &\leqslant \sum_{m=1}^{\infty} t_m |b(\mathbf{u},\mathbf{v},\mathcal{E}_m)| \overline{\left|\int_{\mathbb{R}^3}\mathbf{w}(\mathbf{y})\cdot\mathcal{E}_m(\mathbf{y})d\lambda_3(\mathbf{y})\right|} \\ &\leqslant C\|\mathbf{u}\|_{\mathbb{H}}\|\mathbf{v}\|_{\mathbb{H}}\|\mathbf{w}\|_{\mathbb{H}}.\end{aligned} \tag{3.13}$$

Since $\mathbf{u},\mathbf{v},\mathbf{w}$ are reasonable velocity vector fields, there is a constant M depending on $\mathbf{u}_0,\mathbf{v}_0,\mathbf{w}_0$ and f, such that

$$C\|\mathbf{u}\|_{\mathbb{H}}\|\mathbf{v}\|_{\mathbb{H}}\|\mathbf{w}\|_{\mathbb{H}} \leq M\|\mathbf{u}\|_{\mathbb{H}_{sd}}\|\mathbf{v}\|_{\mathbb{H}_{sd}}\|\mathbf{w}\|_{\mathbb{H}_{sd}}.$$

If $\mathbf{w}=\mathbf{v}=\mathbf{u}$, we have that:

$$\left|\langle B(\mathbf{u},\mathbf{u}),\mathbf{u}\rangle_{\mathbb{H}_{sd}}\right| \leqslant M\|\mathbf{u}\|_{\mathbb{H}_{sd}}^3.$$

This proves (3.11). The proof of (3.12) is a straightforward application of (3.13). □

To compare our results, if $\mathbf{u},\mathbf{v}\in D(\mathbf{A})$, a typical bound in the $\mathbb{H}$ norm for Eq. (3.12) can be found in Sell and You [SY] (see p. 366):

$$\max\{\|B(\mathbf{u},\mathbf{v})\|_{\mathbb{H}},\ \|B(\mathbf{v},\mathbf{u})\|_{\mathbb{H}}\} \leqslant C_0\left\|\mathbf{A}^{5/8}\mathbf{u}\right\|_{\mathbb{H}}\left\|\mathbf{A}^{5/8}\mathbf{v}\right\|_{\mathbb{H}}.$$

3.3.7. Conclusion. In closing, one must have noticed the factor of 3 in Eq. (3.10). Our definition of $\xi_k^i(x)$ can be changed to:

$$\xi_k^i(x) = \alpha_{k,i}^{-1}\chi_k^i(-x) = \begin{cases}\frac{1}{n}e^{\frac{i(x-x^i)}{3}}, & x\in I_k^i\\ 0, & x\notin I_k^i,\end{cases}$$

in order to give us

$$\langle \nu\mathbf{A}\mathbf{u},\mathbf{u}\rangle_{\mathbb{H}_{sd}} = \nu\|\mathbf{u}\|_{\mathbb{H}_{sd}}^2.$$

Thus, the space may be fine-tuned to fit the problem of interest.

3.4. Zachary Spaces

In this section, we briefly discuss two other possible families of spaces that naturally flow from the existence of a Banach space structure for functions with a bounded integral. (We call these spaces Zachary spaces.)

3.4.1. Zachary Functions of Bounded Mean Oscillation $\mathcal{Z}^p[\mathbb{R}^n]$, $1 \leq p \leq \infty$. In this section, we extend the space of functions of bounded mean oscillation.

Definition 3.50. Let $f \in L^1_{\text{loc}}[\mathbb{R}^n]$ and let Q be a cube in $\mathbb{R}^n$.

(1) We define the average of f over Q by

$$\underset{Q}{Avg}\, f = \frac{1}{\lambda_n\left[Q\right]} \int_Q f(\mathbf{y}) d\lambda_n(\mathbf{y}).$$

(2) We defined the sharp maximal function $M^{\#}(f)(\mathbf{x})$, by

$$M^{\#}(f)(\mathbf{x}) = \sup_{Q} \frac{1}{\lambda_n\left[Q\right]} \int_Q \left| f(\mathbf{y}) - \underset{Q}{Avg}\, f \right| d\lambda_n(\mathbf{y}).$$

(3) If $M^{\#}(f)(\mathbf{x}) \in L^\infty[\mathbb{R}^n]$, we say that f is of bounded mean oscillation. More precisely, the space of functions of bounded mean oscillation are defined by:

$$BMO[\mathbb{R}^n] = \left\{ f \in L^1_{loc}[\mathbb{R}^n] : \; M^{\#}(f) \in L^\infty[\mathbb{R}^n] \right\}$$

and

$$\|f\|_{BMO} = \left\| M^{\#}(f) \right\|_{L^\infty}.$$

We may also obtain an equivalent definition of $BMO[\mathbb{R}^n]$ using balls, but for our purposes, cubes are natural (see Grafakos [GRA], p. 546). We note that $BMO[\mathbb{R}^n]$ is not a Banach space and is not separable.

Let $\{\mathcal{E}_k(\mathbf{x})\}$ be the family of generating functions for $KS^2[\mathbb{R}^n]$ and recall that they are the indicator functions for a family of cubes $\{Q_k\}$ centered at each rational point in $\mathbb{R}^n$. Let $f \in L^1_{\text{loc}}[\mathbb{R}^n]$ and define f_{ak} by

$$f_{ak} = \frac{1}{\lambda_n\left[Q_k\right]} \int_{Q_k} f(\mathbf{y}) d\lambda_n(\mathbf{y}) = \frac{1}{\lambda_n\left[Q_k\right]} \int_{\mathbb{R}^n} \mathcal{E}_k(\mathbf{y}) f(\mathbf{y}) d\lambda_n(\mathbf{y}).$$

Definition 3.51. If p, $1 \le p < \infty$ and $t_k = 2^{-k}$, we define $\|f\|_{\mathcal{Z}^p}$ by

$$\|f\|_{\mathcal{Z}^p} = \left\{ \sum_{k=1}^{\infty} t_k \left| \frac{1}{\lambda_n [Q_k]} \int_{Q_k} [f(\mathbf{y}) - f_{ak}] \, d\lambda_n(\mathbf{y}) \right|^p \right\}^{1/p}.$$

The set of functions for which $\|f\|_{\mathcal{Z}^p} < \infty$ is called the Zachary functions of bounded mean oscillation and order p, $1 \le p < \infty$. If $p = \infty$, we say that $f \in \mathcal{Z}^\infty[\mathbb{R}^n]$ if

$$\|f\|_{\mathcal{Z}^\infty} = \sup_k \left| \frac{1}{\lambda_n [Q_k]} \int_{Q_k} [f(\mathbf{y}) - f_{ak}] \, d\lambda_n(\mathbf{y}) \right| < \infty.$$

The following theorem shows how the Zachary spaces are related to the space of functions of Bounded mean oscillation $BMO[\mathbb{R}^n]$. (We omit proofs.)

Theorem 3.52. *If $\mathcal{Z}^p[\mathbb{R}^n]$ is the class of Zachary functions of bounded mean oscillation and order p, $1 \le p \le \infty$, then $\mathcal{Z}^p[\mathbb{R}^n]$ is a linear space and*

(1) $\|\lambda f\|_{\mathcal{Z}^p} \leqslant |\lambda| \, \|f\|_{\mathcal{Z}^p}$.
(2) $\|f + g\|_{\mathcal{Z}^p} \leqslant \|f\|_{\mathcal{Z}^p} + \|g\|_{\mathcal{Z}^p}$.
(3) $\|f\|_{\mathcal{Z}^p} = 0, \Rightarrow f =$ *constant* (*a.s.*).
(4) *The space $\mathcal{Z}^\infty[\mathbb{R}^n] \subset \mathcal{Z}^p[\mathbb{R}^n]$, $1 \le p < \infty$, as a dense continuous embedding.*
(5) *The space $BMO[\mathbb{R}^n] \subset \mathcal{Z}^\infty[\mathbb{R}^n]$, as a dense continuous embedding (i.e., $\|f\|_{\mathcal{Z}^\infty} \le \|f\|_{BMO}$).*

If we consider functions that differ by a constant as equivalent, it is easy to see that $\mathcal{Z}^p[\mathbb{R}^n]$ is a Banach space and $\mathcal{Z}^2[\mathbb{R}^n]$ is a Hilbert space.

We now consider the Carleson measure characterization of $BMO[\mathbb{R}^n]$ which will prove useful in construction another class of Zachary spaces that are Banach spaces (see Grafakos [GRA], p. 540). If $u(\mathbf{x}, t)$ is a solution of the heat equation:

$$u_t - \Delta u = 0, \; u(\mathbf{x}, 0) = f(\mathbf{x}),$$

where $f \in L^1_{loc}[\mathbb{R}^n]$, it can be shown that

$$\|f\|_{BMO} = \sup_{\mathbf{x}, r} \left\{ \frac{1}{\lambda_n [Q(\mathbf{x}, r)]} \int_{Q(\mathbf{x}, r)} \int_0^{r^2} |\nabla u(\mathbf{y}, t)|^2 \, dt d\lambda_n(\mathbf{y}) \right\}^{1/2},$$

where the gradient is in the weak sense. This also means that we can define the norm in $\mathcal{Z}^p[\mathbb{R}^n]$ by

$$\|f\|_{Z^p} = \sup_r \left\{ \sum_{k=1}^{\infty} t_k \left| \frac{1}{\lambda_n\left[Q_k\right]} \int_{Q_k} \int_0^{r^2} \nabla u(\mathbf{y},t) dt d\lambda_n(\mathbf{y}) \right|^p \right\}^{1/p}.$$

We define the class of functions $BMO^{-1}[\mathbb{R}^n]$, as those for which:

$$\|f\|_{BMO^{-1}} = \sup_{\mathbf{x},r} \left\{ \frac{1}{\lambda_n\left[Q(\mathbf{x},r)\right]} \int_{Q(\mathbf{x},r)} \int_0^{r^2} |u(\mathbf{y},t)|^2 \, dt d\lambda_n(\mathbf{y}) \right\}^{1/2} < \infty.$$

It is known that $BMO^{-1}[\mathbb{R}^n]$ is a Banach space in the above norm.

Definition 3.53. We say $f \in \mathcal{Z}^{-p}[\mathbb{R}^n],\ 1 \le p < \infty$ if

$$\|f\|_{Z^{-p}} = \sup_r \left\{ \sum_{k=1}^{\infty} t_k \left| \frac{1}{\lambda_n\left[Q_k\right]} \int_{Q_k} \int_0^{r^2} u(\mathbf{y},s) ds d\lambda_n(\mathbf{y}) \right|^p \right\}^{1/p} < \infty.$$

If $p = \infty$, we say that $f \in \mathcal{Z}^{-\infty}[\mathbb{R}^n]$ if

$$\|f\|_{Z^{-\infty}} = \sup_{k,\,r} \frac{1}{\lambda_n\left[Q_k\right]} \left| \int_{Q_k} \int_0^{r^2} u(\mathbf{y},s) ds d\lambda_n(\mathbf{y}) \right| < \infty.$$

Theorem 3.54. *For the class of spaces $\mathcal{Z}^{-p}[\mathbb{R}^n]$, we have:*

1. *For each p, $1 \le p \le \infty$, $\mathcal{Z}^{-p}[\mathbb{R}^n]$ is a Banach space.*
2. *The space $\mathcal{Z}^{-\infty}[\mathbb{R}^n] \subset \mathcal{Z}^{-p}[\mathbb{R}^n]$, $1 \le p < \infty$, as a dense continuous embedding.*
3. *The space $BMO^{-1}[\mathbb{R}^n] \subset \mathcal{Z}^{-\infty}[\mathbb{R}^n]$, as a dense continuous embedding, $\|f\|_{\mathcal{Z}^{-\infty}} \le \|f\|_{BMO^{-1}}$.*

Proof. The first two are obvious. To prove (3), if $f \in BMO^{-1}[\mathbb{R}^n]$, then

$$\|f\|_{Z^{-\infty}} = \sup_{k,\,r} \frac{1}{\lambda_n\left[Q_k\right]} \left| \int_{Q_k} \int_0^{r^2} u(\mathbf{y},s) ds d\lambda_n(\mathbf{y}) \right|$$

$$= \sup_{k,\,r} \frac{1}{\lambda_n\left[Q_k\right]} \left| \int_{Q_k} \int_0^{r^2} u(\mathbf{y},s) ds d\lambda_n(\mathbf{y}) \right|^{2/2}$$

$$\leqslant \sup_{\mathbf{x},r} \frac{1}{\lambda_n\left[Q(\mathbf{x},r)\right]} \left\{ \int_{Q(\mathbf{x},r)} \int_0^{r^2} |u(\mathbf{y},s)|^2 \, ds d\lambda_n(\mathbf{y}) \right\}^{1/2} = \|f\|_{BMO^{-1}}.$$

□

References

[AX] A.D. Alexandroff, Additive set functions in abstract spaces, I–III. Mat. Sbornik N. S. **8**(50), 307–348 (1940); Ibid. **9**(51), 563–628 (1941); Ibid. **13**(55), 169–238 (1943)

[AL] A. Alexiewicz, Linear functionals on Denjoy-integrable functions. Colloq. Math. **1**, 289–293 (1948)

[ASV] D.D. Ang, K. Schmitt, L.K. Vy, A multidimensional analogue of the Denjoy-Perron-Henstock-Kurzweil integral. Bull. Belg. Math. Soc. Simon Stevin **4**, 355–371 (1997)

[BA] S. Banach, *Théorie des Opérations Linéaires*. Monografj Matematyczn, vol. 1 (Z Subwencji Funduszu Kultury Narodowej, Warsaw, 1932)

[B] V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*. Springer Monographs in Mathematics (Springer, New York, 2010)

[BR] R.G. Bartle, Return to the Riemann integral. Am. Math. Mon. **103**, 625–632 (1996)

[BD] D. Blackwell, L.E. Dubins, On existence and nonexistence of proper, regular conditional distributions. Ann. Probab. **3**, 741–752 (1975)

[BO1] S. Bochner, Additive set functions on groups. Ann. Math. **40**, 769–799 (1939)

[BO2] S. Bochner, Finitely additive integral. Ann. Math. **41**, 495–504 (1940)

[CL] J.A. Clarkson, Uniformly convex spaces. Trans. Am. Math. Soc. **40**, 396–414 (1936)

[CO] H.O. Cordes, *The Technique of Pseudodifferential Operators* (Cambridge University Press, Cambridge, 1995)

[DZ] R.O. Davies, Z. Schuss, A proof that Henstock's integral includes Lebesgue's . Lond. Math. Soc. **2**, 561–562 (1970)

[DFN] B. de Finetti, *Theory of Probability*, vol. I (Wiley, New York, 1974)

[DI] J. Diestel, *Sequences and Series in Banach Spaces.* Graduate Texts in Mathematics (Springer, New York, 1984)

[DU] J. Diestel, J.J. Uhl Jr., *Vector Measures.* Mathematical Surveys, vol. 15 (American Mathematical Society, Providence, RI, 1977)

[DU] L.E. Dubins, Paths of finitely additive Brownian motion need not be bizarre, in *Seminaire de Probabilités XXXIII*, ed. by J. Azéma, M. Émery, M. Ledoux, M. Yor. Lecture Notes in Mathematics, vol. 1709 (Springer, Berlin/Heidelberg, 1999), pp. 395–396

[DUK] L.E. Dubins, K. Prikry, On the existence of disintegrations, in *Seminaire de Probabilités XXIX*, ed. by J. Azéma, M. Émery, P.A. Meyer, M. Yor. Lecture Notes in Mathematics, vol. 1613 (Springer, Berlin/Heidelberg, 1995), pp. 248–259

[DS] N. Dunford, J.T. Schwartz, *Linear Operators Part I: General Theory*, Wiley Classics edition (Wiley Interscience, New York, 1988)

[EV] L.C. Evans, *Partial Differential Equations.* AMS Graduate Studies in Mathematical, vol. 18 (American Mathematical Society, Providence, RI, 1998)

[FK] G. Fichtenholtz, L. Kantorovich, Linear operations in semi-ordered spaces I. Math. Sb. **7**(49), 209–284 (1940)

[GA] G.P. Galdi, *An Introduction to the Mathematical Theory of the Navier-Stokes Equations, Vol. II*, 2nd edn. Springer Tracts in Natural Philosophy, vol. 39 (Springer, New York, 1997)

[GOR] R.A. Gordon, *The Integrals of Lebesgue, Denjoy, Perron and Henstock.* Graduate Studies in Mathematics, vol. 4 (American Mathematical Society, Providence, RI, 1994)

[GRA] L. Grafakos, *Classical and Modern Fourier Analysis* (Person Prentice-Hall, New Jersey, 2004)

[GR] L. Gross, Abstract Wiener spaces, in *Proceedings of 5th Berkeley Symposium on Mathematics Statistics and Probability* (1965), pp. 31–42

[HS] R. Henstock, *The General Theory of Integration* (Clarendon Press, Oxford, 1991)

[HI] T.H. Hildebrandt, On bounded functional operations. Trans. Am. Math. Soc. **36**, 868–875 (1934)

[J] F. Jones, *Lebesgue Integration on Euclidean Space*, revised edn. (Jones and Bartlett Publishers, Boston, 2001)

[KAA] V. Koltchinskii, C.T. Abdallah, M. Ariola, P. Dorato, D. Panchenko, Improved sample complexity estimates for statistical learning control of uncertain systems. IEEE Trans. Autom. Control **45**, 2383–2388 (2000)

[KB] J. Kuelbs, Gaussian measures on a Banach space. J. Funct. Anal. **5**, 354–367 (1970)

[KW] J. Kurzweil, *Nichtabsolut konvergente Integrale.* Teubner-Texte zür Mathematik, Band 26 (Teubner Verlagsgesellschaft, Leipzig, 1980)

[L] P.D. Lax, Symmetrizable linear transformations. Commun. Pure Appl. Math. **7**, 633–647 (1954)

[LE] S. Leader, The theory of L^p-spaces for finitely additive set functions. Ann. Math. **58**, 528–543 (1953)

[GL] G. Leoni, *A First Course in Sobolev Spaces.* AMS Graduate Studies in Mathematics, vol. 105 (American Mathematical Society, Providence, RI, 2009)

[LL] E.H. Lieb, M. Loss, *Analysis.* AMS Graduate Studies in Mathematics, vol. 14 (American Mathematical Society, Providence, RI, 1997)

[McS] E.J. McShane, A unified theory of integration. Am. Math. Mon. **80**, 349–359 (1973)

[NI] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods* (SIAM, Philadelphia, 1992)

[PT] A. Papageorgiou, J.G. Traub, Faster evaluation of multidimensional integrals. Comput. Phys. **11**, 574–578 (1997)

[PTR] S. Paskov, J.G. Traub, Faster valuation of financial derivatives. J. Portf. Manag. **22**, 113–120 (1995)

[PF] W.F. Pfeffer, *The Riemann Approach to Integration: Local Geometric Theory*. Cambridge Tracts in Mathematics, vol. 109 (Cambridge University Press, Cambridge, 1993)

[PF1] W.F. Pfeffer, *Derivation and Integration: Local* (Cambridge University Press, Cambridge, 2001)

[PO] L. Pontryagin, *Topological Groups*. Transl. Russian, Arlen Brown (Gordon and Breach Science Publishers, New York, 1966)

[SK] S. Saks, *Theory of the Integral* (Dover Publications, New York, 1964)

[SC] L. Schwartz, *Théorie des Destributions* (Hermann, Paris, 1966)

[SY] G.R. Sell, Y. You, *Dynamics of Evolutionary Equations*. Applied Mathematical Sciences, vol. 143 (Springer, New York, 2002)

[ST] V. Steadman, Theory of operators on Banach spaces. Ph.D thesis, Howard University, 1988

[TA1] E. Talvila, Integrals and Banach spaces for finite order distributions. Bull. Belg. Math. Soc. Simon Stevin **4**, 355–371 (1997)

[TA2] E. Talvila, The distributional Denjoy integral. Real Anal. Exch. **33**, 51–82 (2008)

[T2] R. Temam, *Infinite Dimensional Dynamical Systems in Mechanics and Physics*. Applied Mathematical Sciences, vol. 68 (Springer, New York, 1988)

[T1] R. Temam, *Navier-Stokes Equations, Theory and Numerical Analysis* (AMS Chelsea Publishing, Providence, RI, 2001)

[TY] L. Tuo-Yeong, Some full descriptive characterizations of the Henstock-Kurweil integral in Euclidean space. Czechoslov. Math. J. **55**, 625–637 (2005)

[TY1] L. Tuo-Yeong, *Henstock-Kurweil Integration on Euclidean Spaces*. Series in Real Analysis, vol. 12 (World Scientific, New Jersey, 2011)

[YH] K. Yosida, E. Hewitt, Finitely additive measures. Trans. Am. Math. Soc. **72**, 46–66 (1952)

Chapter 4

Analysis on Hilbert Space

In this chapter we study operator theory on separable Hilbert spaces. The first part is devoted to some important results on the integration of operator-valued functions. Although some additional material has also been included, the second part is a review of standard theory of operators on Hilbert spaces. The only new material is a recent new spectral representation for linear operators based on the polar decomposition. All results and concepts that are independent of the inner product apply to Banach spaces and will be used in the next chapter without further comment.

4.1. Part I: Analysis on Hilbert Space

4.1.1. Integration of Operator-Valued Functions. In this section, we discuss a few additional topics in integration theory. Our main interest is the class of operator-valued functions f, defined on a measure space $(\Omega, \mathfrak{B}[\Omega], \lambda)$, $\Omega \subset \mathbb{R}$ with values in $\mathcal{B} = L[\mathcal{H}]$. We extend the HK-integral in this setting and prove a version of the Riesz Representation Theorem.

The major problem with integration for operator-valued functions is that these functions need not have a Lebesgue (like) integral

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_4

(see [HP], pp. 71–80). However, they always have a HK-integral. To understand the problem, we need a few definitions.

Definition 4.1. The function $f : \Omega \to \mathcal{B}$ is said to be:

(1) almost surely separably valued (or essentially separably valued) if there exists a subset $N \subset \Omega$ with $\lambda(N) = 0$ such that $f(\Omega \setminus N) \subset \mathcal{B}$ is separable,

(2) countably valued if it assumes at most a countable number of values in $\mathcal{B}$, assuming each value $\neq 0$ on a measurable subset of Ω, and

(3) strongly measurable if there exists a sequence $\{f_n\}$ of countably valued functions converging (a.s.) to f.

(4) Bochner integrable if $\|f\|_{\mathcal{B}}$ is Lebesgue integrable.

(5) Gelfand–Pettis integrable if $\langle f, h' \rangle_{\mathcal{B}}$ is Lebesgue integrable for each $h' \in \mathcal{B}'$.

In order to constructively define the integral one must be able to approximate it with simple functions in either the strong sense (Bochner) or the weak sense (Gelfand–Pettis). However, each simple function must be countably valued and strongly measurable in the first case or countably valued and weakly measurable in the second. In the case of current interest, $\mathcal{B}$ is not separable and a family of operator-valued functions $A : \Omega \to \mathcal{B}$ need not be almost separable valued and hence need not be strongly or weakly measurable. In this section, we provide a useful extension of the HK-integral to operator-valued functions on $\mathbb{R}$, which does not suffer from the above limitations. Since this version of the integral will be our main tool for the Feynman operator calculus, we prove all except the elementary or well-known results. (It should be noted that the theory developed does not depend on the Hilbert space structure.)

Let $[a, b] \subset \mathbb{R}$ and, for each $t \in [a, b]$, let $A(t) \in L(\mathcal{H})$ be a given family of operators.

Recall that, if $\delta(t)$ maps $[a, b] \to (0, \infty)$, and $\mathcal{P} = \{t_0, \tau_1, t_1, \tau_2, \cdots, \tau_n, t_n\}$, where $a = t_0 \leqslant \tau_1 \leqslant t_1 \leqslant \cdots \leqslant \tau_n \leqslant t_n = b$, we say it is a HK-$\delta$ partition provided that, for $0 \leqslant i \leqslant n$, $t_i - t_{i-1} < \delta(\tau_i)$.

Lemma 4.2. *Let $\delta_1(t)$ and $\delta_2(t)$ map $[a, b] \to (0, \infty)$, and suppose that $\delta_1(t) \leqslant \delta_2(t)$. Then, if $\mathcal{P}_1$ is a HK-δ_1 partition, it is also a HK-δ_2 partition.*

Definition 4.3. The family $A(t)$, $t \in [a, b]$, is said to have a (uniform) HK-integral if there is an operator $Q[a, b]$ in $L(\mathcal{H})$ such that, for each $\varepsilon > 0$, there exists a HK-δ partition such that

$$\left\| \sum_{i=1}^{n} \Delta t_i A(\tau_i) - Q[a, b] \right\| < \varepsilon.$$

In this case, we write

$$Q[a, b] = (HK) \int_a^b A(t)dt.$$

Theorem 4.4. *For $t \in [a, b]$, suppose the operators $A_1(t)$ and $A_2(t)$ both have HK-integrals, then so does their sum and*

$$(HK) \int_a^b [A_1(t) + A_2(t)]dt = (HK) \int_a^b A_1(t)dt + (HK) \int_a^b A_2(t)dt.$$

Theorem 4.5. *Suppose $\{A_k(t) \mid k \in \mathbb{N}\}$ is a family of operator-valued functions in $L[\mathcal{H}]$, converging uniformly to $A(t)$ on $[a, b]$, and $A_k(t)$ has a HK-integral $Q_k[a, b]$ for each k; then $A(t)$ has a HK-integral $Q[a, b]$ and $Q_k[a, b] \to Q[a, b]$ uniformly.*

Theorem 4.6. *Suppose $A(t)$ is Bochner integrable on $[a, b]$, then $A(t)$ has a HK-integral $Q[a, b]$ and:*

$$(B) \int_a^b A(t)dt = (HK) \int_a^b A(t)dt. \tag{4.1}$$

Proof. First, let E be a measurable subset of $[a, b]$ and assume that $A(t) = A\chi_E(t)$, where $\chi_E(t)$ is the characteristic function of E. In this case, we show that $Q[a, b] = A\lambda(E)$, where $\lambda(E)$ is the Lebesgue measure of E. Let $\varepsilon > 0$ be given and let D be a compact subset of E. Let $F \subset [a, b]$ be an open set containing E such that $\lambda(F \backslash D) < \varepsilon / \|A\|$; and define $\delta : [a, b] \to (0, \infty)$ such that:

$$\delta(t) = \begin{cases} d(t, [a, b] \backslash F), & t \in E \\ d(t, D), & t \in [a, b] \backslash E, \end{cases}$$

where $d(x\,,\, y) = |x - y|$ is the distance function. Let $\mathcal{P} = \{t_0, \tau_1, t_1, \tau_2, \cdots, \tau_n, t_n\}$ be a HK-δ partition. If $\tau_i \in E$ for $1 \leqslant i \leqslant n$, then $(t_{i-1}, t_i) \subset F$ so that

$$\left\| \sum_{i=1}^{n} \Delta t_i A(\tau_i) - A\lambda(F) \right\| = \|A\| \left[\lambda(F) - \sum_{\tau_i \in E} \Delta t_i \right]. \tag{4.2}$$

On the other hand, if $\tau_i \notin E$ then $(t_{i-1}, t_i) \cap D = \emptyset$ (empty set), then it follows that:

$$\left\| \sum\nolimits_{i=1}^{n} \Delta t_i A(\tau_i) - A\lambda(D) \right\| = \|A\| \left[\sum\nolimits_{\tau_i \notin E} \Delta t_i - \lambda(D) \right]. \quad (4.3)$$

Combining Eqs. (4.2) and (4.3), we have that

$$\left\| \sum\nolimits_{i=1}^{n} \Delta t_i A(\tau_i) - A\lambda(E) \right\| = \|A\| \left[\sum\nolimits_{\tau_i \in E} \Delta t_i - \lambda(E) \right]$$
$$\leqslant \|A\| \left[\lambda(F) - \lambda(E) \right] \leqslant \|A\| \left[\lambda(F) - \lambda(D) \right] \leqslant \|A\| \, \lambda(F \backslash D) < \varepsilon.$$

Now suppose that $A(t) = \sum_{k=1}^{\infty} A_k \chi_{E_k}(t)$. By definition, $A(t)$ is Bochner integrable if and only if $\|A(t)\|$ is Lebesgue integrable with:

$$(B) \int_a^b A(t)dt = \sum\nolimits_{k=1}^{\infty} A_k \lambda(E_k),$$

and (cf. Hille and Phillips [HP])

$$(L) \int_a^b \|A(t)\| dt = \sum\nolimits_{k=1}^{\infty} \|A_k\| \, \lambda(E_k).$$

As the partial sums converge uniformly by Theorem 4.5, $Q[a, b]$ exists and

$$Q[a, b] \equiv (HK) \int_a^b A(t)dt = (B) \int_a^b A(t)dt.$$

Now let $A(t)$ be an arbitrary Bochner integrable operator-valued function in $L(\mathcal{H})$, uniformly measurable and defined on $[a, b]$. By definition, there exists a sequence $\{A_k(t)\}$ of countably valued operator-valued functions in $L(\mathcal{H})$ which converges to $A(t)$ in the uniform operator topology such that:

$$\lim_{k \to \infty} (L) \int_a^b \|A_k(t) - A(t)\| \, dt = 0,$$

and

$$(B) \int_a^b A(t)dt = \lim_{k \to \infty} (B) \int_a^b A_k(t)dt.$$

Since the $A_k(t)$ are countably valued,

$$(HK) \int_a^b A_k(t)dt = (B) \int_a^b A_k(t)dt,$$

so

$$(B) \int_a^b A(t)dt = \lim_{k \to \infty} (HK) \int_a^b A_k(t)dt.$$

We are done if we show that $Q[a,b]$ exists. Since every L-integral is a HK-integral, $f_k(t) = \|A_k(t) - A(t)\|$ has a HK-integral. This means that $\lim_{k\to\infty} (HK)\int_a^b f_k(t)dt = 0$. Let $\varepsilon > 0$ and choose m so large that

$$\left\|(B)\int_a^b A(t)dt - (HK)\int_a^b A_m(t)dt\right\| < \varepsilon/4$$

and

$$(HK)\int_a^b f_k(t)dt < \varepsilon/4.$$

Choose δ_1 so that if $\{t_0, \tau_1, t_1, \tau_2, \cdots, \tau_n, t_n\}$ is a HK-δ_1 partition, then

$$\left\|(HK)\int_a^b A_m(t)dt - \sum\nolimits_{i=1}^n \Delta t_i A_m(\tau_i)\right\| < \varepsilon/4.$$

Now choose δ_2 so that whenever $\{t_0, \tau_1, t_1, \tau_2, \cdots, \tau_n, t_n\}$ is a HK-δ_2 partition,

$$\left\|(HK)\int_a^b f_m(t)dt - \sum\nolimits_{i=1}^n \Delta t_i f_m(\tau_i)\right\| < \varepsilon/4.$$

Set $\delta = \delta_1 \wedge \delta_2$ so that by Lemma 4.2, $\{t_0, \tau_1, t_1, \tau_2, \cdots, \tau_n, t_n\}$ is a HK-δ_1 and HK-δ_2 partition so that:

$$\begin{aligned}
&\left\|(B)\int_a^b A(t)dt - \sum_{i=1}^n \Delta t_i A(\tau_i)\right\| \leqslant \left\|(B)\int_a^b A(t)dt - (HK)\int_a^b A_m(t)dt\right\| \\
&+ \left\|(HK)\int_a^b A_m(t)dt - \sum\nolimits_{i=1}^n \Delta t_i A_m(\tau_i)\right\| \\
&+ \left|(HK)\int_a^b f_m(t)dt - \sum\nolimits_{i=1}^n \Delta t_i f_m(\tau_i)\right| \\
&+ (HK)\int_a^b f_m(t)dt < \varepsilon.
\end{aligned}$$

□

Recall that a function $g : [a,b] \subset \mathbb{R} \to \mathcal{H}$ is of bounded variation or BV, if

$$\sup\left\|\sum_{i=1}^n [g(b_i) - g(a_i)]\right\|,$$

where the supremum is taken over all partitions $\mathbb{P} = \{(a_1,b_1),\ldots,(a_n,b_n)\}$ of nonoverlapping subintervals of $[a,b]$. In this case, we set

$$\sup\left\|\sum_{i=1}^n [g(b_i) - g(a_i)]\right\| = BV_a^b(g).$$

Theorem 4.7. *Let $g : [a, b] \to \mathcal{H}$ be of BV.*

(1) *If h is continuous on $[a, b]$, then*

$$I = HK \int_a^b h(s) dg(s)$$

exists.

(2) *If in addition A is a closed densely defined linear operator on $\mathcal{H}$, $g \in D(A)$ and $Ag(s) = f(s)$ is of BV, then*

$$AI = A \int_a^b h(s) dg(s) = \int_a^b h(s) df(s). \tag{4.4}$$

Proof. Since h is continuous and $g(s)$, $f(s)$ is of BV, we need to only prove the existence of a strong Riemann–Stieltjes integral.

To prove (1), define $S_{\mathbb{P}_k}$ by:

$$S_{\mathbb{P}_k} = \sum_{\mathbb{P}_k} h(s_i) \left[g(b_i) - g(a_i) \right]$$

Since h is continuous, it is uniformly continuous so that, given $\varepsilon > 0$ there exists a $\delta > 0$, such that $|h(s) - h(t)| < \varepsilon$ whenever $|s - t| < \delta$. If $\mathbb{P}_1$, $\mathbb{P}_2$ are partitions such that $\max_i \{b_i - a_i\} = |\mathbb{P}| < \frac{\delta}{2}$, by a standard application of the triangle inequality, we have that:

$$\left| \mathbf{J}_{\mathcal{H}} \left(S_{\mathbb{P}_1} - S_{\mathbb{P}_2} \right) \right| \leqslant 2\varepsilon BV_a^b \{ \mathbf{J}_{\mathcal{H}}(g) \},$$

for all linear functionals $\mathbf{J}_{\mathcal{H}}(\cdot) \in \mathcal{H}'$. Now,

$$\begin{aligned} &BV_a^b \{ \mathbf{J}_{\mathcal{H}}(g) \} \leqslant BV_a^b \left[\Re \{ \mathbf{J}_{\mathcal{H}}(g) \} \right] + BV_a^b \left[\Im \{ \mathbf{J}_{\mathcal{H}}(g) \} \right] \\ &\leqslant 4 \sup_{\mathbb{P}} \left| \mathbf{J}_{\mathcal{H}} \left\{ \sum_i \left[g(b_i) - g(a_i) \right] \right\} \right|, \end{aligned}$$

where the sup is over partitions of $[a, b]$. By definition, there is an M such that $BV_a^b \{ \mathbf{J}_{\mathcal{H}}(g) \} \leq M \| J_{\mathcal{H}} \|$. It follows that

$$\| S_{\mathbb{P}_1} - S_{\mathbb{P}_2} \| = \sup_{\| \mathbf{J}_{\mathcal{H}} \| = 1} \left| \mathbf{J}_{\mathcal{H}} \left(S_{\mathbb{P}_1} - S_{\mathbb{P}_2} \right) \right| \leqslant 2M\varepsilon,$$

so that the strong Riemann–Stieltjes integral exists.

To prove (2), for any $\mathbb{P}$, $AS_{\mathbb{P}}(h, g) = S_{\mathbb{P}}(h, Ag)$ because A is linear. Since we know that

$$\lim_{|\mathbb{P}| \to 0} S_{\mathbb{P}}(h, g) = \int_a^b h(s) dg(s)$$

and Ag is BV. Applying the above for Ag gives us:

$$A \lim_{|\mathbb{P}|\to 0} S_{\mathbb{P}}(h,g) = \lim_{|\mathbb{P}|\to 0} S_{\mathbb{P}}(h,Ag) = \int_a^b h(s)dAg(s) = \int_a^b h(s)df(s).$$

Since A is closed, $\int_a^b h(s)dg(s) \in D(A)$ and Eq. (4.4) is satisfied. □

The next few results are required for a later section on spectral theory. Let $(\Omega, \mathfrak{B}(\Omega), \mu)$ be a measure space, where Ω is a subset of $\mathbb{R}_I^n$ and $\mu = \lambda_\infty$ is Lebesgue measure on $\mathbb{R}_I^n$. We would now like to describe the dual space $L^\infty(\Omega, \mathfrak{B}(\Omega), \mu)^*$, of $L^\infty(\Omega, \mathfrak{B}(\Omega), \mu)$ over $\mathbb{C}$, the complex numbers in a little more detail.

Theorem 4.8. *If $\ell \in L^\infty(\Omega, \mathfrak{B}(\Omega), \mu)^*$, there is a finitely additive complex signed measure μ_ℓ of bounded total variation and absolutely continuous with respect to μ, such that*

$$\ell(\phi) = \int_\Omega \phi(x)d\mu_\ell(x), \quad \phi \in L^\infty[\Omega, \mathfrak{B}[\Omega], \mu],$$

so that $L^\infty(\Omega, \mathfrak{B}(\Omega), \mu)^* = \mathfrak{M}(\Omega, \mathfrak{B}(\Omega), \mu)$.

Proof. By the Jordan Decomposition Theorem, every complex measure μ can be written as $\nu = \nu_1 + \nu_2 + i(\nu_3 + \nu_4)$, where ν_1, ν_3 are positive measures and ν_1, ν_4 are negative measures. Thus, it suffices to prove the theorem when μ_ℓ is real. Let $\ell \in L^\infty(\Omega, \mathfrak{B}(\Omega), \mu)^*$ and, for each $B \in \mathfrak{B}[\Omega]$ set $\mu_\ell(B) = \ell(I_B)$, where I_B is the characteristic function of B. If $B_1,\ B_2 \in \mathfrak{B}[\Omega],\ B_1 \cap B_2 = \emptyset$, then $I_{B_1+B_2} = I_{B_1} + I_{B_2}$ so that

$$\ell(I_{B_1} + I_{B_2}) = \ell(I_{B_1}) + \ell(I_{B_2}) \Rightarrow \mu_\ell\left(B_1 \cup B_2\right) = \mu_\ell\left(B_1\right) + \mu_\ell\left(B_2\right).$$

Since

$$\sup_{B\in\mathfrak{B}} |\mu_\ell(B)| = \sup_{B\in\mathfrak{B}} |\ell(I_B)| \leqslant \|\ell\| \, \|I_B\| < \infty,$$

we see that μ_ℓ is of bounded variation.

Let $\phi \in L^\infty(\Omega, \mathfrak{B}(\Omega), \mu)$ be arbitrary. For any $\varepsilon > 0$, there is a simple function s_ε such that

$$s_\varepsilon = \sum\nolimits_{i=1}^N a_i I_{B_i}, \quad \mu\left(B_i \cap B_j\right) = 0,\ i \neq j, \quad \bigcup_{i=1}^N B_i = \Omega$$

and

$$\left\| \phi - \sum\nolimits_{i=1}^N a_i I_{B_i} \right\| < \varepsilon, \quad \text{so that } \left\| \ell(\phi) - \sum\nolimits_{i=1}^N a_i \mu_\ell(B_i) \right\| < \varepsilon \left\|\ell\right\|.$$

It follows that

$$\ell(\phi) = \int_\Omega \phi(x)d\mu_\ell(x), \quad \text{and} \quad \|\ell\| = \sup_{ess\,\sup|\phi|\leqslant 1} \left| \int_\Omega \phi(x)d\mu_\ell(x) \right|.$$

Finally, since $\mu(B) = 0$ implies that $I_B = 0$ (a.e.), it follows that $\mu_\ell(B) = 0$ so that μ_ℓ is absolutely continuous with respect to μ. □

From here, we see that $L^1(\Omega, \mathfrak{B}(\Omega), \mu)^{**} = \mathfrak{M}(\Omega, \mathfrak{B}(\Omega), \mu)$ and, the injection of $L^1(\Omega, \mathfrak{B}(\Omega), \mu) \to \mathfrak{M}(\Omega, \mathfrak{B}(\Omega), \mu)$ is dense. Since $L^1(\mathbb{R}_I^n, \mathfrak{B}(\mathbb{R}_I^n), \mu)$ is Banach algebra under convolution, it is easy to prove that

Corollary 4.9. *$\mathfrak{M}(\mathbb{R}_I^n, \mathfrak{B}(\mathbb{R}_I^n), \mu)$ is Banach algebra under convolution.*

Recall that if Ω is an open subset of $\mathbb{R}^n$, then $\mathbb{C}_c(\Omega)$ is the set of all continuous functions defined on Ω that vanish outside a compact set.

Corollary 4.10. *If $\phi \in \mathbb{C}_c(\Omega)$, then for each $\ell \in \mathbb{C}_c(\Omega)^*$, there is a countably additive complex measure μ_ℓ such that.*

$$\ell(\phi) = \int_\Omega \phi(x)d\mu_\ell(x).$$

Proof. Since $\phi \in L^\infty[\Omega, \mathfrak{B}[\Omega], \mu]$, we can represent $\ell(\phi)$ as

$$\ell(\phi) = \int_\Omega \phi(x)d\mu_\ell(x),$$

with μ_ℓ finitely additive. As ϕ is continuous, we can extend μ_ℓ to a countably additive measure. To see this, first, assume that μ_ℓ is positive and follow the standard procedure used in Chap. 2 (or use the Daniell method, see Royden [RO]). In the general case, by Jordan's decomposition Theorem, we can write $\mu_\ell = \mu_{\ell_1}^+ + \mu_{\ell_2}^- + i[\mu_{\ell_3}^+ + \mu_{\ell_4}^-]$, where $\mu_{\ell_i}^+$, $i = 1, 3$ are the positive parts and $\mu_{\ell_i}^-$, $i = 2, 4$ are the negative parts. □

Remark 4.11. When Ω has finite measure, the above result can be extended to $\mathbb{C}_0[\Omega]$, the continuous functions that vanish at the boundary, also known as the Riesz Representation Theorem (see Rudin [R1]).

4.2. Part II: Operators on Hilbert Space

Introduction. The need and motivation to learn the basics of Hilbert space operator theory is necessary for any student and/or researcher in the physical sciences, applied mathematics, partial differential equations, or stochastic analysis. In this second part, we develop those aspects that will be necessary for all of the later chapters. For those with a limited background in operator theory and/or its applications, we have included additional material in order to make the presentation self-contained. We also include all proof details but the presentation is rather terse, following a theorem proof protocol, so that consultation of one of the standard references is recommended (i.e., [L1], [R], [RS], [RS], or [YS]).

4.2.1. Basic Results.

Let $\mathcal{H}$ be a Hilbert space and let $\mathbf{V}(\mathcal{H})$ the set of linear contraction operators on $\mathcal{H}$ (i.e., $\|A\|_{\mathcal{H}} \leqslant 1$). We denote by $L[\mathcal{H}]$ the set of bounded linear operators on $\mathcal{H}$ and by $\mathcal{C}(\mathcal{H})$ the set of closed densely defined linear operators on $\mathcal{H}$. The graph of a linear operator A is denoted by: $G(A) = \{(f, Af); f \in D(A)\}$, where $D(A) \subset \mathcal{H}$ is the domain of A, and $\overline{G}(A)$ is its closure in the product space $\mathcal{H}\times\mathcal{H}$. We say that A is a closed linear operator if $G(A) = \overline{G}(A)$. It is easy to show that this is equivalent to the statement that, for each sequence $\{f_n : n \in \mathbb{N}\}$ in $D(A)$ with $f_n \to f$ and $Af_n \to f^*$, we have that $f \in D(A)$ and $Af = f^*$.

Definition 4.12. The numerical range $\mathcal{N}(A)$ of an operator A, in $\mathcal{C}(\mathcal{H})$, is defined by:

$$\mathcal{N}(A) = \{(Af, f)\,|f \in D(A),\ \|f\| = 1\}.$$

The orthogonal complement $\mathcal{M}^{\perp}$ of a linear subspace $\mathcal{M} \subset \mathcal{H}$ is defined by:

$$\mathcal{M}^{\perp} = \{g \in \mathcal{H}\,|\,(f, g)_{\mathcal{H}} = 0,\ \forall f \in \mathcal{M}\}.$$

If M is not closed then it is easy to see that $[\mathcal{M}^{\perp}]^{\perp}$ is the closure, and if it is closed $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$.

Let $\mathcal{H}'$ be the dual space of $\mathcal{H}$ and, for each $g \in \mathcal{H}$, define $\mathbf{J}_g(f) = (f, g)$.

Theorem 4.13. *The mapping $\mathbf{J}_g(\cdot)$ is a conjugate-linear isometry of $\mathcal{H}$ onto $\mathcal{H}'$.*

Proof. It is clear that for each fixed g, $\mathbf{J}_g(\cdot)$ is a continuous linear functional on $\mathcal{H}$, so that $\mathbf{J}_g(\cdot) \in \mathcal{H}'$. Furthermore, by the Schwarz inequality, we have $\|\mathbf{J}_g(f)\|_{\mathcal{H}} \leqslant \|g\|_{\mathcal{H}} \|f\|_{\mathcal{H}}$ and this implies that $\|\mathbf{J}_g(\cdot)\|_{\mathcal{H}} \leqslant \|g\|_{\mathcal{H}}$. On the other hand:

$$\|\mathbf{J}_g(g)\|_{\mathcal{H}} = \|g\|_{\mathcal{H}}^2 \leqslant \|\mathbf{J}_g(\cdot)\|_{\mathcal{H}} \|g\|_{\mathcal{H}} \Rightarrow \|g\|_{\mathcal{H}} \leqslant \|\mathbf{J}_g(\cdot)\|_{\mathcal{H}}.$$

Thus, we see that $\|g\|_{\mathcal{H}} = \|\mathbf{J}_g(\cdot)\|_{\mathcal{H}'}$.

If $\mathbf{L} \in \mathcal{H}'$, let N be the null space of $\mathbf{L}$. It is easy to see that $\mathbf{L} \neq 0 \ \Rightarrow N \neq \mathcal{H}$, so that $N^{\perp} \neq 0$. Let $f \in N^{\perp}$, with $f \neq 0$. Without loss, we can assume that $\mathbf{L}(f) = 1$. Let $h \in \mathcal{H}$, and set $h' = h - \bar{\mathbf{L}}(h)f$ (conjugate). It is easy to see that $\mathbf{L}(h') = 0$, so that $0 = (h', f) = (h, f) - \mathbf{L}(h) \|f\|_{\mathcal{H}}^2$. Thus,

$$\mathbf{L}(h) = \frac{(h, f)}{\|f\|_{\mathcal{H}}^2},$$

so that $\mathbf{L} = \mathbf{J}_g$, where $g = \frac{f}{\|f\|_{\mathcal{H}}^2}$. □

We will denote by $\mathbf{J}$ the mapping from $\mathcal{H}$ onto $\mathcal{H}'$. The second part of the above result is one version of the Riesz Representation Theorem. The following related result will also be useful.

Theorem 4.14 (Lax–Milgram Theorem). *Let $B(f, g)$ be a conjugate bilinear functional on $\mathcal{H} \times \mathcal{H}$ such that there exists δ_1, $\delta_2 > 0$, with $\delta_1 \|f\|^2 \leq B(f, f)$ and $|B(f, g)| \leq \delta_2 \|f\| \|g\|$. Then there is a positive linear operator T with a bounded inverse T^{-1}, such that $(f, g) = B(f, Tg)$, $\|T\| \leq \delta_1^{-1}$ and $\left\|T^{-1}\right\| \leq \delta_2$ for all f, $g \in \mathcal{H}$.*

Proof. If D is the set of all h such that $(f, g) = B(f, h)$, then $D \neq \emptyset$. To see this, suppose $B(f, h) = 0$ for all $f \in \mathcal{H}$, then $0 = B(h, h) \geq \delta_1 \|h\|^2$, so $h = 0$.

Since both (f, g) and $B(f, g)$ are conjugate bilinear functionals, there is a linear operator T, $D(T) = D$, such that $Tg = h$. For $h \in D$, we have

$$\delta_1 \|Tg\|^2 \leqslant B(Tg, Tg) = (Tg, g) \leqslant \|y\| \|Tg\|,$$

so that T is bounded. To see that T is closed, let $g_n \in D$, $g_n \to g$. since T is continuous, $\{Tg_n\}$ is a Cauchy sequence, so that $h = \lim_{n\to\infty} Tg_n$. Since $\mathbf{J}_f(\cdot) = (f, \cdot)$ is continuous, $\lim_{n\to\infty} (f, g_n) = (f, h)$. It follows that $\lim_{n\to\infty} (f, g_n) = B(f, h)$, so that $h \in D$ and $T(g) = h$.

Suppose that $D \neq \mathcal{H}$, then there exists a $0 \neq g_0 \in \mathcal{H},\ g_0 \in D^{\perp}$. Let $L(f) = B(f, g_0)$, $f \in \mathcal{H}$. It is clear that $L(f)$ is continuous, since

$$|L(f)| = |B(f, g_0)| \leqslant \delta_2 \|f\| \|g_0\| .$$

It follows from Theorem 4.13 that there is a g_0^*, such that $B(f, g_0) = (f, g_0^*)$, for all $f \in \mathcal{H}$. But then, $g_0^* \in D$, and $Tg_0^* = g_0$. However, $\delta_1 \|g_0\|^2 \leq B(g_0, g_0) = (g_0, g_0^*) = 0$, so that $g_0 = 0$ contradicting our assumption that $g_0 \neq 0$. It follows that $D = \mathcal{H}$.

Since $Tg = 0 \Rightarrow g = 0$, we see that T is injective and T^{-1} exists. By the same proof above, with T replaced by T^{-1}, shows that it is defined on all of $\mathcal{H}$ and

$$\left|(f, T^{-1}g)\right| = |B(f, g)| \leqslant \delta_2 \|f\| \|g\| ,$$

so that $\left\|T^{-1}\right\| \leq \delta_2$. □

4.2.2. Resolvent and Spectrum.

Definition 4.15. Let A be any operator in $\mathcal{C}(\mathcal{H})$.

(1) The resolvent set, $\rho_{\mathcal{H}}(A)$, is the set of all complex numbers λ such that $(\lambda I - A)$ has a bounded inverse on $\mathcal{H}$.

(2) The complement of $\rho_{\mathcal{H}}(A)$, $\sigma_{\mathcal{H}}(A)$ is called the spectrum of A. It is a disjoint union of three parts: the point, continuous, and residual spectra.

(3) The point spectrum, $\sigma_{\mathcal{H}}^p(A)$, is the set of all complex numbers λ such that $(\lambda I - A)$ has no inverse.

(4) The continuous spectrum, $\sigma_{\mathcal{H}}^c(A)$, is the set of all complex numbers λ such that $(\lambda I - A)$ has an unbounded densely defined inverse.

(5) The residual spectrum, $\sigma_{\mathcal{H}}^r(A)$, is the set of all complex numbers λ such that $(\lambda I - A)$ has an inverse that is not densely defined.

The operator $R(\lambda, A) = (\lambda I - A)^{-1}$, when it exists, is called the resolvent associated with A.

Definition 4.16. If $A \in L[\mathcal{H}]$, the spectral radius r_A is defined by

$$r_A = |\sigma(A)| = \sup_{\lambda \in \sigma(A)} |\lambda| .$$

Example. Let $\mathcal{H} = L^2[a,b],\ -\infty < a < b < \infty$ and define:

(1) $A_1(u) = u'$

(2) $A_2(u) = u',\ u(a) = 0$

(3) $A_3(u) = u',\ u(b) = ku(a),\ k \neq 0$ is constant.

(1) In the first case, $u' - \lambda u = 0$ always has a solution, $u = ce^{\lambda x}$, c arbitrary. It follows that $\sigma_{\mathcal{H}}(A_1) = \mathbb{C}$, the whole complex plane, so that $\rho_{\mathcal{H}}(A_1) = \emptyset$.

(2) In the second case, it is easy to see that $\sigma_{\mathcal{H}}(A_2) = \emptyset$ so that $\rho_{\mathcal{H}}(A_2) = \mathbb{C}$ and $R(\lambda, A_2)(y) = e^{\lambda y} \int_a^y e^{-\lambda x} u(x)dx$ exists for all λ.

(3) In the third case, the point spectrum of A_3,

$$\sigma^p_{\mathcal{H}}(A_3) = \{\lambda_n = \frac{1}{(b-a)}[\ln k + 2in\pi],\ n \in \pm\mathbb{N}\}$$

while the continuous spectrum $\sigma^c_{\mathcal{H}}(A)$ and the residual spectrum $\sigma^r_{\mathcal{H}}(A)$ are both empty; and the resolvent set $\rho_{\mathcal{H}}(A)$ contains all $\lambda \neq \lambda_n$.

Theorem 4.17. *If $A \in L[\mathcal{H}]$, we have:*

(1) *The resolvent set $\rho(A)$ is an open subset of $\mathbb{C}$.*

(2) *The resolvent of A, $R(\lambda, A) = (\lambda I - A)^{-1}$ is an analytic function on $\rho(A)$.*

(3) *If λ, $\mu \in \rho(A)$, then* (First Resolvent Identity)

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda) R(\lambda, A) R(\mu, A).$$

(4) *The spectrum $\sigma(A)$ is a nonempty compact subset of $\mathbb{C}$ and*

$$r_A = \lim_{n\to\infty} \|A^n\|^{1/n}.$$

Proof. To prove (1), let $\lambda \in \rho(A)$ and suppose that $L \in L[\mathcal{H}]$, with $\|L\| < \|R(\lambda, A)\|$, then $R(\lambda, A) - L$ is invertible. To see this, note that because

$$\left\| \frac{L}{R(\lambda, A)} \right\| < 1, \quad \Rightarrow [I - \frac{L}{R(\lambda, A)}]^{-1} = \sum_{n=0}^{\infty} \left(\frac{L}{R(\lambda, A)} \right)^n.$$

If we choose $L = hI$ with h small enough, it follows that $\rho(A)$ is an open subset in $\mathbb{C}$.

To prove (2), note that $\sigma(A)$ is the compliment of $\rho(A)$, so that it is closed. By definition of $R(\lambda, A)$, we see that the series

$$(\lambda I - A)^{-1} = \frac{1}{\lambda}\left(I - \frac{1}{\lambda}A\right)^{-1} = \sum_{n=0}^{\infty} A^n \lambda^{-n-1} \tag{4.5}$$

converges if $\left\|\lambda^{-1}A\right\| < 1$ (i.e., $\|A\| < |\lambda|$). It follows from this that every $\lambda \in \sigma(A)$ is bounded above by $\|A\|$, so that $\sigma(A)$ is compact. Equation (4.5) is a Laurent expansion of $R(\lambda, A)$ around ∞. If we integrate $R(\lambda, A)$ about any contour $C = \{z : |z| > \|A\|\}$, using the Cauchy integral Theorem, we get a nonzero value. On the other hand, if $R(\lambda, A)$ were analytic in all of $\mathbb{C}$, the integral would be zero. It follows that $\sigma(A)$ is not empty.

To prove (3), start with the identity

$$(\mu I - A) - (\lambda I - A) = (\mu - \lambda) I$$

and multiply both sides by $R(\lambda, A)\, R(\mu, A)$.

For a proof of (4), look at the nth term of (4.5). Choose $p < n$ and write $n = kp + r$, $0 \leq r < k$, so that we can write this term as $\left|\lambda^{-(n+1)}\right| \left\|A^{kp+r}\right\| = \left|\lambda^{-(n+1)}\right| \left\|A^{kp}\right\| \|A^r\| \leqslant \left|\lambda^{-(n+1)}\right| \left\|A^k\right\|^p \|A^r\|$. It follows that

$$\begin{aligned}\left\|\sum_{n=0}^{\infty} A^n \lambda^{-(n+1)}\right\| &\leqslant \sum_{n=0}^{\infty} \|A^n\| \left|\lambda^{-(n+1)}\right| \\ &\leqslant \left(\sum_{r=0}^{k-1} \|A^r\| \left|\lambda^{-(r+1)}\right|\right) \sum_{k=p}^{\infty} \left(\left\|A^k\right\| \left|\lambda^{-k}\right|\right)^p\end{aligned}$$

and the series converges absolutely if $\left\|A^k\right\| \left|\lambda^{-k}\right| < 1$. It follows that the series converges absolutely if $\left\|A^k\right\|^{1/k} < |\lambda|$, so that

$$r_A \leqslant \liminf_{k\to\infty} \left\|A^k\right\|^{1/k}. \tag{4.6}$$

If we now let $\varepsilon > 0$ be given and let $C = \{z : |z| = r_A + \varepsilon\}$ be a contour in $\rho(A)$ winding once around $\sigma(A)$. In this case, it is easy to see that

$$\frac{1}{2\pi i} \oint_C (\lambda I - A)^{-1} \lambda^n d\lambda = A^n. \tag{4.7}$$

From here, we get that

$$\|A^n\| \leqslant a\,(r_A + \varepsilon)^{n+1}, \quad a = \sup_{|\lambda| = r_A + \varepsilon} \|R(\lambda, A)\|.$$

It follows that

$$\|A^n\|^{1/n} \leqslant a^{1/n} (r_A + \varepsilon)^{1+1/n} \Rightarrow \limsup_{n\to\infty} \|A^n\|^{1/n} \leqslant (r_A + \varepsilon) .$$

Since this is true for arbitrary $\varepsilon > 0$, along with Eq. (4.6) proves that

$$r_A = \lim_{n\to\infty} \|A^n\|^{1/n} . \tag{4.8}$$

□

Using Eq. (4.7) and the linear nature of the resolvent, we have:

Theorem 4.18. *If $A \in L[\mathcal{H}]$, we have:*

(1) *If f is an analytic function on an open set containing $\sigma(A)$ and C is a contour in $\rho(A)$ winding once around $\sigma(A)$, then the mapping*

$$f(A) = \oint_C (\lambda I - A)^{-1} f(\lambda) d\lambda \tag{4.9}$$

is a homomorphism into $L[\mathcal{H}]$.

(2) *$\sigma[f(A)] = f(\sigma[A])$.*

Proof. It is clear that the map $f \to f(A)$ is linear and continuous. To show that it is multiplicative, let f and g be analytic on some open set $O \supset \sigma(A)$. Let C_1 and C_2 be contours in $O \cap \rho(A)$ with C_1 inside C_2 having no point in common. That is, C_2 winds once around every point of C_1 and C_1 winds zero times around every point of C_2. From (4.9), we have

$$\begin{aligned}
f(A)g(A) &= \oint_{C_2} (\lambda I - A)^{-1} f(\lambda) d\lambda \oint_{C_1} (\mu I - A)^{-1} g(\mu) d\mu \\
&= \oiint\limits_{C_1 \cup C_2} (\lambda I - A)^{-1} (\mu I - A)^{-1} f(\lambda) g(\mu) d\lambda d\mu \\
&= \oiint\limits_{C_1 \cup C_2} (\lambda - \mu)^{-1} \left[(\mu I - A)^{-1} - (\lambda I - A)^{-1} \right] f(\lambda) g(\mu) d\lambda d\mu \\
&= \oint_{C_1} \left[\oint_{C_2} (\lambda - \mu)^{-1} f(\lambda) d\lambda \right] (\mu I - A)^{-1} g(\mu) d\mu \\
&\quad - \oint_{C_2} \left[\oint_{C_1} (\lambda - \mu)^{-1} g(\mu) d\mu \right] (\lambda I - A)^{-1} f(\lambda) d\lambda
\end{aligned} \tag{4.10}$$

Since C_2 winds once around every point of C_1, the first integral above is $f(\mu)$, by Cauchy's integral theorem. On the other hand, since C_1

does not wind around any point in C_2, the μ integration of the second term above is zero, so that

$$f(A)g(A) = \oint_{C_2} (\mu I - A)^{-1} f(\mu)g(\mu)d\mu.$$

To prove (2), suppose $\mu \in \sigma[f(A)]$ and $\mu \neq f(\lambda)$ for any $\lambda \in \sigma(A)$. Then $f(\lambda) - \mu$ is nonzero on $\sigma(A)$, so that $g(\lambda) = (f(\lambda) - \mu)^{-1}$ is analytic on an open set containing $\sigma(A)$, so we can define $g(A)$ by Eq. (4.9). It follows from $(f(\lambda) - \mu)g(\lambda) = 1$ that $(f(A) - \mu I)g(A) = I$, so that $g(A) = (f(A) - \mu I)^{-1}$. Thus, $\mu \notin \sigma[f(A)]$. This is a contradiction, so that $\mu = f(\lambda)$ for some $\lambda \in \sigma(A)$.

If $\mu = f(\lambda)$ for some $\lambda \in \sigma(A)$, then the function

$$h(\alpha) = \frac{f(\alpha) - f(\lambda)}{\alpha - \lambda}$$

is analytic in an open set containing $\sigma(A)$, so that $h(A)$ is well defined by Eq. (4.9). Since $h(\alpha)(\alpha - \lambda) = f(\alpha) - f(\lambda)$, we see that

$$(A - \lambda I)h(A) = f(A) - f(\lambda).$$

Since $\lambda \in \sigma(A)$, we see that $(A - \lambda I)$ is not invertible so that neither is $h(A)$. This shows that $f(\lambda) \in \sigma[f(A)]$. $\square$

4.3. The Adjoint Operator

We now turn to a discussion of the adjoint of a linear operator $A \in \mathcal{C}[\mathcal{H}]$, with dense domain $D(A)$. For each $f \in D(A)$, the relation $(Af, g) = (f, g')$ uniquely determines g' for each g (since $D(A)$ is dense in $\mathcal{H}$).

Definition 4.19. The adjoint of A, A^*, is defined and uniquely determined by the equation $A^* g = g'$.

(Note that A^* is linear and closed whether A is closed or not.)

The relation between A and its adjoint A^* can be simply expressed in terms of graphs. We have $(f, -A^*g) + (Af, g) = 0$, which shows that $(Af, g) \in \mathcal{H} \times \mathcal{H}$ is annihilated by $(f, -A^*g) \in \mathcal{H} \times \mathcal{H}$. Consequently, $G(A) \perp G'(-A^*)$, where $G'(-A^*)$ denotes the inverse graph of $-A^*$, $G'(-A^*) = \{(f, -A^*g) \,|\, g \in D(A^*), f \in D(A)\}$. The inverse graph $G'(-A^*)$ is a closed subspace of $\mathcal{H} \times \mathcal{H}$ since A^* is closed.

There is yet another way to express the relation between A and A^*. Let $\mathbf{J}(\cdot)$ denote the conjugate isomorphism between $\mathcal{H}$ and its dual space $\mathcal{H}'$, so that $(\mathbf{J})(g) = \mathbf{J}_g = (\cdot, g)$. If A' is the dual operator to A, defined on $\mathcal{H}'$, it is not hard to see that $A^* = \mathbf{J}^{-1}A'\mathbf{J}$. It is this form that will be used when we study operator theory on Banach spaces in the next chapter.

If we restrict ourselves to $L[\mathcal{H}]$, we have:

Theorem 4.20. *The mapping $A \to A^*$ is an involution on $L[\mathcal{H}]$, which means that:*

(1) $\|A^*\| = \|A\|$.

(2) $\|A^*A\| = \|A\|^2$.

(3) $(A+B)^* = A^* + B^*$.

(4) $(aA)^* = \bar{a}A^*$.

(5) $(AB)^* = B^*A^*$.

(6) $A^{**} = A$.

The verification of the above properties is an easy exercise and defines $L[\mathcal{H}]$ as a C^*-algebra.

Definition 4.21. Let A be a closed densely defined linear operator on $\mathcal{H}$. We say that:

(1) A is self-adjoint if $D(A) = D(A^*)$ and $Af = A^*f$ for all $f \in D(A)$.

(2) A is normal if $D(A) = D(A^*)$ and $AA^*f = A^*Af$ for all $f \in D(A)$.

(3) A is unitary if $A^*Af = AA^*f = f$ for all $f \in \mathcal{H}$.

(4) A is a projection if $A^2f = Af$ for all $f \in \mathcal{H}$.

(5) A is m-dissipative if, for each $\lambda \in \rho_{\mathcal{H}}(A)$, the range $\mathcal{R}(\lambda I - A)$ is equal to $\mathcal{H}$ and

$$Re(Af, f) \ \leq 0, \text{ for } f \in D(A).$$

(We also say that $-A$ is m-accretive if A is m-dissipative.) The next result is due to von Neumann [VN1].

Theorem 4.22 (von Neumann)**.** *Let $\mathcal{H}$ be a separable Hilbert space and let A be a closed densely defined linear operator on $\mathcal{H}$. Then A has a well-defined adjoint A^* defined on $\mathcal{H}$ such that the following assertions are valid.*

(1) *$A^*A \geqslant 0$ is accretive (i.e., $(A^*Af, f) \geq 0$ for $f \in D(A)$).*

(2) *$(A^*A)^* = A^*A$.*

(3) *The operator $\lambda I + A^*A$ has a bounded inverse for all $\lambda > 0$.*

Proof. From our earlier discussion, it is clear that A^* exists as a closed densely defined linear operator on $\mathcal{H}$. To prove (1), let $f \in D(A)$. Then, from $0 \leqslant \|Af\|_{\mathcal{H}}^2 = (Af, Af)_{\mathcal{H}} = (f, A^*Af)_{\mathcal{H}}$, we see that the operator $A^*A \geqslant 0$.

To prove (2), note that $(f, A^*Af)_{\mathcal{H}} = ((A^*A)^*f, f)_{\mathcal{H}}$ and $(Af, Af)_{\mathcal{H}} = (A^*Af, f)_{\mathcal{H}}$, so that $(A^*A)^* = A^*A$. Since $D(A^*A) = D[(A^*A)^*]$, we see that A^*A is self-adjoint.

To prove (3), observe that, $G(A) \perp G'(-A^*)$, $G(A)$ and $G'(-A^*)$ are complementary subspaces of $\mathcal{H}\times\mathcal{H}$, so that $G(A) \oplus G'(-A^*) = \mathcal{H}\times\mathcal{H}$. Thus, for $\lambda > 0$, any vector in $\mathcal{H} \times \mathcal{H}$ can be written in the form $(Ag, \lambda g) + (f, -A^*f)$ for some $g \in D(A)$, $f \in D(A^*)$. In particular, we choose vectors of the form $(0, h) \in \mathcal{H} \times \mathcal{H}$. In this case, we have: $h = \lambda g - A^*f$ and $0 = Ag + f$. Thus, $f = -Ag$ and $h = (\lambda I + A^*A)g$. Since h can be arbitrarily chosen in $\mathcal{H}$, the range of $(\lambda I + A^*A)$ is the whole space $\mathcal{H}$. Thus, the inverse operator $S = (\lambda I + A^*A)^{-1}$ exists and is determined by the equation $Sh = g$. It is linear and defined on all of $\mathcal{H}$, with

$$(Sh, h)_{\mathcal{H}} = (g, (\lambda I + A^*A)g)_{\mathcal{H}} = \|Ag\|_{\mathcal{H}}^2 + \lambda \|g\|_{\mathcal{H}}^2 \geqslant 0. \tag{4.11}$$

Since (4.11) is real-valued, it follows that S is symmetric. Furthermore, $Sh = 0$ if and only if $h = 0$ and, since $\|Ag\|_{\mathcal{H}}^2 + \lambda \|g\|_{\mathcal{H}}^2 \geqslant \lambda \|g\|_{\mathcal{H}}^2 = \lambda \|Sh\|_{\mathcal{H}}^2$, we have that:

$$\|Sh\|_{\mathcal{H}}^2 = \|g\|_{\mathcal{H}}^2 \leqslant (Sh, h)_{\mathcal{H}} \leqslant \|h\|_{\mathcal{H}} \|Sh\|_{\mathcal{H}}, \tag{4.12}$$

so that $\|Sh\|_{\mathcal{H}} \leqslant \|h\|_{\mathcal{H}}$. Hence, S is a bounded linear (contraction) operator and, from (4.12), we have that $0 \leqslant (Sh, h)_{\mathcal{H}} \leqslant \|h\|_{\mathcal{H}}^2$. As an exercise, it is easy to show that the spectrum of a bounded operator is contained in the closure of its numerical range. From the above result, we see that the spectrum of S lies on the real line between 0 and 1. Therefore, S is positive and, being bounded and symmetric, it is self-adjoint. Thus, we see that $S^{-1} = \lambda I + A^*A$ is also self-adjoint and densely defined (from which, it follows again that A^*A is positive, self-adjoint and densely defined). □

Using properties of the resolvent, it is easy to show that condition (3) is satisfied if $\mathcal{R}(I + A^*A) = \mathcal{H}$ (range).

Definition 4.23. Let A be a closed densely defined linear operator on $\mathcal{H}$. A subspace D is said to be a core for A if $\overline{D} = G(A)$, the graph of A (i.e., if the set of elements $\{g, Ag\},\ g \in D$ is dense in $G(A)$).

The following theorem describes the relationship between A and A^*.

Theorem 4.24. *Let $A \in L[\mathcal{H}]$ then:*

(1) *The null space of A, $N(A) = \mathcal{R}(A^*)^{\perp}$.*

(2) *The null space of A^*, $N(A^*) = \mathcal{R}(A)^{\perp}$.*

(3) *$\mathcal{R}(A)$ is dense in $\mathcal{H}$ if and only if A^* is injective.*

(4) *A is injective if and only if $\mathcal{R}(A^*)$ is dense in $\mathcal{H}$.*

Proof. The proof of (2) and (4) is similar to that of (1) and (3).

For (1), if $f \in N(A)$, $Af = 0$, so that $(Af, g) = 0$ for all $g \in \mathcal{H}$. It follows that $(f, A^*g) = 0$, so that $f \in \mathcal{R}(A^*)^{\perp}$.

To prove (3), $\mathcal{R}(A)$ is dense in $\mathcal{H}$ if and only if $\mathcal{R}(A)^{\perp} = \{0\}$ (since it is closed). Thus, as $N(A^*) = \mathcal{R}(A)^{\perp}$, $N(A^*) = \{0\}$, so that A^* is injective. The opposite direction is obvious. □

4.3.1. The Polar Decomposition of A. In this section, we focus on the square root of an m-accretive operator and use it to derive an important representation for linear operators on $\mathcal{C}(\mathcal{H})$. Recall from Theorem 4.22 that for any closed densely defined linear operator A on $\mathcal{H}$, A^*A is m-accretive. In order to make our approach compatible with that used in partial differential equations and semigroup theory, we work with $S = -A^*A$, so that S is m-dissipative. Without loss, we can assume that S is strictly negative (i.e., there is a $\delta > 0$ and $(Sf, f) < -\delta \|x\|^2$, $f \in D(S)$). Thus, for any α, with $Re(\alpha) > 0$, $(\alpha I - S)^{-1}$ exists and

$$\left\|(\alpha I - S)^{-1}\right\| \leqslant \frac{1}{\mathrm{Re}(\alpha) + \delta}. \tag{4.13}$$

Define the branch of $\alpha^{-1/2}$ so that $Re(\alpha^{-1/2}) > 0$, when $Re(\alpha) > 0$. (This branch is a one-valued function in the α-plane cut along the negative real axis.) The half-plane $\{\alpha : Re(\alpha) > -\delta\} \subset \rho(S)$, so we can choose a path C, which goes from $+\infty \to +\infty$ in $\rho(S)$, turns around the origin in the positive direction and is positive where C touches the negative real axis. We define T by:

$$T = \frac{1}{2\pi i} \oint_C (\alpha I - S)^{-1} \alpha^{-1/2} d\alpha. \tag{4.14}$$

The integral above is absolutely convergent from Eq. (4.13), and we see that $T \in L[\mathcal{H}]$. If we choose a second path C' as in Theorem 4.18

[see Eq. (4.10)] and use the first resolvent identity (Theorem 4.17 (3)) on $(\alpha I - S)^{-1}(\bar{\alpha} I - S)^{-1}$, we have that

$$(T)^2 = \frac{1}{2\pi i}\oint_C (\alpha I - S)^{-1}\,\alpha^{-1} d\alpha = (-S)^{-1}. \tag{4.15}$$

Now note that if $Tf = 0$, then $(-S)^{-1}f = 0$, so that $f = 0$. It follows that T is injective, so it is invertible.

Definition 4.25. We define $|A| = [A^*A]^{1/2} = (-S)^{1/2} = T^{-1}$, so that $[A^*A]^{-1/2} = |A|^{-1} = T$.

It is clear that both $|A|$ and $|A|^{-1}$ commute with $[A^*A]$. In order to obtain the standard representation of $|A|$ as a fractional power of a closed densely defined linear operator, we reduce the path of C to the union of the upper and lower edges of the real axis and use $\alpha^{-1/2} = i\lambda^{-1/2}$, to obtain:

$$|A|^{-1/2} = \frac{1}{\pi}\int_0^\infty (\lambda I + [A^*A])^{-1}\,\lambda^{-1/2} d\lambda. \tag{4.16}$$

It is easy to see that $|A|$ and $[AA^*]^{1/2} = |A^*|$ are nonnegative densely defined closed self-adjoint linear operators on $\mathcal{H}$. In the general case, when A is m-dissipative, Eq. (4.16) becomes ($Re[\lambda] > 0$)

$$(-A)^{-1/2}f = \frac{1}{\pi}\int_0^\infty \lambda^{-1/2}\,(\lambda I - A)^{-1}\,f d\lambda. \tag{4.17}$$

An equally important representation is obtained from Eq. (4.17) by multiplying both sides by A to get, for $f \in D(A)$,

$$(-A)^{1/2}f = \frac{1}{\pi}\int_0^\infty \lambda^{-1/2}\,(\lambda I - A)^{-1}\,Af d\lambda. \tag{4.18}$$

In what follows, we let $T = |A|$ and $\bar{T} = |A^*|$.

Theorem 4.26. *If $A \in \mathcal{C}(\mathcal{H})$, then $D(T) = D(A)$ and:*

(1) *There exists an unique partial isometry U such that*

$$A = UT = \bar{T}U,\ \ A^* = U^*\bar{T} = TU^*.$$

(2) *We also have that $\bar{T} = UTU^*$ and $T = U^*\bar{T}U$.*

(3) *If A is normal (i.e., $A^*A = AA^*$), then $U\,|A| = |A|\,U$ and we can take U to be unitary, $UU^* = U^*U = I$.*

Proof. To prove (1) let $f \in D(A)$, then

$$\|Af\|^2 = (Af, Af) = (f, A^*Af) = (Tf, Tf) = \|Tf\|^2 .$$

This implies that $D(T) = D(A)$ and $Tf \to Af$ defines an isometric map U of the range of $T, \mathcal{R}(T)$, onto the range of $A, \mathcal{R}(A)$, such that $Af = UTf$. The operator U can be extended to an isometric map from the closure, $\overline{\mathcal{R}}(T)$ onto $\overline{\mathcal{R}}(A)$ by continuity. If we define $Ug = 0$ for $g \in \mathcal{R}(T)^{\perp}$, we can extend U to $L[\mathcal{H}]$. This defines U as an unique partial isometry from $\overline{\mathcal{R}}(T)$ onto $\overline{\mathcal{R}}(A)$, so that $A = UT$ with $D(T) = D(A)$. It is easy to see from Theorem 4.20 (5) that $A^* = TU^*$ on $D(A^*)$.

From $A^* = TU^*$, we see that $T = A^*U$, so that

$$AA^* = (UT)(TU^*) = UA^*AU^* \Rightarrow AA^*U = UA^*A.$$

It follows that $\bar{T}U = UT = A$ and $T = U^*\bar{T}U$. Using this last result in $A^* = TU^*$, we see that $A^* = (U^*\bar{T}U)U^* = U^*\bar{T}$.

The proof of (2) follows from $T = U^*\bar{T}U$, obtained above. To prove (3), if A is normal, then $\bar{T} = T$, so that $UT = TU$. From here, we see that U maps the range of A onto the range of A^*. Let M_1 be the closure of the range of T and let $M_2 = M_1^{\perp}$, so that $\mathcal{H} = M_1 \oplus M_2$. It is easy to see that the null space of A, $N(A) = N(A^*) = M_2$. Thus, if we set $U = I$, the identity operator on M_2, we can extend U to all of $\mathcal{H}$ as a unitary operator. □

4.4. Compact Operators

It is fair to say that the inverse (or shifted inverse) of almost all closed densely defined linear (differential) operators encountered in practice are either compact or can be approximated by compact operators. They are also very close to the operators (matrices) studied in elementary linear algebra. This section is devoted to a study of the relevant results on compact operators. We will not use them directly in the following chapters. However, they are just below the surface of all that follows, and one should know as much as possible about them.

Definition 4.27. Let $A \in L[\mathcal{H}]$.

(1) We say that A is of finite rank if $dim[\mathcal{R}(A)] < \infty$ (dimension of the range) and call it the rank of A. We let $\mathbb{F}_{\mathcal{H}}$ be the set of all operators of finite rank in $L[\mathcal{H}]$.

(2) We say that A is compact if and only if for each bounded sequence $\{g_n\} \subset \mathcal{H}$, $\{Ag_n\}$ has is a convergent subsequence. We let $\mathbb{K}_{\mathcal{H}}$ be the set of all compact operators in $L[\mathcal{H}]$.

Lemma 4.28. *If $\mathbb{B}_{\mathcal{H}}$ is the unit ball in $\mathcal{H}$, then the bounded linear operator A is compact if and only if $A(\mathbb{B}_{\mathcal{H}})$ has a compact closure.*

Proof. If for each bounded sequence $\{g_n\} \subset \mathcal{H}$, $\{Ag_n\}$ has is a convergent subsequence, then $A(\mathbb{B}_{\mathcal{H}})$ has a compact closure. On the other hand, if $A(\mathbb{B}_{\mathcal{H}})$ has a compact closure, let $\{g_n\}$ be a bounded sequence in $\mathcal{H}$. Then, the sequence $\{a^{-1}g_n\} \subset \mathbb{B}_{\mathcal{H}}$, where $a = \sup_n \|g_n\|_{\mathcal{H}}$, so that $a^{-1}Ag_n$ has a convergent subsequence. It follows that the set $\{Ag_n\}$ has is a convergent subsequence. $\square$

Lemma 4.29. *Let $M \subset \mathcal{H}$ be a nonempty proper closed linear subspace. Then, given $\varepsilon > 0$, there exists a $f \in M$ such that $\|f\| = 1$ and $1 - \varepsilon \leq dist(f, M)$.*

Proof. Let $g \in \mathcal{H} \setminus M$, with $0 < d = dist(g, M)$ and choose $f_0 \in M$ such that

$$d \leq \|g - f_0\| \leq \frac{d}{1-\varepsilon}, \text{ and set } f = \frac{g - f_0}{\|g - f_0\|}.$$

Using this f, it follows that, for each $h \in M$,

$$\|f - h\| = \left\| \frac{g - f_0}{\|g - f_0\|} - h \right\| \geqslant \frac{d}{\|g - f_0\|} \geqslant 1 - \varepsilon,$$

since $f_0 + h\,\|g - f_0\| \in M$. $\square$

Theorem 4.30. *If the unit ball $\mathbb{B}_{\mathcal{H}}$, of $\mathcal{H}$ is compact, then $\mathcal{H}$ is finite dimensional.*

Proof. If we assume that $\mathcal{H}$ is infinite dimensional, let M_n be a proper increasing sequence of subspaces. By the above lemma, there exists a sequence $\{f_n\}$, with $f_n \in M_n$, $\|f_n\| = 1$ and $\frac{1}{2} \leq dist(f_n, M_{n-1})$. In particular, if $n \neq m$, $\|f_n - f_m\| \geq \frac{1}{2}$. It follows that the sequence $\{f_n\}$ has no convergent subsequence, contradicting our assumption that $\mathbb{B}_{\mathcal{H}}$ is compact. $\square$

Lemma 4.31. *Let $\mathbb{D}$ is a (countable) dense subset of $\mathcal{H}$ and let $\{F_n\}$ be a sequence of mappings on $\mathcal{H}$ such that, for $f \in \mathbb{D}$ the closure of $\{F_n(f)\}$ is compact. Then there is a subsequence $\{F_{n_i}\}$, that converges for each $f \in \mathbb{D}$.*

Proof. Let $\mathbb{D} = \{f_n\}$ and, by the compactness of $\{F_n(f_1)\}$, let $\{F_{n,1}(f_1)\}$ be a convergent subsequence. Now, from $\{F_{n,1}\}$, find a subsequence $\{F_{n,2}\}$, such that $\{F_{n,2}(f_2)\}$ converges. Continuing, we obtain a subsequence $\{F_{n,i}\}$ that converges at f_i for each $f_i \in \mathbb{D}$. If we let $\{F_{n,n}\}$ be the diagonal subsequence, $\{F_{n,n}(f_i)\}$ converges for all $f_i \in \mathbb{D}$, then we have our convergent subsequence. □

Theorem 4.32. *For the set of all compact operators in $\mathbb{K}_{\mathcal{H}} \in L[\mathcal{H}]$, we have:*

(1) *$\mathbb{K}_{\mathcal{H}}$ is a closed subspace of $L[\mathcal{H}]$ in the operator norm.*
(2) *If A is compact and $T \in L[\mathcal{H}]$, then AT and TA are compact operators, so that $\mathbb{K}_{\mathcal{H}}$ is an ideal in $L[\mathcal{H}]$.*
(3) *The operator A is compact if and only if A^* is compact, so that $\mathbb{K}_{\mathcal{H}}$ is a *ideal in $L[\mathcal{H}]$.*

Proof. It is easy to see that $\mathbb{K}_{\mathcal{H}}$ is a linear subspace. To see that it is closed, let $A_n \to A$, so that $\|A_n - A\| \to 0$. Let $\{f_n\}$ be a bounded sequence in $\mathcal{H}$. Following the procedure of Lemma 4.31, let $\{f_n^1\}$ be a subsequence such that $\{A_1 f_n^1\}$ converges. Continuing, we obtain a sequence $\{f_n^k\}$ such that $\{A_k f_n^k\}$ converges for each k. Let $g_n = f_n^n$, so that $\{A_k g_n\}$ converges for each fixed k. Let $\varepsilon > 0$ be given, let $M = \sup \|f_n\|$ and choose N_1 such that, for $k > N_1$, $\|A_k - A\| < \frac{\varepsilon}{4M}$. Then chose N_2 such that $\|A_k g_n - A_k g_m\| < \frac{\varepsilon}{2}$ for $n > N_1$ and $m > N_2$. Then, for k, n, $m > max\{N_1, N_2\}$,

$$\begin{aligned}\|Ag_n - Ag_m\| &\leqslant \|(A - A_k)(g_n - g_m)\| + \|A_k (g_n - g_m)\| \\ &\leqslant 2M\|A - A_k\| + \|(g_n - g_m)\| < \varepsilon.\end{aligned}$$

It follows that the sequence $\{Ag_n\}$ converges, so that A is compact. Thus, $\mathbb{K}_{\mathcal{H}}$ is closed in the $L[\mathcal{H}]$ norm.

To prove (2), let A be compact and $T \in L[\mathcal{H}]$. Let $\{f_n\}$ be a bounded sequence in $\mathcal{H}$ and choose a subsequence $\{f_{n_i}\}$ such that Af_{n_i} converges. It is clear that TAf_{n_i} also converges. For the other case, $\{Tf_n\}$ is bounded, so we can choose a subsequence $\{Tf_{n_i}\}$ such that ATf_{n_i} converges.

To prove (3), let $A = UT$ be the polar decomposition of A. Since $T = U^*A$, T is compact and, since $A^* = TU^*$, it follows from (2) that A^* is compact. □

Recall that a sequence $\{f_n\}$ is said to converge to f weakly ($f_n \xrightarrow{w} f$) if $(f_n, g)_{\mathcal{H}} \to (f, g)_{\mathcal{H}}$ for all $g \in \mathcal{H}$.

Theorem 4.33. *Let $f_n \xrightarrow{w} f$ and let A be compact. Then $\|Af_n - Af\|_{\mathcal{H}} \to 0$ as $n \to \infty$ (i.e., $Af_n \to Af$ strongly).*

Proof. By the Uniform Boundedness Theorem, the sequence $\{\|f_n\|_{\mathcal{H}}\}$ is bounded (see Theorem 1.54). If $g_n = Af_n$, then $(Af_n, \xi) = (f_n, A^*h) \to (f, A^*h)$ for all $h \in \mathcal{H}$. It follows that $g_n \xrightarrow{w} \gamma = Af$. If the sequence $\{g_n\}$ does not converge in norm, then there is a $L > 0$ and a subsequence $\{h_{n_k}\}$ such that $0 < L \leq \|g_{n_k} - g\|_{\mathcal{H}}$. As the subsequence $\{f_{n_k}\}$ is bounded and A is compact, $\{g_{n_k}\}$ has a subsequence that converges to $\bar{g} \neq g$. This is a contradiction since $g_n \xrightarrow{w} g$. Thus, $Af_n \to Af$ in norm. □

Theorem 4.34. *Every compact operator on $\mathcal{H}$ is the norm limit of a sequence of operators of finite rank.*

Proof. Let $\{f_n\}$ be a orthonormal basis for $\mathcal{H}$ and for each n, let M_n be closed subspace the spanned by $\{f_1, \dots f_n\}$ and let P_n be the projection of $\mathcal{H}$ onto M_n. Define λ_n by:

$$\lambda_n = \sup_{f \in M_n^\perp,\ \|f\|=1} \|Af\|_{\mathcal{H}}\,.$$

It is clear that the sequence $\{\lambda_n\}$ is decreasing and has a limit $\lambda \geq 0$. If $\lambda \neq 0$, then, for each n, let $g_n \in M_n^\perp$, $\|g_n\| = 1$, with $\|Ag_n\| \geq \frac{1}{2}\lambda$. Since $g_n \xrightarrow{w} 0$, by the last theorem $Ag_n \xrightarrow{s} 0$. It follows that $\lambda = 0$ and $A_n = AP_n$ is an operator of finite rank, with $A_n f \to Af$, for each $f \in \mathcal{H}$. □

Corollary 4.35. *If $\{\phi_n\}$ is an orthonormal basis for $\mathcal{H}$ and A is compact, then A has the representation:*

$$A = \sum_{n=1}^{\infty} A\phi_n(\cdot, \phi_n).$$

Proof. For any $f \in \mathcal{H}$, we have

$$f = \sum_{n=1}^{\infty} (f, \phi_n)\phi_n,$$

where convergence is in the $\mathcal{H}$ norm. It follows that

$$Af = \sum_{n=1}^{\infty} (f, \phi_n)A\phi_n.$$

□

4.4.1. Fredholm Theory. In this section, we discuss the Fredholm theory and related results for compact linear operators. We begin with an interesting analytic result.

Theorem 4.36 (Analytic Fredholm Theorem)**.** *Let $\mathbb{D}$ be an open connected domain in $\mathbb{C}$, let f be a compact analytic operator-valued map of $\mathbb{D} \to L[\mathcal{H}]$, and let $h(z) = [1 - f(z)]^{-1}$, $z \in \mathbb{D}$. In this case, only one of the following is true:*

(1) *The function $h(z)$ does not exist for any $z \in \mathbb{D}$.*

(2) *There exists a discrete set $E \subset \mathbb{D}$ such that $h(z)$ exist for all $z \in \mathbb{D} \setminus E$. In this case, $h(z)$ is meromorphic in $\mathbb{D}$ (i.e., the ratio of two holomorphic functions), analytic in $\mathbb{D} \setminus E$ and the residues at the poles of $h(z)$ are of finite rank. Furthermore, if $z \in E$, the equation $f(z)\phi = \phi$ has nontrivial solutions.*

Proof. Let $z_0 \in \mathbb{D}$ and, choose $r > 0$ such that $\|f(z) - f(z_0)\| < \frac{1}{2}$, when $\mathbb{D}_r = \{z : |z - z_0| < r\}$. Since $f(z_0)$ is compact, we can find an operator A of finite rank such that $\|f(z_0) - A\| < \frac{1}{2}$. It follows that for $z \in \mathbb{D}_r$, $\|f(z) - A\| < 1$. Thus,

$$\sum_{n=0}^{\infty} (f(z) - A)^n = (I - f(z) + A)^{-1} = g_A(z).$$

It follows that $g_A(z)$ is analytic and since A is of finite rank, by Corollary 4.35, for any $u \in \mathcal{H}$, we can write (with $A\phi_n = \psi_n$)

$$Au = \sum_{n=1}^{N} (u, \phi_n)\psi_n,$$

where $(\phi_1, \phi_2, \ldots, \phi_N)$ is a finite set of orthogonal functions. Let $\phi_n(z) = g_A(z)^*\phi_n$ and define

$$\begin{aligned} g(z)u = A[g_A(z)u] &= \sum_{n=1}^{N} (g_A(z)u, \phi_n)\, \psi_n \\ = \sum_{n=1}^{N} (u, \phi_n(z))\, \psi_n &\Rightarrow g(z) = \sum_{n=1}^{N} (\,\cdot\,, \phi_n(z))\, \psi_n. \end{aligned} \tag{4.19}$$

Now note that

$$\begin{aligned} (I - g(z))\,(I + A - f(z)) &= (I + A - f(z)) - g(z)\,(I + A - f(z)) \\ &= (I + A - f(z)) - A = (I - f(z))\,. \end{aligned}$$

It follows that $(I - f(z))^{-1}$ exists for $z \in \mathbb{D}_r$ if and only if $(I - g(z))$ has an inverse. Furthermore, $f(z)\phi = \phi$ has a nontrivial solution if and only if $g(z)\phi = \phi$ has one.

If $g(z)\phi = \phi$, we see from Eq. (4.19) that

$$\phi = \sum_{k=1}^{N} a_k \psi_k = \sum_{n=1}^{N} (\phi, \phi_n(z))\, \psi_n.$$

It follows that

$$a_n = \sum_{m=1}^{N} (\psi_m, \phi_n(z))\, a_m. \tag{4.20}$$

This is the standard homogeneous problem for N equations in N unknowns. Thus, Eq. (4.20) has a nontrivial solution if and only if

$$d(z) = \det\left[\delta_{nm} - (\psi_m, \phi_n(z))\right] = 0.$$

Since $\phi_n(z)$ is an analytic function in $\mathbb{D}_r$, we see that $d(z)$ is also. Thus, either $E_r = \mathbb{D}_r$ or E_r has a finite number of points in $\mathbb{D}_r$, where $E_r = \{z : \det[\delta_{nm} - (\psi_m, \phi_n(z))] = 0$.

On the other hand, if $z \in \mathbb{D}_r \setminus E_r$, $d(z) \neq 0$ and, from elementary linear algebra, $[I - g(z)]^{-1}$ exists. Furthermore, the inhomogeneous problem $[I - g(z)]\phi = \psi$ has a unique solution. Thus, $[I - g(z)]^{-1}$ exists if and only if $z \notin E_r$.

Since $\mathbb{D}$ is connected, by analytic continuation we can extend $f(z)$ to all of $\mathbb{D}$. We are done since $[I - f(z)]^{-1}$ is meromorphic and the residues at each pole are finite rank operators. □

If we set $f(z) = zA$ in Theorem 4.36, we have:

Theorem 4.37 (The Fredholm Alternative)**.** *If $A \in \mathbb{K}_{\mathcal{H}}$ then either the equation $A\phi = \phi$ has a nontrivial solution or $[I - A]^{-1} \in L[\mathcal{H}]$.*

Theorem 4.38. *For $A \in L[\mathcal{H}]$, we have:*

(1) (The Riesz–Schauder Theorem) *If $A \in \mathbb{K}_{\mathcal{H}}$, then the spectrum $\sigma(A)$ is a discrete set of points in $\mathbb{C}$ with no limit point other than 0. If $0 \neq \lambda \in \sigma(A)$, then the eigenspace for λ is finite dimensional (i.e., λ is an eigenvalue of finite multiplicity).*

(2) (The Hilbert–Schmidt Theorem) *If $A \in \mathbb{K}_{\mathcal{H}}$ is self-adjoint, then the normalized eigenfunctions for A, $\{\phi_n\}$ form a complete orthonormal basis for $\mathcal{H}$, with $A\phi_n = \lambda_n \phi_n$, and $\lambda_n \to 0$, as $n \to \infty$.*

Proof. To prove (1), note that, since $f(z) = zA$ is an entire map of $\mathbb{C} \to \mathbb{K}_{\mathcal{H}}$, it follows that the set $\mathbb{D}_f = \{z : f(z)\psi = \psi \text{ has a nontrivial solution}\}$ is discrete, $z = 0 \notin \mathbb{D}_f$ and, if $\lambda^{-1} \notin \mathbb{D}_f$, then $(\lambda - A)^{-1}$ exists with

$$(\lambda - A)^{-1} = \frac{1}{\lambda}(I - \frac{1}{\lambda}A)^{-1}.$$

Finite dimensionality of the eigenspace for λ follows from compactness.

To prove (2), first note that,

$$\left\|A^2\right\| = \|A^*A\| \leqslant \|A^*\| \, \|A\| = \|A\|^2$$

and

$$\|A\|^2 = \sup_{\|\varphi\|=1} \|A\varphi\|^2 = \sup_{\|\varphi\|=1} (\varphi, A^*A\varphi) \leqslant \|A^*A\| \, ,$$

so that $\left\|A^2\right\| = \|A\|^2$. By induction, we get that $\left\|A^{2n}\right\| = \|A\|^{2n}$, for all n. It follows that the spectral radius of A is $\|A\|$.

For each eigenvalue of A, choose an orthonormal basis for its eigenspace (which is finite dimensional from (1)). Since eigenvectors for distinct eigenvalues are orthogonal, the set of eigenvectors $\{\phi_n\}$ for all λ is an orthonormal set. Let $\mathcal{E}$ be the Hilbert subspace spanned by this family. Since A is self-adjoint, A is invariant on both $\mathcal{E}$ and $\mathcal{E}^{\perp}$. If A_1 is the restriction of A to $\mathcal{E}^{\perp}$, it is clearly self-adjoint and compact. By (1), if there is a $\lambda \neq 0 \in \sigma(A_1)$, it is necessarily an eigenvalue of A_1 and also of A. It follows that $r_{A_1} = 0$, since all eigenvectors of A are in $\mathcal{E}$. Since $r_{A_1} = \|A_1\|$, we see that $A_1 = 0$, so that $\mathcal{E}^{\perp} = \{0\}$. Since $A\phi = 0$ implies that $\phi \in \mathcal{E}$, we see that $\mathcal{E} = \mathcal{H}$, so that the family $\{\phi_n\}$ is complete. Furthermore, since $\sigma(A)$ has no limit point except possibly 0, we see that $\lambda_n \to 0$. □

Theorem 4.39 (Canonical Form for Compact Operators)**.** *If A be compact, then there exists two orthonormal families of functions $\{\phi_n\}$, $\{\psi_n\}$, that need not be complete, such that*

$$Af = \sum_{n=1}^{\infty} \lambda_n (f, \phi_n)\psi_n.$$

Where $\{\lambda_n\}$ are the eigenvalues of $T = [A^*A]^{1/2}$ (called the singular values of A).

Proof. First, note that T is compact and self-adjoint, so that by the last theorem, $\mathcal{H}$ has a orthonormal basis consisting of its eigenvectors

$\{\phi_n\}$, with $T\phi_n = \lambda_n\phi_n$, where the $\{\lambda_n\}$ are the nonzero eigenvalues of T. For $f \in \mathcal{H}$, we have that

$$f = \sum_{k=1}^{\infty} (f, \phi_n)\, \phi_n.$$

It follows that

$$Tf = \sum_{k=1}^{\infty} \lambda_n\, (f, \phi_n)\, \phi_n.$$

Since $A = UT$, if we set $\psi_n = U\phi_n$, the family $\{\psi_n\}$ is orthonormal and,

$$Af = \sum_{n=1}^{\infty} \lambda_n (f, \phi_n) \psi_n.$$

Since U is a partial isometry, the family $\{\psi_n\}$ need not be complete. $\square$

4.4.2. Trace Class Operators.

Definition 4.40. Let $\{\phi_n\}_{n=1}^{\infty}$ be an orthonormal basis for $\mathcal{H}$. If A is any positive linear operator, we define the trace of A, $tr(A)$, by

$$tr(A) = \sum_{n=1}^{\infty} (A\phi_n, \phi_n),$$

whenever the sum converges.

Theorem 4.41. *The number $tr(A)$ is independent of the orthonormal basis chosen and:*

(1) *If $\lambda \geq 0$, then $tr(\lambda A) = \lambda tr(A)$.*

(2) *$tr(A + B) = tr(A) + tr(B)$.*

(3) *If U is a unitary operator, then $tr(UAU^{-1}) = tr(A)$.*

(4) *If $0 \leq A \leq B$, then $tr(A) \leq tr(B)$.*

Proof. If $\{\psi_n\}_{n=1}^{\infty}$ is any other basis, then

$$\sum_{k=1}^{\infty}(A\varphi_k,\varphi_k)=\sum_{k=1}^{\infty}\left(A^{1/2}\varphi_k,A^{1/2}\varphi_k\right)=\sum_{k=1}^{\infty}\left\|A^{1/2}\varphi_k\right\|^2$$
$$=\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\left|\left(A^{1/2}\varphi_k,\psi_j\right)\right|^2$$
$$=\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\left|\left(\varphi_k,A^{1/2}\psi_j\right)\right|^2=\sum_{j=1}^{\infty}\left\|A^{1/2}\psi_j\right\|^2$$
$$=\sum_{j=1}^{\infty}(A\psi_j,\psi_j).$$

The proofs of (1), (2), and (4) are easy. To prove (3), note that $U^*=U^{-1}$ and $\{U^*\varphi_n\}_{n=1}^{\infty}$ is also an orthonormal basis, so that

$$tr\left(UAU^{-1}\right)=\sum_{k=1}^{\infty}(UAU^*\varphi_k,\varphi_k)=\sum_{k=1}^{\infty}(AU^*\varphi_k,U^*\varphi_k)$$
$$=\sum_{k=1}^{\infty}(A\psi_k,\psi_k)=tr\,(A)\,.$$

□

Definition 4.42. An operator $A\in L[\mathcal{H}]$ is called trace class if and only if $tr(|A|)<\infty$. We denote the family of all trace class operators by $\mathbb{S}_1[\mathcal{H}]$.

Let $\{\lambda_n\}$ be the singular values of A, $\lambda_n=(\mu_n)^{1/2}$, where $\{\mu_n\}$ are the eigenvalues of A^*A.

Theorem 4.43. *If A is compact, then $A\in\mathbb{S}_1[\mathcal{H}]$ if and only if $\sum_{n=1}^{\infty}\lambda_n<\infty$, where $\{\lambda_n\}$ are the singular values of A. If $A\in\mathbb{S}_1[\mathcal{H}]$, then A is compact.*

Proof. First, if $A\in\mathbb{S}_1[\mathcal{H}]$ is compact, then $|A|$ is also compact. Thus, by Theorem 4.39, there exists an orthonormal family of functions $\{\phi_n\}$, such that

$$|A|=\sum_{n=1}^{\infty}\lambda_n(\cdot,\phi_n)\phi_n,$$

where $\{\lambda_n\}$ are the eigenvalues of $|A|$. It follows that

$$tr(|A|) = \sum_{n=1}^{\infty} (|A|\,\phi_n, \phi_n), \quad |A| = \sum_{k=1}^{\infty} \lambda_k\,(\cdot, \phi_k)\,\phi_k \Rightarrow$$
$$tr(|A|) = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \lambda_k\,(\phi_k, \phi_n) = \sum_{n=1}^{\infty} \lambda_n.$$

Since $tr(|A|) = \sum_{n=1}^{\infty} \lambda_n < \infty$. This proves the first part.

To prove the second part, assume that $A \in \mathbb{S}_1[\mathcal{H}]$. It follows that $|A|^2 \in \mathbb{S}_1[\mathcal{H}]$, so that $tr(|A|^2) < \infty$. Since $\|A\phi\| = \||A|\,\phi\|$, we have that

$$tr(|A|^2) = \sum\nolimits_{n=1}^{\infty} \left(|A|^2\,\phi_n, \phi_n\right) = \sum\nolimits_{n=1}^{\infty} \||A|\,\phi_n\|^2 < \infty,$$

for any orthonormal basis $\{\phi_n\}$. For each n, let M be the span of $[\phi_1, \ldots, \phi_n]$ and let $\psi \in M^{\perp}$, with $\|\psi\| = 1$. Since $[\phi_1, \ldots, \phi_n, \psi]$ can always be extended to an orthonormal basis, we have that

$$\|A\psi\|^2 \leqslant tr(|A|^2) - \sum\nolimits_{k=1}^{n} \|A\phi_k\|^2.$$

Thus,

$$\sup\left\{\|A\psi\| \,\middle|\, \psi \in [\phi_1, \cdots, \phi_n]^{\perp},\ \|\psi\| = 1\right\} \to 0, \quad n \to \infty,$$

so that $\sum_{k=1}^{n} (\cdot, \phi_k)\,A\phi_k$ converges in norm to A. It follows that A is compact. □

We leave the proof of the next theorem for the reader.

Theorem 4.44. *The family of trace class operators $\mathbb{S}_1[\mathcal{H}]$, with the norm of A defined by $\|A\|_1 = tr(|A|)$, is a Banach space (subspace of $L[\mathcal{H}]$). Furthermore*

(1) $\mathbb{F}_{\mathcal{H}} \subset \mathbb{S}_1[\mathcal{H}]$ *is dense.*
(2) *If $B \in L[\mathcal{H}]$ and $A \in \mathbb{S}_1[\mathcal{H}]$, then AB and $BA \in \mathbb{S}_1[\mathcal{H}]$.*
(3) *If $A \in \mathbb{S}_1[\mathcal{H}]$, then $A^* \in \mathbb{S}_1[\mathcal{H}]$.*

4.4.3. The Schatten Class.

Definition 4.45. For $1 \le p \le \infty$, let

$$\mathbb{S}_p[\mathcal{H}] = \{A \in \mathbb{K}_{\mathcal{H}} : \sum\nolimits_{k=1}^{\infty} \lambda_k^p = tr(|A|^p) < \infty\}.$$

The spaces $\mathbb{S}_p[\mathcal{H}]$ are known as the Schatten class of compact operators. They are the noncommutative analog of the L^p spaces. Proofs of the following can be found in Schatten [SC].

Theorem 4.46. *The family of operators $\mathbb{S}_p[\mathcal{H}]$ is a Banach space with the norm of A defined by $\|A\|_p = [tr(|A|^p)]^{\frac{1}{p}}, 1 \leq p < \infty$. When $p = \infty$, the norm of A defined by $\|A\|_\infty = \|A\|$ (the operator norm of $L[\mathcal{H}]$). Furthermore*

1. *$\mathbb{F}_{\mathcal{H}} \subset \mathbb{S}_p[\mathcal{H}], 1 \leq p \leq \infty$, is dense.*
2. *If $B \in L[\mathcal{H}]$ and $A \in \mathbb{S}_p[\mathcal{H}]$, then AB and $BA \in \mathbb{S}_p[\mathcal{H}]$ (i.e., it is a two-sided ideal).*
3. *If $A \in \mathbb{S}_p[\mathcal{H}]$, then $A^* \in \mathbb{S}_p[\mathcal{H}]$(i.e., it is a two-sided *-ideal).*
4. *If $1 < p < \infty$, then $\mathbb{S}_p[\mathcal{H}]$ is reflexive and its dual space is $\mathbb{S}_q[\mathcal{H}]$, where $\frac{1}{p} + \frac{1}{q} = 1$.*
5. *The dual space of $\mathbb{S}_1[\mathcal{H}]$ is $\mathbb{S}_\infty[\mathcal{H}] = \mathbb{K}_{\mathcal{H}}$.*

In case $p = 2$, $\mathbb{S}_2[\mathcal{H}]$ is also known as the Hilbert–Schmidt class of operators. Since the nonzero singular values of A, A^* and $|A|$ are identical, we have

$$\|A\|_1 = \|A^*\|_1 = \||A|\|_1 = \left\||A|^{1/2}\right\|_2^2.$$

Theorem 4.47. *We can define an inner product on $\mathbb{S}_2[\mathcal{H}]$ by $\langle A, B\rangle = tr(A^*B)$. In addition, we have that:*

1. *$\mathbb{S}_1[\mathcal{H}] \subset \mathbb{S}_2[\mathcal{H}]$ and $\|A\|_2 \leq \|A\|_1$.*
2. *If $A, B \in \mathbb{S}_2[\mathcal{H}]$, then $AB \in \mathbb{S}_1[\mathcal{H}]$ and $\|AB\|_1 \leq \|A\|_2 \|B\|_2$.*
3. *If $T \in \mathbb{S}_1[\mathcal{H}]$, then there exists $A, B \in \mathbb{S}_2[\mathcal{H}]$, with $T = AB$.*

4.5. Spectral Theory

In this section we provide an elementary version of the spectral theorem. Our objective is to provide a simple proof that, in a well-defined sense, every bounded linear operator on $\mathcal{H}$ has a spectral type representation.

To begin, recall that a projection P is a bounded linear operator on $\mathcal{H}$ with $P^2 = P$.

Theorem 4.48. *If P is a projection on $\mathcal{H}$, then the following are equivalent:*

(1) *P is self-adjoint, $P = P^*$*
(2) *P is normal, $PP^* = P^*P$*
(3) $\mathcal{R}(P) = N(P)^{\perp}$
(4) $(Pf, f) = \|Pf\|^2$*, for all* $f \in \mathcal{H}$

Proof. Clearly, (1) ⇒ (2). To show that (2) ⇒ (3), since P is normal,

$$(P^*Pf, f) = (PP^*f, f) \Leftrightarrow (Pf, Pf) = (P^*f, P^*f) \Leftrightarrow$$
$$N(P) = N(P^*) = \mathcal{R}(P)^{\perp}.$$

Since $N(P)$ is closed, $N(P)^{\perp} = \mathcal{R}(P)$ and $\mathcal{H} = N(P) \oplus \mathcal{R}(P)$.

To show that (3) ⇒ (4), for each $f \in \mathcal{H}$ we can write $f = f_1 + f_2$, with $f_1 \in N(P)$, $f_2 \in \mathcal{R}(P)$ and $(f_1, f_2) = 0$. It follows that $Pf = f_2$ and

$$(Pf, f) = (f_2, f_1 + f_2) = (f_2, f_2) = \|f_2\|^2 = \|Pf\|^2 .$$

To see that (4) ⇒ (1), note that $(Pf, f) = \|Pf\|^2 \Rightarrow (f, P^*f) = \|P^*f\|^2$ and $(f, P^*f) = (f_1 + f_2, f_2) = (f_2, f_2) = (Pf, f)$, so that $\|P^*f\|^2 = \|Pf\|^2$. □

Definition 4.49. Let $(\Omega, \mathfrak{B}[\Omega], \mu)$ be a measure space, with $\Omega \subset \mathbb{C}$. A resolution of the identity on $\mathfrak{B}[\Omega]$ is a mapping of $E : \mathfrak{B}[\Omega] \to \mathcal{C}[\mathcal{H}]$ such that, for each $B \in \mathfrak{B}[\Omega]$, $E(B)$ is a self-adjoint projection with the following properties:

(1) $E(\emptyset) = 0, \quad E(\Omega) = 1.$
(2) $E(B \cap B') = E(B)E(B').$
(3) $B \cap B' = \emptyset \Rightarrow E(B \cup B') = E(B) + E(B').$
(4) For each $f, g \in D(E)$, the set function $E_{f,g}(B) = (E(B)f, g)$ is a regular (complex) Borel measure on $\mathfrak{B}[\Omega]$.

From Theorem 4.48, we have the following results:

Theorem 4.50. *Let E is a resolution of the identity on $(\Omega, \mathfrak{B}[\Omega], \mu)$, then:*

(1) *For each $B \in \mathfrak{B}[\Omega]$ and $f \in \mathcal{H}$, we have,*

$$E_{f,f}(B) = (E(B)f, f)_{\mathcal{H}} = (E^2(B)f, f)_{\mathcal{H}} = \|E(B)f\|^2_{\mathcal{H}} .$$

It follows that each $E_{f,f}(\cdot)$ is a positive measure on $\mathfrak{B}[\Omega]$ and, since $E(\Omega) = 1$, its total variation $E_{f,f}(\Omega) = \|E_{f,f}\| = \|f\|^2$.

(2) $E(B')E(B) = E(B')E(B)$.

(3) $B \cap B' = \emptyset \Rightarrow \mathcal{R}[E(B)] \perp \mathcal{R}[E(B')]$.

(4) $E_{f,g}(\cdot)$ *is finitely additive.*

We need the following result on convergence of orthogonal vectors.

Lemma 4.51. *Let $\{\phi_n\}$ be a pairwise sequence of orthogonal vectors in $\mathcal{H}$. The sequence converges weakly if and only if it converges strongly.*

Proof. It is clear that strong convergence implies weak convergence, so we need to only prove the other direction. Suppose that

$$\mathbf{J}_N(\psi) = \sum_{k=1}^{N} (\psi, \varphi_k)$$

converges as $N \to \infty$ for every $\psi \in \mathcal{H}$, then the family $\mathbf{J}_N(\cdot)$ is bounded, by the Uniform Boundedness Theorem. Since

$$\|\mathbf{J}_N(\cdot)\| = \left\| \sum_{k=1}^{N} \varphi_k \right\| = \left\{ \sum_{k=1}^{N} \|\varphi_k\|^2 \right\}^{1/2},$$

we see that the family $\{\phi_n\}$ converges strongly. □

Theorem 4.52. *Let E is a resolution of the identity on $(\Omega, \mathfrak{B}[\Omega], \mu)$.*

(1) *If $f \in \mathcal{H}$, then for each $B \in \mathfrak{B}[\Omega]$, $B \to E(B)f$ is an $\mathcal{H}$-valued measure (countably additive) on $\mathfrak{B}[\Omega]$.*

(2) *If $E(B_k) = 0$ for all k and $B = \bigcup_{k=1}^{\infty} B_k$, then $E(B) = 0$.*

Proof. To prove (1), suppose $\{B_n\}$ is a pairwise sequence of disjoint sets in $\mathfrak{B}[\Omega]$. This implies that $B_n \cap B_m = \emptyset$, when $m \neq n$, so that $E(B_n \cap B_m)f = E(B_n)fE(B_m)f = 0$. Thus, the family $\{E(B_n)f\}$ are orthogonal to each other. From the above lemma, we have that

$$\left(E\left[\bigcup_{k=1}^{\infty} B_k \right] f, g \right) = \sum_{k=1}^{\infty} (E(B_k)f, g),$$

and the series converges in the norm of $\mathcal{H}$.

To prove (2), if $E(B_n) = 0$ for each n, then $E_{f,f}(B_n) = 0$. If we set $B = \bigcup_{k=1}^{\infty} B_k$, then $E_{f,f}(B) = 0$. However,

$$\|E(B)f\|^2 = E_{f,f}(B),$$

so that $E(B) = 0$. □

4.5.1. Spectral Theorem for Self-adjoint Operators.

4.5.1.1. *The Bounded Case.* Let $A \in L[\mathcal{H}]$, with spectrum $\sigma(A)$.

Theorem 4.53. *If f is a continuous function on $\sigma(A)$, the map $f \to f(A)$ is an isometric isomorphism and:*

(1) *$(f+g)(A)=f(A)+g(A)$, $(fg)(A)=f(A)g(A)$.*

(2)
$$\|f(A)\| = \sup_{\lambda\in\sigma(A)} |f(\lambda)| \tag{4.21}$$

(3) *If A is self-adjoint and f is real-valued, then $f(A)$ is also self-adjoint and $\sigma[f(A)] = f(\sigma[A])$.*

Proof. From Theorem 4.18, we see that since (1) holds for polynomials, it also holds for the uniform limit of polynomials.

For (2), since $f(A)$ is the uniform limit of $p_n(A)$, where $\{p_n(A)\}$ is a family of polynomials. It follows that

$$\|f(A)\| = \lim_{n\to\infty} \sup_{\lambda\in\sigma(A)} |p_n(\lambda)| .$$

To prove (3), if A is self-adjoint and f is real-valued, then $(f(A)\phi, \psi) = (\phi, f(A)\psi)$. The second part follows from Theorem 4.18(2). □

Let A be self-adjoint, $\phi, \psi \in \mathcal{H}$ and define a linear functional $\ell_{\phi,\psi}$ on $\mathcal{H}$ by

$$\ell_{\phi,\psi}(f) = (f(A)\phi, \psi) .$$

Since f is continuous, by Corollary 4.10 there is a unique complex measure $\mu_{\phi,\psi}$ such that

$$\ell_{\phi,\psi}(f) = (f(A)\phi, \psi) = \int_{\sigma(A)} f(\lambda) d\mu_{\phi,\psi}(\lambda). \tag{4.22}$$

Theorem 4.54. *If $\mu_{\phi,\psi}$ is the measure defined by (4.22), then*

(1) *$\mu_{\phi,\psi}$ is conjugate bilinear in ϕ, ψ (i.e., $\mu_{\psi,\phi} = \bar{\mu}_{\phi,\psi}$).*

(2) *The measures $\mu_{\phi,\phi}$ are nonnegative.*

(3) *The total variation of* $\mu_{\phi,\psi}$, $BV_{\sigma(A)}(\mu_{\phi,\psi}) \leq \|\phi\|\,\|\psi\|$.

Proof. The proofs of (1) and (2) are easy, while (3) follows from the boundedness of ℓ and the Schwartz inequality. □

If A is self-adjoint and $\Omega = \sigma(A)$, from Theorem 4.50 we see that, for each $B \in \mathfrak{B}(\Omega)$, $\mu_{\phi,\psi}(B)$ is a bounded conjugate bilinear form. From the Lax–Milgram Theorem, there is a bounded operator-valued function $E(B)$, such that

$$\mu_{\phi,\psi}(B) = (E(B)\phi, \psi).$$

The following properties may be easily checked:

Theorem 4.55. *For each* $B \in \mathfrak{B}(\Omega)$, *we have*

(1) $E(B)$ *is self-adjoint,* $E(B)^* = E(B)$.

(2) $\|E(B)\| \leq 1$

(3) *For each* $B \in \mathfrak{B}(\Omega)$, $AE(B) = E(B)A$.

(4) $E(\emptyset) = 0$, $E(\sigma(A)) = 1$.

(5) *If* $B_1 \cap B_2 = \emptyset$, *then* $E[B_1 \cup B_2] = E[B_1] + E[B_2]$ *and* $R(B_1) \perp R(B_2)$.

(6) $E[B_1]E[B_2] = E[B_2]E[B_1]$.

(7) *Each* $E(B)$ *is a orthogonal projection-value measure.*

From Theorems 4.54, 4.55, and Eq. (4.22), we see that:

$$\phi = \int_{\sigma(A)} dE(\lambda)\phi \quad \textit{and } A\phi = \int_{\sigma(A)} \lambda dE(\lambda)\phi. \tag{4.23}$$

The second term of Eq. (4.23) is the spectral representation for a bounded self-adjoint linear operator.

4.5.2. Bounded Deformed Spectral Measure. In this section we take a slightly different approach to the spectral question by noting that any bounded linear operator can be represented as a partial isometry times a self-adjoint operator.

Definition 4.56. Let E be a spectral measure. If U is a partial isometry, we call $E_U = UE$ the deformed spectral measure associated with U.

We now have the following interesting result.

Theorem 4.57. *If $A \in L[\mathcal{H}]$, then there is a partial isometry U and a deformed spectral measure E_U such that*

$$A\phi = \int_0^{r_T} \lambda dE_U(\lambda)\phi, \tag{4.24}$$

where $T = [A^*A]^{1/2}$.

Proof. By the polar decomposition, we can write $A = UT$, where U is a partial isometry and $T = [A^*A]^{1/2}$ is the nonnegative and self-adjoint. This means that $\sigma(T) \subset [0, r_T]$. It follows that, for each $\phi \in \mathcal{H}$, we can represent $T\phi$ as

$$T\phi = \int_0^{r_T} \lambda dE(\lambda)\phi.$$

Since $E(\lambda)\phi$ is a positive vector-valued function of bounded variation and U is a partial isometry, $E_U(\lambda)\phi = UE(\lambda)\phi$ is of bounded variation, with $Var(E_U\phi, \mathbb{R}) \leq Var(E\phi, \mathbb{R})$. Thus, by Theorem 4.7, we have

$$A\phi = UT\phi = \int_0^{r_T} \lambda dUE(\lambda)\phi = \int_0^{r_T} \lambda dE_U(\lambda)\phi.$$

□

Remark 4.58. Let A be self-adjoint, with its spectrum on the negative real axis. In this case, the standard spectral theorem gives us:

$$A = \int_{-r_A}^{0} \lambda dE(\lambda). \tag{4.25}$$

However, the deformed spectral theorem gives

$$A = \int_0^{r_A} \lambda dE_U(\lambda). \tag{4.26}$$

Note that the actual spectrum is not in the interval $[0, r_A]$. Thus, even for self-adjoint operators, the deformed spectral theorem provides a distinct representation. Furthermore, the region of integration is always on the real axis, even if the actual spectrum is in the complex plane.

4.5.2.1. *The Unbounded Case.* We now relax the condition that $A \in L[\mathcal{H}]$ and allow A to be unbounded. In this case, we need a few additional results.

Theorem 4.59. *Let A be any self-adjoint operator in $\mathcal{C}[\mathcal{H}]$. If $\lambda \in \mathbb{C}$, with $Im(\lambda) \neq 0$, then*

(1) *The resolvent* $R(\lambda, A)$ *exists and for each* $f \in D(A)$,

$$\|(\lambda I - A)\, f\| \geqslant |\mathrm{Im}(\lambda)|\, \|f\|\, .$$

(2) *The spectrum of* A, $\sigma(A) \subset \mathbb{R}$, $R(\lambda, A)$ *is normal with* $R(\lambda, A)^* = R(\bar{\lambda}, A)$ *and*

$$\|R(\lambda, A)\| \leqslant \frac{1}{|\mathrm{Im}(\lambda)|}.$$

(3) *If* $-A$ *is positive then for each* $\lambda > 0$, $R(\lambda, A)$ *is self-adjoint and*

$$\|R(\lambda, A)\| \leqslant \frac{1}{\lambda}.$$

(4) *If* $-A$ *is positive, then for* $\lambda > 0$ *and each* $f \in D(A)$,

$$\lim_{\lambda \to \infty} A\lambda R(\lambda, A) f = Af.$$

(5) *If* $-A$ *is positive, then for* $0 < \lambda < \infty$, *the operator*

$$A_\lambda = A\lambda R(\lambda, A)$$

is bounded and self-adjoint.

Proof. To prove (1), if $f \in D(A) = D(\lambda I - A)$, we have

$$\begin{aligned}
&\|(\lambda I - A)\, f\|^2 = ((\lambda I - A)\, f, (\lambda I - A)\, f) \\
&= (\mathrm{Im}(\lambda) f, \mathrm{Im}(\lambda) f) + ((\mathrm{Re}(\lambda) I - A)\, f, (\mathrm{Re}(\lambda) I - A)\, f) \\
&\geqslant (\mathrm{Im}(\lambda) f, \mathrm{Im}(\lambda) f) = \|\mathrm{Im}(\lambda) f\|^2 = |\mathrm{Im}(\lambda)|^2\, \|f\|^2\, .
\end{aligned}$$

To prove (2), it follows from (1) that for $Im(\lambda) \neq 0$ we have

$$\|\mathrm{R}(\lambda, A)\| \leqslant |\mathrm{Im}(\lambda)|^{-1}\, .$$

Since $A = A^*$ is closed and densely defined, we see that $(\lambda I - A)$ is closed and densely defined. From (1) we see that $R(\lambda, A)$ is bounded and that $(\lambda I - A)$ is injective. Thus,

$$N((\lambda I - A)) = \{0\} = N((\lambda I - A)^*).$$

Since $(\lambda I - A)^* = \left(\bar{\lambda} I - A\right)$ and $Im(\bar{\lambda}) \neq 0$, we see that λ, $\bar{\lambda} \in \rho(A)$. It is clear that $R(\lambda, A)$ is normal, $R(\lambda, A)^* = R(\bar{\lambda}, A)$, and $\sigma(A) \subset \mathbb{R}$.

The proof of (3) is like that of (1). If $-A$ is positive and $\lambda > 0$, then

$$\|(\lambda I - A)\, f\|^2 = (\lambda f, \lambda f) + (-Af, -Af) \geqslant |\lambda|^2\, \|f\|^2\, .$$

It follows that $\left\|(\lambda I - A)^{-1}\right\| \leq \lambda^{-1}$.

We prove (4) in two parts. First, if $0 < \lambda \in \rho(A)$ and $f \in D(A)$ we have

$$\begin{aligned}
&R(\lambda, A)\,(\lambda I - A)\,f = f \;\Rightarrow \\
&\lambda R(\lambda, A)f - f = R(\lambda, A)Af \;\Rightarrow \\
&\|\lambda R(\lambda, A)f - f\| \leqslant \|R(\lambda, A)\|\,\|Af\| \leqslant \lambda^{-1}\,\|Af\|
\end{aligned}$$

This last term converges to zero as $\lambda \to \infty$, so that

$$\lim_{\lambda\to\infty} \lambda R(\lambda, A)f = f.$$

Since $D(A)$ is dense, the convergence holds for all $f \in \mathcal{H}$.

For the second part, we see from the last result that

$$\lim_{\lambda\to\infty} \lambda AR(\lambda, A)f = \lim_{\lambda\to\infty} \lambda R(\lambda, A)Af = Af,$$

whenever $f \in D(A)$.

To prove (5), we note that for $-A \geq 0$ and fixed $0 < \lambda < \infty$, we have

$$A_\lambda = \lambda AR(\lambda, A) = \lambda^2 R(\lambda, A) - \lambda I.$$

From here, we see that A_λ is self-adjoint and $\|A_\lambda\| \leqslant 2\lambda$. □

Let $A \in \mathcal{C}[\mathcal{H}]$ be a self-adjoint operator and, without lost, we assume that $-A$ nonnegative, so that its spectrum $\sigma(A) \subset \mathbb{R}^-$.

Theorem 4.60. *There exists a unique regular countably additive projection-valued spectral measure $E(\Omega)$ defined on the Borel sets of $\mathbb{R}$, vanishing on the compliment of the spectrum of A such that, for each $\phi \in D(A)$, we have:*

(1) *$D(A)$ also satisfies*

$$D(A) = \left\{ \phi \in \mathcal{H} \mid \int_{-\infty}^{0} \lambda^2\,(dE(\lambda)\phi, \phi)_{\mathcal{H}} < \infty \right\}$$

and

(2)

$$A\phi = \int_{-\infty}^{0} \lambda dE(\lambda)\phi, \text{ for } \phi \in D(A).$$

(3) *If $g(\cdot)$ is a complex-valued Borel function defined (a.e.) on $\mathbb{R}$, then $g(A) \in \mathcal{C}[\mathcal{H}]$ and, for $\phi \in D(g(A)) = D_g(A)$,*

$$g(A)\phi = \int_{-\infty}^{0} g(\lambda)dE(\lambda)\phi,$$

where

$$D_g(A) = \left\{ \phi \in \mathcal{H} \mid \int_{\sigma(A)} |g(\lambda)|^2 \left(dE(\lambda)\phi, \phi\right)_{\mathcal{H}} < \infty \right\}$$

and $g(A^*) = \bar{g}(A)$.

Proof. From Theorem 4.60, we see that A_μ is a bounded self-adjoint operator, for all $\mu > 0$. Thus, by Theorems 4.54, and 4.55, there is a self-adjoint projection-valued measure $E_\mu(\cdot)$, which vanishes on $\mathbb{R}^+$ such that

$$A_\mu \phi = \int_{-\infty}^{0} \lambda dE_\mu(\lambda)\phi.$$

If we let $\mu = n$, we see that $(A_n\phi, \phi) \to (A\phi, \phi)$ for every $\phi \in D(A)$. Thus,

$$\int_{-\infty}^{0} \lambda \left(dE_n(\lambda)\phi, \phi\right) \to \int_{-\infty}^{0} \lambda \left(dE(\lambda)\phi, \phi\right).$$

Where, for $\phi \in D(A)$ and each $B \in \mathfrak{B}(\mathbb{R}^-)$, $(E(B)\phi, \phi)$ is a measure. Furthermore, from

$$\left(E^2(\lambda)\phi, \phi\right) = \lim_{n\to\infty} \left(E_n^2(\lambda)\phi, \phi\right) = \lim_{n\to\infty} \left(E_n(\lambda)\phi, \phi\right) = \left(E(\lambda)\phi, \phi\right),$$

we see that $E(\cdot)$ is a projection. Since both A_n and A are self-adjoint, we see that $E(\cdot)$ is also. The properties of $E(\cdot)$ now follows from those of $E_n(\cdot)$ with Theorems 4.50 and 4.53. Thus, $E(\cdot)$ is a projection-value measure on $\mathfrak{B}(\mathbb{R}^-)$ and

$$A\phi = \int_{-\infty}^{0} \lambda dE(\lambda)\phi, \text{ for all } \phi \in D(A).$$

Since

$$\left\| \int_{-\infty}^{0} \lambda dE(\lambda)\phi \right\|^2 = \|A\phi\|^2 = (A\phi, A\phi) = \left(A^2\phi, \phi\right) = \int_{-\infty}^{0} \lambda^2 d\left(E(\lambda)\phi, \phi\right),$$

we see that $D(A)$ can also be represented as

$$D(A) = \left\{ \phi \,\middle|\, \int_{-\infty}^{0} \lambda^2 \left(dE(\lambda)\phi, \phi\right) < \infty \right\}.$$

For (3), let $g(\lambda)$ be a complex-valued Borel function defined almost everywhere on $\mathbb{R}^-$. Let

$$g_n(\lambda) = \begin{cases} g(\lambda), & |g(\lambda)| \leqslant n \\ 0 & |g(\lambda)| > n, \end{cases}$$

set

$$D[g(A)] = \left\{ \phi \middle| \lim_{n\to\infty} g_n(A)\phi \quad \text{exists} \right\}$$

and

$$\lim_{n\to\infty} g_n(A)\phi = g(A)\phi.$$

If $B_n = \{\lambda : g(\lambda) \le n\}$, then for $\phi \in D[g(A)]$,

$$\|g(A)\phi\|^2 = \lim_{n\to\infty} \int_{B_n} |g(\lambda)|^2 (dE(\lambda)\phi, \phi) = \int_{-\infty}^{0} |g(\lambda)|^2 (dE(\lambda)\phi, \phi).$$

□

The proof of our next result now follows from the polar decomposition and Theorem 4.61(1) Corollary 4.10.

Theorem 4.61. *If $A \in \mathcal{C}[\mathcal{H}]$, there exists a unique regular countably additive projection-valued deformed spectral measure $E_U(\cdot)$ defined on the Borel sets of $\mathbb{R}^+$, vanishing on the compliment of the spectrum of A such that, for each $\phi \in D(A)$, we have:*

(1) *$D(A)$ also satisfies*

$$D(A) = \left\{ \phi \in \mathcal{H} \mid \int_0^\infty \lambda^2 (dE_U(\lambda)\phi, \phi)_{\mathcal{H}} < \infty \right\}$$

and

(2)

$$A\phi = \int_0^\infty \lambda dE_U(\lambda)\phi, \text{ for } \phi \in D(A).$$

Proof. To prove (1), write $A = UT$, where U is the unique partial isometry and $T = [A^*A]^{1/2}$. By Theorem 4.61, there is a positive spectral measure $E(\cdot)$ such that, for each $x \in D(A) = D(T)$:

$$T\phi = \int_0^\infty \lambda dE(\lambda)\phi. \tag{4.27}$$

Since $E(\lambda)\phi$ is a positive vector-valued function of bounded variation and U is a partial isometry, $E_U(\lambda)\phi = UE(\lambda)\phi$ is of bounded variation, with $Var(E_U\phi, \mathbb{R}) \le Var(E\phi, \mathbb{R})$. Thus, by Theorem 4.8,

$$U \int_0^\infty \lambda dE(\lambda)\phi = \int_0^\infty \lambda dUE(\lambda)\phi.$$

Since $A\phi = UT\phi$, if we set $E_U(\lambda)\phi = UE(\lambda)\phi$, we have from Eq. (4.27),

$$A\phi = \int_0^\infty \lambda dE_U(\lambda)\phi. \tag{4.28}$$

□

Remark 4.62. In general, $Ug(T) \neq g(UT)$, so that a similar result for $g(A)$ with $A \in \mathbb{C}[\mathcal{H}]$ and $A = UT$ does not hold. For example, if we look at $g(A) = A^2$, we see that $g(T) = T^2$, while

$$Ug(T) = UT^2 = AT \neq A^2$$

References

[HP] E. Hille, R.S. Phillips, *Functional Analysis and Semigroups.* American Mathematical Society Colloquium Publications, vol. 31 (American Mathematical Society, Providence, RI, 1957)

[L1] P. Lax, *Functional Analysis* (Wiley-Interscience, New York, 2002)

[RS] M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic, New York, 1972)

[RO] H.L. Royden, *Real Analysis*, 2nd edn. (Macmillan Press, New York, 1968)

[R] W. Rudin, *Functional Analysis* (McGraw-Hill, New York, 1973)

[R1] W. Rudin, *Real Complex Analysis*, 3rd edn. (McGraw-Hill, New York, 1987)

[SC] R. Schatten, *Norm Ideas of Completely Continuous Operators* (Springer, New York, 1960)

[VN1] J. von Neumann, Über adjungierte Funktionaloperatoren. Ann. Math. **33**, 294–310 (1932)

[YS] K. Yosida, *Functional Analysis*, 2nd edn. (Springer, New York, 1968)

Chapter 5

Operators on Banach Space

The Feynman operator calculus and the Feynman path integral develop naturally on Hilbert space. In this chapter we develop the theory of semigroups of operators, which is the central tool for both. In order to extend the theory to other areas of interest, we begin with a new approach to operator theory on Banach spaces. We first show that the structure of the bounded linear operators on Banach space with an S-basis is much closer to that for the same operators on Hilbert space. We will exploit this new relationship to transfer the theory of semigroups of operators developed for Hilbert spaces to Banach spaces. The results are complete for uniformly convex Banach spaces, so we restrict our presentation to that case, with one exception. In the Appendix (Sect. 5.3), we show that all of the results in Chap. 4 have natural analogues for uniformly convex Banach spaces.

5.1. Preliminaries

Let $\mathcal{B}$ be a uniformly convex Banach space with an S-basis. Let $\mathcal{C}[\mathcal{B}]$ be the set of closed densely defined linear operators and let $L[\mathcal{B}]$ be the set of bounded linear operators on $\mathcal{B}$.

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_5

Definition 5.1. A duality map $\mathcal{J} : \mathcal{B} \mapsto \mathcal{B}'$ is a set

$$\mathcal{J}(u) = \left\{ u^* \in \mathcal{B}' \,\middle|\, \langle u, u^* \rangle = \|u\|_{\mathcal{B}}^2 = \|u^*\|_{\mathcal{B}'}^2 \right\}, \ \forall u \in \mathcal{B}.$$

Example 5.2. *Let* Ω *be a bounded open subset of* $\mathbb{R}^n$, $n \in \mathbb{N}$. *If* $u \in L^p[\Omega] = \mathcal{B}$, $1 < p < \infty$, *then*

$$\mathcal{J}(u)(x) = \|u\|_p^{2-p} |u(x)|^{p-2} u(x) = u^* \in L^q[\Omega], \ \tfrac{1}{p} + \tfrac{1}{q} = 1. \tag{5.1}$$

Furthermore,

$$\langle u, u^* \rangle = \|u\|_p^{2-p} \int_\Omega |u(x)|^p \, d\lambda_n(x) = \|u\|_p^2 = \|u^*\|_q^2$$

It can be shown that $L^p[\Omega]$ is uniformly convex and that $u^* = \mathcal{J}(u)$ is uniquely defined for each $u \in \mathcal{B}$. Thus, if $\{u_n\}$ is an S-basis for $L^p[\Omega]$, then the family vectors $\{u_n^*\}$ is an S-basis for $L^q[\Omega] = (L^p[\Omega])'$. The relationship between u and u^* is nonlinear [see Eq. (5.1)]. In the next section we prove the remarkable result that there is another representation of $\mathcal{B}'$, with $u^* = \mathbf{J}_{\mathcal{B}}(u)$ linear, for each $u \in \mathcal{B}$. (However, u^* is no longer a duality mapping.)

5.1.1. The Natural Hilbert Space for a Uniformly Convex Banach Space. We follow the same ideas used in Chap. 3 to embed L^2 in KS^2. However, we take a restricted approach that applies to all uniformly convex Banach spaces with an S-basis. Fix $\mathcal{B}$ and let $\{\mathcal{E}_n\}$ be an S-basis for $\mathcal{B}$. For each n, let $t_n = 2^{-n}$ and for each $\mathcal{E}_n$, let $\mathcal{E}_n^*$ be the corresponding dual vector in $\mathcal{B}'$. For each pair of functions u, v on $\mathcal{B}$, define an inner product by:

$$(u, v) = \sum_{n=1}^{\infty} t_n \langle \mathcal{E}_n^*, u \rangle \overline{\langle \mathcal{E}_n^*, v \rangle}.$$

we let $\mathcal{H}$ be the completion of $\mathcal{B}$ in the induced norm. It is clear that $\mathcal{B} \subset \mathcal{H}$ densely and

$$\begin{aligned} \|u\|_{\mathcal{H}} &= \left[\sum_{n=1}^{\infty} t_n |\langle \mathcal{E}_n^*, u \rangle|^2 \right]^{1/2} \\ &\leqslant \sup_n |\langle \mathcal{E}_n^*, u \rangle| \\ &\leqslant \sup_{\|\mathcal{E}^*\|_{\mathcal{B}'} \leqslant 1} |\langle \mathcal{E}^*, u \rangle| = \|u\|_{\mathcal{B}}, \end{aligned} \tag{5.2}$$

so that the embedding is both dense and continuous. It is clear that $\mathcal{H}$ is unique up to a change of S-basis.

Definition 5.3. If $\mathcal{B}$ be a Banach space, we say that $\mathcal{B}'$ has a Hilbert space representation if there exists a Hilbert space $\mathcal{H}$, with $\mathcal{B} \subset \mathcal{H}$ as a continuous dense embedding and for each $u^* \in \mathcal{B}'$, $u^* = (\cdot, u)_{\mathcal{H}}$ for some $u \in \mathcal{B}$.

Theorem 5.4. *If $\mathcal{B}$ be a uniformly convex Banach space with an S-basis, then $\mathcal{B}'$ has a Hilbert space representation.*

Proof. Let $\mathcal{H}$ be the natural Hilbert space for $\mathcal{B}$ and let $\mathbf{J}$ be the natural linear mapping from $\mathcal{H} \to \mathcal{H}'$, defined by

$$\langle v, \mathbf{J}(u) \rangle = (v, u)_{\mathcal{H}}, \text{ for all } u, v \in \mathcal{H}.$$

It is easy to see that $\mathbf{J}$ is bijective and $\mathbf{J}^* = \mathbf{J}$. First, we note that the restriction of $\mathbf{J}$ to $\mathcal{B}$, $\mathbf{J}_{\mathcal{B}}$, maps $\mathcal{B}$ to a unique subset of linear functionals $\{\mathbf{J}_{\mathcal{B}}(u),\ u \in \mathcal{B}\}$ and, $\mathbf{J}_{\mathcal{B}}(u+v) = \mathbf{J}_{\mathcal{B}}(u) + \mathbf{J}_{\mathcal{B}}(v)$, for each $u, v \in \mathcal{B}$. We are done if we can prove that $\{\mathbf{J}_{\mathcal{B}}(u),\ u \in \mathcal{B}\} = \mathcal{B}'$. For this, it suffices to show that $\mathbf{J}_{\mathcal{B}}(u)$ is bounded for each $u \in \mathcal{B}$. Since $\mathcal{B}$ is dense in $\mathcal{H}$, from equation (5.2) we have:

$$\|\mathbf{J}_{\mathcal{B}}(u)\|_{\mathcal{B}'} = \sup_{v \in \mathcal{B}} \frac{\langle v, \mathbf{J}_{\mathcal{B}}(u) \rangle}{\|v\|_{\mathcal{B}}} \leqslant \sup_{v \in \mathcal{B}} \frac{\langle v, \mathbf{J}_{\mathcal{B}}(u) \rangle}{\|v\|_{\mathcal{H}}} = \|u\|_{\mathcal{H}} \leqslant \|u\|_{\mathcal{B}}.$$

Thus, $\{\mathbf{J}_{\mathcal{B}}(u),\ u \in \mathcal{B}\} \subset \mathcal{B}'$. Since $\mathcal{B}$ is uniformly convex, there is a (unique) one-to-one relationship between $\mathcal{B}$ and $\mathcal{B}'$, so that $\{\mathbf{J}_{\mathcal{B}}(u),\ u \in \mathcal{B}\} = \mathcal{B}'$. □

5.1.2. Construction of the Adjoint on $\mathcal{B}$. We can now show that if $\mathcal{B}'$ has a Hilbert space representation, then each closed densely linear operator on $\mathcal{B}$ has a natural adjoint defined on $\mathcal{B}$.

Theorem 5.5. *Let $\mathcal{B}$ be a uniformly convex Banach space with an S-basis. If $\mathbb{C}[\mathcal{B}]$ denotes the closed densely linear operators on $\mathcal{B}$ and $L[\mathcal{B}]$ denotes the bounded linear operators, then every $A \in \mathcal{C}[\mathcal{B}]$ has a well-defined adjoint $A^* \in \mathcal{C}[\mathcal{B}]$. Furthermore, if $A \in L[\mathcal{B}]$, then $A^* \in L[\mathcal{B}]$ with:*

(1) $(aA)^* = \bar{a}A^*$,

(2) $A^{**} = A$,

(3) $(A^* + B^*) = A^* + B^*$

(4) $(AB)^* = B^*A^*$ *and*

(5) $\|A^*A\|_{\mathcal{B}} \leq \|A\|_{\mathcal{B}}^2$.

Thus, $L[\mathcal{B}]$ is a *algebra.

Proof. Let $\mathbf{J}$ be the natural linear mapping from $\mathcal{H} \to \mathcal{H}'$ and let $\mathbf{J}_{\mathcal{B}}$ be the restriction of $\mathbf{J}$ to $\mathcal{B}$. If $A \in \mathcal{C}[\mathcal{B}]$, then $A'\mathbf{J}_{\mathcal{B}} : \mathcal{B}' \to \mathcal{B}'$. Since A' is closed and densely defined, it follows that $\mathbf{J}_{\mathcal{B}}^{-1}A'\mathbf{J}_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}$ is a closed and densely defined linear operator. We define $A^* = [\mathbf{J}_{\mathcal{B}}^{-1}A'\mathbf{J}_{\mathcal{B}}] \in \mathcal{C}[\mathcal{B}]$. If $A \in L[\mathcal{B}]$, $A^* = \mathbf{J}_{\mathcal{B}}^{-1}A'\mathbf{J}_{\mathcal{B}}$ is defined on all of $\mathcal{B}$. By the Closed Graph Theorem, $A^* \in L[\mathcal{B}]$. The proofs of (1)–(3) are straightforward. To prove (4),

$$\begin{aligned}(BA)^* &= \mathbf{J}_{\mathcal{B}}^{-1}(BA)'\mathbf{J}_{\mathcal{B}} = \mathbf{J}_{\mathcal{B}}^{-1}A'B'\mathbf{J}_{\mathcal{B}} \\ &= \left[\mathbf{J}_{\mathcal{B}}^{-1}A'\mathbf{J}_{\mathcal{B}}\right]\left[\mathbf{J}_{\mathcal{B}}^{-1}B'\mathbf{J}_{\mathcal{B}}\right] = A^*B^*.\end{aligned} \tag{5.3}$$

If we replace B by A^* in Eq. (5.3), noting that $A^{**} = A$, we also see that $(A^*A)^* = A^*A$. To prove (5), we first see that:

$$\langle A^*Av, \mathbf{J}_{\mathcal{B}}(u)\rangle = (A^*Av, u)_{\mathcal{H}} = (v, A^*Au)_{\mathcal{H}},$$

so that A^*A is symmetric. Thus, by Lax's Theorem, A^*A has a bounded extension to $\mathcal{H}$ and $\|A^*A\|_{\mathcal{H}} \leqslant k\,\|A^*A\|_{\mathcal{B}}$, where k is a positive constant. We also have that

$$\|A^*A\|_{\mathcal{B}} \leqslant \|A^*\|_{\mathcal{B}}\|A\|_{\mathcal{B}} \leqslant \|A\|_{\mathcal{B}}^2 . \tag{5.4}$$

It follows that $\|A^*A\|_{\mathcal{B}} \leq \|A\|_{\mathcal{B}}^2$. If equality holds in (5.4), for all $A \in L[\mathcal{B}]$, then it is a C^*-algebra. This is true if and only if $\mathcal{B}$ is a Hilbert space. Thus, in general the inequality in (5.4) is strict. □

5.1.2.1. **Example: Differential Operators**. Let A be a closed densely defined linear operator defined on $L^p[\mathbb{R}^n]$, $1 < p < \infty$, and let A' be the dual defined on $L^q[\mathbb{R}^n]$, $\frac{1}{p} + \frac{1}{q} = 1$. It is easy to show that if A' is densely defined on $L^p[\mathbb{R}^n]$, it has a closed extension to $L^p[\mathbb{R}^n]$ (without using $\mathcal{H}_2 = KS^2[\mathbb{R}^n]$).

Example 5.6. *Let A be a second order differential operator on $L^p[\mathbb{R}^n]$ of the form*

$$A = \sum_{i,j=1}^{n} a_{ij}(\mathbf{x})\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i,j=1}^{n} x_i b_{ij}(\mathbf{x})\frac{\partial}{\partial x_j},$$

where $\mathbf{a}(\mathbf{x}) = [\![a_{ij}(\mathbf{x})]\!]$ and $\mathbf{b}(\mathbf{x}) = [\![b_{ij}(\mathbf{x})]\!]$ are matrix-valued functions in $\mathbb{C}_c^\infty[\mathbb{R}^n \times \mathbb{R}^n]$ (infinitely differentiable functions with compact support). We also assume that for all $\mathbf{x} \in \mathbb{R}^n$ det $[\![a_{ij}(\mathbf{x})]\!] > \varepsilon$ and the imaginary part of the eigenvalues of $\mathbf{b}(\mathbf{x})$ are bounded above by $-\varepsilon$, for some $\varepsilon > 0$. Note, since we don't require $\mathbf{a}$ or $\mathbf{b}$ to be symmetric, $A \neq A'$.

It is well known that $\mathbb{C}_c^\infty[\mathbb{R}^n] \subset L^p[\mathbb{R}^n] \cap L^q[\mathbb{R}^n]$ *for all* $1 < p \leq q < \infty$. *Furthermore, since* A' *is invariant on* $\mathbb{C}_c^\infty[\mathbb{R}^n]$,

$$A' : \mathbb{C}_c^\infty\left[\mathbb{R}^n\right] \subset L^p\left[\mathbb{R}^n\right] \to \mathbb{C}_c^\infty\left[\mathbb{R}^n\right] \subset L^p\left[\mathbb{R}^n\right].$$

It follows that A' *has a closed extension to* $L^p[\mathbb{R}^n]$. *(In this case, we do not need* $\mathcal{H}_2$ *directly, we can identify* $\mathbf{J}_2$ *with the identity on* $\mathcal{H}_2$ *and* A^* *with* A'.*)*

Remark 5.7. For a general A, which is closed and densely defined on $L^p[\mathbb{R}^n]$, we know that it is densely defined on $KS^2[\mathbb{R}^n]$. Thus, it has a well-defined adjoint A^* on $KS^2[\mathbb{R}^n]$. By Theorem 5.5, we can take the restriction of A^* from $KS^2[\mathbb{R}^n]$ to obtain our adjoint on $L^p[\mathbb{R}^n]$.

5.1.2.2. **Example: Integral Operators**. In one dimension, the Hilbert transform can be defined on $L^2[\mathbb{R}]$ via its Fourier transform:

$$\widehat{H(f)} = -i \operatorname{sgn} x \, \hat{f}.$$

It can also be defined directly as principal-value integral:

$$(Hf)(x) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{|x-y| \geqslant \varepsilon} \frac{f(y)}{x-y} dy.$$

For a proof of the following results see Grafakos [GRA, Chap. 4].

Theorem 5.8. *The Hilbert transform on* $L^2[\mathbb{R}]$ *satisfies:*

(1) H *is an isometry,* $\|H(f)\|_2 = \|f\|_2$ *and* $H^* = -H$.
(2) *For* $f \in L^p[\mathbb{R}]$, $1 < p < \infty$, *there exists a constant* $C_p > 0$ *such that,*

$$\|H(f)\|_p \leq C_p \|f\|_p. \tag{5.5}$$

The next result is technically obvious, but conceptually nontrivial.

Corollary 5.9. *The adjoint of* H, H^* *defines a bounded linear operator on* $L^p[\mathbb{R}]$ *for* $1 < p < \infty$, *and* H^* *satisfies Eq. (5.5) for the same constant* C_p.

The Riesz transform, $\mathbf{R}$, is the n-dimensional analogue of the Hilbert transform and its jth component is defined for $f \in L^p[\mathbb{R}^n]$, $1 < p < \infty$, by:

$$R_j(f) = c_n \lim_{\varepsilon \to 0} \int_{|\mathbf{y}-\mathbf{x}| \geqslant \varepsilon} \frac{y_j - x_j}{|\mathbf{y} - \mathbf{x}|^{n+1}} f(\mathbf{y}) d\mathbf{y}, \quad c_n = \frac{\Gamma\left(\frac{N+1}{2}\right)}{\pi^{(n+1)/2}}.$$

Definition 5.10. Let Ω be defined on the unit sphere S^{n-1} in R^n.

(1) The function $\Omega(x)$ is said to be homogeneous of degree n if $\Omega(tx) = t^n \Omega(x)$.

(2) The function $\Omega(x)$ is said to have the cancellation property if

$\int_{S^{n-1}} \Omega(\mathbf{y}) d\sigma(\mathbf{y}) = 0$, where $d\sigma$ is the induced Lebesgue measure on S^{n-1}.

(3) The function $\Omega(x)$ is said to have the Dini-type condition if

$$\sup_{\substack{|\mathbf{x}-\mathbf{y}|\leqslant\delta \\ |\mathbf{x}|=|\mathbf{y}|=1}} |\Omega(\mathbf{x}) - \Omega(\mathbf{y})| \leqslant \omega(\delta) \Rightarrow \quad \int_0^1 \frac{\omega(\delta)d\delta}{\delta} < \infty.$$

A proof of the following theorem can be found in Stein [STE] (see p. 39).

Theorem 5.11. *Suppose that* Ω *is homogeneous of degree* 0*, satisfying both the cancellation property and the Dini-type condition. If* $f \in L^p[\mathbb{R}^n]$, $1 < p < \infty$ *and*

$$T_\varepsilon(f)(\mathbf{x}) = \int_{|\mathbf{y}-\mathbf{x}|\geqslant\varepsilon} \frac{\Omega(\mathbf{y}-\mathbf{x})}{|\mathbf{y}-\mathbf{x}|^n} f(\mathbf{y}) d\mathbf{y}.$$

Then

(1) *There exists a constant* A_p*, independent of both* f *and* ε *such that*

$$\|T_\varepsilon(f)\|_p \leqslant A_p \|f\|_p.$$

(2) *Furthermore,* $\lim_{\varepsilon\to 0} T_\varepsilon(f) = T(f)$ *exists in the* L^p *norm and*

$$\|T(f)\|_p \leqslant A_p \|f\|_p. \tag{5.6}$$

Treating $T_\varepsilon(f)$ as a special case of the Henstock–Kurzweil integral, conditions (1) and (2) are automatically satisfied and we can write the integral as

$$T(f)(\mathbf{x}) = \int_{\mathbb{R}^n} \frac{\Omega(\mathbf{y}-\mathbf{x})}{|\mathbf{y}-\mathbf{x}|^n} f(\mathbf{y}) d\mathbf{y}.$$

For $g \in L^q$, $\frac{1}{p} + \frac{1}{q} = 1$, we have $\langle T(f), g\rangle = \langle f, T^*(g)\rangle$. Using Fubini's Theorem for the Henstock–Kurzweil integral (see [HS]), we have that

Corollary 5.12. *The adjoint of* T, $T^* = -T$ *is defined on* L^p *and satisfies Eq. (5.6)*

It is easy to see that the Riesz transform is a special case of the above theorem and corollary.

Another closely related integral operator is the Riesz potential, $I_\alpha(f)(\mathbf{x}) = (-\Delta)^{-\alpha/2} f(\mathbf{x}),\ 0 < \alpha < n$, is defined on $L^p[\mathbb{R}^n],\ 1 < p < \infty$, by (see Stein [STE], p. 117):

$$I_\alpha(f)(\mathbf{x}) = \gamma^{-1}(\alpha) \int_{\mathbb{R}^n} \frac{f(\mathbf{y})d\mathbf{y}}{|\mathbf{x}-\mathbf{y}|^{n-\alpha}}, \quad \text{and } \gamma(\alpha) = 2^\alpha \pi^{\frac{n}{2}} \frac{\Gamma(\frac{\alpha}{2})}{\Gamma(\frac{n-\alpha}{2})}.$$

Since the kernel is symmetric, application of Fubini's Theorem shows that the adjoint $I_\alpha^* = I_\alpha$ is also defined on $L^p[\mathbb{R}^n]$. Since $(-\Delta)^{-1}$ is not bounded, we cannot obtain L^p bounds for $I_\alpha(f)(\mathbf{x})$. However, if $1/q = 1/p - \alpha/n$, we have the following (see Stein [STE], p. 119)

Theorem 5.13. *If $f \in L^p[\mathbb{R}^n]$ and $0 < \alpha < n,\ 1 < p < q < \infty,\ 1/q = 1/p - \alpha/n$, then the integral defining $I_\alpha(f)$ converges absolutely for almost all $\mathbf{x}$. Furthermore, there is a constant $A_{p,q}$, such that*

$$\|I_\alpha(f)\|_q \leqslant A_{p,q} \|f\|_p. \tag{5.7}$$

5.1.3. Extension of the Adjoint. In this section we discuss an extension of the adjoint for a Banach space $\mathcal{B}$, which need not be uniformly convex. If $\mathcal{B}$ is not uniformly convex, Theorem 5.5 no longer holds and we need $\mathcal{H}_1$. The next theorem shows that, for A bounded, we can always define a reasonable version of the adjoint A^*, which has many of the essential properties that we find for a Hilbert space.

Theorem 5.14. *Let A be a bounded linear operator on $\mathcal{B}$. Then A has a well-defined adjoint A^* defined on $\mathcal{B}$ such that:*

(1) *the operator $A^*A \geq 0$ (accretive),*

(2) *$(A^*A)^* = A^*A$ (naturally self-adjoint), and*

(3) *$I + A^*A$ has a bounded inverse.*

Proof. For $i = 1, 2$, let $\mathbf{J}_i : \mathcal{H}_i \to \mathcal{H}_i'$. As in Theorem 5.5, $\mathbf{J}_i^* = \mathbf{J}_i$. Now, let $A_1 = A_{|\mathcal{H}_1} : \mathcal{H}_1 \to \mathcal{H}_2$, and $A'_1 : \mathcal{H}_2' \to \mathcal{H}_1'$.

It follows that $A'_1\mathbf{J}_2 : \mathcal{H}_2 \to \mathcal{H}_1'$ and $\mathbf{J}_1^{-1}A'_1\mathbf{J}_2 : \mathcal{H}_2 \to \mathcal{H}_1 \subset \mathcal{B}$ so that, if we define $A^* = [\mathbf{J}_1^{-1}A'_1\mathbf{J}_2]_{\mathcal{B}}$, then $A^* : \mathcal{B} \to \mathcal{B}$ (i.e., $A^* \in L[\mathcal{B}]$).

To prove (1), let $g \in \mathcal{B}$, then $(A^*Ag, g)_{\mathcal{H}_2} \geq 0$ for all $g \in \mathcal{B}$. Hence $\langle A^*Ag, g^* \rangle \geq 0$ for all $g^* \in J(g)$ (the duality map of g), so that A^*A is accretive.

To prove (2), we have for $g \in \mathcal{H}_1$,

$$\begin{aligned}(A^*A)^*g &= (\{\mathbf{J}_1^{-1}[\{[\mathbf{J}_1^{-1}A'{}_1\mathbf{J}_2]|_{\mathcal{B}}A\}_1]'\mathbf{J}_2\}|_{\mathcal{B}})g \\ &= (\{\mathbf{J}_1^{-1}[\{A'{}_1[\mathbf{J}_2A_1\mathbf{J}_1^{-1}]|_{\mathcal{B}}\}]\mathbf{J}_2\}|_{\mathcal{B}})g \\ &= A^*Ag.\end{aligned}$$

It follows that the same result holds on all of $\mathcal{B}$.

The proof of (3), that $I + A^*A$ is invertible, follows the same lines as in von Neumann's theorem. □

Since A^*A is self-adjoint on $\mathcal{B}$ (in the sense of (2) above), it is natural to expect that the same is true on $\mathcal{H}_2$. However, this need not be the case. To obtain a simple counterexample, recall that, in standard notation, the simplest class of bounded linear operators on $\mathcal{B}$ is $\mathcal{B} \otimes \mathcal{B}'$, in the sense that:

$$\mathcal{B} \otimes \mathcal{B}' : \mathcal{B} \to \mathcal{B}, \text{ by } Au = (b \otimes l_{b'}(\cdot))u = \left\langle b', u \right\rangle b.$$

Thus, if $l_{b'}(\cdot) \in \mathcal{B}' \backslash \mathcal{H}_2'$, then $J_2\{J_1^{-1}[(A_1)']J_2|_{\mathcal{B}}(u)\}$ is not in $\mathcal{H}_2'$, so that A^*A is not defined as an operator on all of $\mathcal{H}_2$ and thus cannot have a bounded extension.

We now provide the correct extension of Lax's Theorem.

Theorem 5.15. *Let A be a bounded linear operator on $\mathcal{B}$. If $\mathcal{B}' \subset \mathcal{H}_2$, then A has a bounded extension to $L[\mathcal{H}_2]$, with $\|A\|_{\mathcal{H}_2} \le k\|A\|_{\mathcal{B}}$ (for some positive k).*

Proof. We first note that if $g, h \in \mathcal{B}$, then $\mathbf{J}_1^{-1}\mathbf{J}_2(g) = g$ and $(A_1')'h = Ah$. Now let $T = A^*A$, then

$$\begin{aligned}&(Tg, h)_{\mathcal{H}_2} = \langle Tg, \mathbf{J}_2(h)\rangle \\ &= \langle A^*Ag, \mathbf{J}_2(h)\rangle = \left\langle \mathbf{J}_1^{-1}A_1'\mathbf{J}_2(Ag), \mathbf{J}_2(h)\right\rangle \\ &= \left\langle A_1'\mathbf{J}_2(Ag), h\right\rangle = \left\langle \mathbf{J}_2(Ag), (A_1')'h\right\rangle \\ &= \langle Ag, \mathbf{J}_2(Ah)\rangle = \left\langle g, (A_1')\mathbf{J}_2(Ah)\right\rangle \\ &= \left\langle \mathbf{J}_1^{-1}\mathbf{J}_2(g), (A_1')\mathbf{J}_2(Ah)\right\rangle = \left\langle \mathbf{J}_2(g), \mathbf{J}_1^{-1}(A_1')\mathbf{J}_2(Ah)\right\rangle \\ &= (g, Th)_{\mathcal{H}_2}\end{aligned}$$

We can now apply Lax's Theorem to see that, for some k, $\|T\|_{\mathcal{H}_2} = \|A\|_{\mathcal{H}_2}^2 \le k^2\|A\|_{\mathcal{B}}^2$. □

Remark 5.16. Thus, the algebra $L[\mathcal{B}]$ also has a *operation for all Banach spaces with an S-basis and $\mathcal{B}' \subset \mathcal{H}_2$. However, if $\mathcal{B}$ is not uniformly convex and $A \neq B$, B' then, unless

$$(AB\,|_{\mathcal{H}_1})' = (B\,|_{\mathcal{H}_1})'\,(A\,|_{\mathcal{H}_1})', \quad (AB)^* \neq A^*B^*.$$

A natural question is "which Banach spaces with an S-basis have the property that, $\mathcal{B}' \subset \mathcal{H}_2$"? This question has no general answer. However, if $\mathcal{B}$ is one of the following classical Banach spaces and $\mathcal{H}_2 = KS^2[\mathbb{R}^n]$, then $\mathcal{B}' \subset \mathcal{H}_2$ ($\mathcal{H}_1 = GS^2[\mathbb{R}^n]$). A few of the spaces below are not separable (do not have an S-basis).

(1) $\mathbb{C}_b[\mathbb{R}^n]$, the bounded continuous functions on $\mathbb{R}^n$.
(2) $\mathbb{C}_u[\mathbb{R}^n]$, the bounded uniformly continuous functions on $\mathbb{R}^n$.
(3) $\mathbb{C}_0^k[\mathbb{R}^n]$, the continuous functions on $\mathbb{R}^n$, with k derivatives that vanish at infinity.
(4) $L^p[\mathbb{R}^n]$, $1 \leq p \leq \infty$, the Lebesgue integrable functions on $\mathbb{R}^n$ of order p.
(5) $\mathfrak{M}[\mathbb{R}^n]$, the space of finitely additive set functions (measures) on $\mathbb{R}^n$.

We note that both $\mathbb{C}_b[\mathbb{R}^n]$ and $L^\infty[\mathbb{R}^n]$ are nonseparable Banach spaces, with the same dual space $\mathfrak{M}[\mathbb{R}^n] \subset KS^2[\mathbb{R}^n]$ and, the dual space of $\mathbb{C}_u[\mathbb{R}^n]$, $\mathbb{C}_u'[\mathbb{R}^n] \subset \mathfrak{M}[\mathbb{R}^n] \subset KS^2[\mathbb{R}^n]$. In each case, we can use Theorem 5.15.

5.2. Semigroups of Operators

Introduction. Semigroups of operators form the basis for both the Feynman operator calculus and path integral theory of Chaps. 7 and 8. We have restricted our presentation to those aspects that are absolutely necessary and should even be reviewed those with some training in the subject. We provide all of the basic results along with proofs, for those without prior background.

The theory of semigroups of operators is a fairly mature field of study, which has continued to attract the interest of those in analysis, probability theory, partial differential equations, dynamical systems, and quantum theory, in addition to the many areas of applied mathematics. This continued interest is expected because of the simple (conceptual) framework provided, the robustness of the technical methodology, and the wealth of problems and new applications.

Those interested in the finer details are encouraged to seek out the wealth of interesting material by consulting some of the major works in the field. See the standards by Hille and Phillips [HP], Yosida [YS], Kato [K], Pazy [PZ], Goldstein [GS] and the recent ones by Engel and Nagel [EN] and Vrabie [VR]. The book by Vrabie [VR] offers a number of new and interesting applications.

We develop most of the theory for a fixed separable Hilbert space $\mathcal{H}$ over $\mathbb{C}$ and will assume when convenient that $\mathcal{H} = KS^2[\mathbb{R}^n]$. However, we begin with the general theory on a Banach space $\mathcal{B}$.

Definition 5.17. A family of linear operators $\{S(t), 0 \le t < \infty\}$ (not necessarily bounded), defined on $\mathcal{D} \subset \mathcal{B}$, is a semigroup if

(1) $S(t+s)f = S(t)S(s)f$ for $f \in \mathcal{D}$, the domain of the semigroup.
(2) The semigroup is said to be strongly continuous if $\lim_{\tau \to 0} S(t + \tau)f = S(t)f$ for all $f \in \mathcal{D},\ t > 0$.
(3) It is said to be a C_0-semigroup if it is strongly continuous, $S(0) = I,\ \mathcal{D} = \mathcal{B}$ and $\lim_{t \to 0} S(t)f = f$ for all $f \in \mathcal{B}$.
(4) $S(t)$ is a C_0-contraction semigroup if $\|S(t)\|_{\mathcal{B}} \leqslant 1$.
(5) $S(t)$ is a C_0-unitary group if $S^*(t)$ exists and $S(t)S(t)^* = S(t)^*S(t) = I$, and $\|S(t)\|_{\mathcal{B}} = 1$.

Definition 5.18. For a C_0-semigroup $S(t)$, the linear operator A defined by

$$D(A) = \left\{ f \in \mathcal{B} \left| \lim_{t \downarrow 0} \tfrac{1}{t}[S(t)f - f] \text{ exists} \right. \right\} \quad \textit{and}$$

$$Af = \lim_{t \downarrow 0} \tfrac{1}{t}[S(t)f - f] = \left. \frac{d^+ S(t)f}{dt} \right|_{t=0} \quad \text{for} \quad f \in D(A)$$

is the infinitesimal generator of the semigroup $S(t)$ and $D(A)$ is the domain of A.

Lemma 5.19. *Let $S(t)$ be a C_0-semigroup. Then there exist constants $\omega \ge 0$ and $M \ge 1$ such that:*

$$\|S(t)\|_{\mathcal{H}} \leqslant Me^{\omega t}, \quad \text{for} \quad 0 \leqslant t < \infty.$$

Proof. If $\|S(t)\|_{\mathcal{B}}$ is not bounded in any interval $0 \le t \le m,\ m > 0$, then there is a nonnegative sequence t_n such that $\lim_{n\to\infty} t_n = 0$ and $\|S(t_n)\|_{\mathcal{B}} \geqslant n$. By the uniform boundedness theorem it follows that, for some $f,\ S(t)f$ is unbounded. But then $S(t)$ is not strongly continuous (see (3) above). Thus $\|S(t)\|_{\mathcal{B}} \leqslant M$ for $0 \leq t \leq m$. From $\|S(0)\|_{\mathcal{B}} = 1$ and $M \geqslant 1$, we can choose $\omega = m^{-1} log M$. Let $t \geq 0$ be given, then $t = nm + \delta$, where $0 \leq \delta < m$, so, by the semigroup property of $S(t)$, we have:

$$\|S(t)\|_{\mathcal{B}} = \|S(\delta)S(m)^n\|_{\mathcal{H}} \leqslant M^{n+1} \leqslant MM^{t/m} = Me^{\omega t}.$$

□

Theorem 5.20. *Let $S(t)$ be a C_0 contraction semigroup and let A be its infinitesimal generator. Then*

(1) *For all $f \in \mathcal{B}$, we have*

$$\lim_{h \to 0} \tfrac{1}{h} \int_t^{t+h} S(u) f du = S(t) f.$$

(2) *For all $f \in \mathcal{B}$, $\int_0^t S(u) f du \in D(A)$ and,*

$$A \int_0^t S(u) f du = S(t) f - f.$$

(3) *For all $f \in D(A)$,*

$$\frac{d}{dt} S(t) f = AS(t) f = S(t) A f.$$

(4) *For all $f \in D(A)$,*

$$S(t) f - S(u) f = \int_u^t AS(\tau) f d\tau = \int_u^t S(\tau) A f d\tau.$$

(5) *A is closed and $\overline{D(A)} = \mathcal{B}$.*

(6) *The resolvent set $\rho(A)$ of A contains R^+ and, for every $\lambda > 0$,*

$$\|R(\lambda, A)\|_{\mathcal{B}} \leqslant \frac{1}{\lambda}.$$

Proof. The proof of (1) follows from the strong continuity of $S(t)$. To prove (2), let $f \in \mathcal{B}$ and suppose that $h > 0$. Then

$$\frac{S(h) - I}{h} \int_0^t S(u) f du = \tfrac{1}{h} \int_0^t (S(u+h) f - S(u) f) du$$
$$= \tfrac{1}{h} \int_t^{t+h} S(u) f du - \tfrac{1}{h} \int_0^h S(u) f du$$

and, as $h \searrow 0$, the right-hand side tends to $S(t)f - f$. To prove (3), if $f \in D(A)$ and $h > 0$, we have

$$\frac{S(h)-I}{h} S(t)f = S(t)\left(\frac{S(h)-I}{h}\right) f \xrightarrow{h \to 0} S(t)Af.$$

It follows that $S(t)f \in D(A)$ and $S(t)Af = AS(t)f$. This also means that

$$\frac{d^+}{dt} S(t)f = AS(t)f = S(t)Af.$$

To complete our proof, we need to show that, for $t > 0$, the left-hand derivative exists and is equal to $S(t)Af$. To prove this, note that

$$\begin{aligned}
&\lim_{h \searrow 0}\left[\frac{S(t)f - S(t-h)f}{h} - S(t)Af\right] \\
&= \lim_{h \searrow 0} S(t-h)\left(\frac{S(h)f - f}{h} - Af\right) + \lim_{h \searrow 0}\left(S(t-h)Af - S(t)Af\right).
\end{aligned}$$

We are done since the limit of both terms on the right is zero. To prove (4), we need to only look at the integral of $\frac{d}{dt}S(t)f = AS(t)f = S(t)Af$. To prove (5), for each $f \in \mathcal{B}$ set $f_h = \frac{1}{h}\int_0^h S(u)f du$. By (2), $f_h \in D(A)$ and, by (1), $f_h \to f$, so that $\overline{D(A)} = \mathcal{B}$. To prove that A is closed, let $f_n \in D(A)$, $f_n \to f$ and $Af_n \to g$ (as $n \to \infty$). From (4), we have that

$$S(t)f_n - f_n = \int_0^t S(u)Af_n du \to S(t)f - f = \int_0^t S(u)Ag du.$$

If we divide the last integral by t and let $t \searrow 0$, we see from (1) that $f \in D(A)$ and $Af = g$. The proof of (6) requires a little additional work. If $f \in \mathcal{H}$ and $\lambda > 0$, define a bounded linear operator $R(\lambda, A)$ by (the Laplace transform of $S(t)$):

$$R(\lambda, A)f = \int_0^\infty e^{-\lambda t} S(t) f dt.$$

Since the function $t \to S(t)f$ is continuous and uniformly bounded, the integral exists and provides a well-defined linear operator with

$$\|R(\lambda, A)f\|_{\mathcal{B}} \leqslant \int_0^\infty e^{-\lambda t}\,\|S(t)f\|_{\mathcal{B}}\, dt \leqslant \tfrac{1}{\lambda}\,\|f\|_{\mathcal{B}}\,.$$

For $h > 0$,

$$\begin{aligned}
&\frac{S(h)-I}{h} R(\lambda, A)f = \tfrac{1}{h}\int_0^\infty e^{-\lambda t}\left(S(t-h)f - S(t)f\right) \\
&= \frac{e^{-\lambda h}-1}{h}\left(\int_0^\infty e^{-\lambda t} S(t) f dt\right) - \frac{e^{-\lambda h}}{h}\int_0^h e^{-\lambda t} S(t) f dt \underset{h \searrow 0}{\to} \lambda R(\lambda, A)f - f.
\end{aligned}$$

Thus, we see that, for every $\lambda > 0$ and $f \in \mathcal{B}$, $R(\lambda, A)f \in D(A)$ and $AR(\lambda, A)f = \lambda R(\lambda, A)f - f \Rightarrow (\lambda I - A)R(\lambda, A)f = f$. We also have that, for $f \in D(A)$,

$$\begin{aligned} R(\lambda, A)Af &= \int_0^\infty e^{-\lambda t} S(t)Af dt = \int_0^\infty e^{-\lambda t} AS(t)f dt \\ &= A\left[\int_0^\infty e^{-\lambda t} S(t)f dt\right] = AR(\lambda, A)f. \end{aligned}$$

It now follows that $R(\lambda, A)(\lambda I - A)f = f$ for each $f \in D(A)$, so that $R(\lambda, A)$ is the inverse of $(\lambda I - A)$ for all $\lambda > 0$ and

$$\|R(\lambda, A)f\|_{\mathcal{B}} \leqslant \tfrac{1}{\lambda} \|f\|_{\mathcal{B}}.$$

$\square$

Lemma 5.21. *Suppose that $R(\lambda, A) = (\lambda I - A)^{-1}$, where A is a linear operator such that:*

(1) *A is closed and $\overline{D(A)} = \mathcal{B}$.*

(2) *The resolvent set $\rho(A)$ of A contains R^+ and, for every $\lambda > 0$,*

$$\|R(\lambda, A)\|_{\mathcal{B}} \leqslant 1/\lambda.$$

Then $\lim_{\lambda \to \infty} \lambda R(\lambda, A)f = f$ for all $f \in \mathcal{B}$.

Proof. For each $f \in D(A)$, we have that

$$\|\lambda R(\lambda, A)f - f\|_{\mathcal{B}} = \|AR(\lambda, A)f\|_{\mathcal{B}} = \|R(\lambda, A)Af\|_{\mathcal{B}} \leqslant \tfrac{1}{\lambda} \|Af\|_{\mathcal{B}} \xrightarrow{\lambda \to \infty} 0.$$

Since $D(A)$ is dense and $\|\lambda R(\lambda, A)\|_{\mathcal{B}} \leqslant 1$, as $\lambda \to \infty$, $\lambda R(\lambda, A)f \to f$ for each $f \in \mathcal{B}$. $\square$

5.2.1. Hilbert Space. We now look at the case when $\mathcal{B} = \mathcal{H}$ is a Hilbert space.

Definition 5.22. For each $\lambda > 0$, we define the Yosida approximator by: $A_\lambda = \lambda AR(\lambda, A) = \lambda^2 R(\lambda, A) - \lambda I$.

The next result is due to Yosida and applies to generators of strongly continuous semigroups defined on $[0, \infty)$. We will prove a generalized version of the theorem, which applies to strongly continuous semigroups $(0, \infty)$.

Theorem 5.23. *(Yosida) Let A be a closed linear operator with $\overline{D(A)} = \mathcal{H}$. If the resolvent set $\rho(A)$ of A contains R^+ and, for every $\lambda > 0$, $\|R(\lambda, A)\|_{\mathcal{H}} \leqslant \lambda^{-1}$. Then*

(1) $\lim_{\lambda\to\infty} A_\lambda f = Af \quad \text{for } f \in D(A).$

(2) A_λ *is a bounded generator of a contraction semigroup and, for each* $f \in \mathcal{H}$, λ, $\mu > 0$, *we have:*

$$\left\|e^{tA_\lambda} f - e^{tA_\mu} f\right\|_{\mathcal{H}} \leqslant t \left\|A_\lambda f - A_\mu f\right\|_{\mathcal{H}}.$$

If all we know is that A is the generator of a strongly continuous semigroup $S(t) = \exp(tA)$ for $t > 0$, the above result is not enough. Unfortunately, for general strongly continuous semigroups, A may not have a bounded resolvent. The following (artificial example) shows what can (and will) happen in some real cases.

Example 5.24. *Let* $\mathcal{H} = \mathbf{H}_0(\mathbb{R}^n)$ *be the Hilbert space* (over $\mathbb{R}$) *of functions mapping* $\mathbb{R}^n$ *to itself, which vanish at infinity. Consider the Cauchy problem:*

$$\frac{d}{dt}\mathbf{u}(\mathbf{x}, t) = a\left|\mathbf{x}\right|\mathbf{u}(\mathbf{x}, t),\ \ \mathbf{u}(\mathbf{x}, 0) = \mathbf{f}(\mathbf{x}),$$

where $a = \prod_{i=1}^{n} sign(x_i)$. *Let* $S(t)\mathbf{f}(\mathbf{x}) = \mathrm{e}^{ta|\mathbf{x}|}\,\mathbf{f}(\mathbf{x})$, *where* $\mathbf{x} = [x_1, \cdots, x_n]^t$. *It is easy to see that* $S(t)$ *is a semigroup on* $\mathcal{H}$ *with generator* A *such that* $A\mathbf{f}(\mathbf{x}) = a\left|\mathbf{x}\right|\mathbf{f}(\mathbf{x})$. *It follows that* $u(\mathbf{x}, t) = S(t)\mathbf{f}(\mathbf{x})$ *solves the above initial-value problem. If we compute the resolvent, we get that:*

$$R(\lambda, A)\mathbf{f}(\mathbf{x}) = \int_0^\infty e^{-\lambda t}\exp\{-t\left|\mathbf{x}\right|\}\mathbf{f}(\mathbf{x})dt = \frac{1}{\lambda - a\left|\mathbf{x}\right|}\mathbf{f}(\mathbf{x}).$$

It is clear that the spectrum of A *is the real line, so that* $R(\lambda, A)$ *is an unbounded operator for all real* λ. *However, it can be checked that the bounded linear operator*

$$A_\lambda = a\lambda|\mathbf{x}|/[\lambda + |\mathbf{x}|]$$

converges strongly to A *(on* $D(A)$*) as* $\lambda \to \infty$, *and*

$$\lim_{\lambda\to 0} S_\lambda(t)\mathbf{f}(\mathbf{x}) = S(t)\mathbf{f}(\mathbf{x}).$$

As an application of the polar decomposition, the next result shows that the Yosida approach can be generalized in such a way as to give a contractive approximator for all strongly continuous semigroups of operators on $\mathcal{H}$.

For any closed densely defined linear operator A on $\mathcal{H}$, let $T = -[A^*A]^{1/2}$, $\bar{T} = -[AA^*]^{1/2}$. Since $T(\bar{T})$ is m-dissipative, it generates a contraction semigroup. We can now write A as $A = VT$, where $V = -U$ is the unique partial isometry of Chap. 4. Define A_λ by $A_\lambda = \lambda AR(\lambda, T)$. Note that $A_\lambda = \lambda UTR(\lambda, T) = \lambda^2 UR(\lambda, T) - \lambda U$ and, although A does not commute with $R(\lambda, T)$, we have $\lambda AR(\lambda, T) = \lambda R(\lambda, \bar{T})A$.

Theorem 5.25. *(Generalized Yosida) Let A be a closed densely defined linear operator on $\mathcal{H}$. Then*

(1) $A_\lambda = \lambda AR(\lambda, T)$ *is a bounded linear operator and* $\lim_{\lambda\to\infty} A_\lambda f = Af$, *for all* $f \in D(A)$,

(2) $\exp[tA_\lambda]$ *is a bounded contraction for* $t > 0$, *and*

(3) *if* $S(t) = \exp[tA]$ *is defined on* $\mathcal{D}$, $D(A) \subset \mathcal{D}$, *then for* $t > 0$, $f \in \mathcal{D}$, $\lim_{\lambda\to\infty} \|\exp[tA_\lambda]f - \exp[tA]f\|_{\mathcal{H}} = 0$.

Proof. To prove (1), let $f \in D(A)$. Now use the fact that

$$\lim_{\lambda\to\infty} \lambda R(\lambda, \bar{T})f = f$$

and $A_\lambda f = \lambda R(\lambda, \bar{T})Af$. To prove (2), use

$$A_\lambda = \lambda^2 UR(\lambda, T) - \lambda U$$

with $\|\lambda R(\lambda, T)\|_{\mathcal{H}} = 1$, and $\|U\|_{\mathcal{H}} = 1$ to get that

$$\|\exp[t\lambda^2 UR(\lambda, T) - t\lambda U]\|_{\mathcal{H}} \leq \exp[-t\lambda\|U\|_{\mathcal{H}}]\exp[t\lambda\|U\|_{\mathcal{H}}\|\lambda R(\lambda, T)\|_{\mathcal{H}}] \leq 1.$$

To prove (3), let $t > 0$ and $f \in D(A)$. Then

$$\begin{aligned} \|\exp[tA]f - \exp[tA_\lambda]f\|_{\mathcal{H}} &= \|\int_0^t \frac{d}{ds}[e^{(t-s)A_\lambda}e^{sA}]f ds\|_{\mathcal{H}} \\ &\leq \int_0^t \|[e^{(t-s)A_\lambda}(A - A_\lambda)e^{sA}f]\|_{\mathcal{H}} \\ &\leq \int_0^t \|[(A - A_\lambda)e^{sA}f]\|_{\mathcal{H}} ds. \end{aligned}$$

Now use

$$\|[A_\lambda e^{sA}f]\|_{\mathcal{H}} = \|[\lambda R(\lambda, \bar{T})e^{sA}Af]\|_{\mathcal{H}} \leq \|[e^{sA}Af]\|_{\mathcal{H}},$$

to get

$$\|[(A - A_\lambda)e^{sA}f]\|_{\mathcal{H}} \leq 2\|[e^{sA}Af]\|_{\mathcal{H}}.$$

Since $\|[e^{sA}Af]\|_{\mathcal{H}}$ is continuous, by the bounded convergence theorem we have

$$\lim_{\lambda\to\infty} \|\exp[tA]f - \exp[tA_\lambda]f\|_{\mathcal{H}} \leq \int_0^t \lim_{\lambda\to\infty} \|[(A - A_\lambda)e^{sA}f]\|_{\mathcal{H}} ds = 0.$$

Thus, $S(t)f$ exists and the convergence is uniform on bounded intervals for $t > 0$ and all $f \in D(A)$. Since $D(A)$ is dense in $\mathcal{D}$, $S(t)$ can be extended to all of $\mathcal{D}$. □

Remark 5.26. The first result (1) provides an independent proof that every closed densely defined linear operator on a Hilbert space is of first Baire class (may be approximated by bounded linear operators on its domain).

We now turn to the main theorem for semigroups of linear operators.

Theorem 5.27. (Hille–Yosida Theorem) *A linear operator A is the generator of a C_0-semigroup of contractions $S(t)$, $t \geq 0$, if and only if A is closed, densely defined, $\mathbb{R}^+ \subset \rho(A)$ and, for every $\lambda > 0$, $\|R(\lambda, A)\|_{\mathcal{H}} \leq \lambda^{-1}$.*

Proof. The necessity is shown in Theorem 5.23. To prove sufficiency, from Theorem 5.25, we see that, if A is closed and densely defined, with

$$\|R(\lambda, A)\|_{\mathcal{H}} \leq \frac{1}{\lambda}$$

for $\lambda > 0$, then, for $\mu > 0$ we have

$$\left\|e^{tA_\lambda} f - e^{tA_\mu} f\right\|_{\mathcal{H}} \leqslant t \|A_\lambda f - A_\mu f\|_{\mathcal{H}} \leq t \|A_\lambda f - Af\|_{\mathcal{H}} + t \|Af - A_\mu f\|_{\mathcal{H}}.$$

It follows that for $f \in D(A)$, $e^{tA_\lambda} f$ converges as $\lambda \to \infty$ and the convergence is uniform on bounded intervals. Since $\left\|e^{tA_\lambda} f\right\|_{\mathcal{H}} \leq 1$, it follows that $e^{tA_\lambda} f \to S(t)$ for every $f \in \mathcal{H}$. It is clear that $S(t)$ is a semigroup and that $\left\|e^{tA}\right\|_{\mathcal{H}} \leq 1$, with $S(0) = 1$. Thus, $S(t)$ is a C_0-semigroup, since it is strongly continuous. Finally,

$$e^{tA_\lambda} f - f = \int_0^t e^{sA_\lambda} A_\lambda f ds \to \int_0^t e^{sA} Af ds = e^{tA} f - f,$$

so that A is the generator. □

5.2.2. Lumer–Phillips Theory. We now discuss the characterization of an infinitesimal generator of a C_0-semigroup of contractions, due to Lumer and Phillips [LP].

Definition 5.28. Let A be a linear operator on $\mathcal{H}$. A is said to be dissipative if

$$\operatorname{Re} \langle Af, f \rangle \leqslant 0 \text{ for all } f \in D(A).$$

Theorem 5.29. (Lumer–Phillips) *Let A be a linear operator on $\mathcal{H}$; then*

(1) *A is dissipative if and only if*

$$\|(\lambda I - A) f\|_{\mathcal{H}} \geqslant \lambda \|f\|_{\mathcal{H}} \quad \text{for all } f \in D(A) \quad \text{and all } \lambda > 0.$$

(2) *If $D(A)$ is dense in $\mathcal{H}$ and there is a λ_0 such that $Ran(\lambda_0 I - A) = \mathcal{H}$, then A is the generator of a C_0 semigroup of contractions.*

(3) *If A is the generator of a C_0 semigroup of contractions on $\mathcal{H}$, then $Ran(\lambda I - A) = \mathcal{H}$ for all $\lambda > 0$ and A is dissipative.*

Remark 5.30. We note that (2) implies that A is m-dissipative, while (3) asserts that every generator of a contraction semigroup is m-dissipative.

Proof. To prove (1), let A be dissipative, $f \in D(A)$ and $\lambda > 0$. If $\operatorname{Re}\langle Af, f\rangle \leq 0$ then:

$$\|(\lambda I - A) f\| \, \|_{\mathcal{H}} f\| \geq |\langle(\lambda I - A) f, f\rangle| \geq \operatorname{Re}\langle(\lambda I - A) f, f\rangle \geq \lambda \|f\|_{\mathcal{H}}^2.$$

It follows that $\|(\lambda I - A) f\|_{\mathcal{H}} \geq \lambda \|f\|_{\mathcal{H}}$. Conversely, assume that $\lambda \|f\|_{\mathcal{H}} \leq \|(\lambda I - A) f\|_{\mathcal{H}}$ for $f \in D(A)$ and all $\lambda > 0$. If we square both sides, an easy calculation shows that

$$\|Af\|_{\mathcal{H}}^2 - 2\lambda \operatorname{Re}\langle Af, f\rangle \geq 0.$$

Since this is true for all $\lambda > 0$, we see that $\operatorname{Re}\langle Af, f\rangle \leq 0$. To prove (2), note that since A is dissipative we can use (1) for $\lambda > 0$ to get that $\|(\lambda I - A) f\|_{\mathcal{H}} \geqslant \lambda \|f\|_{\mathcal{H}}$ for all $f \in D(A)$. Since $Ran(\lambda_0 I - A) = \mathcal{H}$, with $\lambda = \lambda_0$, it follows that $(\lambda_0 I - A)^{-1}$ is a bounded linear operator. But this means that it is a closed operator, so that $(\lambda_0 I - A)$ and hence A is also a closed operator. Now note that if $Ran(\lambda I - A) = \mathcal{H}$ for every $\lambda > 0$, then $(0, \infty) \subset \rho(\lambda)$ and $\|R(\lambda, A)\|_{\mathcal{H}} \leq \lambda^{-1}$. It will then follow by Theorem 5.27 (Hille–Yosida) that A is the generator of a C_0 contraction semigroup. Thus, we need to show that $Ran(\lambda I - A) = \mathcal{H}$ for every $\lambda > 0$. Let

$$\Lambda = \{\, \lambda : 0 < \lambda < \infty\} \quad \text{and} \quad Ran(\lambda I - A) = \mathcal{H}.$$

If $\lambda \in \Lambda$, $\lambda \in \rho(\lambda)$. As $\rho(\lambda)$ is an open set, there is a nonempty neighborhood of $\lambda \subset \rho(\lambda)$. It follows that the intersection of this neighborhood with $\mathbb{R}$ is in Λ, so that Λ is an open set. If $\lambda_n \in \Lambda$, $\lambda_n \to \lambda > 0$, then, for every $g \in \mathcal{H}$, there exists a $f_n \in D(A)$ such that

$$\lambda_n f_n - A f_n = g. \tag{5.8}$$

Since A is dissipative, we have that $\|f_n\|_{\mathcal{H}} \leq \lambda_n^{-1} \|g\|_{\mathcal{H}} \leq C$ for some $C > 0$. We also have that:

$$\begin{aligned}\lambda_m \|f_n - f_m\|_{\mathcal{H}} &\leq \|\lambda_m (f_n - f_m) - A(f_n - f_m)\|_{\mathcal{H}} \\ &= |\lambda_n - \lambda_m| \|f_n\|_{\mathcal{H}} \leq C |\lambda_n - \lambda_m|,\end{aligned}$$

so that $\{f_n\}$ is a Cauchy sequence. If we let $f_n \to f$, we see from (5.5) that $Af_n \to \lambda f - g$. As A is closed, $f \in D(A)$ and $\lambda f - Af = g$. It follows that $Ran(\lambda I - A) = \mathcal{H}$ and $\lambda \in \Lambda$ so that Λ is also closed in $(0, \infty)$. Since $\lambda_0 \in \Lambda$, we see that $\Lambda \neq \emptyset$ and therefore $\Lambda = (0, \infty)$.

To prove (3), we first observe that if A is the generator of a C_0 contraction semigroup $S(t)$ on $\mathcal{H}$, then it is closed and densely defined. Furthermore, by Theorem 5.27 (Hille–Yosida), $(0, \infty) \subset \rho(A)$ and $Ran(\lambda I - A) = \mathcal{H}$ for all $\lambda > 0$. If $f \in D(A)$ then

$$|\langle S(t)f, f\rangle| \leqslant \|S(t)f\|_{\mathcal{H}} \|f\|_{\mathcal{H}} \leqslant \|f\|_{\mathcal{H}}^2$$

so that

$$\operatorname{Re} \langle S(t)f - f, f\rangle = \operatorname{Re} \langle S(t)f, f\rangle - \|f\|_{\mathcal{H}}^2 \leq 0.$$

If we divide the above equation by $t > 0$ and let $t \downarrow 0$, we get that:

$$\operatorname{Re} \langle Af, f\rangle \leqslant 0,$$

so that A is dissipative. □

The next result follows from the Lumer–Phillips Theorem (see Remark 5.30).

Theorem 5.31. *Suppose A is a densely defined m-dissipative operator. Then A is the generator of a C_0 semigroup $S(t)$ of contraction operators on $\mathcal{H}$.*

Theorem 5.32. *If A is closed and densely defined on $\mathcal{H}$, with both A and A^* dissipative, then A is m-dissipative.*

Proof. It suffices show that $Ran(I - A) = \mathcal{H}$. Since A is both closed and dissipative, $Ran(I - A)$ is closed in $\mathcal{H}$. If $Ran(I - A) \neq \mathcal{H}$ then there is a nonzero $g \in \mathcal{H}$ such that $(f - Af, g) = 0$ for all $f \in D(A)$. This implies that $(g, g - A^*g) = \|g\|^2 - (g, A^*g) = 0$, so that $g - A^*g = 0$. Since A^* is dissipative, from part (1) of Theorem 5.29 (Lumer–Phillips), we must have that $g = 0$. But this is a contradiction since we assumed that $g \neq 0$. □

We now consider an important class of operators which generates C_0-contractions. The next result is due to Vrabie [VR].

Theorem 5.33. *Suppose $-A$ is a closed densely defined positive self-adjoint operator. Then A is the generator of a C_0-contraction semigroup $S(t)$. Furthermore, if $f \in \mathcal{H}$ and $h(t) = S(t)f$, then the problem:*

$$h'(t) = Ah(t), \; h(0) = f, \tag{5.9}$$

has an unique solution

$$h \in D(A) \cap \mathbb{C}^1((0, \infty); \mathcal{H})$$

and

$$\|Ah(t)\|_{\mathcal{H}} \leq \tfrac{1}{2t} \|f\|_{\mathcal{H}}.$$

Proof. First, since $-A$ is a positive, self-adjoint, closed, and densely defined linear operator on $\mathcal{H}$, it follows that both A and $A^* = A$ are dissipative. Hence, by Theorem 5.29, A is m-dissipative so that A generates a C_0-contraction semigroup and for $Re(\lambda) > 0$, $\|R(\lambda, T)\|_{\mathcal{H}} \leq \frac{1}{Re(\lambda)}$.

It is clear that both $S(t)$ and A determine each other uniquely on $D(A)$, so that, at least for $f \in D(A)$, the solution to (5.6) is unique. If $f \in D(A^2)$, we see that, since $(h''(t), h'(t)) = (Ah'(t), h'(t))$, the problem

$$h''(t) = Ah'(t), \; h(0) = f,$$

has an unique solution. Thus, with $(h''(t), h'(t)) = (Ah'(t), h'(t))$ and, for $0 \leq s \leq t$, we have

$$\frac{1}{2} \left\|h'(t)\right\|^2_{\mathcal{H}} - \frac{1}{2} \left\|h'(s)\right\|^2_{\mathcal{H}} = \int_s^t (Ah'(\tau), h'(\tau))d\tau \leq 0$$

(since A is dissipative). This shows that $\|h'(t)\|_{\mathcal{H}}$ is a nonincreasing function. Furthermore,

$$\frac{d}{dt} \|h(t)\|^2_{\mathcal{H}} = 2(Ah(t), h(t)) \tag{5.10}$$

and

$$\frac{d}{dt}(Ah(t), h(t)) = 2(Ah(t), Ah(t)) = 2 \left\|h'(t)\right\|^2_{\mathcal{H}} \geq 0. \tag{5.11}$$

It follows that $(Ah(t), h(t))$ is nondecreasing. If we integrate Eq. (5.7) from $0 \to t$, we have:

$$\|h(t)\|^2_{\mathcal{H}} - \|f\|^2_{\mathcal{H}} = 2 \int_0^t (Ah(\tau), h'(\tau))d\tau \leq 2t(Ah(t), h(t)),$$

$$\Rightarrow \; -t(Ah(t), h(t)) \leq - \int_0^t (Ah(\tau), h'(\tau))d\tau = -\frac{1}{2} \left\|h'(t)\right\|^2_{\mathcal{H}} + \frac{1}{2} \|f\|^2_{\mathcal{H}} \leq \frac{1}{2} \|f\|^2_{\mathcal{H}}.$$

Now recall that $\|h'(t)\|_{\mathcal{H}}$ is a nonincreasing function and integrate equation (5.8) from $0 \to t$ to get

$$(Ah(t), h(t)) - (Af, f) = 2\int_0^t \left\|h'(\tau)\right\|_{\mathcal{H}}^2 d\tau \geq 2t \left\|h'(t)\right\|_{\mathcal{H}}^2.$$

Since $(Ah(t), h(t)) \leq 0$, we see that $2t\,\|h'(t)\|_{\mathcal{H}}^2 \leq (-Af, f)$. If we now multiply both sides of Eq. (5.7) by t and integrate, we see that

$$2t^2 \left\|h'(t)\right\|_{\mathcal{H}}^2 \leq \int_0^t \tau(Ah(\tau), h'(\tau))d\tau = \int_0^t \tau\frac{d}{d\tau}(Ah(\tau), h(\tau))d\tau$$
$$= t(Ah(t), h(t)) - \int_0^t \tau(Ah(\tau), h(\tau))d\tau.$$

Since $t(Ah(t), h(t)) \leq 0$, we see from the inequality above and Eq. (5.7) that $4t^2\,\|Ah(t)\|_{\mathcal{H}}^2 \leq \|f\|_{\mathcal{H}}^2$ so that

$$\|Ah(t)\|_{\mathcal{H}} \leq \frac{\|f\|_{\mathcal{H}}}{2t}.$$ □

The next result shows that we can recover the semigroup as the inverse Laplace transform of the resolvent. It will be important for our study of analytic semigroups in the next section.

Theorem 5.34. *Let A be a closed densely defined dissipative linear operator on $\mathcal{H}$ satisfying:*

(1) *For some $0 < \delta < \pi/2$,*
$$\rho(A) \supset \Sigma_\delta = \{\lambda \,:\, |\arg\lambda| < \pi/2 + \delta\} \cup \{0\}.$$
(2) *The resolvent of A satisfies $\|R(\lambda, A)\| \leqslant 1/|\lambda|$, for each $\lambda \in \Sigma_\delta$, with $\lambda \neq 0$.*

Then A is the generator of a C_0-contraction semigroup $S(t)$, which can be represented as:

$$S(t) = \frac{1}{2\pi i}\int_\Gamma e^{\lambda t}R(\lambda, A)d\lambda, \tag{5.12}$$

where Γ is a smooth curve in Σ_δ going from $\infty e^{-i\theta} \to \infty e^{i\theta}$, for $\pi/2 < \theta < \pi/2+\delta$ and the integral converges in the uniform topology for $t > 0$.

Proof. Let

$$Z(t) = \frac{1}{2\pi i}\int_\Gamma e^{\mu t}R(\mu, A)d\mu. \tag{5.13}$$

Since $\|R(\mu, A)\| \leqslant 1/|\mu|$, we see from the definition of Σ_δ that, for $t > 0$, this integral converges in the uniform norm. In order to see that $Z(t)$

is a semigroup, suppose that $Z(s)$ also has the above representation, with another slightly shifted path Γ' inside Σ_δ. Then

$$Z(s)Z(t) = \left(\frac{1}{2\pi i}\right)^2 \int_{\Gamma'}\int_{\Gamma} e^{\mu t} R(\mu, A) e^{\mu' t} R(\mu', A) d\mu d\mu'$$
$$= \left(\frac{1}{2\pi i}\right)^2 \left[\int_{\Gamma'} e^{\mu' s} R(\mu', A) d\mu' \int_{\Gamma} e^{\mu t}(\mu - \mu')^{-1} d\mu\right.$$
$$\left. - \int_{\Gamma} e^{\mu t} R(\mu, A) d\mu \int_{\Gamma'} e^{\mu' s}(\mu - \mu')^{-1} d\mu'\right],$$

where we have used the resolvent equation, $R(\mu', A)R(\mu, A) = (\mu - \mu')^{-1}R(\mu', A) - R(\mu, A)$, in the second line. If we now use the fact that:

$$\int_{\Gamma'} e^{\mu' s}(\mu - \mu')^{-1} d\mu' = 2\pi i e^{\mu s} \;\&\; \int_{\Gamma'} e^{\mu t}(\mu - \mu')^{-1} d\mu = 0,$$

we get that

$$Z(s)Z(t) = \frac{1}{2\pi i}\int_{\Gamma} e^{\mu(t+s)} R(\mu, A) d\mu = Z(t+s).$$

Since the resolvent uniquely determines the semigroup, we are done if we can show that $R(\lambda, A)$ is the resolvent of $Z(t)$. To do this, use the fact that $R(\lambda, A)$ is analytic in Σ_δ, so that we can shift the path of integration to a new path Γ_t, still inside Σ_δ. We choose $\Gamma_t = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, where $\Gamma_1 = \{re^{-i\theta} : t^{-1} \le r < \infty\}$, $\Gamma_2 = \{t^{-1}e^{i\phi} : -\theta \le \phi \le \theta\}$ and $\Gamma_3 = \{re^{i\theta} : t^{-1} \le r < \infty\}$ without changing the value of the integral. In this case, for the path Γ_3, we have

$$\left\|\frac{1}{2\pi i}\int_{\Gamma_3} e^{\mu t} R(\mu, A) d\mu\right\|_{\mathcal{H}} \leqslant \frac{1}{2\pi}\int_{t^{-1}}^{\infty} e^{-rt\sin(\theta-\pi/2)} r^{-1} dr$$
$$= \frac{1}{2\pi}\int_{\sin(\theta-\pi/2)}^{\infty} e^{-s}s^{-1} ds \leqslant C_1.$$

For the path Γ_2, we see that

$$\left\|\frac{1}{2\pi i}\int_{\Gamma_2} e^{\mu t} R(\mu, A) d\mu\right\|_{\mathcal{H}} \leqslant \frac{1}{2\pi}\int_{-\theta}^{\theta} e^{\cos(\phi)} d\phi \leqslant C_2.$$

The estimate for Γ_1 is like that of Γ_3. This shows that $Z(t)$ is bounded by some constant K for $0 < t < \infty$. Now, if we multiply Eq. (5.10) by $e^{-\lambda t}$ and integrate from 0 to T, using Fubini's Theorem along with the residue theorem, we have

$$\int_0^T e^{-\lambda t} Z(t)dt = \frac{1}{2\pi i}\int_0^T e^{-\lambda t}\left[\int_\Gamma e^{\mu t} R(\mu, A)d\mu\right] dt$$
$$= \frac{1}{2\pi i}\int_\Gamma \left[\int_0^T e^{(\mu-\lambda)t} dt\right] R(\mu, A)d\mu = \frac{1}{2\pi i}\int_\Gamma \frac{\left(e^{(\mu-\lambda)T}-1\right)}{\mu-\lambda} R(\mu, A)d\mu$$
$$= R(\lambda, A) + \frac{1}{2\pi i}\int_\Gamma e^{(\mu-\lambda)T}\frac{R(\mu, A)}{\mu-\lambda}d\mu.$$

However, on Γ,

$$\left\|\frac{1}{2\pi i}\int_\Gamma e^{(\mu-\lambda)T}\frac{R(\mu, A)}{\mu-\lambda}d\mu\right\| \leqslant e^{-\lambda T}\int_\Gamma \frac{d\,|\mu|}{|\mu|\,|\mu-\lambda|} \to 0, \quad T\to\infty.$$

Thus, if we take the limit in our equation, we get

$$\int_0^\infty e^{-\lambda t} Z(t)dt = R(\lambda, A).$$

Since for $Re(\lambda) > 0$, $\frac{1}{|\lambda|} \le \frac{1}{Re(\lambda)}$, we see that $Z(t) = S(t)$ is a contraction semigroup. □

5.2.3. Analytic Semigroups. Let $\Delta = \{w \in \mathbf{C} : \theta_1 < \arg w < \theta_2,\ \theta_1 < 0 < \theta_2\}$. For each $w \in \Delta$, let $S(w)$ be a bounded linear operator on $\mathcal{H}$.

Definition 5.35. The family $S(w)$ is said to be an analytic semigroup on $\mathcal{H}$, for $w \in \Delta$, if

(1) $S(w)f$ is an analytic function of $w \in \Delta$ for each f in $\mathcal{H}$,
(2) $S(0) = I$ and $\lim_{w\to 0} S(w)f = f$ for every $f \in \mathcal{H}$,
(3) $S(w_1 + w_2) = S(w_1)S(w_2)$ for $w_1,\ w_2 \in \Delta$.

Theorem 5.36. *Let $S(t)$ be a C_0-contraction semigroup and let A be the generator of $S(t)$, with $0 \in \rho(A)$. Suppose A satisfies:*

(1) *For $0 < \delta < \pi/2$,*

$$\rho(A) \supset \Sigma_\delta = \{\lambda : |\arg\lambda| < \pi/2 + \delta\} \cup \{0\}.$$

(2) *$\|R(\lambda, A)\| \leqslant M/|\lambda|$ for each $\lambda \in \Sigma_\delta$, with $\lambda \ne 0$.*

Then the following are equivalent:

(1) *$S(t)$ is differentiable for $t > 0$ and there is a constant C such that*

$$\|AS(t)\|_{\mathcal{H}} \le \frac{C}{t} \text{ for } t > 0.$$

(2) *For $t > 0$ and $|z - t| \leqslant Kt$ for some constant K, the series*
$$S(z+t) = S(t) + \sum\nolimits_{n=1}^{\infty} (z^n/n!) S^{(n)}(t)$$
converges uniformly in the above interval.

(3) *$S(t)$ can be extended to a C_0-analytic semigroup $S(z)$, for $z \in \bar{\Delta}_{\delta'}$, with $\bar{\Delta}_{\delta'} = \{z : |\arg z| \leqslant \delta' < \delta\}$.*

Proof. From Eq. (5.9), $S(t) = (1/2\pi i)\int_\Gamma e^{\lambda t} R(\lambda, A) d\lambda$, where Γ is a smooth curve in Σ_δ composed of two rays $\rho e^{i\theta}$ and $\rho e^{-i\theta}$, $0 < \rho < \infty$ and $\pi/2 < \theta < \pi/2 + \delta$ and Γ is oriented so that $\mathrm{Im}(\lambda)$ increases along Γ. The integral converges in the uniform topology for $t > 0$. If we differentiate it formally, we see that:
$$S'(t) = \frac{1}{2\pi i}\int_\Gamma \lambda e^{\lambda t} R(\lambda, A) d\lambda.$$
However, this integral converges in $\mathcal{H}$ for all $t > 0$, since
$$\left\|S'(t)\right\| \leqslant (1/\pi)\int_0^\infty e^{-\rho\cos\theta t} d\rho = \frac{1}{\pi t \cos\theta} = \left(\frac{1}{\pi\cos\theta}\right)\frac{1}{t}. \quad (5.14)$$
Thus, the formal differentiation is justified for $t > 0$ and
$$\|AS(t)\|_{\mathcal{H}} \leq \frac{C}{t}, \text{ where } C = \frac{1}{\pi\cos\theta}.$$
We now prove that $S(t)$ has derivatives of any order, by induction. From above, we know it is true for $k = 1$. Suppose that it is true for $k = n$ and $t \geqslant s$, then
$$S^{(n)}(t) = (AS(t/n))^n = S(t-s)\,(AS(s/n))^n. \quad (5.15)$$
If we differentiate Eq. (5.12) with respect to t we have
$$S^{(n+1)}(t) = (AS(t/n))^n = AS(t-s)\,(AS(s/n))^n.$$
Now set $s = nt/(n+1)$ to get $S^{(n+1)}(t) = [AS(t/(n+1)]^{n+1}$, so that $S(t)$ has derivatives of all orders. If we use this result in Eq. (5.11), and the fact that $n!e^n \geqslant n^n$, we get that:
$$\frac{1}{n!}\left\|S^{(n)}(t)\right\| \leqslant \left(\frac{Ce}{t}\right)^n.$$
Now, consider the power series
$$S(z) = S(t) + \sum\nolimits_{n=1}^{\infty} \frac{S^{(n)}(t)}{n!}(z-t)^n.$$
The series converges uniformly in $L[\mathcal{H}]$ for $|z - t| \leqslant Kt$, where $K = k/eC$, $0 < k < 1$. Thus, $S(z)$ is analytic in $\Delta = \{z : |\arg z| <$

$\arctan K\}$ and hence extends $S(t)$. It is easy to check that $S(z)$ is a C_0-contraction semigroup in any closed subsector $\bar{\Delta}_\varepsilon = \{z : |\arg z| \leqslant \arctan M - \varepsilon\}$ of Δ. □

5.2.4. Perturbation Theory. One of the major concerns for the theory of semigroups of operators is to identify conditions under which the sum of two generators is a generator (when properly understood). We restrict our attention to generators of analytic contraction semigroups. (In practice, by the use of an equivalent norm and a shift in the spectrum, most semigroups of interest can be reduced to contractions.) The next result shows when the sum of generators of analytic contraction semigroups generate an analytic contraction semigroup.

Theorem 5.37. *Let A_0 be an m-dissipative generator of an analytic C_0-semigroup and let A_1 be closed on $\mathcal{H}$, with $D(A_1) \supseteq D(A_0)$. Suppose and there are positive constants $0 \leq \alpha < 1$, $\beta \geq 0$ such that*

$$\|A_1\varphi\| \leqslant \alpha \|A_0\varphi\| + \beta \|\varphi\|, \; \varphi \in D(A_0). \tag{5.16}$$

Then $A = A_0 + A_1$, with domain $D(A) = D(A_1)$, generates an analytic C_0 semigroup.

Remark 5.38. We note that, by the Closed Graph Theorem, it suffices to assume that A_1 is dissipative and $D(A_1) \supseteq D(A_0)$ in order to find constants $0 \leq \alpha < 1$, $\beta \geq 0$ satisfying Eq. (5.13).

Proof. To prove our result, first use the fact that A_0 generates an analytic C_0-semigroup to find a sector Σ in the complex plane, with $\rho(A_0) \supset \Sigma$ ($\Sigma = \{\lambda : \; |\arg \lambda| < \pi/2 + \delta'\}$, for some $\delta' > 0$), and for $\lambda \in \Sigma$, $\|R(\lambda \; :, A_0)\|_{\mathcal{H}} \leqslant |\lambda|^{-1}$. From (5.13), $A_1 R(\lambda, A_0)$ is a bounded operator and:

$$\begin{aligned}
\|A_1 R(\lambda, A_0)\varphi\|_{\mathcal{H}} &\leqslant \alpha \|A_0 R(\lambda, A_0)\varphi\|_{\mathcal{H}} + \beta \|R(\lambda, A_0)\varphi\|_{\mathcal{H}} \\
&\leqslant \alpha \|[R(\lambda, A_0) - I]\,\varphi\|_{\mathcal{H}} + \beta |\lambda|^{-1} \|\varphi\|_{\mathcal{H}} \\
&\leqslant 2\alpha \|\varphi\|_{\mathcal{H}} + \beta |\lambda|^{-1} \|\varphi\|_{\mathcal{H}} .
\end{aligned}$$

Thus, if we set $\alpha = 1/4$ and $|\lambda| > 2\beta$, we have $\|A_1 R(\lambda, A_0)\|_{\mathcal{H}} < 1$ and it follows that the operator $I - A_1 R(\lambda, A_0)$ is invertible. Now it is easy to see that:

$$(\lambda I - (A_0 + A_1))^{-1} = R(\lambda,\, A_0)\,(I - A_1 R(\lambda,\, A_0))^{-1}. \tag{5.17}$$

Using $|\lambda| > 2\beta$, with $|\arg \lambda| < \pi/2 + \delta''$ for some $\delta'' > 0$, and the fact that A_0 and A_1 are m-dissipative generators, we get from (5.14) that

$$\|R(\,\lambda\,,\,A_0 + A_1)\|_{\mathcal{B}} \leqslant |\lambda|^{-1}\,.$$

Thus, A generates a C_0-analytic semigroup. Finally, we note that if $Re(\lambda) > 0$, then $\frac{1}{|\lambda|} \leq \frac{1}{Re(\lambda)}$, so that A also generates a C_0-contraction semigroup. □

Corollary 5.39. *Let A_0 be the generator of an analytic C_0-semigroup and suppose that A_1 is bounded. Then $A_0 + A_1$ is the generator of an analytic C_0-semigroup on $\mathcal{H}$.*

Corollary 5.40. *Let A, A_1 be generators of C_0-contraction semigroups on $\mathcal{H}$ and assume that A_1 is bounded. Then $A + A_1$ is the generator of a C_0-contraction semigroup $S(t)$.*

Theorem 5.25 shows that all closed densely defined linear operators on $\mathcal{H}$ may be approximated by bounded generators of contraction semigroups. This leads to the following result, which shall prove quite useful later.

Theorem 5.41. *Let A_0, A_1* and *$A_0 + A_1$ be generators of contraction semigroups on $\mathcal{H}$, with a common dense domain. Then:*

$$\lim_{\lambda\to\infty} \exp\left\{(A_0 + A_{1,\lambda})\,t\right\}\varphi = \exp\left\{(A_0 + A_1)\,t\right\}\varphi \quad \text{for} \quad t > 0.$$

Proof. The proof is standard. Set $A = A_0 + A_1$, & $A_\lambda = A_0 + A_{1,\lambda}$; then, for $\varphi \in D(A_0) \cap D(A_1)$:

$$\begin{aligned}
&\left\|\left(e^{tA_\lambda} - e^{tA}\right)\varphi\right\|_{\mathcal{H}} = \left\|\int_0^1 \frac{d}{ds}\left[e^{tsA_\lambda}e^{t(1-s)A}\right]\varphi ds\right\|_{\mathcal{H}} \\
&= \left\|t\int_0^1 \left[e^{tsA_\lambda}A_\lambda e^{t(1-s)A} - e^{tsA_\lambda}Ae^{t(1-s)A}\right]\varphi ds\right\|_{\mathcal{H}} \\
&= \left\|t\int_0^1 \left[e^{tsA_\lambda}\left(A_\lambda - A\right)e^{t(1-s)A}\right]\varphi ds\right\|_{\mathcal{H}} \\
&\leqslant t\sup_{s\geq 0}\left\|\left(A_\lambda - A\right)e^{t(1-s)A}\varphi\right\|_{\mathcal{H}} = t\left\|\left(A_{1,\lambda} - A_1\right)e^{t(1-\bar{s})A}\varphi\right\|_{\mathcal{H}},
\end{aligned}$$

where $\bar{s}$ is the point in $[0,1]$ where the sup is attained. The limit of this last term is clearly zero. (Note that A_λ need not commute with A.) □

We reserve our proof of the next result until Chap. 7 (see [K1]). There, we will use it to provide a very general version of the Feynman–Kac formula.

Theorem 5.42. Trotter–Kato product formula *Suppose that* A_0, A_1 *and* $A = \overline{A_0 + A_1}$ *are generators of* C_0*-contraction semigroups* $T_0(t)$, $T_1(t)$ *and* $T(t)$ *on* $\mathcal{H}$. *Then, for* $\varphi \in \mathcal{H}$, *we have*

$$\lim_{n\to\infty} \{T_0(\tfrac{t}{n})T_1(\tfrac{t}{n})\}^n \varphi = T(t)\varphi.$$

5.2.5. Semigroups on Banach Spaces. The purpose of this section is to show that the Hilbert space theory is sufficient for the theory on separable Banach spaces. We assume that "$\mathcal{B}$ is rigged," so that $\mathcal{H}_1 \subset \mathcal{B} \subset \mathcal{H}_2$ as continuous dense embeddings.

Theorem 5.43. *Suppose that* A *generates a* C_0*-contraction semigroup* $T(t)$, *on* $\mathcal{B}$ *and* $\mathcal{B}' \subset \mathcal{H}_2$ *then:*

(1) *A has a closed densely defined extension $\bar{A}$ to $\mathcal{H}_2$, which is also the generator of a C_0-contraction semigroup.*
(2) $\rho(\bar{A}) = \rho(A)$ *and* $\sigma(\bar{A}) = \sigma(A)$.
(3) *The adjoint of $\bar{A}$, $\bar{A}^*$, restricted to $\mathcal{B}$, is the adjoint A^* of A, that is:*
   - *the operator* $A^*A \geqslant 0$,
   - $(A^*A)^* = A^*A$ *and*
   - $I + A^*A$ *has a bounded inverse.*

Proof. Part I
Let $T(t)$ be the C_0-contraction semigroup generated by A. By Theorem 5.15, $T(t)$ has a bounded extension $\bar{T}(t)$ to $\mathcal{H}_2$.

We prove that $\bar{T}(t)$ is a C_0-semigroup. (The fact that it is a contraction semigroup will follow later.) It is clear that $\bar{T}(t)$ has the semigroup property. To prove that it is strongly continuous, use the fact that $\mathcal{B}$ is dense in $\mathcal{H}_2$ so that, for each $g \in \mathcal{H}_2$, there is a sequence $\{g_n\}$ in $\mathcal{B}$ converging to g. We then have:

$$\begin{aligned}
&\lim_{t\to 0} \left\|\bar{T}(t)g - g\right\|_2 \leqslant \lim_{t\to 0} \{\left\|\bar{T}(t)g - \bar{T}(t)g_n\right\|_2 + \left\|\bar{T}(t)g_n - g_n\right\|_2\} + \left\|g_n - g\right\|_2 \\
&\leqslant k\left\|g - g_n\right\|_2 + \lim_{t\to 0}\left\|\bar{T}(t)g_n - g_n\right\|_2 + \left\|g_n - g\right\|_2 \\
&= (k+1)\left\|g - g_n\right\|_2 + \lim_{t\to 0}\left\|T(t)g_n - g_n\right\|_2 = (k+1)\left\|g - g_n\right\|_2,
\end{aligned}$$

where we have used the fact that $\bar{T}(t)g_n = T(t)g_n$ for $g_n \in \mathcal{B}$, and k is the constant in Theorem 5.15. It is clear that we can make the last term on the right as small as we like by choosing n large enough, so that $\bar{T}(t)$ is a C_0-semigroup.

To prove (2), note that if $\bar{A}$ is the extension of A, and $\lambda I-\bar{A}$ has an inverse, then $\lambda I-A$ also has one, so $\rho(\bar{A}) \subset \rho(A)$ and $Ran(\lambda I-A)_{\mathcal{B}} \subset Ran(\lambda I-\bar{A})_{\mathcal{H}_2} \subset \overline{Ran(\lambda I-A)}_{\mathcal{H}_2}$ for any $\lambda \in \mathbb{C}$. For the other direction, since A generates a C_0-contraction semigroup, $\rho(A) \neq \emptyset$. Thus, if $\lambda \in \rho(A)$, then $(\lambda I-A)^{-1}$ is a continuous mapping from $Ran(\lambda I-A)$ onto $D(A)$ and $Ran(\lambda I-A)$ is dense in $\mathcal{B}$. Let $g \in D(\bar{A})$, so that $(g, \bar{A}g) \in \hat{G}(A)$, the closure of the graph of A in $\mathcal{H}_2$. Thus, there exists a sequence $\{g_n\} \subset D(A)$ such that $\|g-g_n\|_G = \|g-g_n\|_{\mathcal{H}_2} + \left\|\bar{A}g-\bar{A}g_n\right\|_{\mathcal{H}_2} \to 0$ as $n \to \infty$. Since $\bar{A}g_n = Ag_n$, it follows that $(\lambda I-\bar{A})g = \lim_{n\to\infty}(\lambda I-A)g_n$. However, by the boundedness of $(\lambda I-A)^{-1}$ on $Ran(\lambda I-A)$, we have that, for some $\delta>0$,

$$\left\|(\lambda I-\bar{A})g\right\|_{\mathcal{H}_2} = \lim_{n\to\infty}\|(\lambda I-A)g_n\|_{\mathcal{H}_2} \geq \lim_{n\to\infty}\delta\,\|g_n\|_{\mathcal{H}_2} = \delta\,\|g\|_{\mathcal{H}_2}.$$

It follows that $\lambda I-\bar{A}$ has a bounded inverse and since $D(A) \subset D(\bar{A})$ implies that $Ran(\lambda I-A) \subset Ran(\lambda I-\bar{A})$, we see that $Ran(\lambda I-\bar{A})$ is dense in $\mathcal{H}_2$ so that $\lambda \in \rho(\bar{A})$ and hence $\rho(A) \subset \rho(\bar{A})$. It follows that $\rho(A)=\rho(\bar{A})$ and necessarily, $\sigma(A)=\sigma(\bar{A})$.

Since A generates a C_0-contraction semigroup, it is m-dissipative. From the Lumer–Phillips Theorem, we have that $Ran(\lambda I-A)=\mathcal{B}$ for $\lambda>0$. It follows that $\bar{A}$ is m-dissipative and $Ran(\lambda I-\bar{A})=\mathcal{H}_2$. Thus, $\bar{T}(t)$ is a C_0-contraction semigroup.

We now observe that the same proof applies to $\bar{T}^*(t)$, so that $\bar{A}^*$ is also the generator of a C_0-contraction semigroup on $\mathcal{H}_2$.

Clearly $\bar{A}^*$ is the adjoint of $\bar{A}$ so that, from von Neumann's Theorem, $\bar{A}^*\bar{A}$ has the expected properties. $\bar{\mathbf{D}}=D(\bar{A}^*\bar{A})$ is a core for $\bar{A}$ (i.e., the set of elements $\{g, \bar{A}g\}$ is dense in the graph, $G[\bar{A}]$, of $\bar{A}$ for $g \in \bar{\mathbf{D}}$). From here, we see that the restriction A^* of $\bar{A}^*$ to $\mathcal{B}$ is the generator of a C_0-contraction semigroup and $\mathbf{D}=D(A^*A)$ is a core for A. The proof of (3) for A^*A now follows. □

Remark 5.44. Theorem 5.43 shows that all C_0-contraction semigroups defined on $\mathcal{B}$ have the same properties as its extension to $\mathcal{H}_2$. Thus, if $\mathcal{B}$ is reflexive or $\mathcal{B}' \subset \mathcal{H}_2$, then all the theorems on $\mathcal{H}_2$ apply to $\mathcal{B}$.

The next result implies that the generalized Yosida Approximation Theorem applies to C_0-semigroups on $\mathcal{B}$

Theorem 5.45. *Let $A \in \mathcal{C}[\mathcal{B}]$ be the generator of a C_0-contraction semigroup. Then there exists an m-accretive operator T and a partial isometry W such that $A = WT$ and $D(A) = D(T)$.*

Proof. The fact that $\mathcal{B}' \subset \mathcal{H}_2$ ensures that A^*A is a closed self-adjoint operator on $\mathcal{B}$ by Theorem 5.40. Furthermore, both A and A^* have closed densely defined extensions $\bar{A}$ and $\bar{A}^*$ to $\mathcal{H}_2$. Thus, the operator $\hat{T} = [\bar{A}^*\bar{A}]^{1/2}$ is a well-defined m-accretive self-adjoint linear operator on $\mathcal{H}_2$, $\bar{A} = \bar{W}\bar{T}$ for some partial isometry $\bar{W}$ defined on $\mathcal{H}_2$, and $D(\bar{A}) = D(\bar{T})$. Our proof is complete when we notice that the restriction of $\bar{A}$ to $\mathcal{B}$ is A and $\bar{T}^2$ restricted to $\mathcal{B}$ is A^*A, so that the restriction of $\bar{W}$ to $\mathcal{B}$ is well defined and must be a partial isometry. The equality of the domains is obvious. □

With respect to our definition of natural self-adjointness, the following related definition is due to Palmer [PL], where the operator is called symmetric. This is essentially the same as a Hermitian operator as defined by Lumer [LU].

Definition 5.46. A closed densely defined linear operator A on $\mathcal{B}$ is called self-conjugate if both iA and $-iA$ are dissipative.

Theorem 5.47. *(Vidav–Palmer) A linear operator A, defined on $\mathcal{B}$, is self-conjugate if and only if iA and $-iA$ are generators of isometric semigroups.*

Theorem 5.48. *The operator A, defined on $\mathcal{B}$, is self-conjugate if and only if it is naturally self-adjoint.*

Proof. Let $\bar{A}$ and $\bar{A}^*$ be the closed densely defined extensions of A and A^* to $\mathcal{H}_2$. On $\mathcal{H}_2$, $\bar{A}$ is naturally self-adjoint if and only if $i\bar{A}$ generates a unitary group, if and only if it is self-conjugate. Thus, both definitions coincide on $\mathcal{H}_2$. It follows that the restrictions coincide on $\mathcal{B}$. □

Additional discussion of the adjoint for operators on Banach spaces can be found in the Appendix (Sect. 5.3).

5.3. Appendix

The appendix is devoted to a number of topics that are not directly related to our main direction, but have independent interest for functional analysis and operator theory. We first discuss the existence of an adjoint for spaces that are not uniformly convex. We then apply our results in subsequent sections to show that the spectral theory that is natural for Hilbert spaces and the Schatten theory of compact operators can also be partially extended to Banach spaces.

5.4. The Adjoint in the General Case

In this section we continue our discussion of the adjoint for an operator on Banach space with an S-basis $\mathcal{B}$, which is not uniformly convex.

5.4.1. The General Case for Unbounded A. A Banach space is said to have the approximation property if every compact operator is the limit of operators of finite rank. It is known that every classical Banach space has the approximation property. However, it is also known that there are separable Banach spaces without the approximation property (see Diestel [**DI**]). Theorem 5.15 tells us that if $\mathcal{B}' \subset \mathcal{H}_2$, then $L[\mathcal{B}] \subset L[\mathcal{H}_2]$ as a continuous embedding. (It's not hard to show that if $\mathcal{B}$ has the approximation property, the embedding is dense.)

Let $A \in \mathcal{C}[\mathcal{B}]$, the closed densely defined linear operators on $\mathcal{B}$. By definition, A is of Baire class one if it can be approximated by a sequence, $\{A_n\}$, of bounded linear operators. In this case, it is natural to define $A^* = s\text{-}\lim A_n^*$ (see below). However, if $\mathcal{B}$ is not uniformly convex there may be operators $A \in \mathcal{C}[\mathcal{B}]$ that are not of Baire class one, so that it is not reasonable to expect Theorem 5.11 to hold for all of $\mathcal{C}[\mathcal{B}]$. First, we note that every uniformly convex Banach space is reflexive. In order to understand the problem, we need the following:

Definition 5.49. A Banach space $\mathcal{B}$ is said to be:

(1) quasi-reflexive if $\dim\{\mathcal{B}''/\mathcal{B}\} < \infty$, and
(2) nonquasi-reflexive if $\dim\{\mathcal{B}''/\mathcal{B}\} = \infty$.

A theorem by Vinokurov et al. [**VPP**] shows that, for every nonquasi-reflexive Banach space $\mathcal{B}$ (for example, $C[0;1]$ or $L^1[\mathbb{R}^n]$, $n \in \mathbb{N}$), there is at least one closed densely defined linear operator A, which is not of Baire class one. It can even be arranged so that A^{-1} is a bounded linear injective operator (with a dense range). This means, in particular, that there does not exist a sequence of bounded linear operators $A_n \in L[\mathcal{B}]$ such that, for $g \in D(A)$, $A_n g \to Ag$, as $n \to \infty$. The following result shows that whenever $\mathcal{B}' \subset \mathcal{H}_2$, every operator of Baire class one has an adjoint.

Theorem 5.50. *If $A \in \mathcal{C}[\mathcal{B}]$ and $\mathcal{B}' \subset \mathcal{H}_2$, then A is in the first Baire class if and only if it has an adjoint $A^* \in \mathcal{C}[\mathcal{B}]$.*

Proof. Let $\mathcal{H}_1 \subset \mathcal{B} \subset \mathcal{H}_2$ and suppose that A has an adjoint $A^* \in \mathcal{C}[\mathcal{B}]$. Let $T = [A^*A]^{1/2}$, $\bar{T} = [AA^*]^{1/2}$. Since T is m-accretive and naturally self-adjoint, for all $\alpha > 0$, $I + \alpha T$ has a bounded inverse $S(\alpha) = (I + \alpha T)^{-1}$. It is easy to see that $AS(\alpha)$ is bounded and, for $g \in D(A)$, $AS(\alpha)g = \bar{S}(\alpha)Ag = (I + \alpha\bar{T})^{-1}Ag$. Using this result, we have:

$$\lim_{\alpha\to 0^+} AS(\alpha)g = \lim_{\alpha\to 0^+} \bar{S}(\alpha)Ag = Ag, \text{ for } g \in D(A).$$

It follows that A is in the first Baire class.

To prove the converse suppose that $A \in \mathcal{C}[\mathcal{B}]$ is of first Baire class. If $\{A_n\}$ is a sequence of bounded linear operators with $A_n g \to Ag$, for all $g \in D(A)$, then each A_n has an adjoint A_n^*. Since $\mathcal{B}' \subset \mathcal{H}_2$, each $A_n\, A_n^*$ has a bounded extension $\bar{A}_n\, \bar{A}_n^*$ to $\mathcal{H}_2$. Furthermore, since A is densely defined, it has a closed densely defined extension $\bar{A}$ on $\mathcal{H}_2$. Let $\bar{A}^*$ be the adjoint of $\bar{A}$. Then, for all $g \in D(A)$, $h \in \mathcal{B}$, we have:

$$\begin{aligned}&\lim_{n\to\infty} (A_n g, h)_{\mathcal{H}_2} = \lim_{n\to\infty} (g, A_n^* h)_{\mathcal{H}_2} = (Ag, h)_{\mathcal{H}_2}\\ &= \left(\bar{A}g, h\right)_{\mathcal{H}_2}\end{aligned}$$

From here, we see that $A^* = \lim_{n\to\infty} A_n^*$ is a densely defined linear operator. If we let $D(A^*) \subset \mathcal{B}$ be the dense set, then for $h \in D(A^*)$

$$\lim_{n\to\infty} (g, A_n^* h)_{\mathcal{H}_2} = \lim_{n\to\infty} (g, A^* h)_{\mathcal{H}_2} = \left(g, \bar{A}^* h\right)_{\mathcal{H}_2},$$

so that A^* is the restriction of $\bar{A}^*$ to $\mathcal{B}$. □

Corollary 5.51. *If $A \in \mathcal{C}[\mathcal{B}]$ is in the first Baire class and $\mathcal{B}' \subset \mathcal{H}_2$, then $A = WT$, where W is a partial isometry and $T = [A^*A]^{1/2}$.*

5.4.1.1. **The Adjoint Is Not Unique.** In this section we show that if A is defined on a fixed Banach space $\mathcal{B}$, then two different Hilbert space riggings can produce two different adjoints for A.

Recall that a regular σ-finite measure on the σ-algebra of Borel sets of a Hausdorff topological space is called a Radon measure, and a function u is of bounded variation on Ω, or $u \in BV[\Omega]$, if $u \in L^1[\Omega]$ and there is a Radon vector measure Du such that

$$\int_\Omega u(x)\nabla\phi(x)dx = -\int_\Omega \phi(x)Du(x),$$

for all functions $\phi \in \mathbb{C}_c^\infty[\Omega, \mathbb{R}^n]$, the $\mathbb{R}^n$-valued infinitely differentiable functions on Ω with compact support. It is easy to see that $W_0^{1,1}[\Omega] \subset BV[\Omega]$. (In this case, we can show that $Du(x) = \nabla u(x)dx$.)

Let us return to the two pair of Hibert spaces $H_0^1[\Omega] \subset \mathbb{C}_0[\Omega] \subset H_0[\Omega]$ and $\mathcal{H}_1[\Omega] \subset \mathbb{C}_0[\Omega] \subset \mathcal{H}_2[\Omega]$ of Example 3.32 in Chap. 3.

Let $A = [-\Delta]$ be defined on $\mathbb{C}_0[\Omega]$, with domain:

$$D_c(A) = \{\Delta u \in \mathbb{C}_0[\Omega] \, |u = 0 \text{ on } \partial\Omega\}.$$

It is easy to see that A extends to a self-adjoint operator on $H_0[\Omega]$, with domain

$$D_2(A) = \{\Delta u \in H_0[\Omega] \, |u=0 \text{ on } \partial\Omega \text{ and, } \nabla u \text{ is absolutely continuous}\}.$$

To begin, we first compute the adjoint A^*, of A directly as an operator on $\mathbb{C}_0[\Omega]$. The dual space of $\mathbb{C}_0[\Omega]$ is $\mathbb{C}_0^*[\Omega] = rca[\Omega]$, the space of regular countable additive measures on Ω.

It follows from

$$\langle Au, v\rangle = -\int_\Omega \Delta u(x)v(x)dx,$$

that

$$\langle u, A^* v\rangle = -\int_\Omega u(x)\Delta v(x)dx$$

and

$$D_c(A^*) = \{u : \Delta u \in BV[\Omega] \mid u = 0 \text{ on } \partial\Omega\},$$

so that $D_c(A) \subset D_c(A^*)$ (proper). Thus, if we restrict A^* to $D_c(A)$ it becomes a self-adjoint operator on $\mathbb{C}_0[\Omega]$ without the rigging.

We now investigate the adjoint obtained from use of the first rigging, $H_0^1[\Omega] \subset \mathbb{C}_0[\Omega] \subset H_0[\Omega]$ (see Barbu [B], p. 4). In this case, $\mathbf{J}_1 = [-\Delta]$ and $\mathbf{J}_2 = \mathbf{I}_2$, the identity operator on $H_0[\Omega]$, so that

$$A_1^* = \mathbf{J}_1^{-1} A_1' \mathbf{J}_2, = \mathbf{I}_2.$$

In the second rigging, $\mathcal{H}_1[\Omega] \subset \mathbb{C}_0[\Omega] \subset \mathcal{H}_2[\Omega]$, constructed in Example 3.10 in Chap. 3, we have

$$A_2^* = \mathbf{J}_1^{-1} A_1' \mathbf{J}_2.$$

In this case,

$$\mathbf{J}_1(v) = \sum\nolimits_{n=1}^\infty t_n^{-1} (e_n, v)_2(\,\cdot\,, e_n)_2, \quad \mathbf{J}_2(v) = \sum\nolimits_{n=1}^\infty t_n \bar{F}_n(v) F_n(\,\cdot\,)$$

and

$$(e_n, v)_2 = \sum\nolimits_{k=1}^\infty t_k \bar{F}_k(v) F_k(e_n) = t_n \bar{F}_n(v),$$

so that $\mathbf{J}_1(v) = \sum_{n=1}^{\infty} \bar{F}_n(v)(\,\cdot,\, e_n)_2$. However,

$$(\,\cdot\,, e_n)_2 = \sum\nolimits_{k=1}^{\infty} t_k \bar{F}_k(e_n) F_k(\,\cdot\,) = t_n F_n(\,\cdot\,), \text{ so that } \mathbf{J}_1 = \mathbf{J}_2.$$

It follows that $\mathbf{J}_2(A_2^* u) = \mathbf{J}_2(Au)$, so that $A_2^* = A = [-\Delta]$, with the same domains.

It follows that the natural adjoint obtained on $\mathbb{C}_0[\Omega]$ coincides with the adjoint constructed from our special rigging. On the other hand, we also see that different riggings can give distinct adjoints. (It is clear that the requirements of Theorem 5.5 are satisfied by both adjoints.)

Definition 5.52. We say that $\mathcal{H}_1$ and $\mathcal{H}_2$ is an adjoint canonical pair for $\mathcal{B}$ if $\mathcal{H}_1 \subset \mathcal{B} \subset \mathcal{H}_2$ as continuous dense embeddings and $\mathcal{B}' \subset \mathcal{H}_2$. In this case, when $A \in \mathcal{C}[\mathcal{B}]$, A^* is called the canonical adjoint.

5.4.2. Operators on $\mathcal{B}$.

Definition 5.53. Let $\mathcal{B}$ have an S-basis, U be bounded, $A \in \mathbb{C}[\mathcal{B}]$ and let $\mathcal{U}$, $\mathcal{V}$ be subspaces of $\mathcal{B}$. Then:

(1) A is said to be naturally self-adjoint if $D(A) = D(A^*)$ and $A = A^*$.
(2) A is said to be normal if $D(A) = D(A^*)$ and $AA^* = A^*A$.
(3) U is unitary if $UU^* = U^*U = I$.
(4) The subspace $\mathcal{U}$ is $\perp$ to $\mathcal{V}$ if, for each $v \in \mathcal{V}$ and $\forall u \in \mathcal{U}$, $\langle v, J(u)\rangle = 0$ and, for each $u \in \mathcal{U}$ and $\forall v \in \mathcal{V}$, $\langle u, J(v)\rangle = 0$ ($J(u)$ respectively $J(v)$ may be multivalued).

The last definition is transparent since, for example,

$$\langle v, J(u)\rangle = 0 \Leftrightarrow \langle v, J_2(u)\rangle = (v, u)_2 = 0 \ \ \forall v \in \mathcal{V}.$$

Thus, orthogonal subspaces in $\mathcal{H}_2$ induce orthogonal subspaces in $\mathcal{B}$.

Theorem 5.54. *(Gram–Schmidt) For each fixed basis $\{\varphi_i,\ 1 \leqslant i < \infty\}$ of $\mathcal{B}$, there is at least one set of dual functionals $\{S_i\}$ such that $\{\{\psi_i\}, \{S_i\}, 1 \leqslant i < \infty\}$ is a biorthonormal set of vectors for $\mathcal{B}$, (i.e., $\langle \psi_i, S_j\rangle = \delta_{ij}$).*

Proof. Since each φ_i is in $\mathcal{H}_2$, we can construct an orthogonal set of vectors $\{\phi_i,\ 1 \leqslant i < \infty\}$ in $\mathcal{H}_2$ by the standard Gram–Schmidt process. Set $\psi_i = \phi_i / \|\phi_i\|_{\mathcal{B}}$, choose $\hat{S}_i \in J(\psi_i)/\|\psi_i\|_{\mathcal{H}}^2$ and restrict it to the subspace $M_i = [\psi_i] \subset \mathcal{B}$. For each i, let $M_i^{\perp}$ be the subspace spanned by $\{\psi_j,\ i \neq j\}$. Now use the Hahn–Banach Theorem to extend $\hat{S}_i$ to S_i, defined on all of $\mathcal{B}$, with $S_i = 0$ on $M_i^{\perp}$ (see Theorem 1.47). From here, it is easy to check that $\{\{\psi_i\}, \{S_i\},\ 1 \leqslant i < \infty\}$ is a biorthonormal set. If $\mathcal{B}$ is reflexive, the family $\{S_i\}$ is unique. □

We close this section with the following observation about the use of $\mathcal{H}_2 = KS^2$, when $\mathcal{B}$ is one of the classical spaces. Let A be any closed densely defined positive naturally self-adjoint linear operator on $\mathcal{B}$ with a discrete positive spectrum $\{\lambda_i\}$. In this case, $-A$ generates a C_0-contraction semigroup, so that it can be extended to $\mathcal{H}_2$ with the same properties. If we compute the ratio $\frac{\langle A\psi, S_\psi \rangle}{\langle \psi, S_\psi \rangle}$ in $\mathcal{B}$, it will be "close" to the value of $\frac{(\bar{A}\psi,\psi)_{\mathcal{H}_2}}{(\psi,\psi)_{\mathcal{H}_2}}$ in $\mathcal{H}_2$. On the other hand, note that we can use the min-max theorem on $\mathcal{H}_2$ to compute the eigenvalues and eigenfunctions of A via $\bar{A}$ exactly on $\mathcal{H}_2$. Thus, in this sense, the min-max theorem holds on $\mathcal{B}$.

5.5. The Spectral Theorem

5.5.1. Background. Dunford and Schwartz define a spectral operator as one that has a spectral family similar to that defined in Theorem 5.29 of Chap. 4, for self-adjoint operators. (A spectral operator is an operator with countably additive spectral measure on the Borel sets of the complex plane.) Strauss and Trunk [**STT**] define a bounded linear operator A, on a Hilbert space $\mathcal{H}$, to be spectralizable if there exists a nonconstant polynomial p such that the operator $p(A)$ is a scalar spectral operator (has a representation as in Eq. (4.27) in Chap. 4). Another interesting line of attack is represented in the book of Colojoară and Foiaş [**CF**], where they study the class of generalized spectral operators. Here, one is not opposed to allowing the spectral resolution to exist in a generalized sense, so as to include operators with spectral singularities.

The following theorem was proven by Helffer and Sjöstrand [HSJ] (see Proposition 7.2):

Theorem 5.55. *Let $g \in \mathcal{C}_0^{\infty}[\mathbb{R}]$ and let $\hat{g} \in \mathcal{C}_0^{\infty}[\mathbf{C}]$ be an extension of g, with $\frac{\partial \hat{g}}{\partial \hat{z}} = 0$ on $\mathbb{R}$. If A is a self-adjoint operator on $\mathcal{H}$, then*

$$g(A) = -\frac{1}{\pi} \iint_{\mathbf{C}} \frac{\partial \hat{g}}{\partial \bar{z}} (z - A)^{-1}\, dxdy.$$

This defines a functional calculus. Davies [DA] showed that the above formula can be used to define a functional calculus on Banach spaces for a closed densely defined linear operator A, provided $\rho(A) \cap \mathbb{R} = \emptyset$. In this program the objective is to construct a functional calculus pre-supposing that the operator of concern has a reasonable resolvent.

5.5.1.1. *Problem.* The basic problem that causes additional difficulty is the fact that many bounded linear operators are of the form $A = B + N$, where B is normal and N is nilpotent (i.e., there is a $k \in \mathbb{N}$, such that $N^{k+1} = 0,\ N^k \neq 0$). In this case, A does not have a representation with a standard spectral measure. On the other hand, $T = [N^*N]^{1/2}$ is a self-adjoint operator, and there is a unique partial isometry W such that $N = WT$. If $\mathbf{E}(\cdot)$ is the spectral measure associated with T, then $W\mathbf{E}(\Omega)x$ is not a spectral measure, but it is a measure of bounded variation. This idea was used in Chap. 4 (Theorem 4.57) to provide an alternate approach to the spectral theory. In this section, we consider the same possibly for operators on Banach spaces.

To begin, we note that in either of the Strauss and Trunk [**STT**], Helffer and Sjöstrand [**HSJ**], or Davies [**DA**] theory, the operator A is in Baire class one. Thus, A has an adjoint, so that, by Corollary 5.51 $A = WT$, where W is a partial isometry and T is a nonnegative self-adjoint linear operator.

5.5.2. Scalar Case.

Theorem 5.56. *If $\mathcal{B}' \subset \mathcal{H}_2$ and $A \in \mathcal{C}[\mathcal{B}]$ is an operator of Baire class one, then there exists a unique vector-valued function $\mathbf{F}_x(\lambda)$ of bounded variation such that, for each $x \in D(A)$, we have:*

(1) *$D(A)$ also satisfies*

$$D(A) = \left\{ x \in \mathcal{B} \mid \int_{|\sigma(A)|} \lambda^2 \left\langle d\mathbf{F}_x(\lambda), x^* \right\rangle < \infty \right\}$$

for each $x^* \in J(x)$ and

(2) $Ax = \lim_{n\to\infty} \int_0^n \lambda d\mathbf{F}_x(\lambda)$, for all $x \in D(A)$.

Proof. Let $A = WT$, where W is the unique partial isometry and $T = [A^*A]^{1/2}$. Let $\bar{T}$ be the extension of T to $\mathcal{H}_2$. It follows that there is a unique spectral measure $\bar{\mathbf{E}}(\Omega)$ such that for each $x \in D(\bar{T})$:

$$\bar{T}x = \lim_{n\to\infty} \int_0^n \lambda d\bar{\mathbf{E}}(d\lambda)x. \tag{5.18}$$

Furthermore, $\bar{\mathbf{E}}(\lambda)x$ is a vector-valued function of bounded variation and, if $\bar{W}$ is the extension of W, $\bar{W}\bar{\mathbf{E}}_x(\lambda)$ is of bounded variation, with $Var(\bar{W}\bar{\mathbf{E}}_x, \mathbb{R}) \leq Var(\bar{\mathbf{E}}_x, \mathbb{R})$. If we set $\bar{\mathbf{F}}_x(\lambda) = \bar{W}\bar{\mathbf{E}}_x(\lambda)$, for each interval $(a, b) \subset [0, \infty)$,

$$\left\{\bar{W} \int_a^b \lambda d\bar{\mathbf{E}}_x(\lambda)\right\} = \int_a^b \lambda d\bar{\mathbf{F}}_x(\lambda).$$

Since $\bar{A}x = \bar{W}\bar{T}x$ and the restriction of $\bar{A}$ to $\mathcal{B}$ is A, we have, for all $x \in D(A)$,

$$Ax = \lim_{n\to\infty} \int_0^n \lambda d\mathbf{F}_x(\lambda). \tag{5.19}$$

This proves (2). The proof of (1) follows from (1) in Theorem 4.61 of Chap. 4 and the definition of x^*. □

5.5.3. General Case. In this section, we assume that for each i, $1 \leq i \leq n$, $n \in \mathbb{N}$, $\mathcal{B}_i = \mathcal{B}$ is a fixed separable Banach space. We set $\mathfrak{B} = \times_{i=1}^n \mathcal{B}_i$, and represent a vector $\mathbf{x} \in \mathfrak{B}$ by $\mathbf{x}^t = [x_1,\ x_2,\ \cdots,\ x_n]$. An operator $\mathbf{A} = [A_{ij}] \in C[\mathfrak{B}]$ is defined whenever $A_{ij} : \mathcal{B} \to \mathcal{B}$, is in $\mathcal{C}[\mathcal{B}]$.

If $\mathcal{B}' \subset \mathcal{H}_2$ and A_{ij} is of Baire class one, then by Theorem 5.54, there exists a unique vector-valued function $F_x^{ij}(\lambda)$ of bounded variation such that, for each $x \in D(A_{ij})$, we have:

(1) $D(A_{ij})$ also satisfies

$$D(A_{ij}) = \left\{ x \in \mathcal{B} \mid \int_0^\infty \lambda^2 \left\langle dF_x^{ij}(\lambda), x^* \right\rangle_{\mathcal{B}} < \infty \right\}$$

for all $x^* \in J(x)$ and

(2)

$$A_{ij}x = \lim_{n\to\infty} \int_0^n \lambda dF_x^{ij}(\lambda), \text{ for all } x \in D(A_{ij}).$$

If we let $d\mathcal{F}(\lambda) = [dF^{ij}(\lambda)]$, then we can represent $\mathbf{A}$ by:

$$\mathbf{Ax} = \lim_{n\to\infty} \int_0^n \lambda d\boldsymbol{\mathcal{F}}(\lambda)\mathbf{x}, \text{ for all } \mathbf{x} \in D(\mathbf{A}).$$

5.6. Schatten Classes on Banach Spaces

In this section, we show how our approach allows us to provide a natural definition for the Schatten class of operators on $\mathcal{B}$. Here, we assume that the reader has at least read the section concerning compact operators on Hilbert spaces in Chap. 4.

5.6.1. Background: Compact Operators on Banach Spaces. Let $\mathbb{K}(\mathcal{B})$ be the class of compact operators on $\mathcal{B}$ and let $\mathbb{F}(\mathcal{B})$ be the set of operators of finite rank. Recall that, for separable Banach spaces, $\mathbb{K}(\mathcal{B})$ is an ideal that need not be the maximal ideal in $L[\mathcal{B}]$. If $\mathbb{M}(\mathcal{B})$ is the set of weakly compact operators and $\mathbb{N}(\mathcal{B})$ is the set of operators that map weakly convergent sequences into strongly convergent sequences, it is known that both are closed two-sided ideals in the operator norm, and, in general, $\mathbb{F}(\mathcal{B}) \subset \mathbb{K}(\mathcal{B}) \subset \mathbb{M}(\mathcal{B})$ and $\mathbb{F}(\mathcal{B}) \subset \mathbb{K}(\mathcal{B}) \subset \mathbb{N}(\mathcal{B})$ (see part I of Dunford and Schwartz [DS], p. 553). For reflexive Banach spaces, $\mathbb{K}(\mathcal{B}) = \mathbb{N}(\mathcal{B})$ and $\mathbb{M}(\mathcal{B})=L[\mathcal{B}]$. For the space of continuous functions $\mathbf{C}[\Omega]$ on a compact Hausdorff space Ω, Grothendieck [GO] has shown that $\mathbb{M}(\mathcal{B})=\mathbb{N}(\mathcal{B})$. On the other hand, it was shown in part I of Dunford and Schwartz [DS] that for a positive measure space, (Ω, Σ, μ), on $\mathbf{L}^1(\Omega, \Sigma, \mu)$, $\mathbb{M}(\mathcal{B}) \subset \mathbb{N}(\mathcal{B})$.

5.6.2. Uniformly Convex Spaces. We assume that $\mathcal{B}$ is uniformly convex, with an S-basis. In operator theoretic language, the interpretation of our S-basis assumption is that the compact operators on $\mathcal{B}$ have the approximation property, namely that every compact operator can be approximated by operators of finite rank. In this section, we will show that, for the class of uniformly convex Banach spaces with an S-basis, $L[\mathcal{B}]$ almost has the same structure as that of $L[\mathcal{H}]$, when $\mathcal{H}$ is a Hilbert space. The difference being that $L[\mathcal{B}]$ is not a C^*-algebra (i.e., $\|A^*A\| = \|A\|^2$, for all $A \in L[\mathcal{B}]$).

In what follows, we fix $\mathcal{H}_2$. Let A be a compact operator on $\mathcal{B}$ and let $\bar{A}$ be its extension to $\mathcal{H}_2$. For each compact operator $\bar{A}$ on $\mathcal{H}_2$, there exists an orthonormal set of functions $\{\bar{\varphi}_n \,|n \geqslant 1\}$ such that

$$\bar{A} = \sum\nolimits_{n=1}^{\infty} \mu_n(\bar{A})\,(\cdot\,, \bar{\varphi}_n)_2\, \bar{U}\bar{\varphi}_n.$$

Where the μ_n are the eigenvalues of $[\bar{A}^*\bar{A}]^{1/2} = |\bar{A}|$, counted by multiplicity and in decreasing order, and $\bar{U}$ is the partial isometry associated with the polar decomposition of $\bar{A} = \bar{U}\,|\bar{A}|$. Without loss, we can assume that the set of functions $\{\bar{\varphi}_n \,|n \geqslant 1\}$ is contained in $\mathcal{B}$

and $\{\varphi_n \,|n \geqslant 1\}$ is normalized version in $\mathcal{B}$. If $\mathbb{S}_p[\mathcal{H}_2]$ is the Schatten Class of order p in $L[\mathcal{H}_2]$, it is well known that if $\bar{A} \in \mathbb{S}_p[\mathcal{H}_2]$, its norm can be represented as:

$$\|\bar{A}\|_p^{\mathcal{H}_2} = \left\{Tr\left[\bar{A}^*\bar{A}\right]^{p/2}\right\}^{1/p} = \left\{\sum_{n=1}^{\infty}\left(\bar{A}^*\bar{A}\bar{\varphi}_n, \bar{\varphi}_n\right)_{\mathcal{H}_2}^{p/2}\right\}^{1/p}$$
$$= \left\{\sum_{n=1}^{\infty}\left|\mu_n\left(\bar{A}\right)\right|^p\right\}^{1/p}.$$

Definition 5.57. We represent the Schatten Class of order p in $L[\mathcal{B}]$ by:

$$\mathbb{S}_p[\mathcal{B}] = \mathbb{S}_p[\mathcal{H}_2]\,|_{\mathcal{B}}\ .$$

Since $\bar{A}$ is the extension of $A \in \mathbb{S}_p[\mathcal{B}]$, we can define A on $\mathcal{B}$ by

$$A = \sum\nolimits_{n=1}^{\infty} \mu_n(A)\left\langle \cdot\ , \varphi_n^*\right\rangle U\varphi_n,$$

where φ_n^* is the unique dual map in $\mathcal{B}'$ associated with φ_n and U is the restriction of $\bar{U}$ to $\mathcal{B}$. The corresponding norm of A on $\mathbb{S}_p[\mathcal{B}]$ is defined by:

$$\|A\|_p^{\mathcal{B}} = \left\{\sum\nolimits_{n=1}^{\infty}\left\langle A^*A\varphi_n, \varphi_n^*\right\rangle^{p/2}\right\}^{1/p}.$$

Theorem 5.58. *Let $A \in \mathbb{S}_p[\mathcal{B}]$, then $\|A\|_p^{\mathcal{B}} = \left\|\bar{A}\right\|_p^{\mathcal{H}_2}$.*

Proof. It is clear that $\{\varphi_n \,|n \geqslant 1\}$ is a set of eigenfunctions for A^*A on $\mathcal{B}$. Furthermore, by Theorem 5.11, A^*A is naturally self-adjoint and, since every compact operator generates a C_0-semigroup, by Theorem 5.40, the spectrum of A^*A is unchanged by its extension to $\mathcal{H}_2$. It follows that $A^*A\varphi_n = |\mu_n(A)|^2\,\varphi_n$, so that

$$\left\langle A^*A\varphi_n, \varphi_n^*\right\rangle = |\mu_n|^2\left\langle\varphi_n, \varphi_n^*\right\rangle = |\mu_n(A)|^2\,,$$

and

$$\|A\|_p^{\mathcal{B}} = \left\{\sum\nolimits_{n=1}^{\infty}\left\langle A^*A\varphi_n, \varphi_n^*\right\rangle^{p/2}\right\}^{1/p} = \left\{\sum\nolimits_{n=1}^{\infty}|\mu_n(A)|^p\right\}^{1/p} = \left\|\bar{A}\right\|_p^{\mathcal{H}_2}.$$

□

It is clear that all of the theory of operator ideals on Hilbert spaces extend to uniformly convex Banach spaces with an S-basis in a straightforward way. We state a few of the more important results to give a sense of the power provided by the existence of adjoints. The first result extends theorems due to Weyl [WY], Horn [HO], Lalesco [LE] and Lidskii [LI]. The proofs are all straightforward, for a given

A extend it to $\mathcal{H}_2$, use the Hilbert space result and then restrict back to $\mathcal{B}$.

Theorem 5.59. *Let $A \in \mathbb{K}(\mathcal{B})$, the set of compact operators on $\mathcal{B}$, and let $\{\lambda_n\}$ be the eigenvalues of A counted up to algebraic multiplicity. If Φ is a mapping on $[0, \infty]$ which is nonnegative and monotone increasing, then we have:*

(1) *(Weyl)*

$$\sum_{n=1}^{\infty} \Phi\left(|\lambda_n(A)|\right) \leqslant \sum_{n=1}^{\infty} \Phi\left(\mu_n(A)\right)$$

and

(2) *(Horn) If $A_1,\ A_2 \in \mathbb{K}(\mathcal{B})$*

$$\sum_{n=1}^{\infty} \Phi\left(|\lambda_n(A_1 A_2)|\right) \leqslant \sum_{n=1}^{\infty} \Phi\left(\mu_n(A_1)\mu_n(A_2)\right).$$

In case $A \in \mathbb{S}_1(\mathcal{B})$, we have:

(3) *(Lalesco)*

$$\sum_{n=1}^{\infty} |\lambda_n(A)| \leqslant \sum_{n=1}^{\infty} \mu_n(A)$$

and

(4) *(Lidskii)*

$$\sum_{n=1}^{\infty} \lambda_n(A) = Tr(A).$$

Simon [SI1] provides a very nice approach to infinite determinants and trace class operators on separable Hilbert spaces. He gives a comparative historical analysis of Fredholm theory, obtaining a new proof of Lidskii's Theorem as a side benefit and some new insights. A review of his paper shows that much of it can be directly extended to operator theory on separable reflexive Banach spaces.

5.6.3. Discussion. On a Hilbert space $\mathcal{H}$, the Schatten classes $\mathbb{S}_p(\mathcal{H})$ are the only ideals in $\mathbb{K}(\mathcal{H})$, and $\mathbb{S}_1(\mathcal{H})$ is minimal. In a general Banach space, this is far from true. A complete history of the subject can be found in the recent book by Pietsch [PI1] (see also Retherford [RE], for a nice review). We limit this discussion to a few major topics in the subject. First, Grothendieck [GO] defined an important class of nuclear operators as follows:

Definition 5.60. If $A \in \mathbb{F}(\mathcal{B})$ (the operators of finite rank), define the ideal $\mathbf{N}_1(\mathcal{B})$ by:

$$\mathbf{N}_1(\mathcal{B}) = \{A \in \mathbb{F}(\mathcal{B}) \mid \mathbf{N}_1(A) < \infty\},$$

where

$$\mathbf{N}_1(A) = \text{glb}\left\{\sum\nolimits_{n=1}^{m} \|f_n\| \, \|\phi_n\| \; \middle| f_n \in \mathcal{B}', \; \phi_n \in \mathcal{B}, \; A = \sum\nolimits_{n=1}^{m} \phi_n \langle \cdot \, , f_n \rangle \right\}$$

and the greatest lower bound is over all possible representations for A.

Grothendieck showed that $\mathbf{N}_1(\mathcal{B})$ is the completion of the finite rank operators and is a Banach space with norm $\mathbf{N}_1(\cdot)$. It is also a two-sided ideal in $\mathbb{K}(\mathcal{B})$. It is easy to show that:

Corollary 5.61. $\mathbb{M}(\mathcal{B}), \mathbb{N}(\mathcal{B})$ *and* $\mathbf{N}_1(\mathcal{B})$ *are two-sided *ideals.*

In order to compensate for the (apparent) lack of an adjoint for Banach spaces, Pietsch [PI2], [PI3] defined a number of classes of operator ideals for a given $\mathcal{B}$. Of particular importance for our discussion is the class $\mathbb{C}_p(\mathcal{B})$, defined by

$$\mathbb{C}_p(\mathcal{B}) = \left\{ A \in \mathbb{K}(\mathcal{B}) \; \middle| \; \mathbb{C}_p(A) = \sum\nolimits_{i=1}^{\infty} [s_i(A)]^p < \infty \right\},$$

where the singular numbers $s_n(A)$ are defined by:

$$s_n(A) = \inf \left\{ \|A - K\|_{\mathcal{B}} \mid \text{rank of } K \leqslant n \right\}.$$

Pietsch has shown that $\mathbb{C}_1(\mathcal{B}) \subset \mathbf{N}_1(\mathcal{B})$, while Johnson et al. [JKMR] have shown that for each $A \in \mathbb{C}_1(\mathcal{B})$, $\sum_{n=1}^{\infty} |\lambda_n(A)| < \infty$. On the other hand, Grothendieck [GO] has provided an example of an operator A in $\mathbf{N}_1(L^{\infty}[0,1])$ with $\sum_{n=1}^{\infty} |\lambda_n(A)| = \infty$ (see Simon [SI], p. 118). Thus, it follows that, in general, the containment is strict. It is known that if $\mathbb{C}_1(\mathcal{B}) = \mathbf{N}_1(\mathcal{B})$, then $\mathcal{B}$ is isomorphic to a Hilbert space (see Johnson et al.). It is clear from the above discussion that:

Corollary 5.62. $\mathbb{C}_p(\mathcal{B})$ *is a two-sided *ideal in* $\mathbb{K}(\mathcal{B})$*, and* $\mathbb{S}_1(\mathcal{B}) \subset \mathbf{N}_1(\mathcal{B})$.

For a given Banach space, it is not clear how the spaces $\mathbb{C}_p(\mathcal{B})$ of Pietsch relate to our Schatten Classes $\mathbb{S}_p(\mathcal{B})$ (clearly $\mathbb{S}_p(\mathcal{B}) \subseteq \mathbb{C}_p(\mathcal{B})$). Thus, one question is that of the equality of $\mathbb{S}_p(\mathcal{B})$ and $\mathbb{C}_p(\mathcal{B})$. (We suspect that $\mathbb{S}_1(\mathcal{B}) = \mathbb{C}_1(\mathcal{B})$.)

Remark 5.63. In closing, we should point out that if $\mathcal{B}$ is not uniformly convex, then for a given $\phi \in \mathcal{B}$ the set $J(\phi) \in \mathcal{B}'$ can be multivalued and there is no unique way to define $\mathbb{S}_p(\mathcal{B})$ (i.e., to choose $\phi^* \in J(\phi)$). If $\mathcal{B}'$ is strictly convex, $J(\phi) \in \mathcal{B}'$ is uniquely defined (single-valued), so that all of our results still hold. However, to our knowledge, all known examples Banach spaces with $\mathcal{B}'$ strictly convex are uniformly convex.

Conclusion. The most interesting aspect of this section is the observation that the dual space of a Banach space can have more than one representation. It is well known that a given Banach space $\mathcal{B}$ can have many equivalent norms that generate the same topology. However, the geometric properties of the space depend on the norm used. We have shown that the properties of the linear operators on $\mathcal{B}$ depend on the family of linear functionals used to represent the dual space $\mathcal{B}'$. This approach offers an interesting tool for a closer study of the structure of bounded linear operators on $\mathcal{B}$.

References

[B] V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*. Springer Monographs in Mathematics (Springer, New York, 2010)

[CF] I. Colojoară, C. Foiaş, *Theory of Generalized Spectral Operators* (Gordon Breach, London, 1968)

[DA] E.B. Davies, *The Functional Calculus*. J. Lond. Math. Soc. **52**, 166–176 (1995)

[DI] J. Diestel, *Sequences and Series in Banach Spaces*. Graduate Texts in Mathematics (Springer, New York, 1984)

[DS] N. Dunford, J.T. Schwartz, *Linear Operators Part I: General Theory*, Wiley Classics edn. (Wiley Interscience, New York, 1988)

[EN] K.J. Engel, R. Nagel, et al., *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics, vol. 194 (Springer, New York, 2000)

[GO] A. Grothendieck, Products tensoriels topologiques et espaces nucleaires. Mem. Am. Math. Soc. **16**, 1–140 (1955)

[GRA] L. Grafakos, *Classical and Modern Fourier Analysis* (Pearson Prentice-Hall, New Jersey, 2004)

[GS] J.A. Goldstein, *Semigroups of Linear Operators and Applications* (Oxford University Press, New York, 1985)

[HHK] H.H. Kuo, *Gaussian Measures in Banach Spaces.* Lecture Notes in Mathematics, vol. 463 (Springer, New York, 1975)

[HO] A. Horn, On the singular values of a product of completely continuous operators. Proc. Natl. Acad. Sci. **36**, 374–375 (1950)

[HP] E. Hille, R.S. Phillips, *Functional Analysis and Semigroups.* American Mathematical Society Colloquium Publications, vol. 31 (American Mathematical Society, Providence, RI, 1957)

[HS] R. Henstock, *The General Theory of Integration* (Clarendon Press, Oxford, 1991)

[HSJ] B. Helffer, J. Sjöstrand, in *Équation de Schrödinger avec champ magnetique et équation de Harper, Schrödinger Operators* (Snderborg, 2988), ed. by H. Holden, A. Jensen. Lecture Notes in Physics, vol. 345 (Springer, Berlin, 1989), pp. 118–197

[JKMR] W.B. Johnson, H. Konig, B. Maurey, J.R. Retherford, Eigenvalues of p-summing and l_p type operators in Banach space. J. Funct. Anal. **32**, 353–380 (1978)

[K] T. Kato, *Perturbation Theory for Linear Operators,* 2nd edn. (Springer, New York, 1976)

[K1] T. Kato, Trotters product formula for an arbitrary pair of selfadjoint contraction semigroups, in *Advances in Mathematics: Supplementary Studies*, vol. 3 (Academic, New York, 1978), pp. 185–195

[KA] S. Kakutani, On equivalence of infinite product measures. Ann. Math. **49**, 214–224 (1948)

[LE] T. Lalesco, Une theoreme sur les noyaux composes. Bull. Acad. Sci. **3**, 271–272 (1914/1915)

[LI] V.B. Lidskii, Non-self adjoint operators with a trace. Dokl. Akad. Nauk. SSSR **125**, 485–487 (1959)

[LP] G. Lumer, R.S. Phillips, Dissipative operators in a Banach space. Pacific J. Math. **11**, 679–698 (1961)

[LU] G. Lumer, Spectral operators, Hermitian operators and bounded groups. Acta. Sci. Math. (Szeged) **25**, 75–85 (1964)

[PI1] A. Pietsch, *History of Banach Spaces and Operator Theory* (Birkhäuser, Boston, 2007)

[PI2] A. Pietsch, Einige neue Klassen von kompacter linear Abbildungen. Revue der Math. Pures et Appl. (Bucharest) **8**, 423–447 (1963)

[PI3] A. Pietsch, *Eigenvalues and s-Numbers* (Cambridge University Press, Cambridge, 1987)

[PL] T.W. Palmer, Unbounded normal operators on Banach spaces. Trans. Am. Math. Sci. **133**, 385–414 (1968)

[PZ] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Applied Mathematical Sciences, vol. 44 (Springer, New York, 1983)

[RE] J.R. Retherford, Applications of Banach ideals of operators. Bull. Am. Math. Soc. **81**, 978–1012 (1975)

[SI] B. Simon, *Trace Ideals and Their Applications.* London Mathematical Society Lecture Notes Series, vol. 35 (Cambridge University Press, New York, 1979)

[SI1] B. Simon, Notes on infinite determinants of Hilbert space operators. Adv. Math. **24**, 244–273 (1977)

[STE] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions* (Princeton University Press, Princeton, 1970)

[STT] V.A. Strauss, C. Trunk, Spectralizable operators. Integr. Equ. Oper. Theory **61**, 413–422 (2008)

[VPP] V.A. Vinokurov, Yu. Petunin, A.N. Pliczko, Measurability and regularizability mappings inverse to continuous linear operators (in Russian). Mat. Zametki. **26**(4), 583–591 (1979). English translation: Math. Notes **26**, 781–785 (1980)

[VR] I. Vrabie, *C_0-Semigroups and Applications.* North-Holland Mathematics Studies, vol. 191 (Elsevier, New York, 2002)

[WY] H. Weyl, Inequalities between the two kinds of eigenvalues of a linear transformation. Proc. Natl. Acad. Sci. **35**, 408–411 (1949)

[YS] K. Yosida, *Functional Analysis*, 2nd edn. (Springer, New York, 1968)

Chapter 6

Spaces of von Neumann Type

This chapter develops the mathematical foundations for the time-evolution of a physical systems as a three-dimensional motion picture (time-ordering). Our objective is to construct the mathematical version of a physical film on which space-time events can evolve. We first construct the film using infinite tensor products of Hilbert spaces, which is natural for physics. Although von Neumann [VN2] did not develop his theory for our purpose, it will be clear that it is natural for our approach. This film, as a Hilbert space, will be used as the ambient space in Chap. 7 for the Feynman (time-ordered) operator calculus. In order to make the theory available for applications beyond physics, we extend von Neumann's method to construct infinite tensor products of Banach spaces. (This approach makes it easy to transfer the operator calculus to the Banach space setting.) We assume that the reader has read Sect. 1.4 of Chap. 1. This section provides a fairly complete introduction to the finite tensor product theory for both Hilbert and Banach spaces.

Summary. In the first section, we begin with a study of continuous tensor products of Hilbert spaces. Since von Neumann's approach is central to our theory and this subject is not discussed in the standard analysis/functional analysis programs, we provide a fairly complete

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_6

exposition. In the second section, we use von Neumann's approach to construct continuous tensor products of Banach spaces. After a few examples in the third section, we discuss operators on continuous tensor products of Banach spaces in the fourth section. In the fifth section we construct our mathematical version of a physical film or Feynman–Dyson space, which is the primary motivation for this chapter. In the last section we define a special operator, called the exchange operator, which will be used in the proof of a generalized version of the Feynman–Kac formula and the interaction representation in the next chapter.

The Appendix (Sect. 6.7) is devoted to a few important applications to infinite dimensional analysis. Here, we discuss a general approach to the Fourier transform, which applies to all Banach spaces with an S-basis. We then discuss infinite sums and products of unbounded operators and use our results to provide a constructive approach to a number of operators in infinitely many variables.

6.1. Infinite Tensor Product Hilbert Spaces

Let $I = [a, b]$, $0 \le a < b \le \infty$, and, in order to avoid trivialities, we always assume that, in any product, all terms are nonzero.

Definition 6.1. If $\{z_\nu\}$ is a sequence of complex numbers indexed by $\nu \in I$,

(1) We say that the product $\prod_{\nu\in I} z_\nu$ is convergent with limit z if, for every $\varepsilon > 0$, there is a finite set $J(\varepsilon)$ such that, for all finite sets $J \subset I$, with $J(\varepsilon) \subset J$, we have $\left|\prod_{\nu\in J} z_\nu - z\right| < \varepsilon$.

(2) We say that the product $\prod_{\nu\in I} z_\nu$ is quasi-convergent if $\prod_{\nu\in I} |z_\nu|$ is convergent. (If the product is quasi-convergent, but not convergent, we assign it the value zero.)

Since I is not countable, we note that

$$0 < \left|\prod\nolimits_{\nu\in I} z_\nu\right| < \infty \text{ if and only if } \sum\nolimits_{\nu\in I} |1 - z_\nu| < \infty.$$

Thus, it follows that convergence implies that at most a countable number of the $z_\nu \neq 1$.

Let $\mathcal{H}_\nu = \mathcal{H}$ be a fixed Hilbert space for each $\nu \in I$ and, for $\{\phi_\nu\} \in \prod_{\nu\in I} \mathcal{H}_\nu$, let Δ_I be those sequences $\{\phi_\nu\}$ such that $\sum_{\nu\in I} |\|\varphi_\nu\|_\nu - 1| < \infty$. Define a functional on Δ_I by

$$\Phi(\psi) = \sum\nolimits_{k=1}^{n} \prod_{\nu\in I} \left\langle \varphi_\nu^k, \psi_\nu \right\rangle_\nu,$$

where $\psi = \{\psi_\nu\}, \{\varphi_\nu^k\} \in \Delta_I$, for $1 \le k \le n$. It is easy to see that this functional is linear in each component. Denote Φ by

$$\Phi = \sum\nolimits_{k=1}^{n} \otimes_{\nu \in I} \varphi_\nu^k.$$

Define the algebraic tensor product, $\otimes_{\nu\in I}\mathcal{H}_\nu$, by

$$\otimes_{\nu\in I}\mathcal{H}_\nu = \left\{ \sum\nolimits_{k=1}^{n} \otimes_{\nu\in I}\varphi_\nu^k \,\middle|\, \{\varphi_\nu^k\} \in \Delta_I, 1 \le k \le n, n \in \mathbb{N} \right\}.$$

We define a linear functional on $\otimes_{\nu\in I}\mathcal{H}_\nu$ by

$$\left(\sum\nolimits_{k=1}^{n} \otimes_{\nu\in I}\varphi_\nu^k, \sum\nolimits_{l=1}^{m} \otimes_{\nu\in I}\psi_\nu^l \right)_\otimes = \sum\nolimits_{l=1}^{m} \sum\nolimits_{k=1}^{n} \prod_{\nu\in I} \left\langle \varphi_\nu^k, \psi_\nu^l \right\rangle_\nu.$$

Lemma 6.2. *The functional* $(\cdot,\cdot)_\otimes$ *is a well-defined mapping on* $\otimes_{\nu\in I}\mathcal{H}_\nu$.

Proof. It suffices to show that, if $\Phi = 0$, then $(\Phi,\Psi)_\otimes = 0$. If $\Phi = \sum_{k=1}^{n} \otimes_{\nu\in I}\varphi_\nu^k$ and $\Psi = \sum_{l=1}^{m} \otimes_{\nu\in I}\psi_\nu^l$, then, with $\psi_l = \{\psi_\nu^l\}$,

$$(\Phi,\Psi)_\otimes = \sum\nolimits_{l=1}^{m} \sum\nolimits_{k=1}^{n} \prod_{\nu\in I} \left\langle \varphi_\nu^k, \psi_\nu^l \right\rangle_\nu = \sum\nolimits_{l=1}^{m} \Phi(\psi_l) = 0.$$

□

Before continuing our discussion of the above functional, we first need to look a little more closely at the structure of the algebraic tensor product space, $\otimes_{\nu\in I}\mathcal{H}_\nu$.

Definition 6.3. Let $\phi = \underset{\nu\in I}{\otimes} \phi_\nu$ and $\psi = \underset{\nu\in I}{\otimes} \psi_\nu$ be in $\otimes_{\nu\in I}\mathcal{H}_\nu$.

(1) We say that ϕ is strongly equivalent to ψ ($\phi \equiv^s \psi$) if and only if $\sum_{\nu\in I} |1 - \langle \phi_\nu, \psi_\nu \rangle_\nu| < \infty$.

(2) We say that ϕ is weakly equivalent to ψ ($\phi \equiv^w \psi$) if and only if $\sum_{\nu\in I} |1 - |\langle \phi_\nu, \psi_\nu \rangle_\nu|| < \infty$.

Lemma 6.4. *We have* $\phi \equiv^w \psi$ *if and only if there exist* z_ν, $|z_\nu| = 1$, *such that* $\underset{\nu\in I}{\otimes} z_\nu\phi_\nu \equiv^s \underset{\nu\in I}{\otimes} \psi_\nu$.

Proof. Suppose that $\underset{\nu\in I}{\otimes} z_\nu\phi_\nu \equiv^s \underset{\nu\in I}{\otimes} \psi_\nu$. Then we have:

$$\sum_{\nu\in I} \left|1 - \left|\langle \phi_\nu, \psi_\nu \rangle_\nu\right|\right| = \sum_{\nu\in I} \left|1 - \left|\langle z_\nu\phi_\nu, \psi_\nu \rangle_\nu\right|\right| \leqslant \sum_{\nu\in I} \left|1 - \langle z_\nu\phi_\nu, \psi_\nu \rangle_\nu\right| < \infty.$$

If $\phi \equiv^w \psi$, set

$$z_\nu = |\langle \phi_\nu, \psi_\nu \rangle_\nu| / \langle \phi_\nu, \psi_\nu \rangle_\nu$$

for $\langle \phi_\nu, \psi_\nu \rangle_\nu \neq 0$, and set $z_\nu = 1$ otherwise. It follows that

$$\sum_{\nu \in I} |1 - \langle z_\nu \phi_\nu, \psi_\nu \rangle_\nu| = \sum_{\nu \in I} |1 - |\langle \phi_\nu, \psi_\nu \rangle_\nu|| < \infty,$$

so that $\underset{\nu \in I}{\otimes} z_\nu \phi_\nu \equiv^s \underset{\nu \in I}{\otimes} \psi_\nu$. □

Theorem 6.5. *The relations defined above are equivalence relations on $\otimes_{\nu \in I} \mathcal{H}_\nu$, which decomposes $\otimes_{\nu \in I} \mathcal{H}_\nu$ into disjoint equivalence classes.*

Proof. Suppose $\underset{\nu \in I}{\otimes} \phi_\nu \equiv^s \underset{\nu \in I}{\otimes} \psi_\nu$. First note that the relation is clearly reflexive. Thus, we need to only prove that it is symmetric and transitive. To prove that the first relation is symmetric, observe that $|1 - \langle \psi_\nu, \phi_\nu \rangle_\nu| = \left|1 - \overline{\langle \phi_\nu, \psi_\nu \rangle_\nu}\right| = \left|\overline{[1 - \langle \phi_\nu, \psi_\nu \rangle_\nu]}\right| = |1 - \langle \phi_\nu, \psi_\nu \rangle_\nu|$. To show that it is transitive, without loss, we can assume that $\|\psi_\nu\|_\nu = \|\phi_\nu\|_\nu = 1$. It is then easy to see that, if $\underset{\nu \in I}{\otimes} \phi_\nu \equiv^s \underset{\nu \in I}{\otimes} \psi_\nu$ and $\underset{\nu \in I}{\otimes} \psi_\nu \equiv^s \underset{\nu \in I}{\otimes} \rho_\nu$, then

$$1 - \langle \phi_\nu, \rho_\nu \rangle_\nu = [1 - \langle \phi_\nu, \psi_\nu \rangle_\nu] + [1 - \langle \psi_\nu, \rho_\nu \rangle_\nu] + \langle \phi_\nu - \psi_\nu, \psi_\nu - \rho_\nu \rangle_\nu .$$

Now $\langle \phi_\nu - \psi_\nu, \phi_\nu - \psi_\nu \rangle_\nu = 2\,[1 - \operatorname{Re} \langle \phi_\nu, \psi_\nu \rangle_\nu] \leqslant 2\,|1 - \langle \phi_\nu, \psi_\nu \rangle_\nu|$, so that $\sum_\nu \|\phi_\nu - \psi_\nu\|_\nu^2 < \infty$ and, by the same observation, $\sum_\nu \|\psi_\nu - \rho_\nu\|_\nu^2 < \infty$. It now follows from Schwartz's inequality that $\sum_\nu \|\phi_\nu - \psi_\nu\|_\nu \|\psi_\nu - \rho_\nu\|_\nu < \infty$. Thus we have that

$$\begin{aligned} &\sum_{\nu \in I} |1 - \langle \phi_\nu, \rho_\nu \rangle_\nu| \\ &\leq \sum_{\nu \in I} |1 - \langle \phi_\nu, \psi_\nu \rangle_\nu| + \sum_{\nu \in I} |1 - \langle \psi_\nu, \rho_\nu \rangle_\nu| \\ &+ \sum_{\nu \in I} \|\phi_\nu - \psi_\nu\|_\nu \|\psi_\nu - \rho_\nu\|_\nu < \infty. \end{aligned}$$

This proves the first case. The proof of the second case (weak equivalence) now follows from the above lemma. □

Theorem 6.6. *Let $\otimes_{\nu\in I}\varphi_\nu$ be in $\otimes_{\nu\in I}\mathcal{H}_\nu$. Then:*

(1) *The product $\prod_{\nu\in I}\|\varphi_\nu\|_\nu$ converges if and only if $\prod_{\nu\in I}\|\varphi_\nu\|_\nu^2$ converges.*

(2) *If $\prod_{\nu\in I}\|\varphi_\nu\|_\nu$ and $\prod_{\nu\in I}\|\psi_\nu\|_\nu$ converge, then $\prod_{\nu\in I}\langle\varphi_\nu,\psi_\nu\rangle_\nu$ is quasi-convergent.*

(3) *If $\prod_{\nu\in I}\langle\varphi_\nu,\psi_\nu\rangle_\nu$ is quasi-convergent, then there exist complex numbers $\{z_\nu\}$, $|z_\nu|=1$, such that $\prod_{\nu\in I}\langle z_\nu\varphi_\nu,\psi_\nu\rangle_\nu$ converges.*

Proof. For the first case, convergence of either term implies that $\{\|\varphi_\nu\|_\nu\,,\ \nu\in I\}$ has a finite upper bound $M>0$. Hence

$$\left|1-\|\varphi_\nu\|_\nu\right| \leqslant \left|1+\|\varphi_\nu\|_\nu\right|\left|1-\|\varphi_\nu\|_\nu\right| = \left|1-\|\varphi_\nu\|_\nu^2\right| \leqslant (1+M)\left|1-\|\varphi_\nu\|_\nu\right|.$$

To prove (2), note that if $J\subset I$ is any finite subset,

$$0\leqslant\left|\prod_{\nu\in J}\langle\varphi_\nu,\psi_\nu\rangle_\nu\right|\leqslant\prod_{\nu\in J}\|\varphi_\nu\|_\nu\prod_{\nu\in J}\|\psi_\nu\|_\nu<\infty.$$

Therefore, $0\leqslant\left|\prod_{\nu\in I}\langle\varphi_\nu,\psi_\nu\rangle_\nu\right|<\infty$ so that $\prod_{\nu\in I}\langle\varphi_\nu,\psi_\nu\rangle_\nu$ is quasi-convergent and, if $0<\left|\prod_{\nu\in I}\langle\varphi_\nu,\psi_\nu\rangle_\nu\right|<\infty$, it is convergent. The proof of (3) now follows directly from the above lemma. □

Definition 6.7. For $\varphi=\underset{\nu\in I}{\otimes}\varphi_\nu\in\otimes_{\nu\in I}\mathcal{H}_\nu$, we define $\mathcal{H}_\otimes^2(\varphi)$ to be the closed subspace generated by the span of all $\psi\equiv^s\varphi$ and we call it the strong partial tensor product space generated by the vector φ.

Theorem 6.8. *For the partial tensor product spaces, we have the following:*

(1) *If $\psi_\nu\neq\varphi_\nu$ occurs for at most a finite number of ν, then $\psi=\underset{\nu\in I}{\otimes}\psi_\nu\equiv^s\varphi=\underset{\nu\in I}{\otimes}\varphi_\nu$.*

(2) *The space $\mathcal{H}_\otimes^2(\varphi)$ is the closure of the linear span of $\psi=\underset{\nu\in I}{\otimes}\psi_\nu$ such that $\psi_\nu\neq\varphi_\nu$ occurs for at most a finite number of ν.*

(3) *If $\Phi=\otimes_{\nu\in I}\varphi_\nu$ and $\Psi=\otimes_{\nu\in I}\psi_\nu$ are in different equivalence classes of $\otimes_{\nu\in I}\mathcal{H}_\nu$, then $(\Phi,\Psi)_\otimes=\prod_{\nu\in I}\langle\varphi_\nu,\psi_\nu\rangle_\nu=0$.*

(4) $\mathcal{H}_\otimes^2(\varphi)^w=\underset{\psi\equiv^w\phi}{\oplus}\left[\mathcal{H}_\otimes^2(\psi)^s\right].$

Proof. To prove (1), let J be the finite set of ν for which $\psi_\nu \neq \varphi_\nu$. Then

$$\begin{aligned}
&\sum_{\nu\in I} |1-\langle\varphi_\nu,\psi_\nu\rangle_\nu| \\
&= \sum_{\nu\in J} |1-\langle\varphi_\nu,\psi_\nu\rangle_\nu| + \sum_{\nu\in I\setminus J} |1-\langle\varphi_\nu,\varphi_\nu\rangle_\nu| \\
&\leqslant c + \sum_{\nu\in I} \left|1-\|\varphi_\nu\|_\nu^2\right| < \infty,
\end{aligned}$$

so that $\underset{\nu\in I}{\otimes}\psi_\nu \equiv \underset{\nu\in I}{\otimes}\varphi_\nu$.

To prove (2), let $\mathcal{H}^2_\otimes(\varphi)^\#$ be the closure of the linear span of all $\psi = \otimes_{\nu\in I}\psi_\nu$ such that $\psi_\nu \neq \varphi_\nu$ occurs for at most a finite number of ν. There is no loss in assuming that $\|\varphi_\nu\|_\nu = 1$ for all $\nu \in I$. It is clear from (1) that $\mathcal{H}^2_\otimes(\varphi)^\# \subseteq \mathcal{H}^2_\otimes(\varphi)$. Thus, we are done if we can show that $\mathcal{H}^2_\otimes(\varphi)^\# \supseteq \mathcal{H}^2_\otimes(\varphi)$. For any vector $\psi = \otimes_{\nu\in I}\psi_\nu$ in $\mathcal{H}^2_\otimes(\varphi)$, $\varphi \equiv \psi$ so that $\sum_{\nu\in I} |1-\langle\varphi_\nu,\psi_\nu\rangle_\nu| < \infty$. If $\|\psi\|_\otimes^2 = 0$ then $\psi \in \mathcal{H}^2_\otimes(\varphi)^\#$, so we can assume that $\|\psi\|_\otimes^2 \neq 0$. This implies that $\|\psi_\nu\|_\nu \neq 0$ for all $\nu \in I$ and $0 \neq \prod_{\nu\in I}(1/\|\psi_\nu\|_\nu) < \infty$; hence, by scaling if necessary, we may also assume that $\|\psi_\nu\|_\nu = 1$ for all $\nu \in I$. Let $0 < \varepsilon < 1$ be given, and choose δ so that $0 < \sqrt{2\delta e} < \varepsilon$ (e is the base for the natural log). Since $\sum_{\nu\in I} |1-\langle\varphi_\nu,\psi_\nu\rangle_\nu| < \infty$, there is a finite set of distinct values $J = \{\nu_1, \cdots, \nu_n\}$ such that $\sum_{\nu\in I-J} |1-\langle\varphi_\nu,\psi_\nu\rangle_\nu| < \delta$. Since, for any finite set of numbers $z_1, \cdots, z_n$, it is easy to see that

$$\left|\prod_{k=1}^n z_k - 1\right| = \left|\prod_{k=1}^n [1+(z_k-1)] - 1\right| \leq \left(\prod_{k=1}^n e^{|z_k-1|} - 1\right),$$

we have that

$$\left|\prod_{\nu\in I\setminus J} \langle\varphi_\nu,\psi_\nu\rangle_\nu - 1\right| \leq \left(\exp\left\{\sum_{\nu\in I\setminus J} |\langle\varphi_\nu,\psi_\nu\rangle_\nu - 1|\right\} - 1\right) \leq e^\delta - 1 \leq e\delta.$$

Now, define $\phi_\nu = \psi_\nu$ if $\nu \in J$, and $\phi_\nu = \varphi_\nu$ if $\nu \in I \backslash J$, and set $\phi_J = \otimes_{\nu \in I} \phi_\nu$ so that $\phi_J \in \mathcal{H}^2_\otimes(\varphi)^\#$ and

$$\begin{aligned}
&\|\psi - \phi_J\|^2_\otimes \\
&= 2 - 2\,\mathrm{Re}\left[\prod_{\nu \in J} \langle \varphi_\nu, \psi_\nu \rangle_\nu \cdot \prod_{\nu \in I-J} \langle \varphi_\nu, \psi_\nu \rangle_\nu\right] \\
&= 2 - 2\,\mathrm{Re}\left[\prod_{\nu \in I} \|\psi_\nu\|^2_\nu \cdot \prod_{\nu \in I-J} \langle \varphi_\nu, \psi_\nu \rangle_\nu\right] \\
&= 2\,\mathrm{Re}\left[1 - \prod_{\nu \in I-J} \langle \varphi_\nu, \psi_\nu \rangle_\nu\right] \leqslant 2e\delta < \varepsilon^2.
\end{aligned}$$

Since ε is arbitrary, ψ is in the closure of $\mathcal{H}^2_\otimes(\varphi)^\#$, so $\mathcal{H}^2_\otimes(\varphi)^\# = \mathcal{H}^2_\otimes(\varphi)$.

To prove (3), first note that if $\prod_{\nu \in I} \|\varphi_\nu\|_\nu$ and $\prod_{\nu \in I} \|\psi_\nu\|_\nu$ converge, then, for any finite subset $J \subset I$,

$$0 \le \left| \prod_{\nu \in J} \langle \varphi_\nu, \psi_\nu \rangle_\nu \right| \le \prod_{\nu \in J} \|\varphi_\nu\|_\nu \prod_{\nu \in J} \|\psi_\nu\|_\nu < \infty.$$

Therefore, $0 \le \left|\prod_{\nu \in I} \langle \varphi_\nu, \psi_\nu \rangle_\nu\right| = \left|(\Phi, \Psi)_\otimes\right| < \infty$ so that $\prod_{\nu \in I} \langle \varphi_\nu, \psi_\nu \rangle_\nu$ is convergent or zero. If $0 < \left|(\Phi, \Psi)_\otimes\right| < \infty$, then $\sum\limits_{\nu \in I} |1 - \langle \phi_\nu, \psi_\nu \rangle_\nu| < \infty$ and, by definition, Φ and Ψ are in the same equivalence class, so we must have $\left|(\Phi, \Psi)_\otimes\right| = 0$. The proof of (4) follows from the definition of weakly equivalent spaces. □

Theorem 6.9. $(\Phi, \Psi)_\otimes$ *is a conjugate bilinear positive definite functional.*

Proof. The first part is trivial. To prove that it is positive definite, let $\Phi = \sum_{k=1}^n \otimes_{\nu \in I} \varphi^k_\nu$, and assume that the vectors $\otimes_{\nu \in I} \varphi^k_\nu, 1 \le k \le n$ are in distinct equivalence classes. This means that, with $\Phi_k = \otimes_{\nu \in I} \varphi^k_\nu$, we have

$$\begin{aligned}
&(\Phi, \Phi)_\otimes \\
&= \left(\sum_{k=1}^n \Phi_k, \sum_{k=1}^n \Phi_k\right)_\otimes \\
&= \sum_{k=1}^n \sum_{j=1}^n (\Phi_k, \Phi_j)_\otimes \\
&= \sum_{k=1}^n (\Phi_k, \Phi_k)_\otimes.
\end{aligned}$$

Note that, from Theorem 6.8 (3), $k \neq j$ implies $(\Phi_k, \Phi_j)_\otimes = 0$. Thus, it suffices to assume that $\otimes_{\nu \in I}\varphi_\nu^k$, $1 \le k \le n$, are all in the same equivalence class. In this case, we have that

$$(\Phi, \Phi)_\otimes = \sum\nolimits_{k=1}^n \sum\nolimits_{j=1}^n \prod_{\nu \in I} \left\langle \varphi_\nu^k, \varphi_\nu^j \right\rangle_\nu,$$

where each product is convergent. It follows that the above will be positive definite if we can show that, for all possible finite sets $J = \{\nu_1, \nu_2 \cdots, \nu_m\}, m \in \mathbb{N}$,

$$\sum\nolimits_{k=1}^n \sum\nolimits_{j=1}^n \prod_{\nu \in J} \left\langle \varphi_\nu^k, \varphi_\nu^j \right\rangle_\nu \ge 0.$$

This is equivalent to showing that the above defines a positive definite functional on $\otimes_{\nu \in J}\mathcal{H}_\nu$, which follows from the standard result for finite tensor products of Hilbert spaces (see Reed and Simon [RS]). □

Definition 6.10. We define $\mathcal{H}_\otimes^2 = \hat{\otimes}_{\nu \in I}\mathcal{H}_\nu$ to be the completion of the linear space $\otimes_{\nu \in I}\mathcal{H}_\nu$, relative to the inner product $(\cdot,\cdot)_\otimes$.

6.2. Infinite Tensor Product Banach Spaces

In this section, we construct infinite tensor product Banach spaces extending the methods of von Neumann [VN2] used in the last section. We call them spaces of von Neumann type in order to distinguish them from a number of other varieties that have been under consideration at various times (see, for example, Guichardet [GU] and Pantsulaia [PA]). For each $t \in I$, let $\mathcal{B}_t$ be a Banach space with an S-basis. We assume that for each t, $\mathcal{H}_t$ with inner product $(\ \cdot\ )_t$ is the canonical Hilbert space constructed from $\mathcal{B}_t$ as in Chap. 5, with $\mathcal{B}_t \subset \mathcal{H}_t$ as a continuous dense embedding.

Using the family $\{\mathcal{H}_t,\ t \in I\}$, construct the continuous tensor product Hilbert space $\mathcal{H}_\otimes^2$ of von Neumann. All vectors $\phi = \underset{t \in I}{\otimes} \phi_t$ under consideration will be either convergent or quasi-convergent. By convention, unless we explicitly say otherwise, ϕ is convergent. Let Δ be the set:

$$\Delta = \left\{ \{\phi_t\} \;\middle|\; 0 \neq \left\| \underset{t \in I}{\otimes} \phi_t \right\|_{\mathcal{H}_\otimes^2} < \infty,\ \sum\nolimits_{t \in I} \left| 1 - \|\phi_t\|_{\mathcal{B}_t} \right| < \infty \right\} \quad (6.1)$$

and define $\underset{t\in I}{\otimes}\mathcal{B}_t$ by:

$$\underset{t\in I}{\otimes}\mathcal{B}_t = \left\{ \phi = \sum_{i=1}^{n} \underset{t\in I}{\otimes}\phi_t^i \,\middle|\, \{\phi_t^i\} \in \Delta,\ i = 1, \cdots, n \right\}.$$

From the definition of $\underset{t\in I}{\otimes}\varphi_t \in \underset{t\in I}{\otimes}\mathcal{B}_t$, we see that the following are well defined:

$$\|\phi\|_\gamma = \inf\left\{ \sum_{k=1}^{m} \prod\nolimits_{t\in I} \left\|\psi_t^k\right\|_{\mathcal{B}_t} \,\middle|\, \sum_{i=1}^{n} \underset{t\in I}{\otimes}\phi_t^i = \sum_{k=1}^{m} \underset{t\in I}{\otimes}\psi_t^k \right\},$$

and

$$\|\phi\|_\lambda = \sup\left\{ \sum_{k=1}^{m} \prod\nolimits_{t\in I} \left|\left\langle \phi_t^i, F_t \right\rangle\right| \,\middle|\, F_t \in \mathcal{B}_t^*,\quad \|F_t\|_{\mathcal{B}_t^*} \leqslant 1,\ \text{for all } t \in I \right\}$$
$$= \sup_{\prod_{t\in I}\|F_t\|_{\mathcal{B}_t^*}\leqslant 1} \left\{ \sum_{k=1}^{m} \prod\nolimits_{t\in I} \left|\left\langle \phi_t^i, F_t \right\rangle\right| \,\middle|\, F_t \in \mathcal{B}_t^*,\ \text{for all } t \in I \right\}.$$

Theorem 6.11. *For all $\phi \in \underset{t\in I}{\otimes}\mathcal{B}_t$, we have:*

(1) *$\|\cdot\|_\lambda$ and $\|\cdot\|_\gamma$ define norms on $\underset{t\in I}{\otimes}\mathcal{B}_t$ with $\|\phi\|_2 \leqslant \|\phi\|_\lambda \leqslant \|\phi\|_\gamma$.*

(2) *If $\phi = \underset{t\in I}{\otimes}\phi_t$, then $\|\phi\|_\lambda = \|\phi\|_\gamma = \prod_{t\in I}\|\phi_t\|_{\mathcal{B}}$ (i.e., $\|\cdot\|_\lambda$ and $\|\cdot\|_\gamma$ are crossnorms).*

Proof. It is easy to check that $\|\cdot\|_\lambda$ and $\|\cdot\|_\gamma$ are norms on $\underset{t\in I}{\otimes}\mathcal{B}_t$. By construction, $\|\phi\|_2 \leqslant \|\phi\|_\lambda$ and $\|\phi\|_2 \leqslant \|\phi\|_\gamma$. Since $\left|\left\langle \phi_t^i\,, F_t\right\rangle\right| \leqslant \left\|\phi_t^i\right\|_{\mathcal{B}_t}$ for all $t \in I$, we have that $\lambda(\phi) \leqslant \gamma(\phi)$. Now, it is easy to see that if $\phi = \underset{t\in I}{\otimes}\phi_t$, then $\|\phi\|_\lambda = \|\phi\|_\gamma = \prod_{t\in I}\|\phi_t\|_{\mathcal{B}}$.

The completions of $\underset{t\in I}{\otimes}\mathcal{B}_t$ with respect to $\|\|_\lambda$ and $\|\|_\gamma$ define the spaces $\underset{t\in I}{\hat{\otimes}}{}^\lambda\mathcal{B}_t \subset \underset{t\in I}{\hat{\otimes}}{}^\gamma\mathcal{B}_t$. We can also construct $\underset{t\in I}{\hat{\otimes}}{}^\alpha\mathcal{B}_t$ so that $\underset{t\in I}{\hat{\otimes}}{}^\lambda\mathcal{B}_t \subset \underset{t\in I}{\hat{\otimes}}{}^\alpha\mathcal{B}_t \subset \underset{t\in I}{\hat{\otimes}}{}^\gamma\mathcal{B}_t$ for all crossnorms α with $\lambda \leqslant \alpha \leqslant \gamma$. For a fixed α, we write $\mathcal{B}_\otimes^\alpha = \underset{t\in I}{\hat{\otimes}}{}^\alpha\mathcal{B}_t$. □

Definition 6.12. We call spaces constructed as above, Banach spaces of von Neumann type.

Definition 6.13. Suppose $F = \otimes_{t\in I} F_t \in \otimes_{t\in I}\mathcal{B}_t^*$, and $\phi = \sum_{i=1}^{n} \underset{t\in I}{\otimes} \phi_t^i \in \underset{t\in I}{\otimes^\alpha} \mathcal{B}_t$; set

$$\alpha^*(F) = \|F\|_{\alpha^*} = \sup_{\phi}\left\{\sum_{i=1}^{n}\prod_{t\in I}\frac{\left|\left\langle \phi_t^i, F_t\right\rangle\right|}{\alpha(\phi)}\right\}.$$

We call α^* the dual norm of α.

Theorem 6.14. *Assume that $\lambda \leqslant \alpha \leqslant \gamma$, then:*

(1) *$\gamma^* \leqslant \alpha^* \leqslant \lambda^*$, and each α^* is a crossnorm,*

(2) *λ and γ are uniform.*

Proof. To prove (1), let $F = \otimes_{t\in I} F_t \in \otimes_{t\in I}\mathcal{B}_t^*$. Then, since:

$$\frac{1}{\left\|\sum_{k=1}^{n}\otimes_{t\in I}\phi_t^k\right\|_{\alpha}} \leqslant \frac{1}{\left\|\sum_{k=1}^{n}\otimes_{t\in I}\phi_t^k\right\|_{\lambda}},$$

by definition of $\|\otimes_{t\in I}F_t\|_{\alpha*}$, we have

$$\begin{aligned}\|\otimes_{t\in I}F_t\|_{\alpha*} &= \sup_{\sum_{k=1}^{n}\otimes_{t\in I}\phi_t^k\neq 0}\frac{\left|\sum_{k=1}^{n}\prod_{t\in I}\left\langle \phi_t^k\,,\,F_\nu\right\rangle\right|}{\left\|\sum_{k=1}^{n}\otimes_{t\in I}\phi_t^k\right\|_{\alpha}}\\ &\leqslant \sup_{\sum_{k=1}^{n}\otimes_{t\in I}\phi_t^k\neq 0}\frac{\left|\sum_{k=1}^{n}\prod_{t\in I}\left\langle \phi_t^k\,,\,F_t\right\rangle\right|}{\left\|\sum_{k=1}^{n}\otimes_{t\in I}\phi_t^k\right\|_{\lambda}} = \|\otimes_{t\in I}F_t\|_{\lambda*}\,,\end{aligned}$$

so that, for $\lambda \leqslant \alpha \leqslant \gamma$, we have $\gamma^* \leqslant \alpha^* \leqslant \lambda^*$ ($\gamma^* = \lambda,\ \lambda^* = \gamma$ on the dual spaces). On the other hand, we have

$$\|\otimes_{t\in I}F_t\|_{\alpha*}\,\|\otimes_{t\in I}\phi_t\|_{\alpha} \geqslant \left|\prod_{t\in I}\langle \phi_t\,,\,F_t\rangle\right|, \text{ which implies}$$

$$\|\otimes_{t\in I}F_t\|_{\alpha*}\prod_{t\in I}\|\phi_t\|_t \geqslant \left|\prod_{t\in I}\langle \phi_t\,,\,F_t\rangle\right|. \text{ Thus, we have}$$

$$\|\otimes_{t\in I}F_t\|_{\alpha*} \geqslant \sup_{\otimes_{t\in I}\phi_t}\left|\prod_{t\in I}\frac{\langle \phi_t\,,\,F_t\rangle}{\|\phi_t\|_t}\right| = \prod_{t\in I}\|F_t\|_{\mathcal{B}*}\,,$$

so that $\|\otimes_{t\in I}F_t\|_{\alpha*} \geqslant \prod_{t\in I}\|F_t\|_{\mathcal{B}*}$. Since it is easy to see that $\|\otimes_{t\in I}F_t\|_{\alpha*} \leqslant \prod_{t\in I}\|F_t\|_{\mathcal{B}*}$, it follows that $\|\otimes_{t\in I}F_t\|_{\alpha*} = \prod_{t\in I}\|F_t\|_{\mathcal{B}*}$. Hence, α^* is a crossnorm for $\gamma^* \leqslant \alpha^* \leqslant \lambda^*$. To prove (2), let $T_t : \mathcal{B}_t \to \mathcal{B}_t$ be a bounded linear operator for each $t \in I$ such that $\prod_{t\in I}\|T_t\|_{\mathcal{B}_t} < \infty$. Let T_t^* be the dual of T_t and suppose $F = \otimes_{t\in I}F_t \in \otimes_{t\in I}\mathcal{B}_t^*$ for λ^*. We then have

$$\begin{aligned}
&\left|(\otimes_{t\in I}F_t)\sum\nolimits_{k=1}^n \otimes_{t\in I}T_t\phi_t^k\right| = \left|(\otimes_{t\in I}T_t^*(F_t))\sum\nolimits_{k=1}^n \otimes_{t\in I}\phi_t^k\right| \\
&\leqslant \|\otimes_{t\in I}T_t^*(F_t)\|_{\lambda*}\left\|\sum\nolimits_{k=1}^n \otimes_{t\in I}\phi_t^k\right\|_\lambda \\
&\leqslant \left(\prod_{t\in I}\|F_t\|_{\mathcal{B}*}\right)\left(\prod_{t\in I}\|T_t\|_{\mathcal{B}}\right)\left\|\sum\nolimits_{k=1}^n \otimes_{t\in I}\phi_t^k\right\|_\lambda.
\end{aligned}$$

Since $\|\otimes_{t\in I}F_t\|_{\lambda*} = \prod_{t\in I}\|F_t\|_{\mathcal{B}*}$, we must have:

$$\left\|\sum\nolimits_{k=1}^n \otimes_{t\in I}T_t\phi_t^k\right\|_\lambda \leqslant \left(\prod_{t\in I}\|T_t\|_{\mathcal{B}}\right)\left\|\sum\nolimits_{k=1}^n \otimes_{t\in I}\phi_t^k\right\|_\lambda,$$

so that λ is uniform. To see that γ is also uniform, note that the same proof for λ applies to γ by symmetry ($\lambda^* = \gamma$) so that

$$\left\|\sum\nolimits_{k=1}^n \otimes_{t\in I}T_t\phi_t^k\right\|_\gamma \leqslant \left(\prod_{t\in I}\|T_t\|_{\mathcal{B}}\right)\left\|\sum\nolimits_{k=1}^n \otimes_{t\in I}\phi_t^k\right\|_\gamma.$$

□

6.2.1. Partial Tensor Product Spaces of Type v. In this section, we construct a class of subspaces that are induced by equivalence relations on $\mathcal{B}_\otimes^\alpha$. The motivation and the construction are the same as the von Neumann's decomposition of the infinite tensor product Hilbert space into orthogonal subspaces of the last section. We assume that $\mathcal{B}_t \subset \mathcal{H}_t$ has an S-basis. Since every $\phi' \in \mathcal{H}_t'$ induces a linear functional on $\mathcal{B}_t$, for each $\phi \in \mathcal{B}_t$, we see that $\mathcal{H}_t' \subset \mathcal{B}_t'$. It is clear that $\mathcal{H}_t'$ is a continuous embedding in $\mathcal{B}_t'$ (which may not be dense), so that $\|(\cdot,\phi)\|_{\mathcal{B}_t'} \le \|(\cdot,\phi)\|_{\mathcal{H}_t'} = \|\phi\|_{\mathcal{H}_t}$. Define ϕ_t^* by

$$\phi_t^* = (\cdot\,,\,\phi_t)_t\left(\frac{\|\phi_t\|_{\mathcal{B}_t}^2}{\|\phi_t\|_{\mathcal{H}_t}^2}\right).$$

It is easy to see that $\langle\phi_t,\phi_t^*\rangle = \|\phi\|_{\mathcal{B}_t}^2$. If $\mathcal{B}$ is uniformly convex, ϕ_t^* is unique. However, if $\mathcal{B}_t$ is not uniformly convex, the set of mappings for a given $\phi_t \in \mathcal{B}_t$ could be uncountable. In this case, we assume a (fixed) choice has been made. The functional ϕ_t^* is called the Steadman map on $\mathcal{B}_t$ associated with $\mathcal{H}_t$. (It is not a duality map.)

Definition 6.15. Let $\phi = \underset{t\in I}{\otimes}\phi_t$ and let $\psi = \underset{t\in I}{\otimes}\psi_t$ be in $\mathcal{B}_\otimes^\alpha$ for $\lambda \leqslant \alpha \leqslant \gamma$.

(1) We say that ϕ is strongly equivalent to ψ, and write $\phi \equiv^s \psi$, if and only if

$$\sum_{t\in I} \left|1 - \langle \phi_t, \psi_t^* \rangle_t\right| < \infty, \text{ and } \sum_{t\in I} \left|1 - \langle \psi_t, \phi_t^* \rangle_t\right| < \infty.$$

(2) We say that ϕ is weakly equivalent to ψ, and write $\phi \equiv^w \psi$, if and only if both

$$\sum_{t\in I} \left|1 - \left|\langle \phi_t, \psi_t^* \rangle_t\right|\right| < \infty, \text{ and } \sum_{t\in I} \left|1 - \left|\langle \psi_t, \phi_t^* \rangle_t\right|\right| < \infty.$$

Theorem 6.16. *We have $\phi \equiv^w \psi$ if and only if there exist z_t, $|\,z_t\,| = 1$, such that $\otimes_{t\in I} z_t\phi_t \equiv^s \otimes_{t\in I} \psi_t$ and $\otimes_{t\in I} \bar{z}_t\psi_t \equiv^s \otimes_{t\in I} \phi_t$.*

Proof. Suppose that $\otimes_{t\in I} z_t\phi_t \equiv^s \otimes_{t\in I} \psi_t$ and $\otimes_{t\in I} \bar{z}_t\psi_t \equiv^s \otimes_{t\in I} \phi_t$. Then we have:

$$\sum_{t\in I} \left|1 - \left|\langle \phi_t, \psi_t^* \rangle_t\right|\right| = \sum_{t\in I} \left|1 - \left|\langle z_t\phi_t, \psi_t^* \rangle_t\right|\right| \leqslant \sum_{t\in I} \left|1 - \langle z_t\phi_t, \psi_t^* \rangle_t\right| < \infty, \text{and}$$

$$\sum_{t\in I} \left|1 - \left|\langle \psi_t, \phi_t^* \rangle_t\right|\right| = \sum_{t\in I} \left|1 - \left|\langle \bar{z}_t\psi_t, \phi_t^* \rangle_t\right|\right| \leqslant \sum_{t\in I} \left|1 - \langle \bar{z}_t\psi_t, \phi_t^* \rangle_t\right| < \infty.$$

If $\phi \equiv^w \psi$, set $z_t = |\langle \phi_t, \psi_t^* \rangle_t| / \langle \phi_t, \psi_t^* \rangle_t$ for $\langle \phi_t, \psi_t^* \rangle_t \neq 0$, and set $z_t = 1$ otherwise. Now

$$z_t = |\langle \phi_t, \psi_t^* \rangle_t| / \langle \phi_t, \psi_t^* \rangle_t = \overline{|\langle \psi_t, \phi_t^* \rangle_t| / \langle \psi_t, \phi_t^* \rangle_t}$$

follows from the definition of the Steadman map. Thus,

$$\sum_{t\in I} \left|1 - \langle z_t\phi_t, \psi_t^* \rangle_t\right| = \sum_{t\in I} \left|1 - \left|\langle \phi_t, \psi_t^* \rangle_t\right|\right| < \infty \text{ and}$$

$$\sum_{t\in I} \left|1 - \langle \bar{z}_t\psi_t, \phi_t^* \rangle_t\right| = \sum_{t\in I} \left|1 - \left|\langle \psi_t, \phi_t^* \rangle_t\right|\right| < \infty,$$

so that $\otimes_{t\in I} z_t\phi_t \equiv^s \otimes_{t\in I} \psi_t$ and $\otimes_{t\in I} \bar{z}_t\psi_t \equiv^s \otimes_{t\in I} \phi_t$. □

In the proof of the next theorem, we appeal to the Gram–Schmidt process on $\mathcal{B}_t$. It is actually proven in the Appendix (Sect. 5.3) of Chap. 5. However, the reader who has not reviewed that section can use the fact that $\mathcal{B}_t \subset \mathcal{H}_t$ as a continuous dense embedding apply the Gram–Schmidt process on $\mathcal{H}_t$ and restrict back to $\mathcal{B}_t$. (We drop the t index on $\mathcal{B}$ in what follows.)

Theorem 6.17. *The relations defined above are equivalence relations on $\mathcal{B}_{\otimes}^{\alpha}$.*

Proof. First note that the relations are clearly symmetric and reflexive, so we need to only show that they are transitive. Suppose that $\underset{t\in I}{\otimes}\phi_t \equiv^s \underset{t\in I}{\otimes}\psi_t$ and $\underset{t\in I}{\otimes}\psi_t \equiv^s \underset{t\in I}{\otimes}\rho_t$, then we know that

$$\prod_{t\in I}\|\phi_t\|_{\mathcal{B}} < \infty,\ \prod_{t\in I}\|\psi_t\|_{\mathcal{B}} < \infty,\ \prod_{t\in I}\|\rho_t\|_{\mathcal{B}} < \infty \text{ and}$$

$$\sum_{t\in I}|1-\langle\phi_t,\psi_t^*\rangle_t| < \infty,\ \ \sum_{t\in I}|1-\langle\psi_t,\phi_t^*\rangle_t| < \infty,$$

so that $\sum_{t\in I}|1-\langle\rho_t,\psi_t^*\rangle_t| < \infty$ and $\sum_{t\in I}|1-\langle\psi_t,\rho_t^*\rangle_t| < \infty$. We need to show that $\sum_{t\in I}|1-\langle\phi_t,\rho_t^*\rangle_t| < \infty$ and $\sum_{t\in I}|1-\langle\rho_t,\phi_t^*\rangle_t| < \infty$. We prove the first case, $\sum_{t\in I}|1-\langle\phi_t,\rho_t^*\rangle_t| < \infty$, as the second is similar. Fix t, and set $\|\phi_t\|_{\mathcal{B}} = 1+\eta$, $\|\psi_t\|_{\mathcal{B}} = 1+\theta$, $\|\rho_t\|_{\mathcal{B}} = 1+\zeta$ and $\langle\phi_t,\psi_t^*\rangle_t = 1+x$, $\langle\psi_t,\rho_t^*\rangle_t = 1+\lambda$. It follows that

$$\max\{|\eta|\,,\ |\theta|\,,\ |\zeta|\,,\ |x|\,,\ |\lambda|\} \leqslant C$$

for some constant C (independent of t). Without loss, we also assume that, except for a finite number of t, $|\theta| \leqslant \frac{1}{2}$ (i.e., $|1-\|\psi_t\|_{\mathcal{B}}| \leqslant \frac{1}{2}$). Using the Gram–Schmidt process in $\mathcal{B}$, write ψ_t, ϕ_t and ρ_t, as:

$$\psi_t = a_{11}\mathrm{u},\quad \psi_t^* = \bar{a}_{11}\mathrm{u}^*,$$
$$\langle\psi_t,\psi_t^*\rangle = \|\psi_t\|_{\mathcal{B}}^2 = |a_{11}|^2 = (1+\theta)^2,$$
$$\phi_t = a_{21}\mathrm{u} + a_{22}\mathrm{v},\quad \phi_t^* = \bar{a}_{21}\mathrm{u}^* + \bar{a}_{22}\mathrm{v}^*,$$
$$\langle\phi_t,\phi_t^*\rangle = \|\phi_t\|_{\mathcal{B}}^2 = |a_{21}|^2 + |a_{22}|^2 = (1+\eta)^2,$$
$$\rho_t = a_{31}\mathrm{u} + a_{32}\mathrm{v} + a_{33}\mathrm{w},\quad \rho_t^* = \bar{a}_{31}\mathrm{u}^* + \bar{a}_{32}\mathrm{v}^* + \bar{a}_{33}\mathrm{w}^*,$$
$$\langle\rho_t,\rho_t^*\rangle = \|\rho_t\|_{\mathcal{B}}^2 = |a_{31}|^2 + |a_{32}|^2 + |a_{33}|^2 = (1+\zeta)^2,$$
$$\langle\phi_t,\psi_t^*\rangle = a_{21}\bar{a}_{11} = 1+x,$$
$$\langle\psi_t,\rho_t^*\rangle = a_{11}\bar{a}_{31} = 1+\lambda.$$

Now,

$$\begin{aligned}
&|1-\langle\phi_t,\rho_t^*\rangle_t| \\
&= |1-(a_{21}\bar{a}_{31} + a_{22}\bar{a}_{32})| \\
&= \left|(a_{21}\bar{a}_{11})(a_{11}\bar{a}_{31})\,|a_{11}|^{-2} - 1 + a_{22}\bar{a}_{32}\right| \\
&\leqslant \left|(a_{21}\bar{a}_{11})(a_{11}\bar{a}_{31})\,|a_{11}|^{-2} - 1\right| + |a_{22}\bar{a}_{32}| \\
&\leqslant \left|(1+|x|)(1+|\lambda|)(1-|\theta|)^{-2} - 1\right| + |a_{22}\bar{a}_{32}|\,. \qquad \text{(A)}
\end{aligned}$$

Looking at the first part of the last inequality and using $|\theta| \leqslant \frac{1}{2}$, we see that

$$\begin{aligned}
&\left|(1+|x|)(1+|\lambda|)(1-|\theta|)^{-2}-1\right| \\
&= \left\{(1+\tfrac{1}{2}|\lambda|)|x| + (1+\tfrac{1}{2}|x|)|\lambda| + (2-|\theta|)|\theta|\right\}(1-|\theta|)^{-2} \\
&\leqslant \left\{(1+\tfrac{1}{2}C)|x| + (1+\tfrac{1}{2}C)|\lambda| + 2|\theta|\right\}(1/2)^{-2} \\
&\leqslant D'\{|x|+|\lambda|+|\theta|\} \\
&\leqslant D'\{|\eta|+|\theta|+|\zeta|+|x|+|\lambda|\},
\end{aligned}$$

where D' is a constant, independent of t. For the last part (of the last inequality in A), we have:

$$\begin{aligned}
|a_{22}|^2 &= \left\{|a_{21}|^2+|a_{22}|^2\right\} - |a_{21}\bar{a}_{11}|^2|a_{11}|^{-2} \leqslant (1+|\eta|)^2-(1-|x|)^2(1+|\theta|)^{-2} \\
&\leqslant D''\{|\eta|+|\theta|+|\zeta|+|x|+|\lambda|\}, \qquad \text{(B)}
\end{aligned}$$

and

$$\begin{aligned}
&|a_{32}|^2 \\
&\leqslant \left\{|a_{31}|^2+|a_{32}|^2+|a_{33}|^2\right\} - |a_{11}\bar{a}_{31}|^2|a_{11}|^{-2} \\
&\leqslant (1+|\zeta|)^2-(1-|\lambda|)^2(1+|\theta|)^{-2} \\
&\leqslant D'''\{|\eta|+|\theta|+|\zeta|+|x|+|\lambda|\}. \qquad \text{(C)}
\end{aligned}$$

Combining terms, we have that:

$$\begin{aligned}
&\left|1-\langle\phi_t,\rho_t^*\rangle_t\right| \\
&\leqslant D\{|\eta|+|\theta|+|\zeta|+|x|+|\lambda|\} \\
&= D\left\{\left|1-\|\phi_t\|_{\mathcal{B}}\right| + \left|1-\|\psi_t\|_{\mathcal{B}}\right| + \left|1-\|\rho_t\|_{\mathcal{B}}\right| + \left|1-\langle\phi_t,\psi_t^*\rangle_t\right| + \left|1-\langle\psi_t,\rho_t^*\rangle_t\right|\right\}.
\end{aligned}$$

Since the constant D is independent of all but a finite number of t, our proof is complete. The proof of weak equivalence now follows from Theorem 6.16. □

Remark 6.18. We note that a proof of the same result for a Hilbert space in the last theorem only required ten lines.

Definition 6.19. Let $\varphi = \underset{t\in I}{\otimes} \varphi_t \in \mathcal{B}_\otimes^\alpha$.

(1) We define $\mathcal{B}_\otimes^\alpha(\varphi)^s$ to be the closed subspace generated by the span of all $\psi \equiv^s \varphi$ and we call it the strong partial tensor product space of type v generated by the vector φ.

(2) We define $\mathcal{B}_\otimes^\alpha(\varphi)^w$ to be the closed subspace generated by the span of all $\psi \equiv^w \varphi$ and we call it the weak partial tensor product space of type v generated by the vector φ.

Theorem 6.20. *With* $\varphi,\ \psi \in \mathcal{B}_{\otimes}^{\alpha}$*:*

(1) *The product* $\prod_{t\in I}\|\varphi_t\|_{\mathcal{B}}$ *converges if and only if* $\prod_{t\in I}\|\varphi_t\|_{\mathcal{B}}^2$ *converges.*

(2) *If* $\prod_{t\in I}\|\varphi_t\|_{\mathcal{B}}$ *and* $\prod_{t\in I}\|\psi_t\|_{\mathcal{B}}$ *converge, then* $\prod_{t\in I}\langle\varphi_t,\psi_t^*\rangle_t$, *and* $\prod_{t\in I}\langle\psi_t,\varphi_t^*\rangle_t$ *are quasi-convergent.*

(3) *If* $\prod_{t\in I}\langle\varphi_t,\psi_t^*\rangle_t$ *and* $\prod_{t\in I}\langle\psi_t,\varphi_t^*\rangle_t$ *are quasi-convergent, then there exist complex numbers* $\{z_t\}$, *with* $|z_t| = 1$, *such that* $\prod_{t\in I}\langle z_t\varphi_t,\psi_t^*\rangle_t$, *and* $\prod_{t\in I}\langle\bar{z}_t\psi_t,\varphi_t^*\rangle_t$ *both converge.*

(4) *If* $\varphi = \otimes_{t\in I}\varphi_t$ *and* $\psi = \otimes_{t\in I}\psi_t$ *are in different equivalence classes of* $\mathcal{B}_{\otimes}^{\alpha}$, *then* $\prod_{t\in I}\langle\varphi_t,\psi_t^*\rangle_t = 0$ *and* $\prod_{t\in I}\langle\psi_t,\varphi_t^*\rangle_t = 0$.

Proof. For the first case, convergence of either term implies that $\{\|\varphi_t\|_{\mathcal{B}},\ t\in I\}$ is bounded by some $M>0$. Hence,

$$\left|1-\|\varphi_t\|_{\mathcal{B}}\right| \leqslant \left|1+\|\varphi_t\|_{\mathcal{B}}\right|\left|1-\|\varphi_t\|_{\mathcal{B}}\right| = \left|1-\|\varphi_t\|_{\mathcal{B}}^2\right| \leqslant (1+M)\left|1-\|\varphi_t\|_{\mathcal{B}}\right|.$$

To prove 2, note that if $J\subset I$ is any finite subset, we have

$$0 \leqslant \left|\prod_{t\in J}\langle\varphi_t,\psi_t^*\rangle_t\right| \leqslant \prod_{t\in J}\|\varphi_t\|_{\mathcal{B}}\prod_{t\in J}\|\psi_t\|_{\mathcal{B}} < \infty.$$

Therefore, $0 \leqslant \left|\prod_{t\in I}\langle\varphi_t,\psi_t^*\rangle_t\right| < \infty$, so that $\prod_{t\in I}\langle\varphi_t,\psi_t^*\rangle_t$ is quasi-convergent, and if $0 < \left|\prod_{t\in I}\langle\varphi_t,\psi_t^*\rangle_t\right| < \infty$, it is even convergent. The proof of 3 follows from Theorem 6.16.

To prove 4, note that if $0 < |\langle\varphi,\psi^*\rangle|, |\langle\psi,\varphi^*\rangle| < \infty$, then φ and ψ are in the same equivalence class, so the only possibilities are $|\langle\varphi,\psi^*\rangle|, |\langle\psi,\varphi^*\rangle| = 0$ or ∞. Thus, it suffices to show that $|\langle\varphi,\psi^*\rangle|, |\langle\psi,\varphi^*\rangle| \neq \infty$, and this fact follows from 2. ☐

We now consider the relationship between $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ and $\mathcal{H}_{\otimes}^{2}(\varphi)^s$. Since $\|\varphi_t\|_{\mathcal{B}}^2\big/\|\varphi_t\|_{\mathcal{H}}^2 \geqslant 1$, we have that $|\langle\psi_t,\varphi_t^*\rangle_t| = \left|(\psi_t,\varphi_t)_{2,t}\right|\left(\|\varphi_t\|_{\mathcal{B}}^2\big/\|\varphi_t\|_{\mathcal{H}}^2\right) \geqslant \left|(\psi_t,\varphi_t)_{2,t}\right|$. Set $\|\varphi_t\|_{\mathcal{B}}^2\big/\|\varphi_t\|_{\mathcal{H}}^2 = a_t$ and $\|\psi_t\|_{\mathcal{B}}^2\big/\|\psi_t\|_{\mathcal{H}}^2 = b_t$.

Theorem 6.21. *Suppose* $\prod_{t\in I}\langle\psi_t,\varphi_t^*\rangle_t$ *and* $\prod_{t\in I}\langle\varphi_t,\psi_t^*\rangle_t$ *both converge.*

(1) *Then* $\prod_{t\in I}(\varphi_t,\psi_t)_{2,t} < \infty$ *(converges) if and only if both* $\prod_{t\in I}a_t$ *and* $\prod_{t\in I}b_t$ *converge.*

(2) *If* $\prod_{t\in I} a_t$ *and* $\prod_{t\in I} b_t$ *converge, then* $\underset{t\in I}{\otimes}\varphi_t \equiv^s_{\mathcal{H}} \underset{t\in I}{\otimes}\psi_t$ *implies* $\underset{t\in I}{\otimes}\varphi_t \equiv^s_{\mathcal{B}} \underset{t\in I}{\otimes}\psi_t$*, so that* $\mathcal{B}^{\alpha}_{\otimes}(\varphi)^s \subset \mathcal{H}^2_{\otimes}(\varphi)^s$ *as a continuous dense embedding.*

Proof. To prove 1, let $J \subset I$ be finite. Then $\left\{\prod_{t\in J} a_t\right\}\prod_{t\in J}(\psi_t,\varphi_t)_{2,t} = \prod_{t\in J}\langle\psi_t,\varphi_t^*\rangle_t$, $\left\{\prod_{t\in J} b_t\right\}\prod_{t\in J}(\varphi_t,\psi_t)_{2,t} = \prod_{t\in J}\langle\varphi_t,\psi_t^*\rangle_t$. It is clear that since both $a_t \geqslant 1$ and $b_t \geqslant 1$, convergence of either term on the left implies convergence of the other term. Thus, $\prod_{t\in I}(\varphi_t,\psi_t)_{2,t}$ converges if and only if $\prod_{t\in I} a_t$ and $\prod_{t\in I} b_t$ converge. The proof of 2 follows from the definition of strong equivalence in both cases. □

The next result is implicit in the above proof.

Corollary 6.22. *If* $\prod_{t\in I}\langle\psi_t,\varphi_t^*\rangle_t$ *and* $\prod_{t\in I}\langle\varphi_t,\psi_t^*\rangle_t$ *converge, then either both* $\prod_{t\in I} a_t$ *and* $\prod_{t\in I} b_t$ *converge or they both diverge.*

Theorem 6.23. *Let* $\psi = \underset{t\in I}{\otimes}\psi_t$ *and* $\varphi = \underset{t\in I}{\otimes}\varphi_t$ *be vectors in* $\mathcal{B}^{\alpha}_{\otimes}$*.*

(1) *If* $\psi_t \neq \varphi_t$ *occurs for at most a finite number of* t*, then* $\psi = \underset{t\in I}{\otimes}\psi_t \equiv^s \varphi = \underset{t\in I}{\otimes}\varphi_t$*.*

(2) *The space* $\mathcal{B}^{\alpha}_{\otimes}(\varphi)^s$ *is the closure of the linear span of* $\psi = \underset{t\in I}{\otimes}\psi_t$ *such that* $\psi_t \neq \varphi_t$ *occurs for at most a finite number of* t*.*

(3) $\mathcal{B}^{\alpha}_{\otimes}(\varphi)^w = \underset{\psi\equiv^w\varphi}{\oplus}\left[\mathcal{B}^{\alpha}_{\otimes}(\psi)^s\right]$.

Proof. To prove 1, let J be the finite set of t for which $\psi_t \neq \varphi_t$. Then

$$\sum_{t\in I}|1-\langle\varphi_t,\psi_t^*\rangle_t| = \sum_{t\in J}|1-\langle\varphi_t,\psi_t^*\rangle_t| + \sum_{t\in I\setminus J}|1-\langle\varphi_t,\varphi_t^*\rangle_t|$$
$$\leqslant c + \sum_{t\in I}\left|1-\|\varphi_t\|_t^2\right| < \infty.$$

To prove 2, let $\mathcal{B}^{\alpha}_{\otimes}(\varphi)^{\#}$ be the closure of the linear span of all $\psi = \underset{t\in I}{\otimes}\psi_t$ such that $\psi_t \neq \varphi_t$ occurs for at most a finite number of t and, without loss, we can assume that $\|\varphi_t\|_t = 1$. It is clear from 1 that $\mathcal{B}^{\alpha}_{\otimes}(\varphi)^{\#} \subseteq \mathcal{B}^{\alpha}_{\otimes}(\varphi)^s$, so that we are done if we can show that $\mathcal{B}^{\alpha}_{\otimes}(\varphi)^{\#} \supseteq \mathcal{B}^{\alpha}_{\otimes}(\varphi)^s$. For any vector $\psi = \underset{t\in I}{\otimes}\psi_t$ in $\mathcal{B}^{\alpha}_{\otimes}(\varphi)^s$, $\varphi \equiv^s \psi$ so that $\sum_{t\in I}|1-\langle\varphi_t,\psi_t^*\rangle_t| < \infty$. If $\|\psi\|^{\alpha}_{\otimes} = 0$ then $\psi \in \mathcal{B}^{\alpha}_{\otimes}(\varphi)^{\#}$, so we can assume that $\|\psi\|_{\otimes}^{2} \neq 0$. This implies that $\|\psi_t\|_t \neq 0$ for all $t \in I$ and $0 \neq \prod_{t\in I}(1/\|\psi_t\|_t) < \infty$. Hence, by scaling if necessary, we may also assume that $\|\psi_t\|_t = 1$ for all $t \in I$. Let $0 < \varepsilon < 1$ be given, and choose δ so that $0 < \sqrt{2\delta e} < \varepsilon$ (e is the base for the natural log). Since $\sum\limits_{t\in I}|1 - \langle\varphi_t, \psi_t^*\rangle_t| < \infty$ and $\sum\limits_{t\in I}|1 - \langle\psi_t, \varphi_t^*\rangle_t| < \infty$, there is a finite set of distinct values $J = \{t_1, \cdots, t_n\}$ such that $\sum\limits_{t\in I\setminus J}|1 - \langle\varphi_t, \psi_t^*\rangle_t| < \delta$ and $\sum\limits_{t\in I\setminus J}|1 - \langle\psi_t, \varphi_t^*\rangle_t| < \delta$. Since, for any finite number of complex numbers $z_1, \cdots, z_n$, it is easy to see that $|\prod_{k=1}^{n} z_k - 1| = |\prod_{k=1}^{n}[1 + (z_k - 1)] - 1| \leqslant \left(\prod_{k=1}^{n} e^{|z_k-1|} - 1\right)$, we have that

$$\left|\prod_{t\in I\setminus J}\langle\varphi_t, \psi_t^*\rangle_t - 1\right| \leqslant \left(\exp\{\sum_{t\in I\setminus J}|\langle\varphi_t, \psi_t^*\rangle_t - 1|\} - 1\right) \leqslant e^{\delta} - 1 \leqslant e\delta$$

and

$$\left|\prod_{t\in I\setminus J}\langle\psi_t, \varphi_t^*\rangle_t - 1\right| \leqslant \left(\exp\{\sum_{t\in I\setminus J}|\langle\psi_t, \varphi_t^*\rangle_t - 1|\} - 1\right) \leqslant e^{\delta} - 1 \leqslant e\delta.$$

Now, define

$$\phi_t = \begin{cases} \psi_t, & t \in J \\ \varphi_t, & t \in I\setminus J, \end{cases}$$

and set $\phi_J = \otimes_{t\in I}\phi_t$ so that $\phi_J \in \mathcal{B}_{\otimes}^{\alpha}(\varphi)^{\#}$. Assume that $\left|\|\psi^*\|_{\otimes} - \|\phi_J^*\|_{\otimes}\right| \neq 0$. In this case, we have

$$\begin{aligned}
&\|\psi - \phi_J\|_{\otimes}\left|\|\psi^*\|_{\otimes} - \|\phi_J^*\|_{\otimes}\right| \leqslant |\langle\psi - \phi_J, \psi^* - \phi_J^*\rangle| \\
&= \left|\prod_{t\in I}\|\psi_t\|_t^2 + \prod_{t\in I}\|\phi_t\|_t^2 - \prod_{t\in I}\langle\psi_t, \phi_t^*\rangle_t - \prod_{t\in I}\langle\phi_t, \psi_t^*\rangle_t\right| \\
&= \left|2 - \prod_{t\in J}\langle\psi_t, \phi_t^*\rangle_t \prod_{t\in I\setminus J}\langle\psi_t, \phi_t^*\rangle_t - \prod_{t\in J}\langle\phi_t, \psi_t^*\rangle_t \prod_{t\in I\setminus J}\langle\phi_t, \psi_t^*\rangle_t\right| \\
&= \left|\left(1 - \prod_{t\in I\setminus J}\langle\psi_t, \phi_t^*\rangle_t\right) + \left(1 - \prod_{t\in I\setminus J}\langle\phi_t, \psi_t^*\rangle_t\right)\right| \leqslant 2e\delta < \varepsilon^2.
\end{aligned}$$

If $\left|\|\psi^*\|_{\otimes} - \|\phi_J^*\|_{\otimes}\right| = 0$, choose $\alpha > 0$ such that

$$\alpha\|\psi - \phi_J\|_{\otimes} \leq |\langle\psi - \phi_J, \psi^* - \phi_J^*\rangle|.$$

In either case, since ε is arbitrary, ψ is in the closure of $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^{\#}$, so that $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^{\#} = \mathcal{B}_{\otimes}^{\alpha}(\varphi)^{s}$. The proof of 3 follows from Theorem 6.20 (4) and the definition of weakly equivalent spaces. $\square$

In the next few sections that follow, we prove all theorems for Banach spaces. In some cases the same proof for a Hilbert space is easier and is left as an exercise (see [GZ1]).

6.2.2. Biorthogonal System for $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^{s}$. We now construct a biorthogonal system for each $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^{s}$. Let Γ be an index set with the dimension of $\mathcal{B}_t$ and let $\{e_{\gamma}^{t},\ (e_{\gamma}^{t})^{*},\ \gamma \in \Gamma\}$ be a biorthogonal system for $\mathcal{B}_t$. Let e_0^t be a fixed unit vector in $\mathcal{B}_t$ and set $E = \otimes_{t\in I} e_0^t$. Let $\mathbf{F}$ be the set of all functions $f : I \to \Gamma \cup \{0\}$ such that $f(t) = 0$ for all but a finite number of t. Let $F(f)$ be the image of $f \in \mathbf{F}$ (i.e., $F(f) = \{f(t),\ t \in I\}$), and set $E_{F(f)} = \otimes_{t\in I} e_{t,f(t)}$, where $f(t) = 0$ implies $e_{t,0} = e_0^t$ and $f(t) = \gamma$ implies $e_{t,\gamma} = e_{\gamma}^{t}$. Let $E_{F(f)}^{*} = \otimes_{t\in I} e_{t,f(t)}^{*}$.

Theorem 6.24. *With the above notation, we have:*

1. *The set $\{E_{F(f)},\ E_{F(f)}^{*},\ f \in \mathbf{F}\}$ is a biorthogonal system for $\mathcal{B}_{\otimes}^{\alpha}(E)^{s}$.*
2. *If I is countable, then $\mathcal{B}_{\otimes}^{\alpha}(E)^{s}$ is separable.*

Proof. To prove 1, note that $E \in \{E_{F(f)},\ f \in \mathbf{F}\}$ and each $E_{F(f)}$ is a unit vector. Also, we have $E_{F(f)} \equiv^{s} E$ and $\left\langle E_{F(f)}, E_{F(g)}^{*}\right\rangle = \left\langle E_{F(g)}, E_{F(f)}^{*}\right\rangle = \prod_{t\in I}\left\langle e_{t,f(t)}, e_{t,g(t)}^{*}\right\rangle = 0$ unless $f(t) = g(t)$ for all t. Hence, the family $\{E_{F(f)},\ f \in \mathbf{F}\}$ is an orthonormal system for $\mathcal{B}_{\otimes}^{\alpha}(E)^{s}$. Let $\mathcal{B}_{\otimes}^{\alpha}(E)^{\#}$ be the completion of the linear span of the family $\{E_{F(f)},\ f \in \mathbf{F}\}$. Clearly, $\mathcal{B}_{\otimes}^{\alpha}(E)^{\#} \subseteq \mathcal{B}_{\otimes}^{\alpha}(E)^{s}$ so we only need to prove that $\mathcal{B}_{\otimes}^{\alpha}(E)^{s} \subset \mathcal{B}_{\otimes}^{\alpha}(E)^{\#}$. By Theorem 6.23(2.), it suffices to prove that $\mathcal{B}_{\otimes}^{\alpha}(E)^{\#}$ contains the closure of the set of all $\varphi = \otimes_{t\in I}\varphi_t$ such that $\varphi_t \neq e_0^t$ occurs for only a finite number of t. Let $\varphi = \otimes_{t\in I}\varphi_t$ be any such vector, and let $J = \{t_1, \cdots, t_n\}$ be the finite set of distinct values of t for which $\varphi_t \neq e_0^t$ occurs. Since $\{e_{\gamma}^{t},\ \gamma \in \Gamma\}$ is a basis for $\mathcal{B}_t$, for each t_i there exist constants $a_{t_i,\gamma}$ such that $\sum_{\gamma\in\Gamma} a_{t_i,\gamma} e_{\gamma}^{t_i} = \varphi_{t_i}$ for $1 \leqslant i \leqslant n$. Let $\varepsilon > 0$ be given. Then, for each t_i there exists a finite subset $N_i \subset \Gamma$ such that $\left\|\varphi_{t_i} - \sum_{\gamma\in N_i} a_{t_i,\gamma} e_{\gamma}^{t_i}\right\|_{\otimes} < \frac{1}{n}(\varepsilon/\|\varphi\|_{\otimes})$. Let $\mathbf{N} = (N_1, \cdots N_n)$ and set $\varphi_{t_i}^{N_i} = \sum_{\gamma \in N_i} a_{t_i,\gamma} e_\gamma^{t_i}$ so that $\varphi^{\mathbf{N}} = \otimes_{t_i \in J} \varphi_{t_i}^{N_i} \otimes (\otimes_{t \in I \setminus J} e_0^t)$ and $\varphi = \otimes_{t_i \in J} \varphi_{t_i} \otimes (\otimes_{t \in I \setminus J} e_0^t)$. It follows that:

$$\begin{aligned} &\left\| \varphi - \varphi^{\mathbf{N}} \right\|_{\otimes} \\ &= \left\| \left[\otimes_{t_i \in J} \varphi_{t_i} - \otimes_{t_i \in J} \varphi_{t_i}^{N_i} \right] \otimes (\otimes_{t \in I \setminus J} e_0^t) \right\|_{\otimes} \\ &= \left\| \otimes_{t_i \in J} \varphi_{t_i} - \otimes_{t_i \in J} \varphi_{t_i}^{N_i} \right\|_{\otimes}. \end{aligned}$$

We can rewrite this as:

$$\begin{aligned} &\left\| \otimes_{t_i \in J} \varphi_{t_i} - \otimes_{t_i \in J} \varphi_{t_i}^{N_i} \right\|_{\otimes} \\ &= \| \varphi_{t_1} \otimes \varphi_{t_2} \cdots \otimes \varphi_{t_n} - \varphi_{t_1}^{N_1} \otimes \varphi_{t_2} \cdots \otimes \varphi_{t_n} \\ &\quad + \varphi_{t_1}^{N_1} \otimes \varphi_{t_2} \cdots \otimes \varphi_{t_n} - \varphi_{t_1}^{N_1} \otimes \varphi_{t_2}^{N_2} \cdots \otimes \varphi_{t_n} \\ &\quad \vdots \\ &\quad + \ \varphi_{t_1}^{N_1} \otimes \varphi_{t_2}^{N_2} \cdots \otimes \varphi_{t_{n-1}}^{N_{n-1}} \otimes \varphi_{t_n} - \varphi_{t_1}^{N_1} \otimes \varphi_{t_2}^{N_2} \cdots \otimes \varphi_{t_n}^{N_n} \Big\|_{\otimes} \\ &\leqslant \sum_{i=1}^{n} \left\| \varphi_{t_i} - \varphi_{t_i}^{N_i} \right\|_{\otimes} \|\varphi\|_{\otimes} \leqslant \varepsilon. \end{aligned}$$

Now, as the tensor product is multilinear and continuous in any finite number of variables, we have:

$$\begin{aligned} &\varphi^{\mathbf{N}} \\ &= \otimes_{t_i \in J} \varphi_{t_i}^{N_i} \otimes (\otimes_{t \in I \setminus J} e_0^t) = \otimes \varphi_{t_1}^{N_1} \otimes \varphi_{t_2}^{N_2} \cdots \otimes \varphi_{t_n}^{N_n} \otimes (\otimes_{t \in I \setminus J} e_0^t) \\ &= \left[\sum_{\gamma_1 \in N_1} a_{t_1,\gamma_1} e_{\gamma_1}^{t_1} \right] \otimes \left[\sum_{\gamma_2 \in N_2} a_{t_2,\gamma_2} e_{\gamma_2}^{t_2} \right] \cdots \otimes \left[\sum_{\gamma_n \in N_n} a_{t_n,\gamma_n} e_{\gamma_n}^{t_n} \right] \otimes (\otimes_{t \in I \setminus J} e_0^t) \\ &= \sum_{\gamma_1 \in N_1 \cdots \gamma_n \in N_n} a_{t_1,\gamma_1} a_{t_2,\gamma_2} \cdots a_{t_n,\gamma_n} \left[e_{\gamma_1}^{t_1} \otimes e_{\gamma_2}^{t_2} \cdots \otimes e_{\gamma_n}^{t_n} \otimes (\otimes_{t \in I \setminus J} e_0^t) \right]. \end{aligned}$$

It is now clear that, by the definition of $\mathbf{F}$, for each fixed set of indices $\gamma_1, \ \gamma_2, \cdots \ \gamma_n$ there exists a function $f : \ I \to \Gamma \cup \{0\}$ such that $f(t_i) = \gamma_i$ for $t_i \in J$ and $f(t) = 0$ for $t \in I \setminus J$. Since each N_i is finite, $\mathbf{N} = (N_1, \cdots N_n)$ is also finite, so that only a finite number of functions are needed. It follows that $\varphi^{\mathbf{N}}$ is in $\mathcal{B}_{\otimes}^{\alpha}(E)^{\#}$, so that φ is a limit point and $\mathcal{B}_{\otimes}^{\alpha}(E)^{\#} = \mathcal{B}_{\otimes}^{\alpha}(E)^{s}$.

To prove 2, note that if each $\mathcal{B}_t$ is separable, the collection of basis sets in Γ is countable. It follows that if I is countable then $\mathbf{F}$ is countable, so that the set of basis vectors of $\mathcal{B}_{\otimes}^{\alpha}(E)^{s}$ is countable and $\mathcal{B}_{\otimes}^{\alpha}(E)^{s}$ is separable. □

Theorem 6.25. *Suppose that* $\|\phi_t\|_{\mathcal{B}} = 1$ *for all* t*, so that* $\|\phi\|_{\otimes}^{\alpha} = 1$*. If* $\{z_t,\ t \in I,\ |z_t| = 1\}$ *is quasi-convergent (but not convergent) and* $\psi = \underset{t\in I}{\otimes} z_t\phi_t$*, then* $\mathcal{B}_{\otimes}^{\alpha}(\phi)^s \perp \mathcal{B}_{\otimes}^{\alpha}(\psi)^s$*.*

Proof. It is clear that $\|\psi\|_{\otimes}^{\alpha} = 1$, so let ϕ_t^* be the Steadman representation for ϕ_t, and ψ_t^* be the Steadman representation for ψ_t. It follows that

$$\langle \phi, \phi^* \rangle_{\otimes}^{\alpha} = \langle \psi, \psi^* \rangle_{\otimes}^{\alpha} = 1$$

with

$$\langle \psi, \phi^* \rangle_{\otimes}^{\alpha} = \prod_{t\in I} z_t \|\phi_t\|_{\mathcal{B}}^2 = \prod_{t\in I} z_t = 0$$

and $\langle \phi, \psi^* \rangle_{\otimes}^{\alpha} = 0$, by symmetry, so that $\phi \perp \psi$. This implies that the set Γ of all linear combinations of vectors $\eta = \underset{t\in I}{\otimes} \eta_t$ with $\eta_t \neq \phi_t$ occurring for only a finite number of t satisfies $\Gamma \perp \psi$. Since Γ is dense in $\mathcal{B}_{\otimes}^{\alpha}(\phi)^s$, we have that $\mathcal{B}_{\otimes}^{\alpha}(\phi)^s \perp \mathcal{B}_{\otimes}^{\alpha}(\psi)^s$. □

6.3. Examples

von Neumann described the decomposition of $\mathcal{H}_{\otimes}^2$ into strong and weak partial tensor product spaces as like a quantum mechanical splitting up. In this section we look at a few examples of strong and weak partial tensor product spaces. Let N be a countable set and, for each $n \in N$, let $(X_n,\ \mathfrak{B}_n,\ m_n)$ be a measure space, where X_n is a complete separable metric space, $\mathfrak{B}_n$ is the Borel σ-algebra generated by the open sets of X_n, and m_n is a probability measure on X_n. For $1 \le p < \infty$, let $\mathbf{L}^p[X_n,\ \mathfrak{B}_n,\ m_n] = \mathbf{L}^p[X_n]$ be the set of complex-valued functions $f(x)$ in X_n such that $|f(x)|^p$ is integrable with respect to m_n. If Δ_p is the natural tensor product norm for L^p spaces, then, for any pair X_m and X_n, $\mathbf{L}^p[X_m] \hat{\otimes}^{\Delta_p} \mathbf{L}^p[X_n] = \mathbf{L}^p[X_m \times X_n]$.

Let $\phi_n \in \mathbf{L}^p[X_n]$ with $\|\phi_n\|_{\mathbf{L}^p[X_n]} = 1$ and, with $\phi = \otimes_{n\in N}\phi_n$, construct the strong tensor product space $\mathbf{L}_{\otimes}^{\Delta_p}[\phi]^s$. Let $X = \prod_{n\in N} X_n$ and $\mathfrak{B} = \hat{\otimes}_{n\in N}\mathfrak{B}_n$ (the smallest σ-algebra containing $\prod_{n\in N}\mathfrak{B}_n$). Recall that a tame function in $\mathbf{L}^p[X]$ is any function $f \in \mathbf{L}^p[X]$ which only depends on a finite number of variables. The next theorem was first proven by Guichardet [GU] for $p = 2$.

Theorem 6.26. $\mathbf{L}_{\otimes}^{\Delta_p}[\phi]^s \cong \mathbf{L}^p[X]$.

Proof. Let $J_M = \{n_1, \cdots n_M\} \subset N$, (where M is finite but arbitrary), let $f(x_{n_1}, \cdots x_{n_M})$ be a tame function in $\mathbf{L}^p[X]$, and define $\tilde{f}(x_{n_1}, \cdots x_{n_M}) = f(x_{n_1}, \cdots x_{n_M}) \otimes (\otimes_{n \in N \setminus J_M} \phi_n)$ so that $\tilde{f}(x_{n_1}, \cdots x_{n_M}) \in \mathbf{L}_{\otimes}^{\Delta_p}[\phi]^s$. Define a function $\Lambda : \mathbf{L}^p[X] \to \mathbf{L}_{\otimes}^{\Delta_p}[\phi]^s$ by $\Lambda(f) = \tilde{f}$. It is easy to check that Λ is well defined and it is easy to see that: (1) $\Lambda(af_1 + bf_2) = a\Lambda(f_1) + b\Lambda(f_2)$ (Λ is a linear mapping); (2) $\|\Lambda(f)\|_{\Delta_p} = \|f\|_{\mathbf{L}^p[X]}$ (Λ is an isometric mapping); and (3) $\Lambda(f_1) = \Lambda(f_2) \Rightarrow f_1 = f_2$ (Λ is a one-to-one mapping). Since the set of tame functions is dense in $\mathbf{L}^p[X]$ and the set of all $\tilde{f}$ is dense in $\mathbf{L}_{\otimes}^{\Delta_p}[\phi]^s$, it follows that, for any f in $\mathbf{L}^p[X]$, we can define $\Lambda(f) = \lim_{k \to \infty} \Lambda(f_k)$, where $\{f_k\}$ is any sequence of tame functions converging to f. Since Δ_p is a faithful norm, the extension to $\mathbf{L}^p[X]$ is one-to-one, so that Λ defines an isometric isomorphism of $\mathbf{L}^p[X]$ onto $\mathbf{L}_{\otimes}^{\Delta_p}[\phi]^s$. □

Remark 6.27. Now observe that this theorem is true if each $X_n = \mathbb{R}$ and each $m_n = \lambda$ (Lebesgue measure). In this case, $\mathbf{L}_{\otimes}^{\Delta_p}[\phi]^s \cong \mathbf{L}^p[\mathbb{R}^\infty, \mathfrak{B}(\mathbb{R}^\infty), \hat{\lambda}_\infty]$, where $\hat{\lambda}_\infty$ is some (unknown) version of Lebesgue measure on $\mathbb{R}^\infty$. It was this observation that led to the work in Chap. 2 (see [GPZ]).

In the next example $\mathcal{B}$ is not separable but is continuously embedded in KS^2, which is separable, so we can construct $\mathcal{B}_{\otimes}^{\alpha}$. For each $i \in I$, let $\mathbb{C}_b[X_i]$ be the space of bounded continuous functions on X_i. It is straightforward to construct $\mathbb{C}_{\otimes}^{\Delta_\infty} = \hat{\otimes}_{i \in I}^{\Delta_\infty} \mathbb{C}_b[X_i]$. If $\phi_i \in \mathbb{C}_b[X_i]$, with $\|\phi_i\|_{C[X_i]} = 1$ and $\phi = \otimes_{i \in I} \phi_i$, we can construct $\mathbb{C}_{\otimes}^{\Delta_\infty}[\phi]^s$. With $X = \prod_{i \in I} X_i$, the next result is proved as in Theorem 6.26.

Theorem 6.28. $\mathbb{C}_{\otimes}^{\Delta_\infty}[\phi]^s \cong \mathbb{C}_b[X]$.

Definition 6.29. Let $h(x) = \otimes_{n=1}^{\infty} h_n(x_n)$, where $h_n(x_n) = \chi_I(x_n)$ and $I = [-\frac{1}{2}, \frac{1}{2}]$. We call $\mathbf{L}_{\otimes}^{\Delta_p}[h]^s$ the canonical representation of $L^p[\mathbb{R}_I^\infty, \mathfrak{B}(\mathbb{R}_I^\infty), \lambda_\infty]$.

It is clear that there is an uncountable number of families of functions $\{g_n\}$, $\|g_n\|_p = 1$ from which we can construct a representation of $L^p[\mathbb{R}_I^\infty, \mathfrak{B}(\mathbb{R}_I^\infty), \lambda_\infty]$. From Theorem 6.20, we see that each such representation will either be equivalent or orthogonal to our canonical one.

Definition 6.30. Let $\mathcal{B}$ be a Banach space with an S-basis. We define $L^p[\mathcal{B}]$ by:

$$L^p[\mathcal{B}](h) \triangleq \{f(x) \in L^p[\mathbb{R}_I^\infty](h) : \ supp(f) \subset \mathcal{B}\},$$

while replacing λ_∞ with $\lambda_{\mathcal{B}}$. (Recall that $\lambda_\infty[\mathcal{B}] = 0$.)

6.3.1. $KS^p[\mathcal{B}]$**.** Let $\{e_n\}$ be a c.o.b for $KS^2[\mathbb{R}]$, let $x = (x_1, x_2, \dots) \in \mathcal{B}$, and define $e_n^i = e_n(x_i)$. Let $e_0^i = h_i(x_i)$ and define $f : \mathbb{N} \to \mathbb{N}$ be any function such that $f(n) = 0$ for all but a finite number of n, (e.f.n). For each such f, let $F(f) = \{f(n) : n \in \mathbb{N})\}$ and define

$$E_{F(f)} = \mathop{\otimes}_{i=1}^{\infty} e_{i,f(i)} : \begin{cases} f(i) = 0 \Rightarrow e_{i,f(i)} = e_0^i \\ f(i) = n \Rightarrow e_{i,f(i)} = e_n^i \end{cases}.$$

Theorem 6.31. *The set* $\{E_{F(f)} : f(n) = 0, (e.f.n)\}$ *is a c.o.b for* $KS^2[\mathbb{R}^\infty](h)$.

Definition 6.32. We define $KS^2\,[\mathcal{B}]\,(h)$ by

$$KS^2\,[\mathcal{B}]\,(h) = \left\{f \in KS^2\,[\mathbb{R}_I^\infty]\,(h) : supp(f) \subset \mathcal{B}\right\},$$

while replacing λ_∞ with $\lambda_{\mathcal{B}}$.

6.4. Operators

In this section, we restrict our discussion to bounded linear operators. In the Appendix, we discuss an interesting class of unbounded linear operators and their relationship to differential equations in infinitely many variables. (In the next section, we study those unbounded operators related to our main objective.)

A vector of the form $\phi = \otimes_{t \in I} \phi_t$, $\|\phi_t\| = 1$ for each t is called a basic vector in $\mathcal{B}_\otimes^\alpha$. We say an operator $\mathcal{A} : \mathcal{B}_\otimes^\alpha \to \mathcal{B}_\otimes^\alpha$ is reducible if the restriction of $\mathcal{A}$ to $\mathcal{B}_\otimes^\alpha[\phi]$ is invariant for every basic vector ϕ. For a particular ϕ, we say that $\mathcal{A}$ is reduced on $\mathcal{B}_\otimes^\alpha[\phi]$.

6.4.1. Bounded Operators on $\mathcal{B}_\otimes^\alpha$. In this section we investigate the class of bounded operators on $\mathcal{B}_\otimes^\alpha$ and their relationship to those on each $\mathcal{B}_t$. Let $L[\mathcal{B}_\otimes^\alpha]$ be the set of bounded operators on $\mathcal{B}_\otimes^\alpha$. For each fixed $t_0 \in I$ and $A_{t_0} \in L(\mathcal{B}_{t_0})$, define $\mathbf{A}(t_0) \in L(\mathcal{B}_\otimes^\alpha)$ by:

$$\mathbf{A}(t_0)(\sum_{k=1}^{N} \otimes_{t \in I} \varphi_t^k) = \sum_{k=1}^{N} A_{t_0} \varphi_{t_0}^k \otimes (\otimes_{t \neq t_0} \varphi_t^k)$$

for $\sum_{k=1}^{N} \otimes_{t \in I} \varphi_t^k$ in $\mathcal{B}_{\otimes}^{\alpha}$ and N finite but arbitrary. Extending to all of $\mathcal{B}_{\otimes}^{\alpha}$ produces an isometric isomorphism of $L[\mathcal{B}_{t_0}]$ into $L[\mathcal{B}_{\otimes}^{\alpha}]$, which we denote by $L[\mathcal{B}(t_0)]$, so that the relationship $L[\mathcal{B}_t] \leftrightarrow L[\mathcal{B}(t)]$ is an isometric isomorphism of algebras. Let $L^{\#}[\mathcal{B}_{\otimes}^{\alpha}]$ be the uniform closure of the algebra generated by $\{L[\mathcal{B}(t)],\ t \in I\}$. It is clear that $L^{\#}[\mathcal{B}_{\otimes}^{\alpha}] \subset L[\mathcal{B}_{\otimes}^{\alpha}]$. It is known that the inclusion becomes equality if and only if I is finite. On the other hand, $L^{\#}[\mathcal{B}_{\otimes}^{\alpha}]$ clearly consists of all operators on $\mathcal{B}_{\otimes}^{\alpha}$ that are generated directly from the family $\{L[\mathcal{B}(t)],\ t \in I\}$ by algebraic and topological processes. Thus, since $L[\mathcal{B}_{\otimes}^{\alpha}] \setminus L^{\#}[\mathcal{B}_{\otimes}^{\alpha}]$ is nonempty when I is infinite, we expect $L[\mathcal{B}_{\otimes}^{\alpha}]$ to contain operators distinct from those of $L^{\#}[\mathcal{B}_{\otimes}^{\alpha}]$. For an example, let $t \in \mathbb{N}$ and define U_n and U by:

$$U_n = \exp\{(-1)^n i\}, \text{ and } U = \hat{\otimes} U_n = \exp\{\sum_{n=1}^{\infty} (-1)^n i\}.$$

It is easy to see that U is unitary and is not reduced on any strong partial tensor product subspace. (It is easy to see that it is always reduced on every weak partial tensor product subspace.) Thus, $U \in L[\mathcal{B}_{\otimes}^{\alpha}] \setminus L^{\#}[\mathcal{B}_{\otimes}^{\alpha}]$.

Let $S_i(t),\ i = 1, 2$ be C_0-contraction semigroups with generators A_i defined on $\mathcal{H}$, so that $\|S_i(t)\|_{\mathcal{H}} \leqslant 1$. Define operators $\mathbf{S}_1(t) = S_1(t) \hat{\otimes} \mathbf{I}_2$, $\mathbf{S}_2(t) = \mathbf{I}_1 \hat{\otimes} S_2(t)$ and $\mathbf{S}(t) = S_1(t) \hat{\otimes} S_2(t)$ on $\mathcal{H} \hat{\otimes} \mathcal{H}$. The proof of the next result is easy.

Theorem 6.33. *The operators* $\mathbf{S}(t)$, $\mathbf{S}_i(t)$ *,* $i = 1, 2$ *are* C_0*-contraction semigroups with generators* $\mathcal{A} = \overline{A_1 \hat{\otimes} \mathbf{I}_2 + \mathbf{I}_1 \hat{\otimes} A_2}$, $\mathcal{A}_1 = A_1 \hat{\otimes} \mathbf{I}_2$, $\mathcal{A}_2 = \mathbf{I}_1 \hat{\otimes} A_2$, *and* $\mathbf{S}(t) = \mathbf{S}_1(t)\mathbf{S}_2(t) = \mathbf{S}_2(t)\mathbf{S}_1(t)$.

Let $S_i(t),\ 1 \leqslant i \leqslant n$ be a family of C_0-contraction semigroups with generators A_i defined on $\mathcal{H}$.

Corollary 6.34. $\mathbf{S}(t) = \hat{\otimes}_{i=1}^{n} S_i(t)$ *is a* C_0*-contraction semigroup on* $\hat{\otimes}_{i=1}^{n} \mathcal{H}$ *and the closure of* $A_1 \hat{\otimes} \mathbf{I}_2 \hat{\otimes} \cdots \hat{\otimes} \mathbf{I}_n + \mathbf{I}_1 \hat{\otimes} A_2 \hat{\otimes} \cdots \hat{\otimes} \mathbf{I}_n + \cdots \mathbf{I}_1 \hat{\otimes} \mathbf{I}_2 \hat{\otimes} \cdots \hat{\otimes} A_n$ *is the generator* $\mathcal{A}$ *of* $\mathbf{S}(t)$.

Returning to our general discussion, let $\mathbf{P}_{\varphi}^{s}$ denote the projection from $\mathcal{B}_{\otimes}^{\alpha}$ onto $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$, and let $\mathbf{P}_{\varphi}^{w}$ denote the projection from $\mathcal{B}_{\otimes}^{\alpha}$ onto $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^w$.

Theorem 6.35. *If* $\mathbf{T} \in L^{\#}(\mathcal{B}_{\otimes}^{\alpha})$*, then* $\mathbf{P}_{\varphi}^{s}\mathbf{T} = \mathbf{T}\mathbf{P}_{\varphi}^{s}$ *and* $\mathbf{P}_{\varphi}^{w}\mathbf{T} = \mathbf{T}\mathbf{P}_{\varphi}^{w}$.

Proof. The weak case follows from the strong case, so we prove that $\mathbf{P}_{\varphi}^{s}\mathbf{T} = \mathbf{T}\mathbf{P}_{\varphi}^{s}$. Since vectors of the form $\Phi = \sum_{i=1}^{L} \otimes_{t\in I}\varphi_t^i$, with $\varphi_t^i = \varphi_t$ for all but a finite number of t, are dense in $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$; it suffices to show that $\mathbf{T}\Phi \in \mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$. Now, $\mathbf{T} \in L^{\#}(\mathcal{B}_{\otimes}^{\alpha})$ implies that there exists a sequence of operators $\mathbf{T}_n$ such that $\|\mathbf{T} - \mathbf{T}_n\|_{\otimes}^{\alpha} \to 0$ as $n \to \infty$, where each $\mathbf{T}_n$ is of the form: $\mathbf{T}_n = \sum_{k=1}^{N_n} a_k^n T_k^n$, with a_k^n a complex scalar, $N_n < \infty$, and each $T_k^n = \hat{\otimes}_{t\in J_k} T_{kt}^n \hat{\otimes}_{t\in I\setminus J_k} I_t$ for some finite set of t-values J_k. Hence,

$$\mathbf{T}_n\Phi = \sum_{i=1}^{L} \sum_{k=1}^{N_n} a_k^n \otimes_{t\in J_k} T_{k,t}^n\varphi_t^i \otimes_{t\in I\setminus J_k} \varphi_t^i.$$

Now, it is easy to see that, for each i, $\otimes_{t\in J_k} T_{k,t}^n\varphi_t^i \otimes_{t\in I\setminus J_k} \varphi_t^i \equiv^s \otimes_{t\in I}\varphi_t$. It follows that $\mathbf{T}_n\Phi \in \mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ for each n, so that $\mathbf{T}_n \in L[\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s]$. As $L[\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s]$ is a norm closed algebra, $\mathbf{T} \in L[\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s]$ and it follows that $\mathbf{P}_{\varphi}^{s}\mathbf{T} = \mathbf{T}\mathbf{P}_{\varphi}^{s}$. □

Let $z_t \in \mathbf{C}$, $|z_t| = 1$, and define $U[\mathbf{z}]$ by: $U[\mathbf{z}] \otimes_{t\in I} \varphi_t = \otimes_{t\in I} z_t\varphi_t$.

Theorem 6.36. *The operator* $U[\mathbf{z}]$ *has a unique unitary extension to* $\mathcal{B}_{\otimes}^{\alpha}$*, which we also denote by* $U[\mathbf{z}]$*, such that:*

(1) $U[\mathbf{z}] : \mathcal{B}_{\otimes}^{\alpha}(\varphi)^w \to \mathcal{B}_{\otimes}^{\alpha}(\varphi)^w$*, so that* $\mathbf{P}_{\varphi}^{w}U[\mathbf{z}] = U[\mathbf{z}]\mathbf{P}_{\varphi}^{w}$.

(2) *If* $\prod_{t\in I} z_t$ *is quasi-convergent but not convergent, then* $U[\mathbf{z}] : \mathcal{B}_{\otimes}^{\alpha}(\varphi)^s \to \mathcal{B}_{\otimes}^{\alpha}(\eta)^s$*, for some* $\eta \in \mathcal{B}_{\otimes}^{\alpha}(\varphi)^w$ *with* $\varphi \perp \eta$.

(3) $U[\mathbf{z}] : \mathcal{B}_{\otimes}^{\alpha}(\varphi)^s \to \mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ *if and only if* $\prod_{t\in I} z_t$ *converges and* $U[\mathbf{z}] = (\prod_{t\in I} z_t)\mathbf{I}_{\otimes}$*. This implies that* $\mathbf{P}_{\varphi}^{s}U[\mathbf{z}] = U[\mathbf{z}]\mathbf{P}_{\varphi}^{s}$.

Proof. For (1), let $\psi = \sum_{k=1}^{N} \otimes_{t\in I}\psi_t^k$, where $\otimes_{t\in I}\psi_t^k \equiv^w \otimes_{t\in I}\varphi_t$, N is arbitrary and $1 \leqslant k \leqslant N$. Then

$$U^*[\mathbf{z}]U[\mathbf{z}]\psi = \sum_{k=1}^{N} \otimes_{t\in I} z_t^* z_t \psi_t^k = \psi = U[\mathbf{z}]U^*[\mathbf{z}]\psi.$$

It is clear that $U[\mathbf{z}]$ is a unitary operator, and since ψ of the above form are dense, $U[\mathbf{z}]$ extends to a unitary operator on $\mathcal{B}_{\otimes}^{\alpha}$. By definition,

$$\sum_{k=1}^{N} \otimes_{t\in I} z_t\psi_t^k \in \mathcal{B}_{\otimes}^{\alpha}(\varphi)^w \quad \text{if} \quad \sum_{k=1}^{N} \otimes_{t\in I}\psi_t^k \in \mathcal{B}_{\otimes}^{\alpha}(\varphi)^w,$$

so that $U[\mathbf{z}] : \mathcal{B}_{\otimes}^{\alpha}(\varphi)^w \to \mathcal{B}_{\otimes}^{\alpha}(\varphi)^w$ and $\mathbf{P}_{\varphi}^{w}U[\mathbf{z}] = U[\mathbf{z}]\mathbf{P}_{\varphi}^{w}$. To prove (2), use Theorem 6.20 to note that $\prod_{t\in I} z_t = 0$ and $\otimes_{t\in I}\psi_t^k \equiv^s \otimes_{t\in I}\varphi_t$

implies that $\otimes_{t\in I} z_t\psi_t^k \in \mathcal{B}_\otimes^\alpha(\eta)^s$ with $\mathcal{B}_\otimes^\alpha(\eta)^s \perp \mathcal{B}_\otimes^\alpha(\varphi)^s$. To prove (3), note that, if $0 < \left|\prod_{t\in I} z_t\right| < \infty$, then $U[\mathbf{z}] = [(\prod_{t\in I} z_t)\mathbf{I}_\otimes]$, so that $U[\mathbf{z}] : \mathcal{B}_\otimes^\alpha(\varphi)^s \to \mathcal{B}_\otimes^\alpha(\varphi)^s$. Now suppose that $U[\mathbf{z}] : \mathcal{B}_\otimes^\alpha(\varphi)^s \to \mathcal{B}_\otimes^\alpha(\varphi)^s$, then $\otimes_{t\in I} z_t\psi_t^k \equiv^s \otimes_{t\in I}\varphi_t$ and so $\prod_{t\in I} z_t$ must converge. Therefore, $U[\mathbf{z}]\psi = [(\prod_{t\in I} z_t)\mathbf{I}_\otimes]\psi$ and $\mathbf{P}_\varphi^s U[\mathbf{z}] = U[\mathbf{z}]\mathbf{P}_\varphi^s$.

It is easy to see that, for each fixed $t \in I$ and any $\mathbf{A}(t) \in L[\mathcal{B}(t)]$, $\mathbf{A}(t)$ commutes with any $\mathbf{P}_\varphi^s$, $\mathbf{P}_\varphi^w$ or $U[\mathbf{z}]$, where φ and $\mathbf{z}$ are arbitrary. □

Theorem 6.37. *Every* $\mathbf{T} \in L^\#[\mathcal{B}_\otimes^\alpha]$ *commutes with all* $\mathbf{P}_\varphi^s$, $\mathbf{P}_\varphi^w$ *and* $U[\mathbf{z}]$, *where* φ *and* $\mathbf{z}$ *are arbitrary. (In particular, every* $\mathbf{T} \in L^\#[\mathcal{B}_\otimes^\alpha]$ *is reducible on* $\mathcal{B}_\otimes^\alpha(\varphi)$ *for all* $\varphi \in \mathcal{B}_\otimes^\alpha$.*)*

Proof. Let $\mathfrak{L}$ be the set of all $\mathbf{P}_\varphi^s$, $\mathbf{P}_\varphi^w$ or $U[\mathbf{z}]$, with φ and $\mathbf{z}$ arbitrary. From the above observation, we see that all $\mathbf{A}(t) \in L[\mathcal{B}(t)]$, $t \in I$ commutes with $\mathfrak{L}$ and hence belongs to its commutator $\mathfrak{L}'$. Since $\mathfrak{L}'$ is a closed algebra, this implies that $L^\#[\mathcal{B}_\otimes^\alpha] \subseteq \mathfrak{L}'$ so that all $\mathbf{T} \in L^\#[\mathcal{B}_\otimes^\alpha]$ commute with $\mathfrak{L}$. □

6.5. The Film

In the world view suggested by Feynman, physical reality is laid out as a three-dimensional motion picture in which we become aware of the future as more and more of the film comes into view. In this section, we construct a mathematical version of Feynman's film for both Hilbert and Banach spaces.

6.5.1. Hilbert Film. We first consider separable Hilbert spaces. Let $\{e_i \mid i \in \mathbb{N}\}$ be a complete orthonormal basis for $\mathcal{H}$ and, for each $t \in I$ and $i \in \mathbb{N}$, let $e_{i,t} = e_i$ and set $E_i = \otimes_{t\in I} e_{i,t}$. The Hilbert space $\widehat{\mathcal{H}}$ generated by the family of vectors $\{E_i,\ i \in \mathbb{N}\}$ is isometrically isomorphic to $\mathcal{H}$ via the mapping $e_i \leftrightarrow E_i$. (For later use, it should be noted that any vector in $\mathcal{H}$ of the form $\varphi = \sum_{k=1}^\infty a_k e_k$ has the corresponding representation in $\widehat{\mathcal{H}}$ as $\hat{\varphi} = \sum_{k=1}^\infty a_k E_k$.) We cannot use $\widehat{\mathcal{H}}$ to construct our operator calculus because it is not invariant for any reasonable class of operators. However, $\widehat{\mathcal{H}}$ is very close to what we need.

Definition 6.38. A film, $\mathcal{FD}_\otimes^2$, is the smallest subspace of $\mathcal{H}_\otimes^2$ containing $\widehat{\mathcal{H}}$, which is an invariant subspace for $L^\#[\mathcal{H}_\otimes^2]$. We call $\mathcal{FD}_\otimes^2$ the Feynman–Dyson space (FD-space) over $\mathcal{H}$.

In order to construct our space, for each i, let $\mathcal{FD}_i^2 = \mathcal{H}_\otimes^2(E_i)$ be the strong partial tensor product space generated by the vector E_i. It is clear that $\mathcal{FD}_i^2$ is the smallest space in $\mathcal{H}_\otimes^2$ containing E_i that is invariant under $L^\#[\mathcal{H}_\otimes^2]$. If we let $\mathcal{FD}_\otimes^2 = \bigoplus_{i=1}^{\infty} \mathcal{FD}_i^2$, we obtain our film, a Hilbert (space) bundle over $I = [a, b]$. It is not a separable space, but the fiber at each time-slice is isomorphic to $\mathcal{H}$.

6.5.2. Banach Film. Let $\{e_i,\ (e_i)^* \mid i \in \mathbb{N}\}$ be a complete biorthonormal system for $\mathcal{B}$. (Recall that $\mathcal{B}$ has an S-basis.) For each $t \in I$ and $i \in \mathbb{N}$, let $e_{i,t} = e_i$, let $E_i = \otimes_{t \in I}\, e_{i,t}$, and set $E_i^* = \otimes_{t \in I}(e_{i,t})^*$. As before, the Banach space $\widehat{\mathcal{B}}$ generated by the family of vectors $\{E_i,\ i \in \mathbb{N}\}$ is isometrically isomorphic to $\mathcal{B}$.

Definition 6.39. A film, $\mathcal{FD}_\otimes^\alpha$, is the smallest subspace of $\mathcal{B}_\otimes^\alpha$ containing $\widehat{\mathcal{B}}$, which is an invariant subspace for $L^\#[\mathcal{B}_\otimes^\alpha]$. We call $\mathcal{FD}_\otimes^\alpha$ the Feynman–Dyson space (FD-space) over $\mathcal{B}$.

If we let $\mathcal{FD}_i{}^\alpha = \mathcal{B}_\otimes^\alpha(E_i)$ be the partial tensor product space generated by the vector E_i, then it is clear that $\mathcal{FD}_i^\alpha$ is the smallest space in $\mathcal{B}_\otimes^\alpha$ which contains the vector E_i. We now set $\mathcal{FD}_\otimes^\alpha = \bigoplus_{i=1}^{\infty} \mathcal{FD}_i^\alpha$.

6.6. Exchange Operator

We now assume that $I = [a, b] \subset \mathbb{R}$ and $L^\#[\mathcal{B}_\otimes^\alpha]$ is the uniform closure of the algebra generated by $\{L[\mathcal{B}(t)],\ t \in I\}$.

Definition 6.40. An exchange operator $\mathbf{E}[t, t']$ on $L^\#[\mathcal{B}_\otimes^\alpha]$ is a linear map defined for pairs t, t' such that:

(1) $\mathbf{E}[t, t'] : L[\mathcal{B}(t)] \to L[\mathcal{B}(t')]$, (isometric isomorphism),

(2) $\mathbf{E}[s, t']\mathbf{E}[t, s] = \mathbf{E}[t, t']$,

(3) $\mathbf{E}[t, t']\mathbf{E}[t', t] = I$,

(4) for $s \neq t,\ t'$, $\mathbf{E}[t, t']\mathcal{A}(s) = \mathcal{A}(s)$, for all $\mathcal{A}(s) \in L[\mathcal{B}(s)]$.

The exchange operator acts to exchange the time positions of a pair of operators in a more complicated expression.

Theorem 6.41 (Existence)**.** *There exists an exchange operator for* $L^\#[\mathcal{B}_\otimes^\alpha]$.

Proof. Define a map $C[t,t'] : \mathcal{B}_{\otimes}^{\alpha} \to \mathcal{B}_{\otimes}^{\alpha}$ (comparison operator) by its action on elementary vectors:

$$C[t,t'] \otimes_{s\in I} \phi_s = (\otimes_{a\leqslant s<t'}\phi_s) \otimes \phi_t \otimes (\otimes_{t'<s<t}\phi_s) \otimes \phi_{t'} \otimes (\otimes_{t<s\leqslant b}\phi_s),$$

for all $\phi = \otimes_{s\in I}\phi_s \in \mathcal{B}_{\otimes}^{\alpha}$. Clearly, $C[t,t']$ extends to an isometric isomorphism of $\mathcal{B}_{\otimes}^{\alpha}$. For $\mathbf{U} \in L^{\#}[\mathcal{B}_{\otimes}^{\alpha}]$, we define $\mathbf{E}[t,t']\,\mathbf{U} = C[t,t']\,\mathbf{U}C[t',t]$. It is easy to check that $\mathbf{E}[\,\cdot\,,\cdot\,]$ satisfies all the requirements for an exchange operator. □

Example 6.42. *Let $U = \otimes_{t\in I}U_t$, so that the action of $\mathbf{E}[t,t']\,\mathbf{U}$ on elementary vectors satisfies:*

$$\begin{aligned}
&\mathbf{E}[t,t']\mathbf{U}\left(\underset{t\in I}{\otimes}\phi_t\right) = C[t,t']\mathbf{U}C[t',t]\left(\underset{t\in I}{\otimes}\phi_t\right)\\
&= C[t,t']\mathbf{U}\left\{\left(\underset{a\leqslant s<t'}{\otimes}\phi_s\right)\otimes\phi_t\otimes\left(\underset{t'<s<t}{\otimes}\phi_s\right)\otimes\phi_{t'}\otimes\left(\underset{t<s\leqslant b}{\otimes}\phi_s\right)\right\}\\
&= C[t,t']\left\{\left(\underset{a\leqslant s<t'}{\otimes}U_s\phi_s\right)\otimes U_{t'}\phi_t\otimes\left(\underset{t'<s<t}{\otimes}U_s\phi_s\right)\otimes U_t\phi_{t'}\otimes\left(\underset{t<s\leqslant b}{\otimes}U_s\phi_s\right)\right\}\\
&= \left\{\left(\underset{a\leqslant s<t'}{\otimes}U_s\phi_s\right)\otimes U_t\phi_{t'}\otimes\left(\underset{t'<s<t}{\otimes}U_s\phi_s\right)\otimes U_{t'}\phi_t\otimes\left(\underset{t<s\leqslant b}{\otimes}U_s\phi_s\right)\right\}.
\end{aligned}$$

6.7. Appendix

The study of infinite tensor products of Banach spaces is an important but neglected area. It offers a natural arena for the constructive, but general study of analysis in infinitely many variables, including partial differential equations and path integrals. In this appendix, we introduce a few topics that have independent interest.

6.7.1. The Fourier Transform Again.

In Chap. 2, we defined the Fourier transform as a mapping from a uniformly convex Banach space to its dual space. This approach exploits the strong relationship between a uniformly convex Banach space and a Hilbert space at the expense of a restricted Fourier transform.

In addition to the definition in Chap. 2, it is also possible to define the Fourier transform, $\mathfrak{F}$, as a mapping on $L^1[\mathbb{R}_I^n]$ to $\mathbb{C}_0[\mathbb{R}_I^n]$ for all n as one fixed linear operator that extends to a definition on $L^1[\mathbb{R}_I^\infty]$. To do this requires a closer look at our Banach spaces defined on $\mathbb{R}_I^\infty$. Recall that $I = [-\frac{1}{2},\frac{1}{2}]$, $\bar{x} = (x_k)_{k=1}^n$, $\hat{x} = (x_k)_{k=n+1}^\infty$ and $h_n(\hat{x}) = \otimes_{k=n+1}^\infty \chi_I(x_k)$ with $I = [-\frac{1}{2},\frac{1}{2}]$. The measurable functions on $\mathbb{R}_I^n$, $\mathcal{M}_I^n$ are defined by $f_n(x) = f_n^n(\bar{x}) \otimes h_n(\hat{x})$, where $f_n^n(\bar{x})$ is measurable on $\mathbb{R}^n$, so that $\mathcal{M}_I^n$ is a partial tensor product subspace generated by the unit vector $h(x) = h_0(\hat{x})$. From this, we see that all of the spaces of functions considered in Chap. 2 are also partial tensor product spaces generated by $h(x)$. In this section we show how the replacement of $L^1[\mathbb{R}_I^n]$, $\mathbb{C}_0[\mathbb{R}_I^n]$ by $L^1[\mathbb{R}_I^n](h)$, $\mathbb{C}_0[\mathbb{R}_I^n](h)$ allows us to offer a different approach to the Fourier transform.

We define $\mathfrak{F}(f_n)(\mathbf{x})$, mapping $\mathbf{L}^1[\mathbb{R}_I^n](h)$ into $\mathbb{C}_0[\mathbb{R}_I^n](\hat{h})$ by

$$\mathfrak{F}(f_n)(\mathbf{x}) = \otimes_{k=1}^n \mathfrak{F}_k(f_n) \otimes_{k=n+1}^\infty \hat{h}_n(\hat{x}), \tag{6.2}$$

where the product of Sinc functions $\hat{h}_n(\hat{x}) = \left[\otimes_{k=n+1}^n \frac{sin(\pi y_k)}{\pi y_k}\right]$ is the Fourier transform of the product $\prod_{k=n+1}^\infty I$ of the interval I.

Theorem 6.43. *The operator $\mathfrak{F}$ extends to a bounded linear mapping of $L^1[\mathbb{R}_I^\infty](h)$ into $\mathbb{C}_0[\mathbb{R}_I^\infty](\hat{h})$.*

Proof. Since

$$\lim_{n\to\infty} L^1[\mathbb{R}_I^n](h) = \bigcup_{n=1}^\infty L^1[\mathbb{R}_I^n](h) = L^1[\mathbb{R}'^\infty_I](h)$$

and $L^1[\mathbb{R}_I^\infty](h)$ is the closure of $L^1[\mathbb{R}'^\infty_I](h)$ in the L^1-*norm* $= \Delta_1$, it follows that $\mathfrak{F}$ is a bounded linear mapping of $L^1[\mathbb{R}'^\infty_I](h)$ into $\mathbb{C}_0[\mathbb{R}_I^\infty](\hat{h})$.

Suppose that $\{f_n\} \subset L^1[\mathbb{R}'^\infty_I](h)$, converges to $f \in L^1[\mathbb{R}_I^\infty](h)$. Since the sequence is Cauchy, $\|f_n - f_m\|_1 \to 0$ as $m,\ n \to \infty$, it follows that

$$|\mathfrak{F}(f_n(\mathbf{x}) - f_m(\mathbf{x}))| \leqslant \int_{\mathbb{R}_I^\infty} |f_n(\mathbf{y}) - f_m(\mathbf{y})|\, d\lambda_\infty(\mathbf{y}) = \|f_n - f_m\|_1 .$$

Thus, $|\mathfrak{F}(f_n(\mathbf{x}) - f_m(\mathbf{x}))|$ is also a Cauchy sequence in $\mathbb{C}_0[\mathbb{R}_I^\infty](\hat{h})$. Since $L^1[\mathbb{R}'^\infty_I](h)$ is dense in $L^1[\mathbb{R}_I^\infty](h)$, it follows that $\mathfrak{F}$ has a bounded extension, mapping $L^1[\mathbb{R}_I^\infty](h)$ into $\mathbb{C}_0[\mathbb{R}_I^\infty](\hat{h})$. □

Corollary 6.44. *The operator $\mathfrak{F}$ extends to a bounded linear mapping of $L^1[\mathcal{B}](h)$ into $\mathbb{C}_0[\mathcal{B}](\hat{h})$.*

Just as for L^2, the Fourier transform is an isometric isomorphism from $KS^2[\mathbb{R}^n]$ onto $KS^2[\mathbb{R}^n]$.

Corollary 6.45. *The operator $\mathfrak{F}$ is an isometric isomorphism of $KS^2[\mathbb{R}_I^\infty](h)$ onto $KS^2[\mathbb{R}_I^\infty](\hat{h})$ and an isometric isomorphism of $KS^2[\mathcal{B}](h)$ onto $KS^2[\mathcal{B}](\hat{h})$.*

Thus, unlike the theory in Chap. 2, the natural interpretation is that the Fourier transform induces a Pontryagin duality like theory that does not depend on the group structure of $\mathbb{R}_I^\infty$, (or $\mathcal{B}$) but depends on the pairing of different function spaces. This approach is direct, constructive, and applies to all separable Banach spaces (with an S-basis). Thus, the group structure of the underlying measure space plays no role.

6.7.2. Unbounded Operators on $\mathcal{B}_\otimes^\alpha$. In this section, we assume that I is countable. For each $i \in I$, let A_i be a closed densely defined linear operator on $\mathcal{B}_i$, with domain $D(A_i)$, and let $\mathbf{A}_i$ be its extension to $\mathcal{B}_\otimes^\alpha$, with domain $D(\mathbf{A}_i) \supset \tilde{D}(\mathbf{A}_i) = D(A_i) \otimes (\otimes_{k \neq i} \mathcal{B}_k)$. The next theorem follows directly from the definition of the tensor product of semigroups and the fact that α is a faithful relative tensor norm.

Theorem 6.46. *Let A_i, $1 \leqslant i \leqslant n$ be generators of a family of C_0-semigroups $S_i(t)$ on $\mathcal{B}_i$ with $\|S_i(t)\|_{\mathcal{B}_i} \leqslant M_i e^{\omega_i t}$. Then $\mathbf{S}_n(t) = \hat{\otimes}_{i=1,n}^\alpha S_i(t)$, defined on $\hat{\otimes}_{i=1,n}^\alpha \mathcal{B}_i$, has a unique extension (also denoted by $\mathbf{S}_n(t)$) to all of $\mathcal{B}_\otimes^\alpha$, such that for all vectors $\sum_{k=1}^K \otimes_{i \in I} \varphi_i^k$ with $\varphi_l^k \in D(A_l)$, $1 \leqslant l \leqslant n$, the infinitesimal generator for $\mathbf{S}_n(t)$ satisfies:*

$$\mathbf{A}^n \left[\sum_{k=1}^K \otimes_{i \in I} \varphi_i^k \right] = \sum_{l=1}^n \sum_{k=1}^K A_l \varphi_l^k (\otimes_{i \in I}^{i \neq l} \varphi_i^k).$$

Definition 6.47. Let $\{A_i\}$, $i \in I$ be a family of closed densely defined linear operators on $\mathcal{B}_i$ and let $\varphi_i \in D(A_i)$ (respectively $\psi_i \in D(A_i)$), with $\|\varphi_i\|_{\mathcal{B}} = 1$ (respectively $\|\psi_i\|_{\mathcal{B}} = 1$), for all $i \in I$.

(1) We say that $\varphi = \otimes_{i \in I} \varphi_i$ is a strong convergence sum (scs)-vector for the family $\{\mathbf{A}_i\}$ if $\lim_{n \to \infty} \sum_{k=1}^n \mathbf{A}_k \varphi = \sum_{k=1}^\infty A_k \varphi_k (\otimes_{i \in I}^{i \neq k} \varphi_i)$ exists.

(2) We say that $\psi = \otimes_{i \in I} \psi_i$ is a strong convergence product (scp)-vector for the family $\{\mathbf{A}_i\}$ if $\lim_{n \to \infty} \prod_{k=1}^n \mathbf{A}_k \psi = \otimes_{i \in I} A_i \psi_i$ exists.

Let $\mathbf{D}_\varphi$ be the linear span of $\{\chi = \otimes_{i \in I} \chi_i,\ \chi_i \in D(A_i)\}$, with $\chi_i = \varphi_i$ (and let $\mathbf{D}_\eta$ be the linear span of $\{\eta = \otimes_{i \in I} \eta_i,\ \eta_i \in D(A_i)\}$, with $\eta_i = \psi_i$) for all $i > L$, where L is arbitrary but finite. Clearly, $\mathbf{D}_\varphi$ is dense in $\mathcal{B}_\otimes^\alpha(\varphi)^s$ ($\mathbf{D}_\eta$ is dense in $\mathcal{B}_\otimes^\alpha(\psi)^s$). If there is a possible chance for confusion, we let $\mathbf{A}_s$, respectively $\mathbf{A}_p$, denote the closure of $\sum_{k=1}^\infty \mathbf{A}_k$ on $\mathcal{B}_\otimes^\alpha(\varphi)^s$ (respectively $\prod_{k=1}^\infty \mathbf{A}_k$ on $\mathcal{B}_\otimes^\alpha(\psi)^s$). It follows that $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ (respectively $\mathcal{B}_{\otimes}^{\alpha}(\psi)^s$) are natural spaces for the study of infinite sums or products of unbounded operators. The notion of a strong convergence sum vector first appeared in Reed [RE].

Definition 6.48. We call $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ a **RS**-space (respectively a **RP**-space $\mathcal{B}_{\otimes}^{\alpha}(\psi)^s$) for the family $\{\mathbf{A}_i\}$.

Let $\{U_k(t)\}$ be a set of unitary groups on $\{\mathcal{H}_k\}$. It is easy to see that $U(t) = \hat{\otimes}_{k=1}^{\infty} U_k(t)$ is a unitary group on $\mathcal{H}_{\otimes}^2$. However, it need not be reduced on any partial tensor product subspace. The following results are due to Streit [ST] and Reed [RE], as indicated.

Theorem 6.49 (Streit)**.** *Suppose $\{\mathbf{A}_k\}$ is a set of self-adjoint linear operators on the space $\mathcal{H}_{\otimes}^2(\varphi)^s$, with corresponding unitary groups $\{U_k(t)\}$. If $U(t) = \hat{\otimes}_{k=1}^{\infty} U_k(t)$, then $\mathbf{P}_{\varphi}^s U(t) = U(t)\mathbf{P}_{\varphi}^s$ (i.e., $U(t)$ is reduced on $\mathcal{H}_{\otimes}^2(\varphi)^s$) and $U(t)$ is a strongly continuous unitary group on $\mathcal{H}_{\otimes}^2(\varphi)^s$ if and only if, for each $c > 0$, the following three conditions are satisfied:*

(1) $\sum_{k=1}^{\infty} |\langle \mathbf{A}_k E_k[-c, c]\varphi_k, \varphi_k\rangle| < \infty$,
(2) $\sum_{k=1}^{\infty} \left|\left\langle \mathbf{A}_k^2 E_k[-c, c]\varphi_k, \varphi_k\right\rangle\right| < \infty$,
(3) $\sum_{k=1}^{\infty} |\langle (I_k - E_k[-c, c])\varphi_k, \varphi_k\rangle| < \infty$,

where $E_k[-c, c]$ are the spectral projectors of $\mathbf{A}_k$ and, in this case, $U(t) = s - \lim_{n\to\infty} \hat{\otimes}_{k=1}^{n} U_k(t)$.

Corollary 6.50. *Conditions 1–3 are satisfied if and only if there exists a strong convergence vector $\varphi = \otimes_{k=1}^{\infty}\varphi_k$ for the family $\{A_k\}$ such that $\varphi_k \in D(A_k)$ and*

$$\sum\nolimits_{k=1}^{\infty} |\langle \mathbf{A}_k \varphi_k, \varphi_k\rangle| < \infty, \ \ \sum\nolimits_{k=1}^{\infty} \|\mathbf{A}_k \varphi_k\|^2 < \infty.$$

Theorem 6.51 (Reed)**.** *$U(t)$ is reduced on $\mathcal{H}_{\otimes}^2(\varphi)^s$ and $U(t)$ is a strongly continuous unitary group on $\mathcal{H}_{\otimes}^2(\varphi)^s$ if and only if $\varphi = \otimes_{k=1}^{\infty}\varphi_k$ is a strong convergence vector for the family $\{A_k\}$ and $\sum_{k=1}^{\infty} |\langle \mathbf{A}_k\varphi_k, \varphi_k\rangle| < \infty$. If each A_k is positive, the statement is true without the absolute value in the above. In either case, $\mathbf{A}$, the closure of $\sum_{k=1}^{\infty} \mathbf{A}_k$, is the generator of $U(t)$.*

The next result strengthens and extends Reed's theorem to contraction semigroups on Banach spaces (e.g., the positivity requirement above can be dropped).

Theorem 6.52. *Let $\{S_k(t)\}$ be a family of strongly continuous contraction semigroups with generators $\{A_k\}$ defined on $\{\mathcal{B}_k\}$, and let $\varphi = \otimes_{k=1}^{\infty}\varphi_k$ be a strong convergence vector for the family $\{A_k\}$. Then $\mathbf{S}(t) = \hat{\otimes}_{k=1}^{\infty} S_k(t)$ is reduced on $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ and is a strongly continuous contraction semigroup. If $\mathbf{S}(t) = \hat{\otimes}_{k=1}^{\infty} S_k(t)$ is reduced on $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ and is a strongly continuous contraction semigroup on $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$, then there exists a strong convergence vector $\psi = \otimes_{k=1}^{\infty}\psi_k \in \mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ for the family $\{A_k\}$.*

Proof. Let $\varphi = \otimes_{k=1}^{\infty}\varphi_k$ be a strong convergence vector for the family $\{A_k\}$. Without loss, we can assume that $\|\varphi_k\| = 1$. Let $\mathbf{S}_n(t) = \hat{\otimes}_{k=1}^{n} S_k(t)\hat{\otimes}(\otimes_{k=n+1}^{\infty} I_k)$ and observe that $\mathbf{S}_n(t)$ is a contraction semigroup on $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ for all finite n. Furthermore, its generator is the closure of $\mathbf{A}^n = \sum_{k=1}^{n} \mathbf{A}_k$, where $\mathbf{A}_k = A_k\hat{\otimes}(\otimes_{i\neq k}^{\infty} I_i)$. If n and m are arbitrary, then

$$\begin{aligned}[\mathbf{S}_n(t) - \mathbf{S}_m(t)]\,\varphi &= \int_0^1 \frac{d}{d\lambda}\left\{\mathbf{S}_n[\lambda t]\mathbf{S}_m[(1-\lambda)t]\right\}\varphi d\lambda \\ &= t\int_0^1 \mathbf{S}_n[\lambda t]\mathbf{S}_m[(1-\lambda)t]\left[\mathbf{A}^n - \mathbf{A}^m\right]\varphi d\lambda,\end{aligned}$$

where we have used the fact that if two semigroups commute, then their corresponding generators also commute. It follows that:

$$\left\|[\mathbf{S}_n(t) - \mathbf{S}_m(t)]\,\varphi\right\| \leqslant t\left\|[\mathbf{A}^n - \mathbf{A}^m]\,\varphi\right\|.$$

Since $\varphi = \otimes_{k=1}^{\infty}\varphi_k$ is a strong convergence vector for the family $\{A_k\}$, it follows that s - $\lim_{n\to\infty}\mathbf{S}_n(t) = \mathbf{S}(t)$ exists on a dense set in $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$. As $\|\mathbf{S}(t)\| \leqslant \overline{\lim}_{n\to\infty}\|\mathbf{S}_n(t)\| < \infty$, we see that $\mathbf{S}(t)$ is bounded. To see that it must be a contraction, choose n so large that $\left\|[\mathbf{S}_n(t) - \mathbf{S}(t)]\,\varphi\right\|_{\otimes} < \varepsilon\|\varphi\|_{\otimes}$. It follows that

$$\|\mathbf{S}(t)\varphi\|_{\otimes} \leqslant \|\mathbf{S}_n(t)\varphi\|_{\otimes} + \left\|[\mathbf{S}_n(t) - \mathbf{S}(t)]\,\varphi\right\|_{\otimes} < \|\varphi\|_{\otimes}(1+\varepsilon).$$

Thus, $\mathbf{S}(t)$ is a contraction operator on $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$. It is easy to check that it is a C_0-semigroup.

Now suppose that $\mathbf{S}(t) = \hat{\otimes}_{k=1}^{\infty} S_k(t)$ is a strongly continuous contraction semigroup which is reduced on $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$. It follows that the generator $\mathbf{A}$ of $\mathbf{S}(t)$ is m-dissipative, and hence defined on a dense domain $D(\mathbf{A})$ in $\mathcal{B}_{\otimes}^{\alpha}(\varphi)^s$ with $\mathbf{S}'(t)\psi = \mathbf{S}(t)\mathbf{A}\psi = \mathbf{A}\mathbf{S}(t)\psi$ for all $\psi \in D(\mathbf{A})$. Since any such ψ is of the form $\psi = \sum_{l=1}^{\infty}\psi^l = \sum_{l=1}^{\infty}\otimes_{k=1}^{\infty}\psi_k^l$, where $\psi^l = \otimes_{k=1}^{\infty}\psi_k^l$ is in $D(\mathbf{A})$. A simple computation shows that $\mathbf{A}\psi^l = \sum_{k=1}^{\infty}\mathbf{A}_k\psi^l$, so that any ψ^l is a strong convergence vector for the family $\{A_k\}$. □

It is easy to see that, in the second part of the theorem, we cannot require that $\varphi = \otimes_{k=1}^{\infty}\varphi_k$ itself be a strong convergence vector for the family $\{A_k\}$ since it need not be in the domain of $\mathbf{A}$. For example, $\varphi_1 \notin D(A_1)$, while $\varphi_k \in D(A_k),\ k \neq 1$.

Example 6.53. *Let $\mathbf{A}_i$ be the generator of a C_0-contraction semigroup $T_i(t)$ on $\mathbb{C}_0[X_i]$ for each $i \in I$, and assume that $T_i(t)$ has the representation:*

$$T_i(t)\varphi_i(\mathbf{x}) = \int_{X_i} K_i[\mathbf{x}, t; \mathbf{y}, 0]\varphi_i(\mathbf{y})dm_i(\mathbf{y}).$$

Where m_i is an associated measure and $K_i[\mathbf{x}, t; \mathbf{z}, s]$ is a kernel function which satisfies

$$\int_{X_i} K_i[\mathbf{x}, t; \mathbf{z}, s]K_i[\mathbf{z}, s; \mathbf{y}, 0]dm_i(\mathbf{z}) = K_i[\mathbf{x}, t + s; \mathbf{y}, 0].$$

Let $\varphi_i \in \ker\{\mathbf{A}_i\}$, with $\|\varphi_i\|_{X_i} = 1$ for each $i \in I$, and note that $\varphi_i \in \ker\{\mathbf{A}_i\} \Rightarrow T_i(t)\varphi_i = \varphi_i$. With $\varphi = \otimes_{i\in I}\varphi_i$, construct $\mathbb{C}_{\otimes}^{\lambda}[\varphi]^s$. It follows that, for any $\psi = \sum_{j=1}^{m} \otimes_{i\in I}\psi_i^j$ with $\psi_i^j \in D(\mathbf{A}_i)$ and $\psi_i^j = \varphi_i$ for all but a finite number of i for each j, we have that the operator

$$\mathbf{A}^n\psi = \sum\nolimits_{k=1}^{n} \mathbf{A}_k\psi = \sum\nolimits_{k=1}^{n} \sum\nolimits_{j=1}^{m} A_k\psi_k^j \otimes_{i\neq k} (\otimes_{i\in I}\psi_i^j)$$

is finite and well defined on a dense set D in $\mathbb{C}_{\otimes}^{\lambda}[\varphi]^s$ and hence has a closure, which we also denote by $\mathbf{A}^n$.

From Theorems 6.33 and 6.52, we have:

Theorem 6.54. *For each n, $\mathbf{A}^n$ is the generator of a C_0-contraction semigroup $\mathbf{T}^n(t)$ on $\mathbb{C}_{\otimes}^{\lambda}[\varphi]^s$ and*

(1) *$s - \lim_{n\to\infty} \mathbf{A}^n = \mathbf{A}$ has a closure which generates a C_0-contraction semigroup $\mathbf{T}(t)$,*

(2) *$s - \lim_{n\to\infty} \mathbf{T}^n(t) = \mathbf{T}(t)$,*

(3) *for all $F(\mathbf{x}) \in \mathbb{C}_0[X]$,*

$$\mathbf{T}(t)F(\mathbf{x}) = \int_X \mathbf{K}[\mathbf{x}, t\,; \mathfrak{D}\mathbf{y}, 0]F(\mathbf{y}),$$

where $\mathbf{x} = (\mathbf{x}_1,\ \mathbf{x}_2,\ \cdots),\ \ \mathbf{y} = (\mathbf{y}_1,\ \mathbf{y}_2,\ \cdots)$, and

$$\mathbf{K}\left[\mathbf{x}, t : \mathfrak{D}\mathbf{y}, 0\right] = \mathop{\otimes}_{i=1}^{\infty} K_i\left[\mathbf{x}_i, t : \mathbf{y}_i, 0\right] dm_i\left(\mathbf{y}_i\right).$$

Example 6.55. *Let* $\{m_i\}$ *be a family of probability measures on* $\mathbb{R}_I$*, and let* m *be the induced version of the family on* $\mathbb{R}_I^\infty$*. Let* $\phi_i = a_i x_i \in \mathbf{L}^p[\mathbb{R}_I, m_i]$*, with* $0 < \prod_{i\in I} |a_i| < \infty$*,* $\|\phi_i\|_{\mathbb{R}_I}^p = 1$ *and construct* $\mathbf{L}_{\otimes}^{\Delta_p}[\phi]^s \cong \mathbf{L}^p[\mathbb{R}_I^\infty, m]$*. Let* $\{\delta_i(x_i)\}$ *be a family of functions such that* $\sum_{i=1}^\infty \|a_i\delta_i\|^p < \infty$ *and define* $A_i = \frac{1}{2}\sigma_{ii}(x_i)\frac{\partial^2}{\partial x_i^2} - \delta_i(x_i)\frac{\partial}{\partial x_i}$*, where* $0 < \sigma_{ii}(x_i)$*. Since* $\frac{\partial \phi_i}{\partial x_i} = D_i\phi_i = a_i$ *and* $\Delta_i\phi_i = \frac{\partial^2\phi_i}{\partial x_i^2} = 0$*, it is easy to see that* $\phi_i \in D(A_i)$ *for each* i*. It follows that* $\phi = \otimes_{i=1}^\infty \phi_i \in \mathbf{L}^p[\mathbb{R}_I^\infty, m]$ *is a strong convergence vector for the family* $\{A_i\}$ *and a strong convergence product vector for the family* $\{D_i\}$*.*

Theorem 6.56. *With the conventions as above:*

(1) *The closure of the operator* $\mathbf{A} = \sum_{i=1}^\infty [\frac{1}{2}\sigma_{ii}(x_i)\frac{\partial^2}{\partial x_i^2} - \delta_i(x_i)\frac{\partial}{\partial x_i}]$ *is a densely defined generator of a contraction semigroup on* $\mathbf{L}^p[\mathbb{R}_I^\infty, m]$*.*

(2) *The closure of* $D = \frac{\partial^\infty}{\partial x_1 \partial x_2 \cdots}$ *is a densely defined linear operator on* $\mathbf{L}^p[\mathbb{R}_I^\infty, m]$*.*

Remark 6.57. Theorem 6.56 can easily be shown to apply to any Banach space with an S-basis, with minor changes. Compare this with Theorem 2.102 of Chap. 2.

Discussion

The following special cases have appeared in the literature:

(1) If, in our definition of $\mathbf{A}$, we set $\delta(x_i) = 0$ and $\sigma_{ii}(x_i) = 2$, we get the natural infinite dimensional Laplacian:

$$\mathbf{A} = \Delta_\infty = \sum\nolimits_{i=1}^\infty \partial^2/\partial x_i^2.$$

(2) If $\delta(x_i) = -b_i x_i$ and $\sigma_{ii}(x_i) = 1$, we get the nonterminating diffusion generator in infinitely many variables (also known as the Ornstein–Uhlenbeck operator):

$$\mathbf{A} = \tfrac{1}{2}\Delta_\infty - B\mathbf{x}\cdot\nabla_\infty = \tfrac{1}{2}\sum\nolimits_{i=1}^\infty \partial^2/\partial x_i^2 - \sum\nolimits_{i=1}^\infty b_i x_i \partial/\partial x_i.$$

(3) If $\delta(x_i) = \frac{-x_i}{c^2}$ and $\sigma_{ii}(x_i) = 2$, we get the infinite dimensional Laplacian of Umemura [UM]:

$$\mathbf{A} = \sum_{i=1}^\infty \left(\frac{\partial^2}{\partial x_i^2} - \frac{x_i}{c^2}\frac{\partial}{\partial x_i}\right).$$

Berezanskii and Kondratyev [BK, pp. 520–521] have also discussed operators analogous to (2) and (3).

Open Problem

In this section, we identify an interesting problem that we believe is worthy of further study.

From our definition of Δ:

$$\Delta = \left\{ \{\phi_\nu\} \mid 0 \neq \|\otimes_{\nu\in I}\phi_\nu\|_{\mathcal{H}^2_\otimes}, \ \& \ \sum\nolimits_{\nu\in I} |1 - \|\phi_\nu\|_{\mathcal{B}}| < \infty \right\},$$

we see that every nonzero basic vector in $\mathcal{B}^\alpha_\otimes$ is nonzero in $\mathcal{H}^2_\otimes$. This raises an important question, but we first need a little background.

Recall that $\left(\mathbf{L}^1[X_i]\right)^{**} = \mathcal{M}[X_i]$, where $\mathcal{M}[X_i]$ is the set of bounded, regular, complex-valued measures on X_i that are absolutely continuous with respect to m_i (see below). We define the (total) variation of μ in $\mathcal{M}[X_i]$ by:

$$|\mu|(X_i) = \sup_{ess.\sup|h(x)|\leqslant 1} \left| \int_{X_i} h(x) d\mu(x) \right|.$$

The sup is over $h \in \mathbf{L}^\infty[X_i]$, and $|\cdot|$ is the induced norm on $\mathcal{M}[X_i]$. Since $\mathcal{M}[X_i]$ is a separable Banach space, construct $\mathcal{H}^1_i \subset \mathcal{M}[X_i] \subset \mathcal{H}^2_i$.

Definition 6.58. If μ, μ' are any two measures in $\mathcal{M}$:

(1) We say that μ' is singular with respect to μ and write it as $\mu' \perp \mu$ if, for each $\varepsilon > 0$, there exists a set $\Omega \subset X_i$ such that $\mu'(\Omega) < \varepsilon$ and $\mu(X_i) \setminus \Omega) < \varepsilon$.

(2) We say that μ' is absolutely continuous with respect to μ and write it as $\mu' \ll \mu$ if, for each set $\Omega \subset X_i$ such that $\mu(\Omega) = 0, \Rightarrow \mu'(\Omega) = 0$.

(3) If $\mu' \ll \mu$ and $\mu \ll \mu'$, we say that μ and μ' are equivalent and write it as $\mu' \approx \mu$.

If we define the square root of a complex function using the principal branch, in the third case, by the Radon–Nikodym theorem there exist (unique) measurable complex-valued functions $p'(x)$, $p(x)$ such that $p'(x) = d\mu'(x)/d\mu(x)^c$, and $p(x) = d\mu(x)/d\mu'(x)^c$, where a^c is the complex conjugate of a. If we set

$$H_i(\mu, \mu') = \int_{X_i} \sqrt{d\mu(x)}\sqrt{d\mu'(x)^c} = \int_{X_i} \sqrt{\left(d\mu(x)/d\mu'(x)^c\right)}\, d\mu'(x)^c$$

$$= \int_{X_i} \sqrt{\left(d\mu'(x)^c/d\mu(x)\right)}\, d\mu(x) = \int_{X_i} \sqrt{\left(d\mu(x)/d\lambda\right)\left(d\mu'(x)^c/d\lambda\right)}\, d\lambda,$$

we obtain a complex version of the Hellinger integral, which defines a complex inner product, where λ is any positive measure with $\mu \ll \lambda$ and $\mu' \ll \lambda$ (for example, $\lambda = m_i \vee \frac{1}{2}|\mu + \mu'|$). In this case, $H_i(\mu, \mu')^c = H_i(\mu', \mu)$ and $\mu' \approx \mu \Rightarrow H_i(\mu, \mu') \neq 0$. It is easy to see that $H_i(\mu, \mu) \leqslant ([|\mu|]^{1/2})^2(X_i) = \|\mu\|_i$, so, without loss, we can assume that $H_i(\mu, \mu') = (\mu, \mu')_{2i}$ is the inner product for our Hilbert space $\mathcal{H}_{2i}$.

If γ is the natural tensor norm for the space of measures, so that $\mathcal{M}[X_i] \hat{\otimes}^\gamma \mathcal{M}[X_j] = \mathcal{M}[X_i \times X_j]$, we can construct $\hat{\otimes}^\gamma_{i \in \mathbb{N}} \mathcal{M}_i = \mathcal{M}^\gamma_\otimes$ so that $\mathcal{H}^1_\otimes \subset \mathcal{M}^\gamma_\otimes \subset \mathcal{H}^2_\otimes$. For each λ_i, $\mu_i \in \mathcal{M}_i$, let λ_i^*, μ_i^* be the Steadman duality maps, where $\langle \mu_i, \lambda_i^* \rangle_i = (\mu_i, \lambda_i)_{2i} \left(\|\lambda_i\|^2_{\mathcal{M}} \big/ \|\lambda_i\|^2_{\mathcal{H}_2} \right)$ and $\langle \lambda_i, \mu_i^* \rangle_i = (\lambda_i, \mu_i)_{2i} \left(\|\mu_i\|^2_{\mathcal{M}} \big/ \|\mu_i\|^2_{\mathcal{H}_2} \right)$. We now have the following problem:

(I) Is it true that for $\mu = \otimes_{i \in \mathbb{N}} \mu_i$, $\lambda = \otimes_{i \in \mathbb{N}} \lambda_i$ in $\mathcal{M}^\gamma_\otimes$ with $\mu_i \approx \lambda_i$ for each $i \in \mathbb{N}$, we have that $\mu \equiv^s \lambda \Leftrightarrow \mu \approx \lambda$ (so that $\mu \in \mathcal{M}^\gamma_\otimes(\lambda)^s$) and $\mu \perp \lambda \Leftrightarrow \mu \notin \mathcal{M}^\gamma_\otimes(\lambda)^s$?

von Neumann [VN2] first mentioned this problem, in a restricted sense, in relation to the decomposition of $\mathcal{H}^2_\otimes$ into orthogonal subspaces and the theory of probability measures on infinite product spaces. (Note that his incomplete direct product is our partial tensor product.) He stated that: "Another application of our theory could be made to the theory of measures in infinite product spaces, which is the basis for the modern theory of probabilities. Here a certain incomplete direct product of $\mathcal{H}^2_\otimes$ is fundamental."

Ten years later, Kakutani [KA], in Chap. 5, published his now famous paper on the equivalence and orthogonality of infinite product measures. In the second paragraph of the introduction to his paper, Kakutani states: "In particular, the introduction of the inner product and isometric embedding of $\mathfrak{M}(\Omega, \mathfrak{B}, m)$ (set of all probability measures on $(\Omega, \mathfrak{B}, m)$) into a general Euclidean space (Hilbert space), as well as the indication of the relationship of this paper with earlier works of E. Hellinger, are due to Professor J. von Neumann."

The space $\mathfrak{M}(\Omega, \mathfrak{B}, m)$ is not a Banach space, but each element has norm 1 in the space of measures and the embedding Hilbert space. In our case:

$$\mathcal{H}^2_\otimes(\mu) \supset \mathcal{M}^\gamma_\otimes(\mu) \supset \mathfrak{M}^\gamma_\otimes(\mu).$$

If $\mathfrak{M}_{\otimes}^{\gamma}(\mu)$ contains an orthonormal basis for $\mathcal{H}_{\otimes}^{2}(\mu)$, for every μ, we would have a positive answer to **(I)**.

If the answer to **(I)** is true, this would explain the appearance of this phenomenon in general and would provide insight into the causes for the failure of certain expected/desired properties of (probability) measures on infinite dimensional spaces. These failures could then be directly linked to the breaking up of the infinite tensor product spaces into orthogonal subspaces as described by Theorem 6.20.

References

[A] R.A. Adams, *Sobolev Spaces* (Academic, New York, 1975)

[BK] Yu.M. Berezanskii, Yu.G. Kondratyev, *Spectral Methods in Infinite Dimensional Analysis* (Naukova Dumka, Kiev, 1988) [Russian]

[DSH] N. Dunford, R. Schatten, On the associate and conjugate space for the direct product of Banach spaces. Trans. Am. Math. Soc. **59**, 430–436 (1946)

[GG] B.R. Gelbaum, J. Gil de Lamadrid, Bases of tensor products of Banach spaces. Pac. J. Math. **11**, 1281–1286 (1961)

[GPZ] T.L. Gill, G. Pantsulaia, W.W. Zachary, Constructive analysis in infinitely many variables. Commun. Math. Anal. **13**, 107–141 (2012)

[GZ1] T.L. Gill, W.W. Zachary, Feynman operator calculus: the constructive theory. Expo. Math. **29**, 165–203 (2011)

[GU] A. Guichardet, *Symmetric Hilbert Spaces and Related Topics*. Lectures Notes in Mathematics, vol. 261 (Springer, New York, 1969)

[IC70] T. Ichinose, On the spectra of tensor products of linear operators in Banach spaces. J. Reine Angew. Math. **244**, 119–153 (1970)

[KP] S. Kwapień, Isomorphic characterization of inner product spaces by orthogonal series with vector valued coefficients. Stud. Math. **44**, 583–595 (1972)

[PA] G. Pantsulaia, *Invariant and Quasiinvariant Measures in Infinite-Dimensional Topological Vector Spaces* (Nova Science Publishers, New York, 2007)

[RE] M.C. Reed, On self-adjointness in infinite tensor product spaces. J. Funct. Anal. **5**, 94–124 (1970)

[RS] M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic, New York, 1972)

[S] R. Schatten, *A Theory of Cross-Spaces* (Princeton University Press, Princeton, 1950)

[ST] L. Streit, Test function spaces for direct product representations. Commun. Math. Phys. **4**, 22–31 (1967)

[VN2] J. von Neumann, On infinite direct products. Compos. Math. **6**, 1–77 (1938)

Chapter 7

The Feynman Operator Calculus

Introduction

In response to the importance of time-ordering in relating the Feynman and Schwinger–Tomonaga theories, Segal [SG], in Chap. 6, suggested that the provision of (real) mathematical meaning for time-ordering is one of the major problems in the foundations for QED.

A number of investigators have attempted to solve this problem using (formal methods of) functional analysis and operator algebras. Miranker and Weiss [MW] showed how the ordering process could be done (in a restricted manner) using the theory of Banach algebras. Nelson [N] also used Banach algebras to develop a theory of "operants" as an alternate (formal) approach. Araki [AK], motivated by the work of Fujiwara [FW], used yet another formal approach to the problem. Other workers include Maslov [M, in Chap. 8], who used the idea of a T-product as an approach to formally order the operators and developed an operational theory. An idea that is closest to that of Feynman and the one discussed in this chapter was developed by Johnson and Lapidus in a series of papers. Their work can be found in the recent book on the subject [JL] in Chap. 8.

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_7

A major difficulty with each approach (other than [JL] in Chap. 8) is the problem of disentanglement, the method proposed by Feynman to relate his results to conventional analysis. Johnson and Lapidus develop a general ordering approach via a probability measure on the parameter space. This approach is also constructive and offers a different perspective on possible frameworks for disentanglement in the Feynman program.

Summary. In this chapter, we first explain what the Feynman–Fujiwara notion of time-ordering means in operational terms. Then we construct the time-ordered integral and extend a few important theorems of semigroup theory to the time-ordered setting.

A general perturbation theory is developed and use it to prove that all theories generated by semigroups are asymptotic in the operator-valued sense of Poincaré. As an application, we prove Dyson's conjecture that the perturbation expansion of QED is asymptotic. This also enables us to develop a general theory for the interaction representation of relativistic quantum theory. We then provide a rigorous development of the disentanglement method suggested by Feynman and Fujiwara to relate his theory to the traditional approach. As an application of this result, we prove that the Trotter–Kato theory is a special case. Finally, we show that the theory can be reformulated as a physically motivated sum over paths.

The operator algebra $L^{\#}[\mathcal{H}_{\otimes}^{2}]$ of the last chapter allows us to give a constructive definition of the formal idea of time-ordering, using the natural order of the interval $I = [a, b] \subset \mathbb{R}$. In particular, for one operator $A(t) \in L[\mathcal{H}]$, $\mathcal{A}(t) \in L^{\#}[\mathcal{H}_{\otimes}^{2}]$ becomes:

$$\mathcal{A}(t) = \overline{\left(\bigotimes_{b \geqslant s > t} I_s\right) \otimes A(t) \otimes \left(\bigotimes_{t > s \geqslant a} I_s\right)},$$

where I_s is the identity operator at time s. It follows that the true operator $A(t)$ only acts at time t, while at all other times in $[a, b]$, $\mathcal{A}(t)$ is the identity operator. This is the exact implementation of Fujiwara's suggestion for the mathematical modeling of Feynman's formal idea.

If we have a family $\{A(t),\ t \in I\} \subset L(\mathcal{H})$, then the operators $\{\mathcal{A}(t),\ t \in I\} \subset L^{\#}(\mathcal{H}_{\otimes}^{2})$ commute when acting at different times. For example, if $t > \tau$, then

$$\mathcal{A}(t)\mathcal{A}(\tau) = \overline{\left(\bigotimes_{b \geqslant s > t} I_s\right) \otimes A(t) \otimes \left(\bigotimes_{t > s \geqslant \tau} I_s\right) \otimes A(\tau) \otimes \left(\bigotimes_{\tau > s \geqslant a} I_s\right)}$$
$$= \mathcal{A}(\tau)\mathcal{A}(t).$$

Thus, our approach is constructive in that we use a sheet of unit operators at every point except at time t, where the true operator is placed, so that operators acting at different times actually commute. A major purpose of this chapter is to show that, using the Feynman–Dyson space, $\mathcal{FD}_{\otimes}^2$, we can lift all of the analysis and operator theory to the time-ordered setting.

In Sect. 7.2 we prove our fundamental theorem showing the existence of time-ordered integrals. This allows us to extend basic semigroup theory to the time-ordered setting, providing, among other results, a time-ordered version of the Hille–Yosida Theorem. In Sect. 7.3 we construct time-ordered evolution operators and prove that they have all the expected properties.

In Sect. 7.4 we provide a precise definition of the term "asymptotic in the sense of Poincaré" for operator-valued functions. We develop a general perturbation theory for time-ordered evolution equations and prove that all theories generated by evolution operators are asymptotic in the operator-valued sense of Poincaré. It is now known from experiment that Hagg's Theorem on the nonexistence of the interaction representation in sharp time does not apply, since there is some time overlap of wave packets. As an application, we give a generalization of the Dyson expansion and provide a general theory for the interaction representation used in relativistic quantum theory, when any time overlap of wave packets is allowed.

7.1. Time-Ordered Operators

7.1.1. Integrals and Generation Theorems. The following notation will be used at various points of this section, so we record the meanings here for reference. (The t value referred to is in our fixed interval I.)

(1) (e.f.o) means: "except for at most one t value";

(2) (e.f.f) means: "except for an at most finite number of t values"; and

(3) (e.f.c) means: "except for an at most countable number of t values."

We assume that, for each $t \in I$, $A(t)$ generates a C_0-semigroup on $\mathcal{H}$. Define $\mathbf{S}_t(\tau)$ by:

$$\mathbf{S}_t(\tau) = \hat{\otimes}_{s\in[b,t)}\mathrm{I}_s \otimes (\exp\{\tau A(t)\}) \otimes \left(\otimes_{s\in(t,a]}\mathrm{I}_s\right). \tag{7.1}$$

We want to briefly investigate the relationship between $S_t(\tau) = \exp\{\tau A(t)\}$ and $\mathbf{S}_t(\tau) = \exp\{\tau \mathcal{A}(t)\}$. By Theorem 6.33 of Chap. 6, we know that $\mathbf{S}_t(\tau)$ is a C_0-semigroup for $t \in I$ if and only if $S_t(\tau)$ is one also. For additional insight, we need a dense core for the family $\{\mathcal{A}(t)\,|t\ \in I\}$, so let $\bar{D} = \underset{t\in I}{\otimes} D(A(t))$ and set $\mathrm{D}_0 = \bar{D} \cap \mathcal{FD}^2_\otimes$. Since $\bar{D}$ is dense in $\mathcal{H}^2_\otimes$, it follows that D_0 is dense in $\mathcal{FD}^2_\otimes$. Using our basis, if $\Phi, \Psi \in \mathrm{D}_0$, $\Phi = \sum_{\mathrm{i}}\sum_{F(f)} a^{\mathrm{i}}_{F(f)} E^{\mathrm{i}}_{F(\mathrm{f})}$, $\Psi = \sum_{\mathrm{i}}\sum_{F(g)} b^{\mathrm{i}}_{F(g)} E^{\mathrm{i}}_{F(g)}$; then, as $\exp\{\tau\mathcal{A}(t)\}$ is invariant on $\mathcal{FD}^i_2$ for each i, we have

$$\langle \exp\{\tau\mathcal{A}(t)\}\Phi,\Psi\rangle = \sum\nolimits_{\mathrm{i}}\sum\nolimits_{F(f)}\sum\nolimits_{F(g)} a^{\mathrm{i}}_{F(f)}\bar{b}^{\mathrm{i}}_{F(g)}\left\langle \exp\{\tau\mathcal{A}(t)\}E^{\mathrm{i}}_{F(f)}, E^{\mathrm{i}}_{F(g)}\right\rangle,$$

and

$$\left\langle \exp\{\tau\mathcal{A}(t)\}E^{\mathrm{i}}_{F(f)}, E^{\mathrm{i}}_{F(g)}\right\rangle = \prod_{s\neq t}\left\langle e^{\mathrm{i}}_{s,f(s)}, e^{\mathrm{i}}_{s,g(s)}\right\rangle \left\langle \exp\{\tau A(t)\}e^{\mathrm{i}}_{t,f(t)}, e^{\mathrm{i}}_{t,g(t)}\right\rangle$$

$$= \left\langle \exp\{\tau A(t)\}e^{\mathrm{i}}_{t,f(t)}, e^{\mathrm{i}}_{t,f(t)}\right\rangle \text{ (e.f.o)},$$

$$= \left\langle \exp\{\tau A(t)\}e^{\mathrm{i}}, e^{\mathrm{i}}\right\rangle \text{ (e.f.f.) implies}$$

$$\langle \exp\{\tau\mathcal{A}(t)\}\Phi,\Psi\rangle = \sum\nolimits_{\mathrm{i}}\sum\nolimits_{F(f)} a^{\mathrm{i}}_{F(f)}\bar{b}^{\mathrm{i}}_{F(f)}\left\langle \exp\{\tau A(t)\}e^{\mathrm{i}}, e^{\mathrm{i}}\right\rangle (e.f.c).$$

Thus, by working on $\mathcal{FD}^2_\otimes$, we obtain a simple direct relationship between the conventional and time-ordered version of a semigroup.

We now consider the general case. Let $\mathcal{A}_z(t) = z\mathcal{A}(t)\mathbf{R}(z,\mathcal{A}(t))$, where $\mathbf{R}(z,\mathcal{A}(t))$ is the resolvent of $\mathcal{A}(t)$.

By Theorem 5.23 of Chap. 5 (Yosida approximator), $A_z(t)$ generates a uniformly bounded semigroup and $\lim_{z\to\infty} A_z(t)\phi = A(t)\phi$ for $\phi \in \mathrm{D}(A(\mathrm{t}))$.

Theorem 7.1. *The operator $\mathcal{A}_z(t)$ satisfies*

(1) $\mathcal{A}(t)\mathcal{A}_z(t)\Phi = \bar{\mathcal{A}}_z(t)\mathcal{A}(t)\Phi$, $\Phi \in D_0$, $\mathcal{A}_z(t)$ *generates a uniformly bounded contraction semigroup on* $\mathcal{FD}^2_\otimes$ *for each* t, *and* $\lim_{z\to\infty}\mathcal{A}_z(t)\Phi = \mathcal{A}(t)\Phi$, $\Phi \in D_0$.

(2) *For each* n, *each set* $\tau_1, \cdots, \tau_n \in I$ *and each set* $a_1, \cdots, a_n$, $a_i \geqslant 0$; $\sum_{i=1}^n a_i\mathcal{A}(\tau_i)$ *generates a* C_0*-semigroup on* $\mathcal{FD}^2_\otimes$.

Proof. The proof of (1) follows from properties of the Yosida approximator and the relationship between $\mathcal{A}(t)$ and $A(t)$. It is an easy computation to check that (2) follows from Theorem 6.33 of Chap. 6, with $\mathbf{S}(t) = \prod_{i=1}^{n} \mathbf{S}_{\tau_i}(a_i t)$. □

We now assume that $A(t)$, $t \in I$, is weakly continuous and that $D(A(t))$ is dense in $\mathcal{H}$. It follows that this family has a weak KH-integral $Q[a,b] = \int_a^b A(t)dt \in C(\mathcal{H})$ (the closed densely defined linear operators on $\mathcal{H}$). Furthermore, it is not difficult to see that $A_z(t)$, $t \in I$, is also weakly continuous and hence the family $\{A_z(t) \mid t \in I\} \subset L(\mathcal{H})$ has a weak HK-integral $Q_z[a,b] = \int_a^b A_z(t)dt \in L(\mathcal{H})$.

Let $\mathrm{P_n}$ be a sequence of HK-partitions for $\delta_n(t): [a,b] \to (0,\infty)$ with $\delta_{n+1}(t) \leq \delta_n(t)$ and $\lim_{n\to\infty} \delta_n(t) = 0$, so that the mesh $\mu_n = \mu(\mathrm{P_n}) \to 0$ as $\mathrm{n} \to \infty$. Set $Q_{z,n} = \sum_{l=1}^{n} A_z(\bar{t}_l)\Delta t_l$, $Q_{z,m} = \sum_{q=1}^{m} A_z(\bar{s}_q)\Delta s_q$; $\mathbf{Q}_{z,n} = \sum_{l=1}^{\mathrm{n}} \mathcal{A}_z(\bar{t}_l)\Delta t_l$, $\mathbf{Q}_{z,m} = \sum_{q=1}^{\mathrm{m}} \mathcal{A}_z(\bar{s}_q)\Delta s_q$; and $\Delta Q_z = Q_{z,n} - Q_{z,m}$, $\Delta \mathbf{Q}_z = \mathbf{Q}_{z,n} - \mathbf{Q}_{z,m}$. Let $\Phi, \Psi \in \mathrm{D}_0$; $\Phi = \sum_i^J \Phi^{\mathrm{i}} = \sum_i^J \sum_{F(f)}^K a_{F(f)}^{\mathrm{i}} E_{F(f)}^{\mathrm{i}}$, $\Psi = \sum_i^L \Psi^{\mathrm{i}} = \sum_i^L \sum_{F(g)}^M b_{F(g)}^{\mathrm{i}} E_{F(g)}^{\mathrm{i}}$. Then we have:

Theorem 7.2 (Fundamental Theorem for Time-Ordered Integrals)**.**

(1) *The family* $\{\mathcal{A}_z(t) \mid t \in I\}$ *has a weak KH-integral and*

$$\langle \Delta\mathbf{Q}_z \Phi, \Psi\rangle = \sum_i^J \sum_{F(f)}^K a_{F(f)}^i \bar{b}_{F(f)}^i \left\langle \Delta Q_z e^i, e^i \right\rangle \ \textit{(e.f.c).} \tag{7.2}$$

(2) *If, in addition, for each* i

$$\sum_{k,}^{n} \Delta t_k \left\| A_z(s_k)e^i - \left\langle A_z(s_k)e^i, e^i \right\rangle e^i \right\|^2 \leqslant M\mu_n^{\delta-1}, \tag{7.3}$$

where M *is a constant,* μ_n *is the mesh of* P_n*, and* $0 < \delta < 1$*, then the family* $\{\mathcal{A}_z(t) \mid t \in I\}$ *has a strong integral,* $\mathbf{Q}_z[t,a] = \int_a^t \mathcal{A}_z(s)ds$*.*

(3) *The linear operator* $\mathbf{Q}_z[t,a]$ *generates a uniformly continuous* C_0*-contraction semigroup.*

Remark 7.3. In general, the family $\{A_z(t) \mid t \in I\}$ need not have a Bochner or Pettis integral. (However, if it has one, our condition (7.3) is automatically satisfied.)

Proof. To prove (1), note that

$$\langle \Delta\mathbf{Q}_z \Phi, \Psi\rangle = \sum_i \sum_{F(f)} \sum_{F(g)} a_{F(f)}^{\mathrm{i}} \bar{b}_{F(g)}^{\mathrm{i}} \left\langle \Delta\mathbf{Q}_z E_{F(f)}^{\mathrm{i}}, E_{F(g)}^{\mathrm{i}} \right\rangle$$

(we omit the upper limit). Now

$$\left\langle \Delta\mathbf{Q}_z E^{\mathrm{i}}_{F(f)}, E^{\mathrm{i}}_{F(g)} \right\rangle = \sum_{l=1}^{\mathrm{n}} \Delta t_l \prod_{\mathrm{t}\neq\bar{t}_l} \left\langle e^{\mathrm{i}}_{t,f(t)}, e^{\mathrm{i}}_{t,g(t)} \right\rangle \left\langle A_z(\bar{t}_l) e^{\mathrm{i}}_{\bar{t}_l,f(\bar{t}_l)}, e^{\mathrm{i}}_{\bar{t}_l,g(\bar{t}_l)} \right\rangle$$

$$-\sum_{q=1}^{\mathrm{m}} \Delta s_q \prod_{\mathrm{t}\neq\bar{s}_q} \left\langle e^{\mathrm{i}}_{t,f(t)}, e^{\mathrm{i}}_{t,g(t)} \right\rangle \left\langle A_z(\bar{s}_q) e^{\mathrm{i}}_{\bar{s}_q,f(\bar{s}_q)}, e^{\mathrm{i}}_{\bar{s}_q,g(\bar{s}_q)} \right\rangle$$

$$= \sum_{l=1}^{\mathrm{n}} \Delta t_l \left\langle A_z(\bar{t}_l) e^{\mathrm{i}}_{\bar{t}_l,f(\bar{t}_l)}, e^{\mathrm{i}}_{\bar{t}_l,f(\bar{t}_l)} \right\rangle$$

$$-\sum_{q=1}^{\mathrm{m}} \Delta s_q \left\langle A_z(\bar{s}_q) e^{\mathrm{i}}_{\bar{s}_q,f(\bar{s}_q)}, e^{\mathrm{i}}_{\bar{s}_q,f(\bar{s}_q)} \right\rangle = \left\langle \Delta Q_z e^i, e^i \right\rangle \text{ (e.f.f).}$$

This gives (7.2) and shows that the family $\{\mathcal{A}_z(\mathrm{t}) \mid t \in I\}$ has a weak HK-integral if and only if the family $\{A_z(\mathrm{t}) \mid t \in I\}$ has one.

To see that condition (7.3) makes $\mathbf{Q}_z$ a strong limit, let $\Phi \in \mathrm{D}_0$. Then

$$\langle \mathbf{Q}_{z,n}\Phi, \mathbf{Q}_{z,n}\Phi \rangle =$$

$$\sum_{i}^{J} \sum_{F(f),F(g)}^{K} a^i_{F(f)} \bar{a}^i_{F(g)} \left(\sum_{k=1}^{n} \sum_{m=1}^{n} \Delta t_k \Delta t_m \left\langle A_z(s_k) E^i_{F(f)}, A_z(s_m) E^i_{F(g)} \right\rangle \right)$$

$$= \sum_{i}^{J} \sum_{F(f)}^{K} \left| a^i_{F(f)} \right|^2$$

$$\times \left(\sum_{k\neq m}^{n} \Delta t_k \Delta t_m \left\langle A_z(s_k) e^i_{s_k f(s_k)}, e^i_{s_k f(s_k)} \right\rangle \left\langle e^i_{s_m f(s_m)}, A_z(s_m) e^i_{s_m f(s_m)} \right\rangle \right)$$

$$+ \sum_{i}^{J} \sum_{F(f)}^{K} \left| a^i_{F(f)} \right|^2 \left(\sum_{k\neq m}^{n} (\Delta t_k)^2 \left\langle A_z(s_k) e^i_{s_k f(s_k)}, A_z(s_k) e^i_{s_k f(s_k)} \right\rangle \right).$$

This can be rewritten as

$$\begin{aligned} \|\mathbf{Q}_{z,n}\Phi\|^2_{\otimes} &= \sum\nolimits_i^J \sum\nolimits_{F(f)}^K \left| a^i_{F(f)} \right|^2 \left\{ \left| \left\langle Q_{z,n} e^i, e^i \right\rangle \right|^2 \right. \\ &\left. + \sum\nolimits_{k=1}^{n} (\Delta t_k)^2 \left(\left\| A_z(s_k) e^i \right\|^2 - \left| \left\langle A_z(s_k) e^i, e^i \right\rangle \right|^2 \right) \right\} \ (e.f.c). \end{aligned} \tag{7.4}$$

First note that:

$$\left\| A_z(s_k) e^i \right\|^2 - \left| \left\langle A_z(s_k) e^i, e^i \right\rangle \right|^2 = \left\| A_z(s_k) e^i - \left\langle A_z(s_k) e^i, e^i \right\rangle e^i \right\|^2,$$

so that the last term in (7.4) can be written as

$$\begin{aligned}&\sum\nolimits_{k=1}^{n} (\Delta t_k)^2 \left(\left\|A_z(s_k)e^i\right\|^2 - \left|\left\langle A_z(s_k)e^i, e^i\right\rangle\right|^2\right)\\&= \sum\nolimits_{k=1}^{n} (\Delta t_k)^2 \left\|A_z(s_k)e^i - \left\langle A_z(s_k)e^i, e^i\right\rangle e^i\right\|^2 \leqslant\ \mu_n^{\delta} M.\end{aligned}$$

We can now use the above result in (7.4) to get

$$\|\mathbf{Q}_{z,n}\Phi\|_{\otimes}^2 \leqslant \sum\nolimits_{i}^{J} \sum\nolimits_{F(f)}^{K} \left|a_{F(f)}^i\right|^2 \left|\left\langle Q_{z,n}e^i, e^i\right\rangle\right|^2 + \mu_n^{\delta} M \quad (e.f.c).$$

Thus, $\mathbf{Q}_{z,n}[t,a]$ converges strongly to $\mathbf{Q}_z[t,a]$ on $\mathcal{FD}_{\otimes}^2$. To show that $\mathbf{Q}_z[t,a]$ generates an uniformly continuous contraction semigroup, it suffices to show that $Q_{\mathrm{z}}[t,a]$ is dissipative. For any Φ in $\mathcal{FD}_{\otimes}^2$,

$$\langle \mathbf{Q}_z[t,a]\Phi, \Phi\rangle = \sum_{i}^{J} \sum_{F(f)}^{K} \left|a_{F(f)}^i\right|^2 \left\langle Q_z e^i, e^i\right\rangle \ (e.f.c)$$

and, for each n, we have

$$\begin{aligned}&\mathrm{Re}\left\langle Q_z[t,a]e^i, e^i\right\rangle = \mathrm{Re}\left\langle Q_{z,n}[t,a]e^i, e^i\right\rangle + \mathrm{Re}\left\langle \left[Q_z[t,a] - Q_{z,n}[t,a]\right] e^i, e^i\right\rangle\\&\leqslant \mathrm{Re}\left\langle \left[Q_z[t,a] - Q_{z,n}[t,a]\right] e^i, e^i\right\rangle,\end{aligned}$$

since $Q_{z,n}[t,a]$ is dissipative. Letting $n \to \infty$ implies that Re $\left\langle Q_z[t,a]e^i, e^i\right\rangle \leqslant 0$, so that $\mathrm{Re}\,\langle \mathbf{Q}_z[t,a]\Phi, \Phi\rangle \leqslant 0$. Thus, $\mathbf{Q}_z[t,a]$ is a bounded dissipative linear operator on $\mathcal{FD}_{\otimes}^2$, which completes our proof. □

We can also prove Theorem 7.2 for the family $\{\mathcal{A}(\mathrm{t}) \mid t \in I\}$. The same proof goes through, but now we restrict to $\mathrm{D}_0 = \bigotimes_{t\in I} \mathrm{D}(A(t)) \cap \mathcal{FD}_{\otimes}^2$. In this case (7.3) becomes:

$$\sum_{k=1}^{n} \Delta t_k \left\|A(s_k)e^i - \left\langle A(s_k)e^i, e^i\right\rangle e^i\right\|^2 \leqslant M\mu_n^{\delta-1}. \tag{7.5}$$

From Eq. (7.4), we have the following important result: (set $\sum_{F(f)}^{K} \left|a_{F(f)}^{\mathrm{i}}\right|^2 = \left|b^i\right|^2$)

$$\|\mathbf{Q}_z[t,a]\Phi\|_{\otimes}^2 = \sum_{\mathrm{i}}^{J} \left|b^i\right|^2 \left|\left\langle Q_z e^i, e^i\right\rangle\right|^2 \ (e.f.c). \tag{7.6}$$

The representation (7.6) makes it easy to prove the next theorem.

Theorem 7.4. *With the conditions of Theorem 7.2, we have:*

(1) $\mathbf{Q}_z[t,s] + \mathbf{Q}_z[s,a] = \mathbf{Q}_z[t,a]$ *(e.f.c)*,

(2) $s\text{ - }\lim\limits_{h\to 0} \frac{\mathbf{Q}_z[t+h,a]-\mathbf{Q}_z[t,a]}{h} = s\text{ - }\lim\limits_{h\to 0} \frac{\mathbf{Q}_z[t+h,t]}{h} = \mathcal{A}_z(t)$ *(e.f.c)*,

(3) $s\text{ - }\lim\limits_{h\to 0} \mathbf{Q}_z[t+h,t] = 0$ *(a.f.c)*,

(4) $s\text{ - }\lim\limits_{h\to 0} \exp\left\{\tau \mathbf{Q}_z[t+h,t]\right\} = I_\otimes$ *(e.f.c)*,$\tau \geqslant 0$.

Proof. In each case, it suffices to prove the result for $\Phi \in \mathrm{D}_0$. To prove (1), use

$$\left\|\left[\mathbf{Q}_z[t,s] + \mathbf{Q}_z[s,a]\right]\Phi\right\|_\otimes^2 = \sum\nolimits_i^J \left|b^i\right|^2 \left|\left\langle \left[Q_z[t,s] + Q_z[s,a]\right]e^i, e^i\right\rangle\right|^2$$
$$= \sum\nolimits_i^J \left|b^i\right|^2 \left|\left\langle Q_z[t,a]e^i, e^i\right\rangle\right|^2 = \left\|\mathbf{Q}_z[t,a]\Phi\right\|_\otimes^2 \text{ (e.f.c)}.$$

To prove (2), use (1) to get that $\mathbf{Q}_z[t+h,a] - \mathbf{Q}_z[t,a] = \mathbf{Q}_z[t+h,t]$ (e.f.c.), so that

$$\lim_{h\to 0}\left\|\frac{\mathbf{Q}_z[t+h,t]}{h}\Phi\right\|_\otimes^2$$
$$= \sum_{\text{i}}^{J} \left|b^i\right|^2 \lim_{h\to 0}\left|\left\langle \frac{Q_z[t+h,t]}{h}e^i, e^i\right\rangle\right|^2 = \left\|\mathcal{A}_z(t)\Phi\right\|_\otimes^2 \text{ (e.f.c.)}.$$

The proof of (3) follows from (2) and the proof of (4) follows from (3). □

The results of the previous theorem are expected if $\mathbf{Q}_z[t,a]$ is an integral in the conventional sense. The important point is that a weak integral on the base space along with (7.3) gives a strong integral on $\mathcal{FD}_\otimes^2$ (note that, by (2) of the last theorem, we also get strong differentiability). This clearly shows that our constructive approach to time-ordering has more to offer than providing a representation space to allow time to act as a place-keeper for operators in a product. It should be observed that, in all results up to now, we have only used the assumption that the family $A(t), t \in I$ is weakly continuous, generates a contraction semigroup and satisfies Eq. (7.5). In what follows, we shall find it convenient to use the fact that each $A(t)$ generates a C_0-contraction semigroup if, for each t, both $A(t)$ and $A^*(t)$ are dissipative. (This is an easier condition to check in practice.)

Theorem 7.5. *With the above assumptions, we have that* $\lim_{z\to\infty} \langle Q_z[t,a]\phi,\psi\rangle = \langle Q[t,a]\phi,\psi\rangle$ *exists for all* $\phi \in D[Q]$, $\psi \in D[Q^*]$. *Furthermore:*

(1) *the operator* $Q[t,a]$ *generates a* C_0*-contraction semigroup on* $\mathcal{H}$,

(2) *for* $\Phi \in D_0$,

$$\lim_{z\to\infty} \mathbf{Q}_z[t,a]\Phi = \mathbf{Q}[t,a]\Phi,$$

and

(3) *the operator* $\mathbf{Q}[t,a]$ *generates a* C_0*-contraction semigroup on* $\mathcal{FD}^2_{\otimes}$,

(4) $\mathbf{Q}[t,s]\Phi + \mathbf{Q}[s,a]\Phi = \mathbf{Q}[t,a]\Phi$ *(e.f.c.)*,

(5)

$$\lim_{h\to 0} [(\mathbf{Q}[t+h,a] - \mathbf{Q}[t,a])/h]\,\Phi = \lim_{h\to 0} [(\mathbf{Q}[t+h,t])/h]\,\Phi = \mathcal{A}(t)\Phi \text{ (e.f.c.)},$$

(6) $\lim_{h\to 0} \mathbf{Q}[t+h,t]\Phi = 0$ *(e.f.c.), and*

(7) $\lim_{h\to 0} \exp\{\tau\mathbf{Q}[t+h,t]\}\,\Phi = \Phi$ *(e.f.c.)*,$\tau \geqslant 0$.

Proof. Since $A_{\mathrm{z}}(t)$, $A(t)$ are weakly continuous and $A_z(t) \xrightarrow{s} A(t)$ for each $t \in I$, given $\varepsilon > 0$ we can choose Z such that, if $z > Z$, then

$$\sup_{s\in[a,b]} |\langle [A(s) - A_z(s)]\,\varphi\,,\,\psi\rangle| < \varepsilon/3(b-a).$$

By uniform (weak) continuity, if $s, s' \in [a,b]$ we can also choose η such that, if $|s - s'| < \eta$,

$$\sup_{z>0} \left|\left\langle \left[A_z(s) - A_z(s')\right] \varphi\,,\,\psi\right\rangle\right| < \varepsilon/3(b-a)$$

and

$$\left|\left\langle \left[A(s) - A(s')\right] \varphi\,,\,\psi\right\rangle\right| < \varepsilon/3(b-a).$$

Now choose $\delta(t) : [a,b] \to (0,\infty)$ so that, for any HK-partition $\mathbf{P}$ for δ, we have that $\mu_n < \eta$, where μ_n is the mesh of the partition. If $Q_{z,n} = \sum_{j=1}^{n} A_z(\tau_j)\Delta t_j$ and $Q_n = \sum_{j=1}^{n} A(\tau_j)\Delta t_j$, we have

$$
\begin{aligned}
&\left|\left\langle \left[Q_z[t,a]-Q[t,a]\right]\varphi\,,\,\psi\right\rangle\right| \leqslant \left|\left\langle \left[Q_n[t,a]-Q[t,a]\right]\varphi\,,\,\psi\right\rangle\right| \\
&\quad+\left|\left\langle \left[Q_{z,n}[t,a]-Q_z[t,a]\right]\varphi\,,\,\psi\right\rangle\right| + \left|\left\langle \left[Q_n[t,a]-Q_{z,n}[t,a]\right]\varphi\,,\,\psi\right\rangle\right| \\
&\quad\leqslant \sum\nolimits_{j=1}^{n}\int_{t_{j-1}}^{t_j}\left|\left\langle \left[A(\tau_j)-A(\tau)\right]\varphi\,,\,\psi\right\rangle\right| d\tau \\
&\quad+\sum\nolimits_{j=1}^{n}\int_{t_{j-1}}^{t_j}\left|\left\langle \left[A_z(\tau_j)-A_z(\tau)\right]\varphi\,,\,\psi\right\rangle\right| d\tau \\
&\quad+\sum\nolimits_{j=1}^{n}\int_{t_{j-1}}^{t_j}\left|\left\langle \left[A(\tau_j)-A_z(\tau_j)\right]\varphi\,,\,\psi\right\rangle\right| d\tau < \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.
\end{aligned}
$$

This proves that $\lim\limits_{z\to\infty}\langle Q_z[t,a]\phi,\psi\rangle = \langle Q[t,a]\phi,\psi\rangle$. To prove (1), first note that $Q[t,a]$ is closable and use

$$
\begin{aligned}
&\operatorname{Re}\langle Q[t,a]\phi,\phi\rangle = \operatorname{Re}\langle Q_z[t,a]\phi,\phi\rangle + \operatorname{Re}\langle\left[Q[t,a]-Q_z[t,a]\right]\phi,\phi\rangle \\
&\quad\leqslant \operatorname{Re}\langle\left[Q[t,a]-Q_z[t,a]\right]\phi,\phi\rangle\,,
\end{aligned}
$$

and let z $\to\infty$ to show that $Q[t,a]$ is dissipative. Then do likewise for $\langle\phi, Q^*[t,a]\phi\rangle$ to show that the same is true for $Q^*[t,a]$ to complete the proof. It is important to note that although $Q[t,a]$ generates a contraction semigroup on $\mathcal{H}$, $exp\{Q[t,a]\}$ does not solve the original initial-value problem [see Eq. (7.7) below].

To prove (2), use (7.6) in the form

$$
\left\|\left[\mathbf{Q}_z[t,a]-\mathbf{Q}_{z'}[t,a]\right]\Phi\right\|_\otimes^2 = \sum_{\mathrm{i}}^{J}\left|b^i\right|^2\left|\left\langle\left[Q_z[t,a]-Q_{z'}[t,a]\right]e^i,e^i\right\rangle\right|^2 . \quad (7.7)
$$

This proves that $\mathbf{Q}_z[t,a]\xrightarrow{\mathrm{s}}\mathbf{Q}[t,a]$. Since $\mathbf{Q}[t,a]$ is densely defined, it is closable. The same method as above shows that it is m-dissipative. Proofs of the other results follow directly from the same results of Theorem 7.4. □

In closing this section, we note that all the results go through for $\mathcal{FD}_\otimes^\alpha$. The clearest way to see this is to extend the family $\mathcal{A}(t)$, $t\in I$, to $\mathcal{FD}_\otimes^2$, use the theorems of this section, and restrict back to $\mathcal{FD}_\otimes^\alpha$.

7.1.2. General Case. We now relax the contraction condition and assume that $A(t), t\in I$ generates a C_0-semigroup on $\mathcal{H}$. We can always shift the spectrum (if necessary) so that $\|\exp\{\tau A(t)\}\|\leqslant M(t)$. We assume that $\sup_J\prod_{i\in J}\|\exp\{\tau A(t_i)\}\|\leqslant M$, where the sup is over all finite subsets $J\subset I$. We remark that if we renorm $\mathcal{H}$ with an equivalent norm; for each t, we can reduce the general case to contractions. (However, the effort does not appear to merit the additional work.)

Theorem 7.6. *Suppose that* $A(t), t \in I$*, generates a* C_0*-semigroup, satisfies (7.3) and has a weak HK-integral,* $Q[t,a]$*, on a dense set* D *in* $\mathcal{H}$*. Then the family* $\mathcal{A}(t), t \in I$ *has a strong HK-integral,* $\mathbf{Q}[t,a]$*, which generates a* C_0*-semigroup on* $\mathcal{FD}_\otimes^2$ *(for each* $t \in I$*) and* $\|\exp\{\mathbf{Q}[t,a]\}\|_\otimes \leqslant M$*.*

Proof. It is clear from part (2) of Theorem 7.5 that $\mathbf{Q}_n[t,a] = \sum_{i=1}^n \mathcal{A}(\tau_i)\Delta t_i$ generates a C_0-semigroup on $\mathcal{FD}_\otimes^2$ and $\|\exp\{\mathbf{Q}_n[t,a]\}\|_\otimes \leqslant M$. If $\Phi \in D_0$, let $\mathbf{P}_m$, $\mathbf{P}_n$ be arbitrary HK-partitions for δ_m, δ_n (of order m and n respectively) and set $\delta(s) = \delta_m(s) \wedge \delta_n(s)$. Since any HK-partition for δ is one for δ_m and δ_n, we have that

$$\begin{aligned}
&\|[\exp\{\tau\mathbf{Q}_n[t,a]\} - \exp\{\tau\mathbf{Q}_m[t,a]\}]\,\Phi\|_\otimes \\
&= \left\|\int_0^\tau \frac{d}{ds}\left[\exp\{(\tau-s)\mathbf{Q}_n[t,a]\}\exp\{s\mathbf{Q}_m[t,a]\}\right]\Phi ds\right\|_\otimes \\
&\leqslant \int_0^\tau \|[\exp\{(\tau-s)\mathbf{Q}_n[t,a]\}\,(\mathbf{Q}_n[t,a]-\mathbf{Q}_m[t,a])\exp\{s\mathbf{Q}_m[t,a]\}\Phi]\|_\otimes \\
&\leqslant M\int_0^\tau \|(\mathbf{Q}_n[t,a]-\mathbf{Q}_m[t,a])\,\Phi\|_\otimes ds \\
&\leqslant M\tau\,\|[\mathbf{Q}_n[t,a]-\mathbf{Q}[t,a]]\,\Phi\|_\otimes + M\tau\,\|[\mathbf{Q}[t,a]-\mathbf{Q}_m[t,a]]\,\Phi\|_\otimes .
\end{aligned}$$

The existence of the weak HK-integral, $Q[t,a]$, on $\mathcal{H}$ satisfying Eq. (7.3) implies that $\mathbf{Q}_n[t,a] \overset{s}{\to} \mathbf{Q}[t,a]$, so that $\exp\{\tau\mathbf{Q}_n[t,a]\}\Phi$ converges as $n \to \infty$ for each fixed $t \in I$; and the convergence is uniform on bounded τ intervals. As $\|\exp\{\mathbf{Q}_n[t,a]\}\|_\otimes \leqslant M$, we have

$$\lim_{n\to\infty} \exp\{\tau\mathbf{Q}_n[t,a]\}\Phi = \mathbf{S}_t(\tau)\Phi,\ \ \Phi \in \mathcal{FD}_\otimes^2.$$

The limit is again uniform on bounded τ intervals. It is easy to see that the limit $\mathbf{S}_t(\tau)$ satisfies the semigroup property, $\mathbf{S}_t(0) = I$, and $\|\mathbf{S}_t(\tau)\|_\otimes \leqslant M$. Furthermore, as the uniform limit of continuous functions, we see that $\tau \to \mathbf{S}_t(\tau)\Phi$ is continuous for $\tau \geqslant 0$. We are done if we show that $\mathbf{Q}[t,a]$ is the generator of $\mathbf{S}_t(\tau)$. For $\Phi \in D_0$, we have that

$$\begin{aligned}
\mathbf{S}_t(\tau)\Phi - \Phi &= \lim_{n\to\infty} \exp\{\tau\mathbf{Q}_n[t,a]\}\Phi - \Phi \\
&= \lim_{n\to\infty} \int_0^\tau \exp\{s\mathbf{Q}_n[t,a]\}\mathbf{Q}_n[t,a]\Phi ds = \int_0^\tau \mathbf{S}_t(\tau)\mathbf{Q}[t,a]\Phi ds.
\end{aligned}$$

Our result follows from the uniqueness of the generator, so that $\mathbf{S}_t(\tau) = \exp\{\tau\mathbf{Q}[t,a]\}$. $\square$

The next result is the time-ordered version of the Hille–Yosida Theorem. We assume that the family $A(t), t \in I$ is closed and densely defined.

Theorem 7.7. *The family $\mathcal{A}(t), t \in I$ has a strong HK-integral, $\mathbf{Q}[t,a]$, which generates a C_0-contraction semigroup on $\mathcal{FD}_{\otimes}^2$ if and only if $\rho(A(t)) \supset (0,\infty)$, $\|R(\lambda : A(t))\| < 1/\lambda$ for $\lambda > 0$; $A(t), t \in I$, satisfies (7.3) and has a densely defined weak HK-integral $Q[t,a]$ on $\mathcal{H}$.*

Proof. In the first direction, suppose $\mathbf{Q}[t,a]$ generates a C_0-contraction semigroup on $\mathcal{FD}_{\otimes}^2$. Then $\mathbf{Q}_n[t,a]\Phi \xrightarrow{s} \mathbf{Q}[t,a]\Phi$ for each $\Phi \in \mathrm{D}_0$ and each $t \in I$. Since $\mathbf{Q}[t,a]$ has a densely defined strong HK-integral, it follows from (7.5) that $Q[t,a]$ must have a densely defined weak HK-integral. Since $\mathbf{Q}_n[t,a]$ generates a C_0-contraction semigroup for each HK-partition of order n, it follows that $\mathcal{A}(t)$ must generate a C_0-contraction semigroup for each $t \in I$. From Eq. (7.1) and the discussion that follows, we see that $A(t)$ must also generate a C_0-contraction semigroup for each $t \in I$. From the conventional Hille–Yosida theorem, the resolvent condition follows.

In the reverse direction, the conventional Hille–Yosida theorem along with the first part of Theorem 7.5 shows that $Q[t,a]$ generates a C_0-contraction semigroup for each $t \in I$. From Theorem 7.6, we see that for HK-partition of order n, $\mathbf{Q}_n[t,a]$ generates a C_0-contraction semigroup. Furthermore, $\mathbf{Q}_{\mathrm{n}}[t,a]\Phi \to \mathbf{Q}[t,a]\Phi$ for each $\Phi \in \mathrm{D}_0$ and $\mathbf{Q}[t,a]$ generates a C_0-contraction semigroup on $\mathcal{FD}_{\otimes}^2$. □

The other generation theorems have a corresponding formulation in terms of time-ordered integrals.

7.2. Time-Ordered Evolutions

As $\mathbf{Q}[t,a]$ and $\mathbf{Q}_{\mathrm{z}}[t,a]$ generate (uniformly bounded) C_0-semigroups, we can set $\mathbf{U}[t,a] = exp\{\mathbf{Q}[t,a]\}$, $\mathbf{U}_z[t,a] = exp\{\mathbf{Q}_z[t,a]\}$. They are C_0-evolution operators and the following theorem generalizes a result due to Hille and Phillips [HP].

Theorem 7.8. *For each n, and $\Phi \in D\left[(\mathbf{Q}[t,a])^{n+1}\right]$, we have: ($w$ is positive and $\mathbf{U}^w[t,a] = exp\,\{w\mathbf{Q}[t,a]\}$)*

$$\mathbf{U}^w[t,a]\Phi = \left\{ I_\otimes + \sum_{k=1}^{n} \frac{(w\mathbf{Q}[t,a])^k}{k!} + \frac{1}{n!}\int\limits_0^w (w-\xi)^n \mathbf{Q}[t,a]^{n+1}\mathbf{U}^\xi[t,a]d\xi \right\}\Phi.$$

Proof. The proof is easy. Start with

$$[\mathbf{U}_{\mathrm{z}}^w[t,a]\Phi - I_\otimes]\,\Phi = \int\limits_0^w \mathbf{Q}_{\mathrm{z}}[t,a]\mathbf{U}_{\mathrm{z}}^\xi[t,a]d\xi\Phi$$

and use integration by parts to get that

$$[\mathbf{U}_z^w[t,a]\Phi - I_\otimes]\,\Phi = w\mathbf{Q}_z[t,a]\Phi + \int\limits_0^w (w-\xi)\left[\mathbf{Q}_{\mathrm{z}}[t,a]\right]^2 \mathbf{U}_z^\xi[t,a]d\xi\Phi.$$

It is clear how to get the nth term. Finally, let $z \to \infty$ to get the result. □

Theorem 7.9. *If* $a < t < b$,

(1) $\lim\limits_{z\to\infty} \mathbf{U}_z[t,a]\Phi = \mathbf{U}[t,a]\Phi,\ \Phi \in \mathcal{FD}_\otimes^2.$

(2)

$$\frac{\partial}{\partial t}\mathbf{U}_z[t,a]\Phi = \mathcal{A}_z(t)\mathbf{U}_z[t,a]\Phi = \mathbf{U}_z[t,a]\mathcal{A}_z(t)\Phi,$$

with $\Phi \in \mathcal{FD}_\otimes^2$, *and*

(3)

$$\frac{\partial}{\partial t}\mathbf{U}[t,a]\Phi = \mathcal{A}(t)\mathbf{U}[t,a]\Phi = \mathbf{U}[t,a]\mathcal{A}(t)\Phi,\ \Phi \in D(\mathbf{Q}[b,a]) \supset D_0.$$

Proof. To prove (1), use the fact that $\mathcal{A}_z(t)$ and $\mathcal{A}(t)$ commute, along with

$$\mathbf{U}[t,a]\Phi - \mathbf{U}_z[t,a]\Phi = \int_0^1 (d/ds)\left(e^{s\mathbf{Q}[t,a]}e^{(1-s)\mathbf{Q}_z[t,a]}\right)\Phi ds$$

$$= \int_0^1 s\left(e^{s\mathbf{Q}[t,a]}e^{(1-s)\mathbf{Q}_{\mathrm{z}}[t,a]}\right)(\mathbf{Q}[t,a] - \mathbf{Q}_z[t,a])\,\Phi ds,$$

so that

$$\lim_{\mathrm{z}\to 0}\|\mathbf{U}[t,a]\Phi - \mathbf{U}_{\mathrm{z}}[t,a]\Phi\| \leqslant M \lim_{\mathrm{z}\to 0}\|\mathbf{Q}[t,a]\Phi - \mathbf{Q}_{\mathrm{z}}[t,a]\Phi\| = 0.$$

To prove (2), use

$$\mathbf{U}_z[t+h,a] - \mathbf{U}_z[t,a] = \mathbf{U}_z[t,a]\left(\mathbf{U}_z[t+h,t] - \mathrm{I}\right) = \left(\mathbf{U}_z[t+h,t] - \mathrm{I}\right)\mathbf{U}_z[t,a],$$

so that

$$(\mathbf{U}_z[t+h,a]-\mathbf{U}_z[t,a])/h=\mathbf{U}_z[t,a]\left[(\mathbf{U}_z[t+h,t]-\mathrm{I})/h\right].$$

Now set $\Phi_z^t=\mathbf{U}_z[t,a]\Phi$ and use Theorem 7.8 with $n=1$ and $w=1$ to get:

$$\mathbf{U}_z[t+h,t]\Phi_z^t=\left\{I_{\otimes}+\mathbf{Q}_z[t+h,t]+\int_0^1(1-\xi)\mathbf{U}_z^{\xi}[t+h,t]\mathbf{Q}_z[t+h,t]^2d\xi\right\}\Phi_z^t,$$

so

$$\begin{aligned}\frac{(\mathbf{U}_z[t+h,t]-\mathrm{I})}{h}\Phi_z^t-\mathcal{A}_z(t)\Phi_z^t&=\frac{\mathbf{Q}_z[t+h,t]}{h}\Phi_z^t-\mathcal{A}_z(t)\Phi_z^t\\&\quad+\int_0^1(1-\xi)\mathbf{U}_z^{\xi}[t+h,t]\frac{\mathbf{Q}_z[t+h,t]^2}{h}\Phi_z^t d\xi.\end{aligned}$$

It follows that

$$\left\|\frac{(\mathbf{U}_z[t+h,t]-\mathrm{I})}{h}\Phi_z^t-\mathcal{A}_z(t)\Phi_z^t\right\|_{\otimes}\leqslant\left\|\frac{\mathbf{Q}_z[t+h,t]}{h}\Phi_z^t-\mathcal{A}_z(t)\Phi_z^t\right\|_{\otimes}$$
$$+\frac{1}{2}\left\|\frac{\mathbf{Q}_z[t+h,t]^2}{h}\Phi_z^t\right\|_{\otimes}.$$

The result now follows from Theorem 7.4, (2) and (3). To prove (3), note that $\mathcal{A}_z(t)\Phi=\mathcal{A}(t)\left\{z\mathbf{R}(z,\mathcal{A}(t))\right\}\Phi=\left\{z\mathbf{R}(z,\mathcal{A}(t))\right\}\mathcal{A}(t)\Phi$, so that $\{z\mathbf{R}(z,\mathcal{A}(t))\}$ commutes with $\mathbf{U}[t,a]$ and $\mathcal{A}(t)$. It is now easy to show that

$$\begin{aligned}&\|\mathcal{A}_z(t)\mathbf{U}_z[t,a]\Phi-\mathcal{A}_{z'}(t)\mathbf{U}_{z'}[t,a]\Phi\|\\&\quad\leqslant\|\mathbf{U}_z[t,a]\left(\mathcal{A}_z(t)-\mathcal{A}_{z'}(t)\right)\Phi\|+\left\|z'\mathbf{R}(z',\mathcal{A}(t))\left[\mathbf{U}_z[t,a]\Phi-\mathbf{U}_{z'}[t,a]\right]\mathcal{A}(t)\Phi\right\|\\&\quad\leqslant M\left\|\left(\mathcal{A}_z(t)-\mathcal{A}_{z'}(t)\right)\Phi\right\|+M\left\|\left[\mathbf{U}_z[t,a]\Phi-\mathbf{U}_{z'}[t,a]\right]\mathcal{A}(t)\Phi\right\|\to0,\ z,z'\to\infty,\end{aligned}$$

so that, for $\Phi\in\mathrm{D}(\mathbf{Q}[b,a])$,

$$\mathcal{A}_z(t)\mathbf{U}_z[t,a]\Phi\to\mathcal{A}(t)\mathbf{U}[t,a]\Phi=\frac{\partial}{\partial t}\mathbf{U}[t,a]\Phi.$$

□

Since, as noted earlier, $exp\{Q[t,a]\}$ does not solve the initial-value problem, we restate the last part of the last theorem to emphasize the importance of this result, and the power of the constructive Feynman theory.

Theorem 7.10. *If $a < t < b$,*

$$\frac{\partial}{\partial t}\mathbf{U}[t,a]\Phi = \mathcal{A}(t)\mathbf{U}[t,a]\Phi = \mathbf{U}[t,a]\mathcal{A}(t)\Phi,\ \Phi \in D_0 \subset D(\mathbf{Q}[b,a]).$$

7.3. Perturbation Theory

In this section, we prove a few results without attempting to be exhaustive. Because of Theorem 5.41, the general problem of perturbation theory can always be reduced to that of the strong limit of the bounded case.

Assume that, for each $t \in I$, $A_0(t)$ is the generator of a C_0-semigroup on $\mathcal{H}$ and that $A_1(t)$ is closed and densely defined. The (generalized) sum of $A_0(t)$ and $A_1(t)$, in its various forms, whenever it is defined (with dense domain), is denoted by $A(t) = A_0(t) \oplus A_1(t)$ (see Kato [KA], and Pazy [PZ]). Let $A_1^n(t) = nA_1(t)R(n, T_1(t))$ be the (generalized) Yosida approximator for $A_1(t)$, where $T_1(t) = -\left[A_1^*(t)A_1(t)\right]^{1/2}$ and set $A_n(t) = A_0(t) + A_1^n(t)$. The first result follows from Theorem 5.34.

Theorem 7.11. *For each n, $A_0(t)+A_1^n(t)$ (respectively $\mathcal{A}_0(t)+\mathcal{A}_1^n(t)$) is the generator of a C_0-semigroup on $\mathcal{H}$ (respectively $\mathcal{FD}_{\otimes}^2$) and:*

(1) *If, for each $t \in I$, $A_0(t)$ generates an analytic or contraction C_0-semigroup, then so does $A_n(t)$ and $\mathcal{A}_n(t)$.*
(2) *If, for each $t \in I$, $A(t) = A_0(t) + A_1(t)$ generates an analytic or contraction C_0-semigroup, then so does $\mathcal{A}(t) = \mathcal{A}_0(t) + \mathcal{A}_1(t)$ and $\exp\{\tau\mathcal{A}_n(t)\} \to \exp\{\tau\mathcal{A}(t)\}$ for $\tau \geqslant 0$.*

We now assume that $A_0(t)$ and $A_1(t)$ are weakly continuous generators of C_0-semigroups for each $t \in I$, and that Eq. (7.3) is satisfied. Then, with the same notation, we have:

Theorem 7.12. *If, for each $t \in I$, $A(t) = A_0(t) \oplus A_1(t)$ generates an analytic or contraction semigroup, then $\mathbf{Q}[t,a]$ generates an analytic or contraction semigroup and $\exp\{\mathbf{Q}_n[t,a]\} \to \exp\{\mathbf{Q}[t,a]\}$.*

Theorem 7.13. *Suppose that $A_0(t)$ and $A_1(t)$ are weakly continuous generators of C_0-contraction semigroups for each $t \in I$ with common dense domains, satisfying Eq. (7.3). If $\mathbf{Q}_0[t,a]$ and $\mathbf{Q}_1[t,a]$ are the corresponding time-ordered generators of contraction semigroups, then*

$$\mathbf{Q}[t,a] = \mathbf{Q}_0[t,a] \oplus \mathbf{Q}_1[t,a]\ (a,s),$$

is the generator of a contraction semigroup on $\mathcal{FD}_{\otimes}^2$.

Proof. Let $\mathbf{Q}_{n,1}[t,a]$ be the Yosida approximator for $\mathbf{Q}_1[t,a]$. It follows that,

$$\mathbf{Q}_n[t,a] = \mathbf{Q}_0[t,a] + \mathbf{Q}_{n,1}[t,a]$$

is the generator of a C_0-contraction semigroup for each n. Furthermore, for any $m,\ n \in \mathbb{N}$ and $\Phi \in D_0$,

$$\begin{aligned}
&\left\| [\exp\{\tau\mathbf{Q}_n[t,a]\} - \exp\{\tau\mathbf{Q}_m[t,a]\}]\,\Phi \right\|_{\otimes} \\
&= \left\| \int_0^\tau \frac{d}{ds}\left[\exp\{(\tau-s)\mathbf{Q}_n[t,a]\}\exp\{s\mathbf{Q}_m[t,a]\}\right]\Phi ds \right\|_{\otimes} \\
&\leq \int_0^\tau \left\| [\exp\{(\tau-s)\mathbf{Q}_n[t,a]\}\exp\{s\mathbf{Q}_m[t,a]\}\left(\mathbf{Q}_n[t,a]-\mathbf{Q}_m[t,a]\right)\Phi] \right\|_{\otimes} \\
&\leq \int_0^\tau \left\| \left(\mathbf{Q}_n[t,a]-\mathbf{Q}_m[t,a]\right)\Phi \right\|_{\otimes} ds \longrightarrow 0,\ n \to \infty.
\end{aligned}$$

Thus, $\exp\{\tau\mathbf{Q}_n[t,a]\}\Phi$ converges as $n \to \infty$ for each fixed $t \in I$; and the convergence is uniform on bounded τ intervals. As $\|\exp\{\mathbf{Q}_n[t,a]\}\|_{\otimes} \leqslant 1$, we have

$$\lim_{n\to\infty} \exp\{\tau\mathbf{Q}_n[t,a]\}\Phi = \mathbf{S}_t(\tau)\Phi,\ \Phi \in \mathcal{FD}_{\otimes}^2.$$

The limit is again uniform on bounded τ intervals. It is easy to see that the limit $\mathbf{S}_t(\tau)$ satisfies the semigroup property, $\mathbf{S}_t(0) = I$, and $\|\mathbf{S}_t(\tau)\|_{\otimes} \leqslant 1$, so that $S_t(\tau)$ is a C_0-contraction semigroup. Furthermore, as the uniform limit of continuous functions, we see that $\tau \to \mathbf{S}_t(\tau)\Phi$ is continuous for $\tau \geqslant 0$. We are done if we show that $\mathbf{Q}[t,a]$ is the generator of $\mathbf{S}_t(\tau)$. For $\Phi \in D_0$, we have that

$$\begin{aligned}
\mathbf{S}_t(\tau)\Phi - \Phi &= \lim_{n\to\infty} \exp\{\tau\mathbf{Q}_n[t,a]\}\Phi - \Phi \\
&= \lim_{n\to\infty} \int_0^\tau \exp\{s\mathbf{Q}_n[t,a]\}\mathbf{Q}_n[t,a]\Phi ds = \int_0^\tau \mathbf{S}_t(\tau)\mathbf{Q}[t,a]\Phi ds,\ (a.s).
\end{aligned}$$

Our result now follows from the uniqueness of the generator, so that $\mathbf{Q}[t,a]$ generates a C_0-contraction semigroup. □

7.4. Interaction Representation

Within the framework of axiomatic field theory, an important theorem of Haag shows that the interaction representation in sharp time does not exist in a rigorous sense (see [HA]). Haag's theorem shows that the equal time commutation relations for the canonical variables of an interacting field are equivalent to those of a free field. Streater and Wightman point out that "...What is even more likely in physically interesting quantum field theories is that equal-time commutation relations will make no sense at all; the field might not be an operator unless smeared in time as well as space."

In this section, we first show that, if one assumes (as Haag did) that operators act in sharp time, then the interaction representation does not exist (see [SW], p. 168). Recent experiments show that there is quantum interference in time for the wave function of a particle (see Horwitz [HW] and references therein). We also show that if there is time smearing, the interaction representation is well defined.

Let us assume that $A_0(t)$ and $A_1(t)$ are weakly continuous generators of C_0-unitary groups for each $t \in I$, $A(t) = A_0(t) \oplus A_1(t)$ is densely defined and Eq. (7.3) is satisfied. Define $\mathbf{U}_n[t,a]$, $\mathbf{U}_0[t,a]$ and $\bar{\mathbf{U}}_0[t,a]$ by:

$$\mathbf{U}_n[t,a] = \exp\{(-i/\hbar)\int_a^t [\mathcal{A}_0(s) + \mathcal{A}_1^n(s)]ds\},$$

$$\mathbf{U}_0[t,a] = \exp\{(-i/\hbar)\int_a^t \mathcal{A}_0(s)ds\},$$

$$\bar{\mathbf{U}}_0[t,a] = \exp\{(-i/\hbar)\int_a^t \mathbf{E}[t,s]\mathcal{A}_0(s)ds\},$$

where $\mathbf{E}[t,s]$ is our exchange operator.

In the first case, using $\mathbf{U}_0[t,a]$, the interaction representation for $\mathcal{A}_1^n(t)$ is given by:

$$\mathcal{A}_{\mathbf{I}}^n(t) = \mathbf{U}_0[a,t]\mathcal{A}_1^n(t)\mathbf{U}_0[t,a] = \mathcal{A}_1^n(t)(a.s.)$$

as $\mathcal{A}_1^n(t)$ commutes with $\mathbf{U}_0[a,t]$ in sharp time. Thus, the interaction representation does not exist. In the last case, we have

$$\mathcal{A}_{\mathbf{I}}^n(t) = \bar{\mathbf{U}}_0[a,t]\mathcal{A}_1^n(t)\bar{\mathbf{U}}_0[t,a],$$

and the terms do not commute. If we set $\Psi_n(t) = \bar{\mathbf{U}}_0[a,t]\mathbf{U}_n[t,a]\Phi$, we have

$$\frac{\partial}{\partial t}\Psi_n(t) = \frac{i}{\hbar}\bar{\mathbf{U}}_0[a,t]\mathcal{A}_0(t)\mathbf{U}_n[t,a]\Phi - \frac{i}{\hbar}\bar{\mathbf{U}}_0[a,t]\left[\mathcal{A}_0(t) + \mathcal{A}_1^n(t)\right]\mathbf{U}_n[t,a]\Phi$$

so that $$\frac{\partial}{\partial t}\Psi_n(t) = -\frac{i}{\hbar}\{\bar{\mathbf{U}}_0[a,t]\mathcal{A}_1^n(t)\bar{\mathbf{U}}_0[t,a]\}\bar{\mathbf{U}}_0[a,t]\mathbf{U}_n[t,a]\Phi$$

and $$i\hbar\frac{\partial}{\partial t}\Psi_n(t) = \mathcal{A}_{\mathbf{I}}^n(t)\Psi_n(t),\ \Psi_n(a) = \Phi.$$

With the stated conditions, we have

Theorem 7.14. *If $Q_1[t,a] = \int_a^t A_1(s)ds$ generates a C_0-unitary group on $\mathcal{H}$, then the time-ordered integral $\mathbf{Q_I}[t,a] = \int_a^t \mathcal{A}_\mathbf{I}(s)ds$, where $\mathcal{A}_\mathbf{I}(t) = \bar{\mathbf{U}}_0[a,t]\mathcal{A}_1(t)\bar{\mathbf{U}}_0[t,a]$ generates a C_0 unitary group on $\mathcal{FD}_\otimes^2$, and*

$$\exp\{(-i/\hbar)\mathbf{Q}_\mathbf{I}^n[t,a]\} \to \exp\{(-i/\hbar)\mathbf{Q_I}[t,a]\},$$

where $\mathbf{Q}_\mathbf{I}^n[t,a] = \int_a^t \mathcal{A}_\mathbf{I}^n(s)ds$, and:

$$i\hbar\frac{\partial}{\partial t}\Psi(t) = \mathcal{A}_\mathbf{I}(t)\Psi(t),\ \Psi(a) = \Phi.$$

Proof. The result follows from an application of Theorem 7.10. □

7.5. Disentanglement

In this section, we relate our results to the conventional approach, where the order of operators is determined by their position on paper. This section is the method of disentanglement as suggested by Feynman and Fujiwara to relate his theory to the standard theory. As an application, we extend the Trotter–Kato Theorem.

Since any closed densely defined generator of a C_0-semigroup may be replaced by its Yosida approximator, we can restrict our study to bounded linear operators. We first need to establish some notation. If $\{A(t),\ t \in I\}$ denotes an arbitrary family of operators in $L[\mathcal{H}]$, the operator $\prod_{t\in I} A(t)$, when defined, is understood in its natural order: $\prod_{b\geqslant t\geqslant a} A(t)$. Let $L[\mathcal{FD}_\otimes^2] \subset L^\#[\mathcal{H}_\otimes^2]$ be the class of bounded linear operators on $\mathcal{FD}_\otimes^2$. It is easy to see that every operator $\mathcal{A} \in L[\mathcal{FD}_\otimes^2]$, which depends on a countable number of elements in I, may be written as:

$$\mathcal{A} = \sum\nolimits_{i=1}^{\infty} a_i \prod_{k=1}^{n_i} \mathcal{A}_i(t_k),$$

where

$$\mathcal{A}_i(t_k) \in L[\mathcal{H}(t_k)],\ k = 1,\ 2, \cdots,\ n_i,\ n_i \in \mathbb{N}.$$

Definition 7.15. The disentanglement morphism, $dT[\,\cdot\,]$, is a mapping from $L[\mathcal{FD}_\otimes^2]$ to $L[\mathcal{H}]$, such that:

$$dT\left[\mathcal{A}\right] = dT\left[\sum\nolimits_{i=1}^{\infty} a_i \prod_{k=1}^{n_i} \mathcal{A}_i(t_k)\right] = \sum\nolimits_{i=1}^{\infty} a_i \prod_{n_i\geqslant k\geqslant 1} A_i(t_k).$$

Theorem 7.16. *The map $dT[\,\cdot\,]$ is a well-defined surjective bounded linear mapping from $L[\mathcal{FD}^2_{\otimes}]$ to $L[\mathcal{H}]$, which is not injective, but $dT[\,\cdot\,]|_{L[\mathcal{H}(t)]} = \mathbf{T}_\theta^{-t}$, where $\mathbf{T}_\theta^t \circ \mathbf{T}_\theta^{-t} = \mathbf{T}_\theta^{-t} \circ \mathbf{T}_\theta^t = \mathbf{I}$.*

Proof. With the stated convention, it is easy to see that $dT[\,\cdot\,]$ is a well-defined bounded, surjective linear mapping. To see that it is not injective, note that $dT[E[t,s]\mathcal{A}(s)] = dT[\mathcal{A}(s)]$, while $E[t,s]\mathcal{A}(s) \in L[\mathcal{H}(t)]$ and $\mathcal{A}(s) \in L[\mathcal{H}(s)]$, so that these operators are not equal when $t \neq s$. To see that $dT[\,\cdot\,]|_{L[\mathcal{H}(t)]} = \mathbf{T}_\theta^{-t}$, we need only show that $dT[\,\cdot\,]$ is injective when restricted to $L[\mathcal{H}(t)]$. If $\mathcal{A}(t)$, $\mathcal{B}(t) \in L[\mathcal{H}(t)]$ and $dT[\mathcal{A}(t)] = dT[\mathcal{B}(t)]$, then $A(t) = B(t)$, by definition of $dT[\,\cdot\,]$, so that $\mathcal{A}(t) = \mathcal{B}(t)$ by definition of $L[\mathcal{H}(t)]$. □

Definition 7.17. A Fujiwara–Feynman algebra ($\mathcal{FF}$-algebra) over $L[\mathcal{H}]$, for the parameter set I, is the quadruple $\left(\left\{\mathbf{T}_\theta^t,\ t \in I\right\},\ L[\mathcal{H}],\ dT[\,\cdot\,],\ L[\mathcal{FD}^2_{\otimes}]\right)$.

We now show that the $\mathcal{FF}$-algebra is universal for time-ordering in the following sense.

Theorem 7.18. *Let $\{A(t)\,|\,t \in I\} \in L[\mathcal{H}]$ be any family of operators. Then the following conditions hold:*

(1) *The time-ordered operator $\mathcal{A}(t) \in L[\mathcal{H}(t)]$ and $dT[\mathcal{A}(t)] = A(t)$, $t \in I$.*

(2) *For any family $\{t_j \,|\, 1 \leqslant j \leqslant n,\ n \in \mathbb{N}\}$, $t_j \in I$ (distinct), the map $\mathop{\times}\limits_{n=1}^{\infty} (A(t_n),\ A(t_{n-1}),\ \cdots,\ A(t_1)) \to \sum_{n=1}^{\infty} a_n \prod\limits_{n \geqslant j \geqslant 1} A(t_j)$ from $\mathop{\times}\limits_{n=1}^{\infty} \left\{ \mathop{\times}\limits_{j=1}^{n} L[\mathcal{H}] \right\} \to L[\mathcal{H}]$ has a unique factorization through $L[\mathcal{FD}^2_{\otimes}]$, so that $\sum_{n=1}^{\infty} a_n \prod\limits_{n \geqslant j \geqslant 1} A(t_j) \in L[\mathcal{H}]$ corresponds to $\sum_{n=1}^{\infty} a_n \prod\limits_{j=1}^{n} \mathcal{A}(t_j)$.*

Proof. $\mathcal{A}(t) = \mathbf{T}_\theta^t[A(t)]$ and $dT[\mathcal{A}(t)] = A(t)$ gives (1).

To prove (2), note that

$$\Theta : \mathop{\times}_{n=1}^{\infty} \left\{ \mathop{\times}_{j=1}^{n} L[\mathcal{H}] \right\} \to \mathop{\times}_{n=1}^{\infty} \left\{ \mathop{\times}_{j=1}^{n} L[\mathcal{H}(t_j)] \right\},$$

defined by

$$\Theta\left[\mathop{\times}_{n=1}^{\infty}\left(A(t_n),\ A(t_{n-1}),\ \cdots,\ A(t_1)\right)\right]=\mathop{\times}_{n=1}^{\infty}\left(\mathcal{A}(t_n),\ \mathcal{A}(t_{n-1}),\ \cdots,\ \mathcal{A}(t_1)\right),$$

is bijective and the mapping

$$\mathop{\times}_{n=1}^{\infty}\left(\mathcal{A}(t_n),\ \mathcal{A}(t_{n-1}),\ \cdots,\ \mathcal{A}(t_1)\right)\to\sum\nolimits_{n=1}^{\infty}a_n\prod_{j=1}^{n}\mathcal{A}(t_j)$$

factors though the tensor algebra $\oplus_{n=1}^{\infty}\left\{\otimes_{j=1}^{n}L[\mathcal{H}(t_j)]\right\}$ via the universal property of that object (see Hu [HU], p. 19). We now note that $\oplus_{n=1}^{\infty}\left\{\otimes_{j=1}^{n}L[\mathcal{H}(t_j)]\right\}\subset L[\mathcal{FD}_{\otimes}^{2}]$. In diagram form we have:

$$\begin{array}{ccc}
\mathop{\times}_{n=1}^{\infty}\left(A(t_n),\ \cdots,\ A(t_1)\right)\in\mathop{\times}_{n=1}^{\infty}\left\{\mathop{\times}_{j=1}^{n}L[\mathcal{H}]\right\} & \xrightarrow{f} & \sum_{n=1}^{\infty}a_n\prod_{n\geqslant j\geqslant 1}A(t_j)\in L[\mathcal{H}] \\
\Theta\downarrow & & \uparrow dT \\
\mathop{\times}_{n=1}^{\infty}\left(\mathcal{A}(t_n),\ \cdots,\ \mathcal{A}(t_1)\right)\in\mathop{\times}_{n=1}^{\infty}\left\{\mathop{\times}_{j=1}^{n}L[\mathcal{H}(t_j)]\right\} & \xrightarrow{f_{\otimes}} & \sum_{n=1}^{\infty}a_n\prod_{j=1}^{n}\mathcal{A}(t_j)\in L[\mathcal{FD}_{\otimes}^{2}]
\end{array}$$

so that $dT\circ f_{\otimes}\circ\Theta=f$. □

Example 7.19. *If $A,B\in L[\mathcal{H}]$ and $s<t$, then $\mathcal{A}(t)\mathcal{B}(s)=\mathcal{B}(s)\mathcal{A}(t)$ and $dT[\mathcal{B}(s)\mathcal{A}(t)]=AB$ while $dT[\mathcal{B}(s)\mathcal{A}(t)-\mathcal{B}(t)\mathcal{A}(s)]=AB-BA$.*

Example 7.20. *Let $\mathcal{A}(t)=\mathbf{T}_{\theta}^{t}[A]$, $\mathcal{B}(t)=\mathbf{T}_{\theta}^{t}[B]$,* with *$I=[0,1]$, where $A,\ B$ are the operators in the last example. Then*

$$\sum_{k=1}^{n}\Delta t_k\left\|A(s_k)e^i-\left\langle A(s_k)e^i,e^i\right\rangle e^i\right\|^2=(b-a)\left\|Ae^i-\left\langle Ae^i,e^i\right\rangle e^i\right\|^2,$$
$$\sum_{k=1}^{n}\Delta t_k\left\|B(s_k)e^i-\left\langle B(s_k)e^i,e^i\right\rangle e^i\right\|^2=(b-a)\left\|Be^i-\left\langle Be^i,e^i\right\rangle e^i\right\|^2,$$

so that the operators are strongly continuous. Hence, $\int_0^1\mathcal{A}(s)ds$, $\int_0^1\mathcal{B}(s)ds$ both exist as strong integrals and

$$e^{\int_0^1[\mathcal{A}(s)+\mathcal{B}(s)]ds}=\exp\{\int_0^1\mathcal{A}(s)ds\}\exp\{\int_0^1\mathcal{B}(s)ds\}\ (a.s). \tag{7.8}$$

Expanding the right-hand side, we obtain:

$$\exp\{\int_0^1 \mathcal{A}(s)ds\}\exp\{\int_0^1 \mathcal{B}(s')ds'\} = \exp\{\int_0^1 \mathcal{A}(s)ds\}\sum\nolimits_{n=0}^{\infty} \frac{\left[\int_0^1 \mathcal{B}(s')ds'\right]^n}{n!}$$
$$= \exp\{\int_0^1 \mathcal{A}(s)ds\} + \exp\{\int_0^1 \mathcal{A}(s)ds\}\int_0^1 \mathcal{B}(s')ds'$$
$$+\tfrac{1}{2}\exp\{\int_0^1 \mathcal{A}(s)ds\}\int_0^1 \mathcal{B}(s')ds'\int_0^1 \mathcal{B}(s'')ds'' + \cdots$$
$$= \exp\{\int_0^1 \mathcal{A}(s)ds\} + \int_0^1 \exp\{\int_0^1 \mathcal{A}(s)ds\}\mathcal{B}(s')ds'$$
$$+\tfrac{1}{2}\int_0^1\int_0^1 \exp\{\int_0^1 \mathcal{A}(s)ds\}\mathcal{B}(s')\mathcal{B}(s'')ds'ds'' + \cdots .$$

Restricting to the second term, we have

$$e^{\int_0^1[\mathcal{A}(s)+\mathcal{B}(s)]ds} = \exp\{\int_0^1 \mathcal{A}(s)ds\}$$
$$+\int_0^1 \exp\{\int_0^{s'} \mathcal{A}(s)ds\}\mathcal{B}(s')\exp\{\int_{s'}^1 \mathcal{A}(s)ds\}ds' + \cdots .$$

Thus, to second order, we have:

$$\exp\{A+B\} = dT\left[\exp\{\int_0^1 [\mathcal{A}(s)+\mathcal{B}(s)]ds\}\right]$$
$$= dT\left[\exp\{\int_0^1 \mathcal{A}(s)ds\}\right] + dT\left[\int_0^1 \exp\{\int_{s'}^1 \mathcal{A}(s)ds\}\mathcal{B}(s')\exp\{\int_0^{s'} \mathcal{A}(s)ds\}ds'\right] + \cdots$$
$$= \exp\{A\} + \int_0^1 \exp\{(1-s)A\}B\exp\{sA\}ds + \cdots .$$

This last example was given by Feynman [F].

At this point, we should revisit the Trotter–Kato product theorem, mentioned in Chap. 5 (see Goldstein [GS], p. 44 and references therein).

Theorem 7.21. *(Trotter) Suppose* A_0, A_1 and $A_0 + A_1$ *generate* C_0-*contraction semigroups* $S(t)$, $T(t)$, $U(t)$ on $\mathcal{H}$. *Then*

$$\lim_{n\to\infty}\left\{S\left(\tfrac{t}{n}\right)T\left(\tfrac{t}{n}\right)\right\}^n = U(t).$$

Remark 7.22. There are cases in which the above limit exists without the assumption that $A_0 + A_1$ generates a C_0-contraction semigroup. In fact, it is possible for the limit to exist while $D(A_0)\cap D(A_1) = \{0\}$. Goldstein [GS] calls the generator C of such a semigroup a generalized or Lie sum and writes it $C = A_0 \oplus_L A_1$(see page 57). Kato [KA1] proves that the limit can exist for an arbitrary pair of self-adjoint contraction semigroups. The fundamental question is: What are the general conditions that makes this possible?

Theorem 7.23 (Generalized Trotter–Kato)**.** *Suppose* A, B and $C = A \oplus_L B$ *generate* C_0*-contraction semigroups* $S(t)$, $T(t)$ and $U(t)$ on $\mathcal{H}$*. Then*

$$d\mathcal{T}\left[\exp\{\int_0^t [\mathcal{A}(s)+\mathcal{B}(s)]\,ds\}\right] = \lim_{n\to\infty} d\mathcal{T}\left[\prod\nolimits_{j=1}^{n} \exp\{\tfrac{t}{n}(\mathcal{A}(\tfrac{jt}{n})+\mathcal{B}(\tfrac{jt'_n}{n}))\}\right]$$
$$= \lim_{n\to\infty} d\mathcal{T}\left[\prod\nolimits_{1}^{n} \exp\{\tfrac{t}{n}(\mathcal{A}(\tfrac{jt}{n})\}\exp\{\mathcal{B}(\tfrac{jt'_n}{n}))\}\right] = \exp\{t(A \oplus_L B)\},$$

where $t'_n = t(1 - \frac{1}{10^{10}}e^{-(n+1)^2})$.

7.6. The Second Dyson Conjecture

In [DY], Freeman Dyson analyzed the renormalized perturbation expansion for quantum electrodynamics and made four conjectures. His second conjecture suggested that the series expansion actually diverges. He concluded that we could at best hope that it is asymptotic. His arguments were based on unconvincing physical considerations and no precise (mathematical) formulation of the problem was possible at that time. However, the calculations of Hurst [HR], Thirring [TH], Peterman [PE], and Jaffe [JA] for specific models all supported Dyson's contention that the renormalized perturbation series may well diverge. In 1996 [DY1] (pp. 13–16), Dyson's views on the perturbation series and renormalization are reiterated: "…in spite of all the successes of the new physics, the two questions that defeated me in 1951 remain unsolved." Here, he is referring to the question of mathematical consistency for the whole renormalization program, and the ability to reliably calculate nuclear processes in quantum chromodynamics. (For other details and references to additional works, see Schweber [SC], Wightman [W], and Zinn-Justin [ZJ].) A satisfactory (mathematical) foundation for quantum field theory is still an open problem. (Many in the mathematics and physics community have become silent on this question.)

In this section we use the Feynman operator calculus to resolve Dyson's second conjecture under conditions that apply to any theory which does not make a radical departure from basic quantum theory (i.e., unitary solution operators). It also applies to the renormalized expansions in some areas of condensed matter physics where the solution operators are contraction semigroups. We begin with an operator version of the notion of asymptotic expansion, as used in ordinary differential equations (see Coddington and Levinson [CL]).

Definition 7.24. The evolution operator $\mathbf{U}^w[t,a] = \exp\{w\mathbf{Q}[t,a]\}$ is said to be asymptotic in the sense of Poincaré if, for each n and each $\Phi_a \in D\left[(\mathbf{Q}[t,a])^{n+1}\right]$, we have

$$\lim_{w\to 0} w^{-(n+1)}\left\{\mathbf{U}^w[t,a] - \sum_{k=1}^{n}\frac{(w\mathbf{Q}[t,a])^k}{k!}\right\}\Phi_a = \frac{\mathbf{Q}[t,a]^{n+1}}{(n+1)!}\Phi_a. \tag{7.9}$$

This is our (unbounded) operator version of an asymptotic expansion in the classical sense.

Theorem 7.25. *Suppose that* $\mathbf{Q}[t,a]$ *generates a contraction* C_0*-semigroup on* $\mathcal{FD}^2_\otimes$ *for each* $t \in I$*. Then:*

(1) *The operator* $\mathbf{U}^w[t,a] = exp\{w\mathbf{Q}[t,a]\}$ *is asymptotic in the sense of Poincaré.*

(2) *For each* n *and each* $\Phi_a \in D\left[(\mathbf{Q}[t,a])^{n+1}\right]$*, we have*

$$\begin{aligned}\Phi(t) &= \Phi_a + \sum_{k=1}^{n} w^k \int_a^t ds_1 \int_a^{s_1} ds_2 \cdots \int_a^{s_{k-1}} ds_k \mathcal{A}(s_1)\cdots\mathcal{A}(s_k)\Phi_a \\ &+ \int_0^w (w-\xi)^n d\xi \int_a^t ds_1 \int_a^{s_1} ds_2 \cdots \int_a^{s_n} ds_{n+1}\mathcal{A}(s_1)\cdots\mathcal{A}(s_{n+1})\mathbf{U}^{\xi}[s_{n+1},a]\Phi_a,\end{aligned} \tag{7.10}$$

where $\Phi(t) = \mathbf{U}^w[t,a]\Phi_a$*.*

Remark 7.26. The above case includes all generators of C_0-unitary groups. Thus, the theorem provides a precise formulation and proof of Dyson's second conjecture: that in general, we can only expect the expansion to be asymptotic. Actually, we prove more since we provide the remainder term, which makes the perturbation expansion (mathematically) exact for all n. However, in actual practice, the expansion may be useless. For example, if bound states are present, all important information resides in the remainder term for every n.

Proof. From Theorem 7.8, we have

$$\mathbf{U}^{w}[t,a]\Phi = \left\{ \sum_{k=0}^{n} \frac{(w\mathbf{Q}[t,a])^{k}}{k!} + \frac{1}{n!} \int_{0}^{w} (w-\xi)^{n} \mathbf{Q}[t,a]^{n+1} \mathbf{U}^{\xi}[t,a] d\xi \right\} \Phi,$$

so that

$$w^{-(n+1)} \left\{ \mathbf{U}^{w}[t,a]\Phi_a - \sum_{k=0}^{n} \frac{(w\mathbf{Q}[t,a])^{k}}{k!} \Phi_a \right\} =$$
$$\frac{(n+1)}{(n+1)!} w^{-(n+1)} \int_{0}^{w} (w-\xi)^{n} d\xi \mathbf{U}^{\xi}[t,a] \mathbf{Q}[t,a]^{n+1} \Phi_a.$$

Replace the right-hand side by

$$\mathbf{I} = \frac{(n+1)}{(n+1)!} w^{-(n+1)} \int_{0}^{w} (w-\xi)^{n} d\xi \left\{ \mathbf{U}_z^{\xi}[t,a] + \left[\mathbf{U}^{\xi}[t,a] - \mathbf{U}_z^{\xi}[t,a] \right] \right\} \mathbf{Q}[t,a]^{n+1} \Phi_a$$
$$= \mathbf{I}_{1,z} + \mathbf{I}_{2,z},$$

where

$$\mathbf{I}_{1,z} = \frac{(n+1)}{(n+1)!} w^{-(n+1)} \int_{0}^{w} (w-\xi)^{n} d\xi \mathbf{U}_z^{\xi}[t,a] \mathbf{Q}[t,a]^{n+1} \Phi_a,$$

and

$$\mathbf{I}_{2,\mathrm{z}} = \frac{(n+1)}{(n+1)!} w^{-(n+1)} \int_{0}^{w} (w-\xi)^{n} d\xi \left[\mathbf{U}^{\xi}[t,a] - \mathbf{U}_{\mathrm{z}}^{\xi}[t,a] \right] \mathbf{Q}[t,a]^{n+1} \Phi_a.$$

Since $\mathbf{U}^{\xi}[t,a] - \mathbf{U}_z^{\xi}[t,a] \to 0$, we see that $\lim_{z\to\infty} \mathbf{I}_{2,z} = 0$. Let $\varepsilon > 0$ be given and choose Z such that $z > Z \Rightarrow \|\mathbf{I}_{2,z}\| < \varepsilon$. Now, use

$$\mathbf{U}_{\mathrm{z}}^{\xi}[t,a] = \mathrm{I}_{\otimes} + \sum\nolimits_{k=1}^{\infty} \frac{\xi^{k} \mathbf{Q}_{\mathrm{z}}^{\mathrm{k}}[t,a]}{k!}$$

for the first term to get that

$$\mathbf{I}_{1,z} = \frac{(n+1)}{(n+1)!} w^{-(n+1)} \int_{0}^{w} (w-\xi)^{n} d\xi \left\{ \mathrm{I}_{\otimes} + \sum\nolimits_{k=1}^{\infty} \frac{\xi^{k} \mathbf{Q}_z^{k}[t,a]}{k!} \right\} \mathbf{Q}[t,a]^{n+1} \Phi_a.$$

If we compute the elementary integrals, we get

$$\mathbf{I}_{1,\mathrm{z}} = \frac{1}{(n+1)!}\mathbf{Q}[t,a]^{n+1}\Phi_a$$
$$+ \sum\nolimits_{\mathrm{k}\,=\,1}^{\infty} \frac{1}{k!n!}\left\{\sum\nolimits_{l=1}^{n} \binom{n}{l} \frac{w^k}{(n+k+1-l)}\right\} \mathbf{Q}_{\mathrm{z}}^k[t,a]\mathbf{Q}[t,a]^{n+1}\Phi_a.$$

Then

$$\left\|\mathbf{I} - \frac{1}{(n+1)!}\mathbf{Q}[t,a]^{n+1}\Phi_a\right\| <$$
$$\left\|\sum\nolimits_{\mathrm{k}\,=\,1}^{\infty} \frac{1}{k!n!}\left\{\sum\nolimits_{l=1}^{n} \binom{n}{l} \frac{w^k}{(n+k+1-l)}\right\} \mathbf{Q}_{\mathrm{z}}^k[t,a]\mathbf{Q}[t,a]^{n+1}\Phi_a\right\| + \varepsilon.$$

Now let $w \to 0$ to get

$$\left\|\mathbf{I} - \frac{1}{(n+1)!}\mathbf{Q}[t,a]^{n+1}\Phi_a\right\| < \varepsilon.$$

Since ε is arbitrary, $\mathbf{U}[t,a] = \exp\{\mathbf{Q}[t,a]\}$ is asymptotic in the sense of Poincaré.

To prove (7.10), let $\Phi_a \in D\left[(\mathbf{Q}[t,a])^{n+1}\right]$ for each $\mathrm{k} \leqslant \mathrm{n} + 1$, and use the fact that (Dollard and Friedman [DF])

$$\begin{aligned}(\mathbf{Q}_{\mathrm{z}}[t,a])^k\,\Phi_a &= \left(\int_a^t \mathcal{A}_{\mathrm{z}}(s)ds\right)^k \Phi_a \\ &= (k!)\int_a^t ds_1 \int_a^{s_1} ds_2 \cdots \int_a^{s_{k-1}} ds_n \mathcal{A}_{\mathrm{z}}(s_1)\mathcal{A}_{\mathrm{z}}(s_2)\cdots\mathcal{A}_{\mathrm{z}}(s_k)\Phi_a.\end{aligned} \tag{7.11}$$

Letting $z \to \infty$ gives the result. □

There are special cases in which the perturbation series may actually converge to the solution. From Theorems 5.33 and 5.36 (1), we know that if $A_0(t)$ is a nonnegative self-adjoint operator on $\mathcal{H}$, then $\exp\{-\tau A_0(t)\}$ is an analytic C_0-contraction semigroup for $\operatorname{Re}\tau > 0$.

Theorem 7.27. *Let* $\mathbf{Q}_0[t,a] = \int_a^t \mathcal{A}_0(s)ds$ *and* $\mathbf{Q}_1[t,a] = \int_a^t \mathcal{A}_1(s)ds$ *be nonnegative self-adjoint generators of analytic* C_0*-contraction semigroups for* $t \in (a,b]$*. Suppose* $D(\mathbf{Q}_1[t,a]) \supseteq D(\mathbf{Q}_0[t,a])$ *and there are positive constants* α, β *such that*

$$\|\mathbf{Q}_1[t,a]\Phi\|_{\otimes} \leqslant \alpha\,\|\mathbf{Q}_0[t,a]\Phi\|_{\otimes} + \beta\,\|\Phi\|_{\otimes}\,, \quad \Phi \in D(\mathbf{Q}_0[t,a]). \tag{7.12}$$

(1) *Then* $\mathbf{Q}[t,a]{=}\mathbf{Q}_0[t,a]{+}\mathbf{Q}_1[t,a]$ *and* $\mathcal{A}_{\mathbf{I}}(t){=}\bar{\mathbf{U}}_0[a,t]\mathcal{A}_1(t)$ $\bar{\mathbf{U}}_0[t,a]$ *both generate analytic* C_0*-contraction semigroups.*

(2) *For each* k *and each* $\Phi_a \in D\left[(\mathbf{Q}_{\mathbf{I}}[t,a])^{k+1}\right]$, *we have that*

$$\mathbf{U}_{\mathbf{I}}^{w}[t,a]\Phi_a = \Phi_a + \sum_{l=1}^{k} w^l \int_a^t ds_1 \int_a^{s_1} ds_2 \cdots \int_a^{s_{k-1}} ds_k \mathcal{A}_{\mathbf{I}}(s_1)\mathcal{A}_{\mathbf{I}}(s_2)\cdots\mathcal{A}_{\mathbf{I}}(s_k)\Phi_a$$
$$+ \int_0^w (w-\xi)^k d\xi \int_a^t ds_1 \int_a^{s_1} ds_2 \cdots \int_a^{s_k} ds_{k+1}\mathcal{A}_{\mathbf{I}}(s_1)\mathcal{A}_{\mathbf{I}}(s_2)\cdots\mathcal{A}_{\mathbf{I}}(s_{k+1})\mathbf{U}_{\mathbf{I}}^{\xi}[s_{k+1},a]\Phi_a.$$

(3) *If* $\Phi_a \in \cap_{k\geqslant 1} D\left[(\mathbf{Q}_{\mathbf{I}}[t,a])^{k}\right]$ *and,* w *small enough, we have*

$$\mathbf{U}_{\mathbf{I}}^{w}[t,a]\Phi_a = \Phi_a + \sum_{k=1}^{\infty} w^l \int_a^t ds_1 \int_a^{s_1} ds_2 \cdots \int_a^{s_{k-1}} ds_k \mathcal{A}_{\mathbf{I}}(s_1)\mathcal{A}_{\mathbf{I}}(s_2)\cdots\mathcal{A}_{\mathbf{I}}(s_k)\Phi_a.$$

Proof. The proof of (1) is almost the same as in Theorem 5.34, so we provide an outline. As there, use the fact that $\mathbf{Q}_0[t,a]$ generates an analytic C_0-contraction semigroup to find a sector Σ in the complex plane, with $\rho(\mathbf{Q}_0[t,a]) \supset \Sigma$ ($\Sigma = \{\lambda : \; |\arg\lambda| < \pi/2 + \delta'\}$, for some $\delta' > 0$), and for $\lambda \in \Sigma$,

$$\|R(\,\lambda\, :\, \mathbf{Q}_0[t,a])\|_{\otimes} \leqslant |\lambda|^{-1}.$$

From (5.14), $\mathbf{Q}_1[t,a]R(\,\lambda\, :\, \mathbf{Q}_0[t,a])$ is a bounded operator and:

$$\begin{aligned}
&\|\mathbf{Q}_1[t,a]R(\,\lambda\, :\, \mathbf{Q}_0[t,a])\Phi\|_{\otimes} \\
&\quad \leqslant \alpha \|\mathbf{Q}_0[t,a]R(\,\lambda\, :\, \mathbf{Q}_0[t,a])\Phi\|_{\otimes} + \beta \|R(\,\lambda\, :\, \mathbf{Q}_0[t,a])\Phi\|_{\otimes} \\
&\quad \leqslant \alpha \|[R(\,\lambda\, :\, \mathbf{Q}_0[t,a]) - \mathbf{I}]\,\Phi\|_{\otimes} + \beta\,|\lambda|^{-1}\,\|\Phi\|_{\otimes} \\
&\quad \leqslant 2\alpha \|\Phi\|_{\otimes} + \beta\,|\lambda|^{-1}\,\|\Phi\|_{\otimes}.
\end{aligned}$$

Thus, if we set $\alpha = 1/4$ and $|\lambda| > 2\beta$, we have

$$\|\mathbf{Q}_1[t,a]R(\,\lambda\, :\, \mathbf{Q}_0[t,a])\|_{\otimes} < 1,$$

and it follows that the operator

$$\mathbf{I} - \mathbf{Q}_1[t,a]R(\,\lambda\, :\, \mathbf{Q}_0[t,a])$$

is invertible. Now it is easy to see that:

$$(\lambda\mathbf{I} - (\mathbf{Q}_0[t,a] + \mathbf{Q}_1[t,a]))^{-1} = R(\,\lambda\, :\, \mathbf{Q}_0[t,a])\,(\mathbf{I} - \mathbf{Q}_1[t,a]R(\,\lambda\, :\, \mathbf{Q}_0[t,a]))^{-1}.$$

It follows that, using $|\lambda| > 2\beta$, with $|\arg \lambda| < \pi/2 + \delta''$ for some $\delta'' > 0$, and the fact that $\mathbf{Q}_0[t, a]$ and $\mathbf{Q}_1[t, a]$ are nonnegative generators, we get that

$$\|R(\lambda : \mathbf{Q}_0[t, a] + \mathbf{Q}_1[t, a])\|_{\otimes} \leqslant |\lambda|^{-1} .$$

Thus $\mathbf{Q}_0[t, a] + \mathbf{Q}_1[t, a]$ generates an analytic C_0-contraction semigroup. The proof of (2) follows from Theorem 7.25. Finally, if w is such that $|\arg w| \leqslant \delta' < \delta$ and $|w - a| \leqslant Ca$ for some constant C, (3) follows from Theorem 5.33 (2). □

There are also cases where the series may diverge, but still respond to some summability method. This phenomenon is well known in classical analysis. In field theory, things can be much more complicated. The book by Glimm and Jaffe [GJ] has a good discussion.

7.7. Foundations for the Feynman Worldview

As discussed earlier, Feynman took a holistic view of physical reality in his development of quantum electrodynamics. He suggested that we view a physical event as occurring on a film which exposes more and more of the outcome as the film unfolds. The purpose of this section is to develop the mathematical framework for a theory of physical measurement on a three-dimensional motion picture.

We would like to begin by investigating what is actually known about our view of the micro-world. The objective is to provide the background for a number of physically motivated postulates that will allow us to develop our theory. This will also make it possible to relate the Feynman operator calculus to the idea of a sum over all paths for a system moving from one space-time point to another.

In spite of the enormous successes of the physical sciences in the last 100 years, our information and understanding of the micro-world is still rather meager. In the macro-world we are quite comfortable with the view that physical systems evolve continuously in time and our results justify this view. Indeed, the success of continuum physics is the basis for a large part of the technical advances in the twentieth century. On the other hand, the same view is also held at the micro-level and, in this case, our position is not very secure. The ability to measure physical events continuously in time at the micro-level must be considered a belief which, although convenient, has no basis in reality.

In order to establish perspective, let us consider a satisfactory and well-justified theory, Brownian motion. This theory lies at the interface between the macro- and the micro-worlds. The careful presentations of this theory make a distinction between the mathematical and the physical foundations of Brownian motion and that distinction is important for our discussion.

When Einstein [EI] began his investigation of the physical issues associated with this phenomenon, he could only assume that physical information about the state of a Brownian particle (position, velocity, etc.) could be known in time intervals that were large compared with the mean time between molecular collisions. It is known that, under normal physical conditions, a Brownian particle receives about 10^{21} collisions per second. To gain perspective, the attosecond is 10^{-18} s, the time it takes light to travel the length of three hydrogen atoms and 12 as is the shortest measured time interval (as of 2010). It follows that Brownian particle receives about 1000 collisions per attosecond, so that we are (as yet) unable to see one collision of a Brownian particle. The smallest known (physical) time is the Planck time defined as:

$$t_p \triangleq \sqrt{\frac{\hbar G}{c^5}} \simeq 5.39106(32) \times 10^{-44}\,\mathrm{s}$$

where $\hbar$ is Planck's constant, G is the gravitational constant, and c is the speed of light. This is approximately 10^{-26} as, a time interval that is far from our ability to measure.

For reasons of simplification, Einstein set the mean time between collisions at zero and obtained the diffusion equation in terms of known physical constants. This allowed him to predict the coefficient of viscosity of the fluid from the diffusion constant, which was later verified by experiment (providing proof that atoms existed). Wiener conducted the first rigorous analysis of Einstein's theory, providing the mathematical theory of Brownian particle and the important Wiener measure. He showed that Einstein's simplification corresponds to the assumption that the ratio of the mass of the particle to the friction of the fluid is zero in the limit (see Wiener et al. [WI]).

Physically the Einstein model was not completely satisfactory since it led to problems of unbounded path length and nondifferentiability at all points. The first problem is physically impossible while the second is physically unreasonable. Of course, this idealization has turned out to be quite satisfactory in areas where the information required need not be very detail, such as large parts of physics, chemistry, biological, and engineering sciences. Ornstein and Uhlenbeck [OU] constructed a model that provides the Einstein view asymptotically, but, in small-time regions, is equivalent to the assumption that the particle travels a linear path between collisions. This model provides finite path length and differentiability. (The theory was later made mathematically rigorous by Doob [DO].) What we do know is that the very nature of the liquid state implies collective behavior among the molecules. This means that we do not know what path the particle travels between collisions. However, since the tools and methods of analysis require some form of continuity, some such (in between observation) assumptions must be made.

Theoretical science is necessarily constrained in attempts to construct mathematical representations of certain restricted portions of physical reality. Simplicity forces one to restrict models to the minimum number of variables, relationships, constraints, etc., which give a satisfactory account of known experimental results and possibly allow for the prediction of heretofore unknown consequences. One important outcome of this approach has been to implicitly eliminate all reference to the background within which all physical systems evolve. In the micro-world, such an action can never be completely justified. We propose to replace the use of mathematical coordinate systems, $\mathbb{R}^3$, by "physical coordinate systems," $\mathbb{R}_p^3$ in order to (partially) remedy this problem. We assume that $\mathbb{R}_p^3(t)$ is attached to an observer (including measuring devices) and is envisioned as $\mathbb{R}^3$ plus all background effects, either local or distant, which affect the observer's ability to obtain precise (ideal) experimental information about the physical world at time t. This in turn affects the observer's ability to construct precise (ideal) representations and make precise predictions about the micro-world. More specifically, we consider the evolution of some micro-system on the interval $I = [a, b]$. Physically this evolution manifests itself as a curve on $\mathbb{X}$, where

$$\prod_{t \in I} \mathbb{R}_p^3(t) = \mathbb{X}.$$

Thus, true physical events occur on $\mathbb{X}$, where actual experimental information is modified by fluctuations in $\mathbb{R}_p^3(t)$, and by the interaction of the micro-system with any measuring equipment. Based on the success of our models, we know that such small changes are in the noise and have no effect on our understanding of macro-systems. However, there is no reason to believe that these fluctuations will not effect micro-systems. Let $\mathbb{X}_o$ represent the observer's space of obtainable information.

In terms of our theoretical representations, we are forced to model physical systems in terms of functions, amplitudes, and/or operator-valued distributions, etc. Thus, there are three spaces, the actual physical space of evolution for the micro-system $\mathbb{X}$, the observer's space of obtainable information concerning this evolution, $\mathbb{X}_o$, and the theoretical representation or mathematical model space of the physical system that is used to explain the observer's experimental information. The lack of distinction between these three spaces seems to be the cause for some of the confusion and lack of clarity. For example, it may be perfectly correct to assume that a particle travels a continuous path on $\mathbb{X}$. However, the assumption that $\mathbb{X}_o$ includes infinitesimal space-time knowledge of this path is not true. This leads to our first postulate

Postulate (1): Physical reality is a continuous process in time.

We thus take this view, fully recognizing that experiment does not provide continuous information about physical reality, and that there is no reason to believe that our mathematical representations (models) contain precise information about the continuous space-time behavior of physical processes.

Before continuing, it will be helpful to have a particular physical picture in mind. For this purpose, we take this picture to be a photograph showing the track left by a π-meson in a bubble chamber, $\pi^+ \to \mu^+ + \nu$ (Fig. 7.1).

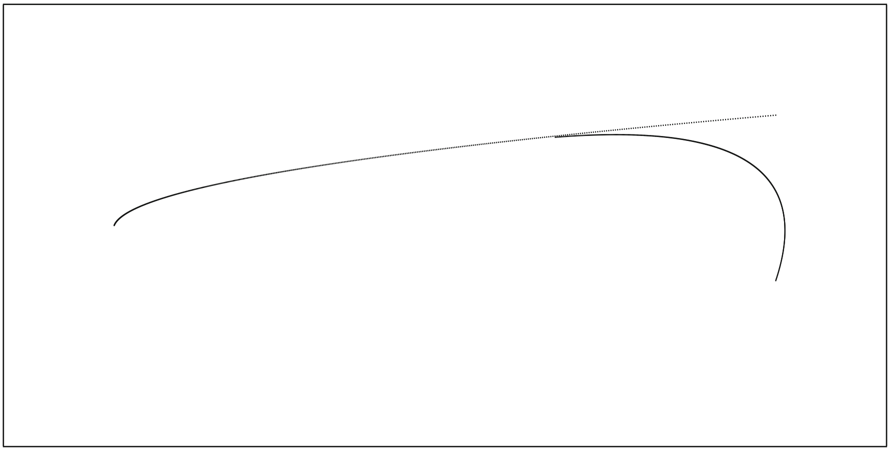

Figure 7.1. Ideal picture of the reaction $\pi^+ \to \mu^+ + \nu$

We further assume that the orientation of our photograph is such that the π-meson enters on the left at time $t = 0$ and the tracks left by the μ-meson disappear on the right at time $t = T$, where T is of the order of 10^{-3}-s, the time exposure for photographic film. Although the neutrino does not appear in the photograph, we also include a track for it.

We have drawn the photograph as if we continuously see the particles in the picture. However, experiment only provides us individual bubbles, which do not necessarily overlap, from which we must extract physical information. A more accurate (though still unrealistic) depiction is given in Fig. 7.2.

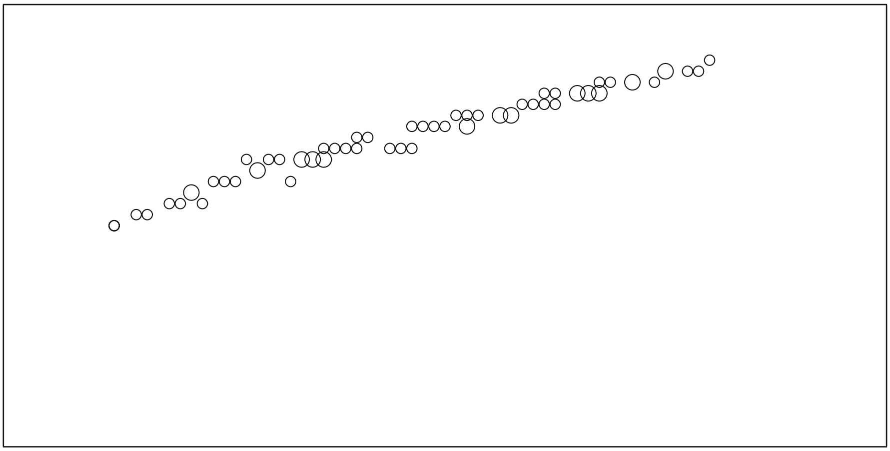

Figure 7.2. More accurate picture of the reaction $\pi^+ \to \mu^+ + \nu$

Let us assume that we have magnified a portion of our photograph to the extent that we may distinguish the individual bubbles created by the π-meson as it passes through the chamber. In Fig. 7.3, we present a simplified model of adjacent bubbles.

Postulate (2): We assume that the center of each bubble represents the average knowable effect of the particle in a symmetric time interval about the center.

By average knowable effect, we mean the average of the physical observables. In Fig. 7.3, we consider the existence of a bubble at time τ_j to be caused by the average of the physical observables over the time interval $[t_{j-1}, t_j]$, where $t_{j-1} = (1/2)[\tau_{j-1} + \tau_j]$ and $t_j = (1/2)[\tau_j + \tau_{j+1}]$. This postulate requires some justification. In general,

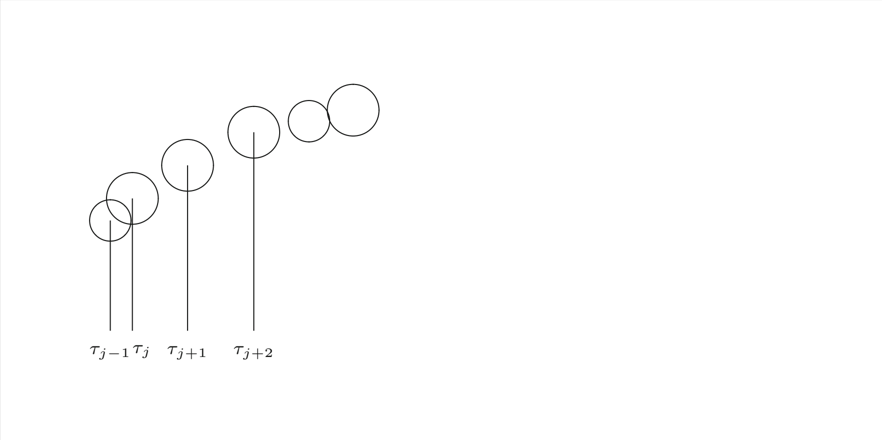

Figure 7.3. Highly magnified view of a portion showing individual bubbles

the resolution of the film and the relaxation time for distinct bubbles in the chamber vapor are limited. This means that if the π-meson creates two bubbles that are closely spaced in time, the bubbles may coalesce and appear as one. If this does not occur, it is still possible that the film will record the event as one bubble because of its inability to resolve events in such small time intervals. Let us now recognize that we are dealing with one photograph so that, in order to obtain all available information, we must analyze a large number of photographs of the same reaction obtained under similar conditions (pre-prepared states). It is clear that the number of bubbles and the time placement of the bubbles will vary (independently of each other) from photograph to photograph. Let λ^{-1} denote the average time for the appearance of a bubble in the film.

Postulate (3): We assume that the number of bubbles in any film is a random variable.

Postulate (4): We assume that, given that n bubbles have appeared on a film, the time positions of the centers of the bubbles are uniformly distributed.

Postulate (5): We assume that $N(t)$, the number of bubbles up to time t in a given film is a Poisson-distributed random variable with parameter λ.

To motivate Postulate 5, recall that τ_j is the time center of the j-th bubble and λ^{-1} is the average (experimentally determined) time between bubbles. The following results can be found in Ross [RO].

Theorem 7.28. *The random variables* $\Delta\tau_j = \tau_j - \tau_{j-1}$ $(\tau_0 = 0)$ *are independent identically distributed random variables of exponential type with mean* λ^{-1}*, for* $1 \leq j \leq n$.

The arrival times $\tau_1, \tau_2, \cdots, \tau_n$ are not independent, but their density function can be computed from

$$\begin{aligned} &Prob\,[\tau_1, \cdots, \tau_n] \\ &= Prob\,[\tau_1]\,Prob\,[\tau_2|\tau_1] \cdots Prob\,[\tau_n|\tau_1, \cdots, \tau_{n-1}]\,. \end{aligned} \tag{7.13}$$

We now use the above theorem to conclude that, for $k \geq 1$,

$$Prob\,[\tau_k|\tau_1, \tau_2, \cdots, \tau_{k-1}] = Prob\,[\tau_k|\tau_{k-1}]\,. \tag{7.14}$$

We don't know this conditional probability. However, the natural assumption is that, given that n bubbles appear, they are equally (uniformly) distributed on the interval. We can now construct what we call the experimental evolution operator. Assume that the conditions for the fundamental theorem are satisfied and that the family $\{\tau_1, \tau_2, \cdots, \tau_n\}$ represents the time positions of the centers of n bubbles in our film of Fig. 7.3.

We now follow physical tradition and replace $\mathcal{A}(t)$ by $H(t)$. Set $a = 0$ and define $\mathbf{Q}_E[\tau_1, \tau_2, \cdots, \tau_n]$ by

$$\mathbf{Q}_E[\tau_1, \tau_2, \cdots, \tau_n] = \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} E[\tau_j, s]H(s)ds. \tag{7.15}$$

Here, $t_0 = \tau_0 = 0$, $t_j = (1/2)[\tau_j + \tau_{j+1}]$, $1 \leq j \leq n$ and $E[\tau_j, s]$ is the exchange operator. The effect of the exchange operator is to concentrate all information contained in $[t_{j-1}, t_j]$ at τ_j. This is how we implement our postulate that the known physical event of the bubble at time τ_j is due to an average of physical effects over $[t_{j-1}, t_j]$ with information concentrated at τ_j. We can rewrite $\mathbf{Q}_E[\tau_1, \tau_2, \cdots, \tau_n]$ as

$$\mathbf{Q}_E[\tau_1, \tau_2, \cdots, \tau_n] = \sum_{j=1}^{n} \Delta t_j \left[\frac{1}{\Delta t_j} \int_{t_{j-1}}^{t_j} E[\tau_j, s]H(s)ds \right].$$

Thus, we indeed have an average as required by Postulate 5. The evolution operator is given by

$$U\left[\tau_1, \tau_2, \cdots, \tau_n\right] = \exp\left\{\sum_{j=1}^{n} \Delta t_j \left[\frac{1}{\Delta t_j}\int_{t_{j-1}}^{t_j} E[\tau_j, s]H(s)ds\right]\right\}.$$

For $\Phi \in \mathcal{FD}_{\otimes}^2$, we define the function $\mathbf{U}[N(t), 0]\Phi$ by:

$$\mathbf{U}[N(t), 0]\Phi = U\left[\tau_1, \tau_2, \cdots, \tau_{N(t)}\right]\Phi. \tag{7.16}$$

The function $\mathbf{U}[N(t), 0]\Phi$ is an $\mathcal{FD}_{\otimes}^2$ random variable which represents the distribution of the number of bubbles that may appear on our film up to time t. In order to relate $\mathbf{U}[N(t), 0]\Phi$ to actual experimental results, we must compute its expected value. Using Postulates 3, 4, and 5, we have

$$\begin{aligned}\bar{\mathbf{U}}_\lambda[t, 0]\Phi &= \mathcal{E}\left\{\mathbf{U}[N(t), 0]\Phi\right\} \\ &= \sum_{n=0}^{\infty} \mathcal{E}\left\{\mathbf{U}[N(t), 0]\Phi \mid \ N(t) = n\right\} \Pr ob[N(t) = n],\end{aligned} \tag{7.17}$$

$$\begin{aligned}&\mathcal{E}\left\{\mathbf{U}[N(t), 0]\Phi \mid \ N(t) = n\right\} \\ &= \int_0^t \tfrac{d\tau_1}{t-0}\int_0^t \tfrac{d\tau_2}{t-\tau_1}\cdots\int_0^t \tfrac{d\tau_n}{t-\tau_{n-1}}\mathbf{U}[\tau_n, \cdots\tau_1]\Phi = \bar{\mathbf{U}}_n[t, 0]\Phi,\end{aligned} \tag{7.18}$$

and

$$\Pr ob[N(t) = n] = \tfrac{(\lambda t)^n}{n!}\exp\{-\lambda t\}. \tag{7.19}$$

The integral in Eq. (7.18) acts to distribute uniformly the time positions τ_j over the successive intervals $[t, \tau_{j-1}]$, $1 \le j \le n$, given that τ_{j-1} has been determined. This is a natural result given our lack of knowledge. The integral in Eq. (7.18) is of theoretical value, but is not easy to compute. Since we are only interested in what happens when $\lambda \to \infty$, and as the mean number of bubbles in the film at time t is λt, we can take $\tau_j = \frac{jt}{n}$, $1 \leqslant j \leqslant n$, $\quad \Delta t_j = \frac{t}{n}$ (for each n). We can now replace $\bar{\mathbf{U}}_n[t, 0]\Phi$ by $\mathbf{U}_n[t, 0]\Phi$; and, with this understanding, we continue to use τ_j, so that

$$\mathbf{U}_n[t, 0]\Phi = \exp\left\{\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j} E[\tau_j, s]H(s)ds\right\}\Phi.$$

We define our experimental evolution operator $\mathbf{U}_\lambda[t,0]\Phi$ by

$$\mathbf{U}_\lambda[t,0]\Phi = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \exp\{-\lambda t\}\mathbf{U}_n[t,0]\Phi. \tag{7.20}$$

We now have the following result, which is a consequence of the fact that Borel summability is regular.

Theorem 7.29. *Assume that the conditions for our fundamental theorem are satisfied. Then*

$$\lim_{\lambda\to\infty} \bar{\mathbf{U}}_\lambda[t,0]\Phi = \lim_{\lambda\to\infty} \mathbf{U}_\lambda[t,0]\Phi = \mathbf{U}[t,0]\Phi. \tag{7.21}$$

Since $\lambda \to \infty \Rightarrow \lambda^{-1} \to 0$, this means that the average time between bubbles is zero (in the limit) so that we get a continuous path. It should be observed that this continuous path arises from averaging the sum over an infinite number of (discrete) paths. (However, from Postulate 1 we assume that the true paths are continuous.) The first term in (7.20) corresponds to the path of a π-meson that created no bubbles (i.e., the photograph is blank). This event has probability $\exp\{-\lambda t\}$ (which approaches zero as $\lambda \to \infty$). The n-th term corresponds to the path of a π-meson that created n bubbles (with probability $\frac{(\lambda t)^n}{n!}\exp\{-\lambda t\}$) etc.

In closing, observe that the ideas of this section are very general and actually independent of the example we used as a pictorial help. We will apply the theory developed here in the next chapter to prove the final remaining conjecture of Dyson concerning QED.

References

[AK] H. Araki, Expansional in Banach algebras. Ann. Sci. École Norm. Sup. Math. **4**, 67–84 (1973)

[CL] A.D. Coddington, N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955)

[DF] J.D. Dollard, C.N. Friedman, *Product Integration with Applications to Differential Equations.* Encyclopedia of Mathematics, vol. 10 (Addison-Wesley, Reading, MA, 1979)

[DO] J.L. Doob, The Brownian movement and stochastic equations. Ann. Math. **43**, 351–369 (1942)

[DY] F. Dyson, The radiation theories of Tomonaga, Schwinger, and Feynman. Phys. Rev. **75**, 486–502 (1949)

[DY1] F. Dyson, *Selected Papers of Freeman Dyson with Commentary* (American Mathematical Society, Providence, RI, 1996)

[EI] A. Einstein, *Investigations on the Theory of the Brownian Movement* (Dover, NY, 1956) [English translation of Einstein's Five Papers on Brownian Motion, ed. by R. Fürth (1926)]

[F] R.P. Feynman, An operator calculus having applications in quantum electrodynamics. Phys. Rev. **84**, 108–128 (1951)

[FW] I. Fujiwara, Operator calculus of quantized operator. Prog. Theor. Phys. **7**, 433–448 (1952)

[GJ] J. Glimm, A. Jaffe, *Quantum Physics. A Functional Integral Point of View* (Springer, New York, 1987)

[GS] J.A. Goldstein, *Semigroups of Linear Operators and Applications* (Oxford University Press, New York, 1985)

[HA] R. Haag, On quantum field theories. Dan Mat. Fys. Medd. **29**, 1–37 (1955)

[HP] E. Hille, R.S. Phillips, *Functional Analysis and Semigroups.* American Mathematical Society Colloquium Publications, vol. 31 (American Mathematical Society, Providence, RI, 1957)

[HW] L.P. Horwitz, On the significance of a recent experiment demonstrating quantum interference in time. Phys. Lett. A. **355**, 1–6 (2006)

[HU] S.T. Hu, *Elements of Modern Algebra* (Holden-Day, San Francisco, CA, 1965)

[HR] C.A. Hurst, The enumeration of graphs in the Feynman-Dyson technique. Proc. Roy. Soc. Lond. **A214**, 44–61 (1952)

[JA] A. Jaffe, Divergence of perturbation theory for bosons. Commun. Math. Phys. **1**, 127 (1965)

[KA] T. Kato, *Perturbation Theory for Linear Operators*, 2nd edn. (Springer, New York, 1976)

[KA1] T. Kato, Trotter's product formula for an arbitrary pair of self-adjoint contraction semigroups, in *Topics in Functional Analysis.* Advances in Mathematics: Supplementary Studies, vol. 3 (Academic, New York, 1978), pp. 185–195

[MW] W.L. Miranker, B. Weiss, The Feynman operator calculus. SIAM Rev. **8**, 224–232 (1966)

[MM] N.F. Mott, H.S. Massey, *Theory of Atomic Collisions* (Clarendon Press, Oxford, 1965)

[N] E. Nelson, Operants: a functional calculus for non-commuting operators, in *Functional Analysis and Related Fields*, ed. by F. Browder (Springer, Berlin, 1970)

[OU] L.S. Ornstein, G.E. Uhlenbeck, Phys. Rev. **36**, 1103 (1930)

[PZ] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Applied Mathematical Sciences, vol. 44 (Springer, New York, 1983)

[PE] A. Petermann, Renormalisation dans les séries divergentes. Helv. Phys. Acta **26**, 291–299 (1953)

[RO] S.M. Ross, *Introduction To Probability Models*, 4th edn. (Academic, New York, 1989)

[SC] S.S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper & Row, New York, 1961)

[SG] I. Segal, *The mathematical theory of quantum fields*, in *Lectures in Modern Analysis and Application II*, (G. T. Tamm editor) Lecture Notes in Math. Vol. 140, Berlin-Heidelberg-New York (1970), pp. 30–57

[SW] R.F. Streater, A.S. Wightman, *PCT, Spin and Statistics and All That* (Benjamin, New York, 1964)

[TH] W. Thirring, On the divergence of the perturbation theory for quantum fields. Helv. Phys. Acta **16**, 33–52 (1953)

[WI] N. Wiener, A. Siegel, B. Rankin, W.T. Martin, *Differential Space, Quantum Systems, and Prediction* (MIT Press, Cambridge, MA, 1966)

[W] A.S. Wightman, in *The Lesson of Quantum Theory*, ed. by J. de Boer, E. Dal, O. Ulfbeck (Elsevier, Amsterdam, 1986), p. 201

[ZJ] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 2nd edn. (Clarendon, Oxford, 1993)

Chapter 8

Applications of the Feynman Calculus

Introduction

This chapter is devoted to a few applications of the Feynman operator calculus. We first consider the theory of linear evolution equations and provide a unified approach to a class of time-dependent parabolic and hyperbolic equations.

We then show that $KS^2[\mathbb{R}^3]$ allows us to construct the elementary path integral in the manner intended by Feynman. We also use the sum over paths theory of the last chapter along with time-ordering to extend the Feynman path integral to a very general setting. We then prove an extended version of the Feynman–Kac theorem. Finally, we prove the last remaining Dyson conjecture concerning the foundations for quantum electrodynamics.

8.1. Evolution Equations

As our first application, we provide a unified approach to a class of time-dependent parabolic and hyperbolic evolution equations. We restrict ourselves to first and second order initial-value problems

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6_8

$\dot{u}(t) = A(t)u(t)$, $u(0) = u_0$, or $\ddot{v}(t) = B^2(t)v(t)$, $v(0) = v_1$ and $\dot{v}(0) = v_2$. In each case, we assume that $A(t)$, $B(t)$ generates a C_0-semigroup for each $t \in I$.

For second order equations, let

$$u(t) = \begin{pmatrix} v(t) \\ \dot{v}(t) \end{pmatrix}, \quad u_0 = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad A(t) = \begin{pmatrix} 0 & I \\ B^2(t) & 0 \end{pmatrix}.$$

We now define a norm on $\mathcal{X} = \mathcal{H} \times \mathcal{H}$ by

$$\left\| \begin{pmatrix} f \\ g \end{pmatrix} \right\|_{\mathcal{X}} = \|f\|_{\mathcal{H}} + \|g\|_{\mathcal{H}}.$$

This makes $\mathcal{X}$ a Hilbert space. It follows that the second order equation on $\mathcal{H}$ becomes the first order equation on $\mathcal{X}$: $\dot{u}(t) = A(t)u(t)$, $u(0) = u_0$. Thus, it suffices to study first order equations. For additional details on this approach, see Yosida [YS] or Goldstein [GS].

In order to prove existence and uniqueness for the initial-value (Cauchy) problem a number of conditions are imposed (see Pazy [PZ], in Chap. 7). The important assumption for the time-ordered theory is a weak continuity condition. (In the following, let $\mathcal{H}$ be a Hilbert space.)

8.2. Parabolic Equations

In the abstract parabolic problem, it is assumed that, on $\mathcal{H}$, the family $A(t)$, $t \in I$, satisfies:

(1) For each $t \in I$, $A(t)$ is densely defined, $R(\lambda;\ A(t))$ exists in a sector $\Sigma = \Sigma(\phi + \pi/2)$ for some ϕ, $0 < \phi < \pi/2$ and a constant ϕ independent of t, such that

$$\|R(\lambda;\ A(t))\| \le 1/|\lambda| \ for\ \lambda \in \Sigma,\ t \in I.$$

(2) The function $A^{-1}(t)$ is continuously differentiable on I.

(3) There are constants $C_1 > 0$ and $\rho :\ 0 < \rho < 1$, such that, for each $\lambda \in \Sigma$ and every $t \in I$, we have

$$\|D_t R(\lambda;\ A(t))\| \le C_1/|\lambda|^{1-\rho}.$$

(4) The function $DA^{-1}(t)$ is Holder continuous in $\mathcal{H}$ and there are positive constants $C_2,\ \alpha$ such that

$$\left\|DA^{-1}(t) - DA^{-1}(s)\right\| \le C_2\,|t-s|^{\alpha},\ s,\ t \in I.$$

The first condition states that $A(t)$ generates an analytic contraction semigroup for each $t \in I$. The four conditions are required to prove the following theorem.

Theorem 8.1. *Let the family $A(t)$, $t \in I$, have a common dense domain and satisfy assumptions (1)–(4). Then the problem*

$$\frac{\partial u(t)}{\partial t} = A(t)u(t), \quad u(a) = u_a,$$

has a unique solution $u(t) = V(t,s)u_a$, for $t,\ s \in I$. Furthermore,

(1) *$V(t,s)$ is strongly continuous on I and continuously differentiable (in the norm of $\mathcal{H}$) with respect to both s and $t \in I$,*

(2) *$V(t,s)\mathcal{H} \subset D(A(t))$,*

(3) *$A(t)V(t,s)$ and $V(t,s)A(s)$ are bounded,*

(4) *$D_tV(t,s) = A(t)V(t,s)$, $D_sV(t,s) = -\overline{V(t,s)A(s)}$, and*

(5) *for $t,\ s \in I$,*

$$\|D_tV(t,s)\| \leq C/(t-s), \quad \|D_sV(t,s)\| \leq C/(t-s).$$

In the proof of this result takes seven pages plus five pages of preparatory work (see page 397). (In Pazy [PZ], in Chap. 7, the proof takes 17 pages.)

Example 8.2. *Let the family of operators $A(t)$, $t \in I = [0,1]$, be defined on $\mathcal{H} = L^2(0,1)$ by:*

$$A(t)u(x) = -\frac{1}{(t-x)^2}u(x).$$

It is easy to see that each $A(t)$ is self-adjoint and $(A(t)u,u) \leq -\|u\|_{\mathcal{H}}^2$ for $u \in D(A(t))$. It follows that the spectrum of $A(t)$, $\sigma(A(t)) \subset (-\infty,-1]$, for $t \in [0,1]$. The first condition is satisfied for any $\phi \in (0,\pi/2)$, while the second condition is clear and makes the fourth condition obvious. For $\lambda \notin (-\infty,-1]$, we have

$$R(\lambda;\ A(t))u(x) = \frac{(t-x)^2}{\lambda(t-x)^2+1}u(x),$$

so that

$$\|R(\lambda;\ A(t))u(x)\|_{\mathcal{H}}^2 = \int_0^1 \frac{(t-x)^4}{[\lambda(t-x)^2+1]^2}u^2(x)dx \leq \frac{1}{|\lambda|^2}\|u\|_{\mathcal{H}}^2.$$

It is now clear that each $A(t)$ *generates a contraction semigroup and*

$$D_tR(\lambda;\ A(t))u(x) = \frac{2(t-x)}{[\lambda(t-x)^2+1]^2}u(x).$$

From here, an easy estimation shows that, for $\lambda \in \Sigma$,

$$\|D_tR(\lambda;\ A(t))\|_{\mathcal{H}} \leq \frac{C}{|\lambda|^{1/2}},$$

so that the third condition follows. The theorem would follow if there was a common dense domain. However, it is not hard to see that $\bigcap_{t\in I} D(A(t)) = \{0\}$.

We now notice that

$$(A(t)-A(s))\,A(\tau)^{-1} = \left[\frac{(\tau-x)^2}{(s-x)} + \frac{(\tau-x)^2}{(t-x)}\right](s-t),$$

so that, for some constants $C>0,\ 0<\beta\leq 1$, *we have*

$$\left\|(A(t)-A(s))\,A(\tau)^{-1}\right\| \leqslant C\,|t-s|^{\beta}\ (a.s)\ \ \textit{for all}\, t,\ s,\ \tau\in[0,1].$$

It follows that the family $A(t),\ t\in[0,1]$, *is strongly continuous and hence satisfies (7.3). Thus, the time-ordered integral exists and generates a contraction semigroup. It is now an exercise to prove that the semigroup is also analytic in the same sector,* Σ.

Returning to the abstract parabolic problem, the conditions used by Pazy [PZ], in Chap. 7, make it easy to see that the $A(t),\ t\in I$, is strongly continuous in general:

(1) For each $t\in I,\ A(t)$ generates an analytic C_0-semigroup with domains $D(A(t)) = D$ independent of t.

(2) For each $t\in I,\ R(\lambda, A(t))$ exists for all λ such that $\operatorname{Re}\lambda \leqslant 0$, and there is an $M>0$ such that:

$$\|R(\lambda, A(t))\| \leqslant M/[|\lambda|+1].$$

(3) There exist constants L and $0<\alpha\leqslant 1$ such that

$$\left\|(A(t)-A(s))\,A(\tau)^{-1}\right\| \leqslant L\,|t-s|^{\alpha}\quad \text{for all}\ t,\ s,\ \tau\in I.$$

From (3), for $\varphi\in D$, we have

$$\begin{aligned}\|[A(t)-A(s)]\,\varphi\| &= \left\|\left[(A(t)-A(s))\,A^{-1}(\tau)\right]A(\tau)\varphi\right\| \\ &\leqslant \left\|(A(t)-A(s))\,A^{-1}(\tau)\right\|\|A(\tau)\varphi\| \leqslant L\,|t-s|^{\alpha}\,\|A(\tau)\varphi\|\,.\end{aligned}$$

Thus, the family $A(t),\ t\in I$, is strongly continuous on D. For comparison with the time-ordered approach, we have:

Theorem 8.3. *Let the family $A(t)$, $t \in I$ be weakly continuous on $\mathcal{H}$ satisfying:*

(1) *For any complete orthonormal basis $\{e^i\}$, for $\mathcal{H}$ and any partition $\mathcal{P}_n$, of I with mesh μ, there is a number δ, with $0 < \delta < 1$ such that:*

$$\sum_{k=1}^{n} \Delta t_k \left\| A(s_k)e^i - \left\langle A(s_k)e^i, e^i \right\rangle e^i \right\|^2 \leqslant C\mu_n^{\delta-1} \tag{8.1}$$

(2) *For each $t \in I$, $A(t)$ generates an analytic C_0-semigroup with dense domains $D(A(t)) = D(t) \subset \mathcal{H}$.*

(3) *For each $t \in I$, $R(\lambda, A(t))$ exists for all λ such that $\operatorname{Re}\lambda \leqslant 0$, and there is an $M(t) > 0$, $t \in I$ such that:*

$$\|R(\lambda, A(t))\| \leqslant M(t)/[|\lambda| + 1],$$

with $\sup_{t\in I} M(t) < \infty$.

Then, for each $\phi \in \mathcal{H}$ the time-ordered family $\mathcal{A}(t)$, $t \in I$ has a strong Riemann integral on $D_0 = \otimes_{t\in I} D(t) \cap \mathcal{H}^2_{\otimes}(\Phi)$, which generates an analytic C_0-semigroup on $\mathcal{H}^2_{\otimes}(\Phi)$, where $\Phi = \otimes_{t\in I}\phi_t$, $\phi_t = \phi$ for all $t \in I$.

Remark 8.4. The left-hand side of Eq. (8.1) could diverge as $\mu \to 0$, but remains finite if the family $A(t)$, $t \in I$ is strongly continuous. If the family $A(t)$, $t \in I$ is not strongly continuous, Eq. (8.1) ensures that weak continuity on $\mathcal{H}$ is sufficient in order for the time-ordered family $\mathcal{A}(t)$, $t \in I$ to have a strong Riemann integral on $\mathcal{H}^2_{\otimes}(\Phi)$, for each Φ. (We do not require a common dense domain.)

8.3. Hyperbolic Equations

In the abstract approach to hyperbolic evolution equations, it is assumed that:

(1) For each $t \in I$, $A(t)$ generates a C_0-semigroup.

(2) For each $t \in I$, $A(t)$ is stable with constants $(M, 0)$ and the resolvent set $\rho(A(t)) \supset (0, \infty)$, $t \in I$, such that:

$$\left\| \prod_{j=1}^{k} \exp\{\tau_j A(t_j)\} \right\| \leqslant M.$$

(3) There exists a Hilbert space $\mathcal{Y}$ densely and continuously embedded in $\mathcal{H}$ such that, for each $t \in I$, $D(A(t)) \supset \mathcal{Y}$ and $A(t) \in L[\mathcal{Y}, \mathcal{H}]$ (i.e., $A(t)$ is bounded as a mapping from $\mathcal{Y} \to \mathcal{H}$), and the function $g(t) = \|A(t)\|_{\mathcal{Y}\to\mathcal{H}}$ is continuous.

(4) The space $\mathcal{Y}$ is an invariant subspace for each semigroup $S_t(\tau) = \exp\{\tau A(t)\}$ and $S_t(\tau)$ is a stable C_0-semigroup on $\mathcal{Y}$ with the same stability constants.

This case is not as easily analyzed as the parabolic case, so we need the following:

Lemma 8.5. *Suppose conditions (3) and (4) above are satisfied with* $\|\varphi\|_{\mathcal{H}} \leqslant \|\varphi\|_{\mathcal{Y}}$. *Then the family* $A(t)$, $t \in I$, *is strongly continuous on* $\mathcal{H}$ *(a.e.) for* $t \in I$.

Proof. Let $\varepsilon > 0$ be given and, without loss, assume that $\|\varphi\|_{\mathcal{H}} \leqslant 1$. Set $c = \|\varphi\|_{\mathcal{Y}}/\|\varphi\|_{\mathcal{H}}$, so that $1 \leqslant c < \infty$. Now

$$\begin{aligned}\|[A(t+h) - A(t)]\,\varphi\|_{\mathcal{H}} &\leqslant \left\{\|[A(t+h) - A(t)]\,\varphi\|_{\mathcal{H}}/\|\varphi\|_{\mathcal{Y}}\right\}\left[\|\varphi\|_{\mathcal{Y}}/\|\varphi\|_{\mathcal{H}}\right] \\ &\leqslant c\,\|A(t+h) - A(t)\|_{\mathcal{Y}\to\mathcal{H}}\,.\end{aligned}$$

Choose $\delta > 0$ such that $|h| < \delta$ implies $\|A(t+h) - A(t)\|_{\mathcal{Y}\to\mathcal{H}} < \varepsilon/c$, which completes the proof. □

Remark 8.6. The important point of this section is that once we know that $A(t)$ generates a semigroup for each t, the only other conditions required are that the family $\{A(t) : t \in I\}$ is weakly continuous and satisfies the growth condition (8.1). However, when the family $\{A(t) : t \in I\}$ is strongly continuous, the growth condition (8.1) is automatically satisfied.

8.4. Path Integrals I: Elementary Theory

Introduction

In this and the next section, we will obtain a general theory for path integrals in exactly the manner envisioned by Feynman. Our approach is distinct from the methods of functional integration, so we do not discuss that subject directly. However, since functional integration represents an important approach to path integrals, a few brief remarks are in order. The methods of functional differentiation and integration were major tools for the Schwinger program in quantum electrodynamics, which was developed in parallel with the Feynman theory (see [DY], in Chap. 7). Thus, these methods were not developed for the study of path integrals. However, historically, path integrals have been studied from the functional integration point of view, and many authors have sought to restrict consideration to the space of continuous functions or related function spaces in their definition of the path integral. The best known example is undoubtedly the Wiener integral [WSRM]. However, from the time-ordering point of view, such a restriction is not natural nor desirable. Thus, our approach does not depend on formulations with countably additive measures. In fact, we take the view that integration theory, as contrasted with measure theory, is the appropriate vehicle for path integrals. Indeed, as shown in [GZ1], there is a one-to-one mapping between path integrals and semigroups of operators that have a kernel representation. In this case, the semigroup operation generates the reproducing property of the kernel.

In their recent (2000) review of functional integration, Cartier and DeWitt-Morette [CDM1] discuss three of the most fruitful and important applications of functional integration to the construction of path integrals. In 1995, the *Journal of Mathematical Physics* devoted a special issue to this subject, Vol. 36, No. 5 (edited by Cartier and DeWitt-Morette). Thus, those with interest in the functional integration approach will find ample material in the above references (see also the book [CDM2]). Both the review and book are excellent on many levels, in addition to the historical information that could only come from one with first-hand information on the evolution of the subject.

8.4.1. Summary. In this section, we restrict our discussion to kernel representations for an interesting class of solutions to partial differential equations. In each case, a path integral representation is fairly straightforward.

We begin with the path integral as first introduced by Feynman [FY1]. The purpose is to show that the simplicity of his original approach becomes possible when the problem is considered on $KS^2[\mathbb{R}^3]$.

Recall that, in elementary quantum theory, the (simplest) problem to solve in $\mathbb{R}^3$ is:

$$i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} - \frac{\hbar^2}{2m}\Delta\psi(\mathbf{x},t) = 0,\ \ \psi(\mathbf{x},s) = \delta(\mathbf{x}-\mathbf{y}), \tag{8.2}$$

with solution

$$\psi(\mathbf{x},t) = K\left[\mathbf{x},\ \ t;\ \mathbf{y},\ \ s\right] = \left[\frac{2\pi i\hbar(t-s)}{m}\right]^{-3/2} \exp\left[\frac{im}{2\hbar}\frac{\left|\mathbf{x}-\mathbf{y}\right|^2}{(t-s)}\right].$$

In his formulation of quantum theory, Feynman wrote the solution to Eq. (8.2) as

$$K\left[\mathbf{x},\, t;\ \mathbf{y},\, s\right] = \int_{\mathbf{x}(s)=y}^{\mathbf{x}(t)=x} \mathcal{D}\mathbf{x}(\tau) \exp\left\{\tfrac{im}{2\hbar}\int_s^t \ \left|\tfrac{d\mathbf{x}}{dt}\right|^2 d\tau\right\}, \tag{8.3}$$

where

$$\begin{gathered}\int_{\mathbf{x}(s)=y}^{\mathbf{x}(t)=x} \mathcal{D}\mathbf{x}(\tau) \exp\left\{\tfrac{im}{2\hbar}\int_s^t \left|\tfrac{d\mathbf{x}}{dt}\right|^2 d\tau\right\} =: \\ \lim_{N\to\infty}\left[\tfrac{m}{2\pi i\hbar\varepsilon(N)}\right]^{3N/2}\int_{\mathbb{R}^3}\prod_{j=1}^{N} d\mathbf{x}_j \exp\left\{\tfrac{i}{\hbar}\sum_{j=1}^{N}\left[\tfrac{m}{2\varepsilon(N)}\left(\mathbf{x}_j-\mathbf{x}_{j-1}\right)^2\right]\right\},\end{gathered} \tag{8.4}$$

with $\varepsilon(N) = (t-s)/N$.

Equation (8.4) represents an attempt to define an integral on the space of continuous paths with values in $\mathbb{R}^3$ (i.e., $\mathbb{C}\left([s,\ t] :\ \mathbb{R}^3\right)$). This approach has a number of well-known mathematical problems:

- The kernel $K\left[\mathbf{x},\ \ t;\ \mathbf{y},\ \ s\right]$ and $\delta(\mathbf{x})$ are not in $L^2[\mathbb{R}^3]$, the standard space for quantum theory.
- The kernel $K\left[\mathbf{x},\ \ t;\ \mathbf{y},\ \ s\right]$ cannot be used to define a measure.

Notwithstanding these problems, the physics community has continued to make progress using this integral and have consistently obtained correct answers, which have been verified whenever independent (rigorous) methods were possible.

In response, the mathematics community has developed a large variety of indirect methods to justify the integral. The recent book by Johnson and Lapidus [JL] discusses many important contributions from the literature.

If we want to retain the approach used by Feynman, the problems identified above must be faced directly. Thus, the natural question is: Does there exist a separable Hilbert space containing $\mathbf{K}\left[\mathbf{x},\, t;\ \mathbf{y},\, s\right]$ and $\delta(\mathbf{x})$? A positive answer is required if the limit in Eq. (8.4) is to make sense. If we also want a space that allows us to include the Feynman, Heisenberg, and Schrödinger representations, we must require that the convolution and Fourier transform exist on the space as bounded linear operators. This requirement is necessary, since the convolution operator is needed for the path integral and the position and momentum operators, $\mathbf{x}, \mathbf{p}$, are canonically conjugate variables (i.e., Fourier transform pairs).

8.4.2. Background. The properties of $KS^2[\mathbb{R}^n]$ derived in Chap. 3 suggest that it is a perfect choice for the Feynman formulation. It is easy to see that both the position and momentum operators have closed, densely defined extensions to $KS^2[\mathbb{R}^n]$. A full theory requires that the Fourier transform, $\mathfrak{F}$, and the convolution operator $\mathfrak{C}$ (as bounded linear operators) have extensions $KS^2[\mathbb{R}^n]$ in order to ensure that both the Schrödinger and Heisenberg theories have faithful representations on $KS^2[\mathbb{R}^n]$. For this, we restate Theorem 5.15 as it applies to $KS^2[\mathbb{R}^n]$.

Theorem 8.7. *Let A be a bounded linear operator on a Banach space $\mathcal{B} \subset KS^2$. If $\mathcal{B}' \subset KS^2$, then A has a bounded extension to $L[KS^2]$, with $\|A\|_{KS^2} \leq k \|A\|_{\mathcal{B}}$ with k constant.*

We can now use Theorem 8.7 to prove that $\mathfrak{F}$ and $\mathfrak{C}$, the Fourier (transform) operator and the convolution operator respectively, defined on $L^1[\mathbb{R}^n]$, have bounded extensions to $KS^2[\mathbb{R}^n]$.

Theorem 8.8. *Both $\mathfrak{F}$ and $\mathfrak{C}$ extend to bounded linear operators on $KS^2[\mathbb{R}^n]$.*

Proof. To prove our result, first note that $C_0[\mathbb{R}^n]$, the bounded continuous functions on $\mathbf{R}^n$ which vanish at infinity, is contained in $KS^2[\mathbb{R}^n]$. Now $\mathfrak{F}$ is a bounded linear operator from $L^1[\mathbb{R}^n]$ to $C_0[\mathbb{R}^n]$, so we can consider it as a bounded linear operator from $L^1[\mathbb{R}^n]$ to $KS^2[\mathbb{R}^n]$. Since $\mathbf{L}^1[\mathbf{R}^n]$ is dense in $KS^2[\mathbb{R}^n]$ and $L^\infty[\mathbf{R}^n] \subset KS^2[\mathbb{R}^n]$, by Theorem 8.7, $\mathfrak{F}$ extends to a bounded linear operator on $KS^2[\mathbb{R}^n]$.

To prove that $\mathfrak{C}$ has a bounded extension, fix g in $L^1[\mathbb{R}^n]$ and define $\mathfrak{C}_g$ on $L^1[\mathbb{R}^n]$ by:

$$\mathfrak{C}_g(f)(\mathbf{x}) = \int g(\mathbf{y}) f(\mathbf{x} - \mathbf{y}) d\mathbf{y}.$$

Once again, since $\mathfrak{C}_g$ is bounded on $L^1[\mathbb{R}^n]$ and $L^1[\mathbb{R}^n]$ is dense in $KS^2[\mathbb{R}^n]$, by Theorem 8.7 it extends to a bounded linear operator on $KS^2[\mathbb{R}^n]$. Now use the fact that convolution is commutative to get that $\mathfrak{C}_f$ is a bounded linear operator on $L^1[\mathbb{R}^n]$ for all $f \in KS^2[\mathbb{R}^n]$. Another application of Theorem 8.7 completes the proof. □

We now return to $\mathfrak{M}[\mathbb{R}^n]$, the space of bounded finitely additive measures on $\mathbb{R}^n$, that are absolutely continuous with respect to Lebesgue measure.

Definition 8.9. A uniformly bounded sequence $\{\mu_k\} \subset \mathfrak{M}[\mathbb{R}^n]$ is said to converge weakly to μ ($\mu_n \xrightarrow{w} \mu$), if, for every bounded uniformly continuous function $h(\mathbf{x})$,

$$\int_{\mathbb{R}^n} h(\mathbf{x}) d\mu_n \to \int_{\mathbb{R}^n} h(\mathbf{x}) d\mu.$$

Theorem 8.10. *If $\mu_n \xrightarrow{w} \mu$ in $\mathfrak{M}[\mathbb{R}^n]$, then $\mu_n \xrightarrow{s} \mu$ (strongly) in $KS^p[\mathbb{R}^n]$.*

Proof. Since the characteristic function of a closed cube is a bounded uniformly continuous function, $\mu_n \xrightarrow{w} \mu$ in $\mathfrak{M}[\mathbb{R}^n]$ implies that

$$\int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) d\mu_n \to \int_{\mathbb{R}^n} \mathcal{E}_m(\mathbf{x}) d\mu$$

for each m, so that $\lim_{n\to\infty} \|\mu_n - \mu\|_{KS^p} = 0$. □

A little reflection gives:

Theorem 8.11. *The space $KS^2[\mathbb{R}^n]$ is a commutative Banach algebra with unit.*

Since $KS^2[\mathbb{R}^n]$ contains the space of measures, it follows that all the approximating sequences for the Dirac measure converge strongly to it in the $KS^2[\mathbb{R}^n]$ topology. For example, $[\sin(\lambda \cdot \mathbf{x})/(\lambda \cdot \mathbf{x})] \in KS^2[\mathbb{R}^n]$ and converges strongly to $\delta(\mathbf{x})$. On the other hand, the function $e^{-2\pi i \mathbf{z}(\mathbf{x}-\mathbf{y})} \in KS^2[\mathbb{R}^n]$, so we can define the delta function directly:

$$\delta(\mathbf{x} - \mathbf{y}) = \int_{\mathbb{R}^n} e^{-2\pi i \mathbf{z}(\mathbf{x}-\mathbf{y})} d\lambda_n(\mathbf{z}),$$

as an HK-integral.

It is easy to see that the Feynman kernel [FH], defined by (with $m = 1$ and $\hbar = 1$):

$$\mathbb{K}_{\mathbf{f}}[t, \mathbf{x}\,;\, s, B] = \int_B (2\pi i(t-s))^{-n/2} \exp\{i|\mathbf{x} - \mathbf{y}|^2 \big/ 2(t-s)\} d\mathbf{y},$$

is in $KS^2[\mathbb{R}^n]$ and $\|\mathbb{K}_{\mathbf{f}}[t, \mathbf{x}\,;\, s, B]\|_{\mathrm{KS}} \leqslant 1$, while $\|\mathbb{K}_{\mathbf{f}}[t, \mathbf{x}\,;\, s, B]\|_{\mathfrak{M}} = \infty$ (the variation norm). Furthermore,

$$\mathbb{K}_{\mathbf{f}}[t, \mathbf{x}\,;\, s, B] = \int_{\mathbb{R}^n} \mathbb{K}_{\mathbf{f}}[t, \mathbf{x}\,;\, \tau, d\mathbf{z}] \mathbb{K}_{\mathbf{f}}[\tau, \mathbf{z}\,;\, s, B], \text{ (HK-integral)}.$$

Remark 8.12. It is not hard to show that $\mathbb{K}_{\mathbf{f}}[t, \mathbf{x}\,;\, s, B]$ generates a finitely additive set function defined on the algebra of sets B, such that $\mathcal{E}_B(|\mathbf{y}|)$ is of bounded variation in the $|\mathbf{y}|$ variable.

Definition 8.13. Let $\mathbf{P}_k = \{t_0, \tau_1, t_1, \tau_2, \cdots, \tau_k, t_k\}$ be a HK-partition of the interval $[0, t]$ for each k, with $\lim_{k\to\infty} \mu_k = 0$ (mesh). Set $\Delta t_j = t_j - t_{j-1}, \tau_0 = 0$ and, for $\psi \in KS^2[\mathbb{R}^n]$, define

$$\int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{0})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}_\lambda \mathbf{x}(\tau)] = e^{-\lambda t} \sum_{k=0}^{[\![\lambda t]\!]} \frac{(\lambda t)^k}{k!} \left\{ \prod_{j=1}^{k} \int_{\mathbb{R}^n} \mathbb{K}_{\mathbf{f}}[t_j, \mathbf{x}(\tau_j)\,;\, t_{j-1}, d\mathbf{x}(\tau_{j-1})] \right\},$$

and

$$\begin{aligned}\int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{0})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau)]\psi[\mathbf{x}(0)] &= \lim_{\lambda\to\infty} \int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{0})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}_\lambda \mathbf{x}(\tau)]\psi[\mathbf{x}(0)] \\ &= \lim_{\lambda\to\infty} e^{-\lambda t} \sum_{k=0}^{[\![\lambda t]\!]} \frac{(\lambda t)^k}{k!} \left\{ \prod_{j=1}^{k} \int_{\mathbb{R}^n} \mathbb{K}_{\mathbf{f}}[t_j, \mathbf{x}(\tau_j)\,;\, t_{j-1}, d\mathbf{x}(\tau_{j-1})]\psi[\mathbf{x}(0)] \right\},\end{aligned} \tag{8.5}$$

whenever the limit exists.

It is easy to see that the limit exists in $KS^2[\mathbb{R}^n]$, whenever we have a reproducing kernel.

Theorem 8.14. *The function* $\psi(\mathbf{x}) \equiv 1 \in KS^2[\mathbb{R}^n]$ *and*

$$\int_{\mathbf{x}(\tau)=\mathbf{y}(\mathbf{s})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau)]\psi[\mathbf{y}(s)] = \mathbb{K}_{\mathbf{f}}[t, \mathbf{x}\,;\, s, \mathbf{y}] = \tfrac{1}{\sqrt{[2\pi i(t-s)]^n}} \exp\{i|\mathbf{x}-\mathbf{y}|^2/2(t-s)\}.$$

The above result is what Feynman was trying to obtain (in this simple case).

8.5. Examples and Extensions

In this section, we provide a few interesting examples. Those with broader interest should consult the references below.

Independent of the mathematical theory, the practical development and use of path integral methods has proceeded at a continuous rate. At this time, it would be impossible to give a survey of the many different types of path integrals and the problems that they have been used to solve. It would be a separate task to provide a reasonable set of references on the subject. However, the following books are suggested for both the material they cover and the references contained in them: Albeverio and H∅egh-Krohn [AH], Cartier and Dewitt-Morette [CDM2], Feynman and Hibbs [FH], Grosche and Steiner [GS], and Kleniert [KL].

8.5.1. The Diffusion Problem. For our first example, let $\mathcal{H} = L^2[\mathbb{R}^3, d\mu]$, where $d\mu = e^{-\pi|\mathbf{x}|^2} d\lambda_3(\mathbf{x})$. The form is nonstandard, but has advantages as discussed in Chap. 2. Consider the problem:

$$\frac{\partial}{\partial t}u(t,\mathbf{x}) = \Delta u(t,\mathbf{x}) - \mathbf{x}\cdot\nabla\mathbf{u}(\mathbf{x},t), \quad u(0,\mathbf{x}) = u_0(\mathbf{x}).$$

This is the Ornstein–Uhlenbeck equation, with solution $(T(t)u_0)(\mathbf{x}) = u(t,\mathbf{x})$, where:

$$(T(t)\mathbf{u}_0)(\mathbf{x}) = \frac{1}{\sqrt{[(1-e^{-t})]^3}}\int_{\mathbb{R}^3} exp\left\{-\pi\frac{\left(e^{-t/2}\mathbf{x}-\mathbf{y}\right)^2}{(1-e^{-t})}\right\}\mathbf{u}_0(\mathbf{y})d\lambda_3(\mathbf{y}).$$

The operator $T(t)$ is a (analytic) contraction semigroup, with generator $D^2 = \Delta - \mathbf{x}\cdot\nabla$. It follows that the kernel is given by

$$\mathbb{K}_{\mathbf{f}}\left[t,\mathbf{x}\,;\,0,d\mathbf{y}\right] = \frac{1}{\sqrt{[(1-e^{-t})]^3}}\exp\left\{-\pi\frac{\left(e^{-t/2}\mathbf{x}-\mathbf{y}\right)^2}{(1-e^{-t})}\right\}d\lambda_3(\mathbf{y}).$$

By Theorem 8.7 $T(t)$ can be extended to $KS^2(\mathbb{R}^3)$, as a C_0-contraction semigroup. It now follows that

$$u(t,\mathbf{x}) = \int_{\mathbf{x}(\tau)=\mathbf{y}(\mathbf{0})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}\left[\mathcal{D}\mathbf{x}(\tau)\right]u[\mathbf{y}(\mathbf{0})].$$

For a more interesting example, let B be a nondegenerate 3×3 matrix with eigenvalues λ such that $Re(\lambda) < 0$, with Q strictly positive definite and set

$$Q_t = \int_0^t e^{sB}Qe^{sB^*}ds.$$

In this case, $\mathcal{H} = L^2[\mathbb{R}^3, d\mu]$, with

$$\mu(d\mathbf{x}) = \frac{1}{\sqrt{(\det Q_\infty}}\exp\left\{-\pi\left\langle Q_\infty^{-1}\mathbf{x},\mathbf{x}\right\rangle\right\}d\lambda_3(\mathbf{x}),$$

and we consider the problem:

$$\frac{\partial}{\partial t}u(t,\mathbf{x}) = \Delta u(t,\mathbf{x}) - B\mathbf{x}\cdot\nabla\mathbf{u}(\mathbf{x},t), \quad u(0,\mathbf{x}) = u_0(\mathbf{x}).$$

This is also a version of the Ornstein–Uhlenbeck equation. (However, since we don't assume that B is symmetric, $A = \Delta - B\mathbf{x}\cdot\nabla$ need not be self-adjoint.) The explicit solution is generated by the contraction semigroup $(T(t)$, where:

$$(T(t)u_0(\mathbf{x}) = \frac{1}{\sqrt{\det Q_t}}\int_{\mathbb{R}^3}\exp\left\{-\pi\left\langle Q_t^{-1}\left(e^{tB}\mathbf{x}-\mathbf{y}\right), e^{tB}\mathbf{x}-\mathbf{y}\right\rangle\right\}u_0(\mathbf{y})d\lambda_3(\mathbf{y}).$$

It follows that

$$\mathbb{K}_{\mathbf{f}}\left[t, \mathbf{x}\,; 0, d\mathbf{y}\right] = \frac{1}{\sqrt{\det Q_t}} \exp\left\{-\pi\left\langle Q_t^{-1}\left(e^{tB}\mathbf{x}-\mathbf{y}\right), e^{tB}\mathbf{x}-\mathbf{y}\right\rangle\right\} d\lambda_3(\mathbf{y}).$$

For this equation, we can also replace $\mathbb{R}^3$ by a separable Hilbert space $\mathcal{H}$ and λ_3 by cylindrical Gaussian measure μ. In this case, B is a symmetric bounded linear operator with spectrum $\sigma(B) < 0$ and $0 < Q_\infty < \infty$ is strictly positive definite. Those with interest in this subject can consult Lorenzi and Bertoldi [LB] for the finite-dimensional case and De Prato [DP] for Hilbert space. In either case, the path integral representation is defined on $KS^2[\mathbb{R}^3]$ or $KS^2[\mathcal{H}]$.

8.5.2. Wave Equation. For this example, write the standard wave equation as

$$\frac{\partial^2\psi}{\partial t^2} - \mathbf{I}c^2\Delta\psi = \frac{1}{\hbar^2}\left[i\hbar\frac{\partial}{\partial t} + \beta\sqrt{-c^2\hbar^2\Delta}\right]\left[-i\hbar\frac{\partial}{\partial t} + \beta\sqrt{-c^2\hbar^2\Delta}\right]\psi = 0.$$

In electromagnetic theory, we only see the wave equation on the left and assume that $\mathbf{I} = 1$. On the right, the β matrix can be of any finite order. Thus, the above equation introduces a rather interesting relationship between quantum theory and the classical wave equation, namely the massless square root equation for any spin. In order to solve this equation, we follow Lieb and Loss [LL], in Chap. 3, and use imaginary time to get:

$$\psi(\mathbf{x}, t) = \frac{it\beta}{\pi^2}\lim_{\varepsilon\to 0}\int_{\mathbb{R}^3}\frac{\psi_0(\mathbf{y})d\mathbf{y}}{\left[\left|\mathbf{x}-\mathbf{y}\right|^2 - t^2\right]^2 + \varepsilon^2} = U(t)\psi_0(\mathbf{x}), \tag{8.6}$$

where $\psi_0(\mathbf{x})$ is the given initial data at time $t = 0$. The convergence factor is necessary for the integral representation because of the light cone problem (in the Lebesgue sense). This is not necessary if we interpret it in the Henstock–Kurzweil sense. We could also compute the solution directly by using the fact that the square root operator is a self-adjoint generator of a unitary group. However, extra work would still be required to obtain the integral representation.

We can now use (8.6) to provide a new representation for the solution of the wave equation. Assume that $\psi(\mathbf{x}, t)\,|_{t=0} = \psi_0(\mathbf{x})$ and $\dot{\psi}(\mathbf{x}, t)\,|_{t=0} = \dot{\psi}_0(\mathbf{x})$ are given (smooth) initial data. Let $A = \beta\sqrt{-c^2\hbar^2\Delta}$ and $\varphi(\mathbf{x}, t) = \left(-i\hbar\partial_t + A\right)\psi(\mathbf{x}, t)$. It follows from this that

$$\varphi(\mathbf{x}, 0) = \varphi_0 = i\hbar\dot{\psi}_0(\mathbf{x}) + A\psi_0(\mathbf{x}).$$

We must now solve:

$$(i\hbar\partial_t + A)\,\varphi(\mathbf{x},t) = 0,\ \ \varphi(\mathbf{x},0) = \varphi_0.$$

The solution to this problem is easily seen to be (8.6), with t replaced by $-t$, so that $\varphi(\mathbf{x},t) = U(-t)\varphi_0$. Using this result, we can now get our new representation. The solution to the wave equation has been reduced to solving:

$$(-i\hbar\partial_t + A)\,\psi(\mathbf{x},t) = U(-t)\varphi_0.$$

Using the method of variation of constants, we have: (see Sell and You [SY], p. 7).

$$\psi(\mathbf{x},t) = U(t)\psi_0 + \int_0^t U(t-s)U(-s)\varphi_0(\mathbf{x})ds.$$

Combining terms, we have:

$$\psi(\mathbf{x},t) = U(t)\psi_0 + \int_0^t U(t-2s)\varphi_0(\mathbf{x})ds. \tag{8.7}$$

It is now easy to check that $\psi(\mathbf{x},0) = \psi_0(\mathbf{x})$ and that $\dot{\psi}(\mathbf{x},0) = \dot{\psi}_0(\mathbf{x})$. We can now use Eq. (8.6) to obtain the explicit representation for a general solution to the wave equation:

$$\begin{aligned}\psi(\mathbf{x},t) &= -\frac{it\beta}{\pi^2}\lim_{\varepsilon\to 0}\int_{\mathbb{R}^3}\frac{\psi_0(\mathbf{y})d\mathbf{y}}{\left[|\mathbf{x}-\mathbf{y}|^2 - t^2\right]^2 + \varepsilon^2}\\ &+\int_0^t\left\{\frac{i(t-2s)\beta}{\pi^2}\lim_{\varepsilon\to 0}\int_{\mathbb{R}^3}\frac{\varphi_0(\mathbf{y})d\mathbf{y}}{\left[|\mathbf{x}-\mathbf{y}|^2 - (t-2s)^2\right]^2 + \varepsilon^2}\right\}ds.\end{aligned} \tag{8.8}$$

We have only worked in $\mathbb{R}^3$. For n arbitrary, the only change (other than initial data) is the kernel. In the general case, we must replace Eq. (8.6) by

$$\psi(\mathbf{x},t) = \frac{it\beta\Gamma\left[\frac{n+1}{2}\right]}{\pi^{(n+1)/2}}\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\frac{\psi_0(\mathbf{y})d\mathbf{y}}{\left[|\mathbf{x}-\mathbf{y}|^2 - t^2\right]^{\frac{(n+1)}{2}} + \varepsilon^2}.$$

Thus, the method is quite general. Recall that the standard approach is based on the method of spherical means (see Evans [EV]). This approach requires different computations depending on the dimension (even or odd). It follows that our approach has some advantages. The path integral representation is straightforward.

8.5.3. The Square-Root Klein–Gordon Equation. The fourth example is taken from [GZ4] and provides another example that is not directly related to a Gaussian kernel. It is shown that if the vector potential $\mathbf{A}$ is constant, $\mu = mc/\hbar$, and $\boldsymbol{\beta}$ is the standard beta matrix, $(I, O : O, -I)$, then the solution to the equation for a spin 1/2 particle in square-root form,

$$i\hbar\partial\psi(\mathbf{x},t)/\partial t = \left\{\boldsymbol{\beta}\sqrt{c^2\left(\mathbf{p}-\tfrac{e}{c}\mathbf{A}\right)^2 + m^2c^4}\right\}\psi(\mathbf{x},t),\ \ \psi(\mathbf{x},0)=\psi_0(\mathbf{x}),$$

is given by:

$$\psi(\mathbf{x},t) = \mathbf{U}[t,0]\psi_0(\mathbf{x}) = \int\limits_{\mathbf{R}^3} \exp\left\{\frac{ie}{2\hbar c}(\mathbf{x}-\mathbf{y})\cdot\mathbf{A}\right\}\mathbb{K}_{\mathbf{f}}\left[\mathbf{x},t\,;\,\mathbf{y},0\right]\psi_0(\mathbf{y})d\mathbf{y},$$

where

$$\mathbb{K}_{\mathbf{f}}\left[\mathbf{x},t\,;\,\mathbf{y},0\right] = \frac{ct\mu^2\beta}{4\pi}\begin{cases} \dfrac{-H_2^{(1)}\left[\mu\left(c^2t^2-\|\mathbf{x}-\mathbf{y}\|^2\right)^{1/2}\right]}{\left[c^2t^2-\|\mathbf{x}-\mathbf{y}\|^2\right]}, & ct < -\left\|\mathbf{x}-\mathbf{y}\right\|, \\ \dfrac{-2iK_2\left[\mu\left(\|\mathbf{x}-\mathbf{y}\|^2-c^2t^2\right)^{1/2}\right]}{\pi\left[\|\mathbf{x}-\mathbf{y}\|^2-c^2t^2\right]}, & c\left|t\right| < \left\|\mathbf{x}-\mathbf{y}\right\|, \\ \dfrac{H_2^{(2)}\left[\mu\left(c^2t^2-\|\mathbf{x}-\mathbf{y}\|^2\right)^{1/2}\right]}{\left[c^2t^2-\|\mathbf{x}-\mathbf{y}\|^2\right]}, & ct > \left\|\mathbf{x}-\mathbf{y}\right\|. \end{cases}$$

The function $K_2(\cdot)$ is a modified Bessel function of the third kind of second order, while $H_2^{(1)}$, $H_2^{(2)}$ are the Hankel functions (see Gradshteyn and Ryzhik [GRRZ]). Thus, we have a kernel that is far from standard. To our knowledge, this representation is new.

8.5.4. Semigroups, Kernels, and Pseudodifferential Operators. In this section, we investigate the general question of the existence of relations of the form:

$$U(t)\phi(\mathbf{x}) = \int_{\mathbb{R}^{\mathbf{n}}} \mathbf{K}(\mathbf{x},\mathbf{t}:\mathbf{y},\mathbf{0})\phi_{\mathbf{0}}(\mathbf{y})\mathbf{dy}, \tag{8.9}$$

between a semigroup of operators $U(t)$, $t \in \mathbb{R}$ and a kernel K. We observe that if a kernel exists, then the semigroup property automatically induces the reproducing property of the kernel and vice versa. Equation (8.9) also leads to a discussion of the close relationship between kernels and the theory of pseudodifferential operators. In this section we show how to associate a reproducing kernel with a large class of semigroups $U(t)$. A more detail discussion of pseudodifferential operators can be found in Treves [TR], Kumano-go [KG], Taylor [TA], Cordes [CO], or Shubin [SHB].

Pseudodifferential operators are a natural extension of linear partial differential operators and interest in them grew out of the study of singular integral operators like the one induce by the square-root operator. The basic idea is that the use of pseudodifferential operators allows one to convert the theory of partial differential equations into an algebraic theory for the characteristic polynomials, or symbols, of the differential operators by means of Fourier transforms.

We begin our study with the definition of hypoelliptic pseudodifferential operators of class $S^m_{\rho,\delta}$ and investigate their basic properties. As noted above, we confine our discussion to Euclidean spaces, $\mathbb{R}^n$, and only consider those parts that pertain to the construction of kernel representations. (Readers interested in more general treatments can consult the cited references.)

Definition 8.15. Recall that a complex-valued function f defined on $\mathbb{R}^n$ is a Schwartz function ($f \in S(\mathbb{R}^n)$ or S) if, for all multi-indices α and β, there exist positive constants $C_{\alpha,\beta}$ such that

$$\sup_{\mathbf{x}\in\mathbb{R}^n} \left| \mathbf{x}^\alpha \partial^\beta f(\mathbf{x}) \right| = C_{\alpha,\beta} < \infty.$$

In what follows, $\mathbb{R}^n_{\mathbf{x}}$ denotes n-dimensional space in the $\mathbf{x}$ variable. For continuity with the literature, we keep the standard notation, where one works on the tangent space of a differential manifold.

Definition 8.16. Let $p(\mathbf{x},\eta)$ be a C^∞ function on $\mathbb{R}^n_{\mathbf{x}} \times \mathbb{R}^n_\eta$.

(1) We say that $p(\mathbf{x},\eta)$ is a symbol of class $S^m_{\rho,\delta}$ ($n \in \mathbb{N},\ 0 \le \delta \le \rho \le 1,\ \delta < 1$) if, for any multi-indices $\alpha,\ \beta$, there exists a constant $C_{\alpha,\beta}$ such that

$$\left| p^{(\alpha)}_{(\beta)}(\mathbf{x},\eta) \right| = C_{\alpha,\beta} \left\langle \eta \right\rangle^{m+\delta|\beta|-\rho|\alpha|},$$

where

$$p^{(\alpha)}_{(\beta)}(\mathbf{x},\eta) = \partial^\alpha_\eta D^\beta_{\mathbf{x}} p(\mathbf{x},\eta),\ \ \langle \eta \rangle = \sqrt{1+\|\eta\|^2},\ \ |\alpha| = \sum\nolimits_{i=1}^n \alpha_i,$$

$$\partial^\alpha_\eta = \partial^{\alpha_1}_{\eta_1}\cdots\partial^{\alpha_n}_{\eta_n},\ \ D^\beta_{\mathbf{x}} = D^{\beta_1}_{x_1}\cdots D^{\beta_n}_{x_n} \text{ and } D_{x_j} = -i\frac{\partial}{\partial x_j}.$$

Also, we set

$$S^{-\infty}_{\rho,\delta} = \bigcap_{m=1}^{\infty} S^m_{\rho,\delta} \text{ and } S^{\infty}_{\rho,\delta} = \bigcup_{m=1}^{\infty} S^m_{\rho,\delta}.$$

(2) A linear operator $P : S(\mathbb{R}^n_{\mathbf{x}}) \to S(\mathbb{R}^n_{\mathbf{x}})$ is said to be a pseudodifferential operator with symbol $p(\mathbf{x}, \eta)$ of class $S^m_{\rho,\delta}$ if, for $u \in S(\mathbb{R}^n_{\mathbf{x}})$, we can write $Pu(\mathbf{x})$ as

$$Pu(\mathbf{x}) = \int_{\mathbb{R}^n} e^{\pi i \mathbf{x}\cdot\eta} p(\mathbf{x}, \eta)\hat{u}(\eta)d\eta,$$

where

$$\hat{u}(\eta) = \mathfrak{F}(u)(\eta) = \int_{\mathbb{R}^n} e^{-\pi i \mathbf{x}\cdot\eta} u(\mathbf{x})d\mathbf{x}$$

is the Fourier transform of $u(\mathbf{x})$.

Whenever $m \leq m'$, $\rho' \leq \rho$, $\delta \leq \delta'$, for any ρ and δ, we have $S^m (= S^m_{1,0}) \subset S^m_{\rho,\delta} \subset S^{m'}_{\rho',\delta'}$. It follows that

$$\bigcap_{m=1}^{\infty} S^m_{\rho,\delta} = \bigcap_{m=1}^{\infty} S^m_{1,0},$$

so that $S^{-\infty} = \bigcap_{m=1}^{\infty} S^m_{\rho,\delta}$. For $p(\mathbf{x}, \eta) \in S^m_{\rho,\delta}$ we define the family of seminorms $|p|_l^{(m)}$, $l = 0, 1, \dots$ by

$$|p|_l^{(m)} = \max_{|\alpha+\beta|=l} \sup_{\mathbb{R}^n_{\mathbf{x}}\times\mathbb{R}^n_{\eta}} \left\{ \left| p^{(\alpha)}_{(\beta)}(\mathbf{x}, \eta) \right| \langle\eta\rangle^{(m+\delta|\beta|-\rho|\alpha|)} \right\}.$$

Then $S^m_{\rho,\delta}$ is a Fréchet space with these seminorms, and we have, for any $p(\mathbf{x}, \eta) \in S^m_{\rho,\delta}$:

$$\left| p^{(\alpha)}_{(\beta)}(\mathbf{x}, \eta) \right| \leqslant |p|^{(m)}_{|\alpha+\beta|} \langle\eta\rangle^{(m+\delta|\beta|-\rho|\alpha|)}.$$

We say that a set $B \subset S^m_{\rho,\delta}$ is bounded in $S^m_{\rho,\delta}$ if $\sup_{p\in B} \left\{ |p|_l^{(m)} \right\} < \infty$.

For $p(\mathbf{x}, \eta) \in S^m_{\rho,\delta}$ we can represent $Pu(\mathbf{x})$, $u \in S(\mathbb{R}^n)$, in terms of oscillatory integrals. These are integrals of the form:

$$Af(\mathbf{x}) = \int_{\mathbb{R}^n} e^{\pi i s(\mathbf{x},\eta)} a(\mathbf{x}, \eta)\hat{f}(\eta)d\eta,$$

where $s(\mathbf{x}, \eta)$ is called the phase function and $a(\mathbf{x}, \eta)$ is called the amplitude function. These functions were first introduced by Lax [LX1] and used to construct asymptotic solutions of hyperbolic differential equations. (In the hands of Hörmander [HO], this later led to the (related) theory of Fourier integral operators.)

We are interested in a restricted class of these integrals.

Definition 8.17. We say that a C^∞ function $a(\zeta, \mathbf{y})$, $(\zeta, \mathbf{y}) \in \mathbb{R}^{2n}_{\zeta,\mathbf{y}} = \mathbb{R}^n_\zeta \times \mathbb{R}^n_\mathbf{y}$ belongs to the class $\mathfrak{A}^m_{\delta,\tau}$ $(m \in \mathbb{N},\ 0 \le \delta < 1,\ 0 \le \tau)$ if, for any multi-indices α, β, there exists a positive constant $C_{\alpha,\beta}$ such that

$$\left|\partial^\alpha_\zeta \partial^\beta_\mathbf{y} a(\zeta, \mathbf{y})\right| \leqslant C_{\alpha,\beta} \langle\zeta\rangle^{(m+\delta|\beta|)} \langle\mathbf{y}\rangle^\tau, \quad l = |\alpha + \beta|.$$

We set

$$\mathfrak{A} = \bigcup_{0\leqslant\delta<1} \bigcup_{m=-\infty}^{\infty} \bigcup_{0\leqslant\tau} \mathfrak{A}^m_{\delta,\tau}.$$

Then we have

Theorem 8.18. *For $a(\zeta, \mathbf{y}) \in \mathfrak{A}^m_{\delta,\tau}$, we define the seminorms $|a|_l$, $l = 0, 1, \dots,$ by*

$$|a|_l = \max_{|\alpha+\beta|\leqslant l} \sup_{\zeta,\mathbf{y}} \left\{ \left|\partial^\alpha_\zeta \partial^\beta_\mathbf{y} a(\zeta, \mathbf{y})\right| \langle\zeta\rangle^{-(m+\delta|\beta|)} \langle\mathbf{y}\rangle^\tau \right\}.$$

(1) *Then $\mathfrak{A}^m_{\delta,\tau}$ is a Frechet space and for $a(\zeta, \mathbf{y}) \in \mathfrak{A}^m_{\delta,\tau}$ we have*

$$\left|\partial^\alpha_\zeta \partial^\beta_\mathbf{y} a(\zeta, \mathbf{y})\right| \leqslant |a|_l \langle\zeta\rangle^{(m+\delta|\beta|)} \langle\mathbf{y}\rangle^\tau, \quad l = |\alpha + \beta|.$$

(2) *If $a, a_1, a_2 \in \mathfrak{A}$, then $\partial^\alpha_\eta \partial^\beta_\mathbf{y} a,\ a_1 + a_2,\ a_1 a_2 \in \mathfrak{A}$.*

Definition 8.19. We say that $B \subset \mathfrak{A}$ is a bounded subset of $\mathfrak{A}$ if there exists $\mathfrak{A}^m_{\delta,\tau}$ such that

$$B \subset \mathfrak{A}^m_{\delta,\tau} \quad \text{and} \quad \sup_{a\in B} \{|a|_l\} < \infty$$

for $l \in \{0\} \cup \mathbb{N}$.

Definition 8.20. For $a(\zeta, \mathbf{y}) \in \mathfrak{A}$, we define the oscillatory integral $O_s(e^{-\pi i \mathbf{y}\dot{\eta}} a) =: O_s$ by

$$O_s = \lim_{\varepsilon\to 0} \iint_{\mathbb{R}^{2n}} e^{-\pi i \mathbf{y}\cdot\eta} \chi(\varepsilon\eta, \varepsilon\mathbf{y}) a(\eta, \mathbf{y})\, d\mathbf{y}\, d\eta,$$

where $\chi(\eta, \mathbf{y}) \in S(\mathbb{R}^{2n}_{\eta,\mathbf{y}})$ and $\chi(0, 0) = 1$.

It is shown in Kumano-go ([KG], p. 47) that O_s is well defined and independent of the choice of $\chi(\eta, \mathbf{y}) \in S(\mathbb{R}^{2n}_{\eta,\mathbf{y}})$ satisfying $\chi(0, 0) = 1$. We note that when $a(\eta, \mathbf{y}) \in L^1(\mathbb{R}^{2n}_{\eta,\mathbf{y}})$, the Lebesgue dominated convergence theorem shows that O_s coincides with the Lebesgue integral $\iint e^{-\pi i \mathbf{y}\cdot\eta} a(\eta, \mathbf{y}) d\mathbf{y} d\eta$.

A fundamental question is: under what general conditions can we expect a given (time-independent) generator of a semigroup to have an associated kernel? Here, we discuss a class of general conditions for unitary groups. It will be clear that the results of this section carry over to semigroups with minor changes.

Let $A(\mathbf{x},\ \mathbf{p})$ denote a $k \times k$ matrix operator $[A_{ij}(\mathbf{x},\ \mathbf{p})]$, $i,\ j = 1, 2, \cdots, k$, whose components are pseudodifferential operators with symbols $a_{ij}(\mathbf{x},\ \boldsymbol{\eta}) \in \mathbb{C}^{\infty}(\mathbf{R^n} \times \mathbf{R^n})$, and we have, for any multi-indices α and β,

$$\left|a^{(\alpha)}_{ij(\beta)}(\mathbf{x}, \boldsymbol{\eta})\right| \leqslant C_{\alpha\beta}(1 + |\boldsymbol{\eta}|)^{m-\xi|\alpha|+\delta|\beta|}, \tag{8.10}$$

where

$$a^{(\alpha)}_{ij(\beta)}(\mathbf{x}, \boldsymbol{\eta}) = \partial^{\alpha}\mathbf{p}^{\beta}a_{ij}(\mathbf{x}, \boldsymbol{\eta}),$$

with $\partial_l = \partial/\partial\eta_l$, and $p_l = (1/i)(\partial/\partial x_l)$. The multi-indices are defined in the usual manner by $\alpha = (\alpha_1, \cdots, \alpha_n)$ for integers $\alpha_j \geq 0$, and $|\alpha| = \sum_{j=1}^{n} \alpha_j$, with similar definitions for β. The notation for derivatives is $\partial^{\alpha} = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$ and $\mathbf{p}^{\beta} = p_1^{\beta_1} \cdots p_n^{\beta_n}$. Here, m, β, and δ are real numbers satisfying $0 \leq \delta < \xi$. Equation (8.10) states that each $a_{ij}(\mathbf{x},\ \boldsymbol{\eta})$ belongs to the symbol class $S^m_{\xi,\delta}$ (see [SH]).

Let $a(\mathbf{x}, \boldsymbol{\eta}) = [a_{ij}(\mathbf{x}, \boldsymbol{\eta})]$ be the matrix-valued symbol for $A(\mathbf{x}, \boldsymbol{\eta})$, and let $\lambda_1(\mathbf{x}, \boldsymbol{\eta}) \cdots \lambda_k(\mathbf{x}, \boldsymbol{\eta})$ be its eigenvalues. If $|\cdot|$ is the norm in the space of $k \times k$ matrices, we assume that the following conditions are satisfied by $a(\mathbf{x}, \boldsymbol{\eta})$. For $0 < c_0 < |\,\boldsymbol{\eta}|$ and $\mathbf{x} \in \mathbb{R}^n$ we have

(1) $\left|a^{(\alpha)}_{(\beta)}(\mathbf{x}, \boldsymbol{\eta})\right| \leq C_{\alpha\beta}\,|a(\mathbf{x}, \boldsymbol{\eta})|\,(1 + |\boldsymbol{\eta}|)^{-\xi|\alpha|+\delta|\beta|}$ (hypoellipticity),

(2) $\lambda_0(\mathbf{x}, \boldsymbol{\eta}) = \max\limits_{1 \leqslant j \leqslant k} \operatorname{Re} \lambda_j(\mathbf{x}, \boldsymbol{\eta}) < 0$,

(3) $\frac{|a(\mathbf{x}, \boldsymbol{\eta})|}{|\lambda_0(\mathbf{x}, \boldsymbol{\eta})|} = O\left[(1 + |\boldsymbol{\eta}|)^{(\xi-\delta)/(2k-\varepsilon)}\right],\ \varepsilon > 0.$

We assume that $A(\mathbf{x}, \mathbf{p})$ is a self-adjoint generator of an unitary group $U(t, 0)$, so that

$$U(t, 0)\psi_0(\mathbf{x}) = exp[(i/\hbar)tA(\mathbf{x}, \mathbf{p})]\psi_0(\mathbf{x}) = \psi(\mathbf{x}, t)$$

solves the Cauchy problem

$$(i/\hbar)\partial\psi(\mathbf{x}, t)/\partial t = A(\mathbf{x}, \mathbf{p})\psi(\mathbf{x}, t), \quad \psi(\mathbf{x}, t) = \psi_0(\mathbf{x}). \tag{8.11}$$

Definition 8.21. We say that $Q(\mathbf{x}, t, \boldsymbol{\eta}, 0)$ is a symbol for the Cauchy problem (8.11) if $\psi(\mathbf{x}, t)$ has a representation of the form

$$\psi(\mathbf{x}, t) = \int_{\mathbb{R}^n} e^{\pi i(\mathbf{x}, \boldsymbol{\eta})} Q(\mathbf{x}, t, \boldsymbol{\eta}, 0)\hat{\psi}_0(\boldsymbol{\eta})d\boldsymbol{\eta}. \tag{8.12}$$

It is sufficient that ψ_0 belongs to the Schwartz space $\mathcal{S}(\mathbb{R}^n)$, which is contained in the domain of $A(\mathbf{x}, \mathbf{p})$, in order that (8.12) makes sense.

Following Shishmarev [SH], and using the theory of Fourier integral operators, we can define an operator-valued kernel for $U(t, 0)$ by

$$K(\mathbf{x}, t\,;\, \mathbf{y}, 0) = \int_{\mathbb{R}^n} e^{\pi i(\mathbf{x}-\mathbf{y}, \boldsymbol{\eta})} Q(\mathbf{x}, t, \boldsymbol{\eta}, 0)d\boldsymbol{\eta},$$

so that

$$\psi(\mathbf{x}, t) = U(t, 0)\psi_0(\mathbf{x}) = \int_{\mathbb{R}^n} K(\mathbf{x}, t\,;\, \mathbf{y}, 0)\psi_0(\mathbf{y})d\mathbf{y}. \tag{8.13}$$

The following results are due to Shishmarev [SH].

Theorem 8.22. *If $A(\mathbf{x}, \mathbf{p})$ is a self-adjoint generator of a strongly continuous unitary group with domain D, $\mathcal{S}(\mathbb{R}^n) \subset D$ in $L^2(\mathbb{R}^n)$, such that conditions (1)–(3) are satisfied, then there exists precisely one symbol $Q(\mathbf{x}, t, \boldsymbol{\eta}, 0)$ for the Cauchy problem (8.11).*

Theorem 8.23. *If we replace condition (3) in Theorem 8.22 by the stronger condition*

(3′) $\frac{|a(\mathbf{x}, \boldsymbol{\eta})|}{|\lambda_0(\mathbf{x}, \boldsymbol{\eta})|} = O\left[(1 + |\boldsymbol{\eta}|)^{(\xi-\delta)/(3k-1-\varepsilon)}\right]$, $\varepsilon > 0, |\boldsymbol{\eta}| > c_0$,

then the symbol $Q(\mathbf{x}, t, \boldsymbol{\eta}, 0)$ of the Cauchy problem (8.11) has the asymptotic behavior near $t = 0$:

$$Q(\mathbf{x}, t, \boldsymbol{\eta}, 0) = exp[-(i/\hbar)ta(\mathbf{x}, \boldsymbol{\eta})] + o(1),$$

uniformly for $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^{\mathbf{n}}$.

Now, using Theorem 8.23 we see that, under the stronger condition (3′), the kernel $K(\mathbf{x}, t\,;\, \mathbf{y}, 0)$ satisfies

$$K(\mathbf{x}, t; \mathbf{y}, 0) = \int_{\mathbb{R}^n} \exp[\pi i(\mathbf{x} - \mathbf{y}, \boldsymbol{\eta}) - (i/\hbar)ta(\mathbf{x}, \boldsymbol{\eta})]d\boldsymbol{\eta} + \int_{\mathbb{R}^n} \exp[\pi i(\mathbf{x} - \mathbf{y}, \boldsymbol{\eta})]o(1)d\boldsymbol{\eta}.$$

From the extension theory of Chap. 5, we see that $A(\mathbf{x}, \mathbf{p})$ has a self-adjoint extension to $KS^2(\mathbb{R}^n)$, which also generates a unitary group.

This means that we can construct a path integral in the same (identical) way as was done for the free-particle propagator (i.e., for all Hamiltonians with symbols in $\mathcal{S}^{m}_{\alpha,\delta}$). Furthermore, it follows that the same comment applies to any Hamiltonian that has a kernel representation, independent of its symbol class. This partially proves a conjecture made in [GZ3], to the effect that there is a kernel for every "physically generated" semigroup.

8.6. Path Integrals II: Time-Ordered Theory

If we want to consider perturbations of the Hamiltonians with various potentials, the normal analytical problems arise. In this case, we must resort to the limited number of Trotter–Kato type results that may apply on $KS^2(\mathbb{R}^n)$. The general question is, "Under what conditions can we expect a path integral to exist?"

8.6.1. Time-Ordered Path Integrals. The results of the last section have direct extensions to time-dependent Hamiltonians, but the operators need not commute. Thus, in order to construct general path integrals, we must use the full power of the time-ordered operator theory in Chap. 7. In this section, we show that the path integral is a special case of the time-ordered operator theory as suggested by Feynman and automatically leads to a generalization and extension of Feynman–Kac theory.

Before proceeding, we should briefly pause for a few words about progress on the development of the Feynman–Kac theory as it relates to nonautonomous systems, evolution processes or time-dependent propagators and their relationship to path integrals and quantum field theory. The major developments in these areas along with many interesting applications can be found in the relatively recent books by: Jefferies [JE], Lorinczi [LO], Gulishashvili and Van Casteren [GC], and Del Moral [MO].

Let $U[t,a]$ be an evolution operator on $KS^2(\mathbb{R}^3)$, with time-dependent generator $A(t)$, which has a kernel $\mathbf{K}[\mathbf{x}(t),t\,;\,\mathbf{x}(s),s]$ such that:

$$\mathbf{K}\left[\mathbf{x}(t),\,t;\,\mathbf{x}(s),\,s\right]=\int_{\mathbf{R}^3}\mathbf{K}\left[\mathbf{x}(t),\,t;\,d\mathbf{x}(\tau),\,\tau\right]\mathbf{K}\left[\mathbf{x}(\tau),\,\tau;\,\mathbf{x}(s),\,s\right],$$

$$U[t,s]\varphi(s)=\int_{\mathbf{R}^3}\mathbf{K}\left[\mathbf{x}(t),\,t;\,d\mathbf{x}(s),\,s\right]\varphi(s).$$

Now let $\mathbf{U}[t,s]$ be the corresponding time-ordered version defined on $\mathcal{FD}_{\otimes}^2 \subset \mathcal{H}_{\otimes}^2$, with kernel $\mathbb{K}_{\mathbf{f}}\,[\mathbf{x}(t),\,t;\,\mathbf{x}(s),\,s]$. Since $\mathbf{U}[t,\tau]\mathbf{U}[\tau,s] = \mathbf{U}[t,s]$, we have:

$$\mathbb{K}_{\mathbf{f}}\,[\mathbf{x}(t),\,t;\,\mathbf{x}(s),\,s] = \int_{\mathbf{R}^3} \mathbb{K}_{\mathbf{f}}\,[\mathbf{x}(t),\,t;\,d\mathbf{x}(\tau),\,\tau]\mathbb{K}_{\mathbf{f}}\,[\mathbf{x}(\tau),\,\tau;\,\mathbf{x}(s),\,s]\,.$$

From our sum over paths representation for $\mathbf{U}[t,s]$, we have:

$$\begin{aligned}\mathbf{U}[t,s]\Phi(s) &= \lim_{\lambda\to\infty} \mathbf{U}_\lambda[t,s]\Phi(s)\\ &= \lim\nolimits_{\lambda\to\infty} \mathrm{e}^{-\lambda(t-s)} \sum_{k=0}^{\infty} \frac{[\lambda\,(t-s)]^k}{k!}\mathbf{U}_k[t,s]\Phi(s),\end{aligned}$$

where

$$\mathbf{U}_k[t,s]\Phi(s) = \exp\left\{(-i/\hbar)\sum_{j=1}^{k}\int_{t_{j-1}}^{t_j} \mathbf{E}[(j/\lambda),\tau]\mathcal{A}(\tau)d\tau\right\}\Phi(s).$$

As in Sect. 8.1, we define $\mathbb{K}_{\mathbf{f}}[\mathcal{D}_\lambda\mathbf{x}(\tau)]$ by:

$$\begin{aligned}&\int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{s})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}_\lambda\mathbf{x}(\tau)]\\ &=: \mathrm{e}^{-\lambda(t-s)}\sum_{k=0}^{n}\frac{[\lambda(t-s)]^k}{k!}\left\{\prod_{j=1}^{k}\int_{\mathbf{R}^3}\mathbb{K}_{\mathbf{f}}[t_j,\mathbf{x}(t_j)\,;\,d\mathbf{x}(t_{j-1}),t_{j-1}]\big|^{(j/\lambda)}\right\},\end{aligned}$$

where $n = [|\lambda(t-s)|]$, the greatest integer in $\lambda(t-s)$, and $|^{(j/\lambda)}$ denotes that the integration is performed in the time-slot (j/λ).

Definition 8.24. We define the Feynman path integral associated with $\mathbf{U}[t,s]$ by:

$$\mathbf{U}[t,s]\Phi(s) = \int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{s})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau)]\Phi(s) = \lim_{\lambda\to\infty}\int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{s})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}_\lambda\mathbf{x}(\tau)]\Phi(s).$$

Theorem 8.25. *For the time-ordered theory, whenever a kernel exists, we have that:*

$$\lim_{\lambda\to\infty}\mathbf{U}_\lambda[t,s]\Phi(s) = \mathbf{U}[t,s]\Phi(s) = \int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{s})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}_\lambda\mathbf{x}(\tau)]\Phi[\mathbf{x}(s)],$$

and the limit is independent of the space of continuous functions.

Let us assume that $A_0(t)$ and $A_1(t)$ are strongly continuous generators of C_0-contraction semigroups, with a common dense domain $D(t)$, for each $t \in E = [a,b]$, and let $\mathcal{A}_{1,\rho}(t) = \rho\mathcal{A}_1(t)\mathbf{R}(\rho,\mathcal{A}_1(t))$ be the Yosida approximator for the time-ordered version of $A_1(t)$, with dense domain $D = \mathcal{F}\mathcal{D}_{\otimes}^2 \cap \otimes_{t \in I} D(t)$. Define $\mathbf{U}^{\rho}[t,a]$ and $\mathbf{U}^0[t,a]$ by:

$$\mathbf{U}^{\rho}[t,a] = \exp\{(-i/\hbar)\int_a^t [\mathcal{A}_0(s) + \mathcal{A}_{1,\rho}(s)]ds\},$$

$$\mathbf{U}^0[t,a] = \exp\{(-i/\hbar)\int_a^t \mathcal{A}_0(s)ds\}.$$

Since $\mathcal{A}_{1,\rho}(s)$ is bounded, $\mathcal{A}_0(s) + \mathcal{A}_{1,\rho}(s)$ is a generator of a C_0-contraction semigroup for $s \in E$ and finite ρ. Now assume that $\mathbf{U}^0[t,a]$ has an associated kernel, so that $\mathbf{U}^0[t,a] = \int_{\mathbb{R}^{3[t,s]}} \mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau); \mathbf{x}(a)]$. We now have the following general result, which is independent of the space of continuous functions.

Theorem 8.26. *(Feynman–Kac)* If $\mathcal{A}_0(s) \oplus \mathcal{A}_1(s)$ is a generator of a C_0-contraction semigroup, then*

$$\lim_{\rho \to \infty} \mathbf{U}^{\rho}[t,a]\Phi(a) = \mathbf{U}[t,a]\Phi(a)$$
$$= \int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{a})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau)] \exp\{(-i/\hbar)\int_a^{\tau} \mathcal{A}_1(s)ds]\}\Phi[\mathbf{x}(a)].$$

Proof. The fact that $\mathbf{U}^{\rho}[t,a]\Phi(a) \to \mathbf{U}[t,a]\Phi(a)$ is clear. To prove that

$$\mathbf{U}[t,a]\Phi(a) = \int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{a})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau)] \exp\{(-i/\hbar)\int_a^t \mathcal{A}_1(s)ds\}\Phi(a),$$

first note that, since the time-ordered integral exists and we are only interested in the limit, we can write for each k

$$U_k^{\rho}[t,a]\Phi(a)$$
$$= \exp\left\{(-i/\hbar)\sum\nolimits_{j=1}^{k}\int_{t_{j-1}}^{t_j} \left[\mathbf{E}[\tau_j, s]\mathcal{A}_0(s) + \mathbf{E}[\tau_j', s]\mathcal{A}_{1,\rho}(s)\right] ds\right\}\Phi(a),$$

where τ_j and τ_j' are distinct points in the interval (t_{j-1}, t_j). Thus, we can also write $U_k^{\rho}[t,a]\Phi(a)$ as

$$\begin{aligned}
&\mathbf{U}_k^{\rho}[t,a] \\
&= exp\left\{\tfrac{-i}{\hbar}\sum_{j=1}^{k}\int_{t_{j-1}}^{t_j}\mathbf{E}[\tau_j,s]\mathcal{A}_0(s)ds\right\}exp\left\{\tfrac{-i}{\hbar}\sum_{j=1}^{k}\int_{t_{j-1}}^{t_j}\mathbf{E}[\tau_j',s]\mathcal{A}_{1,\rho}(s)ds\right\} \\
&= \prod_{j=1}^{k}exp\left\{\tfrac{-i}{\hbar}\int_{t_{j-1}}^{t_j}\mathbf{E}[\tau_j,s]\mathcal{A}_0(s)ds\right\}exp\left\{\tfrac{-i}{\hbar}\sum_{j=1}^{k}\int_{t_{j-1}}^{tj}\mathbf{E}[\tau_j',s]\mathcal{A}_{1,\rho}(s)ds\right\} \\
&= \prod_{j=1}^{k}\int_{\mathbb{R}^3}\mathbb{K}_{\mathbf{f}}\left[t_j,\mathbf{x}(t_j)\ ;\ t_{j-1},d\mathbf{x}(t_{j-1})\right]exp\left\{\tfrac{-i}{\hbar}\sum_{j=1}^{k}\int_{t_{j-1}}^{t_j}\mathbf{E}[\tau_j',s]\mathcal{A}_{1,\rho}(s)ds\right\}.
\end{aligned}$$

If we put this in our experimental evolution operator $\mathbf{U}_\lambda^{\rho}[t,a]\Phi(a)$ and compute the limit, we have:

$$\begin{aligned}
&\mathbf{U}^{\rho}[t,a]\Phi(a) \\
&= \int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{a})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})}\mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau)]\exp\left\{(-i/\hbar)\int_a^t\mathcal{A}_{1,\rho}(s)ds\right\}\Phi(a).
\end{aligned}$$

Since the limit as $\rho\to\infty$ on the left exists, it defines the limit on the right. □

8.6.2. Examples. In this section, we pause to discuss a few examples. Theorem 8.26 is somewhat abstract, so it may not be clear as to its application. Our first example is a direct application of this theorem, which covers all of nonrelativistic quantum theory (i.e., the Feynman formulation of quantum theory).

Let $\triangle$ be the Laplacian on $\mathbf{R}^n$ and let V be any potential such that $A = (-\hbar^2/2)\triangle + V$ generates an unitary group. Then the problem:

$$(i\hbar)\partial\psi(\mathbf{x},t)/\partial t = A\psi(\mathbf{x},t),\quad \psi(\mathbf{x},0)=\psi_0(\mathbf{x}),$$

has a solution with a Feynman–Kac representation.

Our second example is more specific and is due to Albeverio and Mazzucchi [AM]. Their paper provides an excellent view of the power of the approach first introduced by Albeverio and H∅egh-Krohn [AH]. Let $\mathbb{C}$ be a completely symmetric positive definite fourth-order covariant tensor on $\mathbb{R}^n$, let Ω be a symmetric positive definite $n\times n$ matrix, and let λ be a nonnegative constant. It is known [RS1] that the operator

$$\bar{A} = -\tfrac{\hbar^2}{2}\Delta + \tfrac{1}{2}\mathbf{x}\Omega^2\mathbf{x} + \lambda\mathbb{C}[\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}]$$

is a densely defined self-adjoint generator of an unitary group on $L^2[\mathbb{R}^n]$. Using a substantial amount of elegant analysis, Albeverio and Mazzucchi [AM] prove that $\bar{A}$ has a path integral representation as the analytic continuation (in the parameter λ) of an infinite dimensional generalized oscillatory integral.

Our approach to the same problem is both simple and direct using the results of the previous sections. First, since $\bar{A} = \bar{A}^*$ is densely defined on $L^2[\mathbb{R}^n]$, $\bar{A}$ has a closed densely defined self-adjoint extension A to $KS^2[\mathbb{R}^n]$, which generates a unitary group. If we set $V = \frac{1}{2}\mathbf{x}\Omega^2\mathbf{x} + \lambda\mathbb{C}[\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}]$ and $V_\rho = V(I+\rho V^2)^{-1/2}$, $\rho > 0$, it is easy to see that V_ρ is a bounded self-adjoint operator which converges to V on $D(V)$. (This follows from the fact that a bounded (self-adjoint) perturbation of an unbounded self-adjoint operator is self-adjoint.) Now, since $-\frac{\hbar^2}{2}\Delta$ generates a unitary group, $A_\rho = -\frac{\hbar^2}{2}\Delta + V_\rho$ also generates one and converges to A on $D(A)$. Let

$$\mathcal{A}(\tau) = (\hat{\otimes}_{t \geqslant s > \tau} \mathrm{I}_s) \otimes A \otimes (\otimes_{\tau > s \geqslant 0} \mathrm{I}_s),$$

then $\mathcal{A}(t)$ generates a unitary group for each t and $\mathcal{A}_\rho(t)$ converges to $\mathcal{A}(t)$ on $D[\mathcal{A}(t)] \subset \mathcal{FD}^2_\otimes$. We can now apply Theorem 8.26 to get that:

$$\begin{aligned}
&\mathbf{U}[t,a]\Phi \\
&= \int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{a})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau)] \exp\{-(i/\hbar)\int_a^\tau V(s)ds\}\Phi \\
&= \lim_{\rho\to 0} \int_{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{a})}^{\mathbf{x}(\tau)=\mathbf{x}(\mathbf{t})} \mathbb{K}_{\mathbf{f}}[\mathcal{D}\mathbf{x}(\tau)] \exp\{-(i/\hbar)\int_a^\tau V_\rho(s)ds\}\Phi.
\end{aligned}$$

Under additional assumptions, Albeverio and Mazzucchi are able to prove Borel summability of the solution in power series of the coupling constant. With Theorem 7.25 of Chap. 7, we get the Dyson expansion to any order with remainder.

8.7. Dyson's First Conjecture

This section is the last one in the book for two reasons. First, our original objective, leading to most of the work in the book, was to provide an answer this conjecture. The second reason is that this section does not provide any additional mathematics. It essentially gives a physical reinterpretation of the mathematics developed earlier.

At the end of his second paper on the relationship between the Feynman and Schwinger–Tomonaga theories, Dyson explored the difference between the divergent Hamiltonian formalism that one must begin with and the finite S-matrix that results from renormalization (see [DY2]). He takes the view that it is a contrast between a real observer and a fictitious (ideal) observer. The real observer can only determine particle positions with limited accuracy and always gets finite results from his measurements. Dyson then suggests that "... The ideal observer, however, using nonatomic apparatus whose location in space and time is known with infinite precision, is imagined to be able to disentangle a single field from its interactions with others, and to measure the interaction. In conformity with the Heisenberg uncertainty principle, it can perhaps be considered a physical consequence of the infinitely precise knowledge of (particle) location allowed to the ideal observer, that the value obtained when he measures (the interaction) is infinite." He goes on to remark that if his analysis is correct, the problem of divergences is attributable to an idealized concept of measurability. The purpose of this section is to develop the conceptual and technical framework that will allow us to discuss this conjecture.

8.7.1. The S-Matrix. The objective of this section is to provide a formulation of the S-matrix that will allow us to investigate the mathematical sense in which we can believe Dyson's conjecture.

In order to explore this idea, we work in the interaction representation with obvious notation. Replace the interval $[t, 0]$ by $[T, -T]$, $H(t)$ by $\frac{-i}{\hbar}H_I(t)$, and our experimental evolution operator $\mathbf{U}_\lambda[T, -T]\Phi$ by the experimental scattering operator (or S-matrix) $\mathbf{S}_\lambda[T, -T]\Phi$, where

$$\mathbf{S}_\lambda[T, -T]\Phi = \sum_{n=0}^{\infty} \frac{(2\lambda T)^n}{n!} \exp\left[-2\lambda T\right] \mathbf{S}_n[T, -T]\Phi, \tag{8.14}$$

$$\mathbf{S}_n[T, -T]\Phi = \exp\left\{ \frac{-i}{\hbar} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} E[\tau_j, s]H_I(s)ds \right\} \Phi, \tag{8.15}$$

and $H_I(t) = \int_{\mathbb{R}^3} H_I(\mathbf{x}(t), t)d\mathbf{x}(t)$ is the interaction energy. We now give a physical interpretation of our formalism. Rewrite Eq. (8.14) as

$$\mathbf{S}_\lambda[T, -T]\Phi = \sum_{n=0}^{\infty} \frac{(2\lambda T)^n}{n!} \exp\left\{ \frac{-i}{\hbar} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} \left[E[\tau_j, s]H_I(s) - i\lambda\hbar \mathrm{I}_\otimes\right] ds \right\} \Phi. \tag{8.16}$$

In this form, it is clear that the term $-i\lambda\hbar \mathrm{I}_{\otimes}$ has a physical interpretation as the absorption of photon energy of amount $\lambda\hbar$ in each subinterval $[t_{j-1}, t_j]$ (cf. Mott and Massey [MM]). When we compute the limit, we get the standard S-matrix (on $[T, -T]$). It follows that we must add an infinite amount of photon energy to the mathematical description of the experimental picture (at each point in time) in order to obtain the standard scattering operator. This is the ultraviolet divergence and shows explicitly that the transition from the experimental to the ideal scattering operator requires that we illuminate the particle throughout its entire path. Thus, it appears that we have, indeed, violated the uncertainty relation. This is further supported if we look at the form of the standard S-matrix:

$$\mathbf{S}[T, -T]\Phi = \exp\left\{(-i/\hbar)\int_{-T}^{T} H_I(s)ds\right\}\Phi \tag{8.17}$$

and note that the differential ds in the exponent implies perfect infinitesimal time knowledge at each point, strongly suggesting that the energy should be totally undetermined. If violation of the Heisenberg uncertainty relation is the cause for the ultraviolet divergence, then as it is a variance relation, it will not appear in first order (perturbation) but should show up in all higher-order terms. On the other hand, if we eliminate the divergent terms in second order, we would expect our method to prevent them from appearing in any higher order term of the expansion. The fact that this is precisely the case in quantum electrodynamics is a clear verification of Dyson's conjecture.

If we allow T to become infinite, we once again introduce an infinite amount of energy into the mathematical description of the experimental picture, as this is also equivalent to precise time knowledge (at infinity). Of course, this is the well-known infrared divergence and can be eliminated by keeping T finite (see Dahmen et al. [DA]) or introducing a small mass for the photon (see Feynman [FY3]). If we hold λ fixed while letting T become infinite, the experimental S-matrix takes the form:

$$\begin{gathered}\mathbf{S}_\lambda[\infty, -\infty]\Phi = \exp\left\{(-i/\hbar)\sum_{j=1}^{\infty}\int_{t_{j-1}}^{t_j} E[\tau_j, s]H_I(s)ds\right\}\Phi, \\ \bigcup_{j=1}^{\infty}[t_{j-1}, t_j] = (-\infty, \infty), \ \& \ \Delta t_j = \lambda^{-1}.\end{gathered} \tag{8.18}$$

This form is interesting since it shows how a minimal time eliminates the ultraviolet divergence. Of course, this is not unexpected, and has been known at least since Heisenberg [HE] introduced his fundamental length as a way around the divergences. This was a prelude to the various lattice approximation methods. The review by Lee [LE] is interesting in this regard. In closing this section, we record our exact expansion for the S-matrix to any finite order. Let $\mathcal{H}_k = H_I(s_1)\cdots H_I(s_k)$ and let $\Phi \in D\left[(\mathbf{Q}[\infty,-\infty])^{n+1}\right]$, we have

$$\begin{aligned}\mathbf{S}[\infty,-\infty]\Phi(-\infty) &= \sum_{k=0}^{n}\left(\tfrac{-i}{\hbar}\right)^k \int_{-\infty}^{\infty} ds_1 \cdots \int_{-\infty}^{s_{k-1}} ds_k \mathcal{H}_k \Phi \\ &+ \left(\tfrac{-i}{\hbar}\right)^{n+1} \int_0^1 (1-\xi)^n d\xi \int_{-\infty}^{\infty} ds_1 \cdots \int_{-\infty}^{s_n} ds_{n+1}\mathcal{H}_{n+1}\mathbf{S}^{\xi}[s_{n+1},-\infty]\Phi.\end{aligned} \tag{8.19}$$

It follows that (in a theoretical sense) we can consider the standard S-matrix expansion to be exact, when truncated at any order, by adding the last term of Eq. (8.19) to give the remainder. This result also means that whenever we can construct an exact nonperturbative solution, it always implies the existence of a perturbative solution valid to any order. However, in general, only in particular cases can we know if the series at some n (without the remainder) approximates the solution.

In this section we have provided a precise formulation and proof of Dyson's conjecture that the ultraviolet divergence is caused by a violation of the Heisenberg uncertainty relation at each point in time.

In closing, since the time of Dyson's original work, a large amount of progress has been made in understanding the mathematical and physical foundations of relativistic quantum theory. (For a brief discussion including references for further reading, see Gill and Zachary [GZ] and [GZ1].) However, many of the problems encountered by the earlier workers are still with us in one form or another.

References

[AH] S. Albeverio, R. HØegh-Krohn, *Mathematical Theory of Feynman Path Integrals*. Lecture Notes in Mathematics, vol. 523 (Springer, Berlin, 1976)

[AM] S. Albeverio, S. Mazzucchi, Feynman path integrals for polynomially growing potentials. J. Funct. Anal. **221**, 83–121 (2005)

[CDM1] P. Cartier, C. Dewitt-Morette, Functional integration. J. Math. Phys. **41**, 4154–4185 (2000)

[CDM2] P. Cartier, C. Dewitt-Morette, *Functional Integration, Action and Symmetries* (Cambridge University Press, New York, 2006)

[CO] H.O. Cordes, *The Technique of Pseudodifferential Operators* (Cambridge University Press, Cambridge, 1995)

[DA] H.D. Dahmen, B. Scholz, F. Steiner, Infrared dynamics of quantum electrodynamics and the asymptotic behavior of the electron form factor. Nucl. Phys. B **202**, 365 (1982)

[DP] G. Da Prato, *Kolmogorov Equations for Stochastic PDEs*. Advanced Courses in Mathematics CRM Barcelona (Birkhäuser, Boston, 2004)

[MO] P. Del Moral, *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications* (Springer, New York, 2004)

[DY2] F. Dyson, The S-matrix in quantum electrodynamics. Phys. Rev. **75**, 1736–1755 (1949)

[EV] L.C. Evans, *Partial Differential Equations*. AMS Graduate Studies in Mathematics, vol. 18 (American Mathematical Society, Providence, 1998)

[FY1] R.P. Feynman, The principle of least action in quantum mechanics, Ph.D. dissertation, Physics, Princeton University, 1942 (Available from University Microfilms Publications, No. 2948, Ann Arbor, MI)

[FY3] R.P. Feynman, *Quantum Electrodynamics* (Benjamin, New York, 1964)

[FH] R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965)

[GZ3] T.L. Gill, W.W. Zachary, Time-ordered operators and Feynman-Dyson algebras. J. Math. Phys. **28**, 1459–1470 (1987)

[GZ] T.L. Gill, W.W. Zachary, Foundations for relativistic quantum theory I: Feynman's operator calculus and the Dyson conjectures. J. Math. Phys. **43**, 69–93 (2002)

[GZ4] T.L. Gill, W.W. Zachary, Analytic representation of the square-root operator. J. Phys. A Math. Gen. **38**, 2479–2496 (2005)

[GZ1] T.L. Gill, W.W. Zachary, Feynman operator calculus: the constructive theory. Expo. Math. **29**, 165–203 (2011)

[GRRZ] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1965)

[GS] C. Grosche, F. Steiner, *Handbook of Feynman Path Integrals*. Springer Tracts in Modern Physics, vol. 145 (Springer, New York, 1998)

[GC] A. Gulishashvili, J.A. Van Casteren, *Non-autonomous Kato Classes and Feynman-Kac Propagators* (Springer, New York, 2006)

[HE] W. Heisenberg, Die "beobachtbaren Grossen" in der theorie der Elementarteilchen II. Z. Phys. **120**, 673–702 (1943)

[HO] L. Hörmander, Fourier integral operators I. Acta Math. **127**, 79–183 (1971)

[KL] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics and Polymer Physics*, 3rd edn. (World Scientific, Singapore, 2004)

[KG] H. Kumano-go, *Pseudo-Differential Operators* (The MIT Press, Cambridge/London, 1982)

[JE] B. Jefferies, *Evolution Processes and the Feynman-Kac Formulas* (Springer, New York, 2010)

[JL] G.W. Johnson, M.L. Lapidus, *The Feynman Integral and Feynman's Operational Calculus* (Oxford University Press, New York, 2000)

[LX1] P. Lax, *The L^2 Operator Calculus of Mikhlin, Calderón and Zygmund* (Lectures Notes). Courant Institute of Mathematical Sciences (New York University, New York, 1963)

[LE] T.D. Lee, in *The Lesson of Quantum Theory*, ed. by J. de Boer, E. Dal, O. Ulfbeck (Elsevier, Amsterdam, 1986), p. 181

[LB] L. Lorenzi, M. Bertoldi, *Analytical Methods for Markov Semigroups*. Pure and Applied Mathematics (Chapman & Hall/CRC, New York, 2007)

[LO] J. Lorinczi, *Feynman-Kac-type Theorems and Gibbs Measures on Path Space: With Applications to Rigorous Quantum Field Theory*. De Gruyter Studies in Mathematics (De Gruyter, New York, 2008)

[M] V.P. Maslov, *Operational Methods* (Mir, Moscow, 1973). English translation 1976 (revised from the Russian edition)

[MM] N.F. Mott, H.S.W. Massey, *The Theory of Atomic Collisions*. The International Series of Monographs on Physics (Clarendon Press/Oxford University Press, Oxford/New York, 1933)

[RS1] M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic, New York, 1972)

[SY] G.R. Sell, Y. You, *Dynamics of Evolutionary Equations*. Applied Mathematical Sciences, vol. 143 (Springer, New York, 2002)

[SH] I.A. Shishmarev, On the Cauchy problem and T-products for hypoelliptic systems. Math. USSR Izvestiya **20**, 577–609 (1983)

[SHB] M.A. Shubin, *Pseudodifferential Operators and Spectral Theory* (Nauka, Moscow, 1978) (Russian)

[TA] M.E. Taylor, *Pseudodifferential Operators* (Princeton University Press, Princeton, 1981)

[TR] F. Treves, *Introduction to Pseudodifferential and Fourier Integral Operators.* Pseudodifferential Operators, vol. 1. Fourier Integral Operators, vol. 2 (Plenum Press, New York/London, 1980)

[WSRM] N. Wiener, A. Siegel, B. Rankin, W.T. Martin, *Differential Space, Quantum Systems, and Prediction* (The MIT Press, Cambridge, 1966)

[YS] K. Yosida, *Functional Analysis*, 2nd edn. (Springer, New York, 1968)

Index

T.L. Gill, W.W. Zachary, *Functional Analysis and the Feynman Operator Calculus*, DOI 10.1007/978-3-319-27595-6

T

U