

# Functional Analysis and Operator Theory

Nicola Arcozzi



Copyright ©2025 Nicola Arcozzi

**License:**

This work is licensed under the Creative Commons Attribution-Noncommercial Share Alike 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-sa/4.0/> or send a letter to Creative Commons PO Box 1866, Mountain View, CA 94042, USA. You can use, print, duplicate, and share this book as much as you want. You can base your own notes on it and reuse parts if you keep the license the same. The license is CC-BY-NC-SA. Your derivative work must use the same licenses. Derivative works must be prominently marked as such.

## Preface

These lecture notes merge the contents of two versions of a one-semester course in Functional Analysis and Operator Theory I taught in 2024, then in 2025, for the Master in Mathematics at the University of Bologna. In 2024 I did the basics of Gelfand's theory and spectral properties of compact operators, but the spectral theory of self-adjoint operators for the bounded case only. In 2025 I skipped Gelfand's theory and compact operators, and I did instead the spectral theory of unitary and of possibly unbounded self-adjoint operators. All students were familiar with the basis of Hilbert space theory, and the chapter on Hilbert spaces appears here just to keep the notes reasonably self-contained. The structure of the course, on the other hand, stemmed from a [course](#) in Functional Analysis I taught in 2022 for the "SMI" Summer School in Mathematics of Perugia.

The goal of the course, as it happens with most courses we teach, is to give an introduction to the subject which might be useful (i) as general mathematical culture; (ii) for those who will use these tools or some of their many *avatars* in more specialized topics; (iii) as an invitation to study Functional Analysis in greater depth (there are many beautiful treatises, old and new, the study of which might even be pleasant after an introductory course has cleared some ground). Having (i) in mind, I have preferred narrowing the front and moving in depth in the direction of spectral theory, which is one of great achievements of the first thirty or so years of the XX century. Concerning (ii), I unfortunately had to sacrifice the "quick introduction to Calculus of Variations from the viewpoint of functional analysis" I always have in the back of my mind. Time is short and choices have to be made.

The choice of much background material is based on the concrete groups of students attending the class. What some of them had not seen before, I stated and proved (the exception being some notions on Borel measures, which I just enumerated without proofs). All of them were familiar with Fourier transforms, for instance, but many had just a vague idea of what Fourier series are about.

All this accounts for the Frankensteinish structure of the lecture notes, where a number of repetitions and inversions of topics occur. For instance, in the chapter on preliminaries I use Banach-Alaoglu theorem, which appears two chapters downstream. In cases like this, I have preferred to keep most contents of the same kind in the same chapter, where the student can find them better contextualized. The chapters on Hilbert and Banach spaces are copied from notes for other courses and I did not make much of an effort to harmonize them with the rest.

Some chapters have introductions explaining why we study this or that and which sources I am using. Let me just mention the main references:

- (1) Hilbert and Banach spaces: Michael Reed, Barry Simon, [Methods of Modern Mathematical Physics: Functional analysis](#), Academic Press, 1972;
- (2) compact operators and their spectra: Reed and Simon;
- (3) Banach algebras: Functional analysis, by Peter D. Lax, Wiley-Interscience, 2002, see this [review by Meijun Zhu](#);
- (4) spectral theory: for the unitary operators, Michael Taylor's lecture notes [The Spectral Theorem for Self-Adjoint and Unitary Operators \(2018\)](#); for the unbounded, self-adjoint operators, [The spectral theorem and its converses for unbounded symmetric operators \(2011\)](#) by Terence Tao.

Functional analysis, unlike Athena, was not born fully grown and armored from the head of Zeus. Its origins lie in rather concrete and specific problems, especially those concerning integral and differential equations, and calculus of variations. Then quantum mechanics was invented, and it needed what functional analysis was available, and much more. The presence, and the usefulness, of abstract theories emerged gradually, but not so slowly. There are several places where the story is told. One of them is [The Establishment of Functional Analysis](#) (Historia Mathematica 11(3) 1984, p. 258-321) by Garrett Birkhoff and Erwin Kreyszig.

It is a pleasure to thank Nikolaos Chalmoukis, Giovanna Citti, Giovanni Dore, and Davide Guidetti, with whom I had a number of conversations on the material while teaching the class. Some material was included to answer questions from attendees of the Spring 2025 class; often ones I would not have ever thought of (and point in fact to important chapters of the theory).

The L<sup>A</sup>T<sub>E</sub>Xsource is downloadable, in case you wish to rearrange and modify the material according to your taste and needs. For comments and pointing out mistakes, you can write me at [nicola.arcozzi@unibo.it](mailto:nicola.arcozzi@unibo.it).

Bologna 2025  
Nicola Arcozzi



# Contents

Preface	ii
Chapter 1. Some tools from real and complex analysis	1
1.1. The definition of Banach space	1
1.1.1. Normed linear spaces	1
1.1.2. Complete metric spaces and Banach spaces	2
1.2. Borel measures	6
1.2.1. Borel measures on the real line	6
1.2.2. Borel measurable functions and the Lebesgue-Hausdorff theorem	9
1.3. Fourier transforms and series	12
1.3.1. The Fourier transform	13
1.3.2. Fourier series	14
1.4. Basic theory of harmonic functions	16
1.5. The Dirichlet problem for the Laplace equation in the upper half plane	23
1.6. The Fourier transforms of Borel measures	28
1.7. Herglotz representation	30
1.8. Bochner's theorem	33
Chapter 2. Hilbert spaces	39
2.1. Basic geometry of Hilbert spaces and Riesz Lemma	39
2.1.1. Definition and basic properties	39
2.1.2. $L^2$ as a Hilbert space	41
2.1.3. Projections onto subspaces	41
2.1.4. F. Riesz representation in Hilbert spaces	44
2.2. Orthonormal systems	45
2.2.1. Orthogonal vectors	45
2.2.2. Fourier analysis and synthesis of a vector	46
2.2.3. Orthonormal basis in separable Hilbert spaces	47
2.2.3.1. Gram-Schmidt algorithm	47
2.2.3.2. The classification of separable Hilbert spaces	48
2.2.4. Intermezzo: orthonormal basis in general Hilbert spaces	49
2.2.4.1. Existence of o.n.b.	49

2.2.4.2. The dimension of a Hilbert space	49
2.3. The trigonometric system and Fourier series	50
Chapter 3. Banach spaces	53
3.1. Zorn's lemma and some of its consequences	53
3.2. The Hahn-Banach Theorem and some of its consequences	57
3.3. The dual of a Banach space	60
3.4. Weak and weak* topologies, and the Banach-Alaoglu theorem	64
3.4.1. The weak and the weak* topologies	64
3.4.1.1. The weak topology	65
3.4.1.2. The weak* topology	66
3.4.2. Two versions of the Banach-Alaoglu theorem	67
3.4.2.1. Tychonoff's theorem	67
3.4.2.2. Banach-Alaoglu theorem: the topological form	68
3.4.2.3. Banach-Alaoglu theorem: the sequential form	69
3.5. Baire's Theorem and the uniform boundedness principle	70
3.5.1. Banach space-valued holomorphic functions	73
3.6. The Open Mapping Theorem and the Closed Graph Theorem	77
3.7. Integrals of continuous, Banach space valued functions	79
3.8. Sesquilinear forms on Hilbert spaces	81
3.9. The Hilbert space adjoint of an operator	84
3.9.1. Basic properties of the adjoint	84
3.9.2. Spectrum and adjoint	86
3.9.3. The spectrum of a bounded, self-adjoint operator	88
3.9.4. The Hausdorff distance between spectra	90
Chapter 4. Compact operators and their spectra	93
4.1. Preliminaries on compactness	93
4.1.1. Compact subsets of Banach spaces	93
4.1.2. Compact operators: generalities	94
4.2. Compact operators on Hilbert spaces	95
4.2.1. Compactness vs. finite rank	95
4.2.2. Fredholm Theorem and its consequences	96
4.2.3. Self-adjoint compact operators	99
4.3. The matrix representation of bounded operators	100
4.3.1. Bounded operators on separable Hilbert spaces as infinite matrices	101
4.3.2. Hilbert-Schmidt operators on Hilbert spaces	104
4.3.3. The trace of a compact operator	106
4.4. Compact integral operators	106
4.4.1. Integral operators with continuous kernel	106
4.4.2. Hilbert-Schmidt integral operators	108



4.4.3. The Volterra operator	108
4.5. Intermezzo: Sturm-Liouville theory	110
4.5.1. The Sturm-Liouville problem, spaces and operators	110
4.5.2. The solution to the Sturm-Liouville problem	113
4.5.2.1. The solution of the Sturm-Liouville problem when $D$ is $1 - 1$	113
4.5.2.2. The solution of the Sturm-Liouville problem in general	115
Chapter 5. Intermezzo: reproducing kernel Hilbert spaces	117
5.1. Equivalent definitions of RKHS	118
5.2. Some RKHS's	123
5.2.1. The Hardy space $H^2$ and its real counterpart	123
5.2.2. The Dirichlet space $\mathcal{D}(\mathbb{D})$	129
5.3. Multipliers of a RKHS	132
5.3.1. Multipliers in RKHS	132
5.3.2. The multiplier space of $H^2(\mathbb{D})$	134
Chapter 6. Banach algebras	135
6.1. Banach algebras	135
6.1.1. Definition and basic properties	135
6.1.2. Invertibility, resolvent, and spectrum	137
6.1.3. Example: the spectrum of a multiplication operator on $L^2(\nu)$	139
6.1.4. Example: the spectrum of a multiplier on $H^2(\mathbb{D})$	141
6.1.5. The spectral radius	142
6.2. Holomorphic calculus	144
6.2.1. The soon-to-be-lost paradise of polynomial calculus	144
6.2.2. Holomorphic calculus	147
6.3. Intermezzo: holomorphic calculus and Laurent series	150
6.4. Gelfand theory of commutative Banach algebras	152
6.4.1. Multiplicative functionals and ideals	153
6.4.2. The algebra of continuous functions	155
6.4.3. Intermezzo: the algebra of multipliers on a reproducing kernel Hilbert space	155
6.4.4. Gelfand's representation	156
6.4.5. The Gelfand representation of $C(X)$ and other examples	159
6.5. Intermezzo: corona theorems	159
6.5.1. The corona problem in reproducing kernel Hilbert spaces	159
6.5.2. An equivalent, analytic formulation of the corona problem	160
Chapter 7. Spectral theory of unitary, and of self-adjoint operators	163
7.1. Intermezzo: spectral theory and quantum mechanics	164
7.1.1. States and observables	165
7.1.2. Commuting observables	166

7.1.3.	Complete systems of commuting observables	168
7.1.4.	Uncertainty	170
7.1.5.	Evolution and the Schroedinger equation	172
7.2.	The spectral theorem for unitary operators	175
7.2.1.	The spectrum of a unitary operator	175
7.2.2.	The continuous functional calculus for a unitary operator	176
7.2.3.	The spectral measures and the measurable calculus	178
7.2.4.	From the measurable calculus to projection valued measures	182
7.2.5.	From the continuous calculus to the multiplicative form	184
7.2.6.	The spectral theory of a multiplication operator	187
7.2.7.	Some applications of the spectral theorem in its multiplicative form	188
7.2.8.	Some final remarks	190
7.3.	Closed operators and their adjoints	191
7.4.	Resolvent and spectrum	195
7.4.1.	Resolvent and spectrum for a general d.d. symmetric operator	195
7.4.2.	The bounded operator case	200
7.5.	Intermezzo: the momentum operator $\mathbf{D} = i^{-1} \frac{d}{dx}$ on various domains	201
7.5.1.	The real line	201
7.5.2.	$\mathbf{D}$ on some domains of functions on $[-\pi, \pi]$	203
7.5.2.1.	$\mathbf{D}$ on $C_c^1(-\pi, \pi)$	203
7.5.2.2.	$\mathbf{D}$ on $C_{per}^1[-\pi, \pi]$	204
7.5.2.3.	$\mathbf{D}$ on $C_\lambda^1[0, 2\pi]$	204
7.5.2.4.	Final remarks on the finite interval case	205
7.5.2.5.	The case of $[0, \infty)$	205
7.6.	The spectral theorem: measurable calculus form	205
7.6.1.	The measurable calculus for a general self-adjoint operator	206
7.7.	The spectral theorem: multiplication form	216
7.7.1.	The multiplicative form of the spectral theorem	216
7.7.2.	Some consequences	219
7.7.2.1.	The spectral theory of a multiplication operator	219
7.7.2.2.	Some applications of the spectral theorem in its multiplicative form	220
7.8.	One parameter groups of unitary operators and Stone's theorem	221
7.9.	Some applications to bounded and compact self-adjoint operators, and to the spectrum	228
7.9.1.	Bounded operators, continuous calculus, spectrum	228
7.9.2.	Compact self-adjoint operators and their inverses	230
	Index	233

Bibliography

237



## CHAPTER 1

### Some tools from real and complex analysis

This section recalls some results from real and complex analysis which will be used in the sequel. Many of them are probably well known by most students with a bachelor in maths or physics, but perhaps not all of them. When in doubt, I provide proofs. I recommend browsing through the material and, only when needed, reading more carefully the parts which are necessary to prove this or that result in spectral theory. The prerequisites are basic notions from advanced calculus, and real and complex analysis. From advanced calculus: Green's formula, closed and exact fields (or 1-forms); from real analysis: measure theory and Borel measures (including F. Riesz representation theorem for measures, which is the only non-elementary prerequisite), the basic density theorems, Fourier transforms; from complex analysis: the Cauchy-Riemann equations, Cauchy theorem and Cauchy formula, the open mapping theorem, expansion of a holomorphic function as a power series. In one proof we use the Banach-Alaoglu theorem, which is proved in the chapter on Banach spaces. Some basics on Fourier series are also in the chapter on Hilbert spaces, with slightly different proofs.

**Notation.** If  $X$  is a topological space, we denote by  $\overline{A}$  the closure of a subset  $A$  of  $X$ . We make an exception for the complex plane  $\mathbb{C}$ , where the closure of  $A \subseteq \mathbb{C}$  is  $\text{Cl}A$ , while  $\overline{A} = \{\bar{z} : z \in A\}$ .

#### 1.1. The definition of Banach space

**1.1.1. Normed linear spaces.** An especially important family of metric spaces is that of the normed linear spaces. A *normed linear space* is a vector space  $X$  over  $\mathbb{C}$  (or over  $\mathbb{R}$ ) endowed with a *norm*: a function

$$\|\cdot\|: X \rightarrow [0, \infty),$$

satisfying the properties:

- (i)  $\|x\| = 0$  if and only if  $x = 0$ ;
- (ii)  $\|\lambda x\| = |\lambda| \|x\|$  if  $x \in X$  and  $\lambda \in \mathbb{C}$ ;
- (iii)  $\|x + y\| \leq \|x\| + \|y\|$ .

In practice, when we introduce a perspective norm on some concrete vector space  $X$ , we have to verify that it has a finite value for each  $x$  in  $X$ , and we have to verify the conditions above, (iii) being sometimes subtle.

A normed linear space becomes a metric space when endowed with the distance

$$d(x, y) := \|x - y\|.$$

In addition to (i-iii), this distance satisfies properties linking it with the algebraic structure.

$$(iv) \quad d(\lambda x, \lambda y) = |\lambda|d(x, y) \text{ if } x, y \in X \text{ and } \lambda \in \mathbb{C};$$

$$(v) \quad d(x + a, y + a) = d(x, y) \text{ if } x, y, a \in X.$$

The map  $x \mapsto \lambda x$  is a (*complex*) *homothety*, and  $x \mapsto x + a$  is a *translation*.

EXERCISE 1.1. Show that if a metric  $d$  on a vector space  $X$  satisfies (iv-v), then  $\|x\| := d(x, 0)$  defines a norm, and that  $d(x, y) = \|x - y\|$ .

For a continuous function  $f: [a, b] \rightarrow \mathbb{C}$  we define:

$$\|f\|_{L^p} = \left( \int_a^b |f(x)|^p \right)^{1/p} \text{ for } 1 \leq p < \infty, \text{ the } L^p \text{ norm,}$$

$$\|f\|_{L_0^\infty} = \max_{x \in [a, b]} |f(x)| = \sup_{x \in [a, b]} |f(x)|, \text{ the uniform norm.}$$

They make  $(C[a, b], \|\cdot\|_{L^p})$  and  $(C[a, b], \|\cdot\|_u)$  into normed linear spaces.

**1.1.2. Complete metric spaces and Banach spaces.** A basic problem in applications is the convergence of a sequence of objects to an object. Consider for instance the convergence of an algorithm. This notion can be formalized as the convergence of a sequence of points to a point in a metric space. Often, however, the nature of the objects in the sequence is clear, but so is not that of the limiting object. Think of the approximation of  $\sqrt{2}$  by means of decimal (or binary) numbers having finitely many digits. The notion of "familiar" objects converging to a "nonfamiliar", "ghost" one is encoded in the notion of Cauchy sequence. The conceptual tool to make "ghosts" into "real" objects is the completion of a metric space.

A sequence  $\{x_n\}_{n=1}^\infty$  in  $(X, d)$  *converges* to the *limit*  $a \in X$ ,  $\lim_{n \rightarrow \infty} x_n = a$ , if for all  $\epsilon > 0$  there is  $n(\epsilon) > 0$  such that if  $n > n(\epsilon)$ , then  $d(x_n, a) \leq \epsilon$ . The sequence  $\{x_n\}_{n=1}^\infty$  is *Cauchy* if for all  $\epsilon > 0$  there is  $n(\epsilon) > 0$  such that if  $n > n(\epsilon)$  and  $j > 0$  one has  $d(x_n, x_{n+j}) \leq \epsilon$ .

EXERCISE 1.2. Show that, if  $\lim_{n \rightarrow \infty} x_n = x$  and  $\lim_{n \rightarrow \infty} y_n = x$  in  $(X, d)$ , then  $\lim_{n \rightarrow \infty} d(x_n, y_n) = d(x, y)$  exists in  $\mathbb{R}$  in the usual sense.

All convergent sequences are Cauchy, but the opposite implication generally fails. The space  $(X, d)$  is *complete* if all Cauchy sequences in it converge.

For instance, if  $\mathbb{Q} \subset \mathbb{R}$  is the set of the rational numbers and  $\{x_n\}_{n=1}^{\infty}$  is a sequence in  $\mathbb{Q}$  converging to  $\sqrt{2}$ , then (with respect to the usual distance function  $d(x, y) = |x - y|$ ) the sequence  $\{x_n\}_{n=1}^{\infty}$  is Cauchy, but not convergent, in  $\mathbb{Q}$ . Familiar examples of complete metric spaces are  $\mathbb{R}$  and  $\mathbb{C}$ , with respect to the distance  $d(x, y) = |y - x|$ .

Any metric space can be canonically imbedded in a complete one.

**THEOREM 1.1.** *Let  $(X, d)$  be a metric space. Then there exists a complete metric space  $(\tilde{X}, \tilde{d})$ , the **completion** of  $(X, d)$ , and an injective map  $i: X \rightarrow \tilde{X}$  such that:*

$$(i) \quad \tilde{d}(i(x), i(y)) = d(x, y);$$

$$(ii) \quad i(X) \text{ is dense in } \tilde{X}.$$

*The completion is unique in the following sense. For any other complete metric space  $(Z, \delta)$  endowed with an injection  $j: X \rightarrow Z$  satisfying properties (i)  $\delta(j(x), j(y)) = d(x, y)$ , and (ii)  $j(X)$  is dense in  $Z$ , there is a unique map  $F: \tilde{X} \rightarrow Z$  which is a surjective isometry,  $\delta(F(\tilde{x}), F(\tilde{y})) = \tilde{d}(\tilde{x}, \tilde{y})$ .*

**PROOF.** We start with the construction of  $(\tilde{X}, \tilde{d})$ . Let  $C$  be the set of all Cauchy sequences in  $X$  and for  $\{x_n\}, \{y_n\} \in C$  set

$$\{x_n\} \sim \{y_n\} \Leftrightarrow \lim_{n \rightarrow \infty} d(x_n, y_n) = 0.$$

The relation  $\sim$  is an equivalence relation (check it!). Write  $[\{x_n\}]$  for the equivalence class of the Cauchy sequence  $\{x_n\}$ , and let  $\tilde{X} = C / \sim$ , be the corresponding quotient space. Define

$$\tilde{d}([\{x_n\}], [\{y_n\}]) = \lim_{n \rightarrow \infty} d(x_n, y_n).$$

It is easy to see that  $\tilde{d}$  is well defined, and that it defines a distance on  $\tilde{X}$ . (check that it is independent of the representatives). Finally, for  $x \in X$  set  $i(x) = [\{x_n = x\}]$ , the class of the corresponding constant function. Properties (i) and (ii) are easily verified (exercise). Some work is needed to prove completeness, using a diagonal trick.

Let  $\{\tilde{x}_n\}$ ,  $\tilde{x}_n = [\{x_m^n\}_{m=1}^{\infty}]$ , be a Cauchy sequence in  $\tilde{X}$ . For  $k \geq 1$  select integer  $n(k) > n(k-1)$  (if  $n(k-1)$  was already selected) such that, for  $n \geq n(k)$  and  $j \geq 1$ ,

$$\frac{1}{2^k} \geq \tilde{d}(\tilde{x}_n, \tilde{x}_{n+j}) := \lim_{m \rightarrow \infty} d(x_m^n, x_m^{n+j}).$$

Select then  $m(k) > m(k-1)$  such that, for  $i = 0, 1$ ,  $m \geq m(k)$ , and  $j \geq 1$ ,

$$(*) \quad d(x_m^{n(k+i)}, x_{m+j}^{n(k+i)}) \leq \frac{1}{2^k},$$

which we can ask because each  $\{x_m^n\}_{m=1}^\infty$  is Cauchy, and

$$(**) \quad d(x_{m(k)}^{n(k)}, x_{m(k)}^{n(k+1)}) \leq \frac{2}{2^k},$$

which we can ask because  $\frac{1}{2^k} \geq \tilde{d}(\tilde{x}_n, \tilde{x}_{n+j})$ . Set then  $a_k = x_{m(k)}^{n(k)}$ .

The sequence  $\{a_k\}_{k=1}^\infty$  is Cauchy in  $X$ :

$$\begin{aligned} d(a_k, a_{k+1}) &= d(x_{m(k)}^{n(k)}, x_{m(k+1)}^{n(k+1)}) \\ &\leq d(x_{m(k)}^{n(k)}, x_{m(k)}^{n(k+1)}) + d(x_{m(k)}^{n(k+1)}, x_{m(k+1)}^{n(k+1)}) \\ &\leq \frac{2}{2^k} + \frac{1}{2^k} \end{aligned}$$

by  $(*)$  and  $(**)$ , and, by geometric sums,  $d(a_k, a_{k+j}) \leq \frac{6}{2^k}$ . Let  $a = [\{a_k\}]$ .

We want to prove that  $\tilde{d}(\tilde{x}_{n(k)}, a) \rightarrow 0$ . Using the fact that  $\lim_{n \rightarrow \infty} d(x_n, y_n) = \lim_{m \rightarrow \infty} d(x_{j_m}, y_{k_m})$  for all subsequences  $\{j_m\}$  and  $\{k_m\}$  of the positive integers, provided the initial sequences  $\{x_n\}$  and  $\{y_n\}$  are Cauchy (check it), we have

$$\begin{aligned} \tilde{d}(\tilde{x}_{n(k)}, a) &= \lim_{l \rightarrow \infty} d(x_{m(l)}^{n(k)}, x_{m(l)}^{n(l)}) \\ &\leq \limsup_{l \rightarrow \infty} d(x_{m(l)}^{n(k)}, x_{m(k)}^{n(k)}) + \lim_{l \rightarrow \infty} d(x_{m(k)}^{n(k)}, x_{m(l)}^{n(l)}) \\ &\leq \frac{1}{2^k} + \frac{6}{2^k}. \end{aligned}$$

by  $(*)$  and the fact that  $\{a_k = x_{m(k)}^{n(k)}\}$  is Cauchy (the precise value 6 is indeed unimportant).

We now come to uniqueness. Let  $(Z, \delta)$ ,  $j$  as in the hypothesis, and let  $\tilde{x} = [\{x_n : n \geq 1\}]$  be an element in  $\tilde{X}$ , where  $\{x_n\}$  is a Cauchy sequence in  $X$ . Then,  $\{i(x_n) : n \geq 1\}$  is Cauchy in  $\tilde{X}$  and  $\{j(x_n) : n \geq 1\}$  is Cauchy in  $Z$ . Define  $F(\tilde{x}) = \lim_{n \rightarrow \infty} j(x_n)$ , which exists because  $Z$  is complete. The definition is well posed, since for two equivalent Cauchy sequences  $\{x_n\}$ ,  $\{y_n\}$  in  $X$  we have

$$\delta(j(x_n), j(y_n)) = d(x_n, y_n) \rightarrow 0 \text{ as } n \rightarrow \infty,$$

hence, they have the same limit in  $(Z, \delta)$ . Moreover, for  $\tilde{x} = [\{x_n\}]$  and  $\tilde{y} = [\{y_n\}]$  in  $\tilde{X}$ ,

$$\delta(F(\tilde{x}), F(\tilde{y})) = \delta(F([\{x_n\}]), F([\{y_n\}]))$$



$$\begin{aligned}
&= \delta(\lim_{n \rightarrow \infty} j(x_n), \lim_{n \rightarrow \infty} j(y_n)) \\
&= \lim_{n \rightarrow \infty} \delta(j(x_n), j(y_n)) = \lim_{n \rightarrow \infty} d(x_n, y_n) = \lim_{n \rightarrow \infty} \tilde{d}(i(x_n), i(y_n)) \\
&= \tilde{d}(\tilde{x}, \tilde{y}).
\end{aligned}$$

The equalities from second to third, and third to fourth line, follow from Exercise 1.2 applied to  $\delta$  and  $\tilde{d}$ , respectively.

The isometry  $F: \tilde{X} \rightarrow Z$  is surjective. If  $z \in Z$ , by hypothesis there is a sequence  $\{x_n\}$  in  $X$  such that  $\delta(j(x_n), z) \rightarrow 0$ , so that  $\{x_n\}$  is Cauchy in  $X$ . Hence,  $\{i(x_n)\}$  is Cauchy in  $\tilde{X}$  and  $\lim_{n \rightarrow \infty} F(i(x_n)) = \lim_{n \rightarrow \infty} j(x_n) = z$ .  $\square$

A normed linear space is a *Banach space* if it is complete with respect to the distance induced by the norm.

Let  $(X, \|\cdot\|)$  be a normed linear space, and let  $(\tilde{X}, \tilde{d})$  be its completion with respect to the distance  $d(x, y) = \|x - y\|$  on  $X$ . We want to introduce on  $(\tilde{X}, \tilde{d})$  a structure of normed, linear space, in such a way that  $\tilde{X}$  becomes a Banach space. To this aim, we define algebraic operations between equivalence classes of Cauchy sequences,

$$[\{x_n\}] + [\{y_n\}] := [\{x_n + y_n\}], \quad \lambda[\{x_n\}] := [\{\lambda x_n\}];$$

and of the norm:

$$\|[\{x_n\}]\| := \lim_{n \rightarrow \infty} \|x_n\|.$$

**PROPOSITION 1.1.** *Sum, multiplication times scalar, and norm are well defined, and make  $\tilde{X}$  into a normed linear space. Moreover, if  $z, w$  belong to  $\tilde{X}$ , then*

$$\tilde{d}(z, w) = \|z - w\|,$$

*so that, in particular,  $(\tilde{X}, \|\cdot\|)$  is a Banach space. Also, the imbedding map  $i: X \rightarrow \tilde{X}$  is linear.*

**PROOF.** The statement can be split in a number of statements, whose proof is left to the reader.

- (a) About the sum, we have to show that (i)  $\{x_n + y_n\}$  is Cauchy if  $\{x_n\}$  and  $\{y_n\}$  are; (ii) that  $[\{x_n + y_n\}] = [\{z_n + w_n\}]$  if  $[\{x_n\}] = [\{z_n\}]$  and  $[\{y_n\}] = [\{w_n\}]$ .
- (b) Similar statements must be verified for the product of a vector with a scalar.
- (c) We have to verify that for  $[\{x_n\}]$  in  $\tilde{X}$ ,  $\lim_{n \rightarrow \infty} \|x_n\|$  exists and does not depend on the particular representative.

- (d) For  $z, w$  in  $\tilde{X}$ ,  $\tilde{d}(z, w) = \|z - w\|$ .
- (e) The imbedding  $i$  is linear.

□

EXERCISE 1.3. *Prove the assertions above.*

## 1.2. Borel measures

**1.2.1. Borel measures on the real line.** We quickly review some facts from measure theory. A good reference is Gerald B. Folland, [Real Analysis: Modern Techniques and Their Applications](#), Wiley 2nd Edition (1999).

Let  $X$  be a topological space. The *Borel  $\sigma$ -algebra*  $\mathcal{B}(X)$  is the smallest one for which the open sets are measurable. A function  $X \xrightarrow{f} Y$  from  $X$  to another topological space  $Y$  is *Borel measurable* if the preimage  $f^{-1}(A)$  of any open subset  $A$  of  $Y$  is Borel measurable. In particular, a continuous function is measurable. The composition of two Borel measurable functions is measurable (it is not true, in general, that  $g \circ f$  is measurable if  $\mathbb{R} \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R}$ , and  $f$  and  $g$  are Lebesgue measurable: Borel measurability is more stable with respect to composition). A *Borel measure* is a measure on the Borel  $\sigma$ -algebra. (Unless otherwise stated, a measure is positive).

A topological space  $X$  is *locally compact* if any point  $x$  in it has a compact neighborhood. A continuous function  $X \xrightarrow{\varphi} \mathbb{C}$  *vanishes at infinity*,  $\varphi \in C_0(X)$ , if for each  $\epsilon > 0$  there exists  $K$  compact such that  $|\varphi(x)| < \epsilon$  if  $x \in X \setminus K$ . The space  $C_0(X)$  is a Banach space with respect to the *uniform norm*,

$$\|\varphi\|_{L_0^\infty} = \sup_{x \in X} |\varphi(x)|,$$

and the space  $C_c(X)$  of the functions having compact support is dense in it.

A Borel measure on a locally compact space is a *Radon measure* if

- (i)  $\mu(K) < \infty$  if  $K$  is compact;
- (ii)  $\mu(E) = \inf\{\mu(V) : V \supseteq E, V \text{ open in } X\}$  for all measurable  $E$ ;
- (iii)  $\mu(V) = \sup\{\mu(K) : K \subseteq E, K \text{ compact in } X\}$ .

Let  $\mu$  be a finite, positive Borel measure on  $X$  which is bounded on compact sets, and define  $\Lambda : C_0(X) \rightarrow \mathbb{C}$ ,

$$(1.2.1) \quad \Lambda(\varphi) = \int_X \varphi d\mu.$$

Then

- (i)  $\Lambda$  is linear and *positive*: if  $\varphi \geq 0$ , then  $\Lambda(\varphi) \geq 0$ ;
- (ii)  $\Lambda$  is a *bounded functional*,

$$\|\Lambda\|_{C_0(X)^*} := \sup\{|\Lambda(\varphi)| : \varphi \in C_0(X)\} \leq C\|\varphi\|_{L_0^\infty}.$$

The following is a version of the *F.Riesz-Markov-Kakutani representation theorem for measures* provides a converse to the fact just stated.

**THEOREM 1.2.** [*F.Riesz-Markov-Kakutani theorem for positive measures*] *Let  $\Lambda : C_0(X) \rightarrow \mathbb{C}$  be a positive, linear functional. Then, there exists a unique Radon measure  $\mu$  such that (1.2.1) holds. Moreover,  $\mu$  is finite and  $\|\Lambda\|_{C_0(X)^*} = \mu(X)$ .*

Two positive measures  $\mu, \nu$  on the same  $\sigma$ -algebra on  $X$  are *mutually singular*,  $\mu \perp \nu$ , if  $X = A \cup B$  with  $A, B$  measurable,  $A \cap B = \emptyset$ ,  $\mu(B) = 0 = \nu(A)$ .

Radon measures on the real line, or on the unit circle, play an important role in spectral theory, and it can help having an insight on the vastity of their catalogue. *All positive Borel measures on  $\mathbb{R}$  (in fact, on  $\mathbb{C}$ , and on a large class of locally compact spaces) which are finite on the compacts are Radon measures, and viceversa. We will use the terms "finite Borel" and "finite Radon" as synonymous.* If  $\mu$  is one such measure, and it is finite, then it can be uniquely decomposed as the sum of three positive, mutually singular, finite Borel measures,

$$(1.2.2) \quad \mu = \mu_{ac} + \mu_{sc} + \mu_d,$$

where

- (i)  $d\mu_{ac} = f dm$  is the *absolutely continuous part* of  $\mu$ ,  $m$  is the Lebesgue measure and  $f \geq 0$ ,  $f \in L^1(m)$ ;
- (ii)  $\mu_d = \sum_n \lambda_n \delta_{x_n}$ , the *discrete part* of  $\mu$ , is the sum of positive multiples ( $\lambda_n > 0$ ) of point masses  $\delta_{x_n}$  at points  $x_n \in \mathbb{R}$ , with  $\sum_n \lambda_n < \infty$ ;
- (iii)  $\mu_{sc}$ , the *singular continuous part* of  $\mu$ , is singular with respect to the Lebesgue measure and it has no point mass,  $\mu_{sc}(\{x\}) = 0$  for all  $x$ .

The natural Hausdorff measure supported on the Cantor set is a nontrivial example of a continuous, singular measure. If  $\mu$  is a finite Radon measure on  $\mathbb{R}$  and  $\alpha(t) = \mu(-\infty, t]$  is the (increasing, right continuous) distribution function of  $\mu$ , then the decomposition in (1.2.2) corresponds to the decomposition  $\alpha = \alpha_{ac} + \alpha_{sc} + \alpha_d$ , where the functions on the right are increasing,

$\alpha_{ac}$  is *absolutely continuous*,  $\alpha_d$  is *pure jump*, and  $\alpha_{sc}$  is continuous and it has derivative vanishing *a.e.* with respect to Lebesgue's measure.

A *finite complex measure* on a measurable space  $(X, \mathcal{F})$  (where  $\mathcal{F}$  is a  $\sigma$ -algebra) is a map  $\mu : \mathcal{F} \rightarrow \mathbb{C}$  such that, if  $\{E_n\}_{n=1}^\infty$  is a sequence of disjoint measurable sets, then

$$\mu \left( \bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n),$$

and the series converges absolutely. A *finite, complex Borel measure* is one defined on a Borel  $\sigma$ -algebra.

Any finite complex measure can be uniquely be decomposed as a linear combination (*Jordan decomposition*) of four positive finite measures,

$$\mu = \mu_{R+} - \mu_{R-} + i(\mu_{I+} - \mu_{I-}),$$

with the requirement that  $\mu_{R+} \perp \mu_{R-}$  and  $\mu_{I+} \perp \mu_{I-}$ . The four measures are Radon if  $\mu$  is Radon. The integral with respect to  $\mu$  is defined as the linear combination of the integrals with respect to the measures in the decomposition.

The *total variation*  $|\mu|$  of a real valued measure  $\mu = \mu_{R+} - \mu_{R-}$  is the positive measure  $|\mu| = \mu_{R+} + \mu_{R-}$ ; that of a complex measure  $\mu$ , more generally, is defined by

$$|\mu|(E) = \sup \left\{ \sum_{n=1}^{\infty} |\mu(E_j)| : E = \bigcup_{n=1}^{\infty} E_n \text{ is a disjoint decomposition of } E \right\},$$

and clearly  $E$  and all  $E_n$ 's in the definition are measurable. Although it is not obvious, the two definitions agree for a real valued measure.

Complex measures intervene in the following version of F. Riesz representation theorem.

**THEOREM 1.3.** [*F.Riesz-Markov-Kakutani theorem for complex measures*] *Let  $X$  be a locally compact space, and  $\Lambda : C_0(X) \rightarrow \mathbb{C}$  be a bounded, linear functional. Then, there exists a unique finite, complex Radon measure  $\mu$  such that (1.2.1) holds. Moreover,*

$$\|\Lambda\|_{C_0(X)^*} = |\mu|(X).$$

*Conversely, (1.2.1) defines a bounded, linear functional on  $C_0(X)$  if  $\mu$  is finite, complex Radon measure.*

**1.2.2. Borel measurable functions and the Lebesgue-Hausdorff theorem.** The content of this section will only be used to prove some uniqueness statements in spectral theory, and to provide one of two proofs of Stone's theorem. I suggest you just read the statement of theorem 1.4 and of its corollary. There are properties which we will first prove for continuous functions on  $\mathbb{R}$ , and that are preserved under pointwise limits. The punchline is that such properties hold for all Borel measurable functions. My source for this material is [18], S.M. Srivastava [A Course on Borel Sets](#), Springer 1998.

**THEOREM 1.4.** *[Lebesgue-Hausdorff theorem] Let  $\mathcal{F}$  be a class of functions  $\mathbb{R} \xrightarrow{f} \mathbb{C}$  which (i) contains the continuous functions; (ii) is closed under pointwise limits: if  $f_n \in \mathcal{F}$  and  $f_n(x) \rightarrow f(x)$  for all  $x \in \mathbb{R}$ , then  $f \in \mathcal{F}$ .*

*Then,  $\mathcal{F}$  contains all Borel measurable functions.*

Clearly, there is no loss of generality in assuming that we consider functions  $\mathbb{R} \xrightarrow{f} \mathbb{R}$ , and in fact the theorem holds in a much broader context. A slight modification of the proof gives the following variant.

**COROLLARY 1.1.** *Let  $\mathcal{F}$  be a class of functions  $\mathbb{R} \xrightarrow{f} \mathbb{C}$  which (i) contains the bounded continuous functions; (ii) is closed under pointwise limits of uniformly bounded functions: if  $f_n \in \mathcal{F}$ ,  $\|f_n\|_{L_0^\infty} \leq C$ , and  $f_n(x) \rightarrow f(x)$  for all  $x \in \mathbb{R}$ , then  $f \in \mathcal{F}$ .*

*Then,  $\mathcal{F}$  contains  $L_0^\infty(\mathbb{R})$ .*

The smallest class  $\mathbf{B} = \mathbf{B}(\mathbb{R}, \mathbb{R})$  which contains  $C(\mathbb{R})$  and is closed under pointwise limits is called the *Baire class* of  $\mathbb{R}$ . The symbol  $\mathbf{B}^\infty$  is the smallest one containing bounded continuous functions, and closed under pointwise limits of uniformly bounded functions.

In order to prove theorem 1.4 we need to define a hierarchy into the Baire class. If  $\mathcal{A}$  is a class of functions  $f : \mathbb{R} \rightarrow \mathbb{R}$ , let  $\lim \mathcal{A}$  be the set of the pointwise limits of sequences of functions in  $\mathcal{A}$ . If  $\mathcal{A}$  is a class of bounded function,  $\lim^\infty \mathcal{A}$  is defined similarly, but requiring that the functions' sequence is uniformly bounded.

Let  $\omega_1$  be the smallest uncountable ordinal.

(i) Set  $\mathbf{B}_0 = C(\mathbb{R})$ .

(ii) If  $\alpha$  is a countable ordinal,

$$\mathbf{B}_\alpha = \lim \left( \bigcup_{\beta < \alpha} \mathbf{B}_\beta \right).$$

- (iii) The classes  $\mathbf{B}_\alpha^\infty$  are defined similarly, starting with  $\mathbf{B}_0^\infty = C_b(\mathbf{R})$ , the bounded and continuous functions, and using  $\lim^\infty$  instead of  $\lim$ .

With this definition,  $\mathbf{B}_1 \supseteq \mathbf{B}_0$  contains pointwise limits of continuous functions,  $\mathbf{B}_2 \supseteq \mathbf{B}_1$  contains pointwise limits of functions which are of the form

$$f(x) = \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} f_{m,n}(x),$$

where each  $f_{m,n}$  is continuous, and so on. The definition looks more interesting at *limit ordinals*<sup>1</sup> such as

$$\omega_0 = \sup\{n : n \text{ is a finite ordinal}\},$$

the first infinite ordinal. If  $\alpha$  is a limit ordinal, then  $f \in \mathbf{B}_\alpha$  if and only if there is a sequence  $f_n \in \mathbf{B}_{\alpha_n}$  with  $\alpha_n < \alpha$ , such that  $f$  is the pointwise limit of the  $f_n$ 's.

Let  $\omega_1 = \sup\{\alpha : \alpha \text{ is countable ordinal}\}$  be the first uncountable ordinal.

LEMMA 1.1. *We have the decomposition:*

$$\mathbf{B} = \bigcup_{\alpha < \omega_1} \mathbf{B}_\alpha, \text{ and } \mathbf{B}^\infty = \bigcup_{\alpha < \omega_1} \mathbf{B}_\alpha^\infty.$$

PROOF. We first prove the trivial  $\supseteq$ . Surely  $\mathbf{B}_0 = C(\mathbb{R}) \subset \mathbf{B}$ . We have to show that the definition of  $\mathbf{B}_\alpha$  from the form the preceding  $\mathbf{B}_\beta$ 's,  $\beta < \alpha$ , preserves membership in  $\mathbf{B}$  if  $\mathbf{B}_\beta \subseteq \mathbf{B}_\beta$  for all such  $\beta$ 's. Let  $f_n \in \mathbf{B}_{\alpha_n}$  with  $\alpha_n < \alpha$ , and  $f = \lim f_n$ . Then  $f$  is pointwise limit of functions in  $\mathbf{B}$ , hence it belongs to  $\mathbf{B}$ .

To prove  $\subseteq$ , it suffices to show that  $\bigcup_{\alpha < \omega_1} \mathbf{B}_\alpha$  is closed under pointwise limits. Let  $f_n \in \mathbf{B}_{\alpha_n}$  with  $\alpha_n < \omega_1$ , and  $f = \lim f_n$ . The supremum of a countable family of countable ordinals is a countable ordinal<sup>2</sup>,  $\alpha := \sup\{\alpha_n : n \geq 1\} < \omega_1$ . Hence,  $f_n \in \mathbf{B}_\alpha$  for all  $n$ , which implies that  $f \in \mathbf{B}_{\alpha+1}$ , and  $\alpha + 1 < \omega_1$ .

The proof for the class  $\mathbf{B}^\infty$  is the same. □

COROLLARY 1.2. *The Baire class  $\mathbf{B}$  is closed under the following operations.*

<sup>1</sup>Recall that an ordinal  $\alpha$  is a limit ordinal if it is not the successor of any ordinal,  $\alpha \neq \beta + 1$  for all  $\beta < \alpha$ .

<sup>2</sup>We can identify each ordinal  $\alpha$  with the set  $[0, \alpha)$  of the ordinals less than  $\alpha$ :  $\alpha$  is countable if and only if  $[0, \alpha)$  is countable of a set. The ordinal  $\sup_n \alpha_n$  is the one associated with  $\bigcup_{n \geq 1} [0, \alpha_n)$ , which is a countable set. Hence,  $\sup_{n \geq 1} \alpha_n$  is countable.

(i) If  $f, g \in \mathbf{B}$ , then  $\max(f, g), \min(f, g) \in \mathbf{B}$ .

(ii) If  $f, g \in \mathbf{B}$  and  $a, b \in \mathbb{R}$ , then  $af + bg \in \mathbf{B}$ .

Similar statements hold for  $\mathbf{B}^\infty$ .

PROOF. The two statements hold for  $\mathbf{B}_0 = C(\mathbb{R})$ . Suppose they hold in  $\mathbf{B}_\beta$  for all  $\beta \in [0, \alpha)$ , and that  $f, g \in \mathbf{B}_\alpha$ . Then,  $f = \lim f_n$  with  $f_n \in \mathbf{B}_{\alpha_n}$  ( $\alpha_n < \alpha$ ),  $g = \lim g_n$  with  $g_n \in \mathbf{B}_{\beta_n}$  ( $\beta_n < \alpha$ ), so that  $af_n + bg_n \in \mathbf{B}_{\max(\alpha_n, \beta_n)}$  and  $\lim(af_n + bg_n) = af + bg \in \mathbf{B}_\alpha$ , by definition of  $\mathbf{B}_\alpha$ . The same reasoning can be done with  $\max(f, g), \min(f, g)$ .

By transfinite induction, they hold for all  $\mathbf{B}_\alpha$  with  $\alpha < \omega_1$ . By lemma 1.1, they hold the Baire class  $\mathbf{B}$ .  $\square$

Next, we have to formulate a different definition of the Borel class.

PROPOSITION 1.2. Recall that the Borel class  $\mathcal{B} = \mathcal{B}(\mathbb{R})$  on  $\mathbb{R}$  is the smallest  $\sigma$ -algebra containing the open sets. Then,  $\mathcal{B} = \mathcal{C}$ , where  $\mathcal{C}$  is the smallest family of subsets of  $\mathbb{R}$  which contains all open sets and it is closed under countable intersections and countable disjoint unions.

PROOF. It is clear that  $\mathcal{B}$  has the two stated properties, hence  $\mathcal{B} \supseteq \mathcal{C}$ . To show the opposite inclusion, it suffices to show that  $\mathcal{C}$  is closed under complements, if  $A \in \mathcal{C}$ , then  $\mathbb{R} \setminus A \in \mathcal{C}$ . If this holds,  $\mathcal{C}$  is a  $\sigma$ -algebra containing the open sets, hence it contains  $\mathbf{B}$ . Let

$$\mathcal{D} = \{A \in \mathcal{C} : \mathbb{R} \setminus A \in \mathcal{C}\}.$$

Since all closed sets in  $\mathbb{R}$  are countable intersections of open sets (the universal notation for the class of such intersections is  $G_\delta$ ), they belong to  $\mathcal{C}$  and they complements do as well, hence,  $\mathcal{D}$  contains open and closed sets.

We next show that  $\mathcal{D}$  is closed under countable intersection. Suppose  $A_n \in \mathcal{D}$  for  $n \geq 0$ , so that  $A_n, \mathbb{R} \setminus A_n \in \mathcal{D}$ , and  $\bigcap_{n \geq 0} A_n \in \mathcal{C}$ . The sets  $B_0 = \mathbb{R} \setminus A_0$ , and  $B_n = (\mathbb{R} \setminus A_n) \cap A_1 \cap \cdots \cap A_{n-1}$  ( $n \geq 1$ ) are pairwise disjoint and belong to  $\mathcal{D}$ . Then, since  $\mathcal{C}$  is closed under disjoint unions,

$$\mathcal{C} \ni \bigcup_{n \geq 0} B_n = \mathbb{R} \setminus \left( \bigcap_{n \geq 0} A_n \right).$$

Thus,  $\bigcap_{n \geq 0} A_n \in \mathcal{D}$ .

It is now clear that  $\mathcal{D}$  is closed under countable, disjoint unions. If  $B_n \in \mathcal{D}$  are pairwise disjoint,  $n \geq 0$ , then  $\bigcup_{n \geq 0} B_n \in \mathcal{C}$  by definition, and  $\mathbb{R} \setminus (\bigcup_{n \geq 0} B_n) = \bigcap_{n \geq 0} (\mathbb{R} \setminus B_n) \in \mathcal{C}$  by the previous step. Hence,  $\bigcup_{n \geq 0} B_n \in \mathcal{D}$ .

Since  $\mathcal{D}$  contains the open sets and it is closed under countable intersections and countable unions of disjoint sets,  $\mathcal{D} = \mathcal{C}$ . Hence,  $\mathcal{C}$  is also closed under complements, hence  $\mathcal{C} = \mathcal{B}$ .  $\square$

We are now ready to prove the Lebesgue-Hausdorff theorem for characteristic functions of sets.

**PROPOSITION 1.3.** *If  $B$  is a Borel set in  $\mathbb{R}$ , then  $\chi_B \in \mathbf{B}^\infty \subset \mathbf{B}$ .*

**PROOF.** Let

$$\mathcal{U} = \{B \subset \mathbb{R} : \chi_B \in \mathbf{B}^\infty\}.$$

If  $U$  is open, there are closed sets  $F_n$  such that  $F_n \nearrow U$ , and for each of them, by Urysohn lemma, there is  $f_n : \mathbb{R} \rightarrow [0, 1]$  continuous, with  $f_n|_{F_n} = 1$  and  $f_n|_{\mathbb{R} \setminus U} = 0$ . Since  $\{f_n\}$  is uniformly bounded and  $f_n \rightarrow \chi_U$ , we have  $U \in \mathcal{U}$ .

Suppose  $\{B_n : n \geq 0\}$  is a family of disjoint subsets in  $\mathcal{U}$ , having union  $B$ , and set

$$f_n = \sum_{j=0}^n \chi_{B_j}.$$

Then,  $0 \leq f_n \leq 1$ , and  $f_n \rightarrow \chi_B$  by corollary 1.2, hence  $B \in \mathcal{U}$ .

Let  $\{B_n : n \geq 0\}$  is a family of subsets in  $\mathcal{U}$ , having intersection  $B$ , and set

$$f_n = \min_{j \leq n} \chi_{B_j}.$$

Then,  $0 \leq f_n \leq 1$ , and  $f_n \rightarrow \chi_B$  by corollary 1.2, hence  $B \in \mathcal{U}$ .

By proposition 1.2, we have that  $\mathcal{U} = \mathbf{B}^\infty$ .  $\square$

**PROOF OF THE LEBESGUE-HAUSDORFF THEOREM.** The characteristic functions of a Borel set belong to  $\mathbf{B}^\infty$  by proposition 1.3. Any (Borel) measurable function  $f$  is the pointwise limit of finite linear combinations  $f_n$  of characteristic functions of (Borel) measurable sets, hence  $f$  lies in  $\mathbf{B}$ . If  $f$  is also bounded, we can take the  $f_n$  to be uniformly bounded, hence  $f \in \mathbf{B}^\infty$ .  $\square$

### 1.3. Fourier transforms and series

In this subsection we recall without proofs some basic facts in Fourier theory. These are well known by virtually all students, but there are various equally popular normalizations for the Fourier transform and the Fourier series, and it might be helpful to have the formulas for the choice we make here.



**1.3.1. The Fourier transform.** The *Fourier transform*, defined in the first line, enjoys some properties. Below,  $\mathbf{D} = i^{-1} \frac{\partial}{\partial x}$  is the symmetric version of the derivative (the *momentum operator* of quantum mechanics).

$$(1.3.1) \quad \widehat{\varphi}(\omega) = (\mathcal{F}\varphi)(\omega) = \int_{\mathbb{R}} \varphi(x) e^{-i\omega x} dx;$$

$$(1.3.2) \quad \varphi(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\omega) e^{i\omega x} d\omega;$$

$$(1.3.3) \quad \int_{\mathbb{R}} \overline{\varphi(x)} \psi(x) dx = \frac{1}{2\pi} \int_{\mathbb{R}} \overline{\widehat{\varphi}(\omega)} \widehat{\psi}(\omega) d\omega;$$

$$(1.3.4) \quad \widehat{\mathbf{D}\varphi}(\omega) = \omega \widehat{\varphi}(\omega) = M_{\omega} \widehat{\varphi}(\omega).$$

The formula in the third line is the *Plancherel identity*, which is especially interesting when  $\varphi = \psi$ . The definition makes sense for  $f \in L^1(\mathbb{R})$ , but it can be extended to  $L^2(\mathbb{R})$  in the appropriate sense. The two central identities hold, in the appropriate sense, if  $\varphi \in L^2(\mathbb{R})$ , while the fourth holds if, e.g.,  $\varphi \in C_c^1(\mathbb{R})$ . For functions  $f$  in  $L^1(\mathbb{R})$  we the following hold:

$$(1.3.5) \quad \widehat{f} \in C_0(\mathbb{R}) \text{ (i.e. } \widehat{f} \in C(\mathbb{R}) \text{ and } \lim_{\omega \rightarrow \pm\infty} \widehat{f}(\omega) = 0);$$

$$(1.3.6) \quad \widehat{f} = 0 \implies f = 0.$$

The *convolution*  $f * g$  of  $f, g : \mathbb{R} \rightarrow \mathbb{C}$  is defined as

$$(f * g)(x) = \int_{\mathbb{R}} f(x - y) g(y) dy.$$

For  $1 \leq p \leq \infty$  *Young's inequality* holds,

$$(1.3.7) \quad \|f * g\|_{L^p} \leq \|f\|_{L^1} \|g\|_{L^p}.$$

We have the fundamental:

$$(1.3.8) \quad \widehat{f * g}(\omega) = \widehat{f}(\omega) \widehat{g}(\omega).$$

This holds if, e.g.,  $f, g \in L^2 \cup L^1$ .

The convolution between a function in  $C_0(\mathbb{R})$  and a finite Borel measure is defined in the natural way,

$$(1.3.9) \quad (\varphi * \mu)(x) = \int_{\mathbb{R}} \varphi(x - u) d\mu(u).$$

The *Fourier transform of the measure*  $\mu$  is obviously

$$\widehat{\mu}(\omega) = \int_{\mathbb{R}} e^{-i\omega x} d\mu(x),$$

the convolution-to-product relation still holds, with the same proof, if e.g.  $\varphi \in L^1(\mathbb{R}) \cap C_0(\mathbb{R})$ .

$$(1.3.10) \quad \widehat{\varphi * \mu} = \widehat{\varphi} \widehat{\mu}.$$

We might define the convolution between a finite Borel measure and a function in  $L^1(\mathbb{R})$ , or even between finite Borel measures. The definition in (1.3.9) can not be generalized naively, since a function in  $L^1(\mathbb{R})$  is defined a.e with respect to the Lebesgue measure, while the measure  $\mu$  might be orthogonal to the Lebesgue measure.

We end with some elementary, but useful formulas. Let  $(J\varphi)(x) = \varphi(-x)$  the operator "reflection with respect to the origin" in  $\mathbb{R}$ , and  $(C\varphi)(x) = \overline{\varphi(x)}$ . Then,

$$(1.3.11) \quad CJ = JC, \quad C\mathcal{F} = J\mathcal{F}C, \quad J\mathcal{F} = \mathcal{F}J.$$

**1.3.2. Fourier series.** We briefly review some useful formulas concerning the Fourier series of  $\varphi \in L^2[-\pi, \pi]$ , where integrals are taken with respect to the Lebesgue measure. We might identify  $(-\pi, \pi]$  with the *torus*  $\mathbb{T} = \{e^{it} : t \in (-\pi, \pi]\}$  (the unit circle in the complex plane), and we can think of functions on  $(-\pi, \pi]$  as of functions on  $\mathbb{T}$  ( $f(t) \equiv f(e^{it})$ ), or as  $2\pi$ -periodic functions on  $\mathbb{R}$ .

First, we state an important result.

**THEOREM 1.5.** *Let  $e_n(t) = e^{int}$ ,  $e_n : [-\pi, \pi]$ ,  $n \in \mathbb{Z}$ . Then,  $\mathcal{T} = \{(2\pi)^{-1/2}e_n : n \in \mathbb{Z}\}$  is a orthonormal basis for  $L^2[-\pi, \pi]$ .*

If you have never seen any of the numerous proofs of this result, I provide one after we have introduced the Poisson kernel on  $[-\pi, \pi]$ . It is readily verified that  $\mathcal{T}$  is a orthonormal system,

$$(1.3.12) \quad \langle e_m, e_n \rangle_{L^2} = \int_{-\pi}^{\pi} e^{-imt} e^{int} dt = \begin{cases} 2\pi & \text{if } m = n, \\ \left[ \frac{e^{i(n-m)t}}{i(n-m)} \right]_{-\pi}^{\pi} = 0 & \text{if } n \neq m. \end{cases}$$

By general Hilbert space theory, the *trigonometric system*  $\mathcal{T}$  is a orthonormal basis for a closed subspace  $H(\mathcal{T})$  of  $L^2[-\pi, \pi]$ . Theorem 1.5 says that

$$H(\mathcal{T}) = L^2[-\pi, \pi].$$

For a function  $\varphi \in L^1[-\pi, \pi]$ , we define its *Fourier coefficients*  $(\mathcal{F}\varphi)(n) = \widehat{\varphi}(n)$ , and list some of their basis properties.

$$(1.3.13) \quad \widehat{\varphi}(n) = \int_{-\pi}^{\pi} \varphi(x) e^{-inx} dx \text{ for } n \in \mathbb{Z};$$

$$(1.3.14) \quad \varphi(x) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \widehat{\varphi}(n) e^{inx};$$

$$(1.3.15) \quad \langle \varphi, \psi \rangle_{L^2[-\pi, \pi]} = \frac{1}{2\pi} \langle \widehat{\varphi}, \widehat{\psi} \rangle_{\ell^2(\mathbb{Z})};$$

$$(1.3.16) \quad \widehat{\mathbf{D}\varphi}(n) = n\widehat{\varphi}(n) = (M_n\widehat{\varphi})(n).$$

The first equality is the definition of the *Fourier coefficients* of  $\varphi$ . The second holds for  $\varphi \in H(\mathcal{T})$ , and the equality has to be interpreted in the  $L^2$  norm. The third relation is equivalent to the second, and it holds if either  $\varphi$  or  $\psi$  belong to  $H(\mathcal{T})$ . while the last relation holds if  $\varphi \in C_{per}^1$ , the space of the  $C^1$ ,  $2\pi$ -periodic functions  $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ . The second and third relations say that

$$\mathcal{F} : H(\mathcal{T}) \rightarrow \ell^2(\mathbb{Z})$$

is a bijective isometry, but for a factor,

$$\|\mathcal{F}\varphi\|_{\ell^2}^2 = 2\pi \|\varphi\|_{L^2[-\pi, \pi]}^2.$$

This is in perfect agreement with the Fourier transform, and it can be extended to the context of *locally compact Abelian groups*, and to a more general non-commutative universe, with some highly nontrivial adjustments. For the commutative theory, see e.g. [Fourier Analysis on Groups \(1962\)](#), Wiley-Interscience, by Walter Rudin.

The fourth relation in (1.3.13) can be used to give the Fourier series expansion a more concrete meaning. We give here a statement having an elementary proof.

LEMMA 1.2. *Let  $\varphi \in H(\mathcal{T}) \cap C_{per}^2[-\pi, \pi]$ . Then,*

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

*and the series converges uniformly.*

Here  $C_{per}^2([-\pi, \pi])$  is the space of  $2\pi$ -periodic functions in  $C^2$ .

PROOF. The series converges in  $L^2$ , and it converges to  $\varphi$  because  $\varphi \in H(\mathcal{T})$ . Hence a subsequence converges a.e.,

$$(1.3.17) \quad \varphi(t) = \lim_{j \rightarrow +\infty} \frac{1}{2\pi} \sum_{n=-N_j}^{N_j} \widehat{\varphi}(n) e^{int} \text{ a.e. } t$$

for some sequence  $N_j \nearrow \infty$ , and  $|\widehat{\varphi''}(n)| \leq \|\varphi''\|_{L^1[-\pi, \pi]}$ , hence the sequence  $\{\widehat{\varphi''}(n) : n \in \mathbb{Z}\}$  is bounded. On the other hand,

$$-\widehat{\varphi''}(n) = \widehat{\mathbf{D}^2 \varphi}(n) = n^2 \widehat{\varphi}(n),$$

hence, the series

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

converges uniformly. By (1.3.17),  $S_\infty(\varphi)$  converges uniformly to  $\varphi$ .  $\square$

We can define the convolution of functions on  $[-\pi, \pi]$  by thinking of them as  $2\pi$ -periodic functions (or functions on  $\mathbb{T}$ ):

$$(f * g)(t) = \int_{-\pi}^{\pi} f(t-s)g(s)ds.$$

LEMMA 1.3. *If  $f \in H(\mathcal{T})$  and its Fourier series converges absolutely (e.g. if  $f \in C_{per}^2[-\pi, \pi]$ ), and  $g \in L^1[-\pi, \pi]$ , then  $f * g \in H(\mathcal{T})$  and  $\widehat{f * g}(n) = \widehat{f}(n)\widehat{g}(n)$ .*

PROOF.

$$\begin{aligned} (f * g)(t) &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{in(s-t)} g(s) ds \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}(n) \int_{-\pi}^{\pi} e^{-ins} g(s) ds e^{int} \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}(n) \widehat{g}(n) e^{int}, \end{aligned}$$

which converges absolutely to a function in  $H(\mathcal{T})$ , having Fourier coefficients  $\widehat{f}(n)\widehat{g}(n)$ .  $\square$

#### 1.4. Basic theory of harmonic functions

A function  $U : \Omega \rightarrow \mathbb{C}$  is *harmonic* on the open subset  $\Omega \subseteq \mathbb{C}$  if  $U \in C^2(\Omega)$  and

$$(1.4.1) \quad \Delta U(z) := \partial_{xx} U(z) + \partial_{yy} U(z) = 0$$

for all  $z \in \Omega$ . In this two-dimensional context, it is well known that, if  $\Omega$  is also simply connected, then *a real valued  $U$  is harmonic if and only if*

it is the real part of a function  $F$  holomorphic in  $\Omega$  (the "if" holds with no special assumption on  $\Omega$ ). The proof of this fact easily reduces to the

Cauchy-Riemann equations for a holomorphic  $F = U + iV$ ,  $\begin{cases} \partial_x U = \partial_y V \\ \partial_y U = -\partial_x V. \end{cases}$

Equation (1.4.1) implies that  $-\partial_y U dx + \partial_x U dy$  is a closed 1-form, which is then exact. Let  $V$  be one of its potentials. Then,  $\partial_x V = -\partial_y U$  and  $\partial_y V = \partial_x U$ , which are the Cauchy-Riemann equations. Viceversa, that the Cauchy-Riemann equations imply (1.4.1) for both  $U$  and  $V$  is easily verified.

We list here the few facts that we will need.

- (i) **[Series expansion]** If  $U$  is harmonic in  $D(a, R) = \{z : |z - a| < R\}$ , with  $0 < R \leq \infty$ , then  $U$  can be written in the form

$$U(a + re^{it}) = \sum_{n=-\infty}^{+\infty} a_n r^{|n|} e^{int},$$

for  $0 \leq r < R$ . Moreover, if  $0 < r < R$ , then

$$(1.4.2) \quad a_n = \frac{1}{2\pi r^{|n|}} \int_0^{2\pi} U(a + re^{it}) e^{-int} dt.$$

PROOF. We can assume  $a = 0$  and that  $U = \operatorname{Re} F$  is the real part of a holomorphic  $F$ . If  $F(z) = \sum_{n=0}^{\infty} b_n z^n$  is the Taylor expansion of  $F$ , then

$$U(re^{it}) = \frac{F(e^{it}) + \overline{F(e^{it})}}{2} = \operatorname{Re}(b_0) + \sum_{n=-\infty}^{-1} \overline{b_{-n}} r^{|n|} e^{int} + \sum_{n=1}^{\infty} b_n r^n e^{int},$$

The formula for the coefficients follows by integrating term by term.  $\square$

- (ii) **[Liouville theorem]** A bounded harmonic function on  $\mathbb{C}$  is constant.

PROOF. For  $n \neq 0$ , by (1.4.2),

$$|a_n| \leq \frac{\|U\|_{L^\infty}}{r^{|n|}} \rightarrow 0$$

as  $r \rightarrow \infty$ .  $\square$

- (iii) **[Identity principle]** If  $U$  is harmonic on  $\Omega$ , connected, and if  $U$  vanishes on an open set, then it vanishes identically. More generally, if  $U$  is constant on an open subset of  $\Omega$ , then it is globally constant.

PROOF. Let

$$Z = \{a \in \Omega : \exists D(a, r) \subseteq \Omega \text{ such that } U \text{ vanishes on } D(a, r)\}.$$

By (1.4.2), if  $a \in Z$ , then all series coefficients of  $U$  vanish,  $Z$  contains the largest disc centered at  $a$  which is contained in  $\Omega$ . The set  $Z$  is nonempty and open. Let  $Z \ni a_n \rightarrow b \in \Omega$ , and let  $D(b, r) \subset \Omega$ . Pick  $a_n$  such that  $|b - a_n| \leq r/4$ . Then  $b \in D(a_n, r/2) \subseteq Z$ . Being  $\Omega$  connected,  $Z = \Omega$ .  $\square$

A PROOF USING THE OPEN MAPPING THEOREM. If  $\Omega$  is simply connected, then  $U = \operatorname{Re} F$  with  $F$  holomorphic in  $\Omega$ . Since  $F(D(a, r))$  fails to be open, then  $F$ , hence  $U$ , is constant. Suppose  $\Omega$  is just connected, and let  $b \in \Omega$ . Consider a simple, broken line  $l$  in  $\Omega$  which joins  $a$  and  $b$ , and let  $A = \{z \in \Omega : \operatorname{dist}(z, l) < \epsilon\}$ . If  $\epsilon$  is small, then  $A$  is simply connected (exercise) and contained in  $\Omega$ . The previous case applies, then  $U(b) = U(a) = 0$ .  $\square$

- (iv) [**Mean value property.**] *If  $U$  is harmonic in  $\Omega$  and  $\operatorname{Cl} D(a, r) \subset \Omega$ , then*

$$U(a) = \frac{1}{2\pi} \int_0^{2\pi} U(a + re^{it}) dt.$$

PROOF. It follows from (i).  $\square$

- (v) [**Maximum principle**] *Let  $U$  be real valued and continuous on  $\Omega$ , open and connected. If  $U$  satisfies the mean value property and it achieves its maximum, then it is constant.*

PROOF. Let  $M$  be the maximum of  $U$ , and  $Z = \{a \in \Omega : U(z) = M\}$ . By the mean value property and intuition, if  $a \in Z$  and  $D(a, r) \subseteq \Omega$ , then  $D(a, r) \subseteq Z$ . This shows that  $Z$  is open, but  $Z$  is also closed in the relative topology of  $\Omega$ , then  $Z = \Omega$ .  $\square$

- (vi) [**Local maximum principle**] *Let  $U$  be real valued and harmonic on  $\Omega$ , open and connected. If  $U$  has a local maximum in  $\Omega$ , then it is constant.*

PROOF. Suppose  $a$  is a maximum in  $D(a, r) \subset \Omega$ , and let  $F = U + iV$  be holomorphic in  $D(a, r)$ . Then,  $F(D(a, r)) \subseteq C := \{w = u + iv : u \leq U(a)\}$  and  $F(a) \in \partial C$ , which contradicts the open mapping theorem. Thus  $F$  is constant, hence  $U$  is constant on  $D(a, r)$ . By the identity principle,  $U$  is constant.  $\square$

(vii) [**Poisson representation**] If  $U$  is harmonic in  $\Omega$  and  $ClD(a, r) \subset \Omega$ , then for  $z \in D(a, r)$ :

$$(1.4.3) \quad U(a + z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{r^2 - |z|^2}{|re^{it} - z|^2} U(a + re^{it}) dt.$$

The expression

$$P_z^r(e^{it}) = \frac{1}{2\pi} \frac{r^2 - |z|^2}{|re^{it} - z|^2} dt$$

is the *Poisson kernel* at  $z$ .

PROOF. It suffices to show it for  $U = F$  holomorphic, and  $a = 0$ . We start with Cauchy formula,

$$F(z) = \frac{1}{2\pi i} \int_{|w|=r} \frac{F(w)}{w - z} dw = \frac{1}{2\pi} \int_0^{2\pi} \frac{re^{it}}{re^{it} - z} F(re^{it}) dt.$$

The difference between the kernel in the last expression and the Poisson kernel is

$$\begin{aligned} \frac{re^{it}}{re^{it} - z} - \frac{r^2 - |z|^2}{|re^{it} - z|^2} &= \frac{re^{it}(re^{-it} - \bar{z}) - (r^2 - |z|^2)}{(re^{it} - z)(re^{-it} - \bar{z})} \\ &= \frac{|z|^2 - \bar{z}re^{it}}{(re^{it} - z)(re^{-it} - \bar{z})} \\ &= -\frac{\bar{z}}{re^{-it} - \bar{z}} \\ &= -\sum_{n=0}^{\infty} \left(\frac{\bar{z}}{r}\right)^{n+1} e^{i(n+1)t}. \end{aligned}$$

Integrating term by term, we see that

$$\frac{1}{2\pi} \int_0^{2\pi} \left[ \frac{re^{it}}{re^{it} - z} - \frac{r^2 - |z|^2}{|re^{it} - z|^2} \right] F(re^{it}) dt = 0,$$

hence that

$$F(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{re^{it}}{re^{it} - z} F(re^{it}) dt = \frac{1}{2\pi} \int_0^{2\pi} \frac{r^2 - |z|^2}{|re^{it} - z|^2} F(re^{it}) dt,$$

as wished. Since the Poisson kernel is positive (thus real), the relation holds for the real and the imaginary part of  $F$  separately.  $\square$

There are indeed other ways to guess the Poisson kernel and to prove the Poisson representation. Here is one using Fourier series, starting from (i). It does not really make use of the completeness of the trigonometric system, since in the Plancherel formula below one function (there denoted by  $Q$ ) belongs to the span  $H(\mathcal{T})$  of the trigonometric system.

PROOF. Let  $a = 0$  and  $z = \rho e^{is}$ , with  $0 \leq \rho < r$ , and  $\text{Cl}D(0, r) \subset \mathbb{C}$ . By Plancherel's formula,

$$\begin{aligned} U(\rho e^{is}) &= \sum_{n=-\infty}^{+\infty} a_n \rho^{|n|} e^{ins} = \sum_{n=-\infty}^{+\infty} a_n r^{|n|} \overline{(\rho/r)^{|n|} e^{-is}} \\ &= 2\pi \langle Q, U_r \rangle_{L^2[0, 2\pi]}, \end{aligned}$$

where

$$\begin{aligned} U_r(e^{it}) &= \frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} a_n r^{|n|} e^{int} \\ &= \frac{1}{2\pi} U(re^{it}), \end{aligned}$$

and (by computing geometric series)

$$(1.4.4) \quad Q(e^{it}) = \frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} (\rho/r)^{|n|} e^{-ins} e^{int}$$

$$(1.4.5) \quad = \frac{1}{2\pi} \left[ \frac{1}{1 - \frac{\rho}{r} e^{i(t-s)}} + \frac{1}{1 - \frac{\rho}{r} e^{i(t-s)}} - 1 \right]$$

$$(1.4.6) \quad = \frac{1}{2\pi} \frac{r^2 - \rho^2}{|re^{it} - \rho e^{is}|^2}.$$

Summarizing,

$$U(\rho e^{is}) = 2\pi \langle Q, U_r \rangle_{L^2[0, 2\pi]} = \frac{1}{2\pi} \int_0^{2\pi} \frac{r^2 - \rho e^{is}}{|re^{it} - \rho e^{is}|^2} F(re^{it}) dt,$$

as promised. □

(viii) **[The Poisson extension of a continuous function on the circle]** Let  $\varphi = \varphi(t) : \partial D(a, r)$  be continuous, and define

$$(1.4.7) \quad U(a + \rho e^{is}) = \begin{cases} \frac{1}{2\pi} \int_0^{2\pi} \frac{r^2 - \rho^2}{|re^{it} - \rho e^{is}|^2} \varphi(a + re^{it}) dt & \text{if } 0 \leq \rho < r, \\ \varphi(re^{is}) & \text{if } \rho = r. \end{cases}$$



Then,  $U$  is harmonic in  $D(a, r)$  and continuous on  $\text{Cl}D(a, r)$ .

The proof is similar to that of proposition 1.4 below.

PROOF. We have the following three properties.

(a) The function  $z \mapsto P_z^r(e^{is})$  is harmonic and positive in  $D(0, r)$ .

(b) Consider the function  $s \mapsto P_\rho^r(e^{is})$  on  $[-\pi, \pi]$ . Then,

$$\int_{-\pi}^{\pi} P_\rho^r(e^{is}) ds = 1.$$

(c) For each  $\delta > 0$ ,

$$\lim_{\rho \rightarrow r} \sup_{\delta \leq |s| \leq \pi} P_\rho^r(e^{is}) = 0.$$

(a) Positivity is clear, and harmonicity holds because, by (1.4.4),

$$(1.4.8) \quad P_z^r(e^{is}) = \frac{1}{2\pi} \frac{r^2 - |z|^2}{|re^{is} - z|^2} = \sum_{n=0}^{\infty} r^{-|n|} e^{-ins} z^n + \sum_{n=1}^{\infty} r^{-|n|} e^{ins} \bar{z}^n,$$

and the functions  $z^n, \bar{z}^n$  are harmonic.

(b) Integrate the series (1.4.8) term by term.

(c) We have

$$\begin{aligned} \frac{r^2 - \rho^2}{|re^{is} - \rho|^2} &= \frac{r^2 - \rho^2}{r^2 - 2r\rho \cos(s) + \rho^2} = \frac{r^2 - \rho^2}{(r - \rho)^2 + 2r\rho[1 - \cos(s)]} \\ &\leq \frac{r^2 - \rho^2}{(r - \rho)^2 + 2r\rho[1 - \cos(\delta)]} \\ &\rightarrow 0 \text{ as } \rho \rightarrow r. \end{aligned}$$

Differentiating with respect to  $z = \rho e^{is}$  under the integral on the right side of (1.4.7), and using (a), we see that  $U$  is harmonic. We only have to show continuity at points  $a + re^{it}$ . We can assume  $a = 0$ , and verify continuity at  $r$ . Fix  $\epsilon > 0$ . We estimate separately the summands on the right of:

$$|U(\rho e^{is}) - \varphi(r)| \leq |U(\rho e^{is}) - \varphi(re^{is})| + |\varphi(re^{is}) - \varphi(r)|.$$

The second summand can be made less than  $\epsilon$  by choosing  $|s| \leq \eta(\epsilon)$ , by the uniform continuity of  $\varphi$ . For the first, using (b),

$$\begin{aligned} |U(\rho e^{is}) - \varphi(\rho e^{is})| &= \frac{1}{2\pi} \left| \int_{-\pi}^{\pi} \frac{r^2 - \rho^2}{|re^{it} - \rho e^{is}|^2} [\varphi(re^{it}) - \varphi(\rho e^{is})] dt \right| \\ &\leq \frac{\epsilon}{2\pi} \int_{|s-t| \leq \eta(\epsilon)} \frac{r^2 - \rho^2}{|re^{it} - \rho e^{is}|^2} dt \\ &\quad + \frac{2\|\varphi\|_{L^\infty}}{2\pi} \int_{|s-t| \geq \eta(\epsilon)} \frac{r^2 - \rho^2}{|r - \rho e^{i(s-t)}|^2} dt \\ &\leq \epsilon + \frac{\|\varphi\|_{L^\infty}}{\pi} \epsilon, \end{aligned}$$

if we choose  $\rho(\epsilon) \leq \rho < r$ , by (c).  $\square$

- (ix) [**The completeness of the trigonometric system**] We prove here theorem 1.5. Let  $f \in L^2[-\pi, \pi]$  and fix  $\epsilon > 0$ . We want to show that there is a trigonometric polynomial  $p = p(e^{it})$  such that  $\|f - p\|_{L^2} \leq C\epsilon$ . By density of  $C_{per}[-\pi, \pi]$  in  $L^2[-\pi, \pi]$ , there is a continuous function  $\varphi$  such that  $\|f - \varphi\|_{L^2} \leq \epsilon$ . Consider now the Poisson extension of  $\varphi$  in theorem 1.5 (with  $a = 0$  and  $r = 1$ ), and set  $U_\rho(e^{it}) = U(\rho e^{it})$ . There is  $\rho = \rho(\epsilon)$  such that

$$\|U_\rho - \varphi\|_{L^2} \leq \sqrt{2\pi} \|U_\rho - \varphi\|_{L^\infty} \leq \epsilon,$$

since by our estimates the convergence of  $U_\rho$  to  $\varphi$  is uniform.

Last, we have that

$$U(\rho e^{is}) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{\varphi}(n) \rho^{|n|} e^{ins}.$$

Let  $p(e^{is}) = \frac{1}{2\pi} \sum_{n=-N}^N \widehat{\varphi}(n) \rho^{|n|} e^{ins}$ , which is a trigonometric polynomial. Then

$$\|U_\rho - p\|_{L^\infty} \leq \frac{\|\varphi\|_{L^1}}{2\pi} \sum_{|n| > N} \rho^{|n|} \leq \epsilon,$$

provided we choose  $N$  large enough.

- (x) [**The inverse mean value property**] Let  $V$  be continuous on  $\Omega$ , open, and suppose it satisfies the mean value property. Then,  $V$  is harmonic in  $\Omega$ .

PROOF. Consider  $\text{Cl}D(a, r) \subset \Omega$ , and consider the Poisson extension  $U$  of the restriction of  $V$  to  $\partial D(a, r)$ , which is harmonic in  $D(a, r)$ , and continuous in  $\text{Cl}D(a, r)$  by (viii). The function  $V - U$  satisfies the mean value property in  $D(a, r)$  by (iv), hence its maximum is achieved at the boundary by (v). Since the boundary values vanish,  $V \leq U$ . The same reasoning applies to  $U - V$ , then  $V = U$  is harmonic in  $D(a, r)$ .  $\square$

(xi) [**The reflection principle**] Let  $\Omega \subseteq \mathbb{C}_+$  be a domain such that  $\partial\Omega \cap \mathbb{R}$  is a finite union of the closures of the open intervals  $I_1, \dots, I_n$ , and let  $U : \Omega \rightarrow \mathbb{R}$  be harmonic and such that  $\lim_{z \rightarrow x} U(z) = 0$  for all  $x \in \partial\Omega \cap \mathbb{R}$ . Let  $\bar{\Omega} = \{\bar{z} : z \in \Omega\}$  be the reflection of  $\Omega$  in the real axis, and define  $U(\bar{z}) = -U(z)$  for  $\bar{z} \in \bar{\Omega}$ , and  $U = 0$  on  $I_1 \cup \dots \cup I_n$ . Then,  $U$  is harmonic on  $\Omega \cup \bar{\Omega} \cup I_1 \cup \dots \cup I_n$ .

PROOF. We apply (x). The extended function is harmonic in  $\Omega \cup \bar{\Omega}$ , hence it satisfies the mean value property there. We only have to verify the mean value property at the points  $x \in I_1 \cup \dots \cup I_n$ , where the function vanishes: the property there holds by symmetry.  $\square$

### 1.5. The Dirichlet problem for the Laplace equation in the upper half plane

Let  $f : \mathbb{R} \rightarrow \mathbb{C}$  continuous and bounded. A solution of the *Dirichlet problem* for the *Laplace equation* in the upper half plane  $\mathbb{C}_+ = \{z = x + iy : y > 0\}$  with *boundary conditions*  $f$  is a continuous function  $U : \mathbb{C}_+ \cup \mathbb{R} \rightarrow \mathbb{C}$  such that

$$(1.5.1) \quad \Delta U(z) = (\partial_{xx} + \partial_{yy})U(x + iy) = 0 \text{ for } z \in \mathbb{C}_+$$

$$(1.5.2) \quad U(x) = f(x) \text{ for } x \in \mathbb{R}$$

$$(1.5.3) \quad U \text{ is bounded.}$$

For  $z = x + iy \in \mathbb{C}_+$ , let

$$(1.5.4) \quad P_y(x) = \frac{1}{\pi} \frac{y}{x^2 + y^2}.$$

be the *Poisson kernel*.

PROPOSITION 1.4. The unique solution of 1.5.1 is given by

$$(1.5.5) \quad U(x + iy) = (P_y * f)(x) = \frac{1}{\pi} \int_{\mathbb{R}} f(u) \frac{y}{(x - u)^2 + y^2} du.$$

There are various ways to guess the formula (1.5.5) and the form of the Poisson kernel, which are interesting in that they point to extensions and generalizations. We provide one of them here below, and two more after the proof of proposition 1.4.

**The Poisson kernel via Fourier transforms.** By Fourier inversion formula, also assuming for the moment that  $f \in L^1(\mathbb{R})$ ,

$$f(x) = \int_{\mathbb{R}} \widehat{f}(\omega) e^{i\omega x} d\omega$$

is the superposition of the bounded functions  $x \mapsto e^{i\omega x}$  (with  $\omega \in \mathbb{R}$  fixed). If we extend each of them to a bounded harmonic function on the upper half-plane, by superposition we can extend  $f$  to the same. A natural guess is  $z \mapsto e^{i\omega z} = e^{i\omega x} e^{-\omega y}$ , which is holomorphic, hence harmonic, for  $z \in \mathbb{C}$ . We have boundedness on  $\mathbb{C}_+$ , however, only when  $\omega \geq 0$ . When  $\omega < 0$ , the function  $z \mapsto e^{i\omega \bar{z}}$  is bounded on  $\mathbb{C}_+$  and anti-holomorphic, hence harmonic. We are then led to look for a solution to (1.5.1) having the form

$$(1.5.6) \quad U(x + iy) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\omega) e^{-y|\omega|} e^{i\omega x} d\omega = (P_y * f)(x),$$

where

$$(1.5.7) \quad P_y(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-y|\omega|} e^{i\omega x} d\omega$$

$$(1.5.8) \quad = \frac{1}{2\pi} \int_0^\infty e^{iz\omega} d\omega + \frac{1}{2\pi} \int_{-\infty}^0 e^{i\bar{z}\omega} d\omega$$

$$(1.5.9) \quad = \frac{1}{2\pi} \left( \left[ \frac{e^{iz\omega}}{iz} \right]_{\omega=0}^{+\infty} + \left[ \frac{e^{i\bar{z}\omega}}{i\bar{z}} \right]_{\omega=-\infty}^0 \right)$$

$$(1.5.10) \quad = \frac{1}{2\pi i} \left( -\frac{1}{z} + \frac{1}{\bar{z}} \right) = \frac{z - \bar{z}}{2\pi i |z|^2}$$

$$(1.5.11) \quad = \frac{1}{\pi} \frac{y}{x^2 + y^2},$$

which is the formula we wanted.

The *Poisson kernel* satisfies the following crucial properties:

- (i)  $P_y > 0$  and  $\int_{\mathbb{R}} P_y(x) dx = 1$ ;
- (ii) for each  $\delta > 0$ ,  $P_y(x) \rightarrow 0$  as  $y \rightarrow 0$  uniformly in  $|x| \geq \delta$ ;

- (iii)  $z = x + iy \mapsto P_y(x)$  is *harmonic* on  $\mathbb{C}_+$ , and the same holds for  $z \mapsto P_y(x-u)$  for each fixed real  $u$ . This means that  $[\partial_{xx} + \partial_{yy}](P_y(x)) = 0$  for all  $z = x + iy \in \mathbb{C}_+$ .

The first two properties say that  $\{P_y\}_{y>0}$  is an *approximation of the identity* on  $\mathbb{R}$ , whose special nature is encoded in the third. Property (i) follows from the expression for  $P_y$  in the "time"  $x$  variable and from that in the "frequency"  $\omega$  one. Property (ii) is easily verified, and property (iii) follows by differentiating (1.5.7) under the integral sign.

PROOF OF PROPOSITION 1.4. We start by showing that the expression in (1.5.5) is a solution to the Dirichlet problem. By differentiating under the integral and property (iii) of the Poisson kernel,  $U$  is harmonic on  $\mathbb{C}_+$ . It is bounded by property (i) and an elementary estimate. We have to show that for all  $x$  in  $\mathbb{R}$ ,

$$(1.5.12) \quad \lim_{\mathbb{C}_+ \ni z \rightarrow x} U(z) = f(x).$$

We estimate each of the two summands on the right of

$$|U(x + iy) - f(x_0)| \leq |U(x + iy) - U(x_0 + iy)| + |U(x_0 + iy) - f(x_0)|.$$

Fix  $\epsilon > 0$ . For the first term, fix  $\delta \leq 1$  such that  $\sup_{|u| \leq 1} |f(x-u) - f(x_0-u)| \leq \epsilon$  if  $|x - x_0| \leq \delta$ , so that (by uniform continuity of  $f$  on compact intervals)

$$\begin{aligned} |U(x + iy) - U(x_0 + iy)| &= \frac{1}{\pi} \left| \int_{\mathbb{R}} [f(x-u) - f(x_0-u)] \frac{y}{u^2 + y^2} du \right| \\ &\leq \frac{\sup_{|u| \leq 1} |f(x-u) - f(x_0-u)|}{\pi} \\ &\quad + \frac{2\|f\|_{L_0^\infty}}{\pi} \int_{|u| \geq 1} \frac{y}{u^2 + y^2} du \\ &\leq \frac{\epsilon}{\pi} + \frac{2\|f\|_{L_0^\infty}}{\pi} \epsilon, \end{aligned}$$

provided  $0 < y < y_\epsilon$ .

For the second term,

$$\begin{aligned} |U(x_0 + iy) - f(x_0)| &= \frac{1}{\pi} \left| \int_{\mathbb{R}} [f(x_0-u) - f(x_0)] \frac{y}{u^2 + y^2} du \right| \\ &\leq \frac{\epsilon}{\pi} + \frac{2\|f\|_{L_0^\infty}}{\pi} \int_{|u| \geq \delta} \frac{y}{u^2 + y^2} du \\ &\leq \frac{\epsilon}{\pi} + \frac{2\|f\|_{L_0^\infty}}{\pi} \epsilon \end{aligned}$$

provided  $0 < y < y'_\epsilon$ . Overall, if  $0 < y < \max(y_\epsilon, y'_\epsilon)$  and  $|x - x_0| \leq \delta$ , then  $|U(x + iy) - f(x_0)| \leq C\epsilon$ .

To show uniqueness it suffices to show that a bounded function  $U$  which is harmonic in  $\mathbb{C}_+$ , continuous on  $\text{Cl}\mathbb{C}_+$ , and that it vanishes on  $\mathbb{R}$ , has to be constant. Extend such function to one which is defined on the whole of  $\mathbb{C}$  by setting  $U(\bar{z}) = -U(z)$ . By the reflection principle, it is harmonic and bounded on  $\mathbb{C}$ , hence it has to vanish identically by Liouville's theorem.  $\square$

**The Poisson kernel via (at this point, formal) spectral theory.** The operator  $\Delta = \partial_{xx}$  is symmetric on the domain  $C_c^2(\mathbb{R})$  which is dense in  $L^2(\mathbb{R})$ , and  $-\Delta$  is positive,

$$\langle -\Delta\varphi, \varphi \rangle_{L^2} = \int_{\mathbb{R}} |\varphi'(x)|^2 dx \geq 0.$$

We are looking for a function  $U : [0, \infty) \rightarrow L^2(\mathbb{R})$  such that

$$(1.5.13) \quad U''(y) + \Delta U(y) = 0,$$

$$(1.5.14) \quad U(0) = f,$$

$$(1.5.15) \quad U(+\infty) = 0.$$

Interpreting  $\Delta$  as a "negative constant", the first equation in 1.5.13 has formal solution

$$U(y) = e^{-\sqrt{-\Delta}y}A + e^{\sqrt{-\Delta}y}B,$$

with  $A, B$  "constants" in  $L^2(\mathbb{R})$ . Imposing the boundary conditions we have  $B = 0$  (at least on an intuitive level), and  $A = f$ . Thus,

$$(1.5.16) \quad U(y) = e^{-\sqrt{-\Delta}y}f,$$

or, thinking of  $U$  as a function of  $(x, y)$ ,

$$U(x, y) = \left[ e^{-\sqrt{-\Delta}y} f \right] (x).$$

This simple, so far formal, calculation has the advantage that it works for any positive operator  $L$  instead of  $-\Delta$ . For instance,  $L$  might be minus the Laplace-Beltrami operator on a geodesically complete Riemannian manifold, or the generator of a random walk on a graph, or minus a fractional Laplacian, etcetera.

In our case, the Fourier transform provides a "spectral representation" for the operator  $-\Delta$ ;

$$[\mathcal{F}(-\Delta f)](\omega) = \omega^2 [\mathcal{F}f](\omega).$$

which is another way to say that  $-\Delta$  acts like multiplication times  $\omega^2$  on the frequency side. The operator  $e^{-y\sqrt{-\Delta}}$ , then, acts as multiplication times  $e^{-y|\omega|}$ , and we are back to the previous deduction of the the Poisson kernel's expression.

**The Poisson kernel via the fundamental solution for the Laplace operator.** A third way to construct Poisson kernels is by means of Green's formula and the fundamental solution for the Laplace operator. Let  $G(z, w) = \frac{1}{2\pi} \log |z - w|^{-1}$ ,  $z, w \in \mathbb{C}$ , which satisfies

$$(1.5.17) \quad \Delta_z G(z, w) = \delta_z(w),$$

$$(1.5.18) \quad G(z, w) = O_{z \rightarrow \infty}(\log |z|).$$

The first equation means that

$$\int_{\mathbb{C}} G(z, w) \Delta \varphi(w) du dv = \varphi(z)$$

for all  $\varphi \in C_c^2(\mathbb{C})$ .

Let  $\Omega$  be a bounded domain in  $\mathbb{C}$  with smooth boundary  $\partial\Omega$ , and denote by  $\nu$  the external unit normal. Let  $D(z, \epsilon)$  be the disc with center  $z$  and radius  $\epsilon$ . Suppose  $H(z, w)$  is defined for  $z, w \in \text{Cl}\Omega$ ,  $z \neq w$ , and has the properties:

$$(1.5.19) \quad \Delta_w H(z, w) = 0 \text{ if } z \neq w;$$

$$(1.5.20) \quad H(z, w) = O_{w \rightarrow z}(\log |z - w|);$$

$$(1.5.21) \quad \partial_r H(z, z + re^{it}) = \frac{1}{2\pi r} (1 + o_{r \rightarrow 0}(1));$$

$$(1.5.22) \quad H(z, \xi) = 0 \text{ if } \xi \in \partial\Omega \text{ and } z \in \Omega.$$

The first three properties also hold for  $G$ , while the fourth is linked with the domain  $\Omega$ . If a function  $U : \text{Cl}\Omega \rightarrow \mathbb{C}$  is  $C^2$  and  $\Delta U = 0$  in  $\Omega$ , then, by Green's formula,

$$\begin{aligned} 0 &= \int_{\Omega \setminus D(z, \epsilon)} [H(z, w) \Delta U(w) - \Delta_w H(z, w) U(w)] du dv \\ &= \int_{\partial\Omega - \partial D(z, \epsilon)} [H(z, \zeta) \partial_\nu U(\zeta) - \partial_\nu H(z, \zeta) U(\zeta)] d\sigma(\zeta) \\ &= \int_{\partial D(z, \epsilon)} \partial_\nu H(z, \zeta) U(\zeta) d\sigma(\zeta) - \int_{\partial\Omega} \partial_\nu H(z, \zeta) U(\zeta) d\sigma(\zeta) \\ &\rightarrow U(z) - \int_{\partial\Omega} \partial_\nu H(z, \zeta) U(\zeta) d\sigma(\zeta) \end{aligned}$$

as  $\epsilon \rightarrow 0$ . (We did not use the second equation in (1.5.19), which becomes relevant when solving the Poisson equation: the nonhomogeneous version of the Laplace equation). At the end of the day, we have the inspiring formula:

$$(1.5.23) \quad U(z) = \int_{\partial\Omega} \partial_\nu H(z, \zeta) U(\zeta) d\sigma(\zeta),$$

which reconstructs the interior values of a function  $U$  harmonic in  $\Omega$  by its boundary values. We set  $P_z^\Omega(\zeta) := \partial_\nu H(z, \zeta)$  the *Poisson kernel for  $\Omega$  at  $z$* .

This formula can not be immediately applied to the upper-half plane, which is not bounded. We might however borrow the general idea. The function  $H : \mathbb{C}_+ \times \mathbb{C}_+ \setminus \{(z, z) : z \in \mathbb{C}_+\} \rightarrow \mathbb{R}$ ,

$$(1.5.24) \quad H(z, w) = G(z, \bar{w}) - G(z, w) = \frac{1}{2\pi} (\log |z - \bar{w}|^{-1} - \log |z - w|^{-1})$$

satisfies all conditions in (1.5.19) with  $\Omega = \mathbb{C}_+$ . The normal derivative at the boundary is  $-\partial_\nu$ ,

$$\begin{aligned} -\partial_\nu|_{v=0} H(z, u + iv) &= \frac{1}{2\pi} \left( \frac{y}{|z - u|^2} - \frac{-y}{|z - u|^2} \right) \\ &= \frac{1}{\pi} \frac{y}{(x - u)^2 + y^2}. \end{aligned}$$

### 1.6. The Fourier transforms of Borel measures

We start with an approximation lemma for finite Borel measures. I include the approximation by point masses because it has many applications and generalizations, although we will only make use of the approximation by  $L^1$  functions.

**LEMMA 1.4.** *Let  $\mu$  be a finite Borel measure on  $\mathbb{R}$ , not necessarily finite. We can find a sequence  $\{\mu_n\}$  of measures made of a finite number of atoms such that  $\mu_n \xrightarrow[w^*]{n \rightarrow \infty} \mu$ .*

*Another approximation of  $\mu$  is provided by the Poisson integral.*

$$(1.6.1) \quad P_y[\mu] \xrightarrow[w^*]{y \rightarrow 0} \mu.$$

*We also have that  $\|P_y[\mu]\|_{L^1} \leq |\mu|(\mathbb{R})$ , hence  $\mu$  can be  $w^*$ -approximated by  $L^1$  functions.*

**PROOF.** It suffices to show the statement for a positive  $\mu$ , since any finite complex Borel measure is a linear combination of four positive ones. Let  $\alpha_\mu(t) = \alpha(t) = \mu(-\infty, t]$  be the distribution function of  $\mu$ , which is increasing, right continuous,  $\alpha(t) \rightarrow 0$  as  $t \rightarrow -\infty$ , and  $\alpha(t) \rightarrow \mu(\mathbb{R})$  as  $t \rightarrow +\infty$ .



For  $\varphi \in C_0(\mathbb{R})$ , the following limit is easily proved using its uniform continuity:

$$(1.6.2) \quad \int_{\mathbb{R}} \varphi d\mu = \lim_{n \rightarrow \infty} \sum_{j: |j2^{-n}| \leq n} \varphi(j2^{-n}) [\alpha(j2^{-n}) - \alpha((j-1)2^{-n})],$$

(1.6.3)

Equation (1.6.2) provides the desired limit with

$$\mu_n = \sum_{j: |j2^{-n}| \leq n} [\alpha(j2^{-n}) - \alpha((j-1)2^{-n})] \delta_{j2^{-n}}$$

About the second statement, let  $\varphi \in C_0(\mathbb{R})$ . Then,

$$\begin{aligned} \int_{\mathbb{R}} \varphi(x) P_y[\mu](x) dx &= \int_{\mathbb{R}} \varphi(x) \frac{1}{2\pi} \int_{\mathbb{R}} \frac{y}{(x-u)^2 + y^2} d\mu(u) dx \\ &= \int_{\mathbb{R}} \frac{1}{2\pi} \int_{\mathbb{R}} \frac{y}{(x-u)^2 + y^2} \varphi(x) dx d\mu(u) \\ &= \int_{\mathbb{R}} P_y[\varphi](u) d\mu(u). \end{aligned}$$

From the proof of proposition 1.4 it is clear that  $P_y[\varphi]$  converges uniformly to  $\varphi$  as  $y \rightarrow 0$  if  $\varphi$  is bounded and uniformly continuous, which is the present case since  $\varphi \in C_0$ . Then, the last term in the chain of equalities tends to  $\int_{\mathbb{R}} \varphi(x) dx$ , as wished. The estimate  $\|P_y[\mu]\|_{L^1} \leq |\mu(\mathbb{R})|$  holds because the Poisson kernel has integral one.  $\square$

**PROPOSITION 1.5.** *Let  $\mu$  be a finite Borel measure on  $\mathbb{R}$ . Then,  $\hat{\mu} : \mathbb{R} \rightarrow \mathbb{C}$  is bounded and uniformly continuous.*

**PROOF.** We have  $|\hat{\mu}(\omega)| \leq |\mu(\mathbb{R})|$ , hence  $\hat{\mu}$  is bounded. Uniform continuity is proved as in the case of  $L^1$  functions,

$$\begin{aligned} |\hat{\mu}(\omega) - \hat{\mu}(\eta)| &\leq \int_{\mathbb{R}} |e^{i\omega x} - e^{i\eta x}| d|\mu|(x) \\ &\leq 2 \int_{x: |x| > R} d|\mu|(x) + \int_{x: |x| \leq R} |e^{i(\omega-\eta)x}| d|\mu|(x). \end{aligned}$$

For fixed  $\epsilon > 0$ , choose first  $R$  such that the integral in first summand is bounded by  $\epsilon$  (dominated convergence), then choose  $\delta$  such that  $|e^{i(\omega-\eta)x}| \leq \epsilon$  if  $|\omega - \eta| \leq \delta$ .  $\square$

**THEOREM 1.6.** *Let  $\mu$  be a finite Borel measure on  $\mathbb{R}$ . If  $\widehat{\mu} = 0$ , then  $\mu = 0$ .*

**PROOF.** Let  $P_y[\mu] = P_y * \mu$  be the Poisson approximation of  $\mu$ . For all  $\varphi \in C_0(\mathbb{R})$ :

$$\begin{aligned} \int_{\mathbb{R}} P_y[\mu](x) \overline{\varphi(x)} dx &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{P_y[\mu]}(\omega) \widehat{\overline{\varphi}}(\omega) d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\mu}(\omega) e^{-y|\omega|} \widehat{\overline{\varphi}}(\omega) d\omega \\ &= 0. \end{aligned}$$

As  $y \rightarrow 0$ , by lemma 1.4 we have that

$$\int_{\mathbb{R}} \overline{\varphi(x)} d\mu(x) = \lim_{y \rightarrow 0} \int_{\mathbb{R}} P_y[\mu](x) \overline{\varphi(x)} dx = 0.$$

Since this holds for all  $\varphi \in C_0$ ,  $\mu = 0$  as a measure.  $\square$

We close with a useful variation on Plancherel formula, which holds if  $\mu$  is a finite measure and  $\widehat{\varphi} \in L^1(\mathbb{R})$ :

$$(1.6.4) \quad \int_{\mathbb{R}} \varphi(-x) d\mu(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\omega) \widehat{\mu}(\omega) d\omega.$$

In fact,

$$\begin{aligned} \int_{\mathbb{R}} \varphi(-x) d\mu(x) &= \int_{\mathbb{R}} \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\omega) e^{-i\omega x} d\omega d\mu(x) \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\varphi}(\omega) \widehat{\mu}(\omega) d\omega. \end{aligned}$$

### 1.7. Herglotz representation

The theorem of Herglotz we present here will be the key to determine the spectral measures for a self-adjoint operator, which will be the ones associated with holomorphic functions indexed by vectors in a Hilbert space.

Let  $\mu$  be a finite Borel measure on  $\mathbb{R}$ . Its *Poisson integral* on  $\mathbb{C}_+$  is defined in the natural way.

$$(1.7.1) \quad P[\mu](z) := (P_y * \mu)(x) = \int_{\mathbb{R}} P_y(x - u) d\mu(u) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{y}{(x - u)^2 + y^2} d\mu(u).$$

If we want to fix the height  $y$ , we write

$$P_y[\mu](x) := P[\mu](x + iy), \quad P_y[\mu] : \mathbb{R} \rightarrow \mathbb{R}.$$

The function  $P[\mu]$  is (i) harmonic in  $\mathbb{C}_+$  (differentiate under the integral); (ii) positive, if the measure  $\mu$  is positive.

REMARK 1.1. Let  $\mu \geq 0$  be a finite Borel measure on  $\mathbb{R}$ . Then  $P[\mu] : \mathbb{C}_+ \rightarrow \mathbb{R}$  is a positive harmonic function, and

$$(1.7.2) \quad P[\mu](x + iy) \leq \frac{\mu(\mathbb{R})}{\pi y},$$

since  $P_y(x - u) \leq \frac{1}{\pi y}$ .

The remark has a surprising converse.

THEOREM 1.7. [*Herglotz theorem for positive harmonic functions*] Let  $U : \mathbb{C}_+ \rightarrow [0, \infty)$  be a positive, harmonic function, and suppose that  $U(x + iy) \leq \frac{C}{y}$  for some constant independent of  $x, y$ . Then, there exists a positive, finite measure  $\mu$  on  $\mathbb{R}$  such that  $U = P[\mu]$ . Moreover, such measure is unique.

PROOF. We do some bootstrapping followed by an application of the Hahn-Banach theorem. For fixed  $\epsilon > 0$ , let  $U_\epsilon(x) = U(x + i\epsilon)$ , which is positive and harmonic for  $z$  in  $\mathbb{C}_+$ , continuous up to the real line, and bounded,  $U_\epsilon(x) \leq \frac{C}{\epsilon}$ . Since the Poisson kernel is integrable, we can form

$$P[U_\epsilon](z) = P_y[U_\epsilon](x) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{y}{(x - u)^2 + y^2} U_\epsilon(u) du.$$

**Step I.** We will show the semigroup law  $U_{\epsilon+y} = P_y[U_\epsilon]$  for all  $y > 0$ . Let  $G_\epsilon(x + iy) = P_y[U_\epsilon](x)$  and  $H_\epsilon(x + iy) = U_{\epsilon+y}(x)$ . Both functions are positive and harmonic on  $\mathbb{C}_+$ ; they continuous on  $\mathbb{C}_+ \cup \mathbb{R}$ , the closure of the upper-half plane;  $G_\epsilon(x) = U(x + i\epsilon) = H_\epsilon(x)$  for  $x \in \mathbb{R}$ ; they are both bounded. Their difference  $K_\epsilon = H_\epsilon - G_\epsilon$  is then harmonic on the upper-half plane, bounded, continuous on its closure, and vanishing on  $\mathbb{R}$ . We can use the reflection principle to extend  $K_\epsilon$  to a harmonic, bounded ( $K_\epsilon(\bar{z}) := -K_\epsilon(z)$ ) function defined on  $\mathbb{C}$ , which vanishes on the real line. By Liouville theorem, a bounded harmonic function on  $\mathbb{C}$  must be constant, hence  $K_\epsilon(z) = K_\epsilon(0) = 0$  for all complex  $z$ . This proves our claim.

**Step II.** We show that  $1/\pi \int_{\mathbb{R}} U_\epsilon(u) du \leq C$ , where  $C$  is the constant in the hypothesis, which is independent of  $\epsilon > 0$ . In fact, by the assumption on the decay of  $U(x + iy)$  with respect to  $y$ , and using Step I,

$$\begin{aligned} \frac{CR}{\epsilon + R} &\geq RU_{\epsilon+R}(x) = RP_R[U_\epsilon](x) \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{R^2}{(x - u)^2 + R^2} U_\epsilon(u) du \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \left( 1 - \frac{(x - u)^2}{(x - u)^2 + R^2} \right) U_\epsilon(u) du \\ &\rightarrow \frac{1}{\pi} \int_{\mathbb{R}} U_\epsilon(u) du \text{ as } R \rightarrow \infty, \end{aligned}$$

by monotone convergence. We proved the claim, since the limit as  $R \rightarrow \infty$  of the first term in the chain is  $C$ .

**Step III.** Consider the family of finite positive Borel measures  $d\mu_\epsilon(u) = U_\epsilon(u)du$ . Since their norms  $\mu_\epsilon(\mathbb{R}) \leq C$  are uniformly bounded, by the Banach-Alaoglu theorem there exist a sequence  $\epsilon_n \rightarrow 0$  and a positive Borel measure  $\mu$  with  $\mu(\mathbb{R}) \leq C$  such that  $\mu_{\epsilon_n} \xrightarrow{w^*} \mu$ . Recall that (by the Riesz representation of measures) the space of the finite Borel measures on  $\mathbb{R}$  is the dual of  $C_0(\mathbb{R})$ , the space of the continuous functions tending to zero as  $x \rightarrow \pm\infty$ . In particular, for all  $x + iy$  in  $\mathbb{C}_+$ ,

$$\begin{aligned} U(x + iy) &= \lim_{n \rightarrow \infty} U_{\epsilon_n}(x + iy) = \lim_{n \rightarrow \infty} \frac{1}{\pi} \int_{\mathbb{R}} \frac{y}{(x - u)^2 + y^2} U_{\epsilon_n}(u) du \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{y}{(x - u)^2 + y^2} d\mu(u), \end{aligned}$$

as wished.

**Step IV.** We prove uniqueness. The Herglotz representation formula can be written as  $U_y = P_y * \mu$ . It is well known (and easy to verify) that

$$(1.7.3) \quad \widehat{P_y}(\omega) = e^{-y|\omega|}.$$

If  $U_y = P_y * \mu$ , then  $\widehat{U_y}(\omega) = e^{-y|\omega|}\widehat{\mu}(\omega)$ . This shows that  $\widehat{\mu}$  is uniquely determined by  $U$ , and it is known that  $\mu$  is determined by  $\widehat{\mu}$ .  $\square$

**COROLLARY 1.3.** [Herglotz theorem for holomorphic functions  $f : \mathbb{C}_+ \rightarrow \mathbb{C}_+ \cup \mathbb{R}$ ] Let  $f : \mathbb{C}_+ \rightarrow \mathbb{C}_+ \cup \mathbb{R}$  be holomorphic and such that

$$(1.7.4) \quad \operatorname{Im} f(x + iy) \leq \frac{C}{y}.$$

Then, there exist a unique finite, positive Borel measure  $\mu$  and a real number  $a$  such that

$$(1.7.5) \quad f(z) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu(u)}{z - u} + a.$$

**PROOF.** We have that

$$\frac{1}{\pi} \frac{1}{u - (x + iy)} = \frac{1}{\pi} \frac{(u - x) + iy}{(u - x)^2 + y^2} = Q_y(x - u) + iP_y(x - y)$$

is holomorphic with respect to  $z = x + iy$ . If  $f = \varphi + i\psi$  is as in the hypothesis, then,  $\psi$  satisfies the hypothesis of Herglotz theorem, then there exists  $\mu \geq 0$  (Borel, finite) such that  $\psi = P[\mu]$ . The function

$$Q[\mu](x + iy) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{u - x}{(u - x)^2 + y^2} d\mu(u)$$

is harmonic in  $\mathbb{C}$  and  $Q[\mu] + iP[\mu] = Q[\mu] + i\psi$  is holomorphic. Hence,  $Q[\mu] - \varphi$  is a constant.

Uniqueness is easily checked.  $\square$

### 1.8. Bochner's theorem

This is here because it's the key for one of the proofs of Stone's theorem on groups of unitary operators. The result itself is one of the cornerstones of Fourier analysis and has applications to analysis, probability, and statistics.

A function  $f : \mathbb{R} \rightarrow \mathbb{C}$  is *positive definite* if

$$\sum_{j,k=1}^n \varphi(x_k - x_j) c_j \overline{c_k} \geq 0$$

for all choices of  $n$  points  $x_1, \dots, x_n$  in  $\mathbb{R}$  and  $n$  complex scalars  $c_1, \dots, c_n$ . This is the same as saying that for all choices of  $x_1, \dots, x_n$ , the matrix  $[\varphi(x_k - x_j)]_{k,j=1}^n$  is definite positive.

Consider the points  $0, x$ . The matrix

$$\begin{pmatrix} \varphi(0) & \varphi(x) \\ \varphi(-x) & \varphi(0) \end{pmatrix}$$

is positive definite (hence, Hermitian, since the scalar field is  $\mathbb{C}$ ), thus

$$(1.8.1) \quad \varphi(-x) = \overline{\varphi(x)},$$

$$(1.8.2) \quad \varphi(0) \geq |\varphi(x)|.$$

**THEOREM 1.8.** [*Bochner's theorem*] For a function  $f : \mathbb{R} \rightarrow \mathbb{C}$  the following are equivalent:

- (i)  $f$  is continuous and positive definite;
- (ii) there exists a finite, positive Borel measure  $\mu$  on the real line such that  $\hat{\mu} = f$ .

The idea of the proof can be best appreciated in a finite environment. Let  $\mathbb{Z}_N = \{0, 1, \dots, N-1\}$  be the group of the integers modulo  $N$ , and for  $f : \mathbb{Z}_N \rightarrow \mathbb{C}$ , let its *discrete Fourier transform*  $\hat{f} : \mathbb{Z}_N \rightarrow \mathbb{C}$  be

$$\hat{f}(\omega) = \sum_{n=0}^{N-1} e^{-2\pi i n \omega / N} f(n).$$

The usual formulas hold:

$$\begin{aligned} f(n) &= \frac{1}{N} \sum_{\omega=0}^{N-1} e^{2\pi i n \omega / N} \widehat{f}(\omega); \\ \sum_{n=0}^{N-1} \overline{\nu(n)} f(n) &= \frac{1}{N} \sum_{\omega=0}^{N-1} \overline{\widehat{\nu}(\omega)} \widehat{f}(\omega); \\ \widehat{f * \nu}(\omega) &= \widehat{f}(\omega) \widehat{\nu}(\omega), \end{aligned}$$

where

$$(f * \nu)(n) = \sum_{j=0}^{N-1} f(n-j) \nu(j).$$

Using all this, if  $f, c : \mathbb{Z}_N \rightarrow \mathbb{C}$ ,

$$(1.8.3) \quad \frac{1}{N} \sum_{\omega=0}^{N-1} \widehat{f}(\omega) |\widehat{c}(\omega)|^2 = \frac{1}{N} \sum_{\omega=0}^{N-1} \overline{\widehat{c}(\omega)} [\widehat{f\widehat{c}}](\omega)$$

$$(1.8.4) \quad = \sum_{n=0}^{N-1} \overline{c(n)} [f * c](n)$$

$$(1.8.5) \quad = \sum_{m,n=0}^{N-1} f(n-m) \overline{c(n)} c(m).$$

Comparing the first and the last terms in the chain of equalities, it is clear that if  $\widehat{f}$  is positive, then  $f$  is positive definite. Conversely, if  $f$  is positive definite, then the first term is positive for each choice of  $c$ . For any fixed  $\omega_0$  in  $\mathbb{Z}_N$ , we can choose  $c$  such that  $\widehat{c}(\omega) = \delta_{\omega_0, \omega}$ , obtaining  $\widehat{f}(\omega_0) \geq 0$ .

In the general case, the idea is the same, but there are complications. We start with a useful result in linear algebra.

If  $A, B$  are  $N \times N$  matrices with complex entries, their *Schur (or Hadamard) product* is the  $N \times N$  matrix  $A \circ B$  having entries  $[A \circ B]_{ij} = A_{ij} B_{ij}$ . The matrix  $A$  is *positive definite* if

$$0 \leq \sum_{ij} \overline{c_i} A_{ij} c_j = \langle c, Ac \rangle_{\mathbb{C}^N}$$

for all  $c \in \mathbb{C}^N$ . Sometimes this property is called semi-positive definiteness. We say that  $A$  is *strictly positive definite* if  $\langle c, Ac \rangle_{\mathbb{C}^N} > 0$  whenever  $c \neq 0$ .

**THEOREM 1.9** (Schur product theorem). *If  $A, B$  are positive definite  $N \times N$  matrices, then  $A \circ B$  is positive definite.*

For the proof, we lift the problem to a more general one living on the *tensor product*  $\mathbb{C}^N \otimes \mathbb{C}^N$ , the space having as basis  $\{e_i \otimes e_j : 1 \leq i, j \leq N\}$ . The *tensor product*  $\otimes$  is bilinear by definition,

$$\begin{aligned}(ax + by) \otimes z &= a(x \otimes z) + b(y \otimes z), \\ z \otimes (ax + by) &= a(z \otimes x) + b(z \otimes y),\end{aligned}$$

for all  $x, y, z \in \mathbb{C}^N$  and  $a, b \in \mathbb{C}$ ; and no other relation is imposed.

Here is a concrete model of the tensor product. The elements of  $\mathbb{C}^N \otimes \mathbb{C}^N$  might be seen as linear combinations of vectors of the form

$$h \otimes k = hk^t \in \mathbb{C}^{N \times N},$$

where  $k^t$  is the transpose of the column vector  $k$ . This way, we have a linear isomorphism  $\mathbb{C}^N \otimes \mathbb{C}^N \rightarrow \mathbb{C}^{N \times N}$ , which identifies elements in the tensor product with matrices.

A linear map  $A : \mathbb{C}^N \rightarrow \mathbb{C}^N$  induces two linear maps

$$\mathbb{C}^N \otimes \mathbb{C}^N \xrightarrow{A \otimes I, I \otimes A} \mathbb{C}^N \otimes \mathbb{C}^N,$$

where

$$\begin{aligned}(A \otimes I)(x \otimes y) &= (Ax) \otimes y, \\ (I \otimes A)(x \otimes y) &= x \otimes (Ay).\end{aligned}$$

More generally,

$$(A \otimes B)(x \otimes y) = (A \otimes I)(I \otimes B)(x \otimes y) = (Ax \otimes By).$$

In the column $\times$ row model, you might think of  $A \otimes B$  as the linear map  $xy^t \mapsto (Ax)(By)^t = Axy^t B^t$ .

The vector space  $\mathbb{C}^N \otimes \mathbb{C}^N$  is endowed with the inner product making tensor products  $e_i \otimes e_j$  of standard basis elements of  $\mathbb{C}^N$  into a orthonormal basis. This gives

$$(1.8.6) \quad \langle x \otimes u, y \otimes v \rangle = \langle x, y \rangle \cdot \langle u, v \rangle.$$

In fact,

$$\begin{aligned}\langle x \otimes u, y \otimes v \rangle &= \sum_{ijlm} \bar{x}_i u_j \bar{y}_l v_m \langle e_i \otimes e_j, e_l \otimes e_m \rangle \\ &= \sum_{ij} \bar{x}_i u_j \bar{y}_i v_j \\ &= \langle x, y \rangle \cdot \langle u, v \rangle.\end{aligned}$$

In the column $\times$ row model,

$$\langle xu^t, yv^t \rangle = y^t \bar{x} v^t \bar{u}.$$

A consequence of (1.8.6) is that we can replace the standard basis of  $\mathbb{C}^N$  by any other orthonormal basis in the calculations. In fact, (1.8.6) might be taken as the definition of the inner product on  $\mathbb{C}^N \times \mathbb{C}^N$ .

LEMMA 1.5 (Tensor product preserves positivity). *If  $A$  and  $B$  are positive, then  $A \otimes B$  is positive.*

PROOF OF LEMMA 1.5. Choose a orthonormal basis  $\{\epsilon_i\}$  for  $\mathbb{C}^N$  with respect to which  $B$  is diagonal,  $B_{ij} = \delta_{ij} \lambda_i$  with  $\lambda_1, \dots, \lambda_N \geq 0$ . For a vector  $V = \sum_{ij} v_{ij} \epsilon_i \otimes \epsilon_j \in \mathbb{C}^N \otimes \mathbb{C}^N$ , we have

$$\begin{aligned} \langle V, (A \otimes B)V \rangle &= \sum_{ijlm} \bar{v}_{ij} v_{lm} \langle \epsilon_i \otimes \epsilon_j, (A \otimes B)(\epsilon_l \otimes \epsilon_m) \rangle \\ &= \sum_{ijlm} \bar{v}_{ij} v_{lm} \langle \epsilon_i, A\epsilon_l \rangle \langle \epsilon_j, B\epsilon_m \rangle \\ &= \sum_{ijlm} \bar{v}_{ij} v_{lm} A_{il} B_{jm} \\ &= \sum_{jm} B_{jm} \left( \sum_{il} \bar{v}_{ij} v_{lm} A_{il} \right) \\ &= \sum_j \lambda_j \left( \sum_{il} \bar{v}_{ij} v_{lj} A_{il} \right) \\ &\geq 0, \end{aligned}$$

since the sums in the parenthesis are positive and so are the  $\lambda_j$ 's.  $\square$

PROOF OF SCHUR'S THEOREM GIVEN LEMMA 1.5. By the lemma above,

$$0 \leq \langle V, (A \otimes B)V \rangle = \sum_{ijlm} \bar{v}_{ij} v_{lm} A_{il} B_{jm}.$$

Choose in particular  $v_{ij} = \delta_{ij} c_i$ , to obtain

$$0 \leq \sum_{il} \bar{c}_i c_l A_{il} B_{il},$$

as wished.  $\square$



The topic we have just sketched can be interpreted in terms of, and applied to, the theory of *Hilbert spaces with reproducing kernel*, which is a very active area in pure mathematics and in its applications to a number of areas, including statistics and machine learning. See e.g. [An Introduction to the Theory of Reproducing Kernel Hilbert Spaces \(2016\)](#) by Vern I. Paulsen, Mrinal Raghupathi, Cambridge University Press.

PROOF OF BOCHNER'S THEOREM. If  $\mu \geq 0$  is a finite Borel measure, then  $\hat{\mu}$  is continuous and bounded. About positive definiteness,

$$\begin{aligned} \sum_{l,j} \hat{\mu}(\omega_l - \omega_j) c_l \overline{c_j} &= \int_{\mathbb{R}} \sum_{l,j} e^{-i(\omega_l - \omega_j)x} c_l \overline{c_j} d\mu(x) \\ &= \int_{\mathbb{R}} \left| \sum_l e^{-i\omega_l x} c_l \right|^2 d\mu(x) \\ &\geq 0 \quad . \end{aligned}$$

In the opposite direction, suppose that  $f : \mathbb{R} \rightarrow \mathbb{C}$  is positive definite, hence bounded, and continuous. For  $y > 0$ , let  $f_y(\omega) = f(\omega)e^{-y|\omega|}$ , which is still positive definite by Schur's theorem, and it lies in  $L^1(\mathbb{R})$  (hence in  $L^2(\mathbb{R})$ ). Set  $m_y(x) = \frac{1}{2\pi} \int_{\mathbb{R}} f_y(\omega) e^{i\omega x} d\omega = [\mathcal{F}^{-1} f_y](x)$ , so that  $m_y \in L^2(\mathbb{R}) \cap C_0(\mathbb{R})$ . For all functions  $\varphi \in \mathcal{S}(\mathbb{R})$  (the Schwartz class, but if you are not familiar with that object, you might even consider  $\varphi \in C_c^\infty(\mathbb{R})$ ), we have:

$$\begin{aligned} \int_{\mathbb{R}} m_y(x) |\varphi(x)|^2 dx &= \int_{\mathbb{R}} [m_y(x) \varphi(x)] \overline{\varphi(x)} dx \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{m_y} \widehat{\varphi}(\omega) \overline{\widehat{\varphi}(\omega)} d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} [\widehat{m_y} * \widehat{\varphi}](\omega) \overline{\widehat{\varphi}(\omega)} d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \widehat{m_y}(\omega - \eta) \widehat{\varphi}(\omega) \overline{\widehat{\varphi}(\omega)} d\eta d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} f_y(\omega - \eta) \widehat{\varphi}(\omega) \overline{\widehat{\varphi}(\omega)} d\eta d\omega \\ &= \frac{1}{2\pi} \lim_{N \rightarrow \infty} \sum_{l,m=-N^2}^{N^2} f_y(l/N - m/N) \widehat{\varphi}(l/N) \overline{\widehat{\varphi}(m/N)} \frac{1}{N^2} \\ &\geq 0 \end{aligned}$$

because  $f_y$  is positive definite. Hence,  $m_y \geq 0$  is a positive function with  $\|m_y\|_{L^1} = \widehat{m_y}(0) = f(0)$ .

For  $\varphi \in \mathcal{S}(\mathbb{R})$ , let

$$\begin{aligned}\Lambda_y(\varphi) &:= \int_{\mathbb{R}} m_y(x) \varphi(x) dx = \frac{1}{2\pi} \int_{\mathbb{R}} f(\omega) e^{-y|\omega|} \widehat{\varphi}(-\omega) d\omega \\ &\rightarrow \frac{1}{2\pi} \int_{\mathbb{R}} f(\omega) \widehat{\varphi}(-\omega) d\omega.\end{aligned}$$

as  $y \rightarrow 0$ .

On the other hand, by Banach-Alaoglu there exist a sequence  $y_n \rightarrow 0$  and a finite (positive) Borel measure  $\mu$  such that  $\mu(\mathbb{R}) \leq f(0)$  and  $m_{y_n} dm \xrightarrow{w^*} \mu$ , hence

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} m_{y_n}(x) \varphi(x) dx = \int_{\mathbb{R}} \varphi(x) d\mu(x).$$

We have shown that

$$\frac{1}{2\pi} \int_{\mathbb{R}} f(\omega) \widehat{\varphi}(-\omega) d\omega = \int_{\mathbb{R}} \varphi(x) d\mu(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\mu}(\omega) \widehat{\varphi}(-\omega) d\omega$$

holds for all  $\varphi \in \mathcal{S}$  (use (1.6.4) for the second equality), hence  $f = \widehat{\mu}$ .  $\square$

## CHAPTER 2

### Hilbert spaces

Among function spaces, Hilbert spaces are probably the most ubiquitous in mathematics and its applications. At its inception, "Hilbert theory" dealt with concrete  $L^2$  spaces, but soon an abstract theory was developed. The advantage of the abstract theory is that it makes it easier to recognize when a Hilbert space structure underlies a cluster of mathematical objects and phenomena, and it helps in generating Hilbert spaces in which better state, and solve, theoretical, as well as practical problems. Although all Hilbert spaces are isomorphic to some  $L^2$  space, the point of view that knowledge of  $L^2$  is all that's needed is far too simplistic, and basically incorrect. Most Hilbert spaces are spaces of functions defined on some set of points, and such "points" constitute further structure, which often is at the core of the problem we have at hands. In this chapter, however, we are mostly interested in the abstract theory.

Another useful way of thinking is that of viewing at Hilbert spaces as generalizations (often infinite-dimensional, with complex rather than real scalars) of the Euclidean space. Even if in applications each vector in the Hilbert space represents a functions, we can think of each of them as a "point" in a linear space where notions like orthogonality make perfect sense. This intuition is helpful in translating complex phenomena in simple pictures, and pictures into statements (which are sometimes true, sometimes false).

#### 2.1. Basic geometry of Hilbert spaces and Riesz Lemma

**2.1.1. Definition and basic properties.** An *inner product* on a vector space  $V$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ) is a map  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$  such that:

- (i)  $\langle x, x \rangle \geq 0$  and  $\langle x, x \rangle = 0$  if and only if  $x = 0$ ;
- (ii)  $\langle z, \alpha x + \beta y \rangle = \alpha \langle z, x \rangle + \beta \langle z, y \rangle$  for  $x, y, z \in V$  and  $\alpha, \beta \in \mathbb{C}$ ;
- (iii)  $\overline{\langle x, y \rangle} = \langle y, x \rangle$ .

Two vectors  $x, y$  are orthogonal,  $x \perp y$ , if  $\langle x, y \rangle = 0$ . We define  $\|x\| := \langle x, x \rangle^{1/2}$  to be the norm of  $x \in V$ . A simple calculation gives:

LEMMA 2.1. *[Pythagorean relation] Let  $x, y \in V, x \neq 0$ . Then,*

$$\|y\|^2 = \left\| y - \frac{x}{\|x\|} \left\langle \frac{x}{\|x\|}, y \right\rangle \right\|^2 + \left| \left\langle \frac{x}{\|x\|}, y \right\rangle \right|^2.$$

PROOF. Expand the right hand side using the definition of norm and the properties of the inner product.  $\square$

EXERCISE 2.1. *Explain why the lemma is called as it is with a picture, or observing that the vectors  $\frac{x}{\|x\|} \left\langle \frac{x}{\|x\|}, y \right\rangle$  and  $y - \frac{x}{\|x\|} \left\langle \frac{x}{\|x\|}, y \right\rangle$  are orthogonal and their sum is  $y$ .*

COROLLARY 2.1. *[Cauchy-Schwarz inequality] For  $x, y \in V$  we have*

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

*with equality if and only if  $x, y$  are linearly dependent.*

PROOF. If  $x = 0$ , there is nothing to prove. Otherwise, the inequality follows from dropping the first summand from the right of the Pythagorean relation. In case of equality, either  $x = 0$  or the first term in the right of the Pythagorean relation vanishes, in which case  $x, y$  are linearly dependent. Conversely, if they are linearly dependent it is easy to see that equality holds in Cauchy-Schwarz.  $\square$

PROPOSITION 2.1. *The function  $x \mapsto \|x\|$  defines a norm on  $V$ .*

PROOF. It is an immediate consequence of Cauchy-Schwarz:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 \\ &= \|x\|^2 + 2 \operatorname{Re}(\langle x, y \rangle) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

$\square$

An inner product space  $(H, \langle \cdot, \cdot \rangle)$  is a *Hilbert space* if it is complete with respect to the norm induced by the inner product.

EXERCISE 2.2. *Show that the inner product can be written in terms of norms by means of the **polarization identity**:*

$$(2.1.1) \quad \langle x, y \rangle = \frac{1}{4} [(\|x + y\|^2 - \|x - y\|^2) - i(\|x + iy\|^2 - \|x - iy\|^2)].$$

LEMMA 2.2. *[Parallelogram law] Let  $x, y \in V$ . Then,*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

PROOF. Expand the expression on the left using the definition of  $\|\cdot\|$ .  $\square$

EXERCISE 2.3. Show that, if  $\|\cdot\|$  is a norm on a vector space which satisfies the parallelogram law, then it comes from an inner product as defined by the polarization identity.

EXERCISE 2.4. Show that the (two-dimensional) Banach spaces  $\ell^\infty(\{0,1\})$ ,  $\ell^1(\{0,1\})$  are not inner product spaces (i.e. that their norm does not come from an inner product). A **Hint**. The parallelogram law fails.

**2.1.2.  $L^2$  as a Hilbert space.** Let  $(X, \mathcal{F}, \mu)$  be a measure space. Observe that, if  $f, g \in L^2(\mu)$ , then  $\bar{f}g \in L^1(\mu)$ :

$$\int_x |\bar{f}g| d\mu \leq \|f\|_{L^2} \|g\|_{L^2}$$

by Hölder's inequality with  $p = p' = 2$ . We can then define

$$\langle f, g \rangle_{L^2} = \int_X \bar{f}g d\mu,$$

which satisfies the properties of a inner product, with associated norm  $\langle f, f \rangle_{L^2} = \|f\|_{L^2}^2$ . We saw that  $L^2(\mu)$  is complete with respect to this norm.

**2.1.3. Projections onto subspaces.** Let  $M \subseteq H$  be a closed, linear subspace of  $H$ . Its *orthogonal complement*,  $M^\perp = H \ominus M$ , is:

$$M^\perp = \{x \in H : x \perp y \text{ for all } y \in M\}.$$

LEMMA 2.3 (Projection Lemma). Let  $M \subseteq H$  be a closed, linear subspace of  $H$ . Then, for each  $x \in H$  there exists a unique  $u \in M$  such that

$$\|x - u\| = \inf\{\|x - y\| : y \in M\}.$$

PROOF. Let  $\{y_n\}_{n \geq 1}$  be a sequence in  $M$  such that  $\|y_n - x\| \rightarrow \delta := \inf\{\|x - y\| : y \in M\}$ . We show that it is a Cauchy sequence- By the parallelogram law,

$$\begin{aligned} \|y_m - y_n\|^2 &= \|(y_m - x) - (y_n - x)\|^2 \\ &= 2(\|y_m - x\|^2 + \|y_n - x\|^2) - \|y_m + y_n - 2x\|^2 \\ &= 2(\|y_m - x\|^2 + \|y_n - x\|^2) - 4\left\|\frac{y_m + y_n}{2} - x\right\|^2 \\ &\leq 2(\|y_m - x\|^2 + \|y_n - x\|^2) - 4\delta^2 \\ &\quad \text{because } \frac{y_m + y_n}{2} \in M \\ &\rightarrow 4\delta^2 - 4\delta^2 = 0. \end{aligned}$$

We now let  $u = \lim y_n \in M$  (because  $M$  is closed).

If  $u'$  is another point with the same property, then  $\frac{u+u'}{2} \in M$  and

$$\|u - u'\|^2 = 2 \left( \|u - x\|^2 + \|u' - x\|^2 \right) - 4 \left\| \frac{u + u'}{2} - x \right\|^2 \leq 4\delta^2 - 4\delta^2 = 0,$$

hence,  $u = u'$ . □

We call  $u = \pi_M(x)$  the *orthogonal projection* of  $x$  onto  $M$ .

**PROPOSITION 2.2.** *Let  $E \subseteq H$  and define  $E^\perp = \{x \in H : x \perp y \text{ for all } y \in E\}$ . Then,*

- (i)  $E^\perp$  is a closed, linear subspace of  $H$ , and  $E \cap E^\perp = \{0\}$ ;
- (ii) if  $M$  is a linear subspace of  $H$ , then  $M^\perp = \overline{M}^\perp$ ;
- (iii) if  $M$  is a linear subspace of  $H$ , then  $(M^\perp)^\perp = \overline{M}$  is the closure of  $M$  in  $H$ .

**PROOF.** (i) Let  $x, y \in E^\perp$ ,  $\alpha, \beta \in \mathbb{C}$ , and let  $z \in E$ . Then,

$$\langle z, \alpha x + \beta y \rangle = \alpha \langle z, x \rangle + \beta \langle z, y \rangle = 0,$$

hence,  $\alpha x + \beta y \in E^\perp$ .

Let now  $E^\perp \ni x_n$  and  $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$ . Then, for  $z \in E$ ,

$$|\langle z, x \rangle| = |\langle z, x \rangle - \langle z, x_n \rangle| \leq \|z\| \cdot \|x - x_n\| \rightarrow 0,$$

as  $n \rightarrow \infty$ . So,  $|\langle z, x \rangle|$  is smaller than any positive  $\epsilon$ , hence  $\langle z, x \rangle = 0$ , i.e.  $x \in E^\perp$ .

Finally, if  $x \in E \cap E^\perp$ , then  $0 = \langle x, x \rangle = \|x\|^2$ , hence,  $x = 0$ .

(iii)  $(M^\perp)^\perp$  is linear and closed by (i), and  $(M^\perp)^\perp \supseteq M$ : if  $x \in M$ , then, for  $z \in M^\perp$  we have that  $\langle z, x \rangle = 0$ , then  $x \in (M^\perp)^\perp$ . We have thus proved that  $\overline{M}$ , the smallest closed linear subspace containing  $M$ , is a subset of  $(M^\perp)^\perp$ .

Consider  $x \in (M^\perp)^\perp \setminus \overline{M}$ . By the Projection Lemma,  $x = u + v$  with  $u \in \overline{M}$ , and  $v \in \overline{M}^\perp = M^\perp$ . So,  $v \in M^\perp \cap (M^\perp)^\perp$ , hence,  $v = 0$  and  $x \in \overline{M}$ , contradicting the assumption. □

**EXERCISE 2.5.** *Let  $E$  be a subset of  $H$ . Show that  $(E^\perp)^\perp = \overline{\text{span}(E)}$  is the closure of the linear span of  $E$  in  $H$ .*

**THEOREM 2.1. [Orthogonal Decomposition]** *Let  $M \subseteq H$  be a closed, linear subspace of  $H$ . Then, for each  $x \in H$  there is a unique decomposition  $x = u + v$  with  $u \in M$  and  $v \in M^\perp$ .*

PROOF. Let  $u \in M$  be as in the Projection Lemma, and for  $t \in \mathbb{R}$  and  $w \in M$ , so that  $u + tw \in M$ , define  $f(t) = \|u + tw - x\|^2$ . By minimality,  $f'(0) = 0$ , and

$$\begin{aligned} f'(0) &= \left. \frac{d}{dt} \right|_{t=0} (\|u - x + tw\|^2) \\ &= \left. \frac{d}{dt} \right|_{t=0} (\|u - x\|^2 + 2t \operatorname{Re}(\langle u - x, w \rangle) + t^2 \|w\|^2) \\ &= [2 \operatorname{Re}(\langle u - x, w \rangle) + 2t \|w\|^2]_{t=0} \\ &= 2 \operatorname{Re}(\langle u - x, w \rangle). \end{aligned}$$

Hence,  $\operatorname{Re}(\langle u - x, w \rangle) = 0$ . The same reasoning with  $g(t) = \|u + itw - x\|^2$  gives  $\operatorname{Im}(\langle u - x, w \rangle) = 0$ . Hence,  $\langle u - x, w \rangle = 0$  for all  $w \in M$ , i.e.  $v := x - u \in M^\perp$ .

If I have two decompositions  $x = u + v = u' + v'$  with  $u, u' \in M$  and  $v, v' \in M^\perp$ , then  $u - u' = v' - v \in M \cap M^\perp$ , hence  $u - u' = v' - v = 0$ .  $\square$

In the theorem on the orthogonal decomposition of a vector, we saw that, if  $M$  is a closed subspace of a Hilbert space  $H$ , and  $x \in H$ , then there exist unique  $u \in M$  and  $v \in M^\perp$  such that  $x = u + v$ . Define  $\pi_M : H \rightarrow H$  by  $\pi_M(x) = u$ , the (*orthogonal*) *projection* of  $x$  onto  $M$ .

EXERCISE 2.6. Let  $M$  be a closed, linear subspace of a Hilbert space  $H$ .

- (i) Show that  $\pi_M : H \rightarrow H$  is linear and that  $\|\pi_M x\| \leq \|x\|$  for all  $x$  in  $H$ .
- (ii) Show that  $\pi_M^2 = \pi_M$  and that  $\pi_M$  is **self-adjoint**, i.e. that

$$\langle \pi_M x, y \rangle = \langle x, \pi_M y \rangle$$

for  $x, y \in H$ .

- (iii) Verify that  $\pi_M + \pi_{M^\perp} = I$  is the identity operator ( $I(x) = x$ ), and that  $\pi_M \circ \pi_{M^\perp} = 0$ .

It is an interesting fact that the Exercise 2.6 has a converse: all operators sharing the properties of projections are in fact projections on closed subspaces. We can, that is, indifferently work with the *lattice* of the closed subspaces of  $H$ , or with the lattice of the projection operators. Since projections operators are spacial instances of linear operators on  $H$ , this identification gives much more flexibility in theory and calculations. (This is similar to the case of  $\sigma$ -algebras, where characteristic functions of measurable sets are special instances of measurable functions, and measurable functions have a rich structure).

PROPOSITION 2.3. Let  $\pi : H \rightarrow H$  be a linear operator satisfying:

- (i)  $\pi^2 = \pi$ ;

$$(ii) \pi^* = \pi;$$

$$(iii) \|\pi(x)\| \leq C\|x\| \text{ for some } C > 0.$$

Then, there exists a closed, linear subspace  $M$  of  $H$  such that  $\pi = \pi_M$ .

PROOF. Let  $M = \{\pi(x) : x \in H\}$  be the range of  $\pi$ . For  $\pi(x), \pi(y) \in M$  and  $\alpha, \beta \in \mathbb{C}$  we have  $\alpha\pi(x) + \beta\pi(y) = \pi(\alpha x + \beta y) \in M$ , hence  $M$  is a linear space. Let  $M \ni \pi(x_n) \rightarrow y$  in  $H$ . Then,  $\pi(x_n) = \pi^2(x_n) \rightarrow \pi(y)$  because

$$\|\pi^2(x_n) - \pi(y)\| \leq C\|\pi(x_n) - y\| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Thus,  $y = \pi(y) \in M$ , showing that  $M$  is closed.

Finally, let  $x$  be an element of  $H$ , and write  $x = u + v = \pi_M(x) + \pi_{M^\perp}(x)$ , as in Theorem 2.1. By definition of  $M$ ,  $u = \pi(w)$  for some  $w$ . For  $y \in H$  we have:

$$\langle \pi(v), y \rangle = \langle v, \pi(y) \rangle = 0,$$

since  $\pi(y) \in M$  and  $v \in M^\perp$ . Hence,  $\pi(v) = 0$ , and so

$$\pi(x) = \pi(u) = \pi^2(w) = \pi(w) = u = \pi_M(x),$$

showing  $\pi = \pi_M$ . □

**2.1.4. F. Riesz representation in Hilbert spaces.** F. Riesz representation theorem in Hilbert spaces, like the one we saw concerning measures, shows that any element from an abstract collection of objects can be represented as an object from a very special and concrete subclass. The first theorem we saw represented positive functionals on  $C_c(X)$  in terms of measures; the present one represents bounded, linear functionals on  $H$  in terms of inner products.

EXERCISE 2.7. Let  $y \in H$  and define  $T_y(x) = \langle y, x \rangle$ . Then,  $T_y : H \rightarrow \mathbb{C}$  is a linear operator (a linear functional) and  $\|T_y\| = \|y\|$ .

THEOREM 2.2 (F. Riesz Lemma). Let  $T : H \rightarrow \mathbb{C}$  be a bounded, linear functional. Then, there is a unique  $y_T$  in  $H$  such that  $T(x) = \langle y_T, x \rangle$ .

PROOF. Let  $N = \text{Ker } T$ . If  $N = H$ , set  $y_T = 0$ . If not, by the Orthogonal Decomposition Theorem there is some  $x_0 \in N^\perp \setminus \{0\}$ , and we can assume that  $\|x_0\| = 1$ . We write

$$x = \left( x - \frac{T(x)}{T(x_0)} x_0 \right) + \frac{T(x)}{T(x_0)} x_0 = u + \lambda x_0.$$

The second summand is in  $\text{span}(\{x_0\}) \subseteq N^\perp$ , with  $\lambda = \langle x_0, x \rangle$ , while  $u$  clearly lies in  $N$  (incidentally, this shows that  $\text{Ker}(N)$  has codimension one in  $H$ ).

$$T(x) = \lambda T(x_0) = \langle x_0, x \rangle T(x_0) = \left\langle \overline{T(x_0)} x_0, x \right\rangle,$$

hence,  $y_T = \overline{T(x_0)} x_0$ . □



EXERCISE 2.8. *Riesz Lemma provides an identification, in fact an isometry,  $C : T \mapsto y_T$  of the dual space  $H^* = \{T : H \rightarrow \mathbb{C} \text{ such that } \|T\| < \infty\}$  with  $H$  itself. The map  $C$  is conjugate linear,*

$$C(aS + bT) = \bar{a}C(S) + \bar{b}C(T).$$

*Prove this and prove that, if  $A : H \rightarrow K$  is a bounded, linear map between Hilbert spaces, then  $C(TA) = A^*C(T)$ , where the (Hilbert space) adjoint operator  $A^* : K \rightarrow H$  is defined by  $\langle Ax, y \rangle = \langle x, A^*y \rangle$ .*

## 2.2. Orthonormal systems

**2.2.1. Orthogonal vectors.** A family  $S = \{f_a : a \in I\}$  of vectors is an *orthogonal system* if any two vectors in it are orthogonal, and an *orthonormal system* if, in addition, each  $f_a$  has unit norm. An *orthonormal basis* (o.n.b.) for  $H$  is an orthonormal system which is maximal: no other vector can be added to it without breaking the orthonormality condition.

EXERCISE 2.9. . *Show that the orthonormal system  $\{e_a : a \in I\}$  is a o.n.b. for  $H$  if and only if  $\text{span } \{e_a : a \in I\}$  is dense in  $H$ .*

THEOREM 2.3. [Bessel Inequality] *Let  $\{e_i : i = 1, \dots, n\}$  be an orthonormal system in the Hilbert space  $H$ , and let  $x \in H$ . Then, the vector  $x - \sum_{i=1}^n \langle x, e_i \rangle e_i$  is orthogonal to  $\text{span } \{e_1, \dots, e_n\}$ , and*

$$\|x\|^2 \geq \sum_{i=1}^n |\langle e_i, x \rangle|^2.$$

*If  $\{e_a : a \in I\}$  is an o.n.b., the numbers  $\langle e_a, x \rangle = \hat{x}(a)$  are the Fourier coefficients of  $x$  w.r.t. the basis.*

PROOF. The first assertion is clear:

$$\left\langle x - \sum_{i=1}^n \langle x, e_i \rangle e_i, e_j \right\rangle = \langle x, e_j \rangle - \sum_{i=1}^n \langle x, e_i \rangle \langle e_i, e_j \rangle = 0,$$

by orthonormality of the system. As a consequence,

$$\begin{aligned} \|x\|^2 &= \left\| x - \sum_{i=1}^n \langle e_i, x \rangle e_i \right\|^2 + \left\| \sum_{i=1}^n \langle e_i, x \rangle e_i \right\|^2 \\ &= \left\| x - \sum_{i=1}^n \langle e_i, x \rangle e_i \right\|^2 + \sum_{i=1}^n |\langle e_i, x \rangle|^2 \end{aligned}$$

Before proceeding, we need to clarify some concepts about infinite sums. If  $\{c_a : a \in I\}$  is a family of positive numbers, then

$$\sum_{a \in I} c_a := \sup \left\{ \sum_{i=1}^n c_{a_i} : \{a_1, \dots, a_n\} \subseteq I \right\}.$$

□

COROLLARY 2.2. Let  $M = \text{span}\{e_1, \dots, e_n\}$ . Then,

$$\pi_M(x) = \sum_{i=1}^n \langle x, e_i \rangle e_i.$$

PROOF. The first assertion in Bessel's inequality says that  $x = u + v$  with  $u = \sum_{i=1}^n \langle x, e_i \rangle e_i \in M$  and  $v \in M^\perp$ , hence,  $u = \pi_M(x)$ . □

EXERCISE 2.10. Show that, if  $\sum_{a \in I} c_a < \infty$ , then  $\{a \in I : c_a \neq 0\}$  is at most countable. If  $\{c_a : a \in I\}$  is a family of complex numbers, we say that  $\sum_{a \in I} c_a$  converges absolutely if  $\sum_{a \in I} |c_a| < \infty$  converges. In this case we say that  $\sum_{a \in I} c_a$  converges in  $\mathbb{C}$  absolutely, hence irrespective of how  $I$  is ordered (prove it if you have never done it before!).

### 2.2.2. Fourier analysis and synthesis of a vector.

THEOREM 2.4. Let  $\{e_a\}_{a \in I}$  be a o.n.b. of  $H$ . Then,

(i) For all  $x$  in  $H$ ,

$$(2.2.1) \quad x = \sum_{a \in I} \langle e_a, x \rangle e_a \quad (\text{Analysis of } x)$$

converges in  $H$ ;

(ii) we have

$$(2.2.2) \quad \|x\|^2 = \sum_{a \in I} |\langle e_a, x \rangle|^2 \quad (\text{Plancherel Isometry});$$

(iii) if  $\{c_a\}_{a \in I}$  is a sequence in  $H$  such that  $\sum_{a \in I} |c_a|^2$  converges, then

$$(2.2.3) \quad \sum_{a \in I} c_a e_a \text{ converges in } H. \quad (\text{Analysis of } x)$$

PROOF. By Bessel inequality, if  $A \subseteq I$  with  $\#(A) < \infty$ , then  $\sum_{a \in A} |\langle e_a, x \rangle|^2 \leq \|x\|^2$ . As a consequence,  $I(x) := \{a \in I : \langle e_a, x \rangle \neq 0\} = \{a_n\}_n$  is countable, hence, (2.2.2) holds with  $\leq$ . Set now  $x_n = \sum_{j=1}^n \langle e_{a_j}, x \rangle e_{a_j}$ . We show that this defines a Cauchy sequence in  $H$ :

$$\|x_{n+l} - x_n\|^2 = \sum_{j=1}^l |\langle e_{a_j}, x \rangle|^2 \rightarrow 0$$

as  $n \rightarrow \infty$ , by the convergence of the series on the right of (2.2.2). Let  $y = \lim_{n \rightarrow \infty} x_n$  (the limit is taken in  $H$ -norm):

$$\begin{aligned} \langle x - y, e_{a_m} \rangle &= \left\langle x - \lim_{n \rightarrow \infty} \sum_{j=1}^n \langle e_{a_j}, x \rangle e_{a_j}, e_{a_m} \right\rangle \\ &= \lim_{n \rightarrow \infty} \left\langle x - \sum_{j=1}^n \langle e_{a_j}, x \rangle e_{a_j}, e_{a_m} \right\rangle \\ &= \lim_{n \rightarrow \infty} (\langle e_{a_m}, x \rangle - \langle e_{a_m}, x \rangle) = 0. \end{aligned}$$

The same argument shows that  $\langle x - y, e_a \rangle = 0$  if  $a \notin I(x)$ . Hence,  $x - y \perp e_a$  for all elements of the orthonormal basis, hence  $x = y$ , which shows (2.2.1).

We finish the proof of Plancherel's identity:

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \left\| x - \sum_{j=1}^n \langle e_{a_j}, x \rangle e_{a_j} \right\|^2 \\ &= \left( \|x\|^2 - \sum_{j=1}^n |\langle e_{a_j}, x \rangle|^2 \right) \\ &= \|x\|^2 - \sum_{n=1}^{\infty} |\langle e_{a_n}, x \rangle|^2 \\ &= \|x\|^2 - \sum_{a \in I} |\langle e_a, x \rangle|^2, \end{aligned}$$

as wished. □

### 2.2.3. Orthonormal basis in separable Hilbert spaces.

2.2.3.1. *Gram-Schmidt algorithm.* A Hilbert space is *separable* if it is separable as a metric space, i.e. if it has a countable, dense set. Most Hilbert spaces encountered in theory and applications are separable, and in this case the existence of an orthonormal basis is constructive. We will then treat first the separable case, then we will see how things work in the general case, where Zorn's Lemma is required, and we have no way to "see" the basis.

**THEOREM 2.5 (Gram-Schmidt).** *Let  $H$  be a separable Hilbert space. Then,  $H$  has a countable basis.*

**PROOF.** We start with a countable, dense subset  $G = \{g_n : n \geq 1\}$  of  $H$ . Inductively removing vectors linearly dependent from the previously chosen ones, we obtain a linearly independent subfamily  $F = \{f_n : n \geq 1\}$  such that  $\text{span}(F) = \text{span}(G)$  is still dense in  $H$ . We then transform  $F$  into an

orthonormal system:

$$\begin{aligned} e_1 &= \frac{f_1}{\|f_1\|}, \\ e_2 &= \frac{f_2 - \langle e_1, f_2 \rangle e_1}{\|f_2 - \langle e_1, f_2 \rangle e_1\|}, \\ &\dots \\ e_n &= \frac{f_n - \sum_{j=1}^{n-1} \langle e_j, f_n \rangle e_j}{\|f_n - \sum_{j=1}^{n-1} \langle e_j, f_n \rangle e_j\|}, \\ &\dots \end{aligned}$$

The denominators do not vanish because  $F$  is a linearly independent family, and inductively we see that  $\text{span}\{e_1, \dots, e_n\} = \text{span}\{f_1, \dots, f_n\}$ . Hence,  $\{e_n : n \geq 1\}$  is an orthonormal basis for  $H$ .  $\square$

EXERCISE 2.11. Let  $H = L^2[0, 1]$  with the Lebesgue measure, and consider the functions  $1, x, x^2$ . Apply Gram-Schmidt's algorithm to find three orthonormal vectors in  $H$ .

### 2.2.3.2. The classification of separable Hilbert spaces.

THEOREM 2.6. A Hilbert space  $H$  is separable if and only if it has one countable orthonormal basis  $\mathcal{B}$ . If it does, all o.n.b. of  $H$  have the same cardinality. Moreover,

- (i) if  $\sharp(\mathcal{B}) = d < \infty$ , then  $H$  is isometrically isomorphic to  $\mathbb{C}^d$ ;
- (ii) if  $\sharp(\mathcal{B}) = \infty$ , then  $H$  is isometrically isomorphic to  $\ell^2(\mathbb{N})$ .

PROOF. If  $H$  is separable, then Theorem 2.5 provides a countable o.n.b.  $\mathcal{B}$ . Suppose viceversa that  $H$  has a countable basis  $\mathcal{B}$ . If  $\sharp(\mathcal{B}) = d < \infty$ ,  $\mathcal{B} = \{e_1, \dots, e_d\}$ , then each  $x$  in  $H$  can be written as  $x = \sum_{j=1}^d \langle e_j, x \rangle e_j$  by the Analysis part of Theorem 2.4 (or by a much more elementary argument). The map

$$L_{\mathcal{B}} : x \mapsto (\langle e_j, x \rangle)_{j=1}^d$$

is an isometric isomorphism of  $H$  onto  $\mathbb{C}^d$ . Since  $H$  is finite dimensional, all its basis have the same dimension. In the countable case,  $\mathcal{B} = \{e_n\}_{n=1}^{\infty}$ , and we define  $L_{\mathcal{B}} : H \rightarrow \ell^2(\mathbb{N})$  in the same way,  $L_{\mathcal{B}} : x \mapsto \{\langle e_n, x \rangle\}_{n=1}^{\infty}$ . Theorem 2.4 implies that  $L_{\mathcal{B}}$  is a isometric isomorphism.

In the opposite direction, suppose  $H$  has a countable o.n.b., and consider the countable  $Q$  family of the vectors  $q = \sum_{j=1}^n q_j e_j$ , where  $n \geq 1$ , the  $q_j$ 's are complex rationals,  $q_j \in \mathbb{Q} + i\mathbb{Q}$ . If  $x = \sum_{j=1}^{\infty} c_j e_j$  and  $\epsilon > 0$ , then

$$\left\| \sum_{j=1}^{\infty} c_j e_j - \sum_{j=1}^n q_j e_j \right\|^2 = \sum_{j=n+1}^{\infty} |c_j|^2 + \sum_{j=1}^n |c_j - q_j|^2.$$

Choose first  $n$  such that the first sum is dominated by  $\epsilon$ , then  $q_1, \dots, q_n$  such that each summand in the second sum is dominated by  $\epsilon/n$ . This shows separability.

We are left with proving that, if one o.n.b.  $\mathcal{B}$  is infinite countable, then any other o.n.b.  $\mathcal{B}_1$  is. Let  $\mathcal{B}_1 = \{e_a\}_{a \in I}$ , and observe that, for any  $e_a \neq e_b$  in it,  $\|e_a - e_b\|^2 = 2$ . Let  $\{q_n\}_{n=1}^\infty$  be any countable, dense set in  $H$ . Then, for each  $a \in I$  there is  $n = n(a)$  such that  $\|e_a - q_n\| < \sqrt{2}/2$ , and if  $a \neq b$ , then  $n(a) \neq n(b)$ . This provides an injective map  $I \rightarrow \mathbb{N}$ , hence  $I$  is countable.  $\square$

#### 2.2.4. Intermezzo: orthonormal basis in general Hilbert spaces.

2.2.4.1. *Existence of o.n.b.* A widely used version of the Axiom of Choice is Zorn's Lemma, which we are stating below. We work with a partially ordered set  $(A, \leq)$ . A *chain* in  $A$  is a subset  $B \subseteq A$  on which the partial order is in fact a total order: for any  $x, y$  in  $B$ , either  $x \leq y$  or  $y \leq x$ . An element  $m$  in  $A$  is *maximal* for  $C \subseteq A$  if  $x \leq m$  for all  $x$  in  $C$  and if  $n \leq m$  has the same property, then  $n = m$ .

**THEOREM 2.7 (Zorn Lemma).** *Let  $(A, \leq)$  be a (nonempty) partially ordered set with the property that any chain  $B$  in  $A$  has a maximal element in  $A$ . Then,  $A$  has a maximal element.*

As a consequence, we have that an orthonormal basis exists for any Hilbert space.

**THEOREM 2.8.** *Every Hilbert space  $H \neq \{0\}$  has a orthonormal basis.*

**PROOF.** Let  $A$  be the set of all orthonormal systems of  $H$ , ordered by inclusion. It is nonempty because  $\{x/\|x\|\} \in A$  if  $x \neq 0$  is an element of  $H$ . If  $B$  is a chain in  $A$ , then

$$\mathcal{G} = \cup_{\mathcal{F} \in B} \mathcal{F} \in A$$

is a maximal element for  $B$ : it is an orthonormal system, and all orthonormal system containing all  $\mathcal{F}$  in  $B$  contain  $\mathcal{G}$ .

By Zorn's Lemma,  $A$  has a maximal element  $\mathcal{H}$  and, unravelling definitions,  $\mathcal{H}$  is a maximal orthonormal system in  $H$ , hence a basis.  $\square$

2.2.4.2. *The dimension of a Hilbert space.* By Theorem 2.4, if  $H$  be a Hilbert space and  $\mathcal{B} = \{e_a : a \in I\}$  an orthonormal basis for it, then the map

$$L_{\mathcal{B}} : x = \sum_{a \in I} \langle e_a, x \rangle e_a \mapsto \{\langle e_a, x \rangle\}_{a \in I}$$

is an isometric isomorphism from  $H$  onto  $\ell^2(I)$ .

A natural question is whether different orthonormal basis of  $H$  have the same cardinality, as we verified the separable case. The answer is positive.

**THEOREM 2.9.** *Let  $\{e_a : a \in I\}$  and  $\{f_b : b \in J\}$  be o.n.b. of the same Hilbert space  $H$ . Then,  $\sharp(I) = \sharp(J)$ .*

The cardinality in question is the *dimension* of the Hilbert space.

**COROLLARY 2.3.** *All separable Hilbert spaces which are not finite dimensional have countable dimension, and are isomorphic to each other.*

PROOF. If both  $I$  and  $J$  are finite dimensional, the statement is a well know fact from linear algebra (go back to see its proof!).

If  $I$  is finite dimensional then  $J$  is finite dimensional as well. For, suppose  $I$  is finite dimensional and let  $S = \{x \in H; \|x\| = 1\}$  be the unit sphere in  $H$ . Being  $H$  a finite dimensional Euclidean space, and  $S$  both closed and bounded,  $S$  is compact. Suppose now  $J$  is infinite and let  $\{f_{b_n} : n \geq 1\}$  an infinite countable subset of  $\{f_b : b \in J\}$ , which is contained in  $S$  if a subsequence of it (which we might denote the same way) converges to some  $x$  in  $S$ , by continuity of the inner product  $\langle x, f_{b_n} \rangle \rightarrow \|x\|^2 = 1$  as  $n \rightarrow \infty$ . On the other hand,  $x = \sum_{b \in J} \langle x, f_b \rangle f_b$ , hence  $\langle x, f_{b_n} \rangle \rightarrow 0$  as  $n \rightarrow \infty$ , because, for instance,  $\sum_n |\langle x, f_{b_n} \rangle|^2 \leq \|x\|^2 = 1$ .

The remaining case is that in which both  $I$  and  $J$  are infinite. To each  $a \in I$  there corresponds an at most countable subset  $J(a)$  of  $J$  such that

$$e_a = \sum_{b \in J(a)} \langle e_a, f_b \rangle f_b.$$

The number of the involved  $f_b$ 's is

$$\#(\cup_{a \in I} J(a)) \leq \#(J).$$

Suppose a basis element  $f_{b_0}$  was not used. Then,  $f_{b_0}$  is orthogonal to all  $e_a$ 's, hence to  $H$ : contradiction.

We have then that  $\#(J) \leq \#(I)$ , and the opposite inequality holds as well.  $\square$

### 2.3. The trigonometric system and Fourier series

In item (IX) of section 1.3.2 we saw that the trigonometric system  $\{e_n\}_{n=-\infty}^{\infty}$ , where

$$(2.3.1) \quad e_n(t) = \frac{e^{int}}{2\pi}$$

is a complete orthonormal system (an orthonormal basis) for  $L^2(\mathbb{T})$ .

It should be kept in mind that a great amount of work goes on all over the world into finding, studying the special properties of, and applying, new orthonormal basis (ONB) for old and new Hilbert spaces of functions. Broadly speaking, when the functions  $e_n$  in an ONB share some qualitative feature, so do their linear combinations; hence, using that ONB to approximate a function  $f$  means finding closer and closer approximations of  $f$  which "look like" the functions in the basis. To wit, "if all you have in your toolbox is a hammer, then everything you look at seems a hammer". The trigonometric system, however, remains the most natural one, because of its interaction with the group structure of the torus,

$$(2.3.2) \quad e^{in(s+t)} = e^{ins} e^{int}.$$

This is the reason why the trigonometric system *diagonalizes translations*. Let  $\tau_a f(t) = f(t - a) \pmod{2\pi}$ :

$$\tau_a e_n = e^{-ina} e_n;$$

hence the derivative operator, which is a linear combination of "infinitely close" translations,

$$i^{-1} \frac{d}{dt} e_n(t) = \lim_{h \rightarrow 0} \frac{\tau_{-h} e_n(t) - \tau_0 e_n(t)}{ih} = \lim_{h \rightarrow 0} \frac{e^{inh} - 1}{ih} e_n(t) = n e_n(t).$$

In fact, the trigonometric system diagonalizes all "reasonable" linear operators on  $L^2(\mathbb{T})$  which commute with translations. The main example is that of the convolution operators  $f \mapsto f * g$ .

Next, we translate into the special context of the trigonometric system the general Hilbert spaces results we have seen.

(I) **Fourier analysis.** For  $f \in L^2(\mathbb{T})$  and  $N \geq 1$ , let

$$S_N f(t) = \frac{1}{2\pi} \sum_{|n| \leq N} \widehat{f}(n) e^{int},$$

Then.

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

i.e.

$$(2.3.3) \quad f(t) = S_\infty f(t) := \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{int},$$

where the convergence is in  $L^2(\mathbb{T})$ .

(II) **Fourier synthesis.** Let  $\{c_n\}_{n=-\infty}^{\infty} \in \ell^2(\mathbb{Z})$ ,

$$\|\{c_n\}_{n=-\infty}^{\infty}\|_{\ell^2}^2 = \sum_{n \in \mathbb{Z}} |c_n|^2 < \infty.$$

Then, there exists  $f \in L^2(\mathbb{T})$  such that

$$\widehat{f}(n) = c_n.$$

Explicitly,

$$(2.3.4) \quad f(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} c_n e^{int},$$

where the series converges in the  $L^2$  norm.

(III) **Plancherel formula.**

$$\|f\|_{L^2(\mathbb{T})}^2 = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} |c_n|^2.$$

The foundational problem of Fourier theory is understanding for which  $f$ , and in which sense, the series in (2.3.3) converges. For  $f \in L^2(\mathbb{T})$ , we have considered a quantitative answer, using the fact that  $\{e_n\}_{n=-\infty}^{+\infty}$  is an orthonormal basis, and the general theory of Hilbert spaces.

Fourier series provided, in the second half of the XIX century, the motivation for some of the most important innovations in mathematical analysis and in the foundations of mathematics. In S.M. Srivastava [How did Cantor Discover Set Theory and Topology?](#) [Resonance, Vol. 19, p. 977-999, (2014)], for instance, you can find the history of how Cantor came to develop the theory of infinite sets to solve open problems concerning the convergence Fourier series.

It is noteworthy the role of Lebesgue integration. For the *Fourier analysis* of a function, the definition of integral provided by Riemann was considered wholly satisfactory, as it enabled to compute the Fourier coefficients of all functions mathematicians of the age could think of (it was precisely to compute Fourier series that Riemann invented the definition). The problem is with *Fourier synthesis*, since it was not obvious (and in fact it is not true!) that the function  $f$  designed by (2.3.4) is Riemann integrable. The problem is that, as it came to be gradually realized, integrability in the sense of Riemann is not preserved under many limiting procedures, including those which are basilar in Fourier theory. Borel, then Lebesgue came, and this foundational problem was solved.



## CHAPTER 3

### Banach spaces

In this chapter we introduce some basic tools in Banach theory, *Hahn-Banach extension theorem* and *Baire's category theorem and its consequences*. The latter are, on the one hand, *(Banach-Steinhaus) uniform boundedness principle*, and on the other the sequence: *open mapping theorem*, *inverse mapping theorem*, *closed graph theorem*. We will also consider some consequences and applications of these cornerstones of functional analysis.

#### 3.1. Zorn's lemma and some of its consequences

In the proof of Hahn-Banach Theorem we will use Zorn's lemma, which is a consequence of the axiom of choice, and it is in fact equivalent to it. In most courses Zorn's lemma is just stated as a principle, then used alongside the axiom of choice. One reason is that most instructors hope that the relationship between the two statements has been clarified in some other course. The other, probably, is that most instructors are research mathematicians with little taste for nonconstructive principles: they know them and use them, but feel that such principles are a sort of magic of last resort, and having a magic tool generating others, or having a supply of independent ones, does not really make to them much of a difference. A problem with highly nonconstructive existence theorems, in fact, is that we do not know much of the object we have proved the existence of. It is surprising, in view of this fact, that such theorems can be used to obtain more practical statements.

There is a price tag attached to it. Nonconstructive arguments allow us to prove the existence of counterexamples to very reasonable guesses. If we allow the logical basis of the nonconstructive arguments, we have to keep in mind that such monsters exist and have to be taken into account.

So much for the small talk. Before we move to Banach theory, in order to satisfy readers with a taste for logics we state Zorn's lemma and prove that it follows from the axiom of choice (but not the opposite implication).

**THEOREM 3.1 (Zorn's lemma).** *Let  $(X, \leq)$  be a partial order on a set  $X$ , and suppose that each totally ordered subset  $C$  of  $X$  has an upper bound: there is  $u$  in  $X$  such that  $u \geq c$  for all  $c \in C$ . Then,  $X$  has a maximal element  $w$ :  $w \geq x$  for all  $x$  in  $X$ .*

PROOF. While reading the proof, it is very helpful drawing pictures. A partial order is a special instance of a directed graph, and a totally ordered subset of it might be thought of, in pictorial terms, as a subset of the real line (although it is clearly false that all totally ordered sets arise this way, for instance because there is no a priori restriction on their cardinality).

Suppose by contradiction that  $X$  has no maximal element. In particular, by the axiom of choice, for each totally ordered subset  $C$ , including  $C = \emptyset$ , we can choose an upper bound  $u(C)$  in  $X \setminus C$  (if the only upper bound, which exists by hypothesis, belonged to  $C$ , then it would be maximal for  $X$ ). Now, to each subset  $F$  of  $X$  and  $f \in F$  we associate its *tail set*  $F_{<f} = \{x \in F : x < f\}$ .

A subset  $A$  is a  $u$ -set if

- (i)  $A$  is totally ordered;
- (ii)  $A$  contains no infinite descending sequence  $a_1 > a_2 > \dots$  (e.g.  $A$  can have the order type<sup>1</sup> of  $\mathbb{N}$ , but not that of  $-\mathbb{N}$ ), hence, any subset  $A'$  of  $A$  has minimum;
- (iii) if  $a \in A$ , then  $u(A_{<a}) = a$ .

The last property says that  $u$ -sets are linked in a very special way to the choice function  $u$ , and, as we will see below, they share many properties with the ordinals.

If  $A$  is a  $u$ -set and  $a \in A$ , then  $A_{<a}$  has properties (i-iii), hence, it is a  $u$ -set.

If  $A$  and  $B$  are  $u$ -sets, by properties (i) and (ii) they have minimum, and  $\{\min(A)\}, \{\min(B)\}$  are  $u$ -sets. By property (iii),

$$\min(A) = u(\{\min(A)\}_{<\min(A)}) = u(\emptyset) = \dots = \min(B).$$

The proof of Zorn's lemma mostly consists in a "transfinite iteration" of this argument, to show that all  $u$ -sets are tail sets of a largest  $u$ -set  $E$ , and by further applying  $u$  to it we reach a contradiction.

**Claim** Let  $A \neq B$  be  $u$ -sets. Then,  $A = B_{<b}$  for some  $b \in B$ , or  $B = A_{<a}$  for some  $a \in A$ .

Let

$$C = \{c \in A \cap B : A_c = B_c\}.$$

**Sub-claim**  $C = A$  or  $C = A_{<a}$  for some  $a \in A$ .

---

<sup>1</sup>By *order type* of a partially ordered set  $(Z, <)$  we mean the equivalence class of all linearly ordered sets which are in an order-preserving bijection with  $Z$ .

We show the sub-claim first. Suppose  $C \neq A$  and let  $a \in A \setminus C$  be minimal, so that  $A_{<a} \subseteq C$ . In the opposite direction, if  $c \in C \setminus A_{<a} \subseteq A$ , then  $c > a$ . Hence,

$$a \in A_c = B_{<c} \subseteq B,$$

thus  $a \in B$  and  $A_{<a} = B_{<a}$ , implying that  $a \in C$ , which is a contradiction. This implies that  $C = A_{<}$ .

We now prove the claim. If  $A \neq C \neq B$ , by the sub-claim we have that  $A_{<a} = C = B_{<b}$  for some  $a \in A$ ,  $b \in B$ . By property (iii) of the  $u$ -sets,  $a = u(A_{<a}) = u(B_{<b}) = b$ , and we have that

$$a = b \in A \cap B, \text{ and } A_{<a} = B_{<b}.$$

By definition, then,  $a \in C$ , but this contradicts the fact that  $A_{<a} = C$ . The assumption  $A \neq C \neq B$  can not hold, hence we have that, again by the sub-claim, either  $A = C = B_{<b}$  or, the other way around,  $B = C = A_{<a}$ . (Hidden in the proof is that the cases  $a = \min(A)$  or  $b = \min(B)$ , in which  $C = \emptyset$ , are covered because we defined  $u(\emptyset)$ ).

The claim allows us to glue together whatever  $u$ -sets we have into one. Let

$$E = \cup_{A \text{ a } u\text{-set}} A.$$

We first show that, if  $A$  is a  $u$ -set and  $a \in E$ , then  $A_{<a} = E_{<a}$ . The direction  $A_{<a} \subseteq E_{<a}$  follows from  $A \subseteq E$ . In the other direction, let  $x \in E_{<a}$  and let  $B \ni x$  be a  $u$ -set. If  $B \subseteq A$ , then  $x \in A$ . If not, by the claim  $A = B_{<b}$  for some  $b \in B$ , in which case  $x < a < b$ , and so  $x \in B_{<b} = A$ . In both cases,  $x \in A$ , hence,  $E_{<a} \subseteq A$ , which clearly implies  $E_{<a} \subseteq A_{<a}$ .

It is now easy to verify that  $E$  is a  $u$ -set.

- (i) Let  $a, b \in E$ , and  $A \ni a$ ,  $B \ni b$  be  $u$ -sets. Then  $A_{<a} = E_{<a}$  and  $B_{<b} = E_{<b}$  are  $u$ -sets, hence, by the claim above, either they are equal, then  $a = u(A_{<a}) = u(B_{<b}) = b$ , or  $A_{<a} = B_{<c}$  for some  $c$  in  $B_{<b}$ , and  $c = u(B_{<c}) = u(A_{<a}) = a$ , or the other way around. In the second case,  $a = c < b$ , and in the remaining case the opposite inequality holds. Hence,  $E$  is totally ordered.
- (ii) Suppose  $a_1 > a > 2 > \dots$  is a descending sequence in  $E$ , and that  $a_1 \in A$ , a  $u$ -set. For  $i > 1$ ,  $a_i \in E_{a_1} = A_{a_1}$ , hence the sequence, being contained in  $A$ , can not be infinite.
- (iii) If  $a \in E$  and  $A \ni a$  is a  $u$ -set, then  $u(E_{<a}) = u(A_{<a}) = a$ .

Since  $E$  is a  $u$ -set,  $E \cup \{u(E)\}$  is still a  $u$ -set which is not contained in  $E$ , and we have reached a contradiction.  $\square$

In the spirit of the section, we use Zorn's lemma to prove a purely algebraic result.

**THEOREM 3.2.** *[Hamel's basis of a vector space] Any vector space  $X$  over a field  $\mathbb{F}$  has a basis. That is, there is a set  $\{v_a\}_{a \in I}$  of linearly independent vectors of  $X$  such that  $X = \text{span}(\{v_a\}_{a \in I})$ .*

**PROOF.** Consider the set  $\mathcal{V}$  of all families  $V = \{v_b\}_{b \in J}$  of linearly independent vectors of  $X$ , partially ordered by inclusion. If  $\{V_\alpha\}_{\alpha \in H}$  is a totally ordered subfamily in  $\mathcal{V}$ , then  $\bigcup_{\alpha \in H} V_\alpha$  is an element of  $\mathcal{V}$ , and it is an upper bound for  $\{V_\alpha\}_{\alpha \in H}$ . By Zorn's lemma, there exists in  $\mathcal{V}$  a maximal element  $\{v_a\}_{a \in I}$ , which is a basis for  $X$ , since a vector  $u$  which is not in  $\text{span}(\{v_a\}_{a \in I})$  could be added to the family, contradicting maximality.  $\square$

Below we use basic facts about cardinal numbers. For properties of cardinal numbers, and of ordinal ones, you might look into [Reed Solomon, Notes on ordinals and cardinals](#).

**THEOREM 3.3.** *Let  $X$  be a vector space over a field  $\mathbb{F}$ , and let  $U = \{U_a\}_{a \in I}$  and  $V = \{v_b\}_{b \in J}$  be two basis of  $X$ . Then,  $\sharp(I) = \sharp(J)$ ,  $I$  and  $J$  have the same cardinality.*

**PROOF.** We know from linear algebra that  $\sharp(I)$  is finite if and only if  $\sharp(J)$  is, and they coincide. We can suppose then that both index sets are infinite.

To each  $a \in I$  associate the set  $B(a) := \{b_1(a), \dots, b_{n(a)}(a)\} = \{b_1, \dots, b_n\} \subset J$ , where

$$U_a = \sum_{i=1}^n \lambda_i V_{b_i},$$

with  $0 \neq \lambda_i \in \mathbb{F}$ , is the unique expression of  $U_a$  with respect to the basis  $V$ . The map  $B$  is one to one from  $I$  to the family of the finite subsets of  $J$ . Since the latter class has the same cardinality as  $J$ , we have  $\sharp(I) \leq \sharp(J)$ , and the other inequality is proved in the same way.  $\square$

By theorem 3.3, we can attach to any vector space  $X$  the cardinality of one, hence all, of its basis. Such number is called the *algebraic dimension*, or sometimes the *Hamel dimension*, of  $X$ . Let's denote it by  $\dim_{al}(X)$ . The algebraic dimension is clearly a linear invariant: if  $L : X \rightarrow Y$  is a linear bijection, then  $\dim_{al}(X) = \dim_{al}(Y)$ .

Infinite dimensional, but separable, Hilbert spaces have a countable orthonormal basis in the Hilbert sense, but their algebraic dimension is more than countable. This shows that Hilbert theory, and more general Banach theory, or even more generally topological vector spaces theory, is not reducible to linear algebra alone, even at the most basic level of dimension.

PROPOSITION 3.1. *As a vector space, any infinite dimensional Hilbert space  $H$  does not have countable dimension.*

PROOF. Let  $\{v_n\}_{n=1}^\infty$  be any countable, linearly independent subset of  $H$ . Using Gram-Schmidt algorithm, find an orthonormal set  $\{e_m\}_{m=1}^\infty$  such that  $\text{span}(\{v_n\}_{n=1}^\infty) = \text{span}(\{e_m\}_{m=1}^\infty)$ . Consider then

$$x = \sum_{m=1}^{\infty} \frac{1}{m} e_m.$$

Then,  $x \in H \setminus \text{span}(\{e_m\}_{m=1}^\infty) = H \setminus \text{span}(\{v_n\}_{n=1}^\infty)$ . This shows that  $\{v_n\}_{n=1}^\infty$  was not an algebraic basis for  $H$ .  $\square$

The completeness of  $H$  was used in that  $x$  exists in  $H$  since the series converges.

Using Hamel basis, we can construct such monstrous objects as everywhere defined, unbounded linear functionals on a Hilbert space.

PROPOSITION 3.2. *Let  $H$  be an infinite dimensional Hilbert space. Then, there exists a linear functional  $l : H \rightarrow \mathbb{R}$  which is not bounded.*

PROOF. Let  $\{v_\alpha\}_{\alpha \in I}$  be a Hamel basis for  $H$ , let  $\{v_{\alpha_n}\}_{n=1}^\infty$  a countable subset of it, which we might assume to be orthonormal after applying the Gram-Schmidt algorithm to it. We then define  $l$  on the basis by

$$\begin{cases} l(v_\alpha) = 0 & \text{if } \alpha \neq \alpha_n \text{ for all } n \geq 1 \\ l(v_{\alpha_n}) = n. \end{cases}$$

The operator extends to a linear operator  $l$  on  $H$  by linear algebra. On the other hand,  $l$  is not bounded,

$$\|l\|_{H^*} := \sup_{\|x\|=1} |l(x)| \geq |l(v_{\alpha_n})| = n,$$

hence,  $\|l\|_{H^*} = \infty$ .  $\square$

### 3.2. The Hahn-Banach Theorem and some of its consequences

EXERCISE 3.1. *Let  $H$  be a Hilbert space,  $M \subset H$  a closed, linear space, and  $S : M \rightarrow \mathbb{C}$  a (bounded) linear functional on  $M$ . Show that there exists an extension  $T$  of  $S$  to  $H$  with  $\|T\| = \|S\|$ . Moreover, such an extension is unique.*

In general Banach spaces matters are more intricate.

**THEOREM 3.4** (Real Hahn-Banach Theorem). *Let  $X$  be a real vector space, and  $p : X \rightarrow \mathbb{R}$  a real-valued, convex functional:*

$$p(ax + (1 - a)y) \leq ap(x) + (1 - a)p(y)$$

*if  $x, y \in X$  and  $a \in [0, 1]$ . Let  $l : Y \rightarrow \mathbb{R}$  be a linear functional such that for  $x \in Y$*

$$l(y) \leq p(y).$$

*Then, there exists linear  $L : X \rightarrow \mathbb{R}$  extending  $l$  to  $X$ , and such that*

$$L(x) \leq p(x)$$

*on  $X$ .*

**PROOF.** The main step consist in extending  $l$  on  $Y$  to  $\lambda$  on  $\text{span}(Y, z)$ , where  $z \notin Y$ , which we do by a sort of separation of variables. Then we use Zorn's Lemma. We must have  $\lambda(y + az) = l(y) + a\lambda(z)$ , and it suffices to determine  $\lambda(z)$ . We start with an inequality involving  $z$  and  $p$ . Let  $a, b > 0$ ,  $x, y \in Y$ :

$$\begin{aligned} bl(x) + al(y) &= (b + a)l\left(\frac{b}{b + a}x + \frac{a}{b + a}y\right) \\ &\leq (b + a)p\left(\frac{b}{b + a}x + \frac{a}{b + a}y\right) \\ &= (b + a)p\left(\frac{b}{b + a}(x - az) + \frac{a}{b + a}(y + bz)\right) \\ &\leq bp(x - az) + ap(y + bz). \end{aligned}$$

That is,

$$\frac{l(x) - p(x - az)}{a} \leq \frac{p(y + bz) - l(y)}{b}$$

holds for all  $a, b, x, y$  as above. There is then real  $k$  such that  $l(x) - p(x - az) \leq ka$  and  $kb \leq p(y + bz) - l(y)$ , or

$$l(x) - ka \leq p(x - az), \quad l(y) + kb \leq p(y + bz).$$

It follows that  $\lambda(z) = k$  works.

Consider now the set  $A$  having as element  $(Z, \lambda)$  where  $Z$  is a subspace of  $X$  containing  $Y$  and  $\lambda$  is an extension of  $l$  to  $Z$  in such a way  $\lambda(z) \leq p(z)$  on  $Z$ . We partially order  $A$  saying that  $(Z, \lambda) \leq (W, \mu)$  is  $Z$  is a subspace of  $W$  and  $\mu$  extends  $\lambda$ .

Given a chain  $\{(Z_c, \lambda_c) : c \in C\}$  in  $A$ , we have that  $Z = \cup_{c \in C} Z_c$  and  $\lambda$  such that  $\lambda|_{Z_c} = \lambda_c$  are a well defined element in  $A$ , and  $(Z, \lambda)$  is maximal

for  $C$ . By Zorn's Lemma, there is a maximal  $(W, \nu)$  in  $A$ . Now, if  $W \neq X$ , we might apply the one-dimensional extension, and  $(W, \nu)$  would not be maximal. Hence,  $W = X$  and the theorem is proved.  $\square$

When proving Hahn-Banach's theorem for a separable Banach space  $X$  with  $p(x) = \|x\|$ , as it is often the case, Zorn's lemma is not necessary. In fact, we can proceed in a way which is much similar to the Gram-Schmidt construction of an orthonormal basis for a Hilbert space.

**EXERCISE 3.2.** Let  $(X, \|\cdot\|)$  be a separable Banach space and let  $\{x_n\}_{n=1}^\infty$  be a dense set. Let  $Z$  be a subspace of  $X$  and let  $l : Z \rightarrow \mathbb{R}$  be a linear functional such that  $|l(x)| \leq K\|x\|$  for some constant  $K \geq 0$ .

- (i) Let  $Z_0 = \bar{Z}$  be the closure of  $Z$  in  $X$ . Show that  $l$  has a unique extension  $l_0$  to  $Z_0$  with  $|l_0(x)| \leq K\|x\|$ .
- (ii) Let  $x_{i_1}$  be the first  $x_n$  in the sequence such that  $x_n \notin Z_0$ , and let  $Z_1 = \text{span}(Z_0, x_{i_1})$ . Verify that the inductive step in the proof of the real Hahn-Banach theorem shows that  $l_0$  extends to a linear functional  $l_1$  on  $Z_1$  such that  $|l_1(x)| \leq K\|x\|$ .
- (iii) Write down how to iterate the procedure in (ii). What happens if it stops after finitely many steps?
- (iv) Suppose the procedure can be iterated indefinitely, exhausting  $\{x_n\}$ , and let  $Z_\infty = \bigcup_{m=0}^\infty Z_m$ , and  $l_\infty : Z_\infty \rightarrow \mathbb{R}$  be such that  $l_\infty|_{Z_m} = l_m$ . Show that  $Z_\infty$  is dense in  $X$  and  $|l_\infty(x)| \leq K\|x\|$ . Use then again (i) to show that  $l_\infty$  has a unique extension  $\lambda$  to  $X$  which coincides with  $l_\infty$  on  $Z_\infty$  (hence, it coincides with  $l$  on  $Z$ ), and  $|\lambda(x)| \leq K\|x\|$ .

**THEOREM 3.5 (Complex Hahn-Banach Theorem).** Let  $X$  be a complex vector space and  $p : X \rightarrow \mathbb{R}$  be such that

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

whenever  $a, b \in \mathbb{C}$ ,  $|a| + |b| = 1$ , and  $x, y \in X$ . Let  $l : Y \rightarrow \mathbb{C}$  be a linear functional defined on a subspace  $Y$  of  $X$ , satisfying  $|l(y)| \leq p(y)$ . Then, there exists a linear functional  $L : X \rightarrow \mathbb{C}$  which extends  $l$ , and such that  $|L(x)| \leq p(x)$  on  $X$ .

**PROOF.** We consider  $X$  as a real space of "twice the dimension". Let  $\lambda(x) = \text{Re}(l(x))$ , a real functional on  $Y$ , and observe that

$$\lambda(ix) = \text{Re}(l(ix)) = \text{Re}(il(x)) = -\text{Im}(l(x)),$$

so that  $l(x) = \lambda(x) - i\lambda(ix)$  reconstructs  $l$  from  $\lambda$ . By the real Hahn-Banach theorem, there exists a real functional  $\Lambda : X \rightarrow \mathbb{R}$  which extends  $\lambda$  to  $X$ , and such that  $\Lambda(x) \leq p(x)$ . We define  $L(x) = \Lambda(x) - i\Lambda(ix)$ , which defines a complex linear functional on  $X$ :

$$L(ix) = \Lambda(ix) - i\Lambda(-x) = i(\Lambda(x) - i\Lambda(ix)) = iL(x).$$

We also have, for some real  $t$ ,

$$|L(x)| = e^{it}L(x) = L(e^{it}x) = \Lambda(e^{it}x) \leq p(e^{it}x) = p(x),$$

where the third equality holds because  $L(e^{it}x)$  is real, and the last follows from the fact that  $p(e^{it}x) = p(x)$  for all  $t$ 's.  $\square$

### 3.3. The dual of a Banach space

Let  $X$  be a Banach space over  $\mathbb{C}$  (virtually all we say holds for real Banach spaces). Its *dual space*  $X^*$  is the space of the continuous linear functionals on  $X$ . It is a Banach space in itself. In fact more is true. Let  $\mathcal{L}(X, Y)$  be the linear space of the bounded, linear operators  $T : X \rightarrow Y$ , normed with the operator norm,

EXERCISE 3.3. *Show that the operator norm is in fact a norm on  $\mathcal{L}(X, Y)$ .*

THEOREM 3.6. *Let  $X, Y$  be Banach spaces. Then,  $\mathcal{L}(X, Y)$  is complete.*

PROOF. Let  $\{A_n : n \geq 1\}$  be Cauchy in  $\mathcal{L}(X, Y)$ , so that  $\{A_n x : n \geq 1\}$  is Cauchy in  $Y$  for each fixed  $x$  in  $X$ , and  $Ax = \lim_{n \rightarrow \infty} A_n x$  is well defined, and linear. Also,  $|\|A_m\| - \|A_n\|| \leq \|A_m - A_n\|$ , hence  $\|A_n\| \rightarrow C$  has real limit as  $n \rightarrow \infty$ . We have:

$$\|Ax\| = \lim_{n \rightarrow \infty} \|A_n x\| \leq \lim_{n \rightarrow \infty} \|A_n\| \|x\| \leq C \|x\|,$$

hence,  $\|A\| \leq C$ .

We still have to show that  $A_n \rightarrow A$  in  $\mathcal{L}(X, Y)$  norm. But we have:

$$\|(A - A_m)x\| = \lim_{n \rightarrow \infty} \|(A_n - A_m)x\| \leq \lim_{n \rightarrow \infty} \|A_n - A_m\| \cdot \|x\|,$$

hence,

$$\|A - A_m\| \leq \lim_{n \rightarrow \infty} \|A_n - A_m\| \leq \epsilon$$

if  $m \geq m(\epsilon)$ , so  $\|A - A_m\| \rightarrow 0$  as  $m \rightarrow \infty$ .  $\square$

EXERCISE 3.4. *Let  $S \in \mathcal{L}(X, Y)$  and  $T \in \mathcal{L}(Y, Z)$ . Show that  $TS := T \circ S \in \mathcal{L}(X, Z)$  satisfies  $\|TS\| \leq \|T\| \|S\|$ .*

The above inequality, when  $X = Y = Z$ , expresses the fact that  $\mathcal{L}(X) := \mathcal{L}(X, X)$  is a *Banach algebra* with respect to the composition product. It is also a *unital* one, since the identity operator  $Ix = x$  satisfies  $IA = AI$  and  $\|I\| = 1$ .

We will often use without mention the following fact.

EXERCISE 3.5. *Let  $f_0 : X_0 \rightarrow Y$  be a continuous map from a dense subspace  $X_0$  of a metric space  $X$  and a complete metric space  $Y$ . Then,  $f_0$  can be uniquely extended to a continuous map  $f : X \rightarrow Y$ . If  $X$  is a normed space,  $Y$  is a Banach space, and  $f_0$  is linear, then  $f$  is linear and  $\|f\|_{\mathcal{L}(X, Y)} = \|f_0\|_{\mathcal{L}(X_0, Y)}$ .*



On the space  $\mathcal{L}(X, Y)$  we will make use of three different topologies, which naturally arise from the theory of operators and its applications.

(*UOT*) The *uniform operator topology* (*UOT*) is the one induced by the operator norm:

$$T_n \xrightarrow{UOT} T \text{ if and only if } \|T_n - T\| \rightarrow 0.$$

(*SOT*) The *strong operator topology* (*SOT*) is defined by:

$$T_n \xrightarrow{SOT} T \text{ if and only if } \|T_n x - T x\| \rightarrow 0$$

for all  $x$  in  $X$ .

(*WOT*) The *weak operator topology* (*WOT*) is defined by:

$$T_n \xrightarrow{WOT} T \text{ if and only if } l(T_n x - T x) \rightarrow 0$$

for all  $x$  in  $X$  and all  $l \in Y^*$ .

The Hahn-Banach Theorem has important consequences about  $X^*$ .

**COROLLARY 3.1.** *Let  $i : Y \rightarrow X$  be the inclusion in the Banach space  $X$  of  $Y$ , a subspace of it. Then, the restriction map  $l \mapsto l \circ i$  from  $X^*$  to  $Y^*$  is surjective: each  $\lambda$  in  $Y^*$  has an extension  $l$  in  $X^*$ . Moreover, there exists one extension  $l$  such that  $\|l\|_{X^*} = \|\lambda\|_{Y^*}$ .*

**PROOF.** Apply the Hahn-Banach Theorem with  $p(x) = \|\lambda\|_{Y^*} \cdot \|x\|$ .  $\square$

**COROLLARY 3.2.** *Let  $y \in X$ , Banach. Then, there exists  $l \in X^*$ ,  $l \neq 0$ , such that  $l(y) = \|l\|_{X^*} \|y\|$ .*

**PROOF.** Set  $\lambda(ay) = a\|y\|$  on  $\text{span}(y)$ , and extend it to  $l : X \rightarrow \mathbb{C}$  by Hahn-Banach. Since  $|\lambda(ay)| = \|ay\|$ , we have  $\|l\| = \|\lambda\| = 1$ , and  $l(y) = \lambda(y) = \|y\|$ .  $\square$

**COROLLARY 3.3.** *Let  $Z$  be a subspace of a normed linear space  $X$ , and let  $d(y, Z) := \inf\{\|y - z\| : z \in Z\}$ . Then, there is  $L \in X^*$  such that  $L(y) = d(y, Z)$ ,  $\|L\|_{X^*} = 1$ , and  $L(z) = 0$  for  $z \in Z$ .*

**PROOF.** We only have to define such  $L$  on  $\text{span}(Z, y)$ , then use Hahn-Banach. We are forced to define  $L(z + ay) = aL(y) = ad(y, Z)$ . This is a linear functional, and

$$\frac{|L(z + ay)|}{\|z + ay\|} = |a| \frac{d(Z, y)}{\|z + ay\|}$$

$$\begin{aligned}
&= |a| \frac{\inf\{\|w - y\| : w \in Z\}}{\|z + ay\|} \\
&= |a| \frac{\inf\{\|w - y\| : w \in Z\}}{|a| \cdot \left\| -\frac{z}{a} - y \right\|} \\
&\leq 1,
\end{aligned}$$

because  $-z/a \in Z$ . Hence,  $\|L\|_{X^*} \leq 1$ .

To show the opposite inequality, for  $\epsilon > 0$  find  $z$  in  $Z$  such that  $\|y - z\| \leq d(y, Z) + \epsilon$ , so that

$$\frac{|L(z - y)|}{\|z - y\|} \geq \frac{d(y, Z)}{d(y, Z) + \epsilon}.$$

Since  $\epsilon > 0$  is arbitrary,  $\|L\|_{X^*} \geq 1$ . □

EXERCISE 3.6. Give "Hilbertian" elementary proofs of the three corollaries above in the case when  $X$  is a Hilbert space.

The *bidual* of a Banach space  $X$ ,  $X^{**}$ , is the dual of  $X^*$ .

THEOREM 3.7. Let  $X$  be a Banach space and consider the map  $i : X \rightarrow X^{**}$  given by  $[i(x)](\ell) := \ell(x)$  whenever  $x \in X$  and  $\ell \in X^*$ . Then,  $i$  is an isometric imbedding of  $X$  into  $X^{**}$ .

PROOF. We have:

$$|[i(x)](\ell)| = |\ell(x)| \leq \|\ell\|_{X^*} \|x\|_X,$$

hence,  $\|[i(x)]\|_{X^{**}} \leq \|x\|_X$ .

In the other direction, by Corollary 3.2 we can find  $l$  in  $X^*$  with  $\|l\|_{X^*} = 1$ , and such that  $l(x) = \|x\|$ . i.e.,

$$\|i(x)\|_{X^{**}} \geq \frac{|[i(x)](l)|}{\|l\|_{X^*}} = \frac{|l(x)|}{\|l\|_{X^*}} = \|x\|.$$

□

When the map  $i : X \rightarrow X^{**}$  is surjective (hence, a surjective isometry), we say that  $X$  is *reflexive*. Not all Banach spaces are reflexive, as the following exercise shows.

EXERCISE 3.7. Here  $\ell^p = \ell^p(\mathbb{N})$ .

- (i) Show that for each  $\varphi \in \ell^\infty$  the map  $L_\varphi : \psi \mapsto \sum_{n=0}^\infty \varphi(n) \overline{\psi(n)}$ ,  $L_\varphi : \ell^1 \rightarrow \mathbb{C}$ , is a continuous linear functional in  $(\ell^1)^*$  and  $\|L_\varphi\|_{(\ell^1)^*} = \|\varphi\|_\infty$ .
- (ii) Given  $L \in (\ell^1)^*$ , show that there exists  $\varphi \in \ell^\infty$  s.t.  $L = L_\varphi$ . Moreover,  $\|\varphi\|_\infty = \|L\|_{(\ell^1)^*}$ . **Hint.** Make use of  $L(\delta_n)$  as "building blocks" to construct  $\varphi$ . This way, we have isometrically identified  $(\ell^1)^* \equiv \ell^\infty$ .

(iii) Consider the subspace  $C$  of  $\ell^\infty$ ,

$$C = \{\varphi \in \ell^\infty : \exists \lim_{n \rightarrow \infty} \varphi(n)\}.$$

Show that  $C$  is a closed subspace of  $\ell^\infty$ , and that  $L : \varphi \mapsto \lim_{n \rightarrow \infty} \varphi(n)$  is a closed linear functional on  $C$ .

(iv) Show that there is no  $\psi \in \ell^1$  such that  $C(\varphi) = \sum_{n=0}^\infty \psi(n)\varphi(n)$ . Deduce from this that  $\ell^1$  is not reflexive.

EXERCISE 3.8. Let  $\ell_c$  be the space of the sequences  $h : \mathbb{N} \rightarrow \mathbb{C}$  which vanish outside a finite set, and  $\ell$  be the space of all sequences  $h : \mathbb{N} \rightarrow \mathbb{C}$ .

(i) For  $m : \mathbb{N} \rightarrow \mathbb{C}$ , define the multiplication operator with symbol  $m$ ,  $T_m : \ell_c \rightarrow \ell$ , by  $T_m(h)(n) = m(n)h(n)$ . Show that  $\|T_m\|_{\mathcal{B}(\ell^2, \ell^2)} = \|m\|_{\ell^\infty}$ .

(ii) Find a similar statement with  $L^2[0, 1]$  instead of  $\ell^2$ , and prove it.

EXERCISE 3.9. Consider  $\ell^1 = \ell^1(\mathbb{N})$ , its 1-dimensional, closed subspace  $\text{span}(\delta_0)$ , and  $l(a\delta_0) = a$ . Find all extensions  $L$  of  $l$  to a linear functional on  $\ell^1$  satisfying  $\|L\|_{(\ell^1)^*} = \|l\| = 1$ .

How is this different from the Hilbert case?

Another reason why  $\ell^1(\mathbb{N})$  is not reflexive is that, while  $\ell^1(\mathbb{N})$  is separable,  $\ell^\infty(\mathbb{N})$  is not.

THEOREM 3.8. [The predual of a separable space is separable] If  $X$  is a Banach space and  $X^*$  is separable, then  $X$  is separable.

PROOF. We consider the case of a real Banach space, the complex case being identical. If  $X^*$  is separable, it has a dense, countable subset  $\{l_n\}_{n=1}^\infty$ . For each  $n$ , let  $x_n$  be such that  $\|x_n\| = 1$  and  $l_n(x_n) \geq \|l_n\|_{X^*}/2$ . Suppose  $Y_0 = \text{span}_{\mathbb{Q}}\{x_n\}_{n=1}^\infty$ , the linear combinations of the  $x_n$ 's with rational coefficients, is not dense in  $X$ . So,  $\bar{Y}_0 \neq X$ , and we can find  $z$  in  $X \setminus \bar{Y}_0$ . By Corollary 3.3, there exists  $l \in X^*$  such that  $l|_{\bar{Y}_0} = 0$ , and  $\|l\|_{X^*} = 1$ . We have that

$$\|l - l_n\|_{X^*} \geq |(l - l_n)(x_n)| = |l_n(x_n)| \geq \|l_n\|_{X^*}/2.$$

If  $\{l_n\}$  were dense in  $X^*$ , we could find a subsequence  $\{l_{n_k}\}$  such that

$$0 = \lim_{k \rightarrow \infty} \|l - l_{n_k}\|_{X^*} \geq \limsup_{k \rightarrow \infty} \|l_{n_k}\|_{X^*}/2,$$

then,

$$1 = \|l\|_{X^*} = \lim_{k \rightarrow \infty} \|l_{n_k}\|_{X^*} = 0,$$

which is a contradiction.  $\square$

EXERCISE 3.10. (i) Show that  $\ell^\infty(\mathbb{N})$  is not separable, while  $\ell^p(\mathbb{N})$  is separable for  $1 \leq p < \infty$ .

- (ii) Show that  $L^\infty[0, 1]$  is not separable, while  $L^p[0, 1]$  is separable for  $1 \leq p < \infty$  (with respect to the Lebesgue measure).
- (iii) Show that  $C[0, 1]$  is separable (with respect to the uniform norm).
- (iv) Show that a Hilbert space  $H$  is separable if and only if it has a countable basis.
- (v) Find Banach spaces  $X, Y$  with  $X \subset Y$ , such that the imbedding map  $x \mapsto x$  is bounded from  $X$  to  $Y$ ,  $Y$  is separable, and  $X$  is not separable.

In the second part of the next exercise, we need a definition. If  $A : X \rightarrow Y$  is a linear, bounded operator between Banach spaces, its *adjoint*  $T' : Y^* \rightarrow X^*$  is defined on a linear functional  $l \in Y^*$  by  $[T'l](x) = l(Tx)$ .

EXERCISE 3.11. Let  $1 \leq p \leq \infty$  and let  $g : \mathbb{R} \rightarrow \mathbb{C}$  be measurable. On  $L^p(\mathbb{R})$ , consider the multiplication operator  $M_g : f \mapsto gf$ .

- (i) Show that

$$\|M_g\|_{\mathcal{L}(L^p)} := \sup_{0 \neq f \in L^p[0,1]} \frac{\|M_g f\|_{L^p}}{\|f\|_{L^p}} = \|g\|_{L^\infty}.$$

- (ii) Let  $1 \leq p < \infty$  and  $g \in L^\infty(\mathbb{R})$ . After identifying  $[L^p(\mathbb{R})]^* = L^{p'}(\mathbb{R})$  ( $1/p + 1/p' = 1$ ), what is the expression for  $[M_g]'$ , the adjoint of  $M_g$ ?

### 3.4. Weak and weak\* topologies, and the Banach-Alaoglu theorem

Our first experience is that we fix a topology on a set  $Y$  (a notion of points "being close"), and this way we can tell which functions  $f : Y \rightarrow \mathbb{R}$  are continuous (their values do not change "abruptly" from point to point). We might instead want to put in the forefront the functions (the "measurables"), rather than the points. We fix a family  $\mathcal{F}$  of functions  $f : Y \rightarrow \mathbb{R}$ , and require that those functions are continuous (that they do not change abruptly: that we consider two points to be close, that is, if they result to be close in all our observations). The *topology*  $\tau(\mathcal{F})$  generated by  $\mathcal{F}$  is the smallest one making all functions in  $\mathcal{F}$  continuous.

It is easy to show that a basis of neighborhoods for  $\mathcal{F}$  is given by finite intersections of basic sets having the form

$$\mathcal{N}(f, a, r) = f^{-1}(a - r, a + r) = \{y \in Y : |f(y) - a| < r\}.$$

One advantage of such *weak topologies* is that they are most economical in terms of open sets, hence they have the largest family of compact sets, which are good when convergence of (sub)sequences is concerned. We will consider below some important instances of this construction.

#### 3.4.1. The weak and the weak\* topologies.

3.4.1.1. *The weak topology.* Let  $X$  be a Banach space, and let  $X^*$  be its dual. The *weak topology*  $w$  on  $X$  is the coarsest (weaker) which makes all functionals  $l \in X^*$  continuous. By general topology, a basis of neighborhoods for the weak topology is given by the family of the subsets  $N(a; l_1, \dots, l_n; \epsilon)$  (with  $a \in X$ ,  $n \geq 1$ ,  $l_1, \dots, l_n \in X^*$ ,  $\epsilon > 0$ ), where:

$$\begin{aligned}
 N(a; l_1, \dots, l_n; \epsilon) &= \{x \in X : |l_1(x) - l_1(a)| < \epsilon, \dots, |l_n(x) - l_n(a)| < \epsilon\} \\
 &= N(a; l_1; \epsilon) \cap \dots \cap N(a; l_n; \epsilon) \\
 &= [N(0; l_1; \epsilon) + a] \cap \dots \cap [N(0; l_n; \epsilon) + a] \\
 &= [\epsilon N(0; l_1; 1) + a] \cap \dots \cap [\epsilon N(0; l_n; 1) + a] \\
 &= \epsilon N(0; l_1, \dots, l_n; 1) + a \\
 (3.4.1) \quad &= N(0; l_1, \dots, l_n; \epsilon) + a
 \end{aligned}$$

are various ways to write and think of these basic neighborhoods. For instance, it is clear that the topology is invariant under translations and dilations,  $x \mapsto \lambda x + a$  is a homeomorphism whenever  $\lambda \neq 0$ .

If  $(X, w)$  is metrizable, and  $\Omega$  is a metric space, then continuity of  $f : X \rightarrow \Omega$  with respect to the weak topology is equivalent to continuity by sequences:  $x_n \rightarrow x$  in  $(X, w)$  implies  $f(x_n) \rightarrow f(x)$  in  $\Omega$ . Unfortunately  $(X, w)$  might not be metrizable, and we have to use *nets* instead. But some nice features remain.

PROPOSITION 3.3.  $(X, w)$  is a Hausdorff space.

PROOF. By translation invariance, we just have to separate  $a \neq 0$  and  $0$ . By corollary 3.2 to the Hahn-Banach theorem, there is  $l \in X^*$  with  $\|l\|_{X^*} = 1$  and such that  $l(a) = \|a\|$ . If  $N(0; l; \|a\|/2) \ni 0$  and  $N(a; l; \|a\|/2) \ni a$  had a point  $x$  in common,

$$\|a\| = |l(a) - l(0)| \leq |l(a) - l(x)| + |l(x) - l(0)| < \|a\|,$$

a contradiction. □

We say that the sequence  $\{x_n\}$  in  $X$  *converges weakly* to  $a \in X$  if  $l(x_n) \rightarrow l(a)$  for all  $l$  in  $X^*$ . Although this notion is generally weaker than "weak net convergence", it is nonetheless useful in many applications, for instance in calculus of variations. We write  $x_n \xrightarrow{w} a$ , or  $w\text{-}\lim_{n \rightarrow \infty} x_n = a$ . By definition, weak convergence of  $x_n$  to  $a$  means that, for each  $l$  in  $X^*$  and each  $\epsilon > 0$ , there exists  $n(l, \epsilon)$  such that, if  $n > n(l, \epsilon)$ , then  $x_n \in N(l; a; \epsilon)$ .

PROPOSITION 3.4. (1) If  $\|x_n - a\| \rightarrow 0$ , then  $x_n \xrightarrow{w} a$ .

(2) If  $x_n \xrightarrow{w} a$ , then  $\|a\| \leq \liminf_{n \rightarrow \infty} \|x_n\|$ .

PROOF. (1) For any  $l$  in  $X^*$ ,  $|l(x_n) - l(a)| \leq \|l\|_{X^*} \|x_n - a\| \rightarrow 0$ .

(2) Again by corollary 3.2, there is  $l$  with  $\|l\|_{X^*} = 1$  and  $\|a\| = l(a)$ , hence:

$$\|a\| = l(a) = \lim_{n \rightarrow \infty} |l(x_n)| \leq \|l\|_{X^*} \liminf_{n \rightarrow \infty} \|x_n\|_X = \liminf_{n \rightarrow \infty} \|x_n\|_X.$$

□

We can not improve the statement in (2). Let  $H$  be an infinite dimensional Hilbert space and let  $\{e_n\}_{n=1}^\infty$  be a orthonormal system in it. Then,

$$(3.4.2) \quad e_n \xrightarrow{w} 0,$$

but  $\|e_n\| = 1$  for all  $n$ .

3.4.1.2. *The weak\* topology.* Let again  $X$  be a Banach space, and let  $X^*$  be its dual. The *weak\* topology*  $w^*$  on  $X^*$  is the coarsest which makes the functionals  $l \mapsto l(x)$  continuous for all  $x$  in  $X$ . Unless the natural identification  $X \hookrightarrow X^{**}$  is surjective, i.e. unless  $X$  is *reflexive*, the *weak\* topology* on  $X^*$  is weaker than the weak topology.

A basis of neighborhoods for the origin in  $w^*$  is given by finite intersections of sets of the form

$$(3.4.3) \quad N(0; x; \epsilon) = \{l \in X^* : |l(x)| < \epsilon\}.$$

The topology is invariant under translations and a basis at any point is easy to write, as in the case of the weak topology. Its basic properties coincide with those of the weak topology.

PROPOSITION 3.5. (1)  $(X^*, w^*)$  is Hausdorff.

(2) If  $\|x_n - a\| \rightarrow 0$ , then  $x_n \xrightarrow{w^*} a$ .

(3) If  $l_n \xrightarrow{w^*} l$ , then  $\|l\|_{X^*} \leq \liminf_{n \rightarrow \infty} \|l_n\|_{X^*}$ .

PROOF. Property (2) is weaker than the analogous statement for the weak topology. Property (1) is proved like the corresponding statement for the weak topology, using the fact that if  $l \neq 0$  is an element of  $X^*$ , there exists  $x$  in  $X$  with  $\|x\|_X = 1$  and  $|l(x)| \geq \|l\|_{X^*}/2$ . If  $\lambda \in N(0; x; \|l\|_{X^*}/4) \cap N(l; x; \|l\|_{X^*}/4)$ , then

$$\|l\|_{X^*}/2 \leq |l(x)| \leq |(l - \lambda)(x)| + |\lambda(x)| < \|l\|_{X^*}/4 + \|l\|_{X^*}/4,$$

a contradiction.

The proof of (3) follows the same lines as that of the weak topology analog. For  $\epsilon > 0$  there is  $x$  in  $X$  with  $\|x\| = 1$  such that  $|l(x)| \geq \|l\|_{X^*} - \epsilon$ . Then,

$$\|l\|_{X^*} - \epsilon \leq |l(x)| = \lim_{n \rightarrow \infty} |l_n(x)| \leq \liminf_{n \rightarrow \infty} \|l_n\|_{X^*}.$$

Let then  $\epsilon \rightarrow 0$ .

□

**3.4.2. Two versions of the Banach-Alaoglu theorem.** The *raison d'être* of the weak\* topology is related to compactness properties. To keep the exposition self-contained, below you find a proof of the Tychonoff theorem.

3.4.2.1. *Tychonoff's theorem.* If  $\{X_i : i \in I\}$  is a family of sets,  $X_I = \prod_{i \in I} X_i$ , their Cartesian product, is best interpreted as the set of the functions

$$p : I \rightarrow \cup_{i \in I} X_i, \text{ with } p(i) \in X_i \text{ for each } i \in I.$$

We call  $I$  the *domain* of  $p$ .

If each  $X_i$  is a topological space, the *product topology* on  $\prod_{i \in I} X_i$  is the weakest (coarsest, minimal) making all projections  $\pi_i : X_I \rightarrow X_i$ ,  $\pi_i(p) = p(i)$ , continuous. A basis of open sets for it is provided by the sets  $\pi_{i_1}^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_n}^{-1}(U_{i_n})$ , where  $n \geq 1$ ,  $i_1, \dots, i_n \in I$ , and  $U_{i_i}$  is open in  $X_{i_i}$ .

**THEOREM 3.9.** *If each  $X_i$  is compact,  $i \in I$ , then  $\prod_{i \in I} X_i$  is compact.*

We give the first of the three proofs surveyed in [Three Proofs of Tychonoff's Theorem](#) by É. Matheron (2020), which the author labels the *Wisconsin proof*.

Some of the usual notions associated to functions carry over and are useful. If  $J \subset H$ ,  $\Pi_{H,J}p = p|_J$  is the restriction of  $p \in X_H$  to  $X_J$ ,  $\pi_{H,J} : X_H \rightarrow X_J$ ,  $\pi_{H,J} \circ \pi_{J,K} = \pi_{H,K}$ . We set  $\pi_J = \pi_{I,J}$ . If  $p \in X_H$  and  $q \in X_J$  with  $H$  and  $J$  disjoint, then  $p \vee q \in X_{H \cup J}$  is the function such that  $(p \vee q)|_H = p$ ,  $(p \vee q)|_J = q$ . The set  $X_\emptyset$  reduces to the unique *empty function* from  $\emptyset$  to itself, which we denote by  $\emptyset$  (in agreement with the interpretation of functions as particular relations).

We fix a set  $I$ , and let  $\mathbb{P} = \cup_{J \subseteq I} X_J$ , which is partially ordered by the relation  $p \leq q$  if  $q|_J = p$ , where  $J$  is the domain of  $p$ .

**PROOF.** The proof is by contradiction. We suppose that there is a family  $\mathcal{U}$  of open sets in  $X_I$  such that no finite subfamily covers  $X_I$ , and we shall show that  $\mathcal{U}$  itself does not cover  $X_I$  by exhibiting an element  $p \in X_I \setminus \cup_{U \in \mathcal{U}} U$ .

We consider the set  $B$  of the *bad* elements in  $\mathbb{P}$ : those  $p \in X_J$  such that for all open  $V \ni p$  open in  $X_J$ ,  $\pi_J^{-1}(V)$  can not be covered by finitely many sets in  $\mathcal{U}$ . The empty function lies in  $B$ , which is then nonempty. The only open set in  $X_\emptyset$  containing  $\emptyset$ , in fact, is  $X_\emptyset = \{\emptyset\}$  itself, and  $\pi_\emptyset^{-1}(X_\emptyset) = X_I$ , which can not be covered by a finite subfamily of  $\mathcal{U}$  by assumption. We will prove three facts about  $B$ .

(i) *B is downward close:* if  $p \leq q \in B$ , then  $p \in B$ . In fact, if  $H$  is the domain of  $p$  and  $J$  that of  $q$ , and  $V \ni p$  is open in  $X_H$ , then  $\pi_H^{-1}(V) = \pi_J^{-1}(\pi_{H,J}^{-1}(V))$  which can not be covered by a finite subfamily of  $\mathcal{U}$  because  $\pi_{H,J}^{-1}(V) \ni q$  is open and  $q$  is bad. We can assume  $V_a = O_a \times W_a$

(ii) If  $p \in B$  has domain  $J \subset I$  and  $i_0 \in I \setminus J$ , then there is  $a \in X_{i_0}$  such that  $p \vee a \in B$ . Suppose by contradiction that for all  $a$  in  $X_{i_0}$  we have  $p \vee a \notin B$ ; i.e. there exists  $V_a \ni p \vee a$  open in  $X_{J \cup \{i_0\}}$  such that  $\pi_{J \cup \{i_0\}}^{-1}(V_a)$  can be covered by a finite subfamily of  $\mathcal{U}$ . We can assume that  $V_a = O_a \times W_a$  with  $O_a$  open in  $X_J$  and  $W_a \ni a$  open in  $X_{i_0}$ . Since  $X_{i_0}$  is compact, it can be covered by finitely many  $W_{a_1}, \dots, W_{a_n}$ . Let  $O = O_{a_1} \cap \dots \cap O_{a_n} \ni p$ , so that  $O \times W_{a_j} \ni p \vee a_j$  and  $\pi_{J \cup \{i_0\}}^{-1}(O \times W_{a_j})$  can be covered by a finite subfamily of  $\mathcal{U}$ . Now,

$$\pi_J^{-1}(O) = \bigcup_{j=1}^n \pi_{J \cup \{i_0\}}^{-1}(O \times W_{a_j}),$$

and each set in the union can be covered by a finite subfamily of  $\mathcal{U}$ . Hence,  $p \notin B$ .

(iii)  $(B, \leq)$  has a maximal element. If any chain  $C$  in  $B$  has an upper bound in  $B$ , by Zorn's lemma  $B$  has a maximal element. Let  $C$  be a chain in  $B$ , let  $H$  be the union of the domains of all  $q$ 's in  $C$ , and define  $p \in X_H$  to be such that  $p|_J = q$  if  $q \in C$  has domain  $J$ . The function  $p$  is well defined because  $C$  is totally ordered. Thus  $p$  is an upper bound of  $C$  in  $\mathbb{P}$ . It suffices to show that  $p \in B$ . Let  $V \ni p$  be an open subset of  $X_H$ . By definition of the product topology, we can choose  $V \supseteq \pi_{H,F}^{-1}(W) \ni p$  where  $F \subset H$  is finite and  $W \subset X_F$  is open with  $W \ni p|_F$ . Now,  $C$  is a chain and each  $i_l \in F$  belongs to some domain  $J_l$  in the chain, so  $F \subseteq J_1 \cup \dots \cup J_m = J_0 \subseteq H$ , where  $J_0$  is a domain in the chain. Let  $q_0 \in X_{J_0} \cap C$ . Since  $q_0$  is bad,  $p|_F = q_0|_F$  is bad as well by (i), thus  $\pi_F^{-1}(W)$  can not be covered by a finite subfamily of  $\mathcal{U}$ . A fortiori,  $\pi_H^{-1}(V) \supseteq \pi_F^{-1}(W)$  can not be covered by a finite subfamily of  $\mathcal{U}$ , hence  $p$  is bad.

Summarizing,  $B$  has a maximal element  $p$  by (iii), having domain  $I$  by (ii). In particular  $X_I \ni p$  can not be covered by a finite subfamily of  $\mathcal{U}$ , but this is possible only if  $p$  does not belong to any subset of  $\mathcal{U}$ , which henceforth does not cover the whole of  $X_I$ .  $\square$

#### 3.4.2.2. Banach-Alaoglu theorem: the topological form.

**THEOREM 3.10** (Banach-Alaoglu). *Let  $\overline{B}_1$  be the closed unit ball of  $X^*$ . Then,  $\overline{B}_1$  is compact in the weak\* topology.*

**PROOF.** For each  $x \in X$ , consider  $\overline{D(0, \|x\|)}$ , the closed unit disc in the complex plane, and let

$$(3.4.4) \quad K := \prod_{x \in X} \overline{D(0, \|x\|)},$$

endowed with the product topology. Such is the topology having as basis at  $\{f(x) : x \in X\}$  with  $|f(x)| \leq \|x\|$ , finite intersections of sets of the form

$$M(f; a; \epsilon) = \{g : X \rightarrow \mathbb{C} \text{ s.t. } |g(x)| \leq \|x\| \text{ and } |g(a) - f(a)| < \epsilon\}.$$



Since each factor  $\overline{D(0, \|x\|)}$  is compact, by Tychonoff theorem  $K$  is compact, too.

We imbed  $\overline{B}_1$  into  $K$ ,  $\Phi : \overline{B}_1 \rightarrow K$ ,

$$\Phi(l) = \{\{l(x)\}_{x \in X}\}.$$

Unraveling definitions, we have that

$$\Phi(N(l; a; \epsilon)) = M(l; a; \epsilon) \cap \Phi(\overline{B}_1),$$

i.e.  $\Phi$  is a homeomorphism of  $\overline{B}_1$  onto its image into  $K$ . Since the latter is compact, if we show that  $\Phi(\overline{B}_1)$  is closed in  $K$ , then it is compact, hence  $\overline{B}_1$  is compact as well.

Suppose  $f$  lies in the closure of  $\Phi(\overline{B}_1)$ , and consider  $x, y \in X$ . For any  $\epsilon > 0$  there is  $l \in \overline{B}_1$  such that  $|l(x) - f(x)| < \epsilon/3$ ,  $|l(y) - f(y)| < \epsilon/3$ , and  $|l(x+y) - f(x+y)| < \epsilon/3$ , so that

$$|f(x) + f(y) - f(x+y)| < \epsilon.$$

Hence,  $f(x) + f(y) = f(x+y)$ . The same way, one shows that  $f(\lambda x) = \lambda f(x)$  if  $\lambda$  is a complex number. This shows that  $f : X \rightarrow \mathbb{C}$  is a linear functional, which also lies in  $\overline{B}_1$ , since  $|f(x)| \leq \|x\|$ . We have proved that  $\Phi(\overline{B}_1)$  is closed, hence  $K$  is compact.  $\square$

**3.4.2.3. Banach-Alaoglu theorem: the sequential form.** A practical problem with this version of the Banach-Alaoglu theorem is that in general compactness is not in general equivalent to sequential compactness, which is very useful in many applications (measure theory and probability, calculus of variations, PDEs...). The version involving sequential compactness below could be deduced by the previous one, but we provide a direct proof which is similar to the original one for the Ascoli-Arzelà theorem, which is essentially constructive and does not require the Tychonoff theorem.

**THEOREM 3.11** (Banach-Alaoglu, separable pre-dual). *Let  $X$  be a separable Banach space, and let  $\{l_n\}$  be a sequence in  $\overline{B}_1$ , the unit ball in  $X^*$ . then, there exists a subsequence  $\{l_{n_k}\}$  and an element  $l$  in  $\overline{B}_1$  such that  $l_{n_k} \xrightarrow{w^*} l$ .*

**PROOF.** Let  $\{z_j\}_{j=1}^\infty$  be a dense sequence in  $X$ , let  $X_0$  be the dense space it generates in  $X$ , and let  $\{x_m\}_{m=1}^\infty$  be a maximal family of linearly independent vectors in  $\{z_j\}_{j=1}^\infty$ , so that  $X_0 = \text{span}(\{x_m\}_{m=1}^\infty)$ . We will first proceed to define  $l$  on  $X_0$ .

Consider a subsequence  $n^{(1)} = \{n_k^{(1)}\}_{k=1}^\infty$  such that

$$\lim_{k \rightarrow \infty} l_{n_k^{(1)}}(x_1) = a_1$$

exists in  $\mathbb{C}$ . Such sequence exists because  $|l_n(x_1)| \leq \|l_n\|_{X^*} \|x_1\|_X \leq \|x_1\|_X$  is bounded in  $\mathbb{C}$ . Suppose that subsequences  $\{n\}_{n=1}^\infty \supset \{n_k^{(1)}\}_{k=1}^\infty \supset \cdots \supset \{n_k^{(m)}\}_{k=1}^\infty$  have been chosen in such a way (i)  $\lim_{k \rightarrow \infty} l_{n_k^{(j)}}(x_j) = a_j \in \mathbb{C}$  when  $1 \leq j \leq m$ , and (ii)  $n_1^{(1)} < n_1^{(2)} < \cdots < n_1^{(m)}$ . As above, we can choose a subsequence  $\{n_k^{(m+1)}\}_{k=1}^\infty$  of  $\{n_k^{(m)}\}_{k=1}^\infty$  such that  $\lim_{k \rightarrow \infty} l_{n_k^{(m+1)}}(x_{m+1}) = a_{m+1} \in \mathbb{C}$ , and  $n_1^{(m+1)} > n_1^{(m)}$ . The sequence  $\{x_1^{(m)}\}_{m=1}^\infty$  is a subsequence of all subsequences  $\{n_k^{(m)}\}_{k=1}^\infty$ , hence,

$$\lim_{m \rightarrow \infty} l_{n_1^{(m)}}(x_j) = a_j$$

for all  $x_j$ 's. Define  $l(x_j) = a_j$ .

The linear extension of  $l$  to  $X_0$  is forced by its values on the basis elements,

$$l\left(\sum_{j=1}^n c_j x_j\right) := \sum_{j=1}^n c_j l(x_j) = \sum_{j=1}^n c_j \lim_{k \rightarrow \infty} l_{n_k}(x_j) = \lim_{k \rightarrow \infty} l_{n_k}\left(\sum_{j=1}^n c_j x_j\right).$$

As a consequence,

$$(3.4.5) \quad \left| l\left(\sum_{j=1}^n c_j x_j\right) \right| = \lim_{k \rightarrow \infty} \left| l_{n_k}\left(\sum_{j=1}^n c_j x_j\right) \right|$$

$$(3.4.6) \quad \leq \liminf_{k \rightarrow \infty} \|l_{n_k}\|_{X^*} \left\| \sum_{j=1}^n c_j x_j \right\|_X$$

$$(3.4.7) \quad \leq \left\| \sum_{j=1}^n c_j x_j \right\|_X.$$

The unique continuous extension of  $l$  to  $X = \overline{X_0}$  (see Exercise 3.5) satisfies  $\|l\|_{X^*} \leq 1$ .

We have to verify that  $\lim_{k \rightarrow \infty} l_{n_k}(x) = l(x)$  for all  $x$  in  $X$ . This is a simple  $3\epsilon$  argument. For  $x \in X$  and  $x_0 \in X_0$ ,

$$\begin{aligned} |l(x) - l_{n_k}(x)| &\leq |l(x) - l(x_0)| + |l(x_0) - l_{n_k}(x_0)| + |l_{n_k}(x) - l(x)| \\ &\leq 2\|x - x_0\|_X + |l(x_0) - l_{n_k}(x_0)|. \end{aligned}$$

For given  $\epsilon > 0$ , choose  $x_0$  with  $\|x - x_0\|_X \leq \epsilon$ , then  $k(\epsilon) > 0$  such that for  $k > k(\epsilon)$  one has  $|l(x_0) - l_{n_k}(x_0)| < \epsilon$ .  $\square$

### 3.5. Baire's Theorem and the uniform boundedness principle

Let  $(X, d)$  be a metric space. A subset  $A$  of  $X$  is *nowhere dense* in  $X$  if  $\overline{A}$  has empty interior (it does not contain nonempty open subsets).

**THEOREM 3.12.** [*Baire's Theorem*] If  $(X, d)$  is a complete metric space, and  $A_n : n \geq 1$  is a countable union of nowhere dense sets, then  $\cup_{n=1}^{\infty} A_n \neq X$ .

**PROOF.** Since  $\overline{A_1}$  does not contain a nonempty open set, we can find  $\overline{B(x_1, r_1)} \subset X \setminus \overline{A_1}$  with  $r_1 < 1/2$ . Since  $\overline{A_2}$  does not contain a nonempty open set,  $B(x_1, r_1) \setminus \overline{A_2}$  is a nonempty open set, hence, it contains  $\overline{B(x_2, r_2)}$  with  $r_2 < 1/2^2$ . By iteration, we find  $\overline{B(x_n, r_n)} \supseteq \overline{B(x_{n-1}, r_{n-1})}$  with  $r_n < 1/2^n$ , and  $\overline{B(x_n, r_n)} \cap (\overline{A_1} \cup \dots \cup \overline{A_n}) = \emptyset$ .

Since  $X$  is complete, the intersection of the balls  $\overline{B(x_n, r_n)}$  contains (a unique) point  $z \in X \setminus (\cup_{n=1}^{\infty} A_n)$ .  $\square$

The next, important theorem is due to Banach and Steinhaus.

**THEOREM 3.13.** [*Uniform Boundedness Principle*] Let  $\mathcal{F} = \{T\}$  be a family of bounded, linear operators  $T : X \rightarrow Y$  from a Banach space  $X$  to a normed, linear space  $Y$ . Suppose that, for each  $x$  in  $X$ ,  $\|Tx\| \leq C(x)$  independent of  $T \in \mathcal{F}$ , although possibly dependent on  $x$ . Then,  $\mathcal{F}$  is bounded,

$$\|T\| \leq C$$

for some  $C > 0$  independent of  $T \in \mathcal{F}$ .

**PROOF.** For  $m \geq 1$ , let  $A_m = \{x \in X : \sup\{\|Tx\| : T \in \mathcal{F}\} \leq m\}$ . By hypothesis,  $\cup_{m=1}^{\infty} A_m = X$ , hence, some  $A_m$  contains a closed open ball  $\overline{B(z, r)}$ . Since  $z \in A_m$ ,

$$\overline{B(0, r)} \subseteq A_m - z \subseteq A_m + A_m \subseteq A_{2m}.$$

This is what we need, since, for  $\|x\| \leq 1$  and  $T \in \mathcal{F}$ ,

$$\|Tx\| = \frac{1}{r} \|Trx\| \leq \frac{2m}{r}.$$

$\square$

Here is an application of Banach-Steinhaus.

**COROLLARY 3.4.** Let  $T_n \in \mathcal{L}(X, Y)$  be a sequence of bounded, linear maps from  $X$  to  $Y$ , Banach spaces. If for all  $x$  in  $X$  there exists

$$Tx := \lim_{n \rightarrow \infty} T_n x,$$

then  $T \in \mathcal{L}(X, Y)$  is bounded.

PROOF. By hypothesis,  $\{\|T_n x\|, n \geq 1\}$  is bounded for each  $x$ , hence  $\{\|T_n\|, n \geq 1\}$  is bounded by some finite  $C > 0$ . We have, then, if  $\|x\| \leq 1$ ,

$$\|Tx\| = \lim_{n \rightarrow \infty} \|T_n x\| \leq \limsup_{n \rightarrow \infty} \|T_n\| \leq C.$$

Thus,  $\|T\| \leq C$ . □

Banach-Steinhaus allows us to transform (without control of the constants) weak *quantitative* information, into strong quantitative information. The following exercise provides an example.

EXERCISE 3.12. Let  $f : [a, b] \rightarrow X$  a function with values in a Banach space  $X$ , and suppose  $f$  is **weakly Lipschitz**,

$$|l(f(t+h)) - l(f(t))| \leq C(l)|h|$$

for all  $t, t+h \in [a, b]$  and  $l \in X^*$ . Show that  $f$  is Lipschitz: there is a constant  $C > 0$  such that

$$\|f(t+h) - f(t)\| \leq C|h|,$$

We will see another important example in the next subsection, where we treat Banach space valued holomorphic functions.

*Qualitative* information, however, behaves differently. For instance, it is possible to exhibit functions  $f : [0, 1] \rightarrow \ell^2$  which are *weakly continuous* (for all  $h$  in  $\ell^2$ ,  $x \mapsto \langle h, f(x) \rangle_{\ell^2}$  is continuous), but which are not continuous. The following example is modeled on the weak, but not strongly convergent sequence in (3.4.2). Let

$$(3.5.1) \quad [f(x)](n) = \begin{cases} \frac{1}{1+|n-1/x|^2} & \text{if } 0 < x \leq 1, \\ 0 & \text{if } x = 0. \end{cases}$$

Since for  $x \neq 0$  we have  $\|f(x)\|_{\ell^2}^2 \geq \sum_{n=0}^{\infty} \frac{1}{(1+n^2)^2}$ ,  $f$  is not continuous at  $x = 0$ . On the other hand,

$$(3.5.2) \quad \lim_{x \rightarrow 0} \langle h, f(x) \rangle_{\ell^2} = 0.$$

After changing variables to  $y = 1/x \rightarrow \infty$ , and considering WLOG  $h \geq 0$ , we can estimate (using an elementary estimate and Cauchy-Schwarz)

$$\begin{aligned}
\langle h, f(1/y) \rangle_{\ell^2} &= \sum_{0 \leq n \leq y/2} \frac{h(n)}{1 + (n - y)^2} + \sum_{n > y/2} \frac{h(n)}{1 + (n - y)^2} \\
&\leq \frac{C}{y} \sum_{0 \leq n \leq y/2} \frac{h(n)}{(1 + (n - y)^2)^{1/2}} \\
&\quad + 2 \left( \sum_{n > y/2} h(n)^2 \right)^{1/2} \left( \sum_{n \geq 0} \frac{1}{(1 + n^2)^2} \right)^{1/2} \\
&\leq \frac{C}{y} \|h\|_{\ell^2} \left( \sum_{n \geq 0} \frac{1}{(1 + n^2)^2} \right)^{1/2} \\
&\quad + 2 \left( \sum_{n > y/2} h(n)^2 \right)^{1/2} \left( \sum_{n \geq 0} \frac{1}{(1 + n^2)^2} \right)^{1/2}.
\end{aligned}$$

The last expression tends to 0 as  $y \rightarrow \infty$  by monotone convergence. It is easy to see that  $f$  is continuous on  $(0, 1]$  (for instance, by dominated convergence).

**3.5.1. Banach space-valued holomorphic functions.** Complex valued power series of a complex variable define holomorphic functions, which are at the heart of an elegant and powerful theory. At the root of it, is the fact that holomorphic functions can be defined in several, equivalent ways: through power series, through complex integrals, via the Cauchy-Riemann equations, and others. These different viewpoints can be translated to the world of functions with values in Banach spaces  $X$ ; or even better in Banach spaces of the form  $\mathcal{L}(X)$ , where a product is part of the structure; or in Banach algebras, which is the most general Banach structure with products. This way, a number of results, tools, and techniques from holomorphic function theory become available to Functional Analysis, a fact of the greatest importance.

A function  $f : \mathbb{C} \supseteq \Omega \rightarrow X$  defined from a region of the complex plane with values in a Banach space is *holomorphic* if

$$f'(z) := \lim_{h \rightarrow 0 \text{ in } \mathbb{C}} \frac{f(z+h) - f(z)}{h} \in X$$

exists for all  $z$  in  $\Omega$ . The function  $f$  is *weakly holomorphic* if for all  $l$  in  $X^*$  the function

$$z \mapsto l(f(z)) \in \mathbb{C}$$

is holomorphic in the usual sense.

EXERCISE 3.13. Show that a holomorphic function  $f : \Omega \rightarrow X$  is weakly holomorphic.

THEOREM 3.14. Let  $f : \Omega \rightarrow X$  be a map defined from an open subset of  $\mathbb{C}$  with values in a Banach space  $X$ . The following are equivalent.

- (i) The function  $f$  is holomorphic in  $\Omega$ .
- (ii) The function  $f$  is weakly holomorphic in  $\Omega$ .
- (iii) For each  $z_0$  in  $\Omega$ , there is  $r > 0$  such that, for  $|z - z_0| < r$ ,

$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n,$$

where  $a_n \in X$  and the series converges absolutely uniformly for  $|z - z_0| \leq \rho < r$ .

The value of  $r$  can be taken to be that of the larger disc centered at  $z_0$  and contained in  $\Omega$ .

In fact, the series converges to an  $X$ -valued holomorphic function in  $B(z_0, R)$  where

$$(3.5.3) \quad R = \frac{1}{\limsup_{n \rightarrow \infty} \|a_n\|^{1/n}}$$

is the radius of the largest disc centered at  $z_0$  and contained in  $\Omega$ .

The computational part of the proof is contained in the following.

LEMMA 3.1. Consider the power series

$$(3.5.4) \quad g(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$$

with coefficients  $a_n \in X$ , a Banach space, and with radius of convergence as in (3.5.3).

Then,  $g : B(0, R) \rightarrow X$  is strongly holomorphic, and

$$(3.5.5) \quad g'(z) = \sum_{n=1}^{\infty} n a_n(z - z_0)^{n-1},$$

which has the same radius of convergence as  $g$ .

In particular,  $g$  is infinitely differentiable, and

$$(3.5.6) \quad g^{(n)}(z_0) = n! a_n.$$

PROOF OF THE LEMMA. The usual proof from holomorphic theory makes use of tools we do not have, and we do not want to develop. We provide instead an XVIII century style proof which does not require them. We can suppose  $z_0 = 0$ . We start with the estimate

$$(3.5.7) \quad \left| \frac{(z+h)^n - z^n}{h} - nz^{n-1} \right| \leq |h| \frac{n(n-1)}{2} (|z| + |h|)^{n-2},$$

which holds for  $n \geq 1$  and  $z, h \in \mathbb{C}$ . The proof is just a calculation,

$$\begin{aligned} \left| \frac{(z+h)^n - z^n}{h} - nz^{n-1} \right| &= \left| \sum_{j=2}^n \binom{n}{j} z^{n-j} h^{j-1} \right| \\ &\leq |h| \sum_{l=0}^{n-2} \binom{n}{l+2} |z|^{n-2-l} |h|^l \\ &= |h| \sum_{l=0}^{n-2} \frac{n(n-1)}{(l+2)(l+1)} \binom{n-2}{l} |z|^{n-2-l} |h|^l \\ &\leq |h| \frac{n(n-1)}{2} \sum_{l=0}^{n-2} \binom{n-2}{l} |z|^{n-2-l} |h|^l \\ &= |h| \frac{n(n-1)}{2} (|z| + |h|)^{n-2}. \end{aligned}$$

Let now

$$\psi(z) = \sum_{n=1}^{\infty} n a_n z^{n-1},$$

which has the same radius of convergence. For  $|z| < R$  and  $|h| < R - |z|$ , we have:

$$\begin{aligned} \left\| \frac{g(z+h) - g(z)}{h} - \psi(z) \right\| &= \left\| \sum_{n=1}^{\infty} a_n \left( \frac{(z+h)^n - z^n}{h} - nz^{n-1} \right) \right\| \\ &\leq |h| \sum_{n=2}^{\infty} \frac{n(n-1)}{2} \|a_n\| (|z| + |h|)^{n-2} \\ &\rightarrow 0 \end{aligned}$$

as  $h \rightarrow 0$  in  $\mathbb{C}$ , since the series in the line before the last converges. Hence,  $\psi = g'$ . Since  $g'$  has the same radius of convergence as  $g$ , we can iterate the calculation,

$$(3.5.8) \quad g^{(m)}(z) = \sum_{n=m}^{\infty} \frac{n!}{(n-m)!} a_n (z - z_0)^{n-m}, \quad g^{(m)}(z_0) = m! a_m.$$

□

PROOF OF THE THEOREM. (i) implies (ii) by exercise 3.13, and (iii) implies (i) is the lemma above.

We show that (ii) implies (iii). We fix some notation. If  $x \in X$ ,  $\widehat{x} \in X^{**}$  is the functional on  $X^*$  for which  $\widehat{x}(l) = l(x)$ . We set  $\widehat{X} \subseteq X^{**}$  the set of such functionals.

We start by showing that if  $f : \Omega \rightarrow X$  a weakly holomorphic function, and  $\gamma$  is a closed curve in  $\Omega$ , then  $z \mapsto f(z)$  is bounded on  $\gamma$ . In fact, and for  $z$  on  $\gamma$ :

$$\begin{aligned} |\widehat{f(z)}(l)| &= |l(f(z))| \\ &\leq \sup_{z \in \gamma} |l(f(z))| \\ &= C(l), \end{aligned}$$

which is finite because  $z \mapsto l(f(z))$  is continuous on  $\gamma$ . By Banach-Steinhaus theorem,

$$\sup_{z \in \gamma} \|f(z)\|_X = \sup_{z \in \gamma} \|\widehat{f(z)}\|_{X^{**}} < \infty.$$

Let  $B(z_0, r)$  be a disc contained in  $\Omega$ , and let  $\gamma$  be a circle of radius  $\rho < r$  centered at  $z_0$  and contained in  $\Omega$ . Let again  $l$  be in  $X^*$ . Then,  $z \mapsto l(f(z))$  can be expanded as a power series,

$$l(f(z)) = \sum_{n=0}^{\infty} a_n(l)(z - z_0)^n,$$

where

$$a_n(l) = \frac{1}{2\pi i} \int_{\gamma} \frac{l(f(z))}{(z - z_0)^{n+1}} dz.$$

The functional  $l \mapsto a_n(l)$  is linear and

$$|a_n(l)| \leq \sup_{z \in \gamma} \|f(z)\|_X \frac{\|l\|_{X^*}}{\rho^n}.$$

That is,  $a_n \in X^{**}$  with

$$(3.5.9) \quad \|a_n\|_{X^{**}} \leq \frac{\sup_{z \in \gamma} \|f(z)\|_X}{\rho^n}.$$

Thus,

$$(3.5.10) \quad g(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$$



converges for  $|z - z_0| < \rho$  to a function  $g$  with values in  $X^{**}$ . You just have to check that the usual proof in holomorphic function theory works when you have power series with coefficients in a Banach space (geometric series only are involved). The Hadamard formula for the radius of convergence of the series is proved like in holomorphic theory. By the lemma,  $g$  is an  $X^{**}$  valued holomorphic function. We next prove that  $g(z) = \widehat{f(z)}$  is the image in  $X^{**}$  of  $\widehat{f(z)}$ , i.e. that  $[g(z)](l) = l(f(z))$  for all  $l \in X^*$ .

In fact, all  $l$  in  $X^*$  we have

$$[g(z)](l) = \sum_{n=0}^{\infty} a_n(l)(z - z_0)^n = l(f(z)).$$

By the lemma,  $g$  is infinitely differentiable and

$$a_n = \frac{g^{(n)}(z_0)}{n!} = \frac{\widehat{f^{(n)}(z_0)}}{n!} \in X,$$

because  $X$  is closed in  $X^{**}$ , and the derivatives, which are limits of  $X$  valued functions, belong to  $X$ . Thus,  $a_n = \widehat{\alpha_n}$ , with  $\alpha_n \in X$ , and

$$f(z) = \sum_{n=0}^{\infty} \alpha_n (z - z_0)^n,$$

as wished. □

**EXERCISE 3.14.** Let  $\Omega \subseteq \mathbb{C}$  be open, and let  $\lambda : \Omega \rightarrow \mathbb{C}$  be holomorphic, and  $f : \Omega \rightarrow X$  be a Banach space-valued holomorphic function. Show that their product  $\lambda \cdot f : \Omega \rightarrow X$  is holomorphic.

### 3.6. The Open Mapping Theorem and the Closed Graph Theorem

A map  $F : M \rightarrow N$  between metric spaces is *open* if the image of an open set in  $M$  is open in  $N$ .

**EXERCISE 3.15.** If  $T : X \rightarrow Y$  is a linear map between normed, linear spaces, then the following are equivalent:

- (i)  $T$  is open;
- (ii) there is a ball  $B_X(0, r)$  in  $X$  such that  $T(B_X(0, r))$  contains a ball in  $Y$ ;
- (iii) there is a ball  $B_X(0, r)$  in  $X$  such that  $T(B_X(0, r))$  contains a ball centered at 0 in  $Y$ .

Moreover, if  $T$  is open, then it is onto.

**Hint.** (i)  $\implies$  (ii) is clear. For (ii)  $\implies$  (iii), you can show that if  $T(B(0, r)) \supset B(y, R)$ , then  $T(B(0, 2r)) \supset B(0, R)$ . The proof that (iii)  $\implies$  (i) is easy. Also, (iii) implies that  $T$  is onto by the homogeneity of  $T$ .

**THEOREM 3.15.** [*Open Mapping Theorem*] Let  $X, Y$  be Banach spaces, and let  $T : X \rightarrow Y$  be a linear, bounded map from  $X$  onto  $Y$ . Then,  $T$  is open.

**PROOF.** Let  $B_n := B(0, n) \subset X$ . Since  $\cup_n T(B_n) = Y$ , there is  $n$  such that  $\overline{T(B_n)} \supset B(y_0, \epsilon)$ . If  $y \in Y$ , then  $y = (y + y_0) - y_0 \in \overline{T(B_n)} - \overline{T(B_n)} \subset \overline{T(B_{2n})}$  provided  $\|y\| \leq \epsilon$ . i.e.  $B(0, \epsilon/(2n)) = B(0, \eta) \subseteq \overline{T(B_1)}$ .

It suffices then to show that  $\overline{T(B_1)} \subset T(B_2)$ .

Let  $y \in \overline{T(B_1)}$ , and pick  $x_1 \in B_1$  such that  $\|y - Tx_1\| < \eta/2$ , so that  $Tx_1 - y \in B(0, \eta/2) \subseteq \overline{T(B_{1/2})}$ . We can then find pick  $x_2 \in B_{1/2}$  such that  $\|y - Tx_1 - Tx_2\| < \eta/2^2$ .

Iterating, we have  $x_n \in B_{1/2^{n-1}}$  such that  $\|y - T(x_1 + \dots + x_n)\| < \eta/2^n$ . We have that  $x_1 + \dots + x_n \rightarrow x \in B_2$  as  $n \rightarrow \infty$ , and  $T(x) = \lim_{n \rightarrow \infty} T(x_1 + \dots + x_n) = y$ , as wished.  $\square$

**EXERCISE 3.16.** Let  $X = \mathbb{R}^2$  with the Euclidean metric. Show that for all  $\epsilon > 0$  there is a linear bijection  $T : X \rightarrow X$  such that  $\|T\| = 1$ , yet  $T(B(0, 1))$  does not contain  $B(0, \epsilon)$ . That is, the Open Mapping Theorem is not quantitative.

**THEOREM 3.16.** [*Inverse Mapping Theorem*] Let  $T : X \rightarrow Y$  be a bounded bijection of Banach spaces. Then,  $T^{-1} : Y \rightarrow X$  is bounded.

**PROOF.** The fact that  $T$  is open means that for some  $\epsilon > 0$ , if  $\|y\| < \epsilon$  in  $Y$ , then there is  $x \in X$  with  $\|x\| < 1$  and  $Tx = y$ , i.e.  $x = T^{-1}y$ . Said it differently,  $T^{-1}$  maps  $B(o, \epsilon)$  into  $B(0, 1)$ , so that  $\|T^{-1}\| \leq 1/\epsilon$ .  $\square$

**EXERCISE 3.17.** Show that if  $\|\cdot\|_1$  and  $\|\cdot\|_2$  are two norms on  $X$ , if  $\|x\|_2 \leq \|x\|_1$  on  $X$  and  $X$  is Banach with respect to  $\|\cdot\|_1$ , then there is  $C > 0$  such that  $\|x\|_1 \leq C\|x\|_2$  on  $X$ .

**EXERCISE 3.18.** Show that (even!) in  $\mathbb{R}^2$  one can find norms  $\|\cdot\|_1$  and  $\|\cdot\|_2$  with  $\|x\|_2 \leq \|x\|_1$ , yet  $\|x\|_1 \leq C\|x\|_2$  only holds if  $C$  is (arbitrarily) large. That is, the result in the previous exercise is not quantitative.

The *graph* of a function  $f : M \rightarrow N$  is the set  $\Gamma(f) = \{(x, y) \in M \times N : y = f(x)\}$ . If  $T : X \rightarrow Y$  is a linear operator between linear spaces, then  $\Gamma(T)$  is a linear subspace of  $X \times Y$ . If  $X$  and  $Y$  are normed, then  $\|(x, y)\| := \|x\| + \|y\|$  defines a norm on  $X \times Y$ , hence on  $\Gamma(T)$ .

**EXERCISE 3.19.** If  $X$  and  $Y$  are normed linear space and  $T \in \mathcal{L}(X, Y)$ , then  $\Gamma(T)$  is closed.

**THEOREM 3.17.** [*Closed Graph Theorem*] Let  $T : X \rightarrow Y$  be a linear operator defined from a Banach space  $X$  to a Banach space  $Y$ . If  $\Gamma(T)$  is closed, then  $T$  is bounded.

PROOF. By assumption,  $\Gamma(T)$  is closed in a Banach space, hence Banach itself: if  $(x_n, Tx_n) \rightarrow (x, y)$  is Cauchy, then  $y = Tx$ , so  $(x, y) \in \Gamma(T)$ .

Consider the projections  $\pi_X : (x, Tx) \mapsto x$  and  $\pi_Y : (x, Tx) \mapsto Tx$ . Both projections are bounded,  $\pi_X$  is invertible (hence, its inverse is continuous), and  $T = \pi_Y \circ \pi_X^{-1}$ . Hence,  $T$  is bounded.  $\square$

### 3.7. Integrals of continuous, Banach space valued functions

The integral of functions with values on a Banach space can be defined in several ways, choosing which depends on the application we have in mind<sup>2</sup>. Think of Banach space valued holomorphic functions: extending to them the notion of Cauchy integrals requires integrating vector valued functions. Having in mind holomorphic theory, we sketch here a construction of the integral for a continuous function  $f : [a, b] \rightarrow X$ , where  $X$  is Banach. The definition of integral we discuss here does not require any foundational result concerning Banach spaces. We only make use of linearity and completeness.

Let  $f : [a, b] \rightarrow X$  be a continuous function with values in the Banach space  $X$ . Since  $[a, b]$  is compact,  $f$  is uniformly continuous. For fixed  $n \geq 1$ , let

$$(3.7.1) \quad S_n(f) = \sum_{j=1}^{2^n} f\left(\frac{j}{2^n}\right) \frac{b-a}{2^n} \in X.$$

The sequence  $\{S_n(f)\}_{n=1}^\infty$  is Cauchy in  $X$ . The calculation is similar to the one we met when defining the Lebesgue measure. For any fixed  $\epsilon > 0$ , there exists  $\delta > 0$  such that  $\|f(s) - f(t)\| \leq \epsilon$  if  $|s - t| \leq \delta$ . If  $1/2^n \leq \delta$ , then

$$\begin{aligned} \|S_{n+m}(f) - S_n(f)\| &= \left\| \sum_{j=1}^{2^n} \frac{b-a}{2^n} \frac{1}{2^m} \sum_{l=1}^{2^m} \left[ f\left(\frac{j}{2^n}\right) - f\left(\frac{j-1}{2^n} + \frac{l}{2^{n+m}}\right) \right] \right\| \\ &\leq \sum_{j=1}^{2^n} \frac{b-a}{2^n} \frac{1}{2^m} \sum_{l=1}^{2^m} \left\| f\left(\frac{j}{2^n}\right) - f\left(\frac{j-1}{2^n} + \frac{l}{2^{n+m}}\right) \right\| \\ &\leq (b-a)\epsilon. \end{aligned}$$

By definition,

$$(3.7.2) \quad \int_a^b f(t)dt = \lim_{n \rightarrow \infty} S_n(f).$$

---

<sup>2</sup>There are different definitions for the integral of a Banach space valued function. A reasonably general, Lebesgue style one, is provided by *Bochner integrals*. See e.g. [The Bochner integral](#) by Wenjing Wu. The weak version of the Bochner integral is the *Pettis integral*. Here, however, we are integrating continuous functions, and more elementary definitions of integral can be used.

You might complain that this definition depends on the choice of very special partitions of  $[a, b]$  (this makes additivity of the integral problematic, for instance). Fortunately, the dependence is apparent. Given a partition  $a = t_0 < t_1 < \cdots < t_{m-1} < t_m = b$  of  $[a, b]$  and sampling points  $t_{j-1} \leq t_j^* \leq t_j$ , let

$$S(\{t_j\}_{j=0}^m, \{t_j^*\}_{j=1}^m) = \sum_{j=1}^m f(t_j^*)(t_j - t_{j-1}).$$

The *resolution* of the partition is  $\min_{j=1, \dots, m} (t_j - t_{j-1}) = \delta(\{t_j\}_{j=0}^m)$ .

**PROPOSITION 3.6.** *Let  $X$  be Banach and  $f : [a, b] \rightarrow X$  continuous. Then, for all  $\epsilon > 0$  there is  $\delta > 0$  such that if  $\delta(\{t_j\}_{j=0}^m) \leq \delta$ , then*

$$\left\| \int_a^b f(t) dt - S(\{t_j\}_{j=0}^m, \{t_j^*\}_{j=1}^m) \right\| \leq \epsilon.$$

**PROOF.** Let  $n \geq 1$  be such that  $\left\| \int_a^b f(t) dt - S_n(f) \right\| \leq \epsilon$  and that  $\|f(s) - f(t)\| \leq \epsilon$  if  $|s - t| \leq 1/2^n$ . Consider a partition  $\{t_j\}_{j=0}^m$  with resolution less than  $\delta = 1/2^n$ , and a sampling set  $\{t_j^*\}_{j=1}^m$  as above. Below, we denote by  $|I|$  the Lebesgue measure of an interval  $I$ .

$$\begin{aligned} & \left\| S_n(f) - S(\{t_j\}_{j=0}^m, \{t_j^*\}_{j=1}^m) \right\| \\ &= \left\| \sum_{l=1}^{2^n} \sum_{j: [t_{j-1}, t_j] \cap [(l-1)/2^n, l/2^n] \neq \emptyset} (f(l/2^n) - f(t_j^*)) |[t_{j-1}, t_j] \cap [(l-1)/2^n, l/2^n]| \right\| \\ &\leq \sum_{l=1}^{2^n} \sum_{j: [t_{j-1}, t_j] \cap [(l-1)/2^n, l/2^n] \neq \emptyset} \|f(l/2^n) - f(t_j^*)\| \cdot |[t_{j-1}, t_j] \cap [(l-1)/2^n, l/2^n]| \\ &\leq (b-a)\epsilon. \end{aligned}$$

□

**EXERCISE 3.20.** *Let  $\alpha : [a, b] \rightarrow \mathbb{R}$  be increasing (or, more generally, let  $\alpha : [a, b] \rightarrow \mathbb{C}$  be of bounded variation). For  $f : [a, b] \rightarrow X$  continuous (where  $X$  is Banach) provide a definition of the vector valued Stieltjes integral*

$$\int_a^b f(t) d\alpha(t),$$

*and show that it is well defined.*

**COROLLARY 3.5.** *Let  $f : [a, b] \rightarrow X$  be a continuous function with values in a Banach space  $X$ , and let  $T : X \rightarrow Y$  be a bounded operator between  $X$  and another Banach space  $Y$ . Then,*

$$T \left( \int_a^b f(t) dt \right) = \int_a^b T(f(t)) dt.$$

PROOF. It suffices to pass in the limit for  $n \rightarrow \infty$  the two sides of the equality

$$T \left( \sum_{j=1}^m f(t_j^*)(t_j - t_{j-1}) \right) = \sum_{j=1}^m T[f(t_j^*)] (t_j - t_{j-1}).$$

□

We give a sample application to Banach space valued holomorphic functions. The reader might find it interesting and rewarding to translate in the Banach valued world the chapter of a book of complex analysis concerning complex integrals.

**COROLLARY 3.6.** *[Morera's theorem for Banach space valued functions] Let  $f : \Omega \rightarrow X$  be a continuous, Banach space valued function. Then  $f$  is holomorphic if and only if  $\int_{\partial T} f(z)dz = 0$  for all triangles  $T$  contained in  $\Omega$ .*

PROOF. The function  $f$  is holomorphic if and only if it is weakly holomorphic, and by Morera's theorem this holds if and only if

$$0 = \int_{\partial T} l(f(z))dz = l \left( \int_{\partial T} f(z)dz \right)$$

for all triangles and  $l \in X^*$ , the second equality following from corollary 3.5. Thus,  $\int_{\partial T} f(z)dz = 0$ . □

### 3.8. Sesquilinear forms on Hilbert spaces

In this section we consider a few applications of the Banach theory considered above to Hilbert space theory.

Let  $H$  be a Hilbert space. A map  $[\cdot, \cdot] : H \times H \rightarrow \mathbb{C}$ ,  $[x, y] = \langle x, y \rangle$ , is a *sesquilinear form* if it is conjugate linear in the first component and linear in the second.

$$[x, \lambda y + \mu z] = \lambda[x, y] + \mu[x, z], \quad [\lambda y + \mu z, x] = \bar{\lambda}[y, x] + \bar{\mu}[z, x].$$

For instance, the inner product  $\langle \cdot, \cdot \rangle$  is sesquilinear.

Sesquilinear forms are determined by their action on the diagonal.

**PROPOSITION 3.7.** *[Polarization identity] Let  $H$  be a complex Hilbert space. If  $[\cdot, \cdot]$  is a sesquilinear form, then the expression  $[x, y]$  can be expressed as a linear combination of expressions of the form  $[z, z]$ :*

$$(3.8.1) \quad [x, y] = \frac{1}{4}([x+y, x+y] - [x-y, x-y] - i[x+iy, x+iy] + i[x-iy, x-iy]).$$

The proof consists in working out the right hand side of (3.8.1). The polarization identity holds in Hilbert spaces over  $\mathbb{R}$  under the extra hypothesis that  $[x, y] = [y, x]$  is commutative. The form

$$\left[ \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \right] = (x, y) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = xu + xv + yv$$

provides an example of a bi-linear form on  $\mathbb{R}^2$  for which the polarization identity does not hold.

A sesquilinear form  $H \times H \xrightarrow{[\cdot, \cdot]} \mathbb{C}$  is *bounded* if there is a positive constant  $C$  such that

$$|[x, y]| \leq C \|x\| \cdot \|y\|$$

for all  $x, y$  in  $H$ . Using homogeneity in each component, this is equivalent to requiring that there is a (possibly different) constant  $C > 0$  such that

$$\sup_{\|x\|, \|y\| \leq 1} |[x, y]| \leq C.$$

All bounded sesquilinear forms are obtained by "sandwiching" a bounded linear operator between two vectors.

PROPOSITION 3.8. (i) If  $A \in \mathcal{L}(H)$ , then

$$(3.8.2) \quad [x, y] := \langle x, Ay \rangle$$

is a bounded, sesquilinear form:  $|[x, y]| \leq \|A\| \|x\| \|y\|$ .

(ii) Viceversa, for any bounded sesquilinear form  $[\cdot, \cdot]$  there exists  $A \in \mathcal{L}(H)$ .

PROOF. (i) is clear. (ii) The functional  $y \mapsto [x, y]$  is bounded, then, by Riesz theorem for Hilbert spaces, there is  $Ax \in H$  such that  $[x, y] = \langle Ax, y \rangle$ . It is easy to show that  $x \mapsto Ax$  is a linear map from  $H$  to  $H$ . Clearly  $A$  is bounded,

$$\|A\| = \sup_{\|y\| \leq 1} \sup_{\|x\| \leq 1} |[x, y]|.$$

□

Boundedness can be checked separately for each variable.

PROPOSITION 3.9. Let  $[\cdot, \cdot]$  be a sesquilinear form on  $H$ . If  $x \mapsto [x, y]$  is a bounded conjugate-linear functional for each  $y$ , and  $y \mapsto [x, y]$  is a bounded linear functional for each  $x$ , then  $[\cdot, \cdot]$  is bounded.

PROOF. Since  $x \mapsto \overline{[x, y]} = \lambda_y(x)$  is a bounded linear functional, there is  $Ay \in H$  such that  $[x, y] = \langle x, Ay \rangle$ . The map  $A$  is clearly linear,  $A : H \rightarrow H$ , and  $\|\lambda_y\|_{H^*} = \|Ay\|$ . For each  $x$  we have that

$$\sup_{\|y\| \leq 1} |[x, y]| \leq C(x),$$

a constant depending on  $x$ . By Banach-Steinhaus,

$$\infty > \sup_{\|y\| \leq 1} \|\lambda_y\|_{H^*} = \sup_{\|y\| \leq 1} \|Ay\| = \|A\|,$$

as wished.  $\square$

The sesquilinear form  $[\cdot, \cdot]$  is *symmetric* if  $[y, x] = \overline{[x, y]}$ . If there is a linear operator  $A : H \rightarrow H$  such that  $[x, y] = [x, Ay]$ , symmetry translates into:

$$\langle x, Ay \rangle = [x, y] = \overline{[y, x]} = \overline{\langle y, Ax \rangle} = \langle Ax, y \rangle.$$

That is,  $A = A^*$  is self-adjoint.

It is remarkable that all such symmetric sesquilinear forms on a Hilbert space are bounded.

**THEOREM 3.18.** [*Hellinger-Toeplitz*] Let  $A : H \rightarrow H$  be an everywhere defined self-adjoint operator on a Hilbert space  $H$ :

$$\langle Ax, y \rangle = \langle x, Ay \rangle \text{ if } x, y \in H.$$

Then,  $A$  is bounded.

PROOF. Let  $(x_n, Ax_n)$  be a sequence in  $\Gamma(A)$  with  $x_n \rightarrow x$  and  $Ax_n \rightarrow y$ . For all  $z \in H$ :

$$\begin{aligned} \langle z, y \rangle &= \lim_{n \rightarrow \infty} \langle z, Ax_n \rangle = \lim_{n \rightarrow \infty} \langle Az, x_n \rangle \\ &= \langle Az, x \rangle = \lim_{n \rightarrow \infty} \langle z, Ax \rangle, \end{aligned}$$

hence,  $Ax = y$ , so  $\Gamma(A)$  is closed, and we can use the closed graph theorem.  $\square$

A sesquilinear  $[\cdot, \cdot]$  form is *coercive* if

$$(3.8.3) \quad [x, x] \geq c\|x\|^2$$

for some  $c > 0$ .

**THEOREM 3.19 (Lax-Milgram).** *Let  $[\cdot, \cdot]$  be a bounded sesquilinear form with bound  $C$ , which is coercive with constant  $c > 0$ . Then, for all  $w$  in  $H$  there is a unique  $x$  such that:*

$$(3.8.4) \quad [x, y] = \langle w, y \rangle.$$

**PROOF.** We saw earlier that  $[x, y] = \langle Ax, y \rangle$ , with  $A \in \mathcal{L}(H)$ . We have to prove that  $A$  is a bijection.

(i)  **$A$  is injective.** In fact,

$$c\|x\|^2 \leq [x, x] = \langle Ax, x \rangle \leq \|x\| \cdot \|Ax\|,$$

which implies that  $x = 0$  when  $Ax = 0$ .

(ii)  **$\text{Ran}(A)$  is closed in  $H$ .** Let  $Ax_n \rightarrow w$  in  $H$ . In particular,  $\{Ax_n\}$  is Cauchy and, using coercivity as in (i),

$$\|x_n - x_{n+j}\| \leq 1/c \|Ax_n - Ax_{n+j}\|,$$

so that  $\{x_n\}$  is Cauchy, too, hence  $x_n \rightarrow x$ . Since  $A$  is bounded,  $w = \lim_n Ax_n = Ax$ .

(iii)  **$\text{Ran}(A) = H$ .** By (ii) and the general theory, we can write  $H = \text{Ran}(A) \oplus (H \ominus \text{Ran}(A))$ . Let  $k \in H \ominus \text{Ran}(A)$ :

$$c\|w\|^2 \leq [w, w] = \langle Aw, w \rangle = 0,$$

thus  $w = 0$ . Hence,  $\text{Ran}(A) = H$ .

□

### 3.9. The Hilbert space adjoint of an operator

In this section we define the Hilbert space adjoint of a bounded operator, and prove some of its properties. This material is not prerequisite for the chapter on spectral theory.

**3.9.1. Basic properties of the adjoint.** Let  $T \in \mathcal{L}(H)$ . Its *adjoint*  $T^* \in \mathcal{L}(H)$  is defined on  $y \in H$  by:

$$(3.9.1) \quad \langle x, T^*y \rangle = \langle Tx, y \rangle.$$

Equation (3.9.1) defines a linear operator. For  $x, y, z \in H$  and  $a, b \in \mathbb{C}$ ,

$$\langle x, T^*(ay + bz) \rangle = \langle Tx, ay + bz \rangle$$



$$\begin{aligned}
&= a\langle Tx, y \rangle + b\langle Tx, z \rangle \\
&= a\langle x, T^*y \rangle + b\langle x, T^*z \rangle \\
&= \langle x, aT^*y + bT^*z \rangle.
\end{aligned}$$

The map  $T \mapsto T^*$  is conjugate linear,

$$\begin{aligned}
\langle x, (aS + bT)^*y \rangle &= \langle (aS + bT)x, y \rangle \\
&= \bar{a}\langle Sx, y \rangle + \bar{b}\langle Tx, y \rangle \\
&= \langle x, (\bar{a}S^* + \bar{b}T^*)y \rangle.
\end{aligned}$$

The operator  $T$  is *self-adjoint* if  $T^* = T$ , and, more generally, it is *normal* if  $TT^* = T^*T$ .

**THEOREM 3.20.** *Let  $S, T$  be bounded operators on  $H$ .*

- (i)  $(T^*)^* = T$ .
- (ii)  $\|T^*\| = \|T\|$ .
- (iii)  $(TS)^* = S^*T^*$ .
- (iv)  $T$  is invertible if and only if  $T^*$  is invertible, and  $(T^{-1})^* = (T^*)^{-1}$ .
- (v)  $T \mapsto T^*$  is continuous with respect to the uniform and the weak operator topology (but not with respect to the strong operator topology).
- (vi)  $\|T^*T\| = \|T\|^2$ .

**PROOF.** (i) We have

$$\langle x, (T^*)^*y \rangle = \langle T^*x, y \rangle = \overline{\langle y, T^*x \rangle} = \overline{\langle Ty, x \rangle} = \langle x, Ty \rangle.$$

(ii) We have

$$\begin{aligned}
\|T^*\| &= \sup_{\|y\|=1} \|Ty\| = \sup_{\|x\|=1} \sup_{\|y\|=1} |\langle x, T^*y \rangle| \\
&= \sup_{\|x\|=1} \sup_{\|y\|=1} |\langle Tx, y \rangle| \\
&\leq \|T\|,
\end{aligned}$$

showing that  $\|T^*\| \leq \|T\|$ , and so, by (i),  $\|T\| = \|(T^*)^*\| \leq \|T^*\|$ .

(iii) Simple algebra.

(iv) By definition,

$$\langle x, (T^{-1})^*T^*y \rangle = \langle TT^{-1}x, y \rangle = \langle x, y \rangle = \langle T^{-1}Tx, y \rangle = \langle x, T^*(T^{-1})^*y \rangle.$$

(v) For the uniform topology, see (ii). For the weak operator topology,

$$\langle x, (T_n^* - T^*)y \rangle = \langle (T_n - T)x, y \rangle \rightarrow 0$$

if  $T_n \xrightarrow{WOT} T$ .

(vi) The inequality  $\|T^*T\| \leq \|T\|^2$  follows from (ii). In the other direction,

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^*Tx \rangle \leq \|T^*T\| \cdot \|x\|^2.$$

□

About the negative statement on SOT in (iii), see the examples after the definition of the operator topologies.

**COROLLARY 3.7.** (i) *If  $S$  and  $T$  are self-adjoint and commute,  $ST = TS$ , then  $ST$  is self-adjoint.*

(ii) *If  $T$  is self-adjoint, then  $\|T^2\| = \|T\|^2$ , and  $r(T)$ , the spectral radius of  $T$ , is*

$$(3.9.2) \quad r(T) = \|T\|.$$

**PROOF.** (i) is a simple algebraic calculation,

$$\begin{aligned} \langle x, TSy \rangle &= \langle x, STy \rangle = \langle T^*S^*x, y \rangle \\ &= \langle TSx, y \rangle, \end{aligned}$$

shows that  $(TS)^* = TS$ .

The first assertion in (ii) follows from (vi) in the previous theorem, so that

$$r(T) = \lim_{n \rightarrow \infty} \|T^n\|^{1/n} = \lim_{m \rightarrow \infty} \|T^{2^m}\|^{1/2^m} = \|T\|,$$

since  $\|T^{2^m}\| = \|T^{2^{m-1}}\|^2$ . □

### 3.9.2. Spectrum and adjoint.

**LEMMA 3.2.** *The spectrum of  $T^*$  is  $\sigma(T^*) = \overline{\sigma(T)}$ .*

**PROOF.** For complex  $\lambda$ ,  $\lambda I - T$  is invertible if and only if  $(\lambda I - T)^* = \bar{\lambda}I - T^*$  is invertible. □

When we look more closely inside the spectrum, the picture is more complex. For instance, the shift  $\tau_1$  on  $\ell^2(\mathbb{N})$  does not have eigenvalues, while any  $\lambda \in \mathbb{D}$  is an eigenvalue for the the back-shift  $\tau_1^*$ . We split the spectrum of  $T$  into three components.

(i) the *point spectrum*  $\sigma_p(T)$  is the set of the *eigenvalues*  $\lambda$  of  $T$ .

$$(3.9.3) \quad Tv = \lambda v$$

for some  $v \neq 0$  in  $H$ . Any such  $v$  is called an *eigenvector* relative to  $\lambda$ . The closed, linear subspace  $V_\lambda$  of the eigenvectors relative to  $\lambda$  is the *eigenspace* relative to  $\lambda$ . In particular, if  $\ker(T) \neq 0$ , 0 is an eigenvalue with eigenspace  $\ker(T)$ .

(ii) The *continuous spectrum*  $\sigma_c(T)$  is the set of those  $\lambda \in \sigma(T)$  such that  $\lambda I - T$  has dense range in  $H$ .

(iii) The *residual spectrum*  $\sigma_{res}(T)$  of  $T$  is the set of those  $\lambda$  which are not eigenvalues, and such that  $\lambda I - T$  does not have dense range in  $H$ .

Items (ii) and (iii) only happen in infinite dimension. When  $H = \mathbb{C}^N$ , the spectrum is the set of the eigenvalues, and if  $\lambda$  is an eigenvalue of  $A$ , then  $\dim(\text{Ran}(\lambda I - A)) = N - \dim(\ker(\lambda I - A)) < N$ , hence the range of  $\lambda I - A$  is not dense.

PROPOSITION 3.10. (i) If  $\lambda \in \sigma_{res}(T)$ , then  $\bar{\lambda} \in \sigma_d(T)$ .

(ii) If  $\lambda \in \sigma_p(T)$ , then  $\bar{\lambda} \in \sigma_d(T) \cup \sigma_{res}(T)$ .

PROOF. (i) Suppose  $\lambda \in \sigma_{res}(T)$  and consider  $0 \neq k \in H \ominus \text{Cl}(\text{Ran}(\lambda I - T))$ . For all  $x$  in  $H$ ,

$$\langle x, (\bar{\lambda}I - T^*)k \rangle = \langle (\lambda I - T)x, k \rangle = 0,$$

hence,  $T^*k = \bar{\lambda}k$ .

(ii) Let  $k \neq 0$  be such that  $Tk = \lambda k$ . Then, for all  $x$  in  $H$ .

$$\langle (\bar{\lambda}I - T^*)x, k \rangle = \langle x, (\lambda I - T)k \rangle = 0,$$

showing that  $k$  is orthogonal to  $\text{Ran}(\bar{\lambda}I - T^*)$ , which is then not dense in  $H$ .  $\square$

Consider for instance the spectra of the shift  $M_z$  on  $H^2(\mathbb{D})$ , and of its adjoint  $M_z^*$ ,

$$M_z^*f(z) = \frac{f(z) - f(0)}{z}.$$

If  $\lambda \in \mathbb{D}$ , then it is an eigenvalue of  $M_z^*$ . Since  $\sigma_p(M_z) = \emptyset$ , by (ii)  $\bar{\lambda} \in \sigma_{res}(M_z)$ .

If  $|\lambda| = 1$  and  $\lambda \in \sigma_{res}(M_z)$ , by (i) we would have  $\bar{\lambda} \in \sigma_p(M_z^*)$ , which is not true because the non-vanishing solutions of

$$\frac{f(z) - f(0)}{z} = \bar{\lambda}f(z)$$

are multiples of  $f(z) = \frac{1}{1-\bar{\lambda}z} = \sum_{n=0}^{\infty} \bar{\lambda}^n z^n$ , which does not belong to  $H^2$ . Hence,  $\lambda \in \sigma_c(M_z)$ .

Again by (i), if  $|\lambda| = 1$ ,  $\lambda \notin \sigma_{res}(M_z^*)$ , because  $\bar{\lambda}$  is not an eigenvalue of  $M_z$ . Summarizing:

$$(i) \quad \sigma_p(M_z^*) = \mathbb{D}, \sigma_c(M_z^*) = \mathbb{T}, \sigma_{res}(M_z^*) = \emptyset.$$

$$(ii) \quad \sigma_p(M_z) = \emptyset, \sigma_c(M_z) = \mathbb{T}, \sigma_{res}(M_z) = \mathbb{D}.$$

**3.9.3. The spectrum of a bounded, self-adjoint operator.** We have next a foundational result on self-adjoint operators.

**THEOREM 3.21.** *If  $T = T^*$ , then  $\sigma(T) \subset \mathbb{R}$ .*

**PROOF.** We show that if  $a, b \in \mathbb{R}$  and  $b \neq 0$ , then  $a + ib \in \rho(T)$ . Consider the sesquilinear form

$$B(x, y) = \langle (T - (a + ib))x, (T - (a + ib))y \rangle = \langle (T - (a - ib))(T - (a + ib))x, y \rangle,$$

which is coercive since, using again  $T = T^*$ ,

$$\begin{aligned} B(x, x) &= \langle (T - (a + ib))x, (T - (a + ib))x \rangle \\ &= \|(T - a)x\|^2 + b^2\|x\|^2 + ib(-\langle Tx, x \rangle + \langle x, Tx \rangle) \\ &= \|(T - a)x\|^2 + b^2\|x\|^2 \\ &\geq b^2\|x\|^2. \end{aligned}$$

By Lax-Milgram lemma,  $(T - (a - ib))(T - (a + ib))$  is invertible, hence  $T - (a + ib)$  is 1-1 and  $T - (a - ib)$  is onto. Applying the same reasoning  $T - (a - ib)$  instead of  $T - (a + ib)$ , we have that  $T - (a + ib)$  is 1-1 and  $T - (a - ib)$  is onto. Hence,  $T - (a + ib)$  is invertible (incidentally, as well as  $T - (a - ib)$ ).  $\square$

We record here a useful corollary of Lax-Milgram lemma.

**COROLLARY 3.8.** *Let  $\mu \in \mathbb{C}$ . Then,  $\mu \in \sigma(A)$  if and only if there exist vectors  $\{h_j\}_{j=1}^{\infty}$  in  $H$  with  $\|h_j\| = 1$ , such that*

$$(3.9.4) \quad \|(A - \mu I)h_j\| \rightarrow 0.$$

PROOF. Suppose such vectors do not exist. Then, there is  $\epsilon > 0$  such that  $\|(A - \mu I)h\| \geq \epsilon\|h\|$ , hence the sesquilinear form

$$B(h, k) = \langle (A - \mu I)h, (A - \mu I)k \rangle = \langle (A - \mu I)(A - \mu I)h, k \rangle$$

is coercive, and by Lax-Milgram  $(A - \mu I)^2$  is invertible, hence  $A - \mu I$  is invertible,  $\mu \in \rho(A)$ .

In the other direction, suppose  $A - \mu I$  is invertible. For  $\|h\| = 1$ ,

$$\infty > \|(A - \mu I)^{-1}\| \geq \frac{\|(A - \mu I)^{-1}(A - \mu I)h\|}{\|(A - \mu I)h\|} = \frac{1}{\|(A - \mu I)h\|},$$

showing that (3.9.4) does not hold.  $\square$

The spectrum of a self-adjoint operator has no residual part.

COROLLARY 3.9. *If  $T = T^*$ , then  $\sigma_{res}(T) = \emptyset$ .*

PROOF. If  $\lambda \in \sigma_{res}(T)$ , then  $\lambda = \bar{\lambda} \in \sigma_p(T^*) = \sigma_p(T)$ , which is a contradiction.  $\square$

In infinite dimensions, it can happen that  $A : H \rightarrow H$  has an eigenvalue  $\lambda$ , yet  $\text{Ran}(\lambda I - A)$  is the whole of  $H$ : think of the back-shift  $\tau_1^*$ , for which  $\ker(\tau_1^*) \neq 0$  ( $\lambda = 0$  is an eigenvalue), but  $\text{Ran}(\tau_1^*) = \ell^2(\mathbb{N})$ . If  $A = A^*$  is self-adjoint, however, things are similar to the finite dimensional case.

COROLLARY 3.10. *If  $T = T^*$  and  $\lambda \in \sigma_p(T)$ , then  $\text{Ran}(\lambda I - T)$  is not dense in  $H$ .*

PROOF. Since  $\lambda$  is real,  $\lambda I - T$  is self-adjoint, and it suffices to show the statement for  $\lambda = 0$ . We show, more precisely, that if  $0 \neq x \in \ker(T)$ , then  $x \perp \text{Ran}(T)$ , so that the latter is not dense in  $H$ . In fact,

$$\langle Ty, x \rangle = \langle y, Tx \rangle = 0.$$

$\square$

Actually, we can be more precise in locating the spectrum of  $T = T^*$ .

THEOREM 3.22. *If  $T$  is selfadjoint, and*

$$(3.9.5) \quad a(T) := \inf_{\|x\|=1} \langle x, Tx \rangle \text{ and } b(T) := \sup_{\|x\|=1} \langle x, Tx \rangle,$$

*then  $a(T), b(T) \in \sigma(T)$  and  $[a(T), b(T)] \supseteq \sigma(T)$ .*

PROOF. Observe first that  $a(T + cI) = a(T) + c$ ,  $b(T + cI) = b(T) + c$ , and

$$\sigma(T + cI) = \{\lambda : \lambda I - (T + cI) \text{ is not invertible}\} = \sigma(T) + c,$$

and we can suppose  $a(T) = 0$ . If  $\lambda > b(T)$ , then

$$\langle x, (\lambda I - T)x \rangle \geq (\lambda - b(T))\|x\|^2$$

shows that  $B(x, y) := \langle x, (\lambda I - T)y \rangle$  is bounded and coercive, and we have seen before that this implies that  $\lambda I - T$  is invertible, hence that  $\lambda \in \rho(T)$ . The same reasoning applies, changing sign, with  $\lambda < a(T)$ . This shows  $\sigma(T) \subseteq [a(T), b(T)]$ .

On the other hand,  $b(T) \leq \|T\| = r(T)$ , the spectral radius of  $T$ . Hence,  $b(T) \in \sigma(T)$ . The same reasoning, but choosing the normalization  $b(T) = 0$ , shows that  $a(T) \in \sigma(T)$ .  $\square$

**3.9.4. The Hausdorff distance between spectra.** Given two closed sets  $A, B$  in  $\mathbb{R}$  (or in any metric space), their *Hausdorff distance*  $d_H(A, B)$  is defined as

$$(3.9.6) \quad d_H(A, B) = \max \left( \sup_{x \in A} \inf_{y \in B} |x - y|, \sup_{y \in B} \inf_{x \in A} |x - y| \right).$$

To have a better picture of this quantity, let  $d(x, B) = \inf_{y \in B} |x - y|$ . Asking  $\sup_{s \in A} \inf_{y \in B} |x - y| \leq \epsilon$  means that all points of  $A$  are within distance  $\epsilon$  from  $B$ , and asking  $d_H(A, B) \leq \epsilon$  means that this requirement also holds with the roles of  $A$  and  $B$  interchanged. When  $\epsilon = 0$ , we are requiring that  $A \subseteq B$  and  $B \subseteq A$ , i.e.  $A = B$ .

**EXERCISE 3.21.** Let  $A, B$  be subsets, not necessarily closed, of a metric space. Show that  $d_H(A, B) = 0$  if and only if  $\text{Cl}(A) = \text{Cl}(B)$ .

Also, show that  $d_H$  defines a distance, when restricted to nonempty closed sets.

**THEOREM 3.23.** If  $S, T$  are self-adjoint, then

$$(3.9.7) \quad d_H(\sigma(S), \sigma(T)) \leq \|S - T\|.$$

PROOF. The case  $S = T$  is trivial. By compactness of spectra, we can replace  $\sup / \inf$  by  $\max / \min$ . Suppose by contradiction that  $\max_{s \in \sigma(S)} \min_{t \in \sigma(T)} |s - t| > \|S - T\|$ , i.e. that for some  $s \in \sigma(S)$ ,

$$\min_{t \in \sigma(T)} |s - t| > \|S - T\|.$$

Thus  $s \in \rho(T)$ . i.e.  $T - sI$  is invertible, and by the spectral mapping theorem for Banach algebras,

$$\sigma((T - sI)^{-1}) = (\sigma(T) - s)^{-1}.$$

The spectral radius  $r((\sigma(T) - sI)^{-1})$  satisfies, since  $(\sigma(T) - sI)^{-1}$  is self-adjoint:

$$(3.9.8) \quad \begin{aligned} \|(T - sI)^{-1}\| &= r((\sigma(T) - sI)^{-1}) \\ &= \max_{t \in \sigma(T)} |t - s|^{-1} < \|S - T\|^{-1}. \end{aligned}$$

We next decompose  $S - sI$  in a way which allows us to use the information we have gathered so far:

$$S - sI = T - sI + S - T = (T - sI) (I + (T - sI)^{-1}(S - T)).$$

The first factor is invertible, and the second is, too, because by (3.9.8)

$$\|(T - sI)^{-1}(S - T)\| \leq \|(T - sI)^{-1}\| \cdot \|S - T\| < 1.$$

This shows that  $s \notin \sigma(S)$ , a contradiction.  $\square$

We finish this section with some observations on the adjoint.

**EXERCISE 3.22.** *Prove the following properties.*

- (i) Let  $T \in \mathcal{L}(H)$ . Then,  $T_R = \frac{T+T^*}{2}$  and  $T_I = \frac{T-T^*}{2i}$  are self-adjoint, and

$$T = T_R + iT_I.$$

*This decomposition is somehow analogous to the algebraic expression of a complex number.*

- (ii) *The decomposition in (i) is unique. If  $T = A + iB$  with  $A, B$  self-adjoint, then  $A = T_R$  and  $B = T_I$ .*

- (iii) *If  $A$  is self-adjoint, then  $\langle x, Ax \rangle$  is real,*

$$\langle x, Ax \rangle = \overline{\langle Ax, x \rangle} = \overline{\langle x, Ax \rangle}.$$

- (iv)  *$A$  is selfadjoint if and only if  $B = iA$  is **skew-adjoint**,  $B^* = -B$ . In this case,  $\langle x, Bx \rangle$  is purely imaginary.*

- (v)  *$T_I T_R = T_R T_I$  if and only if  $T^* T = T T^*$ , i.e. if  $T$  is **normal**.*

- (vi) *If  $T$  is normal, then  $\|T\| = \|T^* T\|^{1/2} = \sqrt{\|T_I^2 + T_R^2\|}$ .*





## CHAPTER 4

### Compact operators and their spectra

If you are just interested in compact, self-adjoint operators on a Hilbert space, you find a different narrative of the same theory in chapter 7.

#### 4.1. Preliminaries on compactness

**4.1.1. Compact subsets of Banach spaces.** The following lemma shows that closed subspaces of Banach spaces exhibit some orthogonality properties.

**LEMMA 4.1** (Riesz lemma in Banach spaces). *Let  $X$  be a Banach space and  $Y \subset X$  a proper, closed subspace. Let  $0 < \alpha < 1$ . Then, there exists  $z$  in  $X$  with  $\|z\| = 1$  such that  $\alpha \leq d(z, Y) \leq 1$ .*

**PROOF.** Let  $v \in X \setminus Y$ ,  $a := d(v, Y) > 0$ , and choose  $y_0 \in Y$  such that  $a \leq \|v - y_0\| \leq \frac{a}{\alpha}$  (draw a picture to have an intuition of what's going on). Set  $z = \frac{v - y_0}{\|v - y_0\|}$ , so that  $\|z\| = 1$ . For all  $y$  in  $Y$  we have

$$\begin{aligned} \|z - y\| &= \frac{\|v - y_0 - \|v - y_0\|y\|}{\|v - y_0\|} \\ &= \frac{\|v - y_1\|}{\|v - y_0\|} \text{ with } y_1 \in Y \\ &\geq \frac{a}{\|v - y_0\|} \geq \alpha \end{aligned}$$

by our choice of  $y_0$ . □

**COROLLARY 4.1.** *Let  $X$  be a Banach space. The unit ball in  $X$  is compact in the norm topology if and only if  $X$  is finite dimensional.*

**PROOF.** The *if* direction easily follows from Weierstrass theorem on extremals of continuous functions.

In the other direction, suppose  $X$  is infinite dimensional, and construct a sequence of vectors and subspaces as follows.

- (i) Pick any vector  $x_1$  with  $\|x_1\| = 1$ , and let  $Y_1 = \text{span}\{x_1\} \subset X$ .

- (ii) By Riesz lemma, you can pick a second vector  $x_2$  with  $\|x_2\| = 1$  and  $d(x_2, Y_1) \geq 1 - 1/2$ , and let  $Y_2 = \text{span}\{x_1, x_2\} \subset X$ .
- (iii) Iterate, so that  $d(x_n, Y_{n-1}) \geq 1 - 1/n$ , and let  $Y_n = \text{span}\{x_1, \dots, x_n\} \subset X$ .

This way,  $\|x_m - x_n\| > 1/3$  for all  $m \neq n$ . Cover the unit ball by open balls of radius  $1/6$ , so that no one can contain more than one point from the sequence  $\{x_n\}$ ; hence, the unit ball is not compact.  $\square$

**4.1.2. Compact operators: generalities.** Linear functionals on a Banach space  $X$  are maps  $l : X \rightarrow \mathbb{C}$ , with values in the 1-dimensional Banach space  $\mathbb{C}$ . Next level of complexity are bounded, linear maps  $T : X \rightarrow Y$ , where  $Y$  is a finite dimensional Banach space. If  $\{l_n\}_{n=1}^N$  is a basis for  $Y^*$ , then  $C : y \mapsto (l_1(y), \dots, l_N(y)) = \sum_{n=1}^N l_j(y)e_j$  is an isomorphism between  $Y$  and  $\mathbb{C}^N$  ( $\{e_j\}$  is here the standard basis of  $\mathbb{C}^N$ ), and  $C \circ T(x) = \sum_{n=1}^N l_j(Tx)e_j$  identifies  $T$  with a map from  $X$  to  $\mathbb{C}^N$ . All this makes sense with little changes if  $T(X)$  is finite dimensional in  $Y$ , a Banach space. In this case we say that  $T$  is *finite rank*.

Now the closed unit ball  $\overline{B_1}$  in  $X$  is not compact if  $X$  is infinite dimensional and Banach; but if  $T$  is finite rank,  $T(\overline{B_1})$  is both bounded and closed in  $Y$ , hence it is compact.

We say that an operator  $T : X \rightarrow Y$  between Banach spaces is *compact* if  $\overline{T(B_1)}$  is compact in  $Y$ .

EXERCISE 4.1. Let  $m : \mathbb{N} \rightarrow \mathbb{C}$  be bounded, and consider the bounded operator  $M_b \varphi = m\varphi$ ,  $M_m : \ell^2 \rightarrow \ell^2$ . Show that  $M_m$  is finite rank if and only if  $m(j) = 0$  for all, but a finite set of  $j$ 's in  $\mathbb{N}$ ; and that it is compact if and only if  $\lim_{j \rightarrow \infty} m(j) = 0$ .

Recall that a sequence  $\{x_n\}$  in a Banach space  $X$  converges to  $x$  weakly,  $x_n \xrightarrow{w} x$ , if

$$\lim_{n \rightarrow \infty} l(x_n) = l(x)$$

for all  $l \in X^*$ .

The theorem below is crucial in many applications, e.g. to Calculus of Variations.

**THEOREM 4.1.** *If  $T$  is compact, it maps weakly convergent sequences to norm convergent sequences.*

**PROOF.** By Banach-Steinhaus,  $x_n \xrightarrow{w} x$  implies that  $\|x_n\| \leq C$  is norm bounded, hence

$$l(Tx_n) - l(Tx) = (T'l)(x_n - x) \rightarrow 0$$

as  $n \rightarrow \infty$  for all  $l \in Y^*$ . This means that  $T(x_n) \xrightarrow{w} T(x)$  weakly in  $Y$ , hence that  $\|Tx_n\| \leq K$  is bounded, too.

On the other hand, if by contradiction  $\|Tx_n - Tx\| \not\rightarrow 0$ , then  $\|Tx_{n_j} - Tx\| \geq \epsilon$  for a subsequence, and by compactness of  $T$  there exists a sub-subsequence  $\{Tx_{n_{j_k}}\}$  strongly convergent to some  $y \neq Tx$  in  $Y$ . But then  $Tx_{n_{j_k}} \xrightarrow{w} y \neq Tx$ , contradicting the assumption.  $\square$

The opposite implication holds, for instance, if  $X$  is a reflexive space.

**THEOREM 4.2.** *Let  $X \equiv X^{**}$  be a reflexive Banach space, and  $T : X \rightarrow Y$  be an operator with values in a Banach space  $Y$ . If  $T$  maps weakly convergent sequences to norm convergent sequences, then  $T$  is compact.*

**PROOF.** Let  $\bar{B}$  be the unit ball in  $X$  and suppose that  $\square$

The spectrum of a compact operator on an infinite dimensional Banach space always contains 0.

**PROPOSITION 4.1.** *Let  $T : X \rightarrow X$  be a compact operator on an infinite dimensional Banach space  $X$ . Then,  $T$  is not surjective. In particular,  $0 \in \sigma(T)$ .*

**PROOF.** Let  $B$  be the unit ball in  $X$ . Suppose  $T$  is surjective. By the open mapping theorem,  $T(\bar{B})$  contains a closed ball  $\epsilon\bar{B}$ , which is compact, since  $T(\bar{B})$  is compact. By corollary 4.1,  $X$  is finite dimensional.  $\square$

## 4.2. Compact operators on Hilbert spaces

**4.2.1. Compactness vs. finite rank.** The picture in the Hilbert space case is much cleaner.

**THEOREM 4.3.** *Let  $T \in \mathcal{L}(H)$  be a compact operator on a separable Hilbert space. Then, there are finite rank operators  $T_n$  on  $H$  such that  $\|T_n - T\| \rightarrow 0$ .*

**PROOF.** Let  $\{e_n\}$  be a o.n.b. for  $H$  and set

$$\lambda_n := \sup_{\|x\|=1, x \in \text{span}\{e_1, \dots, e_n\}^\perp} \|Tx\| \geq \lambda_{n+1}.$$

Let  $\lambda = \lim_n \lambda_n$ . We start by proving that  $\lambda = 0$ . Pick first  $x_n \in \text{span}\{e_1, \dots, e_n\}^\perp$  such that  $\|Tx_n\| \geq \lambda_n/2 \geq \lambda/2$ . Observe that  $x_n \xrightarrow{w} 0$ :

$$|\langle x_n, y \rangle| = \left| \langle x_n, \sum_{j=1}^{\infty} \langle y, e_j \rangle e_j \rangle \right|$$

$$\begin{aligned}
&= \left| \sum_{j=n+1}^{\infty} \langle x_n, e_j \rangle \overline{\langle y, e_j \rangle} \right| \\
&\leq \left( \sum_{j=n+1}^{\infty} |\langle x_n, e_j \rangle|^2 \right)^{1/2} \left( \sum_{j=n+1}^{\infty} |\langle y, e_j \rangle|^2 \right)^{1/2} \\
&\leq \|x_n\| \left( \sum_{j=n+1}^{\infty} |\langle y, e_j \rangle|^2 \right)^{1/2} \\
&\rightarrow 0
\end{aligned}$$

as  $n \rightarrow \infty$ . By the previous theorem,  $\|Tx_n\| \rightarrow 0$  because  $T$  is compact, hence,  $\lambda = 0$ .

Let now  $\pi_n$  be the projection onto  $\text{span}(e_1, \dots, e_n)$ , and

$$\begin{aligned}
T_n h &= (T\pi_n)h \\
&= T \left( \sum_{j=1}^n \langle e_j, h \rangle e_j \right) \\
&= \sum_{j=1}^n \langle e_j, h \rangle T e_j,
\end{aligned}$$

which has finite rank, and satisfies, with  $\pi_n^\perp$  being the projection onto  $\text{span}(e_1, \dots, e_n)^\perp$ :

$$\begin{aligned}
\|T - T_n\| &= \|T(I - \pi_n)\| \\
&= \|T\pi_n^\perp\| \\
&= \lambda_n \rightarrow 0
\end{aligned}$$

as  $n \rightarrow \infty$ . □

#### 4.2.2. Fredholm Theorem and its consequences.

**THEOREM 4.4.** *Let  $T$  be a compact operator on a separable Hilbert space  $H$ . Then,  $(I - zT)^{-1}$  exists for  $z \in \mathbb{C} \setminus D$ , where  $D$  is a discrete set in  $\mathbb{C}$ .*

*Moreover,  $z \in D$  if and only if the equation  $x = zTx$  has a solution  $x \neq 0$ .*

The idea of the proof consists in first localizing the problem, then reducing it, in substance, to a similar problem concerning finite rank operators. The latter can be rephrased in terms of zeros of a holomorphic determinant, and this is where the discrete set makes its appearance.

**PROOF.** We start with a formal calculation also involving a finite rank operator  $F$ ,

$$I - zT = (I - zT + F) - F$$

$$(4.2.1) \quad = (I - F(I - zT + F)^{-1})(I - zT + F),$$

provided  $(I - zT + F)^{-1}$  exists. The advantage of the first factor in the last expression over  $I - zT$  is that the former has the form  $I - [\textbf{finite rank operator}]$ , on which we can use tools from ordinary linear algebra.

We start by proving the theorem on discs. Let  $z_0$  a complex number, and consider  $z$ 's such that

$$\|zT - z_0T\| = |z - z_0| \cdot \|T\| < 1/2,$$

and a finite rank  $F$  such that

$$\|z_0T - F\| < 1/2,$$

so that  $(I - zT + F)^{-1} = \sum_{n=0}^{\infty} (zT - F)^n$  exists by the convergence of the Neumann series. Define then  $g(z) = F(I - zT + F)^{-1}$ , a finite rank operator which holomorphically depends on  $z$ , and note that above we have computed:

$$I - zT = (I - g(z))(I - zT + F),$$

the second factor on the right being invertible. Then,

- $x = g(z)x$  has a solution  $x \neq 0$  if and only if  $y = zTy$  has a solution  $y \neq 0$ ;
- $I - g(z)$  is invertible if and only if  $I - zT$  is invertible.

Let  $e_1, \dots, e_n$  be a bases for  $\text{Ran}(F)$ , and observe that a solution to  $x = g(z)x$  must have the form  $x = \sum_{j=0}^n a_j e_j$ , since it belongs to  $\text{Ran}(g(z)) \subseteq \text{Ran}(F)$ . On the other hand, the operator  $F$  must have the form:

$$Fx = \sum_{j=1}^n \langle f_j, x \rangle e_j$$

for some  $f_1, \dots, f_n \in H$ . Plug this into the equation,

$$\begin{aligned} \sum_{j=0}^n a_j e_j &= x = g(z)x \\ &= \sum_{l=0}^n a_l g(z)e_l \\ &= F((I - zT + F)^{-1}e_l) \\ &= \sum_{l=0}^n a_l \sum_{j=1}^n \langle f_j, (I - zT + F)^{-1}e_l \rangle e_j \end{aligned}$$

$$= \sum_{j=1}^n \left( \sum_{l=0}^n a_l \langle [(I - zT + F)^{-1}]^* f_j, e_l \rangle \right) e_j.$$

i.e. we look for a nonzero solution to

$$(4.2.2) \quad a_j = \sum_{l=0}^n a_l \langle [(I - zT + F)^{-1}]^* f_j, e_l \rangle = \sum_{l=0}^n a_l \langle f_j(z), e_l \rangle, \quad j = 1, \dots, n.$$

Such solution exists if and only if  $d(z) \neq 0$ , where

$$d(z) = \det [\delta_{j,l} - \langle f_j(z), e_l \rangle]_{j,l=1,\dots,n}.$$

The matrix has holomorphic entries, hence, one of the two can happen:  $d(z) \equiv 0$  in  $B(z_0, \frac{1}{2\|T\|})$ , or, there is a discrete subset  $D_{z_0}$  of  $B(z_0, \frac{1}{2\|T\|})$  such that  $d(z) = 0$  for  $z \in D_{z_0}$ .

We claim that if  $d(z) \neq 0$ , then  $I - g(z)$  is invertible, and so is  $I - zT$ . Fix  $u \in H$  and look for solution  $x = u + \sum_{j=0}^n a_j e_j$  to the equation

$$x - g(z)x = u,$$

i.e.

$$\begin{aligned} g(z)u &= \sum_{j=0}^n a_j e_j - g(z) \left( \sum_{l=0}^n a_l e_l \right) \\ &= \sum_{j=0}^n a_j e_j - \sum_{j=1}^n \left( \sum_{l=0}^n a_l \langle [(I - zT + F)^{-1}]^* f_j, e_l \rangle \right) e_j, \end{aligned}$$

which is a nonhomogeneous version of (4.2.2), having solution if  $d(z) \neq 0$ .

At this point, we have also shown that  $(I - zT)^{-1}$  does not exist if and only if  $x = zTx$  has a nonzero solution.

We have now to pass from local to global. The subset  $D'$  of the limit (accumulation) points of  $D$  is closed by general topology, and it is open by what we have just seen, hence it is either empty or the whole complex plane. Since  $\{1/z : z \in D\} \subseteq \sigma(T)$  is nonempty and compact,  $D' = \emptyset$ . Hence,  $D$  is discrete.  $\square$

**COROLLARY 4.2.** *[Fredholm alternative] Let  $T$  be a compact operator on a separable Hilbert space. If  $I - T$  is not invertible, then  $x = Tx$  has a nonzero solution.*

**PROOF.** Apply the theorem to  $z = 1$ .  $\square$

**THEOREM 4.5.** [*Riesz-Schauder Theorem*] *Let  $T$  be a compact operator on a separable Hilbert space  $H$ . Then,  $\sigma(T)$  can only have 0 as accumulation point (equivalently:  $\sigma(T) \setminus \{0\}$  is discrete in  $\mathbb{C} \setminus \{0\}$ ).*

*Moreover, the eigenspace  $E_\lambda$  relative to a nonzero  $\lambda \in \sigma(T)$  is finite dimensional.*

**PROOF.** For  $\lambda \neq 0$ , we have  $\lambda I - T = \lambda \left( I - \frac{1}{\lambda} T \right)$ , which is invertible for  $1/\lambda$  in a discrete set in  $\mathbb{C}$ , i.e. for  $\lambda$  in a bounded, discrete set in  $\mathbb{C} \setminus \{0\}$ , having 0 as accumulation point if the set of the  $1/\lambda$ 's was infinite.

Suppose that  $E_\lambda$  is infinite dimensional for an eigenvalue  $\lambda \neq 0$ , observe that  $E_\lambda$  is closed, and let  $\mathcal{E} = \{e_n\}_{n=1}^\infty$  be an o.n.b. for it. Then,  $\mathcal{E}$  is bounded, but  $T\mathcal{E} = \lambda\mathcal{E}$  is not precompact (as we have seen in an exercise).  $\square$

**4.2.3. Self-adjoint compact operators.** If we put together what we have seen so far on the topic, we have the fundamental:

**THEOREM 4.6.** [*Hilbert-Schmidt Theorem*] *Let  $A$  be a compact self-adjoint operator on a (infinite dimensional) separable Hilbert space  $H$ . Then  $\sigma(A) \subset \mathbb{R}$ . Moreover, there is a o.n.b.  $\{e_n\}_{n=1}^\infty$  and a real sequence  $\{\lambda_n\}_{n=1}^\infty$  of  $H$  (possibly with repetitions) such that*

$$Ae_n = \lambda_n e_n$$

and

$$(4.2.3) \quad \lim_{n \rightarrow \infty} \lambda_n = 0.$$

*If  $\lambda_n \neq 0$ , the eigenspaces  $E_{\lambda_n} = \{x \in H : Ax = \lambda_n x\}$  are finite dimensional.*

**PROOF.** Since  $A = A^*$ ,  $\sigma(A) \subset \mathbb{R}$ . For an eigenvalue  $\mu \neq 0$  of  $A$ , the image of the closed unit ball  $B$  in  $H$  contains  $|\mu|$  times the unit ball in  $E_\mu$ , thus the latter has to be compact, which only is if  $\dim(E_\mu) < \infty$ .

By Fredholm theorem,  $\sigma(A) \setminus \{0\} = \{\mu_m\}$ , where  $1/\mu_m$  has no point of accumulation. If the number of such  $\mu_m$ 's is finite, then  $A$  is finite rank. If it is infinite, then  $\mu_m \rightarrow 0$ .

Arrange the elements in decreasing order  $|\mu_1| \geq |\mu_2| \geq \dots > 0$ , where it might happen that  $\mu_i = -\mu_{i-1} > 0$  for some  $i$ 's. Let  $M_i = \sharp(E_{\mu_i})$ . Consider then a orthonormal basis  $e_1^i, \dots, e_{M_i}^i$  for each  $E_{\mu_i}$ , and a possibly countable orthonormal basis  $f_1, f_2, \dots$  of  $\ker(A)$ . For both the cases where the number of the nonzero eigenvalues is finite, or infinite, we can arrange all these vectors in a sequence  $\{e_n\}$  with  $Ae_n = \lambda_n$  (and  $\lambda_n$  is one of the  $\mu_i$ 's, or  $\lambda_n = 0$ ), and  $\lambda_n \rightarrow 0$  as  $n \rightarrow \infty$ .

Finally, if  $\lambda_m \neq \lambda_n$ , and  $Ax = \lambda_m x$ ,  $Ay = \lambda_n y$ , then

$$\lambda_m \langle x, y \rangle = \langle Ax, y \rangle = \langle x, Ay \rangle = \lambda_n \langle x, y \rangle,$$

hence,  $\langle x, y \rangle = 0$ . In particular,  $\{e_n\}$  is a orthonormal system. If it were not complete, then

$$K := \overline{\text{span}\{e_n\}} \subset H, \text{ with } A(K) = K.$$

We have then  $A : H \ominus K \rightarrow H \ominus K$ . In fact,  $\langle e_n, x \rangle = 0$  implies that  $\langle e_n, Ax \rangle = \langle Ae_n, x \rangle = \lambda_n \langle e_n, x \rangle = 0$ . Clearly  $A|_{H \ominus K} : H \ominus K \rightarrow H \ominus K$  is self-adjoint and compact. Since  $\ker(A) \subseteq K$ ,  $\ker(A|_{H \ominus K}) = 0$ . Then,  $A|_{H \ominus K}$  has an eigenvalue  $\mu \in \mathbb{R} \setminus \{0\}$ . The eigenvalue  $\mu$  is not even one of the previous nonzero eigenvalues, otherwise its eigenspace in  $H \ominus K$  would lie in  $E_{\mu_i} \subseteq K$  for some  $i$ . It is an eigenvalue for  $A$ , hence  $\{\mu_i\} \cup \{0\}$  did not exhaust  $\sigma(A)$ , and we have reached a contradiction.  $\square$

**COROLLARY 4.3.** *Let  $A$  be a compact self-adjoint operator on a (infinite dimensional) separable Hilbert space  $H$ . Then,*

$$(4.2.4) \quad H = \overline{\text{Ran}(A)} \oplus \ker(A).$$

**PROOF.** Let  $\{\lambda_n\}$  be the sequence of the nonzero eigenvalues of  $A$ . By Hilbert-Schmidt theorem, each  $x$  in  $H$  can be written as

$$x = \sum_n \Pi_{E_{\lambda_n}} x + \Pi_{\ker(A)} x,$$

and (i)  $Ax = \sum_n \lambda_n \Pi_{E_{\lambda_n}} x$ ; (ii)  $\sum_n \Pi_{E_{\lambda_n}} x \perp \Pi_{\ker(A)} x$ . Hence,

$$\text{Ran}(A) = \overline{\bigoplus_n E_{\lambda_n}} \perp \ker(A).$$

$\square$

### 4.3. The matrix representation of bounded operators

Any separable, infinite dimensional Hilbert space is unitarily isomorphic to  $\ell^2 = \ell^2(\mathbb{N}_*)$  (where  $\mathbb{N}_* = \{1, 2, \dots\}$  might be replaced by any infinite, countable set). To any linear operator  $A : H \rightarrow H$  we can then associate the infinite matrix  $[A_{m,n}]_{m,n=1}^\infty$ , where  $A_{m,n} = \langle e_m, Ae_n \rangle = \overline{A_{n,m}^*}$ . We can perform on infinite matrices representing bounded operators many of the calculations we are used to do in the finite dimensional case, which is often useful. In order to make calculations into proofs, it has to be clear in which sense the involved infinite sums converge, and which manipulations are licit. In this section we will see some results, examples, and counterexamples.

This material is not needed in the sequel, but it can be useful in working with concrete operators.



**4.3.1. Bounded operators on separable Hilbert spaces as infinite matrices.** If  $A \in \mathcal{L}(H)$  is bounded, and  $x = \sum x_n e_n$ , then, by the continuity of  $A$  and of the inner product,

$$(4.3.1) \quad \langle e_m, Ax \rangle = \left\langle e_m, A \sum_n x_n e_n \right\rangle = \left\langle e_m, \sum_n x_n A e_n \right\rangle$$

$$(4.3.2) \quad = \sum_n \langle e_m, A e_n \rangle x_n = \sum_n A_{m,n} x_n.$$

Since  $x \mapsto \langle e_m, Ax \rangle$  is a bounded linear functional on  $H$ , the row sequence  $\{A_{m,n}\}_{n=1}^\infty$  belongs to  $\ell^2$ . This also implies that the numerical series above converge absolutely. We will write  $A$  in terms of its rows and columns,

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \dots \end{pmatrix} = (A^1 | A^2 | \dots).$$

Basically, we have proved the following.

**PROPOSITION 4.2.** *The infinite matrix  $A = [A_{m,n}]$  represents a bounded operator on  $\ell^2$  if and only if:*

- (i) *each row of  $A$  belongs to  $\ell^2$ , and*
- (ii) *for all  $x \in \ell^2$ ,  $(A_{m,n} x_n)_{m=1}^\infty = (\langle A_m^t, x \rangle)_{m=1}^\infty$  (where  $^t$  denotes the transpose) belongs to  $\ell^2$ , and the operator  $(x_n)_{n=1}^\infty \mapsto (A_{m,n} x_n)_{m=1}^\infty$ .*

*Moreover, (i) and (ii) might be equivalently required to hold for the columns instead of the rows.*

An immediate consequence of proposition 4.2 is the following version of Fubini theorem for infinite sums, and the usual row-times-product rule for the product of matrices.

**COROLLARY 4.4** (Change of order of summation in infinite sums). *Let  $A$  be a bounded operator on  $\ell^2$  and  $x, y \in \ell^2$ . Then,*

$$(4.3.3) \quad \sum_n \left( \sum_m \overline{y_m} A_{m,n} \right) x_n = \sum_m \overline{y_m} \left( \sum_n A_{m,n} x_n \right).$$

*If  $A$  and  $B$  are bounded operators on  $\ell^2$ , then*

$$(4.3.4) \quad (BA)_{m,n} = \sum_j B_{m,j} A_{j,n}.$$

PROOF. The identity (4.3.3) follows from

$$\begin{aligned}
 \sum_m \overline{y_m} \left( \sum_n A_{m,n} x_n \right) &= \langle y, Ax \rangle \\
 &= \langle A^* y, x \rangle = \sum_n \left( \sum_m \overline{A_{n,m}^* y_m} \right) x_n \\
 &= \sum_n \left( \sum_m \overline{y_m} A_{m,n} \right) x_n.
 \end{aligned}$$

The row-times-product rule for the product of operators follows from Plancherel formula (used in the equality from first to second line):

$$\begin{aligned}
 (BA)_{m,n} &= \langle e_m, BAe_n \rangle = \langle B^* e_m, Ae_n \rangle \\
 &= \sum_j \overline{(B^* e_m)_j} (Ae_n)_j = \sum_j \langle e_j, B^* e_m \rangle \langle e_j, Ae_n \rangle \\
 &= \sum_j \langle e_m, Be_j \rangle \langle e_j, Ae_n \rangle \\
 &= \sum_j B_{m,j} A_{j,n}.
 \end{aligned}$$

□

The relation (4.3.3) can be used to change order of summation when several operators are composed. For instance, if  $A$  and  $B$  are bounded and  $x \in \ell^2$ , then

$$(4.3.5) \quad \sum_n \left( \sum_j B_{m,j} A_{j,n} \right) x_n = \sum_j B_{m,j} \left( \sum_n A_{j,n} x_n \right),$$

since  $(B_{m,j})_{j=1}^\infty$  lies in  $\ell^2$ .

We might prove (4.3.5) by a different argument:

$$\begin{aligned}
 \sum_n \left( \sum_j B_{m,j} A_{j,n} \right) x_n &= \sum_n (BA)_{m,n} x_n = (BAx)_m \\
 &= [B(Ax)]_m = \sum_j B_{m,j} (Ax)_j \\
 &= \sum_j B_{m,j} \left( \sum_n A_{j,n} x_n \right).
 \end{aligned}$$

EXERCISE 4.2. Let  $U$  be a bounded operator on  $\ell^2$ ,  $U = (U^1|U^2|\dots) = \begin{pmatrix} U_1 \\ U_2 \\ \dots \end{pmatrix}$ . Show that  $U$  is unitary if and only if  $U_1^t, U_2^t, \dots$  are an orthonormal basis for  $\ell^2$ , if and only if  $U^1, U^2, \dots$  are an orthonormal basis for  $\ell^2$ .

One might think that (4.3.3) follows from Fubini theorem for series. The following exercise shows that this is not the case. In fact, Fubini theorem requires that

$$(4.3.6) \quad \sum_{m,n} |y_m| |A_{m,n}| |x_n| < \infty,$$

which is a condition involving the absolute values of the matrix entries. The requirement (4.3.6) is that  $\tilde{A} = [|A_{m,n}|]_{m,n}$  defines a bounded operator on  $\ell^2$ .

PROPOSITION 4.3. (i) Let  $\{x_n\}_{n=1}^\infty$  be a sequence of positive numbers,  $x(n) \geq 0$ . Then,  $x \in \ell^2$  if and only if

$$(4.3.7) \quad \sum_n x(n)y(n) < \infty$$

for all positive  $y \in \ell^2$ .

(ii) Let  $A = [A(m,n)]_{m,n=1}^\infty$  be an infinite matrix with positive entries,  $A(m,n) \geq 0$ . Then,  $A$  defines a bounded operator on  $\ell^2$  if and only if

$$(4.3.8) \quad \sum_{m,n} x(m)A(m,n)y(n) < \infty$$

for all positive  $x, y \in \ell^2$ .

PROOF. (i) The only if direction follows from Cauchy-Schwarz. Suppose (4.3.7) holds. Then, it is easy to show that  $x(n) \rightarrow 0$  as  $n \rightarrow \infty$ . Suppose by contradiction that  $\sum_n x(n)^2 = \infty$ : we will find a sequence  $\epsilon(n) \searrow 0$  such that (a)  $\sum_n x(n)^2 \epsilon(n) = \infty$ , but (b)  $\sum_n x(n)^2 \epsilon(n)^2 < \infty$ . If we do this, the sequence  $y(n) = x(n)\epsilon(n)$  belongs to  $\ell^2$ , but fails to satisfy (4.3.7). Let's do the construction, assuming with no loss of generality that  $x(n) \leq 1$  for all  $n$ . Find an increasing sequence  $n_j$  ( $n_1 = 0$ ) such that

$$1 \leq \sum_{n=n_{j-1}+1}^{n_j} x(n)^2 \leq 2,$$

and let  $\epsilon(n) = \frac{1}{j}$  for  $n_{j-1} + 1 \leq n \leq n_j$ . Then, both (a) and (b) hold.

(ii) Again, the only if direction is elementary. Suppose now (4.3.8) holds. By (i),  $m \mapsto \sum_n A(m, n)y(n)$  and  $n \mapsto \sum_m A(m, n)x(m)$  are  $\ell^2$  sequences,

$$\sum_m \left( \sum_n A(m, n)y(n) \right)^2 < \infty, \quad \sum_n \left( \sum_m A(m, n)x(m) \right)^2 < \infty.$$

Thus, both  $A$  and  $A^*$  are everywhere defined, and so is  $A^*A$ , which is self-adjoint. By the Hellinger-Toeplitz theorem,  $A^*A$  is bounded, hence

$$\|Ax\|_{\ell^2}^2 = \langle x, A^*Ax \rangle \leq \|A^*A\|_{\mathcal{L}(\ell^2)} \|x\|_{\ell^2}^2,$$

which shows that  $A$  is bounded.  $\square$

**EXERCISE 4.3.** *Construct a bounded operator on  $\ell^2(\mathbb{N})$  such that (4.3.6) does not hold. Here is possible pathway.*

- (i) Consider the **discrete Fourier transform**  $\mathcal{F}_N : \ell^2(\{0, 1, \dots, N-1\})$ ,

$$(4.3.9) \quad (\mathcal{F}_N x)_m := \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{2\pi i m n / N} x_n.$$

That is, the matrix elements of  $\mathcal{F}_N$  are  $(\mathcal{F}_N)_{m,n} = \frac{e^{2\pi i m n / N}}{\sqrt{N}}$ .

Show that  $\mathcal{F}_N$  is a unitary operator.

- (ii) Consider the operator  $A_N$  having as entries the moduli of the entries of  $\mathcal{F}_N$ ,  $(A_N)_{m,n} = \frac{1}{\sqrt{N}}$ . Show that  $\|A_N\| = \sqrt{N}$ .
- (iii) Construct a unitary operator  $A$  on  $\ell^2(\mathbb{N})$  in block diagonal form, where the blocks are  $\mathcal{F}_N$  with increasing  $N$ . Show that  $A$  does not satisfy (4.3.6).
- (iv) Modify the construction to construct an operator  $A$  which is bounded and self-adjoint, yet (4.3.6) fails.

**4.3.2. Hilbert-Schmidt operators on Hilbert spaces.** Let  $A$  be a bounded operator on a separable Hilbert space  $H$ ,  $\{e_n\}$  be a orthonormal basis for  $H$ , and  $\{A_{m,n}\}_{m,n}$  be the matrix coefficients of  $A$  with respect to the basis. The *Hilbert-Schmidt norm* of  $A$  is

$$(4.3.10) \quad \|A\|_{HS}^2 := \sum_{m,n} |A_{m,n}|^2.$$

**PROPOSITION 4.4.** *The Hilbert-Schmidt norm of  $A$  is independent of the basis.*

PROOF. The statement is equivalent to show that, if  $U : \ell^2 \rightarrow \ell^2$  is unitary, then

$$\|U^*AU\|_{HS} = \|A\|_{HS},$$

which is what we do below. The matrix coefficients of  $U^*AU$  are

$$(U^*AU)_{m,n} = \sum_{i,j} U_{m,i}^* A_{i,j} U_{j,n} = \sum_{i,j} \overline{U_{i,m}} A_{i,j} U_{j,n}.$$

We can freely change the order of summation by corollary 4.4:

$$\begin{aligned} \|U^*AU\|_{HS}^2 &= \sum_{m,n,i,j,k,l} U_{i,m} \overline{A_{i,j}} U_{j,n} \overline{U_{l,m}} A_{l,k} U_{k,n} \\ &= \sum_{i,j,k,l} \left( \sum_m U_{i,m} \overline{U_{l,m}} \right) \overline{A_{i,j}} \left( \sum_n \overline{U_{j,n}} U_{k,n} \right) A_{l,k} \\ &= \sum_{i,j,k,l} \delta_{i,l} \overline{A_{i,j}} \delta_{j,k} A_{l,k} \\ &= \sum_{i,k} \overline{A_{i,k}} A_{i,k} \\ &= \|A\|_{HS}^2. \end{aligned}$$

□

A bounded operator  $A$  is in the *Hilbert-Schmidt class*,  $A \in HS(H)$ , if  $\|A\|_{HS} < \infty$ . The map  $A \mapsto [A_{m,n}]$  is unitary identification of  $HS$  with  $\ell^2(\mathbb{N}_* \times \mathbb{N}_*)$ . In particular,  $HS(H)$  has a Hilbert space structure. We also have

$$(4.3.11) \quad \|A\| \leq \|A\|_{HS}.$$

Also observe that membership in  $HS(H)$  only depends on the moduli  $|A_{m,n}|$  of the matrix coefficients (compare with exercise 4.3). We have the following important result.

**THEOREM 4.7.** *Hilbert-Schmidt operators are compact.*

PROOF. For  $M \geq 1$  fixed, let  $A^{(M)}$  be the infinite matrix with

$$A_{m,n}^{(M)} = \begin{cases} A_{m,n} & \text{if } 1 \leq m \leq M \\ 0 & \text{if } m > M. \end{cases}$$

Then  $A^{(M)}$  represents a finite rank operator, and

$$\|A - A^{(M)}\|^2 \leq \|A - A^{(M)}\|_{HS}^2 = \sum_{m>M} \sum_{n \geq 1} |A_{m,n}|^2 \rightarrow 0$$

as  $M \rightarrow \infty$  by dominated convergence. □

EXERCISE 4.4. (i) *Characterize self-adjoint Hilbert-Schmidt operators in terms of their spectrum.*

(ii) *Find a compact operator which is not Hilbert-Schmidt.*

The Hilbert-Schmidt class can be introduced in a coordinate-free way, but we will not do it here.

**4.3.3. The trace of a compact operator.** Let  $A \geq 0$  be a positive, compact operator on a separable Hilbert space  $H$ , and let  $\{e_n\}$  be an orthonormal basis of  $H$ . The *trace* of  $A$  is

$$(4.3.12) \quad \text{Trace}(A) = \sum_{n=1}^{\infty} \langle e_n, Ae_n \rangle \geq 0.$$

It readily verifies that the trace is a unitary invariant. We use matrix notation  $\langle e_n, Ae_n \rangle = A_{nn}$ . If  $U = [U_{ij}]$  is unitary, and  $B = U^*AU$ , then

$$\begin{aligned} \sum_n B_{nn} &= \sum_{njk} U_{nj}^* A_{jk} U_{kn} = \sum_{jk} A_{jk} \sum_n U_{nj}^* U_{kn} \\ &= \sum_{jk} A_{jk} \delta_{jk} \\ &= \sum_n A_{nn}. \end{aligned}$$

#### 4.4. Compact integral operators

The study of compact operators had its inception in the study of integral operators, especially those arising from the analysis of differential equations. In this section, we consider some important examples. We are interested in operators having the form

$$(4.4.1) \quad T_k f(x) = \int_Y k(x, y) f(y) d\nu(y),$$

where the input is  $f : Y \rightarrow \mathbb{C}$  is defined on a measure space  $(Y, \nu)$ ;  $k : X \times Y \rightarrow \mathbb{C}$  has some regularity properties; and the output is  $T_k f : X \rightarrow \mathbb{C}$ . The function  $k$  is the *kernel* of the integral operator  $K$ .

**4.4.1. Integral operators with continuous kernel.** We consider here the case where the kernel  $k$  is continuous. Statement and proof are on closed intervals, and the extension to a more general framework is left as an exercise in critical analysis of proofs.

THEOREM 4.8. Let  $k : [a, b] \times [c, d] \rightarrow \mathbb{C}$  be continuous, and define

$$T_k f(x) = \int_c^d k(x, y) f(y) dy.$$

Then,

- (i)  $T_k : C[a, b] \rightarrow C[a, b]$  is compact.
- (ii) For  $1 \leq p \leq \infty$ ,  $T_k : L^p[a, b] \rightarrow C[a, b]$  is compact.
- (iii) For  $1 \leq p < \infty$ ,  $T_k : L^p[a, b] \rightarrow L^p[a, b]$  is compact.

PROOF. The main tool here is Ascoli-Arzelà theorem and we use that kernel  $k$  is uniformly continuous on  $[a, b] \times [c, d]$  by compactness. We start with (i), proving at the same time that  $K$  maps  $C[a, b]$  into itself, and that it is a compact operator. For real  $\delta$  and  $f$  continuous with  $\|f\|_{L^\infty} \leq 1$ , in fact,

$$\begin{aligned} |T_k f(x + \delta) - T_k f(x)| &\leq \int_c^d |k(x + \delta, y) - k(x, y)| \cdot |f(y)| dy \\ &\leq \epsilon(d - c) \|f\|_{L^\infty} \leq \epsilon[d - c], \end{aligned}$$

for any  $\epsilon > 0$ , provided  $|\delta| \leq \eta(\epsilon)$  is such that  $|k(x + \delta, y) - k(x, y)| \leq \epsilon$  for all  $(x, y) \in [a, b] \times [c, d]$ . Hence, the family  $\{T_k f : \|f\|_{L^\infty} \leq 1\}$  is equicontinuous, and also uniformly bounded because

$$|T_k f(x)| \leq (d - c) \max_{(x, y) \in [a, b] \times [c, d]} |k(x, y)| \cdot \|f\|_{L^\infty}.$$

By the Ascoli-Arzelà theorem, the family has compact closure in  $C[a, b]$  (with respect to the norm topology).

The proof of (ii) is similar. Using Hölder inequality with  $1/p + 1/p' = 1$ , if  $\|f\|_{L^p} \leq 1$ ,

$$\begin{aligned} |T_k f(x + \delta) - T_k f(x)| &\leq \int_c^d |k(x + \delta, y) - k(x, y)| \cdot |f(y)| dy \\ &\leq \left( \int_c^d |k(x + \delta, y) - k(x, y)|^{p'} dy \right)^{1/p'} \|f\|_{L^p} \\ &\leq (d - c)^{1/p'} \epsilon \|f\|_{L^p}, \end{aligned}$$

by choosing  $\delta$  as in (i). We have equicontinuity, and uniform boundedness is proved similarly.

(iii) follows from (ii) and the fact that the identity  $I : C[a, b] \rightarrow L^1[a, b]$  is bounded, and the composition of a compact map and a bounded map is compact.  $\square$

**4.4.2. Hilbert-Schmidt integral operators.** Let  $k : [c, d] \times [a, b] \rightarrow \mathbb{C}$ ,  $k \in L^2([c, d] \times [a, b])$ . To each function  $f \in L^2[a, b]$  we associate, as before,

$$(4.4.2) \quad T_k f(x) = \int_a^b k(x, y) f(y) dy, \quad T_k f : [c, d] \rightarrow \mathbb{C}.$$

**THEOREM 4.9.** *The operator  $T_k : L^2[a, b] \rightarrow L^2[c, d]$  is a Hilbert-Schmidt operator.*

**PROOF.** Boundedness follows, as earlier, from Cauchy-Schwarz and Fubini:

$$\begin{aligned} \|T_k f\|_{L^2}^2 &= \int_c^d \left| \int_a^b k(x, y) f(y) dy \right|^2 dx \\ &\leq \int_c^d \left( \int_a^b |k(x, y)|^2 dy \cdot \int_a^b |f(z)|^2 dz \right) dx \\ &= \iint_{[c, d] \times [a, b]} |k(x, y)|^2 dx dy \cdot \int_a^b |f(z)|^2 dz. \end{aligned}$$

Let now  $\{e_n\}$  be a orthonormal basis for  $L^2[a, b]$  and  $\{\epsilon_m\}$  be one for  $L^2[c, d]$ , so that  $\{e_n(y)\epsilon_m(x)\}_{m,n}$  is a orthonormal basis for  $L^2([a, b] \times [c, d])$ , and we can expand

$$\begin{aligned} k(x, y) &= \sum_{m,n} k_{n,m} e_n(y) \epsilon_m(x), \\ \|k\|_{L^2([a, b] \times [c, d])}^2 &= \sum_{m,n} |k_{n,m}|^2 < \infty. \end{aligned}$$

Then, if  $f(y) = \sum_n f_n \cdot e_n(y)$ ,

$$K f(x) = \sum_m \epsilon_m(x) \sum_n k_{n,m} f_n.$$

With respect to the basis  $\{e_n\}$  and  $\{\epsilon_m\}$ , the operator  $T_k$  is represented by the Hilbert-Schmidt matrix  $[k_{n,m}]_{m,n=1}^\infty$ , hence it is a Hilbert-Schmidt operator.  $\square$

Statement and proof hold for more general measure spaces.

**4.4.3. The Volterra operator.** The *Volterra operator*  $V$  is defined as

$$(4.4.3) \quad V f(x) = \int_0^x f(t) dt,$$

where  $f \in L^1[0, 1]$ . It has a holomorphic avatar, with  $f : \mathbb{D} \rightarrow \mathbb{C}$ , which we will consider below. We will see some basic properties of  $V$  as operator on  $L^2[0, 1]$ .



THEOREM 4.10. (i) *The operator  $V$  is compact on  $L^2[0, 1]$  (in fact, it is a Hilbert-Schmidt integral operator). Its range is*

$$\{F \in AC[0, 1] : F(0) = 0 \text{ and } F' \in L^2[0, 1]\},$$

*where  $AC[0, 1]$  is the class of the absolutely continuous functions.*

(ii)  *$V$  does not have eigenvalues, and  $\sigma(V) = \{0\}$ .*

(iii) *The adjoint  $V^*$  of  $V$  is*

$$(4.4.4) \quad V^*g(x) = \int_x^1 g(y)dy,$$

*and*

$$(4.4.5) \quad V^*Vf(x) = \int_0^1 (1 - \max(x, y))f(y)dy.$$

PROOF. We show (i), (iii) first. Since we can write  $Vf(x) = \int_0^1 k(x, y)f(y)dy$  with  $k(x, y) = \chi(0 \leq y \leq x \leq 1)$ , and  $k \in L^2[0, 1]^2$ ,  $V$  is a Hilbert-Schmidt operator. Equations (4.4.4) and (4.4.5) follow by, respectively, observing that  $V^*$  has kernel  $k^*(x, y) = k(y, x) = \chi(0 \leq x \leq y \leq 1)$ , and computing the kernel

$$h(x, y) = \int_0^1 k^*(x, u)k(u, y)du.$$

About (ii),  $\lambda$  is an eigenvalue of  $V$  if and only if the equation

$$F = \lambda F'$$

has a nonzero solution  $F$  with  $F(0) = 0$ . You might worry about the fact that a priori derivatives exist just almost everywhere; but  $Vf = \lambda f$  implies that  $f$  is continuous, hence it is  $C^1$ . For  $\lambda = 0$ ,  $F = 0$ , hence  $\ker(F) = \{0\}$ . For the other  $\lambda$ 's,  $F(x) = Ce^{x/\lambda}$ , which satisfies  $F(0) = 0$  only if  $F = 0$ .  $\square$

The *holomorphic Volterra operator* is defined on holomorphic functions  $f : \mathbb{D} \rightarrow \mathbb{C}$ ,

$$(4.4.6) \quad Vf(z) = \int_0^z f(w)dw.$$

We use the same symbol  $V$ , although the operator is not the one we had considered earlier. We might consider it as an operator defined on the Taylor coefficients of  $f$ ,

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \implies Vf(z) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} z^{n+1},$$

which we might write as

$$(4.4.7) \quad \widehat{V}(\{a_n\}_{n=0}^\infty) = \left\{ \frac{a_{n-1}}{n} \right\}_{n=0}^\infty,$$

or  $\widehat{V} = \tau_1 \circ S$ , where  $S(\{a_n\}) = \{a_n/(n+1)\}$  is compact.

**THEOREM 4.11.** (i) *The operator  $V$  is a Hilbert-Schmidt (hence, compact) on  $H^2(\mathbb{D})$ .*

(ii)  *$V$  does not have eigenvalues, and  $\sigma(V) = \{0\}$ .*

(iii) *The adjoint  $V^*$  of  $V$  is*

$$(4.4.8) \quad \widehat{V}^*(\{b_n\}_{n=0}^\infty) = \left\{ \frac{b_{n+1}}{n+1} \right\}_{n=0}^\infty.$$

**EXERCISE 4.5.** *Prove theorem 4.11. Write explicit expressions for  $\widehat{V}^*\widehat{V}$ , and for  $\widehat{V}\widehat{V}^*$ .*

#### 4.5. Intermezzo: Sturm-Liouville theory

<sup>1</sup>. In this section we see an application of the spectral theory for self-adjoint compact operators to ODE's. The differential operator appearing in the ODE turns out to be the inverse of a (compact) integral operator. The eigenvalues of the latter will have a role in the analysis of the former.

**4.5.1. The Sturm-Liouville problem, spaces and operators.** In its general form, Sturm-Liouville theory deals with the differential equations

$$(4.5.1) \quad -(py')' + qy = \lambda\rho y,$$

with  $y \in C^2[a, b]$ ;  $\lambda \in \mathbb{C}$ ;  $p \in C^1[a, b]$  real valued with  $p(x) \neq 0$  in  $[a, b]$ ;  $q, \rho \in C[a, b]$  real valued with  $\rho(x) \neq 0$  in  $[a, b]$ ; and non-degenerate boundary conditions

$$(4.5.2) \quad \alpha_0 y(a) + \alpha_1 y'(a) = 0 = \beta_0 y(b) + \beta_1 y'(b),$$

where  $(\alpha_0, \alpha_1) \neq (0, 0) \neq (\beta_0, \beta_1)$  have real entries. Such conditions might be read as orthogonality relations in  $\mathbb{R}^2 \subset \mathbb{C}^2$ . The problem is finding information about *eigenvalues*: the  $\lambda$ 's such that (4.5.1), (4.5.2) has a nontrivial solution  $y$ .

---

<sup>1</sup>For this section I basically follow the exposition in [AN INTRODUCTION TO STURM-LIOUVILLE THEORY](#) by ERIK BÉDOS

We consider here the case  $p \equiv 1 \equiv \rho$ , which contains the main features of the theory:

$$(4.5.3) \quad \begin{aligned} \lambda y &= -y'' + qy, \\ 0 &= \alpha_0 y(a) + \alpha_1 y'(a) = \beta_0 y(b) + \beta_1 y'(b). \end{aligned}$$

Let  $F \subset C^2[a, b]$  be the space of the functions satisfying the boundary conditions (4.5.2), and  $D : F \rightarrow C[a, b]$  be the operator in (4.5.3),

$$(4.5.4) \quad Dy = -y'' + qy.$$

The Sturm-Liouville problem, then, consists in finding for which  $\lambda \in \mathbb{C}$  the equation

$$(4.5.5) \quad Dy = \lambda y, \text{ where } y \in F$$

has a solution  $y \neq 0$ .

From elementary results on linear ODEs we have the following.

LEMMA 4.2. *For  $\lambda \in \mathbb{C}$ , let  $S_\lambda = \{y \in C^2[a, b] : D(y) = \lambda y\}$ . Then,  $\dim_{\mathbb{C}}(S_\lambda) = 2$ .*

*In fact, for  $c \in [a, b]$  and  $\lambda \in \mathbb{C}$  fixed, the map  $L_{\lambda, c} : S_\lambda \rightarrow \mathbb{C}^2$ ,  $L_{\lambda, c}y = (y(c), y'(c))$ , is an isomorphism.*

Given  $y_1, y_2 \in C^2[a, b]$ , the Wronsky determinant  $W(y_1, y_2) \in C^1[a, b]$  is

$$W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}.$$

The following *Lagrange identity* is elementary.

LEMMA 4.3. *For  $f, g \in C^2[a, b]$ ,*

$$(4.5.6) \quad W(f, g)' = D(f)g - fD(g).$$

*If  $y_1, y_2$  are solution of  $D(y) = 0$ , then,  $W(y_1, y_2) = W$  is constant.*

Having two linearly independent solutions of  $D(y) = 0$ , which is (4.5.3) with  $\lambda = 0$ , using variation of parameters it is easy to write the general solution of

$$(4.5.7) \quad -y'' + qy = f,$$

where  $f \in C[a, b]$ . The solution can be written in a way which is symmetric with respect to the interval's endpoints.

LEMMA 4.4. *Let  $u, v$  be two linearly independent solutions of  $-y'' + qy = 0$ , and let  $f$  be continuous. Let  $W$  be the Wronski determinant of  $u, v$ . Then, a solutions of (4.5.7) can be written in the form*

$$(4.5.8) \quad y(x) = \int_a^b G(x, t) dt, \text{ where } G(x, t) = \frac{1}{W} \begin{cases} u(x)v(t) & \text{if } a \leq t \leq x \leq b, \\ u(t)v(x) & \text{if } a \leq x \leq t \leq b \end{cases}$$

*is symmetric,  $G(x, t) = G(t, x)$ , and continuous on  $[a, b] \times [a, b]$ .*

PROOF. We look for a solution of (4.5.7) of the form  $y = cu + dv$  for some functions  $c, d$  to be chosen. We have:

$$(4.5.9) \quad \begin{aligned} y &= cu + dv, \text{ so that} \\ y' &= cu' + dv' \text{ if } c'u + d'v = 0, \text{ and} \\ y'' &= cu'' + dv'' - f \text{ if } u'c' + v'd' = -f. \end{aligned}$$

It is immediately verified that, with these requirements,  $y$  in fact solves (4.5.7). What we need is

$$\begin{cases} c'u + d'v &= 0 \\ u'c' + v'd' &= -f \end{cases}, \text{ i.e. } \begin{pmatrix} u & v \\ u' & v' \end{pmatrix} \begin{pmatrix} c' \\ d' \end{pmatrix} = \begin{pmatrix} 0 \\ -f \end{pmatrix}.$$

Solving, we obtain

$$\begin{cases} c' &= \frac{fv}{W} \\ d' &= -\frac{fu}{W} \end{cases}$$

Usually, the next step would be writing the general solution in the form

$$\tilde{y} = \frac{1}{W} \int_{x_0}^x (u(x)v(t) - v(x)u(t))f(t)dt + \rho u(x) + \sigma v(x),$$

with real  $\rho, \sigma$ , but this expression is unnatural for our boundary conditions, since we have just one endpoint  $x_0$  to play with. We consider instead the

two "halves" of  $\tilde{y}$  separately, and we see what can be done with them.

$$\begin{aligned}
 A(x) &= \frac{u(x)}{W} \int_a^x v(t)f(t)dt \text{ and } B(x) = \frac{v(x)}{W} \int_b^x u(t)f(t)dt, \\
 A'(x) &= \frac{u'(x)}{W} \int_a^x v(t)f(t)dt + \frac{(uvf)(x)}{W} \text{ and} \\
 B'(x) &= \frac{v'(x)}{W} \int_b^x u(t)f(t)dt + \frac{(uvf)(x)}{W} \\
 A''(x) &= \frac{u''(x)}{W} \int_a^x v(t)f(t)dt + \frac{[u'vf + (uvf)'](x)}{W} \text{ and} \\
 (4.5.10) \quad B''(x) &= \frac{v''(x)}{W} \int_b^x u(t)f(t)dt + \frac{[v'uf + (uvf)'](x)}{W}.
 \end{aligned}$$

The function  $y$  in (4.5.8) is  $y(x) = A(x) - B(x)$ , and it satisfies

$$\begin{aligned}
 -y'' + qy &= \frac{(-u'' + qu)(x)}{W} \int_a^x v(t)f(t)dt - \frac{(-v'' + qv)(x)}{W} \int_b^x u(t)f(t)dt \\
 &\quad + \frac{(v'u - uv')(x)f(x)}{W} \\
 &= f,
 \end{aligned}$$

since  $v'u - uv' = W$ . □

#### 4.5.2. The solution to the Sturm-Liouville problem.

4.5.2.1. *The solution of the Sturm-Liouville problem when  $D$  is  $1 - 1$ .* Here we suppose that  $D : F \rightarrow C[a, b]$  is  $1 - 1$ . In general this is not the case, as the example  $D(y) = -y'' + y$  with conditions  $y(0) = y(\pi) = 0$  clearly illustrates:  $D(\sin x) = 0$  and  $\sin$  verifies the boundary conditions. Observe that the same operator is  $1 - 1$  on  $F$  if the conditions are  $y(0) = y(\pi/2) = 0$ . Injectivity of  $D$  on  $F$  means that the boundary problem

$$D(y) = 0, \quad \alpha_0 y(a) + \alpha_1 y'(a) = 0 = \beta_0 y(b) + \beta_1 y'(b)$$

only has the trivial solution  $y = 0$ .

We introduce spaces where either of the boundary conditions hold for the equation  $D(y) = \lambda y$  (but here we are interested in  $\lambda = 0$ ),

$$L_\lambda = \{y \in C^2[a, b] : D(y) = \lambda y \text{ and } \alpha_0 y(a) + \alpha_1 y'(a) = 0\},$$

$$R_\lambda = \{y \in C^2[a, b] : D(y) = \lambda y \text{ and } \beta_0 y(b) + \beta_1 y'(b) = 0\},$$

so that  $E_\lambda = L_\lambda \cap R_\lambda$  is the space of the solutions of (4.5.3). By lemma 4.2,  $\dim(L_\lambda) = \dim(R_\lambda) = 1$ , hence  $\dim(E_\lambda) \leq 1$ . We are here assuming that  $\dim(E_0) = 0$ , hence, there exist  $v \in L_0$  and  $u \in R_0$  which are linearly independent.

PROPOSITION 4.5. *Suppose  $D$  is  $1-1$  and let  $v \in L_0$ ,  $u \in R_0$ . Then, the solution of  $D(y) = f$  in lemma 4.4 is the only one satisfying the boundary conditions.*

PROOF. Uniqueness is clear, since  $D$  is  $1-1$  on  $F$ . We have to verify that the boundary conditions are satisfied.

$$\begin{aligned} A(a) &= 0 \text{ and } B(a) = \frac{v(a)}{W} \int_b^a u(t)f(t)dt, \\ A(b) &= \frac{u(b)}{W} \int_a^b v(t)f(t)dt \text{ and } B(b) = 0, \\ A'(a) &= \frac{(uvf)(a)}{W} \text{ and } B'(a) = \frac{v'(a)}{W} \int_b^a u(t)f(t)dt + \frac{(uvf)(a)}{W} \\ A'(b) &= \frac{u'(b)}{W} \int_a^b v(t)f(t)dt + \frac{(uvf)(b)}{W} \text{ and } B'(b) = \frac{(uvf)(b)}{W}, \end{aligned}$$

hence

$$\begin{aligned} \alpha_0 y(a) + \alpha_1 y'(a) &= -\frac{\alpha_0 v(a)}{W} \int_b^a u(t)f(t)dt - \frac{\alpha_1 v'(a)}{W} \int_b^a u(t)f(t)dt \\ &= 0 \end{aligned}$$

because  $v \in L_0$ , and similarly

$$\begin{aligned} \beta_0 y(b) + \beta_1 y'(b) &= \frac{\beta_0 u(b)}{W} \int_a^b v(t)f(t)dt + \frac{\beta_1 u'(b)}{W} \int_a^b v(t)f(t)dt \\ &= 0 \end{aligned}$$

because  $v \in R_0$ . □

The result can be rephrased in a more functional form.

COROLLARY 4.5. *If  $D : F \rightarrow C[a, b]$  is  $1-1$ , then it is a bijection, and its inverse  $T_G : C[a, b] \rightarrow F$  has the expression*

$$(4.5.11) \quad T_G(f) = \int_a^b G(x, t)f(t)dt,$$

where  $G : [a, b] \times [a, b] \rightarrow \mathbb{R}$  is that given in (4.5.8), and  $u \in R_0$ ,  $v \in L_0$ .

PROOF. We have verified above that  $D \circ T_G(f) = f$ . Since  $D$  is injective,  $T_G$  is surjective, and we also have  $T_G \circ D(y) = y$  for  $y \in F$ . □

Since the operator  $T_G$  extends to a compact, self-adjoint operator on  $L^2[a, b]$ , we are now in a position where the Hilbert-Schmidt theorem can be applied.

THEOREM 4.12. *Suppose the operator  $D$  is  $1-1$  on  $F$ , and consider the Sturm-Liouville problem in (4.5.5),*

$$(4.5.12) \quad D(y) = \lambda y, \quad y \in F.$$

*Then, the following hold.*

- (i) *The set  $\{\lambda_n\}$  of the  $\lambda$ 's for which there is a solution  $y \neq 0$  is countable, each eigenvalue is real and nonzero, and  $|\lambda_n| \rightarrow \infty$ .*
- (ii) *For each  $n$ , the eigenspace  $E_{\lambda_n} = \{y : (4.5.12) \text{ holds}\}$  is 1-dimensional.*
- (iii) *Let  $e_n$  be a unit vector in  $E_{\lambda_n}$ . Then,  $\{e_n\}$  is a orthonormal basis for  $L^2[a, b]$ .*

PROOF. The space  $F$  is dense in  $L^2[a, b]$  (see exercise below) and  $F = T_G(C[a, b])$ , hence,

$$L^2[a, b] = \overline{F} \subseteq \overline{T : G(L^2[a, b])} \subseteq L^2[a, b].$$

Thus,  $\overline{T : G(L^2[a, b])}$ , and corollary 4.3 implies that  $\ker(T_G) = L^2[a, b] \ominus \overline{T_G(L^2[a, b])} = 0$ .

We apply the spectral theorem to  $T_G$ , and find eigenvalues  $\{\mu_n\}$  with  $\mu_n \neq 0$ , and  $\mu_n \rightarrow 0$  as  $n \rightarrow \infty$ . Let  $f_n \in L^2[a, b]$  be an eigenfunction for  $\mu_n$ ,

$$f_n = 1/\mu_n T_G(f_n), \text{ hence, } f_n \in C[a, b], \text{ and } f_n = 1/\lambda_n^2 T_G(T_G f_n) \in F.$$

Applying  $D$  to both sides of the first equality,

$$D(f_n) = 1/\mu_n f_n.$$

Set  $\lambda = \lambda_n = 1/\mu_n$ . Then (4.5.12) holds for  $y = f_n$ . The eigenspace  $E_{\lambda_n}$  is one dimensional, as observed earlier.

We have to make sure that  $\{\lambda_n\}$  exhausts the eigenvalues of  $D$ . If there were another  $\lambda \neq 0$  for which  $D(y) = \lambda y$  has a nonzero solution in  $F$ , then  $y = \lambda T_G(y)$ , and  $1/\lambda$  would be an other eigenvalue for  $T_G$ , which is not possible.

Finally, by normalizing the  $f_n$ 's,  $e_n = f_n/\|f_n\|_{L^2}$ , we find a orthonormal basis of (real valued) eigenfunctions of  $D$ .  $\square$

4.5.2.2. *The solution of the Sturm-Liouville problem in general.* Suppose  $D(y) = -y'' + qy$  does not define a  $1-1$  operator on  $F$ . The results of the previous section, though, still holds if there is  $\mu \in \mathbb{R}$  such that  $D_\mu(y) = -y'' + (q - \mu)y$  is injective (we just have to shift the old  $\lambda$ 's by  $\mu$ ). Hence, we are done if we prove the following.

LEMMA 4.5. *There exists  $\mu \in \mathbb{R}$  which is not an eigenvalue of  $D$  on  $F$ .*

PROOF. If there are uncountably many eigenvalues, then there exists an uncountable orthonormal system in  $L^2[a, b]$ , which is separable. Contradiction.  $\square$

We have then the complete theorem.

THEOREM 4.13. *Consider the Sturm-Liouville problem in (4.5.5),*

$$(4.5.13) \quad D(y) = \lambda y, \quad y \in F.$$

*Then, the following hold.*

- (i) *The set  $\{\lambda_n\}$  of the  $\lambda$ 's for which there is a solution  $y \neq 0$  is countable, each eigenvalue is real and nonzero, and  $|\lambda_n| \rightarrow \infty$ .*
- (ii) *For each  $n$ , the eigenspace  $E_{\lambda_n} = \{y : (4.5.13) \text{ holds}\}$  is 1-dimensional.*
- (iii) *Let  $e_n$  be a unit vector in  $E_{\lambda_n}$ . Then,  $\{e_n\}$  is a orthonormal basis for  $L^2[a, b]$ .*



## CHAPTER 5

### Intermezzo: reproducing kernel Hilbert spaces

The *structure* of a complex Hilbert space  $(H, \langle \cdot, \cdot \rangle)$  includes a vector space  $H$  and an inner product  $\langle \cdot, \cdot \rangle$  on  $H$ , plus the requirement that  $H$  is complete with respect to the norm associated to the inner product (this is not especially demanding, since, if  $H$  is not complete, its completion is a Hilbert space). All orthonormal basis of  $H$  have the same cardinality, the (*Hilbert*) *dimension* of  $H$ ; and if  $H_1, H_2$  are Hilbert, there exists a bounded, linear bijection  $T : H_1 \rightarrow H_2$  if and only if  $H_1$  and  $H_2$  have the same dimension. Moreover,  $T$  can be taken to be *unitary*,

$$(5.0.1) \quad \langle Tx, Ty \rangle_{H_2} = \langle x, y \rangle_{H_1}.$$

The dimension, that is, is a *unitary invariant* of the Hilbert space, and it completely classifies the elements of the Hilbert class modulo unitary equivalence.

In mathematics and its applications, however, the elements of the Hilbert spaces we work with are typically *functions*  $f$ , which are defined on some space  $X$ , taking values at its *points*. The expression *function Hilbert space* is often used. The points of  $X$  (and, more generally, subsets of  $X$ ) provide extra structure. The structure, that is, has the form  $(H, \langle \cdot, \cdot \rangle, X)$ .

The points of  $X$  can enter the picture in many different ways. Consider the (unitarily equivalent) Hilbert spaces  $\ell^2(\mathbb{N})$  and  $L^2[0, 1]$ . The elements of the first space are functions  $n \mapsto f(n)$  defined on  $\mathbb{N}$ , while the elements of the second, we learn in Lebesgue theory, are *equivalence classes* of functions  $x \mapsto f(x)$ . In the second case, the difficulty is that  $\|f\|_{L^2} = 0$  does not imply that  $f = 0$ , but just that  $f(x) = 0$  *a.e.x*.

This fact is not just an artifact depending on our construction of  $L^2[0, 1]$ : it is a phenomenon which is intrinsic to  $L^2[0, 1]$  as a function Hilbert space: the value of a function  $f \in L^2[0, 1]$  at a point  $x \in [0, 1]$  can not be detected using the Hilbert structure. Suppose  $f : [0, 1] \rightarrow \mathbb{C}$  is continuous, so that  $f(1)$  can be defined unambiguously. In fact,  $\eta_1 : f \mapsto f(1)$  is a linear functional defined on  $C[0, 1]$ . Consider now the functions  $f_n(x) = x^n$ . We have that:

$$\eta_1(f_n) = 1, \text{ yet } \lim_{n \rightarrow \infty} \|f_n - 0\|_{L^2} = 0, \text{ and } \eta_1(0) = 0.$$

The *evaluation functional*  $\eta_1$ , that is, is not bounded in the  $L^2$ -norm on  $C[0, 1]$ , and in no way can be extended to a bounded, linear functional on  $L^2[0, 1]$ , no matter how  $L^2[0, 1]$  is thought of. (We might use Zorn's lemma to extend  $\eta_1$  to an *unbounded* linear functional on  $L^2[0, 1]$ , with no reasonable relation to evaluation at  $x = 1$  of functions which are not continuous).

EXERCISE 5.1. Let  $E \subseteq [0, 1]$  be measurable with positive measure, and define  $\eta_E : L^2[0, 1] \rightarrow \mathbb{C}$  be the mean value of  $f$  on  $E$ ,

$$\eta_E(f) = \frac{\int_E f dm}{m(E)},$$

which might be thought of as an "average evaluation" on the set  $E$ . Show that  $\|\eta_E\|_{H^*} = m(E)^{-1/2} \rightarrow \infty$  as  $m(E) \rightarrow 0$ .

In  $\ell^2(\mathbb{N})$  we have a completely different story, since, for any positive integer  $m$ ,  $|\eta_m(f)| = |f(m)| \leq \|f\|_{\ell^2}$ : evaluating  $f$  at the point  $m$  is a functional having finite energy, so to speak.

These examples are rather trivial, and do not provide sufficient motivation to develop a theory. In the 1930' Hilbert spaces were still a novelty, they provided a unified approach to a variety of problems, and unified in a simple way a great deal of apparently distant mathematical objects, phenomena, and techniques. Stefan Bergman, while developing ways to efficiently approximate conformal mappings, introduced a Hilbert space of holomorphic functions, which nowadays goes under his name: the *Bergman space*. Soon after, Nachman Aronszajn framed Bergman's ideas in a general theory, presented in [Theory of reproducing kernels](#) (1950), which is still the one of best sources to start studying the subject.

From its inception, RKHS theory followed several paths: holomorphic function theory in one, then several variables; and statistics. At the end of the past century there was an explosion of new interest in both directions: in statistics in view of applications to machine learning, and in mathematical analysis after several breakthroughs concerning specific problems, and the realization that underlying there were structural properties of general families of RHHS's (Pick spaces, De Branges spaces...).

The renewed interest, and the speed with which results are nowadays communicated, opened new areas, perspectives, and cross-contamination between different fields.

### 5.1. Equivalent definitions of RKHS

In this subsection we see different, equivalent definitions of a reproducing kernel Hilbert space.

Let  $X$  be a set and  $H$  a Hilbert space of functions  $f : X \rightarrow \mathbb{C}$ . We say that  $H$  has *bounded point evaluation* if for each  $x$  in  $X$  there the evaluation functional  $\eta_x : f \mapsto f(x)$  is bounded on  $H$ .

By Riesz Representation Theorem, this holds if and only if for each  $x$  there is a (unique) vector  $k_x$  in  $H$  such that

$$f(x) = \langle k_x, f \rangle.$$

Moreover, choosing  $f = k_x$  we see that  $\|\eta_x\|_{H^*} = \|k_x\|_H$ .

**An example where point evaluation is not bounded** We use Zorn's lemma. Consider in  $\ell^2(\mathbb{N})$  the orthonormal basis  $\{\delta_n : n \in \mathbb{N}\}$  given by the Dirac's deltas, and use Zorn's lemma to extend it to a Hamel (algebraic) basis  $\{v_\alpha\}_{\alpha \in I}$ . Define then the non-bounded, linear functional  $\lambda : \ell^2 \rightarrow \mathbb{C}$  by its action on the Hamel basis,

$$\lambda(v_\alpha) = \begin{cases} n & \text{if } v_\alpha = e_n \text{ for some } n, \\ 0 & \text{otherwise.} \end{cases}$$

Let then  $X = \mathbb{N} \cup \{\lambda\}$  and to each  $h \in \ell^2(\mathbb{N})$  associate  $f_h : X \rightarrow \mathbb{C}$ ,

$$f_h(x) = \begin{cases} \langle \delta_n, h \rangle_{\ell^2} & \text{if } x = n, \\ \lambda(h) & \text{if } x = \lambda. \end{cases}$$

The space  $H$  of such functions is unitarily equivalent to  $\ell^2(\mathbb{N})$  with respect to the inner product  $\langle f_h, f_k \rangle_H := \langle h, k \rangle_{\ell^2(\mathbb{N})}$ , and evaluation at  $\lambda$ ,  $f_h(\lambda) = \lambda(h)$  is not bounded.

The functions  $\{k_x\}_{x \in X}$  are the *kernel functions* for  $H$ , which is then called a *Reproducing kernel Hilbert space*, with kernel  $k : X \times X \rightarrow \mathbb{C}$  defined by

$$k(x, y) := \langle k_x, k_y \rangle = k_y(x).$$

**PROPOSITION 5.1.** *The following properties hold.*

- (i)  $k(x, x) = \|k_x\|^2 \geq 0$ , with equality if and only if  $f(x) = 0$  for all  $f$  in  $H$  (and such points can be removed).
- (ii)  $k_x(y) = k(y, x) = \overline{k(x, y)} = \overline{k_y(x)}$ .
- (iii) If  $x_1, \dots, x_n \in X$  and  $c_1, \dots, c_n \in \mathbb{C}$ , then

$$\sum_{i,j=1}^n c_i \overline{c_j} k(x_j, x_i) = \left\| \sum_{i=1}^n c_i k_{x_i} \right\|^2 \geq 0.$$

*Property (iii), by definition, say that  $k : X \times X \rightarrow \mathbb{C}$  is a positive semi-definite function.*

**EXERCISE 5.2.** *Prove the proposition.*

Property (iii) obviously implies that  $k(x, x) \geq 0$ , hence that  $\overline{k(x_1, x_2)} = k(x_2, x_1)$ . In fact, setting  $c_1 = 1$  and  $c_2 = i$ :

$$0 \leq k(x_1, x_1) + ik(x_1, x_2) - ik(x_2, x_1) + k(x_2, x_2)$$

which implies that  $\operatorname{Re}(k(x_1, x_2)) = \operatorname{Re}(k(x_2, x_1))$ , and setting  $c_1 = c_2 = 1$ ,

$$0 \leq k(x_1, x_1) + k(x_1, x_2) + k(x_2, x_1) + k(x_2, x_2)$$

which implies that  $\operatorname{Im}(k(x_1, x_2)) + \operatorname{Im}(k(x_2, x_1)) = 0$ . If in the definition we had just considered real scalars  $c_1, \dots, c_n$ , as it is sometimes done in the real theory, dealing with real kernels  $k$ , then the condition  $k(x, y) = k(y, x)$  does not follow from the real version of (iii), as illustrated by the matrix

$$K = \begin{pmatrix} 1 & 1/2 \\ 0 & 1 \end{pmatrix},$$

which can be seen as a non-symmetric kernel on a space  $X$  with two points, having an associated quadratic form which is positive definite.

It is noteworthy that to each positive semi-definite function  $k : X \times X \rightarrow \mathbb{C}$  we can associate a RKHS of which  $k$  is the reproducing kernel.

**THEOREM 5.1** (Moore–Aronszajn theorem). *Let  $k : X \times X \rightarrow \mathbb{C}$  be a positive semi-definite function on some set  $X$ , and define on  $\operatorname{span}(k_x : x \in X)$  by*

$$(5.1.1) \quad \left\langle \sum_{i=1}^m a_i k_{x_i} \left| \sum_{j=1}^n b_j k_{y_j} \right. \right\rangle := \sum_{i=1}^m \sum_{j=1}^n \bar{a}_i b_j k(x_i, y_j).$$

*Then, the completion  $H$  of  $\operatorname{span}(k_x : x \in X)$  with respect to the inner product is a reproducing kernel Hilbert space having kernel  $k$ .*

**PROOF. Step I.** Since  $k$  is positive semi-definite,  $\langle \cdot, \cdot \rangle$  defines an inner product on the function space  $\operatorname{span}(k_x : x \in X)$  provided we show that, for  $f \in \operatorname{span}(k_x : x \in X)$ ,  $\langle f, f \rangle = 0$  implies that  $f(x) = 0$  for all  $x$  in  $X$ . Let  $f = \sum_{i=1}^n a_i k_{x_i}$  and suppose that  $\langle f, f \rangle = 0$ . If  $x = x_{n+1}$  is any other point in  $X$ , we let  $a_{n+1} = 0$ .

The matrix  $K = [k(x_i, x_j)]_{i,j=0}^{n+1}$  is Hermitian and positive, then it can be written as  $K = U^* \Delta U$ , where  $\Delta$  is the diagonal matrix having diagonal entries

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0 = \lambda_{m+1} = \dots = \lambda_{n+1}.$$

With  $a \in \mathbb{C}^{n+1}$  and  $a^* = \bar{a}^t$ , We are assuming that

$$0 = a^* K a = (Ua)^* \Delta (Ua) = b^* \Delta b = \sum_{l=1}^m \lambda_l |b_l|^2,$$

which implies that  $b_1 = \dots = b_m = 0$ , hence  $\Delta b = 0$ , so

$$0 = U^* \Delta b = K a.$$

We have, then, for  $j = 1, \dots, x_{n+1}$ :

$$\begin{aligned} f(x_j) &= \sum_{i=1}^n a_i k_{x_i}(x_j) = \sum_{i=1}^n a_i k(x_j, x_i) = \sum_{i=1}^n a_i K(j, i) \\ &= (K a)_j = 0. \end{aligned}$$

Hence,  $f = 0$  as a function.

**Step II.** The completion  $H$  of  $\text{span}(k_x : x \in X)$  under the norm induced by the inner product is a Hilbert space, having as generators the functions  $k_x$ . Each element  $f$  of it can be interpreted as a function on  $X$ ,

$$f(x) := \langle k_x, f \rangle.$$

The function  $k_x \in H$  is then reproducing for the value of  $f$  at  $x$ , and, as we saw above, the corresponding reproducing kernel is  $\langle k_x, k_y \rangle = k(x, y)$ , as stated.  $\square$

The theorem above is somehow deceptive. Its proof, in fact, gives little hint on how to find a "natural" expression for the inner product associated with a positive semi-definite function. We will see below that  $k(z, w) = \frac{1}{1-\bar{w}z}$  defines a positive semi-definite function, hence a kernel, on the unit disc  $\mathbb{D}$  of the complex plane, but at this moment is not clear how the inner product really looks like. We have, however, a characterization of reproducing Hilbert spaces based on orthonormal basis, which is sometimes useful in this respect.

**THEOREM 5.2.** *Let  $H$  be a reproducing kernel Hilbert space on some set  $X$ , and let  $\{e_\alpha\}_{\alpha \in I}$  be a orthonormal basis of it. Then, for  $y \in X$ , the series*

$$(5.1.2) \quad k_y := \sum_{\alpha \in I} \overline{e_\alpha(y)} e_\alpha$$

*converges in  $H$ , and it converges to the reproducing function at  $y$ .*

*Viceversa, suppose  $H$  is a Hilbert function space on  $X$  and that the series (5.1.2) converges in  $H$ . Then  $k(x, y) = k_y(x)$  reproduces the values of functions in  $\text{span}(e_\beta : \beta \in I)$ :*

$$(5.1.3) \quad k(x, y) = \sum_{\alpha \in I} \overline{e_\alpha(y)} e_\alpha(x), \quad \langle k_x, e_\alpha \rangle = e_\alpha(x) \text{ for } \alpha \in I.$$

Moreover, the series in (5.1.3) converges pointwise for all  $x, y \in X$ . After associating to each function  $f$  in  $H$  its modification  $\tilde{f}(x) := \langle k_x, f \rangle$ , the space  $\tilde{H}$  of such  $\tilde{f}$ 's is a reproducing Hilbert space with kernel  $k = k(x, y)$ .

The reproducing kernel is uniquely determined by the structure of RKHS, hence it does not depend on the orthonormal basis under consideration. Also, we have further evidence that  $L^2[0, 1]$  is not a RKHS. The series

$$y \mapsto \sum_{n=-\infty}^{+\infty} e^{-2\pi i n y} e^{2\pi i n x}$$

is not, in fact, convergent in  $L^2[0, 1]$ .

PROOF. Observe that, by basic properties of the orthonormal basis, the series on the right of (5.1.3) converges if and only if

$$\sum_{\alpha \in I} e_{\alpha}(y)^2 \text{ converges in } \mathbb{R}_+,$$

or, equivalently, if for all  $x, y \in X$

$$\sum_{x, y} \overline{e_{\alpha}(x)} e_{\alpha}(y) \text{ converges absolutely in } \mathbb{C}.$$

Suppose  $H$  is a RKHS with kernel  $k$ . Then,

$$k_y = \sum_{\alpha \in I} \langle e_{\alpha}, k_y \rangle e_{\alpha} = \sum_{\alpha \in I} \overline{\langle k_y, e_{\alpha} \rangle} e_{\alpha} = \sum_{\alpha \in I} \overline{e_{\alpha}(y)} e_{\alpha}$$

converges in  $H$ .

Suppose, viceversa, that the series in (5.1.2) converges in  $H$ . Then,

$$(5.1.4) \quad \langle k_y, e_{\beta} \rangle = \left\langle \sum_{\alpha \in I} \overline{e_{\alpha}(y)} e_{\alpha} \mid e_{\beta} \right\rangle = \sum_{\alpha \in I} e_{\alpha}(y) \langle e_{\alpha} \mid e_{\beta} \rangle = e_{\beta}(y).$$

Thus,  $k_y$  is the reproducing function for the basis elements, hence for all elements in  $\text{span}(e_{\beta} : \beta \in I)$ . Also, the function  $k(x, y) = \langle k_x, k_y \rangle$  is positive semi-definite:

$$\begin{aligned} \left\langle \sum_{i=1}^n c_i k_{x_i} \mid \sum_{j=1}^n c_j k_{x_j} \right\rangle &= \sum_{i,j=1}^n \sum_{\alpha \in I} \overline{c_i c_j} e_{\alpha}(x_i) \overline{e_{\alpha}(x_j)} \\ &= \sum_{\alpha \in I} \left| \sum_{i=1}^n c_i e_{\alpha}(x_i) \right|^2 \\ &\geq 0. \end{aligned}$$

By Moore-Aronszajn theorem,  $k$  is the reproducing kernel of  $\tilde{H}$  and  $\tilde{e}_{\alpha} = e_{\alpha}$  by (5.1.4).  $\square$

Finally, we can see reproducing kernel Hilbert spaces "from the inside" of a given, possibly abstract Hilbert space. This viewpoint is especially useful in applications to machine learning. Observe that to a RKHS  $H$  on  $X$  with reproducing functions  $\{k_x : x \in X\}$  we can associate the map  $x \mapsto k_x$  (the *feature map* of machine learning) which associates to each point in  $X$  a vector in  $H$ . Below, we see that the viceversa holds as well.

**THEOREM 5.3.** *Let  $H$  be a Hilbert space, and let  $\{v_x : x \in X\}$  be a family of vectors in  $H$ , indexed by some set  $X$ , such that  $\text{span}(v_x : x \in X)$  is dense in  $H$ . To each  $h$  in  $H$  associate  $f_h : x \mapsto \langle v_x, h \rangle_H$ , let  $H_X$  be the vector space of such functions, and set  $\langle f_h, f_k \rangle_{H_X} := \langle h, k \rangle_H$ . Then,  $H_X$  is a RKHS having  $\{f_{v_x} : x \in X\}$  as kernel functions.*

**PROOF.** By definition,  $f_{v_x}(y) = \langle v_y, v_x \rangle_H$ , and  $(x, y) \mapsto k(x, y) := \langle v_x, v_y \rangle_H$  is positive definite on  $X$ , as is easily verified. Still by definition we have that

$$f_h(x) = \langle v_x, h \rangle_H = \langle f_{v_x}, f_h \rangle_{H_X},$$

hence,  $f_{v_x}$  is the reproducing function for  $x$  in  $H_X$ , as wished.  $\square$

At the end of the day, we have several, equivalent viewpoints on RKHS's.

- (i) Hilbert function spaces with bounded point evaluation.
- (ii) Hilbert function spaces with a reproducing kernel.
- (iii) Positive semi-definite functions on sets.
- (iv) Hilbert function spaces where the values at each point of an orthonormal basis' elements are square integrable.
- (v) Families of vectors in a Hilbert space.

## 5.2. Some RKHS's

**5.2.1. The Hardy space  $H^2$  and its real counterpart.** We denote by  $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$  the unit disc in the complex plane. Let  $H^2 = H^2(\mathbb{D})$  be the space of the power series  $f(z) = \sum_{n=0}^{\infty} a_n z^n$ ,  $z \in \mathbb{C}$ , for which the norm

$$\left\| \sum_{n=0}^{\infty} a_n z^n \right\|_{H^2}^2 = \|\{a_n\}_{n=0}^{\infty}\|_{\ell^2}^2 = \sum_{n=0}^{\infty} |a_n|^2$$

is finite. In analogy with Fourier theory, in fact consistently with it, we write  $a_n = \widehat{f}(n)$ .

**LEMMA 5.1.** *If  $f \in H^2$ , then the series defining  $f$  converges on  $\mathbb{D}$ .*

PROOF. If  $m, j \geq 1$ ,

$$\begin{aligned} \left| \sum_{n=m+1}^{m+j} a_n z^n \right|^2 &\leq \sum_{n=m+1}^{m+j} |a_n|^2 \cdot \sum_{n=m+1}^{m+j} |z|^{2n} \\ &\leq \sum_{n=m+1}^{m+j} |a_n|^2 \cdot \frac{|z|^{2(m+1)}}{1 - |z|^2} \end{aligned}$$

if  $|z| < 1$ , hence Cauchy criterion is verified.  $\square$

By the theory of series in a complex variable,  $f(z)$  defines a holomorphic function in  $\mathbb{D}$ .

PROPOSITION 5.2. *The reproducing kernel of the Hardy space is*

$$k_z(w) = \frac{1}{1 - \bar{z}w}.$$

*In particular, the best estimate for the value of a Hardy function at  $z \in \mathbb{D}$  is:*

$$|f(z)| \leq \frac{\|f\|_{H^2}}{(1 - |z|^2)^{1/2}}.$$

PROOF.

$$\begin{aligned} \sum_{n=0}^{\infty} \widehat{f}(n) z^n &= f(z) \\ &= \langle f | k_z \rangle_{H^2} = \left\langle \sum_{n=0}^{\infty} \widehat{f}(n) w^n \middle| \sum_{n=0}^{\infty} \widehat{k}_z(n) w^n \right\rangle_{H^2} \\ &= \sum_{n=0}^{\infty} \widehat{f}(n) \overline{\widehat{k}_z(n)}, \end{aligned}$$

which is equivalent to  $\overline{\widehat{k}_z(n)} = z^n$ , hence,

$$k_z(w) = \sum_{n=0}^{\infty} \widehat{k}_z(n) w^n = \sum_{n=0}^{\infty} \bar{z}^n w^n = \frac{1}{1 - \bar{z}w}.$$

$\square$

An interesting variation on  $H^2$  is  $h^2$ , the *real (or harmonic) Hardy space*: the space populated by the series  $u(re^{it}) = \sum_{n=-\infty}^{+\infty} a_n r^{|n|} e^{int}$ ,  $0 \leq r < 1$ ,



such that  $\|u\|_{h^2}^2 = \sum_{n=-\infty}^{+\infty} |a_n|^2 < \infty$ . Another way to write a function in  $h^2$  is

$$(5.2.1) \quad u(z) = \sum_{n=0}^{\infty} a_n z^n + \sum_{n=1}^{\infty} a_{-n} \bar{z}^n.$$

We argue as above in order to find the reproducing kernel:

$$u(w) = \sum_{n=0}^{\infty} a_n \bar{w}^n + \sum_{n=1}^{\infty} a_{-n} \bar{w}^n = \langle k_w, u \rangle_{h^2},$$

where

$$\begin{aligned} k_w(z) &= \sum_{n=0}^{\infty} z^n \bar{w}^n + \sum_{n=1}^{\infty} a_{-n} \bar{z}^n w^n \\ &= \frac{1}{1 - \bar{w}z} + \frac{1}{1 - w\bar{z}} - 1 \\ &= \frac{1 - |z|^2 |w|^2}{|1 - \bar{w}z|^2}. \end{aligned}$$

The reason why  $h^2$  is also called *harmonic* Hardy space is that its elements are functions which are harmonic on  $\mathbb{D}$ . Recall that a function  $u : \Omega \rightarrow \mathbb{R}$  defined on an open set  $\Omega \subseteq \mathbb{C}$  is *harmonic* if it is  $C^2(\Omega)$  (it has continuous, partial second order derivatives), and it satisfies the *Laplace equation*,

$$(5.2.2) \quad 0 = \Delta u := \partial_{xx} u + \partial_{yy} u.$$

A complex valued function is harmonic if its real and imaginary parts are. By Cauchy-Riemann equations, for instance, a holomorphic function is harmonic. In particular,  $z \mapsto z^n = u + iv$  is harmonic on  $\mathbb{C}$ , and for the same reason  $z \mapsto \bar{z}^n = u - iv$  is harmonic, too. Any function  $u$  of the form (5.2.1), if the series converges in  $\mathbb{D}$  (hence, by the theory of power series, it converges absolutely and uniformly on any compact in  $\mathbb{D}$ ), is harmonic, because we can indefinitely differentiate the series under the sum.

Let's verify this for the derivative w.r.t.  $x$ . We have that  $\partial_x(z^n) = \partial_x(x+iy)^n = n z^{n-1}$ , and  $\partial_x(\bar{z}^n) = n \bar{z}^{n-1}$ . The power series of the derivatives,

$$w(z) := \sum_{n=1}^{\infty} n a_n z^{n-1} + \sum_{n=1}^{\infty} n a_{-n} \bar{z}^{n-1},$$

has the same radius of convergence as that representing  $u$ , hence, by a well know results of advanced calculus,  $w(z) = \partial_x u(z)$ . The reasoning is the same with  $y$  instead of  $x$ , and

$$\partial_y u(z) = i \sum_{n=1}^{\infty} n a_n z^{n-1} - i \sum_{n=1}^{\infty} n a_{-n} \bar{z}^{n-1}.$$

Now,  $\partial_x u$  and  $\partial_y u$  satisfy the same hypothesis of  $u$ , and we can take further partial derivatives.

With this at hands,

$$\Delta u(z) = \sum_{n=0}^{\infty} a_n \Delta(z^n) + \sum_{n=1}^{\infty} a_{-n} \Delta(\bar{z}^n) = 0.$$

We introduced the Hardy space from the viewpoint of the "Fourier side" (which engineers consider the *time domain*:  $\mathbb{N}$  for the holomorphic, and  $\mathbb{Z}$  for the harmonic case), imposing a size condition on the coefficients of the power series. The *raison d'être* of representing the sequence  $\{a_n\}$  by means of the holomorphic (or harmonic) function  $f(z)$  of which it provides the Taylor coefficients at the origin (its *generating function*, as it is named in combinatorics, or its  *$z$ -transform*, as it is called in engineering, where the set of the admissible  $z$ 's is called the *frequency domain*), is that many algebraic and, more important, quantitative properties of the sequence, and of the operators acting on it, emerge in the complex domain, as we will see later on. To start, we give a characterization of the  $H^2$  norm in "frequency space".

We denote by  $\mathbb{T} : \partial\mathbb{D} = \{e^{it} : t \in [0, 2\pi)\}$  the *torus*, i.e, the unit circle in the complex plane. Sometimes we identify  $t \leftrightarrow e^{it}$ , and  $[0, 2\pi) \leftrightarrow \mathbb{T}$ .

**THEOREM 5.4.** *Let  $f : \mathbb{D} \rightarrow \mathbb{C}$  be a holomorphic function. Then,*

$$(5.2.3) \quad \|f\|_{H^2}^2 = \sup_{0 \leq r < 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^2 dt = \lim_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^2 dt.$$

Moreover, if  $f_*(e^{it}) = \sum_{n=0}^{\infty} a_n e^{int}$ , where  $a_n$  is the  $n^{\text{th}}$  coefficients in the power expansion of  $f$  at the origin, and  $f_r(e^{it}) = f(re^{it})$ , then

$$(5.2.4) \quad \lim_{r \rightarrow 1} \|f_r - f_*\|_{L^2(0, 2\pi)} = 0.$$

In particular,

$$(5.2.5) \quad \|f\|_{H^2}^2 = \frac{1}{2\pi} \int_0^{2\pi} |f_*(e^{it})|^2 dt.$$

In (5.2.4) we have that the boundary values  $f_*$  of  $f(r \cdot)$  as  $r \rightarrow 1$  exist in the  $L^2$  sense. A cornerstone of  $H^2$ -theory is that  $\lim_{r \rightarrow 1} f(re^{it})$  exists for a.e.  $t \in [0, 2\pi)$ . We are not concerned, here, with these "hard analysis" developments of the theory, which are, however, important also for the more functional analytic side.

PROOF. For each  $0 \leq r < 1$ ,  $f_r$  is continuous on  $\mathbb{T}$ , and by Plancherel isometry we have that

$$\frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^2 dt = \sum_{n=0}^{\infty} |a_n|^2 r^{2n},$$

so that (5.2.3) follows by dominated convergence, and  $f \in H^2(\mathbb{D})$  if and only if  $f_* \in L^2(\mathbb{T})$ . Still by Plancherel,

$$\|f_r - f\|_{L^2(0,2\pi)}^2 = \sum_{n=0}^{\infty} |a_n|^2 (1 - r^n)^2 \rightarrow 0, \text{ as } r \rightarrow 1$$

by dominated convergence, which shows (5.2.4).  $\square$

For the harmonic Hardy space we have a similar statement, which immediately follows from the holomorphic one provided you know the following.

**THEOREM 5.5.** *Let  $\Omega$  be a simply connected set in  $\mathbb{C}$ , and  $u : \Omega \rightarrow \mathbb{R}$  be harmonic. Then, there exists  $v$  harmonic in  $\Omega$  such that  $f = u + iv$  is holomorphic in  $\Omega$ . Moreover, there is a unique such  $v$  with  $v(z_0) = 0$ , where  $z_0$  is any distinguished point in  $\Omega$ .*

*In particular,  $u = \frac{f+\bar{f}}{2}$  can be written as the sum of a holomorphic and an anti-holomorphic function.*

Applying the theorem to real and negative parts of  $u$ , and using known properties of holomorphic functions, we have the following.

**COROLLARY 5.1.** *Let  $\Omega$  be a simply connected set in  $\mathbb{C}$ . The functions  $u : \Omega \rightarrow \mathbb{C}$  which are harmonic in  $\Omega$  are exactly those having the form  $u(z) = f(z) + g(\bar{z})$ , with  $f, g$  holomorphic in  $\Omega$ . Moreover,  $f$  and  $g$  are uniquely determined once we require  $g(z_0) = 0$  at a distinguished point  $z_0$  in  $\Omega$ .*

*If  $\Omega = \mathbb{D}$ , then*

$$(5.2.6) \quad u(re^{it}) = \sum_{n=-\infty}^{+\infty} a_n r^{|n|} e^{int},$$

*where both the series with negative  $n$ 's and positive  $n$ 's converge in  $\mathbb{D}$ . The coefficients themselves are given by*

$$a_n = \frac{r^{-|n|}}{2\pi} \int_0^{2\pi} u(re^{it}) e^{-int} dt,$$

*and are independent of  $0 \leq r < 1$ .*

PROOF OF THE THEOREM. We want  $u + iv$  to satisfy Cauchy-Riemann equations, i.e.  $\partial_x v = -\partial_y u$  and  $\partial_y v = \partial_x u$ , or

$$(5.2.7) \quad dv = -\partial_y u dx + \partial_x u dy.$$

Now, the 1-form on the right hand side of (5.2.7) is closed because  $\partial_y(-\partial_y u) = \partial_x(\partial_x u)$  by harmonicity of  $u$ . Poincaré theorem ensures that the form is exact, hence

$$v(z) := \int_{\gamma(z_0, z)} (-\partial_y u dx + \partial_x u dy)$$

is well defined, independently of the chosen integration path  $\gamma(z_0, z)$ , as long as it starts in  $z_0$  and ends in  $z$ .

Moreover,  $u + iv$  satisfies the Cauchy-Riemann equation, since  $\nabla v = (-\partial_y u, \partial_x u)$ , hence  $u + iv$  is holomorphic.

Uniqueness holds because, in a connected domain, a function is determined by its gradient, modulo an additive constant.  $\square$

We have then the counterpart of theorem 5.4.

THEOREM 5.6. *Let  $u : \mathbb{D} \rightarrow \mathbb{C}$  be a harmonic function, with series expansion like in (5.2.6). Then,*

$$(5.2.8) \quad \|u\|_{H^2}^2 = \sup_{0 \leq r < 1} \frac{1}{2\pi} \int_0^{2\pi} |u(re^{it})|^2 dt = \lim_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} |u(re^{it})|^2 dt.$$

Moreover, if

$$u_*(e^{it}) = \sum_{n=-\infty}^{\infty} a_n e^{int}$$

and  $u_r(e^{it}) = u(re^{it})$ , then

$$(5.2.9) \quad \lim_{r \rightarrow 1} \|u_r - u_*\|_{L^2(0, 2\pi)} = 0.$$

In particular,

$$(5.2.10) \quad \|h\|_{H^2}^2 = \frac{1}{2\pi} \int_0^{2\pi} |u_*(e^{it})|^2 dt.$$

The proof of the theorem easily follows from theorem 5.5, corollary 5.1, and theorem 5.2.3.

**5.2.2. The Dirichlet space  $\mathcal{D}(\mathbb{D})$ .** Consider the subspace  $\mathcal{D} = \mathcal{D}(\mathbb{D})$  of the Hardy space  $H^2(\mathbb{D})$  having as elements functions  $f(z) = \sum_{n=0}^{\infty} a_n z^n$  such that

$$(5.2.11) \quad \|f\|_{\mathcal{D}}^2 = \sum_{n=0}^{\infty} (n+1)|a_n|^2 = \|f\|_{H^2}^2 + [f] < \infty.$$

The expression  $[f]^{1/2}$ , which is the "largest part" of the norm, is just a seminorm, since it vanishes on constant functions. We will see below that is is a *conformal invariant*.

**EXERCISE 5.3.** Show that  $\mathcal{D}$  has reproducing kernel  $k(z, w) = \frac{1}{\bar{w}z} \log \frac{1}{1-\bar{w}z}$ , where  $\log$  is the determination of the complex logarithm which takes real values on the positive axis.

The study of the Dirichlet space started in 1940 with Arne Beurling's<sup>1</sup> [Ensembles exceptionnels](#).

**LEMMA 5.2.** Let  $f$  be a function which is holomorphic in the unit disc. Then,

$$(5.2.12) \quad [f] = \frac{1}{\pi} \int_{\mathbb{D}} |f'(z)|^2 dx dy, \quad z = x + iy.$$

**PROOF.** For  $f(z) = \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} a_n r^n e^{int}$ , using the orthogonality of the trigonometric system,

$$\begin{aligned} \int_{\mathbb{D}} |f'(z)|^2 dx dy &= \int_0^{2\pi} dt \int_0^1 \left| \sum_{n=1}^{\infty} n a_n r^{n-1} e^{i(n-1)t} \right|^2 r dr \\ &= \int_0^1 \left( \int_0^{2\pi} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} m n a_m \bar{a}_n r^{m-n} e^{i(m+n-2)t} dt \right) r dr \\ &= 2\pi \sum_{n=1}^{\infty} n^2 |a_n|^2 \int_0^1 r^{2n-2} r dr \\ &= \pi \sum_{n=0}^{\infty} n |a_n|^2 \\ &= \pi [f]. \end{aligned}$$

□

---

<sup>1</sup>Arne Beurling (1905-1986) was one of the great analysts of the middle part of the XX century, and he gave fundamental contributions to a number of areas. This [article](#) summarizes his role in the Swedish cryptanalysis program during WWII.

The first geometric interpretation of  $[f]$  is in terms of areas.

PROPOSITION 5.3. *We have that*

$$(5.2.13) \quad [f] = \frac{1}{\pi} \text{Area}(f(\mathbb{D}))$$

is  $1/\pi$  times the area of  $f(\mathbb{D})$ , computed with multiplicities (that is, if  $f$  is a  $n$  – to – 1 map from open  $A \subset \mathbb{D}$  to  $f(A)$ , then the area of  $f(A)$  appears multiplied times  $n$  in the right hand side of (5.2.13)).

PROOF. The function  $f = u + iv$  is holomorphic, hence  $f' = \partial_x f = \partial_x u + i\partial_x v$  and, when  $f$  is thought of as a map from  $\mathbb{R}^2$  to itself, its Jacobian is, by Cauchy-Riemann equations,

$$Jf = \begin{pmatrix} \partial_x u & \partial_y u \\ \partial_x v & \partial_y v \end{pmatrix} = \begin{pmatrix} \partial_x u & -\partial_x v \\ \partial_x v & \partial_x u \end{pmatrix},$$

hence,

$$\det(Jf) = (\partial_x u)^2 + (\partial_x v)^2 = |f'|^2.$$

By the change of variables formula for double integrals,

$$\int_{\mathbb{D}} |f'(z)|^2 dx dy = \int_{\mathbb{D}} \det(Jf)(z) dx dy = \int_{f(\mathbb{D})} du dv = \text{Area}(f(\mathbb{D})),$$

as wished. □

In order to draw some further geometric information on the seminorm  $[f]^{1/2}$ , we review some facts concerning the unit disc. Any biholomorphic map  $\varphi : \mathbb{D} \rightarrow \mathbb{D}$  (in literature such maps are often called *automorphisms* of  $\mathbb{D}$ , or *Möbius maps*) can be uniquely expressed in the form

$$(5.2.14) \quad \varphi(z) = v \frac{a - z}{1 - \bar{a}z},$$

where  $v, a$  are complex numbers,  $|a| < 1$  and  $|v| = 1$ . Since  $f \circ \varphi$  is a map having the same image as  $f$  (counting multiplicities), we have that

$$(5.2.15) \quad [f \circ \varphi] = [f],$$

i.e. the quantity  $[f]$  is a *conformal invariant*.

EXERCISE 5.4. Let  $p_n(z) = z^n$ , and  $f \in \mathcal{D}$ . Show that  $[f \circ p_n] = n[f]$ .

Another interpretation of the Dirichlet seminorm comes from hyperbolic geometry, and it is based of a couple of "magic relations" concerning automorphisms of the disc. If  $\varphi$  is the map in (5.2.14)

$$(5.2.16) \quad \begin{aligned} 1 - |\varphi(z)|^2 &= \frac{(1 - |a|^2)(1 - |z|^2)}{|1 - \bar{a}z|^2} \\ |\varphi'(z)|^2 &= \frac{1 - |a|^2}{|1 - \bar{a}z|^2} = \frac{1 - |\varphi(z)|^2}{1 - |z|^2}. \end{aligned}$$

EXERCISE 5.5. *Verify the equations (5.2.16).*

Consider on  $\mathbb{D}$  the Riemannian metric

$$ds^2 = \frac{4|dz|^2}{(1 - |z|^2)^2},$$

which is called the *hyperbolic metric*, or *Poincaré metric* on  $\mathbb{D}$ . The area form associated to  $ds^2$  is

$$dA_{hyp}(z) = \frac{4dxdy}{(1 - |z|^2)^2}.$$

Let  $d$  be the Riemannian distance associated to the metric,

$$d(z, w) = \log \frac{1 + \left| \frac{z-w}{1-\bar{w}z} \right|}{1 - \left| \frac{z-w}{1-\bar{w}z} \right|}.$$

(Look for a proof on the web, or produce your own).

The relations (5.2.16) imply that the automorphisms of  $\mathbb{D}$  are isometries of  $ds^2$ :

$$\frac{4|d\varphi(z)|^2}{(1 - |\varphi(z)|^2)^2} = \frac{4|dz|^2}{(1 - |z|^2)^2}.$$

Now, for a holomorphic map  $f : \mathbb{D} \rightarrow \mathbb{C}$ , the *hyperbolic distortion* is (written in a somehow old fashioned way) the ratio between the infinitesimal increment of  $f$  in the Euclidean plane and that of the independent variable in the disc endowed with the hyperbolic distance:

$$\delta_{hyp}(f) := \frac{|f(z + dz) - f(z)|}{d(z + dz, z)} = \frac{|f'(z)| \cdot |dz|}{\frac{2|dz|}{1 - |z|^2}} = (1 - |z|^2)|f'(z)|/2.$$

Since they express quantities of hyperbolic geometry, the hyperbolic area element and the hyperbolic distortion of a holomorphic function  $f : \mathbb{D} \rightarrow \mathbb{C}$  are invariant under any automorphism  $\varphi$ ,

$$dA_{hyp}(\varphi(z)) = dA_{hyp}(z), \text{ and } \delta_{hyp}^2(f \circ \varphi)(z) = \delta_{hyp}^2(f)(z).$$

Hence, the same holds for

$$\pi[f] = \int_{\mathbb{D}} |f'(z)|^2 dx dy = \int_{\mathbb{D}} \delta_{hyp}^2(f) dA_{hyp}.$$

That is,

$$\pi[f \circ \varphi] = \pi[f].$$

This new, and different, argument proving the invariance of the Dirichlet seminorm, has the advantage that it allows us to write down a number of different, conformally invariant quantities. For instance,

$$(5.2.17) \quad \|f\|_{\mathcal{B},*} := \sup_{z \in \mathbb{D}} (1 - |z|^2) |f'(z)| = 2 \sup_{z \in \mathbb{D}} \delta_{hyp}(f)(z),$$

the *Bloch seminorm* of the holomorphic function  $f$ , is conformally invariant,

$$(5.2.18) \quad \|f \circ \varphi\|_{\mathcal{B},*} = \|f\|_{\mathcal{B},*}.$$

The *Bloch space*  $\mathcal{B} = \mathcal{B}(\mathcal{D})$  is the space of functions for which  $\|f\|_{\mathcal{B},*} < \infty$ , and it is a Banach space under the norm  $\|f\|_{\mathcal{B}} = \|f\|_{\mathcal{B},*} + |f(0)|$  (which is *not* conformally invariant), where the term  $|f(0)|$  was added to avoid that the norm of a constant vanishes.

### 5.3. Multipliers of a RKHS

The translation of convolution on the "time side" into products on the "frequency side" suggests that multiplication operators should play a prominent role in reproducing Hilbert space theory. This happens, in fact, even outside Fourier theory proper. In this section we see the definition and the first properties of such multiplication operators in general, then we characterize the multipliers of the (holomorphic) Hardy space.

**5.3.1. Multipliers in RKHS.** Let  $H$  be a reproducing kernel Hilbert space on a set  $X$ , with reproducing kernel  $k(x, y) = k_y(x)$ . A function  $b : X \mapsto \mathbb{C}$  is a *multiplier* of  $H$  if the linear operator  $M_b : f \mapsto bf$  maps  $H$  into  $H$ .

LEMMA 5.3. *If  $b$  is a multiplier, then  $M_b$  is bounded on  $H$ .*

PROOF. We show that the graph of  $M_b$  is closed in  $H \times H$ . Suppose  $f_n \rightarrow f$  and  $bf_n \rightarrow g$  in  $H$ . By Cauchy-Schwarz and the reproducing property,

$$|f_n(x) - f(x)| = |\langle k_x, f_n \rangle - \langle k_x, f \rangle| \leq \|k_x\| \cdot \|f_n - f\| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Then,

$$g(x) = M_b^* k_x = \lim_{n \rightarrow \infty} \langle k_x, bf_n \rangle$$



$$\begin{aligned}
&= \lim_{n \rightarrow \infty} b(x) f_n(x) \\
&= \overline{b(x)} f(x).
\end{aligned}$$

Hence,  $bf \in H$ , proving the closedness of  $M_b$ .  $\square$

We denote by  $\mathcal{M}(H)$  the *multiplier space* of  $H$ , and by  $\|b\|_{\mathcal{M}(H)} := \|M_b\|_{\mathcal{B}(H)}$  the *multiplier norm* of  $b$ .

**THEOREM 5.7.** *Let  $b \in \mathcal{M}(h)$  and let  $M_b^*$  be the Hilbert space adjoint of the corresponding multiplication operator. Then, for all  $x$  in  $X$ ,*

$$(5.3.1) \quad M_b^* k_x = \overline{b(x)} k_x.$$

*Viceversa, if  $T : H \rightarrow H$  is a bounded operator having each kernel function  $k_x$  as eigenfunctions, with eigenvalue  $\overline{b(x)}$ , then  $T^* = M_b$  is a multiplication operator.*

**PROOF.** About the direct statement,

$$\begin{aligned}
M_b^* k_x(y) &= \langle k_y, M_b^* k_x \rangle \\
&= \langle M_b^* k_y, k_x \rangle \\
&= \overline{\langle k_x, M_b k_y \rangle} \\
&= \overline{\langle k_x, b k_y \rangle} = \overline{b(x) k_y(x)} \\
&= \overline{b(x)} k_x(y).
\end{aligned}$$

For the converse one, the operator  $T$  satisfies

$$\begin{aligned}
(T^* f)(x) &= \langle k_x | T^* f \rangle \\
&= \langle T k_x | f \rangle \\
&= \langle \overline{b(x)} k_x | f \rangle \\
&= b(x) \langle k_x, f \rangle \\
&= b(x) f(x),
\end{aligned}$$

it is then multiplication times  $b$ .  $\square$

**COROLLARY 5.2.** *We have the estimate  $\|b\|_{\mathcal{M}(H)} \geq \sup_{x \in X} |b(x)|$ .*

**PROOF.** In fact,

$$\|M_b\|_{\mathcal{B}(H)} = \|M_b^*\|_{\mathcal{B}(H)} \geq \sup_{x \in X} \frac{\|M_b^* k_x\|}{\|k_x\|} = \sup_{x \in X} \left( |\overline{b(x)}| \frac{\|k_x\|}{\|k_x\|} \right) = \sup_{x \in X} |b(x)|.$$

$\square$

The fact that the reproducing functions are eigenvalues of  $M_b^*$ , and that  $\langle k_x | k_y \rangle = k(x, y)$  does not vanish identically, but for trivial examples, says that multiplication operators in RKHS are very far from being self-adjoint.

**EXERCISE 5.6.** *The multiplier space  $\mathcal{M}(H)$  of a RKHS  $H$  is complete with respect to the multiplier norm.*

It is reasonable asking if, in general,  $\|b\|_{\mathcal{M}(H)} = \sup_{x \in X} |b(x)|$ , or at least if we can provide some estimate  $\|b\|_{\mathcal{M}(H)} \approx \sup_{x \in X} |b(x)|$ . The answer is negative, and much research goes into finding sharp estimates of the multiplier norm for specific RKHS's in terms of concrete "size" information concerning the multiplier. In the Hardy space  $H^2(\mathbb{D})$ , however multiplier norm and sup norm coincide, as we show in the next subsection.

**5.3.2. The multiplier space of  $H^2(\mathbb{D})$ .** We denote by  $H^\infty(\mathbb{D})$  the space of the bounded, holomorphic functions on the unit disc. We endow it with the sup norm,

$$\|b\|_{H^\infty(\mathbb{D})} = \sup_{|z| < 1} |b(z)|.$$

**THEOREM 5.8.** *A function  $b : \mathbb{D} \rightarrow \mathbb{C}$  is a multiplier of  $H^2(\mathbb{D})$  if and only if it is holomorphic and bounded. Moreover,*

$$(5.3.2) \quad \|b\|_{\mathcal{M}(H^2(\mathbb{D}))} = \|b\|_{H^\infty(\mathbb{D})}.$$

**PROOF.** Since  $1 \in H^2(\mathbb{D})$ , if  $b$  is a multiplier, then  $b = b \cdot 1 \in H^2$ , hence it is holomorphic. We already know from the general theory that  $\|b\|_{\mathcal{M}(H^2(\mathbb{D}))} \geq \|b\|_{H^\infty(\mathbb{D})}$ . In the other direction,

$$\begin{aligned} \|bf\|_{H^2}^2 &= \sup_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} |b(re^{it})f(re^{it})|^2 dt \\ &\leq \|b\|_{H^\infty(\mathbb{D})}^2 \sup_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^2 dt \\ &= \|b\|_{H^\infty(\mathbb{D})}^2 \|f\|_{H^2}^2. \end{aligned}$$

Passing to sup over  $f$ 's with  $\|f\|_{H^2} = 1$ , we see that  $\|b\|_{\mathcal{M}(H^2(\mathbb{D}))} \leq \|b\|_{H^\infty(\mathbb{D})}$ .  $\square$

## CHAPTER 6

### Banach algebras

#### 6.1. Banach algebras

**Notation** In this chapter we denote by  $Cl(E)$  the closure of a subset  $E$  in  $\mathbb{C}$ . We write  $\overline{E} = \{\bar{z} : z \in E\}$  to denote the set of the complex conjugates of elements in  $E$ .

**6.1.1. Definition and basic properties.** An *associative algebra* is a vector space  $\mathcal{A}$  endowed with a product  $\cdot : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$  such that, for  $A, B, C \in \mathcal{A}$  and  $\lambda \in \mathbb{C}$ ,

- (a)  $(AB)C = A(BC)$ ,
- (b)  $\lambda(AB) = (\lambda A)B = A(\lambda B)$ ,
- (c)  $(A + B)C = AB + BC$  and  $A(B + C) = AB + AC$ .

An element  $1 \in \mathcal{A}$  is a *unit* if  $1A = A1 = A$  for  $A \in \mathcal{A}$ .

The algebra  $\mathcal{A}$  is *normed* when endowed with a norm  $A \mapsto \|A\|$  such that

$$(6.1.1) \quad |AB| \leq |A| \cdot |B|.$$

It is a *Banach algebra* if it is complete with respect to the given norm. If the Banach algebra has a unit  $I$ , we require that

$$|I| = 1.$$

As we have seen before, examples of Banach algebras are given by  $\mathcal{L}(X)$ , where  $X$  is a Banach space. Even the finite dimensional Banach algebra of the  $2 \times 2$  complex matrices is instructive in guessing which phenomena might occur.

If a Banach algebra  $\mathcal{A}$  does not have a unit, we can add it in a canonical way. Let  $\tilde{\mathcal{A}} = \mathcal{A} \times \mathbb{C}$  containing elements  $(A, z) \equiv A + 1z$  with the product

$$(A + Iz)(B + 1w) := AB + zB + wA + Izw.$$

Then,  $\tilde{\mathcal{A}}$  is an associative algebra with unit  $I$ ,  $\mathcal{A}$  has codimension one in  $\tilde{\mathcal{A}}$ , and it becomes a normed algebra under the norm

$$\|A + Iz\| := |A| + |z|.$$

EXERCISE 6.1. Show that  $(\tilde{\mathcal{A}}, \|\cdot\|)$  is a Banach algebra if  $(\mathcal{A}, |\cdot|)$  is.

From now on, all our algebras will have a unit, unless otherwise stated. Here are some notable examples of Banach algebras.

- (1) Let  $X$  be a Banach space. Then,  $\mathcal{L}(X)$ , the algebra of the bounded operators on  $X$ , is a Banach algebra.
- (2) In particular,  $\mathcal{L}(H)$  is a Banach algebra when  $H$  is a Hilbert space.
- (3) By Young's inequality,  $L^1(\mathbb{R})$  with the convolution product is a (commutative) Banach algebra. The same holds for  $\ell^1(\mathbb{Z})$  and, more generally, for  $L^1(G)$  (with respect to Haar measure), when  $G$  is a locally compact group (but commutativity ceases to hold in general, unless  $G$  is a commutative group).
- (4) The  $n \times n$  matrices with complex entries form a Banach algebra with respect to the operator norm.
- (5) Let  $\Omega$  be a locally compact topological space. Then,  $C_0(\Omega)$  (continuous functions vanishing at infinity) and  $C_b(\Omega)$  (bounded, continuous functions) are (commutative) Banach algebras.  $C_0(\Omega)$  has unit if and only if  $\Omega$  is compact.
- (6) The multiplier space  $\mathcal{M}(H)$  of a reproducing kernel Hilbert space  $H$  is a (commutative) Banach algebra.
- (7) In particular,  $H^\infty(\mathbb{D})$  is a (commutative) Banach algebra.

An element  $A$  in  $\mathcal{A}$  is *invertible* if there is  $E$  in  $\mathcal{A}$  such that  $AE = EA = I$ . We denote it by  $A^{-1}$ . It might happen that  $A$  has just a *left inverse*  $E$ ,  $EA = I$ , or a *right inverse*  $F$ ,  $AF = I$ . If they both exist,  $E = F$  is an inverse for  $A$ , since

$$E = EI = E(AF) = (EA)F = IF = F.$$

The shift  $\tau_1$  on  $\ell^2(\mathbb{N})$ ,  $\tau_1 f(n) = \begin{cases} 0 & \text{if } n = 0 \\ f(n-1) & \text{if } n \geq 1 \end{cases}$  has  $\tau_1^* f(n) = f(n-1)$ , the *back-shift*, as left inverse, but it does not have a right inverse:

$$\tau_1^* \tau_1 = I, \text{ but } \tau_1 \tau_1^* = I - \delta_0.$$

On the algebraic side, we have some useful properties.

THEOREM 6.1. (1) If  $A, B \in \mathcal{A}$  are invertible, then  $AB$  is invertible. If  $AB = BA$ , then  $B^{-1}A^{-1} = A^{-1}B^{-1}$ .

(2) If  $AB = BA$  and  $AB$  is invertible, then  $A$  and  $B$  are invertible.

The example above shows that  $AB = BA$  is necessary in (2), since  $\tau_1^* \tau_1 = I$ , but neither factor is invertible.

PROOF. (1) We have

$$(AB)B^{-1}A^{-1} = I = B^{-1}A^{-1}(AB).$$

The second assertion is an easy exercise. (2) The elements  $(AB)^{-1}A$  and  $A(AB)^{-1}$  are left and right inverse of  $B$ ,

$$[(AB)^{-1}A]B = (AB)^{-1}(AB) = I = (AB)(AB)^{-1} = (BA)(AB)^{-1} = B[A(AB)^{-1}],$$

hence, they coincide and are a bilateral inverse for  $B$ .  $\square$

**6.1.2. Invertibility, resolvent, and spectrum.** The next result has both an algebraic and a topological nature.

**THEOREM 6.2.** *The set of the invertible elements in  $\mathcal{A}$  is open. More precisely, if  $A_0$  is invertible and  $|A - A_0| < |A_0^{-1}|^{-1}$ , then  $A$  is invertible.*

PROOF. We start by elements which are close to the identity,  $A = I - H$  with  $|H| < 1$ . The idea is mimicking the usual geometric series argument in the complex plane,

$$\frac{1}{1 - h} = \sum_{n=0}^{\infty} h^n.$$

The series  $\sum_{n=0}^{\infty} H^n$  converges for  $|H| < 1$ , and using the telescopic property of the sum,

(6.1.2)

$$(I - H) \sum_{n=0}^{\infty} H^n = \sum_{n=0}^{\infty} (H^n - H^{n+1}) = I - H^{m+1} + \sum_{n=m+1}^{\infty} (H^n - H^{n+1}) = I,$$

and similarly  $\sum_{n=0}^{\infty} H^n (I - H) = I$ .

Suppose  $A_0$  is invertible. Then,  $A = A_0 - H = A_0(I - A_0^{-1}H)$  is invertible if  $|A - A_0| = |H| < |A_0^{-1}|^{-1}$ , with inverse

$$A^{-1} = (I - A_0^{-1}H)^{-1} A_0^{-1} = \sum_{n=0}^{\infty} (A_0^{-1}H)^n A_0^{-1},$$

by what we have seen above.  $\square$

The series in (6.1.2) is called the *Neumann series* of  $H$ .

The *resolvent* of an element  $A$  of  $\mathcal{A}$  is

$$(6.1.3) \quad \rho(A) := \{\lambda \in \mathbb{C} : \lambda I - A \text{ is invertible}\} \subset \mathbb{C},$$

and the *spectrum* of  $A$  is

$$(6.1.4) \quad \sigma(A) := \{\lambda \in \mathbb{C} : \lambda I - A \text{ is not invertible}\} = \mathbb{C} \setminus \rho(A).$$

THEOREM 6.3. (1) For any  $A$ ,  $\rho(A)$  is open and  $\sigma(A)$  is compact.

(2) The **resolvent function** of  $A$ ,  $R_\lambda(A) : \rho(A) \rightarrow \mathbb{C}$ ,

$$(6.1.5) \quad R_\lambda(A) := (\lambda I - A)^{-1}$$

is holomorphic on  $\rho(A)$ .

PROOF. (1)  $\lambda \in \rho(A)$  if and only if  $\lambda I - A$  is invertible. If  $h \in \mathbb{C}$ ,

$$(\lambda - h)I - A = (\lambda I - A) - hI,$$

which is invertible by theorem 6.2, provided  $|h| < |(\lambda I - A)^{-1}|^{-1}$ . We so have that  $\rho(A)$  is open, hence that  $\sigma(A)$  is closed. On the other hand,  $\lambda I - A$  is invertible if  $|\lambda| > |A|$ , hence  $\rho(A) \supseteq \mathbb{C} \setminus \overline{B(0, |A|)}$ , and  $\sigma(A) \subseteq \overline{B(0, |A|)}$  is compact.

(2) With  $A$  and  $\lambda \in \rho(A)$  as in (1), and complex  $h$ ,

$$\begin{aligned} R_{\lambda-h}(A) &= [(\lambda - h)I - A]^{-1} \\ &= [(\lambda I - A) - hI]^{-1} = (\lambda I - A)^{-1} [I - h(\lambda I - A)^{-1}]^{-1} \\ &= (\lambda I - A)^{-1} \sum_{n=0}^{\infty} (\lambda I - A)^{-n} h^n, \end{aligned}$$

which defines a holomorphic function of the variable  $h$  for  $|h| < |(\lambda I - A)^{-1}|^{-1}$ .  $\square$

COROLLARY 6.1. For the distance of  $\lambda \in \rho(A)$  to  $\sigma(A)$  we have the estimate

$$\min\{|\lambda - z| : z \in \sigma(A)\} \geq |(\lambda I - A)^{-1}|^{-1}.$$

THEOREM 6.4. For any  $A$  in  $\mathcal{A}$ ,  $\sigma(A) \neq \emptyset$ .

PROOF. We have

$$(6.1.6) \quad R_\lambda(A) = (\lambda I - A)^{-1} = \lambda^{-1} (I - A/\lambda)^{-1} = \sum_{n=0}^{\infty} A^n \lambda^{-n-1},$$

which converges for  $|\lambda| > |A|$  and it is a Laurent series. Its coefficients can be computed via complex integrals on contours  $\gamma$  in  $|\lambda| > |A|$  and, in particular,

$$(6.1.7) \quad I = \frac{1}{2\pi i} \int_{\gamma} (\lambda I - A)^{-1} d\lambda.$$

If the spectrum were empty,  $\lambda \mapsto R_{\lambda}(A)$  would be holomorphic on the whole plane and bounded by (6.1.6), hence constant, and again by (6.1.6) we have  $\lim_{\lambda \rightarrow \infty} (\lambda I - A)^{-1} = 0$ ,

$$R_{\lambda}(A) = \frac{\sum_{n=0}^{\infty} A^n \lambda^{-n}}{\lambda} \rightarrow 0 \text{ as } \lambda \rightarrow \infty.$$

Hence  $(\lambda I - A)^{-1} = 0$  for all  $\lambda$ , which contradicts (6.1.7).  $\square$

**6.1.3. Example: the spectrum of a multiplication operator on  $L^2(\nu)$ .** Many spectral theorems for operators  $L$  defined on a Hilbert space have the form:  *$L$  is unitarily equivalent to a multiplication operator on some  $L^2$  space.* Here we consider the bounded operators having this form.

Let  $(X, \nu)$  be a measure space in which for each measurable  $E$ ,

$$\nu(E) = \sup_{F \subseteq E, \nu(F) < \infty} \nu(F).$$

Let  $\varphi : X \rightarrow \mathbb{C}$  be a measurable function. Denote by  $M_{\varphi} : f \mapsto \varphi f$  the corresponding multiplication operator. The (natural) domain of  $M_{\varphi}$  is  $D(M_{\varphi}) = \{g \in L^2(\nu) : M_{\varphi}g \in L^2(\nu)\}$ .

LEMMA 6.1. (i)  $D(M_{\varphi}) = \{g : \int_X (1 + |\varphi|^2)|g|^2 d\nu\} < \infty$ . Endowed with the norm

$$\|g\|_{D(M_{\varphi})}^2 = \int_X (1 + |\varphi|^2)|g|^2 d\nu,$$

$D(M_{\varphi})$  is a Hilbert space, which is dense in  $L^2(\nu)$  with respect to the  $L^2(\nu)$ -norm.

(ii)  $M_{\varphi} \in \mathcal{L}(L^2(\nu))$  if and only if  $\varphi \in L^{\infty}(\nu)$ . In fact,

$$\|M_{\varphi}\|_{\mathcal{L}(L^2(\nu))} = \|\varphi\|_{L^{\infty}(\nu)}.$$

(iii)  $\varphi \mapsto M_{\varphi}$  is an isometric algebra homomorphism from  $L^{\infty}(\nu)$  into  $\mathcal{L}(L^2(\nu))$ ; in particular,  $M_{\varphi}M_{\psi} = M_{\varphi\psi}$ . Moreover, it is a  $*$ -homomorphism:  $M_{\varphi}^* = M_{\bar{\varphi}}$ . Also, the operator  $M_{\varphi}$  is **normal**,  $M_{\varphi}M_{\varphi}^* = M_{\varphi}^*M_{\varphi}$ .

PROOF. (i)  $\int_X (1 + |\varphi|^2)|g|^2 d\mu < \infty$  if and only if  $\int_X |g|^2 d\mu < \infty$  and  $\int_X \varphi g|^2 d\mu < \infty$ , if and only if  $g \in L^2(\nu)$  and  $M_\varphi g \in L^2(\nu)$ . Hence, the identification of  $D(M_\varphi)$ . The norm  $\|g\|_{D(M_\varphi)}$  is the  $L^2$  norm with respect to the measure  $(1 + |\varphi|^2)d\nu$ , which defines a Hilbert space.

(ii) Since  $\|\varphi g\|_{L^2(\nu)} \leq \|\varphi\|_{L^\infty(\nu)} \|g\|_{L^2(\nu)}$ , we have  $\|M_\varphi\|_{\mathcal{L}(L^2(\nu))} \leq \|\varphi\|_{L^\infty(\nu)}$ . In the other direction, we can suppose  $\|\varphi\|_{L^\infty(\nu)} > 0$ . Let  $0 < \lambda < \|\varphi\|_{L^\infty(\nu)}$ , and let  $E \subseteq \{x \in X : |\varphi(x)| > \lambda\}$  such that  $0 < \nu(E) < \infty$ . Then  $\nu(E) > 0$  and

$$\frac{\|\varphi \chi_E\|_{L^2(\nu)}^2}{\|\chi_E\|_{L^2(\nu)}^2} \geq \lambda^2,$$

hence,  $\|M_\varphi\|_{\mathcal{L}(L^2(\nu))} \geq \|\varphi\|_{L^\infty(\nu)}$ .

(iii) The proof is an easy exercise in bringing abstract definitions down to earthly objects.  $\square$

Let  $\varphi : X \rightarrow \mathbb{C}$  be measurable. Its *essential range* is

$$\text{essRan}(\varphi) = \{w \in \mathbb{C} : \nu(\varphi^{-1}(D(w, \epsilon))) > 0 \text{ for all } \epsilon > 0\}.$$

An arbitrarily small perturbation of a point in the essential range of  $\varphi$ , that is, belongs to the range of  $\varphi$  with "positive probability".

THEOREM 6.5. *If  $\varphi \in L^\infty(\nu)$ , then*

$$(6.1.8) \quad \sigma(M_\varphi) = \text{essRan}(\varphi).$$

PROOF. If  $w \notin \text{essRan}(\varphi)$ , then there is  $\epsilon > 0$  such that  $|\varphi(x) - w| \geq \epsilon$  for a.e.  $x$ . Then,  $(\varphi - w)^{-1} \in L^\infty(\nu)$ ,  $\|(\varphi - w)^{-1}\|_{L^\infty(\nu)} \leq 1/\epsilon$ , hence,  $M_{(\varphi - w)^{-1}} = (M_\varphi - wI)^{-1}$  is the inverse of  $M_\varphi - wI$ . This shows that  $w \in \rho(M_\varphi)$ .

In the opposite direction, let  $w \in \text{essRan}(\varphi)$  and fix any  $\epsilon > 0$ . Then, there  $E_\epsilon \subseteq X$  such that  $0 < \nu(E_\epsilon) < \infty$  and  $|\varphi(x) - w| \leq \epsilon$  for all  $x \in E_\epsilon$ . Then,  $\|(\varphi - w)^{-1}\|_{L^\infty(\nu)} \geq 1/\epsilon$ . This shows that  $M_\varphi - wI$  does not have a bounded inverse, hence that  $w \in \sigma(M_\varphi)$ .  $\square$

Some purely operator theoretic properties of multiplication operators are easily spotted and proved.

A complex number  $\lambda$  is an *eigenvalue* of a bounded operator  $L : H \rightarrow H$  ( $H$  Hilbert) if there exists  $0 \neq h \in H$  such that

$$Lh = \lambda h.$$

We say that  $h$  is an *eigenvector* relative to  $\lambda$ , and the linear space  $E_\lambda$  of such eigenvectors is the *eigenspace* relative to the eigenvector  $\lambda$ . Clearly,  $\lambda \in \sigma(L)$ .



EXERCISE 6.2. Let  $\varphi \in L^\infty(\nu)$ . Then,  $\lambda$  is an eigenvector for  $\lambda$  if and only if there is a measurable set  $A$  with  $0 < \nu(A)$  such that  $\varphi(x) = \lambda$  for all  $x \in A$ .

An operator  $U : H \rightarrow H$  is *unitary* if  $UU^* = U^*U = I$ . A bounded operator  $A : H \rightarrow H$  is *self-adjoint* if and only if  $A = A^*$ .

EXERCISE 6.3. Let  $\varphi \in L^\infty(\nu)$ . Show that  $M_\varphi$  is unitary if and only if  $|\varphi(x)| = 1$  a.e. in  $X$ . Show that  $M_\varphi$  is self-adjoint if and only if  $\varphi(x) \in \mathbb{R}$  a.e. in  $X$ . Deduce that if  $M_\varphi$  is self-adjoint, then  $M_{e^{i\varphi}}$  is self-adjoint.

**6.1.4. Example: the spectrum of a multiplier on  $H^2(\mathbb{D})$ .** If instead of  $L^2(\nu)$  we consider one of its subspaces, the spectrum of a multiplication operator might be very different from what we have seen in §6.1.3.

Let's compute spectrum and resolvent of the back-shift operator  $\tau_1^*$  on  $\ell^2(\mathbb{N})$ . We start from the most obvious enemies of invertibility, which are *eigenvalues*, i.e.  $\lambda$ 's for which  $\tau_1^*\varphi = \lambda\varphi$  has a nonvanishing solution  $\varphi \in \ell^2$ . Such  $\varphi$  are the solution to

$$\varphi(n+1) = \tau_1^*\varphi(n) = \lambda\varphi(n),$$

that is  $\varphi(n) = c\lambda^n$  for some  $c \in \mathbb{C}$ , which lies in  $\ell^2$  if and only if  $|\lambda| < 1$ . The same argument shows that  $\lambda$  is not an eigenvalue if  $|\lambda| \geq 1$ . The *point spectrum*  $\sigma_p(\tau_1^*)$  of the back-shift, the set of its eigenvalues, is the open disc  $\mathbb{D}$ . But the spectrum is closed, hence  $\sigma(\tau_1^*)$  contains the closed unit disc  $\overline{\mathbb{D}}$ .

Consider now  $|\lambda| > 1$ . Since  $\|\tau_1^*\| \leq 1 < |\lambda|$ , the inverse of  $I - \frac{\tau_1^*}{\lambda}$  exists, hence  $\lambda \in \rho(\tau_1^*)$ . We have proved:

PROPOSITION 6.1. *The spectrum of the back-shift is  $\sigma(\tau_1^*) = \overline{\mathbb{D}}$ . Its point spectrum is  $\sigma_p(\tau_1^*) = \mathbb{D}$ .*

We will say more on the role of  $\mathbb{T}$  in the spectrum when we do the spectral theory of operators on a Hilbert space.

In frequency space, the eigenvectors corresponding to  $\lambda$  are constant multiples of

$$f_\lambda(z) = \sum_{n=0}^{\infty} \lambda^n z^n = \frac{1}{1 - \lambda z} = k_{\overline{\lambda}}(z),$$

is the reproducing kernel at  $\overline{\lambda}$ ! In fact, we showed that  $M_z^* k_{\overline{\lambda}} = \lambda k_{\overline{\lambda}}$ . Our result was for general multipliers, and we obtain a general result as a consequence of it.

THEOREM 6.6. (1) *Let  $H$  be a RKHS on a set  $X$ , and let  $b$  be a multiplier. Then,*

$$\sigma(M_b^*) \supseteq Cl(\overline{b(X)}).$$

(2) If  $b \in H^\infty(\mathbb{D})$ , then

$$\sigma(M_b^*) = Cl(\overline{b(\mathbb{D})}).$$

PROOF. (1) holds because each  $\overline{b(x)}$  is an eigenvalue of  $M_b^*$ , corresponding to the eigenvector  $k_x$ . (2) In  $H^2(\mathbb{D})$  it is easy to see that there are no other elements in the spectrum. We use the fact that  $\sigma(M_b^*) = \overline{\sigma(M_b)}$ . Suppose  $\lambda \notin \overline{b(\mathbb{D})}$ . Then,  $\frac{1}{\lambda - b} \in H^\infty(\mathbb{D})$ . In fact we can say more:

$$0 < c \leq \left| \frac{1}{\lambda - b(z)} \right| \leq C < \infty.$$

This implies that  $M_{(\lambda - b)^{-1}} = M_{b - \lambda}^{-1}$  is an inverse for  $M_{\lambda - b}$ , hence,  $\lambda \in \rho(M_b)$ .  $\square$

1

**6.1.5. The spectral radius.** The *spectral radius*  $r(T)$  of  $T$  is

$$(6.1.11) \quad r(T) = \max\{|\lambda| : \lambda \in \sigma(T)\} \leq |T|.$$

To compute it, we have a beautiful formula.

<sup>1</sup>We can do the same calculations in frequency domain. The shift operator corresponds to  $M_z$ , multiplication times the  $z$  variable on  $H^2(\mathbb{D})$ , and the back-shift to

$$\begin{aligned} \langle g | M_z^* f \rangle_{H^2} &= \langle M_z g | f \rangle_{H^2} = \lim_{r \rightarrow 1} \frac{1}{2\pi} \int_{|z|=r} \overline{zg(z)} f(z) \frac{dz}{iz} \\ &= \lim_{r \rightarrow 1} \frac{1}{2\pi} \int_{|z|=1} \overline{zg(z)} (f(z) - f(0)) \frac{dz}{iz} \\ &= \lim_{r \rightarrow 1} \frac{r^2}{2\pi} \int_{|z|=1} \overline{g(z)} \frac{f(z) - f(0)}{z} \frac{dz}{iz} \\ &= \left\langle g \left| \frac{f - f(0)}{z} \right. \right\rangle_{H^2}, \end{aligned}$$

then

$$(6.1.9) \quad M_z^* f(z) = \frac{f(z) - f(0)}{z}.$$

EXERCISE 6.4. Deduce (6.1.9) from the  $\ell^2(\mathbb{N})$  expression of  $\tau_1^*$ .

The eigenvalues  $\lambda$  of  $M_z^*$  are those for which

$$(6.1.10) \quad \frac{f(z) - f(0)}{z} = \lambda f(z)$$

has a solution  $f \in H^2$ . The equation has solution  $f(z) = \frac{f(0)}{1 - \lambda z}$ , which is holomorphic in  $\mathbb{D}$  if and only if  $|\lambda| \leq 1$ , and it belongs to  $H^2(\mathbb{D})$  if and only if  $|\lambda| < 1$ .

THEOREM 6.7 (Gelfand). *We have*

$$(6.1.12) \quad r(T) = \lim_{n \rightarrow \infty} |T^n|^{1/n}.$$

PROOF. The Laurent expansion of  $R_\lambda(T)$  at infinity is, as we have seen before,

$$(6.1.13) \quad (\lambda I - T)^{-1} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}},$$

As in the case of holomorphic functions of one variable, one shows that the radius of convergence of the series with respect to  $1/\lambda$  is

$$\frac{1}{\limsup_{n \rightarrow \infty} |T^n|^{1/n}}.$$

Hence,  $R_\lambda(T)$  is holomorphic for  $|\lambda| > \limsup_{n \rightarrow \infty} |T^n|^{1/n}$ . If it were holomorphic for  $|\lambda| > Q$  with  $Q < \limsup_{n \rightarrow \infty} |T^n|^{1/n}$ , arguing as in holomorphic function theory we would deduce (using Cauchy formula) that the series in (6.1.13) would converge for  $|\lambda| > Q$ , contradicting formula for the radius of convergence. Hence,

$$r(T) = \limsup_{n \rightarrow \infty} |T^n|^{1/n}.$$

We then show that  $\lim_{n \rightarrow \infty} |T^n|^{1/n}$  exists. Observe first that  $|T^{m+n}| \leq |T^m| |T^n|$ , hence that  $a_n := \log |T^n|$  satisfies  $a_{m+n} \leq a_m + a_n$ . For  $n = mq + r$ ,  $r = 0, \dots, m-1$ , we have

$$\frac{a_n}{n} \leq \frac{qa_m}{mq+r} + \frac{a_r}{mq+r} = \frac{a_m}{m+r/r} + \frac{a_r}{mq+r},$$

thus,

$$\limsup_{n \rightarrow \infty} \frac{a_n}{n} \leq \frac{a_m}{m},$$

for each fixed  $m$ , hence,

$$\limsup_{n \rightarrow \infty} \frac{a_n}{n} \leq \liminf_{m \rightarrow \infty} \frac{a_m}{m}.$$

Then,  $\lim_{n \rightarrow \infty} \frac{a_n}{n} = \lim_{n \rightarrow \infty} \log(|T^n|^{1/n})$  exists.  $\square$

## 6.2. Holomorphic calculus

If  $A$  is an element of an algebra  $\mathcal{A}$  on  $\mathbb{C}$ ,  $q \in \mathbb{C}[z]$  is a polynomial with complex coefficients, the definition of  $q(A) \in \mathcal{A}$  is given in the obvious way,

$$(6.2.1) \quad q(z) = \sum_{j=0}^n \alpha_j z^j \mapsto \Phi_A(p) := q(A) = \sum_{j=0}^n \alpha_j A^j.$$

Interesting facts happen when  $A \neq 0$ , yet  $p(A) = 0$  for some  $p \neq 0$ , as we see below. In this section, we extend the definition from polynomials to holomorphic functions. We first review some basic algebra in order to have a basic theory to make comparisons with.

**6.2.1. The soon-to-be-lost paradise of polynomial calculus.** Given an element  $A$  in an algebra  $\mathcal{A}$  and a polynomial  $p(z)$  with complex coefficients, we can compute  $p(A)$  in the obvious way. This operation gives a algebra homomorphism  $\Phi_A : p \mapsto p(A)$  from  $\mathbb{C}[z]$  to  $\mathcal{A}$ . Actually, we can be more precise. Let  $\mathbb{C}[A]$  be the smallest sub-algebra of  $\mathcal{A}$  which contains  $A$  and  $I$  (which is a commutative algebra). Then,  $\Phi_A$  is a surjective homomorphism from  $\mathbb{C}[z]$  to  $\mathbb{C}[A]$ .

In general,  $\Phi_A$  is far from being injective. Let  $\mathcal{A}$  be the algebra of the  $n \times n$  matrices with complex coefficients. Since  $\mathcal{A}$  has dimension  $n^2$ , while  $\mathbb{C}[z]$  has infinite dimension,  $\Phi_A$  has an infinite dimensional kernel. By contrast, let  $\mathcal{A} = H^\infty(\mathbb{D})$  and let  $A(z) = z$ , the shift. Then,  $p(A)(z) = p(z)$ , which we might think of as multiplication times  $p(z)$ , hence  $\Phi_A = Id$  is the identity. Similarly, if  $\mathcal{A} = C[0, 1]$  is the algebra of the continuous functions on  $[0, 1]$ , and  $A(x) = x$ ,  $\Phi_A(p)(x) = p(x)$ .

Let  $\mathcal{I} = \ker \Phi_A$ , which is an *ideal* in  $\mathbb{C}[z]$ . Recall that all ideals in the algebra of polynomials are *principal ideals*, i.e. there is a *generating polynomial*  $p$ , uniquely determined but for a multiplicative constant, such that  $\mathcal{I} = \langle p \rangle := \{p \cdot q : q \in \mathbb{C}[z]\}$ . To see this, it suffices to show that for  $a, b \in \mathcal{I}$ ,  $GCD(a, b) \in \mathcal{I}$ , but this follows from the Euclidean algorithm to compute the greatest common divisor of two polynomials. If  $\ker \Phi_A = \langle p \rangle$ , then  $p(A) = 0$  and  $q(A) \neq 0$  for all polynomials which are not multiple of  $p$  (in particular, for all proper factors of  $p$ ).

We have then a complete classification of the algebras  $\mathbb{C}[A]$ , since they are all isomorphic to  $\mathbb{C}[z]/\langle p \rangle$  for a uniquely determined, monic  $p$ , or (when  $\ker \Phi_A = 0$ )  $\mathbb{C}[A]$  is isomorphic to  $\mathbb{C}[z]$ . The isomorphism is canonical:  $q(A) \mapsto [q]_{\text{mod } p}$ .

Summarizing, we have the first two items of the following statement.

**THEOREM 6.8.** (i) *Let  $A \in \mathcal{A}$ . Then, there is a unique minimal monic polynomial  $p(z)$  such that  $p(A) = 0$ . Let  $[q]_{\text{mod } p} = \langle p \rangle + q$*

denote the equivalence class of  $q$  in  $\mathbb{C}[z]/\langle p \rangle$ . Then, the map

$$(6.2.2) \quad [q] \mapsto q(A)$$

is an isomorphism of  $\mathbb{C}[z]/\langle p \rangle$  onto  $\mathbb{C}[A]$ .

(ii) In (i),  $\dim(\mathbb{C}[A]) = \deg(p)$ , if  $\deg(p) \neq 0$ , and  $\dim(\mathbb{C}[A]) = \infty$  if  $\deg(p) = 0$ .

(iii) Viceversa, let  $p$  be a monic polynomial with  $\deg(p) \geq 1$ , let  $\mathcal{A} = \mathbb{C}[z]/\langle p \rangle$ , and set  $A = [z] \bmod p$ . Then,  $q(A) = [q] \bmod p$ .

PROOF. The third one follows from the fact that  $\pi_p : q \mapsto [q] \bmod p$  is a homomorphism of algebras,  $\pi_p : \mathbb{C}[z] \rightarrow \mathbb{C}[z]/\langle p \rangle$ .  $\square$

The nonconstant polynomial  $p(z)$  can be written as

$$(6.2.3) \quad p(z) = (z - \lambda_1)^{m_1} \dots (z - \lambda_l)^{m_l},$$

where  $\lambda_1, \dots, \lambda_l$  are the distinct roots of  $p$ , and  $m_j \geq 1$  is the multiplicity of  $\lambda_j$ .

Before we proceed, let's consider two instructive examples. If all multiplicities are one, then we can consider the diagonal  $l \times l$  matrix

$$\text{Diag}(\lambda_1, \dots, \lambda_l) = \begin{pmatrix} \lambda_1 & 0 & \dots \\ 0 & \lambda_2 & \dots \\ 0 & 0 & \dots \\ \dots & & \end{pmatrix}$$

for which

$$p(\text{Diag}(\lambda_1, \dots, \lambda_l)) = \begin{pmatrix} p(\lambda_1) & 0 & \dots \\ 0 & p(\lambda_2) & \dots \\ 0 & 0 & \dots \\ \dots & & \end{pmatrix} = 0,$$

and  $p$  is clearly the minimal polynomial with this property.

Consider  $p(z) = (z - 1)^2$ . A "concrete" model of an algebra  $\mathcal{A}$  and of an element  $A$  in it, such that  $\mathcal{A}$  is isomorphic to  $\mathbb{C}[z]/\langle p \rangle$  is the algebra of the  $2 \times 2$  matrices generated by the identity and

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

which satisfies  $(A - I)^2 = 0$ , but  $I \neq A$ . In general, a representation of this kind can be obtained in terms of the Jordan decomposition of a matrix.

An element in  $\mathbb{C}[z]/\langle p \rangle$  can be thought of as a polynomial with  $\deg(r) < \deg(p)$ . For any polynomial  $a \in \mathbb{C}[z]$ ,  $a = qp + r$ , and what  $a$  and  $r$  share (a set of invariants for  $[a]_{\text{mod } p}$ ) are the  $m_1 + \dots + m_l$  complex numbers:

$$(6.2.4) \quad \begin{aligned} & a(\lambda_1) = r(\lambda_1), a'(\lambda_1) = r'(\lambda_1), \dots, a^{(m_1-1)}(\lambda_1) = r^{(m_1-1)}(\lambda_1), \\ & \dots \\ & a(\lambda_l) = r(\lambda_l), a'(\lambda_l) = r'(\lambda_l), \dots, a^{(m_l-1)}(\lambda_l) = r^{(m_l-1)}(\lambda_l). \end{aligned}$$

These numbers, in turn, are sufficient to determine  $r$ . There are interpolating formulas dating back to Lagrange, reconstructing  $r$  by its values on the zeros of  $p$ , if all zeros have multiplicity one, or by values of  $r$  and its derivatives in the case of higher multiplicity (see [this](#) lecture).

The set  $\{\lambda_1, \dots, \lambda_l\}$  is the (*algebraic*) *spectrum* of the algebra  $\mathbb{C}[z]/\langle p \rangle = \mathbb{C}[A]$ . We have just seen that the spectrum alone is not sufficient to determine the algebra  $\mathbb{C}[A]$ , but that we also need the multiplicities attached to each of the points in it. But for the constant  $p$ , that is, the algebras  $\mathbb{C}[A]$  are parametrized by sets of the form

$$\{(\lambda_1, m_1), \dots, (\lambda_l, m_l)\},$$

with the  $\lambda_j$ 's complex, and the  $m_j \geq 1$ 's positive integers.

An interesting special case is that of the *Hermitian matrices*, those matrices  $A$  such that  $\langle z, Aw \rangle_{\mathbb{C}^N} = \langle Az, w \rangle_{\mathbb{C}^N}$ . The infinite dimensional version of them are the self-adjoint operators, to be discussed in the chapter on spectral theory. A basic fact of linear algebra is that, after a suitable unitary change of coordinates (i.e. passing from  $A$  to  $U^*AU$ , where  $UU^* = U^*U = I$ ), all such matrices can be written as diagonal matrices with real entries, the elements  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_N$  on the diagonal being the eigenvalues. If  $\lambda_1 < \dots < \lambda_L$  are the *distinct* eigenvalues, then

$$(6.2.5) \quad A = \Delta((\lambda_1, m_1), \dots, (\lambda_L, m_L)),$$

where  $m_i$  is the multiplicity of  $\lambda_i$ . Let  $E(\lambda_i)$  be the eigenspace corresponding to the eigenvalue  $\lambda_i$ , and  $\Pi_{E(\lambda_i)}$  be the (orthogonal) projection onto it. Then, (6.2.5) can be written as:

$$(6.2.6) \quad A = \sum_{i=1}^L \lambda_i \Pi_{E(\lambda_i)},$$

which is the *spectral resolution* of  $A$ . The *projection valued measure* (p.v.m.) corresponding to  $A$  is the family  $\{\Pi_{\lambda_i} : 1 \leq i \leq L\}$ , which might be written

as a singular measure on  $\mathbb{R}$ , having values in the space of projection,

$$(6.2.7) \quad \Pi = \sum_{i=1}^L \Pi_{E(\lambda_i)} \delta_{\lambda_i}, \text{ i.e. } \Pi(B) = \sum_{\lambda_i \in B} \Pi_{E(\lambda_i)},$$

where  $\delta_{\lambda_i}$  is the Dirac mass at  $\lambda_i$ , and  $B$  is a Borel set in  $\mathbb{R}$ .

**6.2.2. Holomorphic calculus.** For a power series like

$$(6.2.8) \quad f(z) = \sum_{n=0}^{\infty} \alpha_n z^n,$$

the natural definition of  $f(A)$  is:

$$(6.2.9) \quad f(A) = \sum_{n=0}^{\infty} \alpha_n A^n.$$

Let  $R$  be the radius of convergence of  $f(z)$ . By comparison with power series, (6.2.9) converges if  $R > |A|$ , and, in fact, for  $R > r(A)$ , the spectral radius of  $A$ . Power series are not the only way to represent holomorphic functions. For instance, we have Cauchy formula, which is more flexible since it deals with regions which are more general than discs.

Let  $A$  be an element of a Banach algebra  $\mathcal{A}$ ,  $\Omega \supset \sigma(A)$  an open set in the complex plane, and  $f : \Omega \rightarrow \mathbb{C}$  a function which is holomorphic in  $\Omega$ . Let  $\gamma$  be a loop (or a formal sum of loops) in  $\Omega$  such that  $\text{Ind}(\gamma, \zeta) = 1$  for all  $\zeta$  in  $\sigma(A)$ , i.e. such that  $\gamma$  circles around each  $\zeta$  in the spectrum of  $A$  exactly once, counterclockwise. We define, then

$$(6.2.10) \quad f(A) := \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} f(z) dz.$$

We have already seen that the expression  $z \mapsto (zI - A)^{-1}$  defines an  $\mathcal{A}$ -valued holomorphic function in  $\mathbb{C} \setminus \sigma(A)$ , and the integral in (6.2.10) does not depend on the curve  $\gamma$  by Cauchy integral theorem. The element  $f(A)$ , that is, is defined if  $f$  is holomorphic on (any open set containing)  $\sigma(A)$ .

If we replace  $A$  by  $w$ , the formula gives back  $f(w)$ , and for the moment this is justification enough for the definition.

We need a relation which has an independent interest.

LEMMA 6.2. [*Resolvent identity*] If  $z, w \in \rho(A)$ , then

$$(6.2.11) \quad (zI - A)^{-1}(wI - A)^{-1} = (w - z)^{-1}[(zI - A)^{-1} - (wI - A)^{-1}].$$

PROOF. Multiply both sides of (6.2.11) by all inverses in sight:

$$(w - z)I = (wI - A) - (zI - A),$$

which in fact holds.  $\square$

THEOREM 6.9. (i) *If  $q$  is a polyomial, or more generally a function of the form (6.2.8), the definition of  $q(A)$  coincides with that given by (6.2.1), or (6.2.9) in the general case.*

(ii) *The map  $f \mapsto f(A)$  is an algebra homomorphism.*

(iii) **(Spectral mapping theorem)**  $\sigma(f(A)) = f(\sigma(A))$ .

(iv)  $(g \circ f)(A) = g(f(A))$ , where  $g : E \rightarrow \mathbb{C}$  is holomorphic on an open set  $E \supseteq f(\Omega) \supset f(\sigma(A))$ .

PROOF. (i) Let  $L_f(A)$  be the algebra element in (6.2.10): we have to show that  $L_q(A) = q(A)$ . We can assume  $\gamma$  to be a circle containing  $\sigma(q)$  in its interior. By linearity, it suffices to show it for powers  $q(z) = z^m$ . By developing a Neuman series, which converges of  $\gamma$ , we have

$$\begin{aligned} l_q(A) &= \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} z^m dz \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi i} \int_{\gamma} z^{m-1-n} dz A^n \\ &= A^m, \end{aligned}$$

as wished.

(ii) We have to show that  $L_{fg} = L_f L_g$ . Cosider a curve  $\gamma$  as in the definition, and another curve  $\delta$  with the same properties, but outside  $\gamma$ , so that  $\text{Ind}(\delta, z) = 1$  for all  $z$  on  $\gamma$ , and  $\text{Ind}(\gamma, w) = 0$  for all  $w$  on  $\delta$ . Using the resolvent identity from first to second line,

$$\begin{aligned} L_f(A)L_g(A) &= \frac{1}{(2\pi i)^2} \int_{\gamma} \int_{\delta} (zI - A)^{-1} (wI - A)^{-1} f(z)g(w)dw dz \\ &= \frac{1}{(2\pi i)^2} \int_{\gamma} \int_{\delta} (w - z)^{-1} [(zI - A)^{-1} - (wI - A)^{-1}] f(z)g(w)dw dz \\ &= \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} f(z) \left( \frac{1}{2\pi i} \int_{\delta} (w - z)^{-1} g(w)dw \right) dz \\ &\quad - \frac{1}{2\pi i} \int_{\delta} \left( \frac{1}{2\pi i} \int_{\gamma} (w - z)^{-1} f(z)dz \right) (wI - A)^{-1} g(w)dw. \end{aligned}$$



The inner integral in the second summand vanishes, while the inner integral in the first is  $g(z)$ , hence,

$$\begin{aligned} L_f(A)L_g(A) &= \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} f(z)g(z)dz \\ &= L_{fg}(A). \end{aligned}$$

(iii) If  $\mu \notin f(\sigma(A))$ , then  $f - \mu$  does not vanish on  $\sigma(A)$ , hence  $g = (f - \mu)^{-1}$  defines a function which is holomorphic in a neighborhood containing  $\sigma(A)$ . We can then define  $g(A)$ , and by (ii)

$$g(A)(f(A) - \mu I) = I = (f(A) - \mu I)g(A),$$

i.e.  $\mu \notin \sigma(f(A))$ . This shows  $\sigma(f(A)) \subseteq f(\sigma(A))$ .

In the other direction, suppose that  $\lambda \in \sigma(A)$  and observe that

$$h(z) = \frac{f(z) - f(\lambda)}{z - \lambda}$$

defines a function which is holomorphic on an open set containing  $\lambda$ , with  $h(\lambda) = f(\lambda)$ . Again by (ii) we can define  $h(A)$  and

$$h(A)(A - \lambda I) = f(A) - f(\lambda)I.$$

The first term is not invertible because  $\lambda \in \sigma(A)$ , thus the second term is not invertible either, showing that  $f(\lambda) \in \sigma(f(A))$ , hence  $f(\sigma(A)) \subseteq \sigma(f(A))$ .

(iv) Our hypothesis guarantee that  $g$  is holomorphic on an open set  $\Omega$  containing  $f(\sigma(A)) = \sigma(f(A))$ . Let  $\gamma$  be a simple curve in  $\Omega$  such that  $\text{Ind}(\gamma, \xi) = 1$  for all  $\xi$  in  $f(\sigma(f(A)))$ , and let  $E = \text{Interior}(\gamma) \supset \sigma(f(A))$  be the open region inside  $\gamma$ . Then,  $f^{-1}(E)$  is an open set containing  $A$  in its interior. Let  $\delta$  be a curve in  $f^{-1}(B)$  such that  $\text{Ind}(\delta, \zeta) = 1$  for all  $\zeta$  in  $A$ , and observe that  $f(z) \neq w$  for  $z$  in  $f^{-1}(E)$  and  $w$  on  $\delta$ , hence,  $z \mapsto (w - f(z))^{-1}$  is holomorphic on  $f^{-1}(E)$ .

Using the definition of the holomorphic function of an operator in the first two and in the last equality:

$$\begin{aligned} g(f(A)) &:= \frac{1}{2\pi i} \int_{\gamma} (wI - f(A))^{-1} g(w)dw \\ &= \frac{1}{2\pi i} \int_{\gamma} \left( \frac{1}{2\pi i} \int_{\delta} (w - f(z))^{-1} (Iz - A)^{-1} dz \right) g(w)dw \\ &= \frac{1}{2\pi i} \int_{\delta} \left( \frac{1}{2\pi i} \int_{\gamma} (w - f(z))^{-1} g(w)dw \right) (Iz - A)^{-1} dz \\ &= \frac{1}{2\pi i} \int_{\delta} g(f(z))(Iz - A)^{-1} dz \end{aligned}$$

$$\begin{aligned}
& \text{by Cauchy formula} \\
& = (g \circ f)(A).
\end{aligned}$$

□

### 6.3. Intermezzo: holomorphic calculus and Laurent series

Let  $A$  be an element of a Banach algebra  $\mathcal{A}$  having spectrum  $\sigma(A) \subseteq \text{Cl}B(0, r) = \{z \in \mathbb{C} : |z| \leq R\}$ , let  $\Omega = \Omega(r, R) = \{z \in \mathbb{C} : 0 \leq r < |z| \leq R\}$ , and let  $f$  be holomorphic in  $\Omega$ . If  $\gamma$  is a loop in  $\Omega$  such that  $\text{Ind}(\gamma, \zeta) = 1$  for all  $\zeta \in \text{Cl}B(0, r)$ , then the expression

$$(6.3.1) \quad L_A(f, \gamma) = \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} f(z) dz$$

certainly makes sense and it defines an element  $L_A(f, \gamma) \in \mathcal{A}$ . Can we identify in some way  $L_A(f, \gamma)$  with  $f(A)$ ? We provide here an example of a function  $f \neq 0$  such that  $L_A(f, \gamma) = 0$ .

Let  $\mathcal{A} = \mathbb{C}^{2 \times 2}$  be the space of the  $2 \times 2$  matrices with complex entries, let  $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  be the projection onto the first coordinate, so that  $\sigma(A) = \{0, 1\}$ , and let  $f(z) = \frac{1}{z}$ , which is holomorphic in  $\Omega(1, \infty)$ . If  $\gamma(t) = p \cdot e^{it}$  ( $t \in [0, 2\pi]$ ,  $p > 1$  fixed) is as above, then  $L_A(f, \gamma)$  is well defined, although it can not be  $A^{-1}$ , since  $A$  is not invertible. We claim that  $L_A(f, \gamma) = 0$ .

The function  $g$ ,

$$g(w) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - w} dz = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{(z - w)z}$$

vanishes in the interior of  $\gamma$ . In fact, by the residue theorem,

$$g(w) = \text{Res} \left( \frac{1}{(z - w)z}, z = 0 \right) + \text{Res} \left( \frac{1}{(z - w)z}, z = w \right) = \frac{1}{-w} + \frac{1}{w} = 0.$$

We can now compute

$$\begin{aligned}
\frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} f(z) dz &= \frac{1}{2\pi i} \int_{\gamma} \left[ z \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right]^{-1} \frac{dz}{z} \\
&= \frac{1}{2\pi i} \int_{\gamma} \begin{pmatrix} z-1 & 0 \\ 0 & z \end{pmatrix}^{-1} \frac{dz}{z} \\
&= \frac{1}{2\pi i} \int_{\gamma} \begin{pmatrix} \frac{1}{(z-1)z} & 0 \\ 0 & \frac{1}{z^2} \end{pmatrix} dz \\
&= \begin{pmatrix} g(1) & 0 \\ 0 & g(0) \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}$$

All this seems rather *ad hoc*, but it can be rephrased in a more conceptual way. Recall that any function  $f$  which is holomorphic in the annulus  $\Omega(r, R)$  can be expanded in a Laurent series there,

$$(6.3.2) \quad f(z) = \sum_{n=-\infty}^{+\infty} a_n z^n,$$

where  $f_+(z) = \sum_{n=0}^{+\infty} a_n z^n$  converges in  $B(0, R)$  and  $f_-(z) = \sum_{n=-\infty}^{-1} a_n z^n$  converges in  $\mathbb{C} \setminus [\text{Cl}(B(0, r))]$ . Then, with the hypothesis on the objects contained in (6.3.1),

$$(6.3.3) \quad L_A(f, \gamma) = f_+(A)$$

is the element provided by the "projection" of  $f$  onto the functions holomorphic in  $B(0, R)$ . Here is the proof (I leave it to you to verify all needed convergences).

$$\begin{aligned}
L_A(f, \gamma) &= \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} f(z) dz \\
&= \sum_{n=-\infty}^{+\infty} a_n \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} z^n dz \\
&= \sum_{n=-\infty}^{-1} a_n \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} z^n dz + \sum_{n=0}^{+\infty} a_n A^n \\
&= \sum_{n=-\infty}^{-1} a_n \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} z^n dz + f_+(A).
\end{aligned}$$

In the third equality we have used the holomorphic calculus. To finish, observe that for  $n < 0$

$$\begin{aligned}
 \frac{1}{2\pi i} \int_{\gamma} (zI - A)^{-1} z^n dz &= \frac{1}{2\pi i} \int_{\gamma} (I - z^{-1}A)^{-1} z^{n-1} dz \\
 &= \sum_{k=0}^{\infty} \frac{1}{2\pi i} \int_{\gamma} A^k z^{-k} z^{n-1} dz \\
 &= \sum_{k=0}^{\infty} A^k \frac{1}{2\pi i} \int_{\gamma} z^{n-1-k} dz \\
 &= 0,
 \end{aligned}$$

because  $n - 1 - k < 0$ . The second equality seems to require  $|z| > |A|$ , but if you look more closely, you realize that  $|z| > r(A)$  suffices ( $r(A)$  is the spectral radius of  $A$ ), and since  $r(A) \leq r < p$  by hypothesis, we are done.

#### 6.4. Gelfand theory of commutative Banach algebras

There are many interesting examples of commutative Banach algebras with unit. Here are some examples, not always independent of each other.

- (i)  $C_b(X)$ , the algebra of the bounded and continuous functions on a topological space  $X$  with respect to the uniform norm.
- (ii)  $L^{\infty}(\mu)$  on a measure space  $(X, \mu)$  with respect to the  $L^{\infty}$  norm.
- (iii)  $C^1[0, 1]$ , the space of the continuously differentiable functions on  $[0, 1]$ , with respect to the norm  $\|f\|_{C^1} = \|f\|_C + \|f'\|_C$ .
- (iv)  $\mathcal{M}(H)$ , the multiplier space of a reproducing kernel Hilbert space  $H$ , is a commutative Banach algebra.
- (v) In particular,  $H^{\infty}(\mathbb{D})$ , the multiplier space of  $H^2(\mathbb{D})$ , is a commutative Banach algebra.

In example (iv) you might worry about completeness. If you use both completeness of  $\mathcal{L}(H)$  and reproducing functions, worries evaporate.

EXERCISE 6.5.  $\mathcal{M}(H)$  is complete.

Gelfand theory identifies a commutative Banach algebra  $\mathcal{A}$  with a subalgebra of  $C(X)$ , the algebra of the continuous functions on a compact set  $X$ . The points of such set are the maximal ideals of  $\mathcal{A}$ . This section is devoted to precisely stating, and proving, this fact.

**6.4.1. Multiplicative functionals and ideals.** We start with some definitions. Let  $\mathcal{A}$  be an algebra. a *multiplicative functional*  $p$  on  $\mathcal{A}$  is a homomorphism of algebras  $p : \mathcal{A} \rightarrow \mathbb{C}$ . We include in the definition that  $p \neq 0$  is not the trivial homomorphism. The set of the nonzero algebra homomorphisms  $\Phi : \mathcal{A} \rightarrow \mathcal{L}$  is denoted by  $\text{Hom}(\mathcal{A}, \mathcal{L})$ , so that  $\text{Hom}(\mathcal{A}, \mathbb{C})$  is the set of the multiplicative functionals.

A *left ideal*  $\mathcal{I}$  of an algebra  $\mathcal{A}$  is a vector subspace of  $\mathcal{A}$  such that  $A\mathcal{I} \subseteq \mathcal{I}$  for all  $A$  in the algebra; it is a *right ideal* if  $\mathcal{I}A \subseteq \mathcal{I}$  instead; it is a *bilateral ideal* if it is both right and left. Since we deal with commutative Banach algebras, we simply call  $\mathcal{I}$  and *ideal*. A bilateral ideal  $\mathcal{I}$  is *maximal* if it is not contained in any other proper, maximal ideal.

An element  $A$  in a algebra  $\mathcal{A}$  with unit  $I$  is *invertible* if  $A^{-1}$  exists in  $\mathcal{A}$ :  $AA^{-1} = A^{-1}A = I$ .

We list in an exercise some basic, purely algebraic facts about algebras.

EXERCISE 6.6. *Let  $\mathcal{A}$  be an algebra. Prove the following.*

- (i) *Show that a left ideal that contains  $I$ , or an invertible element, is  $\mathcal{A}$ .*
- (ii) *Let  $\Phi : \mathcal{A} \rightarrow \mathcal{L}$  a homomorphism of algebras. Show that  $\ker \Phi$  is a bilateral ideal.*
- (iii) *If  $p$  is a multiplicative functional and  $\mathcal{A}$  has unity, then  $p(I) = 1$ .*
- (iv) *If  $A$  is invertible in the algebra with unity  $\mathcal{A}$  and  $p \in \text{Hom}(\mathcal{A}, \mathbb{C})$ , then  $p(A) \neq 0$ .*
- (v) *Conversely to (ii), if  $\mathcal{I}$  is a bilateral ideal of  $\mathcal{A}$ , then  $\mathcal{I}$  is the kernel of the algebra homomorphism*

$$\pi : \mathcal{A} \rightarrow \mathcal{A}/\mathcal{I} := \{\pi(A) = A + \mathcal{I} : A \in \mathcal{A}\}.$$

*On the quotient we define the product  $(A + \mathcal{I})(B + \mathcal{I}) = AB + \mathcal{I}$ , which is well defined (prove it).*

- (vi) *If  $\Phi \in \text{Hom}(\mathcal{A}, \mathcal{L})$  and  $\mathcal{I}$  is a bilateral ideal in  $\mathcal{L}$ , then  $\Phi^{-1}(\mathcal{I})$  is a bilateral ideal in  $\mathcal{A}$ .*
- (vii) *By Zorn's lemma, in a commutative algebra with identity a proper ideal  $\mathcal{I}$  is always contained in a maximal ideal.*

At the end of the section we will see that (iv) has an inverse in commutative Banach algebras: if  $p(A) \neq 0$  for all  $p \in \text{Hom}(\mathcal{A}, \mathbb{C})$ , then  $A$  is invertible.

We close this algebraic introduction with a lemma.

LEMMA 6.3. *Let  $\mathcal{M}$  be a maximal, bilateral ideal in  $\mathcal{A}$ , an algebra with unity. Then,  $\mathcal{L} = \mathcal{A}/\mathcal{M}$  is a **division algebra**: all nonzero elements in it have an inverse.*

PROOF. Let  $\pi : \mathcal{A} \rightarrow \mathcal{L}$  be the projection. By contradiction, let  $0 \neq \pi(B)$ , with  $\pi(B)$  noninvertible in  $\mathcal{L}$ , and let  $\mathcal{L} \supset \mathcal{I}$  be the ideal generated by  $B$ , which is proper since  $I + \mathcal{M} \notin \mathcal{I}$ . Now,  $\pi^{-1}(\mathcal{I})$  is a proper ideal of  $\mathcal{A}$ ,  $I \notin \pi^{-1}(\mathcal{I})$ , and it properly contains  $\mathcal{M}$ :  $B \notin \mathcal{M}$ , but  $B \in \pi^{-1}(\mathcal{I})$ . Hence,  $\mathcal{M}$  was not maximal.  $\square$

We now come to analytic-algebraic results.

LEMMA 6.4. *Let  $\mathcal{A}$  be a commutative Banach algebra, and  $p \in \text{Hom}(\mathcal{A}, \mathbb{C})$ . Then,  $|p(A)| \leq |A|$ .*

PROOF. Suppose that  $|p(A)| > 1$  for some  $A$ , and let  $B = \frac{A}{p(A)}$ , so that  $|B| < 1$ . Then,  $I - B$  is invertible (by a Neumann series), hence,

$$0 \neq p(I - B) = 1 - \frac{p(A)}{p(A)} = 0,$$

contradiction.  $\square$

The lemma self-improves to the following inequality.

COROLLARY 6.2. *Let  $\mathcal{A}$  be a commutative Banach algebra, and  $p \in \text{Hom}(\mathcal{A}, \mathbb{C})$ . Then,  $|p(A)| \leq r(A)$ .*

PROOF. We have

$$|p(A)| = |p(a^n)|^{1/n} \leq |A^n|^{1/n} \rightarrow r(A)$$

as  $n \rightarrow \infty$ .  $\square$

LEMMA 6.5. *Let  $\mathcal{I}$  be a proper ideal in a commutative Banach algebra  $\mathcal{A}$  with unity. Then, its closure  $\overline{\mathcal{I}}$  is a proper ideal.*

PROOF. The continuity of the algebra operations ensure that  $\overline{\mathcal{I}}$  is an ideal, hence it suffices to show that  $I \notin \overline{\mathcal{I}}$ . Since elements  $A$  with  $|I - A| < 1$  are invertible,  $\{A : |I - A| < 1\} \cap \mathcal{I} = \emptyset$ , and this shows that  $I \notin \overline{\mathcal{I}}$ .  $\square$

COROLLARY 6.3. *A maximal ideal in  $\mathcal{A}$  is closed.*

**6.4.2. The algebra of continuous functions.** We consider here the special case where  $\mathcal{A} = C(X)$  is the algebra of the continuous functions on a Hausdorff, compact space  $X$ , endowed with the uniform norm. In this case, the evaluations at point,

$$(6.4.1) \quad \eta_x : f \mapsto f(x)$$

is a multiplicative functional, and  $\ker(\eta_x) = \mathcal{M}_x = \{f \in C(X) : f(x) = 0\}$  is a maximal ideal in  $C(X)$ . Maximality, which we obtained by abstract reasoning, can be rephrased as follows. If  $a \in X$  and  $h \in C(X)$  is such that  $h(a) \neq 0$ , then there are  $\lambda \in \mathbb{C}$  and  $f \in C(X)$  with  $f(a) = 0$  such that  $1 = \lambda h + f$ . In fact, we obtain this by letting  $\lambda = 1/h(a)$  and  $f = 1 - h/h(a)$ .

**THEOREM 6.10.** *All maximal ideals in  $C(X)$  have the form  $\mathcal{M}_a$  for some  $a \in X$ .*

**PROOF.** By contradiction, let  $\mathcal{M}$  be a maximal ideal which is not of the form  $\mathcal{M}_a$  for any  $a$ . Then, for each  $x$  in  $X$ ,  $\mathcal{M}$  contains a function  $f_x$  with  $f_x(x) \neq 0$  (otherwise  $\mathcal{M}_x \subseteq \mathcal{M}$ ), and we can assume that  $f_x(x) > 0$ . After multiplying  $f_x$  times a cut-off function  $h_x$  in  $C(X)$ ,  $h_x f_x \in \mathcal{M}$ , we can also assume that  $h_x f_x \geq 0$  on  $X$  and  $h_x(x) f_x(x) > 0$ . Let  $U_x = \{y \in X : h_x f_x(y) > 0\}$ . By compactness,  $X = \bigcup_{i=1}^n U_{x_i}$ , and so, by Weierstrass theorem,

$$c \leq \sum_{i=1}^n h_{x_i} f_{x_i} \leq C.$$

That is,  $(\sum_{i=1}^n h_{x_i} f_{x_i})^{-1} \in C(X)$ , hence,

$$\mathcal{M} \ni \sum_{i=1}^n \frac{h_{x_i}}{\sum_{j=1}^n h_{x_j} f_{x_j}} f_{x_i} = 1.$$

This shows that  $\mathcal{M} = C(X)$ . □

**EXERCISE 6.7.** *Show that the maximal ideal space of the algebra  $C^1[0, 1]$  is formed by the elements  $\mathcal{M}_x \cap C^1[0, 1]$ , as  $x$  ranges in  $[0, 1]$ .*

**6.4.3. Intermezzo: the algebra of multipliers on a reproducing kernel Hilbert space.** Other notable examples of ideals in algebras are found in multiplier algebras  $\mathcal{M}(H)$  of RKHSs  $H$  on a set  $X$ :

$$(6.4.2) \quad \mathcal{M}_x = \{f \in \mathcal{M}(H) : f(x) = 0\},$$

where  $x \in X$ . Observe that  $\mathcal{M}_x$  is proper since  $1 \notin \mathcal{M}_x$ .

More generally, for  $x_1, \dots, x_n \in X$ , we can define

$$\mathcal{M}_{\{x_1, \dots, x_n\}} = \{f \in \mathcal{M} : f(x_1) = \dots = f(x_n) = 0\}.$$

If  $E \subset F \subset X$ , then  $\mathcal{M}_E \supseteq \mathcal{M}_F$ .

We recall a result from basic linear algebra.

**LEMMA 6.6.** *Let  $V$  be a vector space over  $\mathbb{C}$  and  $l : V \rightarrow \mathbb{C}$  be linear,  $l \neq 0$ . Then,  $\ker(l)$  is a maximal proper subspace of  $V$ .*

**PROOF.** Let  $u, v \in V \setminus \ker(l)$ . Then,  $l(u)v - l(v)u = w \in \ker(l)$ , i.e.  $v = \frac{l(v)}{l(u)}u + \frac{z}{l(u)} \in \text{span}(\ker(l), u)$ , showing that  $V = \text{span}(\ker(l), u)$ .  $\square$

**PROPOSITION 6.2.** *The evaluation map  $\eta_x : f \mapsto f(x)$  is a multiplicative functional, hence it is bounded. The ideal  $\mathcal{M}_x$  is maximal, hence closed.*

**PROOF.** The functional  $f \mapsto f(x)$  is well defined by definition of multiplier ( $f$  is a function) and it is multiplicative. Hence, it is a contraction by lemma 6.4. Its kernel,  $\mathcal{M}_x$ , is maximal by lemma 6.6.  $\square$

In the chapter on reproducing kernel Hilbert spaces we had a different proof that  $\eta_x$  is a contraction on  $\mathcal{M}(H)$ , which followed from the fact that the conjugates of values of  $f$ , if  $f$  is a multiplier, are eigenvalues of the multiplier operator's adjoint.

**6.4.4. Gelfand's representation.** We return to the general theory.

**THEOREM 6.11.** *Let  $\mathcal{I}$  be a proper, closed ideal in a commutative Banach algebra  $\mathcal{A}$ . Then,  $\mathcal{A}/\mathcal{I}$  is a Banach algebra with respect to its quotient norm as Banach space,*

$$|A + \mathcal{I}| := \inf\{|A + N| : N \in \mathcal{I}\}.$$

**PROOF.**

$$\begin{aligned} |AB + \mathcal{I}| &= \inf\{|AB + N| : N \in \mathcal{I}\} \\ &\leq \inf\{|AB + AN + BM + MN| : M, N \in \mathcal{I}\} \\ &\leq \inf\{|A + M| \cdot |B + N| : M, N \in \mathcal{I}\} \\ &= \inf\{|A + M| : M \in \mathcal{I}\} \inf\{|B + N| : N \in \mathcal{I}\} \\ &= |A + \mathcal{I}| \cdot |B + \mathcal{I}|. \end{aligned}$$

$\square$

Next, we have a useful characterization of the complex field.

**THEOREM 6.12 (Mazur).** *Let  $\mathcal{A}$  be a Banach algebra which unit and which is a division algebra. Then,  $\mathcal{A}$  is isometrically isomorphic to  $\mathbb{C}$ .*



PROOF. Let  $A \neq 0$  be an element in  $\mathcal{A}$ . The spectrum  $\sigma(A)$  is nonempty, hence there is  $\alpha$  such that  $\alpha I - A$  is not invertible. Since  $\mathcal{A}$  is a division algebra,  $A = \alpha I$ , and  $\sigma(A) = \{A\}$ . We then have a map  $\Phi : \mathcal{A} \rightarrow \mathbb{C}$ ,  $\Phi(A) = \Phi(\alpha I) = \alpha$ , which is an isometric algebra isomorphism.  $\square$

Finally can state and prove the inverse of property (iv) in exercise 6.6.

THEOREM 6.13. *Let  $\mathcal{A}$  be a commutative Banach algebra with unit.*

- (i) *If  $A$  is noninvertible, then there exists  $p \in \text{Hom}(\mathcal{A}, \mathbb{C})$  such that  $p(A) = 0$ .*
- (ii) *To each maximal ideal  $\mathcal{M}$  in  $\mathcal{A}$  there corresponds a unique  $p \in \text{Hom}(\mathcal{A}, \mathbb{C})$  such that  $\ker(p) = \mathcal{M}$ .*

PROOF. (i) By (vii) in exercise 6.6,  $A$  is contained in a maximal ideal  $\mathcal{M}$ . Then,

$$p_{\mathcal{M}} : \mathcal{A} \rightarrow \mathcal{A}/\mathcal{M} \equiv \mathbb{C}$$

provides an element in  $\text{Hom}(\mathcal{A}, \mathbb{C})$  with  $p_{\mathcal{M}}(A) = 0$ .

(ii) To each maximal  $\mathcal{M}$  we associate  $p_{\mathcal{M}}$ , and the functional  $p$  is uniquely determined since  $\mathcal{A} = \text{span}(\mathcal{M}, I)$  and  $p(I) = 1$ .  $\square$

Let  $\Omega = \Omega(\mathcal{A})$  be the set of the maximal ideal in  $\mathcal{A}$ , a commutative Banach algebra with unity. By theorem 6.13 we have a well defined map

$$(6.4.3) \quad p : \Omega \times \mathcal{A} \rightarrow \mathbb{C}, \quad p(\mathcal{M}, A) = p_{\mathcal{M}}(A).$$

We might look at it as a map  $\mathcal{G} : \mathcal{A} \mapsto p(\cdot, A) = p^A$ ,

$$(6.4.4) \quad \mathcal{G} : \mathcal{A} \rightarrow \Omega^{\mathbb{C}}, \quad \mathcal{G}(A) : \Omega \rightarrow \mathbb{C}, \quad [\mathcal{G}(A)](\mathcal{M}) = p_{\mathcal{M}}(A).$$

The map  $\mathcal{G}$ , representing elements of the Banach algebra  $\mathcal{A}$  as functions on its maximal ideal space  $\Omega$ , is the *Gelfand representation* of  $\mathcal{A}$ . We summarize the properties we have so far seen concerning it.

THEOREM 6.14. (i) *The Gelfand representation  $\mathcal{G}$  is a homomorphism of algebras,  $\mathcal{G} \in \text{Hom}(\mathcal{A}, \Omega^{\mathbb{C}})$ .*

(ii) *The representation is contractive in the sense that  $\sup_{\mathcal{M} \in \Omega} |[\mathcal{G}(A)](\mathcal{M})| \leq |A|$ .*

(iii)  $\sigma(A) = [\mathcal{G}(A)](\Omega) = \{p(\mathcal{M}, A) : \mathcal{M} \in \Omega\}$ .

(iv)  $\mathcal{G}(I) = 1$  is the constant function.

(v) *The functions  $\mathcal{G}(A)$  separate the points of  $\Omega$ : if  $\mathcal{M} \neq \mathcal{N}$  are maximal ideals, then  $[\mathcal{G}(A)](\mathcal{N}) = p(\mathcal{M}, A) \neq p(\mathcal{N}, A) = [\mathcal{G}(A)](\mathcal{N})$ .*

PROOF. □

The heart of Gelfand theory consists in making the representation sharp. In order to do this, we have to make  $\Omega$  into a topological space.

Consider on  $X$  the *Gelfand topology*  $\tau$ , the coarsest making all maps  $p_A : \Omega \rightarrow \mathbb{C}$  continuous. A basis of neighborhoods for  $\mathcal{M}_0 \in \Omega$  in  $\tau$  are finite intersections of the sets

$$N(A; U) = \{\mathcal{M} \in \Omega : p(\mathcal{M}, A) \in U\},$$

where  $A \in \mathcal{A}$  and  $U \subseteq \overline{D(0, |A|)}$  is open in the Euclidean topology of  $\overline{D(0, |A|)} = \{z \in \mathbb{C} : |z| \leq |A|\}$ .

**THEOREM 6.15.** *The maximal ideal space  $\Omega$  is compact with respect to the Gelfand topology.*

The proof is similar to the one of the Banach-Alaoglu theorem.

PROOF. Consider the space

$$Z = \Pi_{A \in \mathcal{A}} \overline{D(0, |A|)}$$

endowed with the product topology, which is the coarsest making the projections  $\pi_{A_0} : Z \rightarrow \overline{D(0, |A_0|)}$ ,  $\pi_{A_0}(\{z_A\}_{A \in \mathcal{A}}) = z_{A_0}$ . A basis of neighborhoods for the product topology are finite intersections of sets of the form

$$\mathcal{N}(A; U) = \{\{z_A\}_{A \in \mathcal{A}} : z_{A_0} \in U\},$$

with  $A_0$  and  $U$  as above. By Tikhonov theorem,  $Z$  is compact.

Consider then the map  $\Phi : \Omega \rightarrow Z$

$$\Phi(\mathcal{M}) = \{p(\mathcal{M}, A) : A \in \mathcal{A}\}.$$

By the way the topologies on  $\Omega$  and  $Z$  are defined,  $\Phi$  is a homomorphism onto its image. To show that  $\Omega$  is compact, then, it suffices to show that  $\Phi(\Omega)$  is closed in  $Z$ .

Suppose  $\{z_A\}_{A \in \mathcal{A}}$  lies in the closure of  $\Phi(X)$ , and consider  $A, B \in \mathcal{A}$ . For any  $\epsilon > 0$  there is  $\mathcal{M} \in X$  such that  $|p(\mathcal{M}, A) - z_A| < \epsilon/3$ ,  $|p(\mathcal{M}, B) - z_B| < \epsilon/3$ , and  $|p(\mathcal{M}, A+B) - z_{A+B}| < \epsilon/3$ , so that

$$|z_A + z_B - z_{A+B}| < \epsilon.$$

Hence,  $z_A + z_B = z_{A+B}$ . The same way, one shows that  $z_{cA} = cz_A$  if  $c$  is a complex number, and  $z_{AB} = z_A z_B$ . This shows that  $p : A \mapsto z_A$  is a multiplicative functional from  $\mathcal{A}$  to  $\mathbb{C}$ . We know that  $p = p(\mathcal{M}, \cdot)$  for some  $\mathcal{M}$  in  $\Omega$ , hence that  $\{z_A\}_{A \in \mathcal{A}} = \Phi(\mathcal{M})$ . □

**6.4.5. The Gelfand representation of  $C(X)$  and other examples.**

When  $\mathcal{A} = C(X)$ , the space of the continuous functions on a compact, Hausdorff space  $X$  with topology  $\sigma$ , the Gelfand representation is just the identity.

We saw that there is a bijection between  $X$  and the maximal ideal space,  $a \mapsto \mathcal{M}_a$ . The multiplicative functional associated to  $\mathcal{M}_a$  is evaluation at  $a$ ,  $\eta_a(f) = f(a)$ . the Gelfand transform, then acts as

$$[\mathcal{G}(f)](\mathcal{M}_a) = \eta_a(f) = f(a)$$

is nothing but  $f$  itself. The topology  $\tau$  on  $X$  interpreted as a maximal ideal space is the coarsest making each  $f \in C(X)$  continuous, hence  $\tau \subseteq \sigma$  and  $Id : (X, \sigma) \rightarrow (X, \tau)$  is continuous. But both  $\sigma$  and  $\tau$  make  $X$  compact, hence sets which are closed for  $\sigma$  are also compact for  $\sigma$ , then they are compact, hence closed for  $\tau$ . This shows that  $\sigma \subseteq \tau$ , hence that  $\sigma = \tau$ .

**EXERCISE 6.8.** *Show that the Gelfand representation applied to  $C^1[0, 1]$  just identifies  $C^1[0, 1]$  with a particular subalgebra of  $C[0, 1]$ .*

The space  $H^\infty[0, 1]$  is a Banach algebra, hence it has a Gelfand representation. Its maximal ideal space is huge, but there seems to be no concrete, known example of maximal ideal.

**6.5. Intermezzo: corona theorems****6.5.1. The corona problem in reproducing kernel Hilbert spaces.**

Let  $H$  be a reproducing Hilbert space of functions on a set  $X$ , and let  $\mathcal{M}(H)$  be its multiplier algebra. In many important cases, the multiplier algebra is too small even to separate points on  $X$ . In fact, it might very well be that  $\mathcal{M}(H)$  reduces to the constant functions. In many cases of interest, however, including the Hardy space  $H = H^2(\mathbb{D})$ , the multiplier algebra is large. For instance,  $f(z) = z$  already separates points in the disc.

Suppose  $\mathcal{M}(H)$  separates points of  $X$ , and consider its maximal ideal space  $\Omega = \Omega(\mathcal{M}(H))$ . Since point evaluations are multiplicative functionals, we have a map  $i : X \rightarrow \Omega$  associating to  $x$  the maximal ideal  $\mathcal{M}_x = \{f \in \mathcal{M}(H) : f(x) = 0\}$ . The map is injective because  $\mathcal{M}(H)$  separates points, and we can thus view  $X$  a subset of  $\Omega$ , which is compact with respect to the Gelfand topology.

Let  $\overline{X}$  be the closure of  $X$  in  $\Omega$ . The *corona* of  $\Omega(\mathcal{M}(H))$  is  $\Omega(\mathcal{M}(H)) \setminus \overline{X}$ . We say that  $\Omega(\mathcal{M}(H))$  has a corona if  $\Omega(\mathcal{M}(H)) \setminus \overline{X}$  is nonempty.

In 1941 Gelfand published his article on the representation of commutative Banach algebras, and in the same year Kakutani asked if  $H^\infty(\mathbb{D})$ , the multiplier space of  $H^2(\mathbb{D})$ , has a corona. The question remained open for twenty years, during which several reformulations of it were found. In 1962, Carleson showed that, as conjectured by Kakutani,  $H^\infty(\mathbb{D})$  has no corona.

Carleson proof was hard, and it involved a number of other foundational results. In 1979, Tom Wolff came with a much simplified proof, which he never published, but which has become the standard one in monographs<sup>2</sup>. This is not the end of the story, because the maximal ideal space of  $H^\infty(\mathbb{D})$  is *not* homeomorphic to the closed unit disc, as one might naively guess by comparison with the algebra of the continuous functions on a compact set. A concrete understanding of it came later.

After Carleson's result, the corona theorem was extended to many other multiplier spaces, and examples of multiplier spaces where it does not hold where found. The area is an active one, and new results and open problems abound. In this section, we limit ourselves to show that the corona theorem is equivalent to an analytic statement, which is the starting point of most of the "corona proofs" in literature. We do it for  $H^\infty$ , but the strategy has general value.

**6.5.2. An equivalent, analytic formulation of the corona problem.** Let  $\mathcal{G}(H^\infty(\mathcal{D}))$  be the maximal ideal space of  $H^\infty(\mathbb{D})$ . Each  $\mathcal{M}$  in it is the kernel of a unique multiplicative functional  $p_{\mathcal{M}} : H^\infty(\mathbb{D}) \rightarrow \mathbb{C}$ . For  $f$  in  $H^\infty(\mathbb{D})$ , we write

$$f(\mathcal{M}) := p_{\mathcal{M}}(f),$$

thus thinking of  $f$  as a function on  $\mathcal{G}(H^\infty(\mathcal{D}))$ .

LEMMA 6.7 (An equivalent statement of the corona theorem). *The following are equivalent.*

- (i) *The corona theorem holds in  $H^\infty(\mathbb{D})$ .*
- (ii) *Suppose that for all  $n \geq 1$ , if  $f_1, \dots, f_n \in H^\infty(\mathbb{D})$  there is  $j$  such that  $|f_j(z)| \geq \delta > 0$  for some  $\delta > 0$  independent of  $z \in \mathbb{D}$ . Then there exist  $g_1, \dots, g_n \in H^\infty(\mathbb{D})$  such that*

$$(6.5.1) \quad f_1 g_1 + \dots + f_n g_n = 1,$$

*i.e. the ideal generated by  $f_1, \dots, f_n$  is the whole of  $H^\infty(\mathbb{D})$ ,  $\mathcal{M}(f_1, \dots, f_n) = H^\infty(\mathbb{D})$ .*

The implication we need for proving the corona theorem is (ii)  $\implies$  (i). It is useful, however, knowing that the two properties are equivalent. Property (ii) is obvious if we drop the requirement that  $g_1, \dots, g_n$  be holomorphic:

$$1 = \frac{f_1 \overline{f_1} + \dots + f_n \overline{f_n}}{f_1 \overline{f_1} + \dots + f_n \overline{f_n}}$$

---

<sup>2</sup>See e.g. [Paul Koosis, Introduction to  \$H^p\$  Spaces](#), 2<sup>nd</sup> edition (1999) Cambridge University Press.

$$= f_1 \frac{\overline{f_1}}{\sum_{j=1} |f_j|^2} + \cdots + f_n \frac{\overline{f_n}}{\sum_{j=1} |f_j|^2},$$

and  $g_i = \frac{\overline{f_i}}{\sum_{j=1} |f_j|^2}$  is  $C^\infty$  and bounded on  $\mathbb{D}$  by our assumptions on the  $f_j$ 's.

PROOF. (ii)  $\implies$  (i). Suppose  $H^\infty(\mathbb{D})$  is not dense in  $\mathcal{G}(H^\infty(\mathbb{D}))$  with respect to the Gelfand topology. Then, there are a maximal ideal  $\mathcal{M}$ ,  $\delta > 0$ , and  $h_1, \dots, h_n \in H^\infty(\mathbb{D})$  such that, for all  $z \in \mathbb{D}$ ,  $\mathcal{M}_z \notin N(\mathcal{M}; h_1, \dots, h_n; \delta)$ , i.e.

$$|h_1(\mathcal{M}) - h_1(z)| \geq \delta, \text{ or } |h_2(\mathcal{M}) - h_2(z)| \geq \delta, \text{ or } \dots, \text{ or } |h_n(\mathcal{M}) - h_n(z)| \geq \delta.$$

So,  $f_j := h_j - h_j(\mathcal{M}) \in H^\infty(\mathbb{D})$ ,  $j = 1, \dots, n$ , satisfy

$$f_j(\mathcal{M}) = 0,$$

but for each  $z$  in the disc there is  $f_j$  such that  $|f_j(z)| \geq \delta$ . By (ii), we find  $g_1, \dots, g_n \in H^\infty(\mathbb{D})$  such that  $f_1 g_1 + \cdots + f_n g_n = 1$ , hence

$$1 = p_{\mathcal{M}}(1) = \sum_{j=1}^n p_{\mathcal{M}}(f_j) p_{\mathcal{M}}(g_j) = 0$$

and we have found a contradiction.

(i)  $\implies$  (ii). Suppose the corona theorem holds and let  $f_1, \dots, f_n$ , and  $\delta > 0$ , be as in (ii). Suppose by contradiction that for no  $g_1, \dots, g_n \in H^\infty(\mathbb{D})$  we have that  $f_1 g_1 + \cdots + f_n g_n = 1$ , i.e. that

$$\mathcal{M}(f_1, \dots, f_n) = \{f_1 g_1 + \cdots + f_n g_n : g_1, \dots, g_n \in H^\infty(\mathbb{D})\}$$

is a proper ideal in  $H^\infty(\mathbb{D})$ . Let  $\mathcal{M}$  be a maximal ideal in  $H^\infty(\mathcal{D})$  which contains  $\mathcal{M}(f_1, \dots, f_n)$ , and let  $p_{\mathcal{M}}$  be the corresponding multiplicative functional, so that  $p_{\mathcal{M}}(f_j) = 0$  for  $j = 1, \dots, n$ . Since the corona theorem holds, there exists  $z \in \mathbb{D}$  such that  $\mathcal{M}_z \in N(\mathcal{M}; f_1, \dots, f_n; \delta)$ , i.e. such that for  $j = 1, \dots, n$ ,

$$\delta > |p_{\mathcal{M}}(f_j) - p_{\mathcal{M}_z}(f_j)| = |f_j(z)|,$$

which contradicts our assumption on  $f_1, \dots, f_n$ .  $\square$

A critical analysis of the proof shows that we used very little of the structure of  $H^\infty(\mathcal{D})$ .

**THEOREM 6.16.** *Let  $\mathcal{A}$  be a Banach algebra of functions  $f : X \rightarrow \mathbb{C}$ , defined on some set  $X$ , and suppose that the **evaluation functionals**  $\eta_x : \mathcal{A} \rightarrow \mathbb{C}$ ,*

$$\eta_x(f) = f(x)$$

*are bounded on  $\mathcal{A}$  for all  $x$  in  $X$ . Then, the following are equivalent.*

(i) The set  $\{\ker(\eta_x) : x \in X\}$  is dense in  $\mathcal{G}(\mathcal{A})$ .

(ii) Suppose that for all  $n \geq 1$ , if  $f_1, \dots, f_n \in \mathcal{A}$  there is  $j$  such that  $|f_j(x)| \geq \delta > 0$  for some  $\delta > 0$  independent of  $x \in X$ . Then there exist  $g_1, \dots, g_n \in \mathcal{A}$  such that

$$(6.5.2) \quad f_1 g_1 + \dots + f_n g_n = 1,$$

i.e. the ideal generated by  $f_1, \dots, f_n$  is the whole of  $\mathcal{A}$ ,  $\mathcal{M}(f_1, \dots, f_n) = \mathcal{A}$ .

The theorem can be in particular be applied to the multiplier space of a reproducing kernel Hilbert space. In fact, most proofs of corona theorems try to show (ii), which is an analytic statement.

EXERCISE 6.9. Prove theorem 6.16.

The case  $n = 1$  in (6.5.2) is already an interesting one. It says that if  $f$  lies in the multiplier space, and  $\mathcal{M}(H)$  and  $|f|$  is bounded from below, then  $1/f \in \mathcal{M}(H)$ .

## CHAPTER 7

### Spectral theory of unitary, and of self-adjoint operators

Most of this chapter is based on the expository [The spectral theorem and its converses for unbounded symmetric operators \(2011\)](#) by Terence Tao. The advantage of this approach is that we do not need any previous version of the spectral theorem: not that for bounded self-adjoint operators, or more generally that for the normal operators, and not even that for the finite dimensional case. In fact, the self-adjoint bounded case, and the self-adjoint compact case, can be recovered as special cases, the peculiarities of which can be easily spotted and proved using rather basic measure theory. The pedagogical risk, on the other hand, is that the unbounded case presents many subtle points which have to be untangled before reaching the heart of the matter.

A good source for a more comprehensive treatment of spectral theory using Herglotz representation is [Mathematical Methods in Quantum Mechanics With Applications to Schrödinger Operators \(2014\)](#) by Gerald Teschl (Graduate Studies in Mathematics, Volume 157, Amer. Math. Soc.). A new approach to the whole matter is in [The Functional Calculus Approach to the Spectral Theorem \(2020\)](#) by Markus Haase, [arXiv:2003.06130](#). Another recommended reading is Michael Taylor's [Spectral Theory](#).

A spectral theorem with a simpler proof is the one for unitary operators. I have included it because the result is important by itself, and it can serve as a starter, its proof having several ideas in common with the one for self-adjoint operators. The variant I present here is from Michael Taylor's lecture notes [The Spectral Theorem for Self-Adjoint and Unitary Operators](#).

The spectral theorem which is considered here tells the life and works of *one* self-adjoint operator  $L$ , and of a commutative Banach algebra it generates (in an appropriate way). In fact, the generated algebra is a *von Neumann algebra*. Here is a short list of topics we do not cover.

- (1) If  $L, M$  are *commuting* bounded, self-adjoint operators,  $LM = ML$ , can they be simultaneously represented in the same model (can they be simultaneously diagonalized)? The answer is yes, but you have to find it elsewhere.

- (2) What about normal operators? See Michael Reed, Barry Simon, [Methods of Modern Mathematical Physics: Functional analysis](#), Academic Press, 1972, chapter VII, exercise 5 p. 246.
- (3) On the way towards our spectral theorem we will make a number of arbitrary choices. Is there a *canonical way* to represent a self-adjoint operator? Yes, up to a point, but you have to learn it somewhere else (e.g. Reed and Simon's Functional Analysis).
- (4) An  $N \times N$  Hermitian matrix is not in general unitarily determined by its eigenvalues: we also need to know their multiplicity. Is there a multiplicity theory even in the more general Hilbert universe? Again, the answer is positive, but for the details you have to find other sources. e.g. [Introduction to Hilbert Space and the Theory of Spectral Multiplicity](#), New York, Chelsea Pub. Co. 1957, by Paul R. Halmos.

### 7.1. Intermezzo: spectral theory and quantum mechanics

In this "motivational" section I try to sketch the reason why the notions of spectral theory are foundational in quantum mechanics. I do everything in a finite dimensional Hilbert space, so what is needed to follow is just some notion of linear algebra, and you do not need to know the theory which lies ahead. You also might read the section after you have read the rest of the chapter. Or you might skip altogether, if you so wish.

Spectral theory had its origins in the study of integral equations, with foundational contributions of Fredholm, then Hilbert, and it first appeared in its infinite dimensional form. In the course of the 1920's, physicists found that it provided the mathematical framework for quantum mechanics, which was being developed in those years, and both finite and infinite dimensional theories play there a role. It is a really amazing coincidence that the word *spectrum* introduced by Hilbert in his work turned out to represent, in the light of the quantum mechanical interpretation, the spectrum of electromagnetic radiation physicists were accustomed to. It was von Neumann who wrote the foundations of quantum mechanics in the language of abstract Hilbert spaces, which has since become standard.

I know very little physics, and the interested student is invited to compare what is to follow with some authoritative source. I suggest the video lectures [The theoretical minimum: quantum mechanics](#) by Leonard Susskind, or, even better, the book with the same title by Susskind and Art Friedman (2014). I like the book because: (i) the mathematics involved are kept to a minimum, so the math student can quickly browse through the preliminaries of each chapter and speedily get to (ii) the conceptual exposition of



the physics, which are presented in rigorous terms, (iii) in a way which is not encumbered by a long discussion of experimental results and history of concepts. Susskind's text, that is, is a fast vehicle for those who want to proceed from operator theory to its motivations and applications in theoretical physics.

Of course, at some point experiments and history of concepts have to be considered, as always in natural sciences. For instance, the [Lecture notes on operator algebras and their application in physics](#) (2022) by Alexander Kegeles provide a more advanced, but nice and readable, introduction to the quantum mechanical interpretation of operator theory, and to operator theory itself, for Master students in mathematics.

I thank Giacomo De Palma for correcting a conceptual mistake in the definition of "complete system of commuting observables".

**7.1.1. States and observables.** A *physical system* is described by a (complex!) Hilbert space  $H$ , whose vectors represent the *states* of the system (or, better: states are vectors of unitary norm). Here we consider  $H$  to be finite dimensional. Any self-adjoint operator  $A$  represents an *observable* of such system, and its eigenvalues  $\lambda_i$ 's are the possible *measures* of the observable:  $\lambda_i$  is the value of the observation represented by  $A$  when the system is in a state  $e \in E(\lambda_i)$ , where  $E(\lambda_i)$  is, recall, the eigenspace relative to  $\lambda_i$ .

Consider two states  $x, y$ . They are *distinguishable* if there is an observable  $A$  which returns (with certainty: see below) different values for the two states. i.e. there is a self-adjoint operator  $A$  such that  $Ax = \lambda x$ ,  $Ay = \mu y$ ,  $\lambda \neq \mu$ . This happens if and only if  $x$  and  $y$  are orthogonal:  $\langle x|y \rangle = 0$ .

More generally, if the system lies in the state  $h \in H$ ,  $\|h\| = 1$ , then the *probability* that the measure  $X$  (a random variable) of the observable  $A$  has the value  $\lambda_i$  is

$$(7.1.1) \quad \mathbb{P}_h(X = \lambda_i) = \|\Pi_{E(\lambda_i)} h\|^2.$$

The states, that is, encode probabilities of observables. If  $h = \alpha k$  with  $\alpha \in \mathbb{C} \setminus \{0\}$ ,  $h$  and  $k$  return the same probabilities, and they can not be distinguished. The state space really is the projectivization  $(H \setminus \{0\})/\mathbb{C}$  of  $H$ . Many properties of the system, however, depend on the linear structure of  $H$ . Some examples of this will be presented below.

With this probabilistic interpretation, the *expected value* (or *mean*) of the measure in the state  $h$  is

$$\mathbb{E}_h(X) = \sum_i \lambda_i \mathbb{P}_h(X = \lambda_i) = \sum_i \lambda_i \|\Pi_{E(\lambda_i)} h\|^2$$

$$\begin{aligned}
(7.1.2) \quad &= \left\langle h, \sum_{i=1}^L \lambda_i \Pi_{E(\lambda_i)} h \right\rangle \\
&= \langle h, Ah \rangle.
\end{aligned}$$

The variance of  $X$  is called the *uncertainty* of  $A$ , and it can be expressed in a number of ways,

$$\begin{aligned}
(7.1.3) \quad \text{Var}_h(X) &= \mathbb{E}_h[X - \mathbb{E}_h X]^2 \\
&= \sum_i [\lambda_i - \sum_j \lambda_j \mathbb{P}_h(X = \lambda_j)]^2 \mathbb{P}_h(X = \lambda_i) \\
&= \mathbb{E}_h X^2 - (\mathbb{E}_h X)^2 \\
&= \langle h, A^2 h \rangle - \langle h, Ah \rangle^2 \\
&= \langle h, [A - \langle h, Ah \rangle I]^2 h \rangle \\
&= \langle h, A^2 h \rangle,
\end{aligned}$$

where the last expression holds if we put the mean of  $A$  to zero.

For vectors  $h$  in  $E(\lambda_i)$  there is no uncertainty:  $\mathbb{P}_h(X = \lambda_i) = 1$ ,  $\mathbb{P}_h(X = \lambda_j) = 0$  for  $j \neq i$ . Such vectors are called *pure states* for the observable  $A$ . A consequence of the spectral theorem for self-adjoint matrices is that there is a orthonormal bases of pure states; any state is *super-position* (linear combination) of pure states.

**7.1.2. Commuting observables.** After the measure of  $A$  in the state  $h$  is performed, and the eigenvalue  $\lambda_i$  is its outcome, the system *collapses* to the state  $\Pi_{E(\lambda_i)} h$  (which has to be properly normalized by considering  $\Pi_{E(\lambda_i)} h / \|\Pi_{E(\lambda_i)} h\|$ )<sup>1</sup>. Suppose we have a second observable  $B$ , with associated random variable  $Y$ , eigenvalues  $\mu_k$ , and eigenspaces  $F(\mu_k)$ . If the initial state is  $h$ , and the measurement  $A$  is performed with outcome  $\lambda_i$ , then the observation of  $B$  is performed in the state  $\Pi_{E(\lambda_i)} h$ . By (7.1.1), this means that:

$$\begin{aligned}
(7.1.4) \quad \mathbb{P}_h(Y = \mu_k | X = \lambda_i) &= \mathbb{P}_{\Pi_{E(\lambda_i)} h}(Y = \mu_k) \\
&= \frac{\|\Pi_{F(\mu_k)} \Pi_{E(\lambda_i)} h\|^2}{\|\Pi_{E(\lambda_i)} h\|^2}, \text{ and} \\
\mathbb{P}_h(Y = \mu_k \text{ after } X = \lambda_i) &= \frac{\|\Pi_{F(\mu_k)} \Pi_{E(\lambda_i)} h\|^2}{\|\Pi_{E(\lambda_i)} h\|^2}.
\end{aligned}$$

Here we can see a peculiar feature of quantum theory. If the measurement of  $A$  did not affect the system in those parts measured by  $B$ , we would have

<sup>1</sup>You might be worried about the denominator: where does  $h$  collapse when  $\Pi_{E(\lambda_i)} h = 0$ ? That case has null probability,  $\mathbb{P}_h(X = \lambda_i) = \|\Pi_{E(\lambda_i)} h\|^2 = 0$ , and we can simply ignore it.

equality from first to second line in:

$$\begin{aligned}
 \|\Pi_{F(\mu_k)}h\|^2 &= \mathbb{P}_h(Y = \mu_k) \\
 &= \sum_i \mathbb{P}_h(Y = \mu_k \text{ after } X = \lambda_i) \\
 (7.1.5) \qquad &= \sum_i \|\Pi_{F(\mu_k)}\Pi_{E(\lambda_i)}h\|^2,
 \end{aligned}$$

simply because the events  $\{X = \lambda_i\}$  exhaust all possible outcomes of the measure of  $A$ . This is certainly the case if  $[A, B] := AB - BA = 0$ , since in that case

$$(7.1.6) \qquad \Pi_{F(\mu_k)}\Pi_{E(\lambda_i)} = \Pi_{E(\lambda_i)}\Pi_{F(\mu_k)},$$

and the equality between the first and the last terms in (7.1.5) would follow from the fact that  $\{\Pi_{E(\lambda_i)}\}_i$  is pvm, so that

$$(7.1.7) \qquad \mathbb{P}_h(Y = \mu_k) = \sum_i \mathbb{P}_h(Y = \mu_k \text{ after } X = \lambda_i).$$

We say that the the observables represented by  $A$  and  $B$  are *compatible* if (7.1.7) holds. We wrote "after", not "and", because the system, "after" measuring  $X$ , is in a different state. The event  $\{Y = \mu_k, \text{ in the state } h\}$ , that is, is not a priori (and typically is not, as we shall see below), the union  $\cup_i \{Y = \mu_k \text{ after } X = \lambda_i, \text{ in the state } h\}$ .

In the finite dimensional case, compatibility is easily characterized in terms of spectral theory. The extension of this and other notion to infinite dimensions requires much more technicalities, although one can guess what the statements are.

**THEOREM 7.1 (Compatibility theorem).** *The following are equivalent.*

- (i) *The observables represented by  $A$  and  $B$  are compatible.*
- (ii)  *$A$  and  $B$  commute.*
- (iii) *There is a common system of eigenvectors for  $A$  and  $B$  (i.e.  $A$  and  $B$  can be simultaneously diagonalized).*

**PROOF.** (ii)  $\implies$  (i). By the spectral theorem,  $[A, B] = 0$  if and only if  $[f(A), B] = 0$  for all  $f : \sigma(A) \rightarrow \mathbb{C}$ , if and only if  $[f(A), g(B)] = 0$  for all  $f : \sigma(A) \rightarrow \mathbb{C}$  and  $g : \sigma(B) \rightarrow \mathbb{C}$ , which implies that  $[\Pi_{F(\mu_k)}, \Pi_{E(\lambda_i)}] = 0$  for all  $i, j$  (and by inserting this in the spectral resolutions of  $A$  and  $B$ , we see that in fact the last condition is equivalent to (ii)). We have observed above that (i) follows.

(iii)  $\implies$  (ii) is clear: with the respect to the common eigenbasis, both  $A$  and  $B$  are diagonal matrices.

(i)  $\implies$  (iii). Let  $x$  be a vector in  $F(\mu_k)$ . By (7.1.5),

$$\begin{aligned} \sum_i \|\Pi_{F(\mu_k)} \Pi_{E(\lambda_i)} x\|^2 &= \|\Pi_{F(\mu_k)} x\|^2 \\ &= \|x\|^2 = \sum_i \|\Pi_{E(\lambda_i)} \Pi_{F(\mu_k)} x\|^2 \\ &= \sum_i \|\Pi_{E(\lambda_i)} x\|^2. \end{aligned}$$

Each summand of the first term is less or equal to each sum of the last term, then

$$\|\Pi_{F(\mu_k)} \Pi_{E(\lambda_i)} x\|^2 = \|\Pi_{E(\lambda_i)} x\|^2.$$

This can only happen if  $\Pi_{E(\lambda_i)} x \in F(\mu_k)$ . i.e.  $\Pi_{E(\lambda_i)} : F(\mu_k) \rightarrow F(\mu_k)$  is the projection onto  $F(\mu_k) \cap E(\lambda_i)$ . Since the latter spaces form an orthogonal decomposition of  $F(\mu_k)$ , there is an orthonormal basis  $\{e_j^k\}_{j=1}^{\dim(F(\mu_k))}$  of  $F(\mu_k)$ , with respect to which each  $\Pi_{E(\lambda_i)}|_{F(\mu_k)}$  is diagonal. Hence, with respect to the basis  $\bigcup_k \{e_j^k\}_{j=1}^{\dim(F(\mu_k))}$  both  $A$  and  $B$  are diagonal.  $\square$

**COROLLARY 7.1.** *Two observables  $A$  and  $B$  are compatible if and only if for all states  $h$ :*

$$\mathbb{P}_h(Y = \mu_k \text{ after } X = \lambda_i) = \mathbb{P}_h(X = \lambda_i \text{ after } Y = \mu_k).$$

**EXERCISE 7.1.** *Prove corollary 7.1.*

**7.1.3. Complete systems of commuting observables.** The space  $H$  expresses all possible states of the system, and two states are physically distinct if it is possible to devise a measurement which assumes different values (or, at least, different expectations of such values) if the system is in one or the other state. In modelling physical systems, it is important that there are no redundancies: the state space  $H$  must be large enough to accommodate all physical states of the system, and no more.

In the finite dimensional case, a single observable  $A$  does the job if and only if its eigenvalues are simple,  $m_1 = \dots = m_L = 1$ . If  $m_i \geq 2$ , in fact, there are distinct pure states  $h, k \in E(\lambda_i)$  which return the same values for the observable  $A$ . The problem arises, then, of understanding when a set  $\{A_1, \dots, A_M\}$  of commuting observables, possibly having eigenvalues with higher multiplicities (that is, some eigenvalues are *degenerate*) has the property that, for each eigenvalue  $\lambda$ , the intersection  $E_1(\lambda) \cap \dots \cap E_M(\lambda)$  of the eigenspaces  $E_j(\lambda)$ 's corresponding to the  $A_j$ 's is at most one dimensional.

Such set  $\{A_1, \dots, A_M\}$  is called a *complete system of commuting operators*. In the finite dimensional theory these problems are elementary. In the infinite dimensional case they are less elementary and they are the subject of the *multiplicity theory*, which is best seen having at hands the multiplication form of the spectral theorem, which we will see later on.

We see here how the multiplication form of the spectral theorem looks in the finite dimensional case. We saw in (6.2.6) and (6.2.7) the pvm associate to a self-adjoint operator on a finite dimensional  $H$ . If  $\dim(E(\lambda_i)) = 1$  for all  $i$ 's, and  $e_i$  is a unitary vector in  $E(\lambda_i)$ , map  $U : e_i \mapsto (f_i : \lambda_j \mapsto \delta_{ij})$ , where  $\delta_{ij}$  is the Kroenecker symbol,

$$Ue_i(\lambda_j) = \delta_{i,j}, \quad U\left(\sum_i a_i e_i\right)(\lambda_j) = a_j,$$

defines a unitary map from  $H$  to  $L^2(\mathbb{R}, \mu)$ , where  $\mu = \sum_i \delta_i$ . We have then:

$$(7.1.8) \quad (UAU^{-1}\varphi)(\lambda_j) = \lambda_j \varphi(\lambda_j),$$

i.e. the operator  $A$  is unitarily equivalent to the operator "multiplication times  $\lambda$ " on  $L^2(\mu)$ . Let's verify (7.1.8).

$$\begin{aligned} \varphi &= \sum_i \varphi(\lambda_i) f_i, \\ U^{-1}\varphi &= \sum_i \varphi(\lambda_i) e_i, \\ AU^{-1}\varphi &= \sum_i \lambda_i \varphi(\lambda_i) e_i, \\ (UAU^{-1}\varphi)(\lambda_j) &= \sum_i \lambda_i \varphi(\lambda_i) (f_i(\lambda_j)) \\ &= \lambda_j \varphi(\lambda_j). \end{aligned}$$

In the general case, eigenvalues  $\lambda_i$  can have multiplicity  $m_i \geq 2$ , and we have to split the above procedure accordingly. Let  $m = \max(m_1, \dots, m_M)$  be the maximum of multiplicities. Rearrange so that  $m = m_1 \geq m_2 \geq \dots \geq m_M$ . For each eigenspace  $E(\lambda_k)$  consider a orthonormal basis  $\{e_j^k\}_{j=1}^{m_k}$  (the choice of which is not canonical); pick  $m$  copies of  $\mathbb{R}$ , and on the  $l^{th}$  copy consider the measure

$$\mu_l = \sum_{k:m_k \geq l} \delta_{\lambda_k}.$$

This way,  $\delta_{\lambda_k}$  appears in the measures  $\mu_1, \dots, \mu_{m_k}$ , and not in  $\mu_{m_k+1}, \dots, \mu_m$ . Let  $k(l)$  to be the largest  $k$  for which  $m_k \geq l$ .

Pick then a vector  $x = \sum_k \sum_{i=1}^{m_k} a_i^k e_i^k$ . Associate to it the family  $Ux := \{\varphi_l\}_{l=1}^m \in L^2(\mu_1) \oplus \cdots \oplus L^2(\mu_m)$ , where

$$\varphi_l = \sum_{k:m_k \geq l} a_l^k f_l^k,$$

and where, as before,  $f_l^k(\lambda_j) = \delta_{k,j}$  for  $1 \leq j \leq k(l)$ . This provides a constructive proof of the following.

**THEOREM 7.2** (Spectral theorem finite dimensions: multiplicative form). *Let  $A$  be a self-adjoint operator on a finite dimensional space. Then, there are measures  $\mu_1, \dots, \mu_M$  on  $\mathbb{R}$  (sums of Dirac deltas at disjoint real points), and there is a unitary operator*

$$U : H \rightarrow L^2(\mu_1) \oplus \cdots \oplus L^2(\mu_M), \quad Ux = \varphi = (\varphi_1, \dots, \varphi_M),$$

such that

$$(UAU^{-1}\varphi)_l(\lambda) = \lambda\varphi_l(\lambda).$$

Moreover,  $\dim(H)$  is the total number of the Dirac's deltas involved.

We can rephrase the theorem in a language involving just one measure space. Let  $X = \mathbb{R} \times \{1, \dots, M\}$ , endowed with the topology generated by the open sets of each factors. Identify  $\mu_j$  with the corresponding measure on  $\mathbb{R} \times \{j\}$ , and let  $\mu = \mu_1 + \cdots + \mu_M$ , which is a measure on  $X$ . Consider then the unitary map

$$V : L^2(\mu_1) \oplus \cdots \oplus L^2(\mu_M) \rightarrow L^2(\mu), \quad V(\varphi_1, \dots, \varphi_M)(\lambda, j) = \varphi_j(\lambda).$$

Then,

$$VUAU^{-1}V^{-1}\psi(\lambda, j) = g(\lambda, j)\psi(\lambda, j), \quad \text{where } g(\lambda, j) = \lambda.$$

This is a special case of the multiplicative form of the spectral theorem, stating that bounded, self-adjoint operators on Hilbert spaces can be always be represented as multiplication operators on (Borel)  $L^2$  spaces.

The presence of several summands is indication of how  $\{A\}$  alone fails to be a complete system of commuting observables. Multiplicity theory provides a more quantitative measure of the same.

**7.1.4. Uncertainty.** We saw that commuting self-adjoint operators  $A$  and  $B$  are compatible in the sense the measure of  $A$  does not affect the measure of  $B$ . *Compatibility* should not be confused with *independence* of the random variables  $X, Y$  associated with  $A, B$  with respect to the probability measure  $\mathbb{P}_h$  associated with a state  $h \in H$ , as the case  $A = B$  clearly illustrates.

Lack of commutativity, however, has probabilistic consequences.

**THEOREM 7.3** (Heisenberg inequality). *Suppose that  $A$  and  $B$  are self-adjoint and that they have zero mean with respect to a state  $h$ ,  $0 = \langle h, Ah \rangle = \langle h, Bh \rangle$ . Then,*

$$(7.1.9) \quad 1/4 |\langle h, [A, B]h \rangle|^2 \leq \langle h, A^2 h \rangle \langle h, B^2 h \rangle = \text{Var}_h(X) \text{Var}_h(Y).$$

Actually, we will prove a slightly sharper inequality. The main consequence of (7.1.9) is that  $[A, B]$  sets a lower bound to the simultaneous concentration of  $X$  and  $Y$  around their means.

**PROOF.** We set  $\langle A \rangle_h = \langle h, Ah \rangle$  (in physics the subscript  $h$  is implicit). If  $A$  and  $B$  are self-adjoint,  $[A, B]$  is a measure of how much  $AB$  fails to be self-adjoint:  $AB - (AB)^* = AB - BA = [A, B]$ . We can write  $AB$  in "algebraic form",

$$AB = \frac{\{A, B\}}{2} + i \frac{[A, B]}{2i} = S + iT, \text{ with } S = S^* \text{ and } T = T^*.$$

where we use the notation  $\{A, B\} = AB + BA$ . An easy calculation gives

$$\langle h, 1/2\{A - \langle A \rangle_h, B - \langle B \rangle_h\}h \rangle = \langle h, [1/2\{A, B\} - \langle A \rangle_h \langle B \rangle_h]h \rangle,$$

and clearly

$$\langle h, [A - a, B - b]h \rangle = \langle h, [A, B]h \rangle$$

for all constants  $a, b$ . We have then:

$$(7.1.10) \quad \begin{aligned} |1/(2i) \langle h, [A, B]h \rangle|^2 &\leq |\langle h, [1/2\{A, B\} - \langle A \rangle_h \langle B \rangle_h]h \rangle|^2 + |1/(2i) \langle h, [A, B]h \rangle|^2 \\ &= |\langle h, (A - \langle A \rangle_h)(B - \langle B \rangle_h)h \rangle|^2 \\ &= |\langle (A - \langle A \rangle_h)h, (B - \langle B \rangle_h)h \rangle|^2 \\ &\leq \|(A - \langle A \rangle_h)h\|^2 \|(B - \langle B \rangle_h)h\|^2 \\ &= \text{Var}_h(X) \text{Var}_h(Y), \end{aligned}$$

□

The stronger inequality, starting from the right hand side of the first inequality in (7.1.10), is known as the *Robertson-Schroedinger uncertainty relation*. One of its advantages is that it does not become trivial in the case  $[A, B] = 0$ ; e.g. when  $A = B$ . In a sense it carries both classical and quantum probabilistic information.

**7.1.5. Evolution and the Schroedinger equation.** Quantum physical systems evolve, and, if they are isolated, they evolve according to the following rules. First a bit of notation. If  $x$  is a state of the system,  $U(t)x$  is the state of that system at time  $t > 0$ . This way,  $U(0) = I$  is the identity.

(a) Distinguishable states remain distinguishable:

$$\langle x, y \rangle = 0 \implies \langle U(t)x, U(t)y \rangle = 0.$$

(b) If  $\|x\| = 1$ , we normalize  $U(t)x$  to have  $\|U(t)x\| = 1$ .

(c)  $U(t+s)x = U(t)U(s)x$ , which is self-explanatory: "the evolution of the evolution..."

(d)  $U(t)$  is linear. Together with (b), this implies that  $\|U(t)x\| = \|x\|$ . Linearity is a nontrivial requirement, whose roots are in experiments.

(e) Some continuity is required...

In purely mathematical terms, (a-e) mean:

- (i) each  $U(t)$  is a unitary operator,  $U(t)^*U(t) = U(t)U(t)^* = I$ ;
- (ii)  $t \mapsto U(t)$  defines a semigroup of such operators;
- (iii) some continuity...

We consider here the special case where

$$(7.1.11) \quad U(t) = e^{itA}, \quad t \geq 0, \quad \text{with } A \text{ bounded.}$$

**THEOREM 7.4.** *Let  $A$  be a self-adjoint operator on a Hilbert space  $H$ .*

- (i)  $U(t)$  is unitary.
- (ii)  $t \mapsto U(t)$  is a semi-group for all bounded operators  $A$ .
- (iii) The Schroedinger equation

$$(7.1.12) \quad \frac{1}{i} \frac{dU(t)}{dt} = AU(t)$$

*holds in the norm topology:*

$$\lim_{\Delta \rightarrow 0} \frac{U(t+\Delta) - U(t)}{\Delta} = AU(t)$$

*in the operator norm.*



(iv) *The semigroup is **strongly continuous**,*

$$(7.1.13) \quad \lim_{\Delta \rightarrow 0} \|U(\Delta)x - x\| = 0$$

*for all  $x$  in  $H$ .*

PROOF. Since the functions  $z \mapsto e^{tz}$  are entire, we can use the holomorphic calculus. We can write

$$U(t) = e^{itA} = \sum_{n=0}^{\infty} \frac{(itA)^n}{n!},$$

where the series converges in the operator norm, and all usual manipulations of series hold, provided we take into account the eventual break of commutativity (which here does not occur). We have  $\|U(t)\| \leq e^{t\|A\|}$ . Item (ii) follows immediately.

It is easily verified that

$$U(t)^* = e^{-itA}.$$

and this gives (i):

$$\begin{aligned} \langle U(t)x, U(t)y \rangle &= \langle U(t)^*U(t)x, y \rangle \\ &= \langle x, y \rangle. \end{aligned}$$

Since  $U(t + \Delta) = U(\Delta)U(t)$ , it suffices to verify (iii) at  $t = 0$ , which we can do using series:

$$\begin{aligned} \frac{1}{i} \frac{U(\Delta) - I}{\Delta} - A &= \sum_{n=1}^{\infty} \frac{A^n (i\Delta)^{n-1}}{n!} - A \\ &= (i\Delta) \sum_{n=2}^{\infty} \frac{A^n (i\Delta)^{n-2}}{n!}, \end{aligned}$$

which tends to zero in the operator norm as  $\Delta \rightarrow 0$ . (iv) is an even weaker statement.  $\square$

All this is very nice, but also highly unsatisfactory. The Schroedinger equation for a free particle, in fact, is (but for the Plank constant) of the form (7.1.12), but with  $H = L^2(\mathbb{R}^3)$  and  $A = \Delta_{\mathbb{R}^3}$ , the Laplace operator, which is *not* bounded (and not even everywhere defined, in fact) on  $H$ . This provided the main motivation for von Neumann to develop the spectral theory for unbounded self-adjoint operators, which also has applications all over mathematics. One of main items in the unbounded case is the following theorem of Stone, which extends theorem 7.4.

**THEOREM 7.5** (Stone's theorem). *Let  $t \mapsto U(t)$  be a strongly continuous semigroup of unitary operators on  $H$  (that is, (iv) in theorem 7.4 holds). Let  $D(A)$  be the set of those  $x$  in  $H$  such that*

$$\lim_{\Delta \rightarrow 0} \frac{U(\Delta)x - x}{i\Delta} = Ax$$

*exists in norm.*

*Then,  $D(A)$  is dense in  $H$ . The operator  $A : D(A) \rightarrow H$  is self-adjoint, and  $U(t) = e^{itA}$ .*

*Viceversa, if  $A$  is self-adjoint (and possibly unbounded), then  $U(t) = e^{itA}$  defines a strongly continuous unitary semigroup.*

Of course, we should define what an *unbounded self-adjoint operator* is, and what we mean by  $e^{itA}$ , which requires a more general version of the spectral theorem. Such tools were developed by von Neumann, in fact, in the late 20's of the XX century, and they are the main topic we discuss in this chapter.

If the system starts in the state  $x$  and the dynamics are given by  $A$ , at time  $t$  the system is in the state  $e^{itA}x$ . Suppose we wish to perform a measure of the observable  $B$  at time  $t$ . How does the expectation of the random variable  $Y$  associate to  $B$  change over time?

$$\mathbb{E}_{U(t)x} Y = \langle U(t)x, BU(t)x \rangle = \langle x, U(-t)BU(t)x \rangle = \langle x, e^{-itA}Be^{itA}x \rangle,$$

hence (at least formally),

$$\frac{d}{dt} \mathbb{E}_{U(t)x} Y = (-iu) \langle x, e^{-itA}(AB - BA)e^{itA}x \rangle = -i \langle U(t)x, [A, B]U(t)x \rangle.$$

In particular, the mean remains constant if and only if  $[A, B] = 0$ : the observables must be compatible. The case  $A = B$  is interpreted in physics as *conservation of energy*.

Schroedinger's evolution operator can be written in terms of the pvm associated with  $A$ ,

$$A = \int_{-\infty}^{+\infty} \lambda \Pi_{d\lambda}.$$

By the spectral theorem in pvm form,

$$(7.1.14) \quad e^{itA} = \int_{-\infty}^{+\infty} e^{i\lambda t} \Pi_{d\lambda},$$

which makes sense for  $t \in \mathbb{R}$ , and it defines a notion of Fourier transform for the pvm  $\Pi$ ,  $\widehat{\Pi}(t) = e^{itA}$ . If you are familiar with Bochner's theorem, and

you observe that  $\Pi$  exhibits some positivity ( $\Pi_E \geq 0$  as operator), then you expect as well its "Fourier transform" to have a property similar to "positive definiteness":

$$\sum_{j,k=1}^n c_j \overline{c_k} e^{i(t_j - t_k)A} \geq 0$$

as operator, for any choice of complex  $\alpha_1, \dots, \alpha_n$  and real  $t_1, \dots, t_n$ . The guess is correct,

$$\left\langle x, \sum_{j,k=1}^n c_j \overline{c_k} e^{i(t_j - t_k)A} x \right\rangle = \left\| \sum_{j=1}^n c_j e^{it_j A} x \right\|^2 \geq 0.$$

At some point one should talk about entanglement, but you will have to read it elsewhere.

## 7.2. The spectral theorem for unitary operators

Here we follow one of the two proofs in Michael Taylor's lecture notes [The Spectral Theorem for Self-Adjoint and Unitary Operators](#). The proof only requires basic Fourier series.

Recall that a linear operator  $H \xrightarrow{U} H$  is *unitary* iff and only if  $U^*U = UU^* = I$  is the identity. This is equivalent to the property of preserving the inner product,

$$(7.2.1) \quad \langle Ux, Uy \rangle = \langle x, y \rangle,$$

whenever  $x, y \in H$ . We have clearly that  $\|U\| = 1$ , but there are plenty of operators having norm one which are not unitary.

An abundant source of unitary operators is provided by specific multiplication operators on  $L^2$  spaces. If  $(X, \nu)$  is a measure space and  $\varphi : X \rightarrow \mathbb{R}$  is measurable, then

$$(7.2.2) \quad M_{e^{i\varphi}} : f \mapsto e^{i\varphi} f$$

is unitary on  $L^2(\nu)$ . Also observe that  $M_{e^{i\varphi}}^* = M_{e^{-i\varphi}}$ . At the end of the day, we learn that all unitary operators are unitarily equivalent to some  $M_{e^{i\varphi}}$  for some measure space  $(X, \nu)$ .

**7.2.1. The spectrum of a unitary operator.** Recall that  $\mathbb{T} = \{e^{it} : t \in (-\pi, \pi]\}$  is the unit circle in the complex plane (the *torus*).

**PROPOSITION 7.1.** *If  $H \xrightarrow{U} H$  is unitary, then  $\sigma(U) \subseteq \mathbb{T}$ .*

**PROOF.** If  $|\lambda| > 1$ , then  $U - \lambda I = \lambda(U/\lambda - I)$  is invertible because  $\|U/\lambda\| < 1$ , hence  $\lambda \in \rho(U)$ . If  $|\lambda| < 1$ , then  $U - \lambda I = U(I - \lambda U^*)$  is invertible because  $U$  and  $I - \lambda U^*$  are, so again  $\lambda \in \rho(U)$ .  $\square$

**7.2.2. The continuous functional calculus for a unitary operator.**

If  $p(e^{it}) = p(e^{it}) = \sum_{n=-N}^N a_n e^{int}$  is a trigonometric polynomial, it natural to define

$$(7.2.3) \quad p(U) = \sum_{n=-N}^N a_n U^n.$$

The obvious *polynomial calculus* defined by (7.2.3) generates a homomorphism of algebras,  $(pq)(U) = p(U)q(U)$ .

This definition immediately extends to functions  $f : \mathbb{T} \rightarrow \mathbb{C}$  whose Fourier series converges absolutely,

$$(7.2.4) \quad \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} |\widehat{f}(n)| = \|f\|_{A(\mathbb{T})} < \infty.$$

The space of such functions is called **Wiener algebra**. However, we just want to have a dense subspace of  $L^2(\mathbb{T})$ , so we consider  $f \in C^2(\mathbb{T}) := C_{per}^2$ . For such  $f$ 's, define

$$(7.2.5) \quad f(U) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{f}(n) U^n,$$

which converges in operator norm. For such " $C^2(\mathbb{T})$  calculus" the following properties are easily established.

- (i) **If  $f \in C^2(\mathbb{T})$ , then  $f(U)^* = \overline{f}(U)$ , where  $\overline{f}(e^{it}) := \overline{f(e^{it})}$ .** Using that

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

in fact,

$$\begin{aligned} f(U)^* &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \overline{\widehat{f}(n)} U^{*n} \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \widehat{\overline{f}}(-n) U^{-n} \\ &= \overline{f}(U). \end{aligned}$$

- (ii) **If  $f, g \in C^2(\mathbb{T})$ , then  $(fg)(U) = f(U)g(U)$ .** Absolute convergence of the corresponding Fourier series ensures that the following equalities hold:

$$\begin{aligned} (fg)(U) &= \frac{1}{(2\pi)^2} \sum_{m,n \in \mathbb{Z}} \widehat{f}(n-m) \widehat{g}(m) U^n \\ &= \frac{1}{2\pi} \sum_{m \in \mathbb{Z}} \left( \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \widehat{f}(n-m) U^{n-m} \right) \widehat{g}(m) U^m \\ &= f(U)g(U). \end{aligned}$$

- (iii) **If  $f \geq 0$  and  $f \in C^2(\mathbb{T})$ , then  $\langle x, f(U)x \rangle \geq 0$  for all  $x$  in  $H$ .** If  $\epsilon > 0$ , then  $f + \epsilon \in C^2(\mathbb{T})$  is strictly positive, hence  $\sqrt{f + \epsilon} \in C^2(\mathbb{T})$ . Since  $\sqrt{f + \epsilon}$  is real valued, by (i)  $(\sqrt{f + \epsilon})(U)$  is self-adjoint. Then,

$$\begin{aligned} \langle x, f(U)x \rangle + \epsilon \|x\|^2 &= \langle x, [f(U) + \epsilon I]x \rangle = \langle x, (\sqrt{f + \epsilon})(U)(\sqrt{f + \epsilon})(U)x \rangle \\ &= \langle (\sqrt{f + \epsilon})(U)x, (\sqrt{f + \epsilon})(U)x \rangle = \|(\sqrt{f + \epsilon})(U)x\|^2 \\ &\geq 0. \end{aligned}$$

Hence,  $\langle x, f(U)x \rangle \geq 0$ .

- (iv) **If  $f \in C^2(\mathbb{T})$ , then  $\|f(U)x\| \leq \|f\|_{L_0^\infty} \|x\|$ .** The function  $g(e^{it}) = \|f\|_{L_0^\infty}^2 - |f|^2$  is positive, hence, by (iii),

$$\begin{aligned} 0 &\leq \langle x, (\|f\|_{L_0^\infty}^2 - |f|^2)(U)x \rangle = \|f\|_{L_0^\infty}^2 \|x\|^2 - \langle x, (\overline{f}f)(U)x \rangle \\ &= \|f\|_{L_0^\infty}^2 \|x\|^2 - \langle x, f^*(U)f(U)x \rangle \\ &= \|f\|_{L_0^\infty}^2 \|x\|^2 - \|f(U)x\|^2. \end{aligned}$$

Item (iv) means that the regularity assumption  $f \in C^2$  is somewhat redundant, since only the  $L^\infty$  norm of  $f$  is relevant in order to estimate  $\|f(U)\|$ . Based on this, we can extend the  $C^2(\mathbb{T})$  calculus to a *continuous calculus*.

**THEOREM 7.6.** *Let  $U$  be a unitary operator on a Hilbert space  $H$ . There is a unique homomorphism of algebras  $\Lambda : C(\mathbb{T}) \rightarrow \mathcal{B}(H)$ ,  $\Lambda(f) = f(U)$ , such that*

- (i) *if  $e(e^{it}) = e^{it}$  is the identity function on  $\mathbb{T}$ , then  $\Lambda(e) = U$  (hence,  $\Lambda(p) = p(U)$  when  $p$  is a trigonometric polynomial);*
- (ii)  *$\Lambda$  is continuous when  $\mathcal{B}(H)$  is endowed with the uniform operator topology (the one induced by the operator norm),  $\|\Lambda(f)\| \leq \|f\|_{L_0^\infty}$ .*

Moreover, if  $f \in C(\mathbb{T})$ ,

$$(iii) \quad f(U)^* = \overline{f}(U);$$

$$(iv) \quad \text{if } f \text{ is real valued, then } f(U) = f(U)^*, \text{ and, if } f \geq 0, \text{ then } f(U) \geq 0;$$

PROOF. The map  $\Lambda$  was introduced above on the subspace of the trigonometric polynomial in accordance to (i). Given  $f \in C^2(\mathbb{T})$ , and

$$(S_N f)(e^{it}) = \frac{1}{2\pi} \sum_{|n| \leq N} \widehat{f}(n) e^{int},$$

we have that  $\|S_N f - f\|_{L^\infty} \rightarrow 0$  as  $N \rightarrow \infty$ , hence, by item (iv) above,

$$\lim_{N \rightarrow \infty} \|\Lambda(S_N f) - f(U)\| = \lim_{N \rightarrow \infty} \|(S_N f)(U) - f(U)\| = 0,$$

where  $f(U)$  was defined earlier; hence,  $f(U) = \Lambda(f)$ .

If  $g \in C(\mathbb{T})$  and  $\{f_n\}$  is a sequence in  $C^2(\mathbb{T})$  converging to  $g$  uniformly, then  $\{\Lambda(f_n)\}$  is a Cauchy sequence  $\mathcal{B}(H)$  with respect to the norm topology, hence it converges to some operator  $A \in \mathcal{B}(H)$ . By requirement (ii),  $A = \Lambda(g)$ . It is readily verified that any other sequence converging to  $g$  provides the same definition of  $\Lambda(g)$ . The fact that  $\Lambda$  is a homomorphism of algebras is routine.

This shows existence and uniqueness of a map  $\Lambda$  with the properties (i), (ii). The proof of (iii) and (iv) is left as an exercise.  $\square$

The map  $\Lambda$  needs not being injective. If  $H$  is finite dimensional, for instance,  $C(\mathbb{T})$  is infinite dimensional, whereas  $\mathcal{L}(H)$  is finite dimensional.

**7.2.3. The spectral measures and the measurable calculus.** Fix a unitary operator  $U$  on a Hilbert space  $H$ . By theorem 7.6, for each  $x \in H$ , the map

$$(7.2.6) \quad \Lambda_{x,x} : f \mapsto \langle x, f(U)x \rangle, \quad \Lambda_{x,x} : C(\mathbb{T}) \rightarrow \mathbb{C},$$

is linear, continuous, and positive. By the Riesz-Markov-Kakutani representation theorem for measures, there exists a (unique) finite Borel measure  $\mu_{x,x}$  on  $\mathbb{T}$  such that

$$(7.2.7) \quad \langle x, f(U)x \rangle = \int_{\mathbb{T}} f(e^{it}) d\mu_{x,x}(e^{it}).$$

By polarization, for each  $x, y \in H$  we can find a complex valued, finite Borel measure  $\mu_{x,y}$  such that

$$(7.2.8) \quad \langle x, f(U)y \rangle = \int_{\mathbb{T}} f(e^{it}) d\mu_{x,y}(e^{it}).$$

The measures  $\mu_{x,y}$  are the *spectral measures* associated with  $U$ . Observe that the association is *canonical*: they are uniquely determined by  $U$ . We will need the following.

LEMMA 7.1. *If  $U$  is unitary,  $x, y \in H$ , and  $f \in C(\mathbb{T})$ , then*

$$(7.2.9) \quad f d\mu_{x,y} = d\mu_{x,f(U)y}.$$

PROOF. For any other continuous  $g$ ,

$$\int_{\mathbb{T}} f g d\mu_{x,y} = \langle x, g(U) f(U) y \rangle = \int_{\mathbb{T}} g d\mu_{x,f(U)y}.$$

□

The map  $(x, y) \mapsto \mu_{x,y}$  is a sesquilinear map with values in the finite, complex measures on  $\mathbb{T}$ .

We have, in particular,

$$\int_{\mathbb{T}} e^{it} d\mu_{x,y}(e^{it}) = \langle x, U y \rangle,$$

and  $\mu_{x,y}(\mathbb{T}) = \langle x, y \rangle$ . The advantage of the spectral measures is that we can define a *measurable calculus*.

THEOREM 7.7. *[Measurable calculus for a unitary operator] There is a unique linear map  $\Lambda : L_0^\infty(\mathbb{T}) \rightarrow \mathcal{B}(H)$ ,  $\Lambda(f) = f(U)$ , such that:*

- (i)  $\Lambda(U) = f(U)$  agrees with the previously defined one when  $f \in C(\mathbb{T})$ ;
- (ii)  $\Lambda$  is continuous in the sense that if  $L^\infty(\mathbb{T}) \ni f_n \rightarrow f \in L^\infty(\mathbb{T})$  pointwise and  $\|f_n\|_{L_0^\infty} \leq C$  is bounded, then  $\Lambda(f_n) \rightarrow \Lambda(f)$  in the **weak operator topology**.

Moreover,  $\Lambda$  is a  $*$ -algebra homomorphism ( $\Lambda(fg) = \Lambda(f)\Lambda(g)$ , and  $\Lambda(f)^* = \Lambda(\bar{f})$ ) and

$$(7.2.10) \quad \langle x, f(U)y \rangle = \int_{\mathbb{T}} f d\mu_{x,y}.$$

PROOF. We begin by showing that a linear map  $\Lambda : L_0^\infty(\mathbb{T}) \rightarrow \mathcal{B}(H)$  with the properties (i) and (ii) exists. For  $f \in L_0^\infty$  and  $x, y \in H$ , let

$$[x, y]_f = \int_{\mathbb{T}} f d\mu_{x,y}.$$

By uniqueness of the spectral measures,  $[x, y]_f$  is sesquilinear, and

$$|[x, y]_f| \leq C \|f\|_{L_0^\infty}$$

if  $\|x\|, \|y\| \leq 1$ . By homogeneity,  $|[x, y]_f| \leq C \|f\|_{L_0^\infty} \|x\| \cdot \|y\|$  for all  $x, y$ , hence there exists a unique operator  $\Lambda(U) = f(U) \in \mathcal{B}(H)$  such that

$$[x, y]_f = \langle x, f(U)y \rangle.$$

The map  $\Lambda : f \mapsto f(U)$  just defined agrees with the previous definition when  $f$  is continuous, so (i) holds. Moreover, it is clearly linear in the  $f$  variable.

Suppose  $L^\infty(\mathbb{T}) \ni f_n \rightarrow f \in L_0^\infty$  pointwise, and that they are uniformly bounded. For each  $x, y$  in  $H$ , by dominated convergence,

$$\lim_{n \rightarrow \infty} \langle x, f_n(U)y \rangle = \lim_{n \rightarrow \infty} \int_{\mathbb{T}} f_n d\mu_{x,y} = \int_{\mathbb{T}} f d\mu_{x,y} = \langle x, f(U)y \rangle,$$

i.e.  $f_n(U) \rightarrow f(U)$  is the weak operator topology, i.e. property (ii) holds.

We show uniqueness. If (i) and (ii) hold for a linear map  $\tilde{\Lambda}$ , the space  $\mathcal{A}$  of functions  $f : \mathbb{T} \rightarrow \mathbb{C}$  in  $L_0^\infty(\mathbb{T})$  such that  $\tilde{\Lambda}(f) = \Lambda(f)$ : (1) contains  $C(\mathbb{T})$ , and (2) is closed under pointwise limits of uniformly bounded functions. By the corollary 1.1 to the Lebesgue-Hausdorff theorem,  $\mathcal{A} = L_0^\infty(\mathbb{T})$ .

We are left with proving that  $\Lambda$  is an  $*$ -algebra homomorphism. We start by extending lemma 7.1 to  $f \in L_0^\infty(\mathbb{T})$ . For  $n \in \mathbb{Z}$ , let  $e_n(t) = e^{int}$ . We start by showing that

$$\mu_{x,Uy} = \mu_{U^{-1}x,y}.$$

In fact, for  $n \in \mathbb{Z}$ ,

$$\int_{\mathbb{T}} e^{imt} d\mu_{x,Uy} = \langle x, U^m Uy \rangle = \langle U^{-1}x, U^m y \rangle = \int_{\mathbb{T}} e^{imt} d\mu_{U^{-1}x,y},$$

and a finite Radon measure  $\mu$  on  $\mathbb{T}$  is uniquely characterized by its Fourier coefficients  $\hat{\mu}(n) = \int_{\mathbb{T}} e^{-int} d\mu$ ,  $n \in \mathbb{Z}$ .

Now, for  $x, y \in H$ , and  $f \in L_0^\infty$ ,

$$\begin{aligned} \int_{\mathbb{T}} e^{int} f(e^{it}) d\mu_{x,y}(e^{it}) &= \int_{\mathbb{T}} f(e^{it}) d\mu_{x,U^n y}(e^{it}) = \int_{\mathbb{T}} f(e^{it}) d\mu_{U^{-n}x,y}(e^{it}) \\ &= \langle U^{-n}x, f(L)y \rangle = \langle x, U^n f(L)y \rangle \\ &= \int_{\mathbb{T}} e^{int} d\mu_{x,f(U)y}(e^{it}), \end{aligned}$$

which implies

$$(7.2.11) \quad f(e^{it}) d\mu_{x,y}(e^{it}) = d\mu_{x,f(U)y}(e^{it}).$$



With this at hands,

$$\begin{aligned}
 \langle x, (fg)(U)y \rangle &= \int_{\mathbb{T}} (fg)(e^{it}) d\mu_{x,y}(e^{it}) = \int_{\mathbb{T}} f(e^{it})g(e^{it}) d\mu_{x,y}(e^{it}) \\
 &= \int_{\mathbb{T}} f(e^{it}) d\mu_{x,g(U)y}(e^{it}) = \int_{\mathbb{T}} f(e^{it}) d\mu_{x,f(U)g(U)y}(e^{it}) \\
 &= \langle x, f(U)g(U)y \rangle.
 \end{aligned}$$

2

As a byproduct of the proof, we have that  $(fg)(U) = f(U)g(U)$  if  $f \in L^\infty(\mathbb{T})$  and  $g \in C(\mathbb{T})$ ,

$$\langle x, (fg)(U)y \rangle = \int_{\mathbb{T}} fg d\mu_{x,y} = \langle x, g(U)f(U)y \rangle.$$

We can now prove that  $(fg)(U) = f(U)g(U)$  for  $f, g \in L^\infty_0(\mathbb{T})$ ,

$$\begin{aligned}
 \langle x, (fg)(U)y \rangle &= \int_{\mathbb{T}} fg d\mu_{x,y} = \int_{\mathbb{T}} f d\mu_{x,g(U)y} = \int_{\mathbb{T}} d\mu_{x,f(U)g(U)y} \\
 &= \langle x, f(U)g(U)y \rangle.
 \end{aligned}$$

To show that  $\Lambda$  is a  $*$ -homomorphism,

$$\begin{aligned}
 \langle x, f(U)^*y \rangle &= \langle f(U)x, y \rangle = \overline{\langle y, f(U)x \rangle} \\
 &= \overline{\int_{\mathbb{T}} f d\mu_{y,x}} = \int_{\mathbb{T}} \overline{f} d\overline{\mu_{y,x}} \\
 &= \int_{\mathbb{T}} \overline{f} d\mu_{x,y} = \langle x, \overline{f}(U)y \rangle,
 \end{aligned}$$

as wished. □

---

<sup>2</sup>A different proof of (7.2.11), using the Lebesgue-Hausdorff theorem. For  $x, y \in H$ , let

$$\mathcal{A}_{x,y} = \{f \in L^\infty_0(\mathbb{T}) : f d\mu_{x,y} = d\mu_{x,f(U)y}\}.$$

We know that  $C(\mathbb{T}) \subseteq \mathcal{A}_{x,y}$ , and to apply the Lebesgue-Hausdorff theorem we only need to show that  $f \in \mathcal{A}_{x,y}$  if  $f$  is the pointwise limit of a uniformly bounded sequence  $f_n \in \mathcal{A}_{x,y}$ . For  $g \in C(\mathbb{T})$ , by dominated convergence and lemma 7.1,

$$\begin{aligned}
 \int_{\mathbb{T}} fg d\mu_{x,y} &= \lim_{n \rightarrow \infty} \int_{\mathbb{T}} f_n g d\mu_{x,y} = \lim_{n \rightarrow \infty} \int_{\mathbb{T}} g d\mu_{x,f_n(U)y} \\
 &= \lim_{n \rightarrow \infty} \langle x, g(U)f_n(U)y \rangle = \lim_{n \rightarrow \infty} \langle g(U)^*x, f_n(U)y \rangle \\
 &= \langle g(U)^*x, f(U)y \rangle = \langle x, g(U)f(U)y \rangle \\
 &= \int_{\mathbb{T}} g d\mu_{x,f(U)y},
 \end{aligned}$$

hence,  $f \in \mathcal{A}_{x,y}$ . Thus, lemma 7.1 holds for  $f \in L^\infty_0(\mathbb{T})$ .

**7.2.4. From the measurable calculus to projection valued measures.** Let  $E$  be a Borel subset of  $\mathbb{T}$ . We define  $\mu(E) = \chi_E(U)$ ,

$$(7.2.12) \quad \langle x, \mu(E)y \rangle = \int_E d\mu_{x,y} = \mu_{x,y}(E).$$

The map  $E \mapsto \mu(E)$  defines a function from the Borel  $\sigma$ -algebra into the set of the projections of  $H$ . In fact,

$$\begin{aligned} \mu(E)^* &= \overline{\chi_E}(U) = \chi_E(U) = \mu(E), \\ \mu(E)^2 &= \chi_E^2(U) = \chi_E(U) = \mu(E). \end{aligned}$$

**THEOREM 7.8.** *The map  $E \mapsto \mu(E)$  is a **projection valued measure** on  $E$ . That is,*

- (i)  $\mu(\emptyset) = 0$  and  $\mu(\mathbb{T}) = I$ ;
- (ii) if  $\{E_n\}_{n=1}^\infty$  is a sequence of disjoint, Borel measurable sets in  $\mathbb{T}$ , then

$$\mu\left(\bigcup_{n=1}^\infty E_n\right) = \sum_{n=1}^\infty \mu(E_n)$$

*converges in the **strong operator topology**.*

Having a measure, we expect to have an associated version of integral. In this case, we would like to integrate functions  $f : \mathbb{T} \rightarrow \mathbb{C}$ ,

$$(7.2.13) \quad \int_{\mathbb{T}} f(e^{it}) d\mu(e^{it}) \in \mathcal{L}(H).$$

This can be done in a rigorous way. See e.g. §3.4 in [Spectral theory in Hilbert spaces \(2009\)](#), ETH Zürich lecture notes, by Emmanuel Kowalski. We will give the expression in (7.2.13) the weak meaning

$$(7.2.14) \quad \left\langle x, \left( \int_{\mathbb{T}} f(e^{it}) d\mu(e^{it}) \right) y \right\rangle := \int_{\mathbb{T}} f d\mu_{x,y}.$$

**PROOF.** Property (i) is clear. Let  $\{E_n\}$  be like in (ii) and  $E = \bigcup_{n=1}^\infty E_n$ , and let  $y \in H$ . The vectors  $\mu(E_n)y$ ,  $n \geq 1$ , are orthogonal and

$$\sum_{n=1}^\infty \|\mu(E_n)y\|^2 \leq \|y\|^2,$$

hence

$$\sum_{n=1}^\infty \mu(E_n)y$$

converges in  $H$ . Since for all  $x \in H$

$$\begin{aligned}\langle x, \mu(E)y \rangle &= \int_E d\mu_{x,y} = \sum_{n=1}^{\infty} \int_{E_n} d\mu_{x,y} = \sum_{n=1}^{\infty} \langle x, \mu(E_n)y \rangle \\ &= \left\langle x, \sum_{n=1}^{\infty} \mu(E_n)y \right\rangle,\end{aligned}$$

we have that  $\sum_{n=1}^{\infty} \mu(E_n) = \mu(E)$ .  $\square$

The p.v.m.  $\mu$  and the spectral measures  $\mu_{x,y}$  are related by (7.2.12). The advantage of  $\mu$  is that we can use it to define null-sets and an  $L^\infty$  norm. A Borel subset  $E$  of  $\mathbb{T}$  is a *null-set* for  $U$  if  $\mu(E) = 0$ . The countable union of null-sets is obviously a null-set. A property  $\mathcal{P}(e^{it})$  ( $e^{it} \in \mathbb{T}$ ) holds  $\mu$ -almost everywhere if the set where it fails is a null-set for  $\mu$ . A function  $\varphi : \mathbb{T} \rightarrow \mathbb{C}$  belongs to  $L^\infty(\mu)$  if there is  $C \leq 0$  such that  $|\varphi(e^{it})| \leq C$   $\mu$ -a.e.. The minimum value of  $C$  for which this holds is the  $L^\infty(\mu)$ -norm of  $\varphi$ ,  $\|\varphi\|_{L^\infty(\mu)}$ .

Observe that other  $L^p(\mu)$ -norms can not be defined so easily. The advantage of the  $L^\infty$ -norm is that it is based on the primitive notion of null-set, not on the more sophisticated notion of the measure of a set ( $\mu$  is a measure, but with values in the space of operators).

We now relate being null for  $\mu$  with being null for  $\mu_{x,x}$ . In analogy with the case of finite positive Borel measures, we define the *support* of a p.v.m.  $\mu$  on  $\mathbb{T}$  to be the smallest closed subset  $F$  of  $\mathbb{T}$  such that  $\mu(\mathbb{T} \setminus F) = 0$ .

**THEOREM 7.9.** [*Nullsets for the p.v.m. associated with a unitary operator*]

(i) For a Borel subset  $E$  of  $\mathbb{T}$  the following are equivalent:

- (a)  $\mu(E) = 0$ ;
- (b)  $\mu_{x,x}(E) = 0$  for all  $x \in H$ ;
- (c)  $\mu_{x,y}(E) = 0$  for all  $x, y \in H$ ;
- (d)  $\mu_{e_\alpha, e_\beta}(E) = 0$  for all  $e_\alpha, e_\beta$  in an orthonormal basis of  $H$ .

(ii) For a Borel function  $\varphi : \mathbb{T} \rightarrow \mathbb{C}$ ,

$$\|\varphi\|_{L^\infty(\mu)} = \sup_{x \in H} \|\varphi\|_{L^\infty(\mu_{x,x})}.$$

(iii) If  $f \in L_0^\infty(\mathbb{T})$ , then  $\|f(U)\| = \|f\|_{L^\infty(\mu)}$ .

PROOF. (i) If  $\mu(E) = 0$ , then  $\mu_{x,y}(E) = \langle x, \mu(E)y \rangle = 0$  for all  $x, y$ , which by polarization is the same as  $\mu_{x,x}(E) = 0$  for all  $x$ . If  $\langle x, \mu(E)y \rangle = 0$  for all  $x, y$ , on the other hand,  $\mu(E)y = 0$  for all  $y$ , then  $\mu(E) = 0$ . This shows (a)  $\iff$  (b)  $\iff$  (c). On the other hand, (d) states that all "matrix coefficients"  $\langle e_\alpha, \mu(E)e_\beta \rangle = \mu_{e_\alpha, e_\beta}(E)$  of  $\mu(E)$  vanish, hence that  $\mu(E) = 0$ .

(ii) For  $\alpha > 0$ , using (i), we have

$$\begin{aligned} \|\varphi\|_{L^\infty(\mu)} \leq \alpha &\iff \mu\{t : |\varphi(e^{it})| > \alpha\} = 0 \\ &\iff \text{for all } x \in H : \mu_{x,x}\{t : |\varphi(e^{it})| > \alpha\} = 0 \\ &\iff \sup_{x \in H} \|\varphi\|_{L^\infty(\mu)} \leq \alpha, \end{aligned}$$

hence  $\|\varphi\|_{L^\infty(\mu)} = \sup_{x \in H} \|\varphi\|_{L^\infty(\mu)}$ .

(iii) In one direction, for  $x \in H$  we have

$$\begin{aligned} \|f(U)x\|^2 &= \langle x, |f|^2(U)x \rangle = \int_{\mathbb{T}} |f|^2 d\mu_{x,x} \\ &\leq \|f\|_{L^\infty(\mu_{x,x})}^2 \|x\|^2 \leq \|f\|_{L^\infty(\mu)}^2 \|x\|^2 \text{ by (ii),} \end{aligned}$$

hence,  $\|f(U)\| \leq \|f\|_{L^\infty(\mu)}$ .

In the other direction, let  $0 \leq \lambda < \|f\|_{L^\infty(\mu)}$  (we can assume  $\|f\|_{L^\infty(\mu)} > 0$ , since if  $\|f\|_{L^\infty(\mu)} = 0$  there is nothing to prove). By (ii), there is  $x$  in  $H$  such that  $\lambda < \|f\|_{L^\infty(\mu_{x,x})}$ . Hence, there exists a Borel set  $E$  such that  $\mu_{x,x}(E) > 0$  and  $|f(e^{it})| > \lambda$  for  $e^{it}$  in  $E$ . Thus,

$$\begin{aligned} 0 &< \int_E (|f|^2 - \lambda^2) d\mu_{x,x} = \int_{\mathbb{T}} (|f|^2 - \lambda^2) \chi_E^2 d\mu_{x,x} \\ &= \int_{\mathbb{T}} (|f|^2 - \lambda^2) d\mu_{\chi_E(U)x, \chi_E(U)x} \\ &= \langle \chi_E(U)x, |f|^2(U) \chi_E(U)x \rangle - \lambda^2 \langle \chi_E(U)x, \chi_E(U)x \rangle \\ &= \|f(U)[\chi_E(U)x]\|^2 - \lambda^2 \|\chi_E(U)x\|^2. \end{aligned}$$

The strict inequality rules out the possibility that  $\chi_E(U)x = 0$ , and the inequality itself implies that

$$\lambda \|\chi_E(U)x\| < \|f(U)[\chi_E(U)x]\|,$$

which implies that  $\lambda < \|f(U)\|$ . Hence,  $\|f\|_{L^\infty(\mu)} \leq \|f(U)\|$ .  $\square$

### 7.2.5. From the continuous calculus to the multiplicative form.

Let's go back to the continuous calculus. A vector  $x \in H$  is *cyclic* for the unitary operator  $U$  if the linear space

$$\{f(U)x : f \in C(\mathbb{T})\}$$

is dense in  $H$ . The space  $V = \{f(U)x : f \in C(\mathbb{T})\}$  is *invariant for the action of  $U$* ,

$$UV \subseteq V.$$

In fact,  $V$  is invariant under the action of  $f(U)$  if  $f \in C(\mathbb{T})$ ,  $[f(U)](V) \subseteq V$ .

**PROPOSITION 7.2.** *Suppose  $U$  has a cyclic vector  $x$ . Then, the map  $\Theta : C(\mathbb{T}) \rightarrow H$ ,  $\Theta(f) = f(U)x$ , extends to a unitary map from  $L^2(\mathbb{T}, \mu_{x,x})$  to  $H$ , such that*

$$(7.2.15) \quad ((\Theta^{-1}U\Theta)f)(e^{it}) = e^{it}f(e^{it}).$$

That is,  $U$  is unitarily equivalent to multiplication times  $e^{it}$  on  $L^2(\mathbb{T}, \mu_{x,x})$ ,

$$\begin{array}{ccc} L^2(\mathbb{T}, \mu_{x,x}) & \xrightarrow{\Theta} & H \\ M_{e^{it}} \downarrow & & \downarrow U \\ L^2(\mathbb{T}, \mu_{x,x}) & \xrightarrow{\Theta} & H \end{array}$$

**PROOF.** The map  $\Theta$  is an  $L^2(\mu_{x,x})$  isometry on  $C(\mathbb{T})$ ,

$$\|f(U)x\|^2 = \langle x, f(U)^*f(U)x \rangle = \langle x, |f|^2(U)x \rangle = \int_{\mathbb{T}} |f|^2 d\mu_{x,x}.$$

As such,  $\Theta$  extends to an isometry of  $\overline{C(\mathbb{T})}^{L^2(\mu_{x,x})}$ , the closure of  $C(\mathbb{T})$  in the  $L^2(\mu_{x,x})$  norm. The hypothesis that  $x$  is cyclic means that  $\Theta(C(\mathbb{T}))$  is dense in  $H$ , hence that  $\Theta$  is a surjective isometry.  $\square$

We can then use some kind of finite or infinite induction, to have a similar result even when we do not have a cyclic vector at hands. We consider the case where  $H$  is separable. The general case can be dealt with by Zorn's lemma.

**THEOREM 7.10.** *[Spectral theorem for unitary operators: multiplicative form] Let  $U$  be a unitary operator on a separable Hilbert space  $H$ . Then, there exists a locally compact space  $X$ , a finite Borel measure  $\nu$  on  $X$ , and a continuous function  $\varphi : X \rightarrow \mathbb{R}$ , and a unitary map  $\Theta : L^2(\nu) \rightarrow H$ , such that*

$$(7.2.16) \quad (\Theta^{-1}U\Theta)f = e^{i\varphi}f,$$

*is multiplication times  $e^{i\varphi}$ .*

PROOF. Let  $\{x_n : n \geq 1\}$  be a orthogonal basis for  $H$ , normalized to have  $\sum_{n \geq 1} \|x_n\|^2 < \infty$ . Let  $P(\mathbb{T})$  be the space of the trigonometric polynomials. We construct a family of closed, mutually orthogonal subspaces  $H_n$ , whose sum exhausts  $H$ .

- Let  $y_1 = x_1$  and  $H_1 = \overline{\{p(U)y_1 : p \in P(\mathbb{T})\}}^H$ .
- Let  $y_2 = x_2 - \pi_{H_1}x_2 \in H_1^\perp$ , and  $H_2 = \overline{\{p(U)y_2 : p \in P(\mathbb{T})\}}^H$ .
- If  $H_1, \dots, H_{n-1}$  have been chosen, let  $y_n = x_n - \pi_{H_1 \oplus \dots \oplus H_{n-1}}x_n \in (H_1 \oplus \dots \oplus H_{n-1})^\perp$ , and  $H_n = \overline{\{p(U)y_n : p \in P(\mathbb{T})\}}^H$ .

Pick  $p(U)y_n \in H_n$  and  $q(U)y_m \in H_m$  with  $m < n$ . Then,

$$\langle p(U)y_n, q(U)y_m \rangle = \langle y_n, (\bar{p}q)(U)y_m \rangle = 0,$$

because  $\bar{p}q$  is a trigonometric polynomial, hence,  $(\bar{p}q)(U)y_m \in H_m$ , to which  $y_n$  is orthogonal. Since the elements of the form  $p(U)y_n$  are dense in  $H_n$  and those of the form  $q(U)y_m$  are dense in  $H_m$ , we have that  $H_n \perp H_m$ .

After exhausting the list of the  $x_n$ 's, we have that  $\bigoplus_{n \geq 1} H_n$  is dense in  $H$ . Some of  $H_n$ 's might be trivial. In this case, we can remove them and renumber the sequence. The nontrivial  $H_n$ 's might be finitely many.

For fixed  $n$ , we are in the situation described by proposition 7.2. The map  $\Theta_n : P(\mathbb{T}) \rightarrow H_n$ ,

$$\Theta_n(p) = p(U),$$

extends to a unitary map from  $\Theta_n : L^2(\mathbb{T}, \mu_{y_n, y_n}) \rightarrow H_n$ , and

$$[(\Theta_n^{-1}U\Theta_n)f](e^{it}) = e^{it}f(e^{it}).$$

Also,  $\mu_{y_n, y_n}(\mathbb{T}) = \|y_n\|^2 \leq \|x_n\|^2$ .

Let  $A$  be the (discrete) set of the indices  $n$  (which might be finite or infinite), and let  $X = \mathbb{T} \times A$ , with the product topology, which is locally compact. Let  $\nu$  be the Borel measure which restricts to  $\mu_{y_n, y_n}$  on the component  $\mathbb{T} \times \{n\}$ , and observe that  $\nu$  is finite. We can naturally identify  $L^2(X, \nu) = \bigoplus_{n \in A} L^2(\mathbb{T}, \mu_{y_n, y_n})$ . Finally, define  $\Theta : L^2(X, \nu) \rightarrow H$  summand by summand,

$$\Theta|_{L^2(\mathbb{T}, \mu_{y_n, y_n})} = \Theta_n.$$

If  $F : \mathbb{T} \times A \rightarrow \mathbb{C}$ , then

$$(7.2.17) \quad [(\Theta^{-1}U\Theta)F](e^{it}, n) = e^{it}F(e^{it}, n).$$

The operator  $(\Theta^{-1}U\Theta)$ , that is, acts like multiplication times  $e^{it}$  on each "cl-open" component  $\mathbb{T} \times \{n\}$  of  $X$ .  $\square$

The representation we have provided of  $U$  as a multiplication operator is not canonical, even when  $H = \mathbb{C}^2$ . Consider for instance

$$U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The vector  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$  is cyclic, and this provides a representation with just an "invariant subspace". The vectors  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  are not, and they provide a representation with two "invariant subspaces".

The spaces  $H_n$  in the proof of theorem 7.10 are clearly invariant for  $U$ .

**7.2.6. The spectral theory of a multiplication operator.** The multiplicative form of the spectral theorem tells us that any unitary operator is unitarily equivalent to a multiplication operator  $M_\eta$  of the following form, for which we can explicitly write down the spectrum, the spectral measures, and so on.

(a) *We have a locally compact space  $X$  and a finite positive Borel measure  $\nu$ , which we can assume to be finite, and a surjective local homeomorphism  $\eta : X \rightarrow \mathbb{T}$ , such that  $M_\eta : L^2(\nu) \rightarrow L^2(\nu)$  is the operator  $M_\eta f = \eta f$ .*

(b)  $\sigma(M_\eta) = \overline{\eta(\text{supp}(\nu))} \subseteq \mathbb{T}$  is the spectrum of  $M_\eta$ .

PROOF. Let  $z \in \mathbb{T}$ .

$$\begin{aligned} z \in \rho(M_\eta) &\iff (z - M_\eta)^{-1} \in L^\infty(\nu) \\ &\iff |\eta(x) - z| \geq \epsilon \text{ } \nu - a.e. \text{ for some } \epsilon > 0 \\ &\iff \eta^{-1}(D(z, \epsilon)) \cap \text{supp}(\nu) = \emptyset \text{ for some } \epsilon > 0 \\ &\iff D(z, \epsilon) \cap \eta(\text{supp}(\nu)) = \emptyset \text{ for some } \epsilon > 0 \\ &\iff z \notin \overline{\eta(\text{supp}(\nu))}. \end{aligned}$$

□

(c) *Let  $f \in L^2(\nu)$ . The spectral measure  $\mu_{f,f}$  is such that, for any trigonometric polynomial  $p : \mathbb{T} \rightarrow \mathbb{C}$ ,*

$$\int_X p(\eta) |f|^2 d\nu = \langle f, p(M_\eta) f \rangle_{L^2(\nu)} = \int_{\mathbb{T}} p d\mu_{f,f}.$$

That is,  $d\mu_{f,f} = \eta_*(|f|^2 d\nu)$ . (Recall that the *push-forward* of a measure  $\nu$  on  $X$  under a map  $\eta : X \rightarrow Y$  is the measure defined by  $\int_Y \varphi d(\eta_*\nu) := \int_X (\varphi \circ \eta) d\nu$ . Of course, this needs some measurability assumptions, that is our case are easily verified). Similarly,  $d\mu_{f,g} = \eta_*(\bar{g}f d\nu)$ .

- (d) The p.v.m.  $\mu$  maps Borel measurable sets  $E$  in  $\mathbb{T}$  to projections of  $L^2(\nu)$  in the following fashion:

$$\int_X \bar{f}[\mu(E)g] d\nu = \langle f, \mu(E)g \rangle_{L^2(\nu)} = \mu_{f,g}(E) = \int_X \chi_E(\eta) \bar{f}g d\nu,$$

$$\text{i.e. } (\chi_E(\eta) = \chi_E \circ \eta)$$

$$(7.2.18) \quad \mu(E)g = \chi_E(\eta)g, \quad \mu(E) = M_{\chi_E(\eta)}.$$

- (e) A Borel set  $E \subseteq \mathbb{T}$  is a null-set for  $\mu$  if and only if  $0 = \chi_E \circ \eta = \chi_{\eta^{-1}(E)} \nu$  - a.e., i.e. if and only if  $\nu(\eta^{-1}(E)) = 0$ .

- (f) By (e) and the topological properties of  $\eta$ ,  $\overline{\eta(\text{supp}(\nu))} = \text{supp}(\mu)$ .

PROOF.

$$\begin{aligned} z \notin \text{supp}(\mu) &\iff \mu(D(z, \epsilon)) = 0 \text{ for some } \epsilon > 0 \\ &\iff \nu(\eta^{-1}(D(z, \epsilon))) = 0 \text{ for some } \epsilon > 0, \text{ by (e)} \\ &\iff \eta^{-1}(D(z, \epsilon)) \cap \text{supp}(\nu) = \emptyset \text{ for some } \epsilon > 0 \\ &\iff D(z, \epsilon) \cap \eta(\text{supp}(\nu)) = \emptyset \text{ for some } \epsilon > 0 \\ &\iff z \notin \overline{\eta(\text{supp}(\nu))}. \end{aligned}$$

□

- (g) By (b) and (f),

$$(7.2.19) \quad \text{supp}(\mu) = \sigma(M_\eta).$$

**7.2.7. Some applications of the spectral theorem in its multiplicative form.** A notable advantage of the multiplicative form of the spectral theory is that we can now use measure theory to study functional analytic properties of unitary operators. We start by highlighting the role of the spectrum.



**THEOREM 7.11.** *[The spectrum and the support of the projection valued measure] Let  $U : H \rightarrow H$  be unitary,  $\{\mu_{x,x} : x \in H\}$  be the family of the spectral measures for  $U$ , and  $\mu$  be the projection valued measure associated with  $U$ . We have*

$$(7.2.20) \quad \text{supp}(\mu) = \sigma(U).$$

**PROOF OF THEOREM 7.11.** By the multiplicative form of the spectral theorem, it suffices to show it for multiplication operators. This is done in (7.2.19).  $\square$

**COROLLARY 7.2.** *With the same notation of theorem 7.11, we also have the following.*

- (i) *For a function  $f \in L^\infty(\mu)$ ,  $f(U) = 0$  if and only if  $f$  vanishes  $\mu$ -a.e on  $\sigma(U)$ . Thus, the map  $f \mapsto f\chi_{\sigma(U)}$  is the identity of  $L^\infty(\mu)$ . To highlight this fact, we also write  $L^\infty(\mu) = L^\infty(\sigma(U))$ , and for  $\varphi : \sigma(U) \rightarrow \mathbb{C}$  measurable and bounded, we write  $\varphi(U) = f(U)$ , where  $f$  is any measurable extension of  $\varphi$  to  $\mathbb{T}$ .*
- (ii) *If  $\{f_n : n \geq 1\}$  is a sequence in  $L^\infty(\sigma(U))$ , if  $\|f_n\|_{L^\infty(\sigma)} \leq C$  is uniformly bounded, and  $f_n \rightarrow f$   $\mu$ -a.e., then  $f_n(U) \rightarrow f(U)$  in the strong operator topology.*
- (iii) *By Tietze extension theorem, the operator  $f \mapsto f|_{\sigma(U)} = \Psi f$  maps  $C(\mathbb{T})$  onto  $C(\sigma(U))$ . By (i),  $f(U) = 0$  if and only if  $\Psi f = 0$ . For  $g \in C(\sigma(U))$ , then, we can write  $g(U) := f(U)$ . Now, if  $\{g_n : n \geq 1\}$  is a sequence in  $C(\sigma(U))$  converging to  $g$  uniformly, then  $g_n(U)$  converges to  $g(U)$  uniformly.*

*Moreover,  $f \mapsto f(U)$  is an isometry of  $C(\sigma(U))$  into  $\mathcal{L}(H)$ .*

Here is another sample application. We denote by  $\mathcal{U}(H)$  the set of the unitary operators on  $H$ , which is also a group under composition.

**PROPOSITION 7.3.** *The unitary group  $\mathcal{U}(H)$ , endowed with the strong operator topology, is arc-connected.*

**PROOF.** We can assume that  $H = L^2(\nu)$  and that  $Uf = e^{i\varphi}f$ , where  $\varphi : X \rightarrow \mathbb{R}$  is measurable. For  $t \in \mathbb{R}$ , let

$$U_t f = e^{it\varphi} f,$$

$U = U_1$ . Each  $U_t$  is unitary,  $U_0 = I$ , and it is easy to prove that  $t \mapsto U_t$  is strongly continuous from  $\mathbb{R}$  to  $\mathcal{U}(L^2(\nu))$ . (We have  $U_{s+t} = U_s U_t$ :  $t \mapsto U_t$  is a strongly continuous one parameter group in  $\mathcal{U}(L^2(\nu))$ ).

Since any  $U$  can be connected to  $I$ , two such  $U$ 's can be connected to each other.  $\square$

In general, there are many arcs of this group type connecting  $U$  to the identity. We have in fact that

$$e^{i\varphi} = e^{i(\varphi+2\pi n)},$$

where  $n : X \rightarrow \mathbb{Z}$  is any measurable function. The maps  $t \mapsto e^{it(\varphi+2\pi n)}$  provide one parameter groups joining  $U$  and  $I$ , and in general they are different from each other.

**7.2.8. Some final remarks.** The various forms we have seen of the spectral theorem for unitary operators, and their proofs, suggest a general framework, strategy, and objectives.

- (i) Given the operator  $L$  on the Hilbert space  $H$ , we see if we can perform some simple algebraic operations on it. In this case, we can consider linear combinations of  $U^n$  with complex coefficients, where  $n$  can be both positive or negative. This provides a "restricted calculus", that in our case involves trigonometric polynomials  $p(e^{it})$ , which become the operators  $p(L)$ .

We will see when working with unbounded operators that finding a "restricted calculus" is not easy, since unbounded operators can not be in general composed without restricting more and more the domain. We will use instead the *Cauchy kernel*  $u \mapsto \frac{1}{u-z}$  ( $u \in \mathbb{R}$ ,  $z \in \mathbb{C} \setminus \mathbb{R}$ ), and the operators  $(L - z)^{-1}$  (the "resolvent map").

- (ii) If we have some information about convergence, we can extend the "restricted calculus" to a "continuous calculus". Here we had, for each  $f \in C(\mathbb{T})$ , a corresponding operator  $f(L)$ . The good news is that bounded, linear functionals on nice spaces of continuous functions are in a bijection with finite Borel measures by Riesz theorem. For instance, this is the case with the functionals

$$f \mapsto \langle x, f(L)y \rangle.$$

The "spectral" measures we have found are defined by the relation

$$\langle x, f(L)y \rangle = \int_{\mathbb{T}} f d\mu_{x,y}.$$

With unbounded operators things will be trickier. Lacking a polynomial, even less a continuous calculus, we have to resort to different results from real and complex analysis. Following Tao and one

stream of literature, we will use *Herglotz representation theorem*, associating to each positive harmonic function defined on a half-plane (and having suitable decay) a positive measure on the half-plane's boundary. This will produce "spectral measures", satisfying

$$\langle x, (L - z)^{-1}y \rangle = \int_{\mathbb{R}} \frac{d\mu_{x,y}(u)}{u - z}.$$

In both cases,

$$(7.2.21) \quad \langle x, Ly \rangle = \int_{\mathbb{R}} u d\mu_{x,y}(u) :$$

the operator  $L$  becomes, in the world of each spectral measure, multiplication times the coordinate of the space.

- (iii) One way to exploit the state of things described in (ii) is by extending the integrals to Borel measurable functions,

$$\langle x, m(L)y \rangle := \int_X m(u) d\mu_{x,y}(u),$$

for  $m$  Borel measurable and bounded. Here  $X$  is the "spectral set": the natural subset of the complex plane containing the spectra of the operators  $L$  belonging to the class we are studying. The most interesting instance is that of the characteristic functions,  $m = \chi_E$ , where  $E$  is measurable in  $X$ . Since  $\chi_E^2 = \chi_E$  and  $\chi_E$  is real valued, if we have an effective calculus, then  $\chi_E(L)^2 = \chi_E(L)$  and  $\chi_E(L)^* = \chi_E(L)$ ; i.e.  $\mu(E) := \chi_E(L)$  is a projection onto a closed subspace of  $H$ ! We have, that is, *projection valued measures*. By the way, these are of the utmost importance in quantum mechanics.

- (iv) A different strategy consists in starting again with (ii). We have realized that for elements of  $H$  having the form  $f(L)x$  (with  $x \in H$  fixed), the map  $\Theta : f \mapsto f(L)x$  is an isometry from  $L^2(\mu_{x,x})$  onto a "cyclic" subspace of  $H$ , and that  $\Theta$  intertwines  $L$  with a multiplication operator on  $L^2(\mu_{x,x})$ . This is the main building block of the multiplicative form of the spectral theorem.

### 7.3. Closed operators and their adjoints

Let  $D$  be a subspace of  $H$  and  $D \xrightarrow{L} H$  be a linear operator. We say that  $L$  is *densely defined* (d.d.) if  $D$  is dense in  $H$ . The *graph*  $\Gamma(L)$  of  $L$  is

$$\Gamma(L) = \{(f, Lf) : f \in D\} \subseteq H \times H.$$

The Cartesian product  $H \times H = H \oplus H$  is itself a Hilbert space with the inner product

$$\langle (f, g), (h, k) \rangle_{H \oplus H} = \langle f, h \rangle_H + \langle g, k \rangle_H.$$

The operator  $L$  is *closed* if  $\Gamma(L)$  is closed in  $H \times H$ , i.e. if

$$D \ni f_n \rightarrow f \text{ and } Lf_n \rightarrow g \text{ in } H, \text{ implies that } f \in D \text{ and } g = Lf.$$

REMARK 7.1. Let  $D \xrightarrow{L} H$  be closed, injective, and having closed range (this holds if, for instance,  $L$  is a bijection). By the closed graph theorem,  $\text{Ran}(L) \xrightarrow{L^{-1}} D$  is bounded.

EXAMPLE 7.1. The operator  $(Lf)(x) = xf(x)$ , having domain  $D = \{f \in L^2(\mathbb{R}) : xf \in L^2(\mathbb{R})\}$  is closed.

An *extension*  $D_1 \xrightarrow{L_1} H$  of an operator  $D \xrightarrow{L} H$  is a linear operator  $L_1$  with domain  $D_1 \supseteq D$  such that  $D_1|_D = L$ .

The operator  $D \xrightarrow{L} H$  is *symmetric* if for all  $f, g \in D$

$$\langle f, Lg \rangle = \langle Lf, g \rangle.$$

LEMMA 7.2. If  $D \xrightarrow{L} H$  is d.d. and symmetric, then

$$\overline{\Gamma(L)} = \Gamma(\overline{L})$$

is the graph of a closed, symmetric linear operator  $L$ , the **closure** of  $L$ . The domain of  $\overline{L}$  is denoted by  $\overline{D}$ . Beware!  $\overline{D}$  is not the closure of  $D$  in  $H$  (which is  $H$  itself), but rather  $\pi_1(\overline{\Gamma(L)})$ . That is,

$$\overline{D} = \{f \in H : \exists D \ni f_n \rightarrow f \text{ such that } \exists \lim_{n \rightarrow \infty} f_n \in H\}.$$

If  $L_1$  is a closed operator extending  $L$ , then  $\Gamma(L_1) \supseteq \Gamma(\overline{L}) \supseteq \Gamma(L)$ . In this sense  $\overline{L}$  is the minimal closed extension of  $L$ .

PROOF. Suppose  $\Gamma(L) \ni (f_n, Lf_n) \rightarrow (f, g)$ , and set  $\overline{L}f = g$ . The definition is well posed if it does not depend on the sequence  $\{f_n\}$  in  $D$ , but just on  $f$ , and this happens if and only if  $(h_n, Lh_n) \rightarrow (0, p)$  with  $h_n \in D$  implies that  $p = 0$ . For  $k \in D$  we have that

$$\langle k, p \rangle = \lim_n \langle k, Lh_n \rangle = \lim_n \langle Lk, Lh_n \rangle = 0,$$

hence  $p \in D^\perp = H^\perp = 0$ , since  $D$  is dense in  $H$ . Let  $\pi_1(f, g) = f$  be the projection onto the first coordinate in  $H \times H$ . We have defined  $\overline{D} \xrightarrow{\overline{L}} H$  on

$$\overline{D} := \{f \in H : \exists D \ni f_n \rightarrow f \text{ such that } \lim_n Lf_n \text{ exists in } H\} = \pi_1(\overline{\Gamma(L)}),$$

and  $\Gamma(\bar{L}) = \overline{\Gamma(L)}$ .

The operator  $\bar{L}$  is linear. If  $D \ni f_n \rightarrow f$ ,  $Lf_n \rightarrow g = \bar{L}f$ ,  $D \ni h_n \rightarrow h$ , and  $Lh_n \rightarrow k = \bar{h}$ , then  $L(af_n + bh_n) \rightarrow ag + bk = \overline{af + bh}$ , where  $a, b$  are scalars. Also,  $\bar{L}$  is symmetric:

$$\langle \bar{L}f, h \rangle = \lim_n \langle Lf_n, h_n \rangle = \lim_n \langle f_n, Lh_n \rangle = \langle f, \bar{L}h \rangle.$$

□

The *adjoint* of a d.d. linear operator  $D \xrightarrow{L} H$  is defined as in the bounded operator case, although we have to be careful with the domains. Let

$$D^* = \{g \in H : f \mapsto \langle g, Lf \rangle \text{ is bounded on } D\}.$$

For each  $g \in D^*$ , by density of  $D$  in  $H$  and Riesz theorem, there exists a unique  $L^*g \in H$  such that

$$\langle L^*g, L \rangle = \langle g, Lf \rangle.$$

It is an easy exercise showing that  $D^*$  is a subspace of  $H$  and  $L^*$  is a linear operator.

**REMARK 7.2.** *If  $L$  is d.d. and symmetric, then  $D^* \supseteq D$  and  $L^*$  extends  $L$ . In fact, for  $g \in D$ ,  $f \mapsto \langle Lf, g \rangle = \langle f, Lg \rangle$  is bounded, and  $L^*g = Lg$ .*

**LEMMA 7.3.** *If  $D \xrightarrow{L} H$  is d.d., then  $L^*$  is closed.*

**PROOF.** Let  $\tilde{\Gamma}(T) = \{(Tf, f) : f \in E\}$  be the co-graph of an operator  $E \xrightarrow{T} H$ . We prove that:

$$(7.3.1) \quad \tilde{\Gamma}(-L^*) := \{(-L^*g, g) : g \in D^*\} = \Gamma(L)^\perp \text{ in } H \oplus H.$$

On the one hand, for  $f \in D$ , we have

$$\langle (-L^*g, g), (f, Lf) \rangle = -\langle L^*g, f \rangle + \langle g, Lf \rangle = 0,$$

which shows that  $\tilde{\Gamma}(-L^*) \subseteq \Gamma(L)^\perp$ . Conversely,  $(h, g) \in \Gamma(L)^\perp$  if and only if  $0 = \langle h, f \rangle + \langle g, Lf \rangle$  for all  $f$  in  $D$ , but this implies that  $f \mapsto \langle g, Lf \rangle$  is bounded on  $D$  (hence,  $g \in D^*$ ), and that  $h = -L^*g$ . Hence,  $(h, g) \in \tilde{\Gamma}(-L^*)$ . By (7.3.1),  $\tilde{\Gamma}(-L^*)$  is closed, which implies that  $\Gamma(L^*)$  is closed, as wished. □

**LEMMA 7.4.** *Let  $L$  be d.d. and symmetric. Then,  $L^{**} = \bar{L}$  and  $L^* = \bar{L}^*$ .*

PROOF. By lemma 7.2 we have

$$\begin{aligned}\tilde{\Gamma}(-L^*) &= \Gamma(L)^\perp = \overline{\Gamma(L)}^\perp = \Gamma(\overline{L})^\perp \\ &= \tilde{\Gamma}(-\overline{L}^*),\end{aligned}$$

hence,  $L^* = \overline{L}^*$ , which is our second thesis. As a consequence, using lemma 7.2 twice,

$$\Gamma(L^{**}) = \tilde{\Gamma}(-L^*)^\perp = \tilde{\Gamma}(-\overline{L}^*)^\perp = \Gamma(\overline{L}),$$

hence,  $L^{**} = \overline{L}$ . □

EXAMPLE 7.2. *Unbounded operators provide examples of closed subspaces  $M, N \subset \mathcal{H}$  of a Hilbert space  $\mathcal{H}$  such that the orthogonal projection  $\pi_M$  onto  $M$  maps  $N$  onto a dense subspace  $\pi_M(N) \neq M$ . This fact is not evident from pictures. A general machine to produce counterexamples is by considering  $\mathcal{H} = H \times H$ , where  $H$  is infinite dimensional,  $M = H$  (say, the first component), and  $N = \Gamma(L)$ , where  $D \xrightarrow{L} H$  is a d.d., closed, unbounded operator (so that  $D$ , by the Hellinger-Toeplitz theorem, is strictly contained in  $H$ ). Then,  $\pi_1(\Gamma(L)) = D$  is not the whole range of  $\pi_1$ , although it is dense in it.*

A consequence of lemma 7.3 is that, if  $L$  is d.d. and symmetric, then  $L^*$  is a closed extension of  $L$ , hence of  $\overline{L}$ . We might hope that, under these assumptions,  $L^* = \overline{L}$ . Unfortunately this is not the case.

REMARK 7.3. *There are d.d., symmetric operators such that  $L^*$  is not symmetric. Since for all such operators  $L^* = \overline{L}^*$  by lemma 7.4, such examples arise even for some d.d., symmetric, closed operators. See the section on the momentum operator for a concrete instance.*

A d.d. defined operator  $D \xrightarrow{L} H$  is *self-adjoint* (s.a.) if and only if  $L = L^*$ :  $D = D^*$  and  $L = L^*$ . In particular, self-adjoint operators are closed (because  $L^*$  is closed) and symmetric: if  $f \mapsto \langle g, Lf \rangle$  is bounded on  $D$ , then  $g \in D^* = D$ , and  $\langle g, Lf \rangle = \langle L^*g, f \rangle = \langle Lg, f \rangle$ .

Self-adjointness is much stronger than it looks.

REMARK 7.4. *Let  $Dx \xrightarrow{L} H$  be self-adjoint, and let  $D \subseteq E \xrightarrow{M} H$  be a symmetric extension of  $L$ . Then,  $E = D$ . That is,  $L$  can not be extended to a larger domain while preserving symmetry.*

*The proof is easy. If  $y \in E$ , then*

$$x \mapsto \langle y, Lx \rangle = \langle y, Mx \rangle = \langle My, x \rangle$$

*defines a bounded functional on  $D$ , hence,  $y \in D = D^*$ .*

Self-adjoint operators, that is, are maximal in the lattice of the symmetric operators, ordered by extension. In particular, the domain of a self-adjoint operator must be "best possible".

A d.d. symmetric operator  $D \xrightarrow{L} H$  is just *essentially self-adjoint* if its closure  $\bar{L}$  is self-adjoint.

REMARK 7.5. *A d.d. symmetric  $D \xrightarrow{L} H$  is essentially self-adjoint if and only if  $\bar{D} = D^*$ .*

The next proposition gives a first, abstract criterion for self-adjointness.

PROPOSITION 7.4. *Let  $D \xrightarrow{L} H$  be d.d., closed, symmetric. Then,  $L$  is self-adjoint if and only if  $L^*$  is symmetric.*

PROOF. The only if direction is obvious. In the other direction, suppose  $L^*$  is symmetric. We claim that, then,  $D \subseteq D^*$  and  $\Gamma(L) \subseteq \Gamma(L^*)$ . If  $g \in D$ , in fact, then  $|\langle g, Lf \rangle| = |\langle Lg, f \rangle| \leq \|Lg\| \cdot \|f\|$ , hence  $g \in D^*$ . For such  $g$ 's,  $\langle Lg, f \rangle = \langle g, Lf \rangle = \langle L^*g, f \rangle$  for all  $f$  in  $D$ , hence  $L^*g = Lg$ . This shows that  $D \subseteq D^*$ , hence that  $\Gamma(L) \subseteq \Gamma(L^*)$ . We want to show the opposite inclusion.

Both  $\Gamma(L)$  and  $\Gamma(L^*)$  are closed. Let  $(g, L^*g) \in \Gamma(L^*) \ominus \Gamma(L)$ . For  $f \in D$ ,

$$0 = \langle (g, L^*g), (f, Lf) \rangle = \langle g, f \rangle + \langle L^*g, Lf \rangle,$$

then,  $f \mapsto \langle L^*g, Lf \rangle = -\langle g, f \rangle$  is bounded, hence  $L^*g \in D^*$  and  $L^{**}g = -g$ . Then,

$$\|L^*g\|^2 = \langle L^{**}g, g \rangle = -\|g\|^2,$$

which implies that  $g = 0$ . Thus,  $\Gamma(L) = \Gamma(L^*)$ , and  $L$  is self-adjoint.  $\square$

## 7.4. Resolvent and spectrum

**7.4.1. Resolvent and spectrum for a general d.d. symmetric operator.** Consider  $D \xrightarrow{L} H$  symmetric, d.d., and closed, and  $z \in \mathbb{C}$ . We say that  $z \in \sigma(L)$  belongs to the *spectrum* of  $L$  if  $L - zI = L - z : D \rightarrow H$  is not a bijection. This can happen because  $L - zI$  is not injective, or because it is not surjective. If it is injective, we can still define  $R(z) = (L - z)^{-1} : \text{Ran}(L - z) \rightarrow D$ .

LEMMA 7.5. *Let  $D \xrightarrow{L} H$  be closed, and  $z \in \mathbb{C}$ . If the inverse  $(L - z)^{-1} : H \rightarrow D$  exists, then it is bounded.*

PROOF. It sufficed to prove that  $\Gamma((L - z)^{-1}) = \tilde{\Gamma}(L - z)$  is closed, and apply the closed graph theorem. Suppose  $f_n \rightarrow f$ , and  $(L - z)f_n \rightarrow g$  in  $H$ , so that  $Lf_n \rightarrow zf + g$ . Since  $L$  is closed,  $Lf = zf + g$ , i.e.  $g = (L - z)f$ , showing that  $\tilde{\Gamma}(L - z)$  is closed.  $\square$

As in the case of bounded operators, or of Banach algebras, we deduce that the *resolvent set*  $\rho(L) = \mathbb{C} \setminus \sigma(L)$  of  $L$  is open.

**EXERCISE 7.2.** Let  $D \xrightarrow{L} H$  be closed. Use a Neumann series to show that  $\rho(L)$  is open. Deduce that the map  $z \mapsto R(z)$ ,  $R : \rho(L) \rightarrow \mathcal{B}(H)$ , is holomorphic.

**EXERCISE 7.3.** Let  $M_x$  be the position operator  $M_x f(x) = xf(x)$  on  $L^2(\mathbb{R})$ , with domain  $D = \{f \in L^2 : \int_{\mathbb{R}} (1+x^2)|f(x)|^2 dx < \infty\}$ . Show that  $M_x$  is self-adjoint and that its spectrum is  $\sigma(M_x) = \mathbb{R}$  (which is not compact).

The map  $z \mapsto R(z) = (L - z)^{-1}$  is the *resolvent function* of  $L$ ,  $R : \rho(L) \rightarrow \mathcal{B}(H)$ .

In order to define  $R(z) = (L - z)^{-1} : \text{Ran}(L - z) \rightarrow D$  we just need  $L - z$  to be injective, and not necessarily surjective. In the case of symmetric operators we have some quantitative information.

**LEMMA 7.6.** Let  $D \xrightarrow{L} H$  be d.d., symmetric and closed, and let  $z \in \mathbb{C} \setminus \mathbb{R}$ . Then,  $\text{Ran}(L - z)$  is closed,  $R(z) : \text{Ran}(L - z) \rightarrow D$ , and

$$(7.4.1) \quad \|R(z)\| \leq \frac{1}{|\text{Im}(z)|}.$$

**PROOF.** By symmetry of  $L$ , for  $f \in D$  we have  $\langle f, Lf \rangle \in \mathbb{R}$ , hence,

$$\begin{aligned} |-\text{Im}z| \cdot \|f\|^2 &= |\text{Im}\langle f, (L - z)f \rangle| \\ &\leq \|(L - z)f\| \cdot \|f\|, \end{aligned}$$

showing that

$$(7.4.2) \quad \|(L - z)f\| \geq |\text{Im}z| \cdot \|f\|.$$

This implies that  $L - z$  is injective.

If  $D \ni f_n$  and  $(L - z)f_n \rightarrow g$ , so that  $\{(L - z)f_n\}$  is Cauchy, then  $\{f_n\}$  is Cauchy as well by (7.4.2). Thus  $Lf_n \rightarrow g + zf = Lf$ , because  $L$  is closed, so  $g = (L - z)f$ . This shows that  $\text{Ran}(L - z)$  is closed.

Finally, equation (7.4.2) is equivalent to (7.4.1).  $\square$

**PROPOSITION 7.5.** Let  $D \xrightarrow{L} H$  be d.d., symmetric, and closed.

(i) If  $z, w \in \mathbb{C}$  and  $R(z), R(w)$  are everywhere defined (i.e.  $z, w \in \rho(L)$ ), then

$$(7.4.3) \quad R(z) - R(w) = (z - w)R(z)R(w).$$

In particular  $R(z)R(w) = R(w)R(z)$  and  $R(z)R(w)$



(ii) If  $z \in \mathbb{C} \setminus \mathbb{R}$ , and  $R(z), R(\bar{z}) \in \rho(L)$ , then  $R(z)^* = R(\bar{z})$ .

PROOF. (i) Formally,

$$\begin{aligned} R(z) - R(w) &= (L - z)^{-1} - (L - w)^{-1} \\ &= (L - z)^{-1}[(L - w) - (L - z)](L - w)^{-1} \\ &= (L - z)^{-1}(z - w)(L - w)^{-1} \\ &= (z - w)R(z)R(w). \end{aligned}$$

The expression after the first equality is well defined, since

$$H \xrightarrow{(L-w)^{-1}} D \xrightarrow{L} H \xrightarrow{(L-z)^{-1}} D.$$

(ii) We have  $D \ni u = R(z)x, v = R(\bar{z})y$  if and only if  $x = (L - z)u$ ,  $y = (L - \bar{z})v$ , hence

$$\begin{aligned} \langle x, R(\bar{z})y \rangle &= \langle (L - z)u, v \rangle = \langle u, (L - \bar{z})v \rangle \\ &= \langle R(z)x, y \rangle. \end{aligned}$$

The equalities hold for all  $x, y$  in  $H$  because  $R(z)$  and  $R(\bar{z})$  are defined on  $H$ . Hence  $R(z)^* = R(\bar{z})$ .  $\square$

The operator  $D \xrightarrow{L} H$  is *positive* if  $\langle Lf, f \rangle \geq 0$  for all  $f \in H$ .

REMARK 7.6. If  $H$  is, as in our case, a Hilbert space on  $\mathbb{C}$ , then positive operators are symmetric. In fact,

$$\langle f, Lf \rangle = \overline{\langle f, Lf \rangle} = \langle Lf, f \rangle,$$

and by polarization it follows that  $\langle f, Lg \rangle = \overline{\langle g, Lf \rangle}$ . It suffices to show that the polarization identity holds even if an operator is sandwiched between the vectors,

$$\langle x, Ly \rangle = \frac{1}{4}[\langle x+y, L(x+y) \rangle - \langle x-y, L(x-y) \rangle - i\langle x+iy, L(x+iy) \rangle + i\langle x-iy, L(x-iy) \rangle].$$

The same is not true for Hilbert spaces over  $\mathbb{R}$ . For instance, the matrix  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  defines a positive operator from  $\mathbb{R}^2$  to itself, which is not symmetric.

LEMMA 7.7. Let  $D \xrightarrow{L} H$  be a d.d. positive, closed operator, and let  $z \in \mathbb{C}$  with  $\operatorname{Re} z < 0$ . Then,  $R(z)$  is defined on  $\operatorname{Ran}(L - z)$ , which is closed, and

$$\|R(z)\| \leq \frac{1}{|\operatorname{Re} z|}.$$

The proof is similar to that of lemma 7.6.

PROOF. For  $f \in D$ ,  
 (7.4.4)  
 $-\operatorname{Re}(z)\|f\|^2 \leq \langle (L-z)f, f \rangle \leq \|(L-z)f\| \cdot \|f\|$ , hence,  $|\operatorname{Re} z| \cdot \|f\| \leq \|(L-z)f\|$ .

As in the previous lemma this implies that  $L-z$  is injective, that  $\operatorname{Ran}(L-z)$  is closed, and the estimate on  $\|R(z)\|$ .  $\square$

We have a first, useful criterion for self-adjointness.

PROPOSITION 7.6. *Let  $D \xrightarrow{L} H$  be d.d., symmetric, and closed, let  $D^* \xrightarrow{L^*} H$  be its adjoint, and let  $z \in \mathbb{C} \setminus \mathbb{R}$ .*

- (i)  $z \in \rho(L)$  if and only if  $L^* - \bar{z}$  is injective;
- (ii) if  $L = L^*$  is self-adjoint, then  $z \in \rho(L)$ ;
- (iii) if  $z, \bar{z} \in \rho(L)$ , then  $L$  is self-adjoint.

COROLLARY 7.3. *Let  $D \xrightarrow{L} H$  be d.d., symmetric, and closed. Then,  $L$  is self-adjoint if and only if  $\pm i \in \rho(L)$ , i.e. if and only if  $R(i), R(-i)$  are defined on  $H$ .*

PROOF. (i)  $R(z)$  is not defined on  $H$  if and only if  $\operatorname{Ran}(L-z) \neq H$ , i.e. if and only if there exists  $0 \neq v \perp \operatorname{Ran}(L-z)$ , since we proved in lemma 7.7 that  $\operatorname{Ran}(L-z)$  is closed. For all  $x \in D$ , that is,  $\langle v, (L-z)x \rangle = 0$ . Thus,

$$|\langle v, Lx \rangle| = |z| \cdot |\langle v, x \rangle| \leq |z| \cdot \|v\| \cdot \|x\|,$$

which shows that  $v \in D^*$ , the domain of  $L^*$ . Moreover,

$$\langle (L^* - \bar{z})v, x \rangle = \langle v, (L-z)x \rangle = 0$$

for all  $x \in D$ , hence  $(L^* - \bar{z})v = 0$  because  $D$  is dense, showing that  $L^* - \bar{z}$  is not injective.

Viceversa, if  $0 \neq v \in \ker(L^* - \bar{z})$  for some  $v \in D^*$ , then for  $x \in D$

$$0 = \langle (L^* - \bar{z})v, x \rangle = \langle v, (L-z)x \rangle,$$

i.e.  $v \in D^\perp$ , contradicting the fact that  $L$  is d.d.

(ii) By (i), if  $z \notin \rho(L)$  then there is  $0 \neq v \in D^* = D$  such that  $(L - \bar{z})v = 0$ , which contradicts the inequality  $\|(L - \bar{z})v\| \geq |\operatorname{Im} z| \cdot \|v\| \neq 0$  in lemma 7.6.

(iii) Let  $g \in D^*$ . For all  $f \in H$ , then, using proposition 7.5 in the first equality,

$$\begin{aligned}\langle f, R(\bar{z}(L^* - \bar{z})g) \rangle &= \langle R(z)f, (L^* - \bar{z})g \rangle = \langle (L - z)R(z)f, g \rangle \\ &= \langle f, g \rangle.\end{aligned}$$

The second equality holds because  $R(z)f \in D(L - z)$ , the domain of  $L - z$ . Thus  $R(\bar{z})(L^* - \bar{z})g = g$ , hence  $g \in \text{Ran}(R(\bar{z})) = D$ . This shows that  $D^* \subseteq D$ , and the opposite inclusion holds for all closed, d.d., symmetric operators.  $\square$

The basic criterion of self-adjointness in corollary 7.3 is less innocent than it seems. It requires, in fact, to show that for all  $a$  in  $H$  the equations

$$(L \pm i)x = a$$

have a solution  $x \in D$ . If  $L$  is a differential operator, we need to solve differential equations, and the domain  $D$  of the operator has to be chosen in such a way that the solution  $x$  exists for all data  $a$ . Since  $L \pm i$  must be invertible, the domain for which  $L$  is self-adjoint, if it exists, it is uniquely determined. In concrete situations, this and other criteria require tools from "hard analysis".

**PROPOSITION 7.7.** *Let  $D \xrightarrow{L} H$  be d.d., closed and symmetric, and  $z \in \mathbb{C} \setminus \mathbb{R}$ .*

- (i) *If  $z \in \rho(L)$ , then  $w \in \rho(L)$  whenever  $|z - w| < |\text{Im}z|$ .*
- (ii) *If  $z \in \rho(L)$  and  $w \in \mathbb{C} \setminus \mathbb{R}$  is on the same side of  $z$  with respect to the real line, then  $w \in \rho(L)$ .*

**PROOF.** Item (i) can be easily proved by a Neumann series argument, and item (ii) is a consequence of (i) and of the fact that half-planes are path-connected.  $\square$

For positive operators we have an analogous proposition.

**PROPOSITION 7.8.** *Let  $D \xrightarrow{L} H$  be d.d., closed, symmetric, and positive, and let  $w \in \mathbb{C} \setminus [0, +\infty)$ . Then,  $w \in \rho(L)$  if and only if  $L$  is self-adjoint.*

**IF.** By corollary 7.3,  $L$  is self-adjoint if and only if  $\pm i \in \rho(L)$ . Draw a curve from  $i$  to  $w$  in the complex plane, which does not meet  $[0, \infty)$ . We can use Neumann series on overlapping discs centered on the curve to show that  $w \in \rho(L)$ , using the fact that  $\|R(z)\| \leq \frac{1}{|\text{Im}z|}$  if  $z \notin \mathbb{R}$ , and that  $\|R(z)\| \leq \frac{1}{|\text{Re}z|}$  if  $\text{Re}z < 0$ .

In the opposite direction, we use the same reasoning on curves starting at  $w$  and ending at  $\pm i$ , respectively, showing that  $\pm i \in \rho(L)$ , hence that  $L$  is self-adjoint.  $\square$

Sometimes it is easier working with essentially self-adjoint operators, which do not require optimal information on the domain.

**PROPOSITION 7.9.** *Let  $D \xrightarrow{L} H$  be d.d., symmetric.*

- (i)  *$L$  is essentially self-adjoint if and only if  $\text{Ran}(L \pm i)$  are dense in  $H$ .*
- (ii) *If  $L \geq 0$ , then  $L$  is essentially self-adjoint if and only if  $\text{Ran}(L + 1)$  is dense in  $H$ .*
- (iii)  *$L$  is essentially self-adjoint if and only if  $\ker(L^* \pm i) = 0$ .*

**PROOF.** (i)  $L$  is essentially self-adjoint  $\iff \overline{L}$  is self-adjoint  $\iff \overline{L} \pm i$  are onto  $\iff$  for all  $y$  in  $H$  there is  $\bar{x}$  in  $\overline{D}$  such that  $(\overline{L} \pm i)\bar{x} = y \iff$  for all  $y \in H$  there is a sequence  $D \ni x_n \rightarrow \bar{x}$  such that  $Lx_n \rightarrow z = L\bar{x}$ , so that  $y = z \pm i\bar{x}$ . The latter holds if and only if for all  $y$  in  $H$  we have  $y = \lim(Lx_n \pm ix_n)$ , i.e. if and only if  $y \in \overline{\text{Ran}(L \pm i)}$ . This means that  $\text{Ran}(L \pm i)$  are dense in  $H$ .

(ii) The proof is similar, using the criterion of self-adjointness for positive operators.

(iii)  $\ker(L^* \pm i) \ni h$  if and only if  $f \mapsto \langle h, Lf \rangle$  is bounded on  $D$  and

$$\langle h, Lf \rangle = \langle L^*h, f \rangle = \mp i \langle h, f \rangle = \langle h, \pm if \rangle,$$

i.e.  $\langle h, (L \mp i)f \rangle = 0$  for all  $f \in D$ :  $\text{Ran}(L \pm i)$  is dense, then, if and only if  $\ker(L^* \pm i) = 0$ .  $\square$

**7.4.2. The bounded operator case.** If  $H \xrightarrow{L} H$  is bounded and self-adjoint (the Hellinger-Toeplitz theorem says that if  $L$  is symmetric, then it is bounded), the results obtained so far are easily summarized. For  $z \in \rho(L)$ , let  $R(z) = (L - z)^{-1}$ .

- (i)  $\sigma(L) \subset \mathbb{R}$  is compact (we already proved compactness in the context of Banach algebras; the proof consisted in showing that the Neumann series for  $R(z) = (L - z)^{-1} = z^{-1}(L/z - I)^{-1}$  converges for large  $z$ ). If  $L \geq 0$ , then  $\sigma(L) \subset [0, \infty)$ . The map  $z \mapsto R(z)$  is holomorphic from  $\rho(L)$  to  $\mathcal{H}$ .

(ii) We have the estimates:

$$\|R(z)\| \leq \frac{1}{|\operatorname{Im} z|} \text{ if } L \text{ is self-adjoint,}$$

$$\|R(x)\| \leq \frac{1}{|x|} \text{ if } x < 0 \text{ and } L \geq 0.$$

(iii) We still have the identities for the resolvent:

$$R(z) - R(w) = (z - w)R(z)R(w),$$

$$R(z)^* = R(\bar{z}).$$

There are two reasons for a real  $\lambda$  to belong to  $\sigma(L)$ .

- (i)  $L - \lambda$  is not injective: there exists  $v \neq 0$  in  $H$  such that  $Lv = \lambda v$ . In this case we say that  $\lambda$  is an *eigenvalue* of  $L$ , and  $v$  is an *eigenvector* corresponding to  $\lambda$  (or that  $v \in E_\lambda$  belongs to the *eigenspace* relative to  $\lambda$ ).
- (ii)  $L - \lambda$  is not surjective.

### 7.5. Intermezzo: the momentum operator $\mathbf{D} = i^{-1} \frac{d}{dx}$ on various domains

In this section we discuss the different operator theoretic properties of the *momentum operator*  $\mathbf{D} = i^{-1} \frac{d}{dx}$  when it is defined on this or that domain. The right theoretical (as well as practical) framework to place the corresponding phenomenology is that of the von Neumann index theorem. An in depth discussion of the topic for physicists, with plenty of examples, is in [Self-adjoint extensions of operators and the teaching of quantum mechanics \(2001\)](#), Guy BONNEAU Jacques FARAUT, Galliano VALENT. The proof of the index theorem can be found, e.g., in M. Reed, B. Simon [Methods of modern mathematical physics: II Fourier Analysis, Self-Adjointness \(1975\)](#), Academic Press, p. 135-141.

We will use the same symbol  $\mathbf{D}$  for the operator on different domains. We will, however, always specify its domain, not to cause confusion.

**7.5.1. The real line.** The operator  $(\mathbf{D}, C_c^1(\mathbb{R})), C_c^1(\mathbb{R}) \xrightarrow{\mathbf{D}=i^{-1}d/dx} L^2(\mathbb{R})$ , is symmetric and has graph

$$\Gamma(\mathbf{D}, C_c^1(\mathbb{R})) = \{(\varphi, \mathbf{D}\varphi) : \varphi \in C_c^1(\mathbb{R})\}.$$

Since the Fourier transform is unitary (by Plancherel), we can consider the multiplication operator  $M_\omega = \widehat{\mathbf{D}}$  in place of  $\mathbf{D}$ , where the starting domain is  $\mathcal{F}(C_c^1(\mathbb{R}))$ .

Let

$$(7.5.1) \quad H^1(\mathbb{R}) = \{\varphi \in L^2(\mathbb{R}) : M_\omega \widehat{\varphi} \in L^2(\mathbb{R})\}$$

which is normed by

$$\|\varphi\|_{H^1}^2 = \int_{\mathbb{R}} (1 + \omega^2) |\widehat{\varphi}(\omega)|^2 d\omega.$$

This is the *Sobolev space*  $H^1(\mathbb{R})$  in disguise, but we do not need here being aware of it. For  $\varphi \in C_c^1(\mathbb{R})$ ,

$$\|\varphi\|_{H^1}^2 = \|\varphi\|_{L^2}^2 + \|\varphi'\|_{L^2}^2,$$

which is the norm induced by the  $L^2 \times L^2$  norm on the graph  $\Gamma(\mathbf{D}, C_c^1(\mathbb{R}))$ .

PROPOSITION 7.10. *The closure of  $\mathcal{F}(C_c^1(\mathbb{R})) \xrightarrow{M_\omega} L^2(\mathbb{R})$  is*

$$\mathcal{F}(H^1(\mathbb{R})) \xrightarrow{M_\omega} L^2(\mathbb{R}).$$

Moreover, this extension is self-adjoint.

PROOF. We show first that  $\overline{\Gamma(M_\omega, \mathcal{F}(C_c^1(\mathbb{R})))} \subseteq \Gamma(M_\omega, \mathcal{F}(H^1))$ . Suppose  $(\psi, \eta)$  is in the closure of  $\Gamma(M_\omega, \mathcal{F}(C_c^1(\mathbb{R})))$ :

$$\mathcal{F}(C_c^1(\mathbb{R})) \ni \psi_n \rightarrow \psi \text{ and } M_\omega \psi_n \rightarrow \eta \text{ in } L^2.$$

Then, there is a subsequence  $\psi_{n_j}(\omega) \rightarrow \psi(\omega)$  a.e.  $\omega$ , and a smaller subsequence  $\omega \psi_{n_{j_l}}(\omega) \rightarrow \eta(\omega)$ . Hence,  $\omega \psi(\omega) = \eta(\omega)$  a.e.  $\omega$ . Hence,  $(\psi, \eta) \in \Gamma(M_\omega, \mathcal{F}(H^1))$ .

In the opposite direction, consider  $\psi = \widehat{\varphi} \in \mathcal{F}(H^1(\mathbb{R}))$ . We use an approximation of the identity. Let  $h \in C_c^\infty(\mathbb{R})$  with support in  $[-1, 1]$ ,  $h \geq 0$  and such that  $\int_{\mathbb{R}} h(x) dx = 1$ . Let  $h_n(x) = nh(nx)$ , so that  $\widehat{h_n}(0) = 1$ ,  $|\widehat{h_n}(\omega)| \leq 1$ , and  $\widehat{h_n}(\omega) = \widehat{h}(\omega/n) \rightarrow 1$  uniformly on compact sets. The function  $h_n * \varphi$  lies in  $L^2(\mathbb{R}) \cap C^\infty(\mathbb{R})$ . Also,  $(h_n * \varphi)' \in L^2(\mathbb{R})$  because

$$\|(h_n * \varphi)'\|_{L^2}^2 = \frac{1}{2\pi} \|\omega \widehat{h_n} \widehat{\varphi}\|_{L^2}^2 \leq \frac{1}{2\pi} \|\omega \widehat{\varphi}\|_{L^2}^2 < \infty.$$

We have

$$\|M_\omega \widehat{\varphi} - M_\omega \widehat{h_n * \varphi}\|_{L^2} = \|\widehat{\varphi} \omega (\widehat{h_n} - 1)\|_{L^2} \rightarrow 0 \text{ as } n \rightarrow \infty$$

by dominated convergence, since  $\omega \widehat{\varphi}(\omega) \in L^2$ . Then,

$$\|h_n * \varphi - \varphi\|_{H^1}^2 = \|h_n * \varphi - \varphi\|_{L^2}^2 + \frac{1}{2\pi} \|M_\omega \widehat{\varphi} - M_\omega \widehat{h_n * \varphi}\|_{L^2}^2 \rightarrow 0 \text{ as } n \rightarrow \infty.$$

The function  $\psi = h_n * \varphi$  might not lie in  $C_c(\mathbb{R})$ . Fix a function  $\eta \in C^\infty(\mathbb{R})$  such that  $0 \leq \eta \leq 1$ ,  $\eta(x) = 1$  for  $|x| \leq 1$  and  $\eta(x) = 0$  for  $|x| \geq 2$ , and let  $\psi_n(x) = \psi(x)\eta(x/n) = \psi(x)\eta_n(x)$ . Then,  $\psi_n \in C_c^1(\mathbb{R})$ ,

$$\|\psi - \psi_n\|_{L^2} \rightarrow 0 \text{ as } n \rightarrow \infty$$

by dominated convergence, and

$$\begin{aligned} \|(\psi - \psi_n)'\|_{L^2} &= \|\psi'(1 - \eta_n) - \psi\eta_n'\|_{L^2} \\ &\leq \|\psi'(1 - \eta_n)\|_{L^2} + \|\psi\eta_n'\|_{L^2}. \end{aligned}$$

Since  $\psi' \in L^2$ , the first term tends to zero as  $n \rightarrow \infty$  by dominated convergence; while the second term goes to zero as  $n \rightarrow \infty$  because

$$\|\psi\eta_n'\|_{L^2}^2 = \frac{1}{n^2} \int_{n \leq |x| \leq 2n} |\psi(x)|^2 \eta'(x/n) dx \rightarrow 0 \text{ as } n \rightarrow \infty.$$

We have finished the proof that  $\overline{\Gamma(M_\omega, \mathcal{F}(C_c^1(\mathbb{R})))} = \Gamma(M_\omega, \mathcal{F}(H^1))$ :  $\Gamma(M_\omega, \mathcal{F}(H^1))$  is (modulo a unitary map) the closure of  $(\mathbf{D}, C_c^1(\mathbb{R}))$ .

To finish the proof, we have to show that  $\mathcal{F}(H^1) \xrightarrow{\mathbf{D}} L^2(\mathbb{R})$  is self-adjoint. Let  $g \in L^2(\mathbb{R})$  be such that the following map is bounded on  $\mathcal{F}(H^1)$ :

$$\begin{aligned} f &\mapsto \frac{1}{2\pi} \langle g, M_\omega f \rangle_{L^2} = \frac{1}{2\pi} \int_{\mathbb{R}} \overline{g(\omega)} \omega f(\omega) d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \overline{\omega g(\omega)} f(\omega) d\omega. \end{aligned}$$

By density of  $H^1(\mathbb{R})$  in  $L^2(\mathbb{R})$ , it must be  $M_\omega g \in L^2(\mathbb{R})$ , hence,  $g \in \mathcal{F}(H^1)$ , which proves self-adjointness.  $\square$

### 7.5.2. $\mathbf{D}$ on some domains of functions on $[-\pi, \pi]$ .

7.5.2.1.  $\mathbf{D}$  on  $C_c^1(-\pi, \pi)$ . The operator  $C_c^1(-\pi, \pi) \xrightarrow{\mathbf{D}} L^2[-\pi, \pi]$  is densely defined and it is symmetric,

$$(7.5.2) \quad \int_{-\pi}^{\pi} \overline{\varphi(x)} i^{-1} \psi'(x) dx = i^{-1} [\overline{\varphi(x)} \psi(x)]_{-\pi}^{\pi} + \int_{-\pi}^{\pi} \overline{i^{-1} \varphi'(x)} \psi(x) dx = \int_{-\pi}^{\pi} \overline{i^{-1} \varphi'(x)} \psi(x) dx.$$

Its closure is symmetric as well. We show that it is not self-adjoint.

For a complex number  $k$ , let  $e_k(x) = e^{kx}$ . Then,  $e_k \in L^2[-\pi, \pi]$  and  $\mathbf{D}e_k = i^{-1} k e_k$ . The function  $e_k$  belongs to the domain of  $\mathbf{D}^*$ . In fact,

$$\varphi \mapsto \langle e_k, \mathbf{D}\varphi \rangle_{L^2} = \int_{-\pi}^{\pi} e^{\bar{k}x} i^{-1} \varphi'(x) dx = i\bar{k} \int_{-\pi}^{\pi} e^{\bar{k}x} i^{-1} \varphi(x) dx$$

is bounded in  $L^2$ -norm on  $C_0^1(-\pi, \pi)$ . If  $D^*$  were self-adjoint, it would be symmetric, but this would imply that

$$i\bar{j}\langle e_j, e_k \rangle_{L^2} = \langle D^* e_j, e_k \rangle_{L^2} = \langle e_j, D^* e_k \rangle_{L^2} = i^{-1}k\langle e_j, e_k \rangle_{L^2},$$

which implies that  $j = -k$  for all  $j, k$  real, which is absurd.

The closure of  $C_c^1(-\pi, \pi) \xrightarrow{\mathbf{D}} L^2[-\pi, \pi]$  provides an example of closed, densely defined, symmetric operator which is not self-adjoint. We will see below, however, that this operator has infinitely many self-adjoint extensions.

**7.5.2.2.  $\mathbf{D}$  on  $C_{per}^1[-\pi, \pi]$ .** We consider here  $C_{per}^1[-\pi, \pi] \xrightarrow{\mathbf{D}} L^2[-\pi, \pi]$ , which is densely defined and symmetric in virtue of the periodicity: the boundary term in (7.5.2) vanish for different reasons. Clearly  $C_{per}^1[-\pi, \pi] \supset C_c^1(-\pi, \pi)$ , hence, the closure of  $C_{per}^1[-\pi, \pi] \xrightarrow{\mathbf{D}} L^2[-\pi, \pi]$  is a closed, symmetric extension of the closure of  $C_c^1(-\pi, \pi) \xrightarrow{\mathbf{D}} L^2[-\pi, \pi]$ .

We can reason like we did with  $C_c(\mathbb{R})$ , with some simplifications. Let  $H_{per}^1$  be the space of those  $\varphi \in L^2[-\pi, \pi]$  such that  $M_n \widehat{\varphi} \in \ell^2(\mathbb{Z})$ , where  $(M_n h)(n) = n \cdot h(n)$ . We show first that  $C_{per}^1$  is dense in  $H^1$  with respect to the norm

$$\|\varphi\|_{H^1}^2 = \frac{1}{2\pi} \|\widehat{\varphi}\|_{\ell^2}^2 + \frac{1}{2\pi} \|M_n \widehat{\varphi}\|_{\ell^2}^2 = \frac{1}{2\pi i} \sum_n (1 + n^2) |\widehat{\varphi}(n)|^2.$$

Let  $\varphi_N(x) = \frac{1}{2\pi} \sum_{|n| \leq N} \widehat{\varphi}(n) e^{inx}$ , a trigonometric polynomial which surely sits in  $C_{per}^1$ .

$$\|M_n \widehat{\varphi} - M_n \widehat{\varphi_N}\|_{\ell^2}^2 = \sum_{|n| > N} n^2 |\widehat{\varphi}(n)|^2 \rightarrow 0 \text{ as } n \rightarrow \infty,$$

and *a fortiori*  $\|\varphi - \varphi_N\|_{L^2} \rightarrow 0$  as  $N \rightarrow \infty$ .

As we did before, one first shows that, modulo a passage to the frequency side, the closure of  $\Gamma(\mathbf{D}, C_{per}^1)$  can be identified with  $\Gamma(M_n, \mathcal{F}(H^1))$ ; then that  $H^1 \xrightarrow{\mathbf{D}} L^2$  is self-adjoint. Details are left to the non-indolent reader.

**7.5.2.3.  $\mathbf{D}$  on  $C_\lambda^1[0, 2\pi]$ .** The operator  $\mathbf{D}$  is symmetric also on the space

$$C_\lambda^1[0, 2\pi] = \{\varphi \in C^1[0, 2\pi] : \varphi(2\pi) = e^{i\lambda} \varphi(0)\},$$

where  $\lambda \in \mathbb{R}$  (and we might consider  $\lambda \in [0, 2\pi)$ ). The map  $\varphi \mapsto e^{-i\lambda x} \varphi = h$  maps  $C_\lambda^1$  onto  $C_{per}^1$  and we have

$$(\mathbf{D}\varphi)(x) = (\lambda h(x) + (\mathbf{D}h)(x))e^{i\lambda x},$$



hence,

$$\begin{aligned}\|\varphi\|_{L^2}^2 &= \frac{1}{2\pi} \sum_n |\widehat{h}(n)|^2, \\ \|\mathbf{D}\varphi\|_{L^2}^2 &= \frac{1}{2\pi} \sum_n (\lambda + n)^2 |\widehat{h}(n)|^2.\end{aligned}$$

For  $\varphi \in C_\lambda^1$  we have then that the norm induced by the graph of  $\mathbf{D}$  is

$$\|\varphi\|_{H_\lambda^1}^2 = \|\varphi\|_{L^2}^2 + \|\mathbf{D}\varphi\|_{L^2}^2 = \frac{1}{2\pi} \sum_n [1 + (\lambda + n)^2] |\widehat{h}(n)|^2.$$

We call  $H_\lambda^1$  the completion of  $C_\lambda^1$  with respect to this norm. The operator  $\mathbf{D}$  acts on  $\widehat{h}$  as  $M_{\lambda+n}$ , multiplication times  $\lambda+n$ . It is now routine verifying that  $M_{\lambda+n}$  is unitarily equivalent to the closure of  $\mathbf{D}$ , and that it is self-adjoint.

7.5.2.4. *Final remarks on the finite interval case.* The closure of the operator  $\mathbf{D}_0 = C_0^1 \xrightarrow{\mathbf{D}} L^2$  is not self-adjoint, but we have found a one-parametr family of self-adjoint extensions of it. The adjoint  $\mathbf{D}_0^*$  of  $\mathbf{D}_0$ , which is not symmetric, is a common extension of all these self-adjoint extensions we have found.

7.5.2.5. *The case of  $[0, \infty)$ .* Finally, we mention the case of  $C_c^1(0, \infty) \xrightarrow{\mathbf{D}} L^2(0, \infty)$ , which is symmetric and densely defined. To show that its closure is not self-adjoint, use a reasoning similar to that we used with  $C_c(0, 2\pi)$ , using the functions  $u_\lambda(x) = e^{-\lambda x}$  ( $\lambda > 0$ ). von Neumann's index theory shows that, in fact, this operator does not have any self-adjoint extension.

## 7.6. The spectral theorem: measurable calculus form

In this section we discuss the *measurable calculus form* of the spectral theorem: if  $D \xrightarrow{L} H$  is self-adjoint and  $m \in L_0^\infty(\mathbb{R})$ , we can make sense of the expression  $H \xrightarrow{m(L)} H$  as a bounded operator. The map  $m \mapsto m(L)$  is a  $*$ -homomorphism, and in fact a  $*$ -isomorphism after a notion of "almost everywhere" is set up. This is a far reaching extension of the holomorphic calculus in two respects: we go beyond holomorphic functions; we can deal with unbounded operators (in particular, with linear differential operators). The advantage of the measurable calculus form is that it is canonical: the operator  $m(L)$  is uniquely determined by  $m$  and  $L$ .

The statement, however, is long and somehow intricate. In the next section we will translate it into a more transparent form, showing that  $L$  is unitarily equivalent to a multiplication operator on a nice measure space. The price to pay is that we will not have universality. We will show how

to find a unitary map  $\Theta : H \rightarrow L^2(\nu)$ , where  $\nu$  is a suitable measure on a suitable measurable space, but neither  $\nu$ , nor  $\Theta$ , are uniquely determined: a number of arbitrary choices will be made.

**7.6.1. The measurable calculus for a general self-adjoint operator.** Let  $D \xrightarrow{L} H$  be self-adjoint and  $f \in H$ . Define:

$$(7.6.1) \quad F_{f,f}(z) = \langle f, R(z)f \rangle, \quad z \in \rho(L).$$

**THEOREM 7.12.** (i)  $F_{f,f}$  is holomorphic on  $\rho(L) \supseteq \mathbb{C} \setminus \mathbb{R}$ . Moreover,  $F_{f,f}(\bar{z}) = \overline{F_{f,f}(z)}$ .

(ii) We have the estimate

$$|F_{f,f}(z)| \leq \frac{\|f\|}{|\operatorname{Im}(z)|}.$$

(iii) If  $z \in \mathbb{C}_+$ , then  $F_{f,f}(z) \in \mathbb{C}_+ \cup \mathbb{R}$ .

(iv) We have

$$\lim_{y \rightarrow +\infty} i^{-1}yF_{f,f}(iy) = \|f\|^2.$$

(v) There exists  $\mu_{f,f} \geq 0$ , a finite Borel measure on  $\mathbb{R}$ , with

$$(7.6.2) \quad \langle f, R(z)f \rangle = F_{f,f}(z) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,f}(u)}{u - z}$$

for all  $z \in \mathbb{C} \setminus \mathbb{R}$ . Moreover,  $\frac{\mu_{f,f}(\mathbb{R})}{\pi} = \|f\|^2$ .

(vi) If  $L \geq 0$ , then  $\operatorname{supp}(\mu_{f,f}) \subseteq [0, +\infty)$ .

The measures  $\{\mu_{f,f} : f \in H\}$  are the *spectral measures* for the operator  $L$ . By polarization, we can extend the family to  $\{\mu_{f,g} : f, g \in H\}$ , where

$$(7.6.3) \quad F_{f,g}(z) := \langle f, R(z)g \rangle = \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,g}(u)}{u - z},$$

for  $z \in \mathbb{C} \setminus \mathbb{R}$ .

**PROOF.** (i) and (ii). We proved that the function  $z \mapsto R(z)$  is holomorphic from  $\mathbb{C} \setminus \mathbb{R}$  to  $(H)$ , hence  $F_{f,f}$  is holomorphic, and

$$|F_{f,g}| \leq \|R(z)\| \cdot \|f\| \cdot \|g\| \leq \frac{\|f\| \cdot \|g\|}{|\operatorname{Im} z|}.$$

(iii). If  $z = x + iy$  with  $y > 0$ , then

$$\operatorname{Im} F_{f,f}(z) = \operatorname{Im} \langle f, R(z)f \rangle = \operatorname{Im} \langle (L-z)R(z)f, R(z)f \rangle = y \langle R(z)f, R(z)f \rangle \geq 0.$$

(iv) and (v). By (ii), (iii), and Herglotz theorem, there exists a positive, finite, Borel measure  $\mu_{f,f}$  on the real line such that (7.6.2) holds for  $z \in \mathbb{C}_+$ . Since  $R(z)^* = R(\bar{z})$  and the measure is real valued, the relation extends to  $\mathbb{C}_-$ . After splitting the Cauchy kernel in its real and imaginary parts, we have

$$i^{-1}yF_{f,f}(iy) = \frac{i^{-1}}{\pi} \int_{\mathbb{R}} \frac{uy}{u^2 + y^2} d\mu_{f,f}(u) + \frac{1}{\pi} \int_{\mathbb{R}} \frac{y^2}{u^2 + y^2} d\mu_{f,f}(u) = i^{-1}A(y) + B(y).$$

For the imaginary part we have

$$A(y) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{t}{t^2 + 1} d\mu_{f,f}(ty) \rightarrow 0 \text{ as } y \rightarrow \infty$$

because  $\mu_{f,f}$  is finite and  $h(t) = \frac{1}{1+t^2} \in C_0(\mathbb{R})$ . We are left with the task of proving that  $B(y) \rightarrow \|f\|^2$  as  $y \rightarrow \infty$ . We start by computing

$$\begin{aligned} 2i\operatorname{Im} F_{f,f}(iy) &= \langle f, R(iy)f \rangle - \overline{\langle f, R(iy)f \rangle} = \langle f, [R(iy) - R(-iy)]f \rangle \\ &= \langle f, 2iyR(-iy)R(iy)f \rangle \text{ by the resolvent identity} \\ &= 2iy\|R(iy)f\|^2. \end{aligned}$$

Pick  $f = R(i)g \in D$  (recall that  $[R(i)](H) = D$ ):

$$\begin{aligned} y\operatorname{Im} F_{f,f}(iy) &= y^2\|R(iy)R(i)g\|^2 \\ &= y^2\left\| \frac{[R(iy) - R(i)]g}{iy - i} \right\|^2 \text{ by the resolvent identity} \\ &= \frac{y^2}{(y-1)^2} \| [R(iy) - R(i)]g \|^2 \\ &\rightarrow \|R(i)g\|^2 \end{aligned}$$

since  $\|R(iy)\| \leq \frac{1}{y}$ . More generally, if  $f, g \in H$ , then

$$\begin{aligned} |y(\|R(iy)f\| - \|R(iy)R(i)g\|)| &\leq y^2\|R(iy)f - R(iy)R(i)g\|^2 \\ &\leq y^2\|R(iy)\|^2\|f - R(i)g\|^2 \\ &\leq \|f - R(i)g\|^2, \end{aligned}$$

because  $\|R(iy)\| \leq 1/y$ . Thus,

$$\begin{aligned} |y\|R(iy)f\| - \|f\|| &\leq |y(\|R(iy)f\| - \|R(iy)R(i)g\|)| \\ &\quad + |y\|R(iy)R(i)g\| - \|R(i)g\|| + \|R(i)g - f\| \\ &\leq 2\|f - R(i)g\| + |y\|R(iy)R(i)g\| - \|R(i)g\||. \end{aligned}$$

For fixed  $\epsilon > 0$  choose first  $g$  such that  $\|f - R(i)g\| \leq \epsilon$ , then  $y_\epsilon$  such that  $\|y\|R(iy)R(i)g\| - \|R(i)g\| \leq \epsilon$  for  $y > y_\epsilon$ . This shows that  $\|y\|R(iy)f\| \rightarrow \|f\|$ , as wished.

(vi). This property will be more easily proved after we have the spectral theory in its multiplication form.  $\square$

LEMMA 7.8. *If  $\mu, \nu$  are finite Borel measures on  $\mathbb{R}$  such that  $\int_{\mathbb{R}} \frac{d\mu(u)}{u-z} = \int_{\mathbb{R}} \frac{d\nu(u)}{u-z}$  for all  $z \in \mathbb{C} \setminus \mathbb{R}$ , then  $\mu = \nu$ .*

PROOF. For  $z = x + iy$ ,  $y > 0$ , let  $h_y(x) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu(u)}{u-z}$ , so that

$$k_y(x) := \frac{h_y(x) - h_{-y}(x)}{2i} = \frac{1}{\pi} \int_{\mathbb{R}} \frac{y d\mu(u)}{(u-x)^2 + y^2} = (P_y * \mu)(x),$$

hence,  $\widehat{k}_y(\omega) = \widehat{\mu}(\omega) e^{-y|\omega|}$ . If  $h_y = h_{-y} = 0$ , then  $\widehat{\mu} = 0$ , hence  $\mu = 0$ .  $\square$

As usual, uniqueness leads to formulas.

PROPOSITION 7.11. *The map  $(f, g) \mapsto \mu_{f,g}$  from  $H \times H$  to the space of the finite Borel measures on  $\mathbb{R}$  is sesquilinear and*

$$\overline{\mu_{f,g}} = \mu_{g,f}.$$

PROOF. We just prove the last item, the other proofs being similar.

$$\begin{aligned} \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,g}(u)}{u-z} &= \langle f, R(z)g \rangle = \overline{\langle g, R(\bar{z})f \rangle} \\ &= \overline{\frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{g,f}(u)}{u-\bar{z}}} \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\overline{\mu_{g,f}}(u)}{u-z}. \end{aligned}$$

$\square$

Since  $\frac{\mu_{f,f}(\mathbb{R})}{\pi} = \|f\|^2$ , by polarization we have  $\frac{\mu_{f,g}(\mathbb{R})}{\pi} = \langle f, g \rangle$ . We denote by  $L_0^\infty(\mathbb{R})$  the space of the Borel measurable, bounded functions on  $\mathbb{R}$ , normed with the sup norm. The subscript means that we do not have a notion of almost everywhere, and neither of essential supremum. For  $m \in L_0^\infty$ , we define  $\Gamma(f, g)$  by

$$(7.6.4) \quad \Gamma(f, g) := \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,g}(u), \quad \Gamma : H \times H \rightarrow \mathbb{C}.$$

By proposition 7.11,  $\Gamma$  is sesquilinear, and

$$|\Gamma(f, g)| \leq \|m\|_{L_0^\infty} \frac{|\mu|_{f,g}(\mathbb{R})}{\pi} \leq C \|m\|_{L_0^\infty} \|f\| \cdot \|g\|.$$

By a corollary of the Banach-Steinhaus theorem. there is a bounded operator  $m(L)$  on  $H$  such that  $\Gamma(f, g) = \langle f, m(L)g \rangle$ ; i.e.

$$(7.6.5) \quad \langle f, m(L)g \rangle = \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,g}(u).$$

The map  $m \mapsto m(L)$  is the *measurable calculus* for  $L$ . We already know how it forks for some maps  $m$ :

$$(7.6.6)$$

$$m(x) = 1 \implies m(L) = I \text{ is the identity operator : } \langle f, g \rangle = \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{f,g}(u);$$

$$(7.6.7)$$

$$m(x) = (u - z)^{-1} \implies m(L) = R(z) \text{ if } z \notin \mathbb{R} : \langle f, R(z)g \rangle = \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{u - z} d\mu_{f,g}(u).$$

We will split the measurable calculus theorem into two statements.

**THEOREM 7.13.** [*Spectral theorem, measurable calculus II*] *There exists a unique linear map  $\Lambda : L_0^\infty(\mathbb{R}) \rightarrow \mathcal{L}(H)$  such that*

(i) *for  $z \in \mathbb{C} \setminus \mathbb{R}$ ,*

$$\langle f, \Lambda(R(z))g \rangle = \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,g}(u)}{u - z};$$

(ii) *if  $m_n$  is a uniformly bounded sequence in  $L_0^\infty(\mathbb{R})$ , and  $m_n \rightarrow m$  pointwise, then  $\Lambda(m_n) \xrightarrow{WOT} \Lambda(m)$ .*

*The map is the one in (7.6.5),*

$$\langle f, \Lambda(m)g \rangle = \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,g}(u).$$

**PROOF.** The map  $m \mapsto m(L)$  has properties (i,ii). To prove uniqueness, we check that the hypothesis of the Lebesgue-Hausdorff theorem are verified. By (i) and linearity,

$$\begin{aligned} \left\langle f, \Lambda \left( \frac{R(z) - R(\bar{z})}{2i} \right) g \right\rangle &= \frac{1}{2\pi i} \int_{\mathbb{R}} \left( \frac{1}{u - z} - \frac{1}{u - \bar{z}} \right) d\mu_{f,g}(u) \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{y}{(x - u)^2 + y^2} d\mu_{f,g}(u) \\ &= P_y[\mu_{f,g}](x). \end{aligned}$$

Let  $\mathbf{P}(z) = \frac{R(z) - R(\bar{z})}{2i} = \Lambda(\tau_x P_y)$ , where  $\tau_x \psi(u) = \psi(u - x)$ . Then,

$$\langle f, \mathbf{P}(z)g \rangle = P_y[\mu_{f,g}](x).$$

For  $\varphi \in C_c(\mathbb{R})$ ,

$$P_y[\varphi](u) = \lim_{N \rightarrow \infty} \frac{1}{\pi} \sum_{n=-N}^N \frac{y}{(u - n/N)^2 + y^2} \varphi(n/N) \frac{1}{N},$$

where the convergence holds at least pointwise. Using (ii) in the last equality,

$$\begin{aligned} \frac{1}{\pi} \int_{\mathbb{R}} P_y[\varphi](u) d\mu_{f,g}(u) &= \frac{1}{\pi} \int_{\mathbb{R}} P_y[\varphi](u) d\mu_{f,g}(u) \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \left( \lim_{N \rightarrow \infty} \frac{1}{\pi} \sum_{n=-N}^N \frac{y}{(u - n/N)^2 + y^2} \right. \\ &\quad \left. \varphi(n/N) \frac{1}{N} \right) d\mu_{f,g}(u) \\ &= \left\langle f, \lim_{N \rightarrow \infty} \frac{1}{\pi} \sum_{n=-N}^N \mathbf{P}(-n/N + iy) \varphi(n/N) \frac{1}{N} g \right\rangle \\ &= \left\langle f, \lim_{N \rightarrow \infty} \frac{1}{\pi} \sum_{n=-N}^N \Lambda(\tau_{n/N} P_y \varphi) \varphi(n/N) \frac{1}{N} g \right\rangle \\ &= \langle f, \Lambda(P_y[\varphi])g \rangle. \end{aligned}$$

Using again (ii),

$$\begin{aligned} \frac{1}{\pi} \int_{\mathbb{R}} \varphi(u) d\mu_{f,g}(u) &= \frac{1}{\pi} \int_{\mathbb{R}} \lim_{y \rightarrow 0^+} P_y[\varphi](u) d\mu_{f,g}(u) \\ &= \left\langle f, \Lambda\left(\lim_{y \rightarrow 0^+} P_y[\varphi]\right)g \right\rangle \\ &= \langle f, \Lambda(\varphi)g \rangle. \end{aligned}$$

Thus,

$$\langle f, \Lambda(\varphi)g \rangle = \frac{1}{\pi} \int_{\mathbb{R}} \varphi(u) d\mu_{f,g}(u)$$

for  $\varphi \in C_c(\mathbb{R})$ . By (ii) again, this extends to  $C_0(\mathbb{R})$ , and we can apply the Lebesgue-Hausdorff theorem.  $\square$

**THEOREM 7.14.** *[Spectral theorem, measurable calculus II] Let  $D \xrightarrow{L} H$  be self-adjoint.*

- (i) The map  $m \mapsto m(L)$  satisfies  $m(L)^* = \overline{m}(L)$ .
- (ii)  $d\mu_{f,R(z)g}(u) = \frac{1}{u-z}d\mu_{f,g}(u)$  for all  $z \in \mathbb{C} \setminus \mathbb{R}$ .
- (iii)  $d\mu_{f,m(L)g}(u) = m(u)d\mu_{f,g}(u)$  for all  $m \in L_0^\infty$ .
- (iv)  $m \mapsto m(L)$  is multiplicative (it is a  $*$ -morphism of  $*$ -algebras).
- (v)  $\|m(L)\| \leq \|m\|_{L_0^\infty}$ .
- (vi) For  $E \subseteq \mathbb{R}$ , a Borel set, define  $\mu(E) = \chi_E(L)$ . Then,  $\mu$  is a **projection valued measure** (p.v.m.):
  - (a)  $\mu(E) : H \rightarrow H$  is an orthogonal projection;
  - (b) if  $E = \cup_{n=1}^\infty E_n$  with the  $E_n$ 's Borel, disjoint, then  $\mu(E) = \sum_{n=1}^\infty \mu(E_n)$ , where the series converges in the strong operator topology.
  - (c)  $\mu(\emptyset) = 0$  and  $\mu(\mathbb{R}) = I$ .
- (vii) If  $E$  is bounded, then  $\text{Ran}(\mu(E)) \subseteq D$ .
- (viii) For  $f \in H$  and  $g \in D$ ,  $d\mu_{f,Lg}(u) = u d\mu_{f,g}(u)$ .
- (ix) Let  $D_c = \cup_{N=1}^\infty \text{Ran}(\mu[-N, N])$ . Then,  $D_c$  is dense in  $D$  and  $D_c \xrightarrow{L} D_c$ .
- (x) We have  $D = \{f \in H : \frac{1}{\pi} \int_{\mathbb{R}} |u|^2 d\mu_{f,f}(u) < \infty\}$ . In particular,  $D \xrightarrow{m(L)} D$  if  $m \in L_0^\infty$ .
- (xi) If  $m, m' \in L_0^\infty$ ,  $m'(u) = um(u)$ , then
  - (a) if  $f \in H$ , then  $m(L)f \in D$  and  $m'(L)f = Lm(L)f$ ;
  - (b) if  $f \in D$ , then  $m'(L)f = m(L)Lf$ .

PROOF. (i). Linearity is clear. Also,

$$\begin{aligned}
 \langle f, m(L)^* g \rangle &= \overline{\langle g, m(L)f \rangle} = \overline{\frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{g,f}(u)} \\
 &= \frac{1}{\pi} \int_{\mathbb{R}} \overline{m(u)} d\overline{\mu_{g,f}(u)} = \frac{1}{\pi} \int_{\mathbb{R}} \overline{m}(u) d\mu_{f,g}(u) \\
 &= \langle f, \overline{m}(L)^* g \rangle.
 \end{aligned}$$

(ii) We have

$$\begin{aligned} \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,R(z)g}(u)}{u-w} &= \langle f, R(w)R(z)f \rangle = \frac{\langle f, [R(w) - R(z)]f \rangle}{w-z} \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{w-z} \left( \frac{1}{u-w} - \frac{1}{u-z} \right) d\mu_{f,g}(u) \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{u-w} \frac{1}{u-z} d\mu_{f,g}(u), \end{aligned}$$

hence,

$$d\mu_{f,R(z)g}(u) = \frac{1}{u-z} d\mu_{f,g}(u).$$

(iii). From (ii),

$$d\mu_{f,R(z)g} = \frac{d\mu_{f,g}(u)}{u-z} = \frac{\overline{d\mu_{g,f}(u)}}{u-\bar{z}} = \overline{d\mu_{g,R(\bar{z})f}(u)} = d\mu_{R(\bar{z})f,g},$$

hence,

$$\begin{aligned} \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,m(L)g}(u)}{u-z} &= \langle f, R(z)m(L)g \rangle = \langle R(\bar{z})f, m(L)g \rangle \\ &= \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{R(\bar{z})f,g}(u) = \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,R(z)g}(u) \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{m(u)}{u-z} d\mu_{f,g}(u). \end{aligned}$$

(iv). By (iii),

$$d\mu_{f,m'(L)m(L)g}(u) = m'(u) d\mu_{f,m(L)g}(u) = m'(u)m(u) d\mu_{f,g}(u) = d\mu_{f,[m'm](L)g}(u).$$

(v).

$$\begin{aligned} \|m(L)f\|^2 &= \langle f, m(L)^* m(L)f \rangle = \langle f, \overline{m}(L)m(L)f \rangle = \langle f, |m|^2(L)f \rangle \\ &= \frac{1}{\pi} \int_{\mathbb{R}} |m(u)|^2 d\mu_{f,g}(u) \\ &\leq \|m\|_{L_0^\infty}^2 \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{f,g}(u) = \|m\|_{L_0^\infty}^2 \|f\|^2. \end{aligned}$$

(vi). Since

$$\mu(E)^* = \overline{\chi_E}(L) = \chi_E(L) = \mu(E),$$

the operator  $\mu(E)$  is self-adjoint, and

$$\mu(E)^2 = \chi_E^2(L) = \chi_E(L) = \mu(E),$$



hence,  $\mu(E)$  is an orthogonal projection. Also, if  $E \cap F = \emptyset$ , then

$$\mu(E \cup F) = \chi_{E \cup F}(L) = \chi_E(L) + \chi_F(L) = \mu(E) + \mu(F),$$

thus

$$\begin{aligned} \left\| \left( \mu(E) - \sum_{n=1}^N \mu(E_n) \right) f \right\|^2 &= \left\| \left( \chi_E(L) - \chi_{\bigcup_{n=1}^N E_n}(L) \right) f \right\|^2 \\ &= \left\| \chi_{\bigcup_{n=N+1}^\infty E_n}(L) f \right\|^2 \\ &= \frac{1}{\pi} \int_{\bigcup_{n=N+1}^\infty E_n} d\mu_{f,f} = \sum_{n=N+1}^\infty \frac{1}{\pi} \int_{E_n} d\mu_{f,f} \\ &= \sum_{n=N+1}^\infty \left\| \mu(E_n) f \right\|^2 \rightarrow 0 \text{ as } N \rightarrow \infty \end{aligned}$$

because

$$\sum_{n=1}^\infty \left\| \mu(E_n) f \right\|^2 = \dots = \frac{1}{\pi} \int_E d\mu_{f,f} = \langle f, \mu(E) f \rangle \leq \|\mu(E)\|^2 \cdot \|f\|^2 \leq \|f\|^2,$$

hence the tail of the series tends to zero.

(vii). Recall that if  $z \notin \mathbb{R}$ , then  $H \xrightarrow{L} D$  is a bijection. It suffices to show, then, that for each  $f \in D$  there is  $h \in H$  such that  $\mu(E)f = R(z)h$ . For now only formally, such  $h$  is given by  $h = R(z)^{-1}\mu(E)f$ , so that

$$\langle g, h \rangle = \langle g, R(z)^{-1}\mu(E)f \rangle = \langle g, (L - z)\mu(E)f \rangle = \frac{1}{\pi} \int_{\mathbb{R}} (u - z)\chi_E(u) d\mu_{g,f}(u).$$

With this euristics at hands, we set  $m(u) = (u - z)\chi_E(u)$ ,  $m \in L_0^\infty$ , and compute:

$$\begin{aligned} \langle g, R(z)m(L)f \rangle &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{m(u)}{u - z} d\mu_{g,f}(u) = \frac{1}{\pi} \int_{\mathbb{R}} \chi_E(u) d\mu_{g,f}(u) \\ &= \langle g, \mu(E)f \rangle, \end{aligned}$$

which gives  $h = m(L)f$ .

(viii). If  $g \in D$  and  $z \notin \mathbb{R}$ , then  $g = R(z)h$  for some  $h \in H$ , and  $Lg = LR(z)h$ . Set  $m(u) = \frac{u}{u-w}$ , with  $w \notin \mathbb{R}$ . On the one hand,

$$\begin{aligned} \langle f, R(w)h \rangle &= \langle f, R(w)(L - z)g \rangle = \langle f, R(w)Lg \rangle - z\langle f, R(w)g \rangle \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,Lg}(u)}{u - w} - \frac{z}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,g}(u)}{u - w}. \end{aligned}$$

On the other hand,

$$\begin{aligned}
\langle f, R(w)h \rangle &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,h}(u)}{u-w} = \frac{1}{\pi} \int_{\mathbb{R}} \left( \frac{u}{u-z} - \frac{z}{u-z} \right) \frac{d\mu_{f,h}(u)}{u-w} \\
&= \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,R(z)h}(u) - \frac{z}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,R(z)h}}{u-w} \\
&= \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,g}(u) - \frac{z}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,g}}{u-w}.
\end{aligned}$$

Comparing the two expressions we have

$$\frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,Lg}(u)}{u-w} = \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,g}(u) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{u d\mu_{f,g}(u)}{u-w}(u),$$

which shows that  $d\mu_{f,Lg}(u) = u d\mu_{f,g}(u)$ , as wished.

As byproduct of the proof, we have that

$$(7.6.8) \quad \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{f,R(w)Lg}(u) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{d\mu_{f,Lg}(u)}{u-w} = \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,g}(u) = \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{f,m(L)g}(u),$$

hence,  $R(w)Lg = m(L)g$  for all  $g \in D$ . We have then:

$$(7.6.9) \quad R(w)L = m(L)|_D.$$

In particular,  $m(L)(D) \subseteq D$ , since the range of  $R(w)$  is  $D$ .

(ix). Let  $E_n = [-n, n]$ , so that by (vii) we have that  $\mu(E_n)(H) = \text{Ran}(\mu(E_n)) \subseteq D$ . If  $f = \mu(E_n)h$ , by (viii)

$$\begin{aligned}
d\mu_{g,Lf}(u) &= u d\mu_{g,f}(u) = u d\mu_{g,\chi_{E_n}h}(u) = u \chi_{E_n}(u) d\mu_{g,h}(u) \\
&= m(u) d\mu_{g,h}(u) \text{ with } m(u) = u \chi_{E_n}(u), \text{ bounded,} \\
&= m(u) \chi_{E_n}(u) d\mu_{g,h}(u) = m(u) d\mu_{g,f}(u) \\
&= d\mu_{g,m(L)f},
\end{aligned}$$

so that  $Lf = m(L)f = \mu(E_n)m(L)f \in \text{Ran}(\mu(E_n))$ . This shows that  $\text{Ran}(\mu(E_n)) \xrightarrow{L} \text{Ran}(\mu(E_n))$ , hence that  $D_c$  is mapped by  $L$  into itself.

We have to show that  $D_c$  is dense in  $H$ . For  $f \in H$ , in fact,

$$\|\mu(E_n)f - f\|^2 = \int_{\mathbb{R} \setminus E_n} d\mu_{f,f} \rightarrow 0$$

as  $n \rightarrow \infty$  by dominated convergence.

(x). If  $f \in D$ , then, by (viii).

$$\begin{aligned} \infty &> \|Lf\|^2 = \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{Lf, Lf}(u) = \frac{1}{\pi} \int_{\mathbb{R}} u d\mu_{Lf, f}(u) \\ &= \frac{1}{\pi} \int_{\mathbb{R}} u^2 d\mu_{f, f}(u). \end{aligned}$$

In the opposite direction, we could try the following:

$$\begin{aligned} \infty &> \frac{1}{\pi} \int_{\mathbb{R}} u^2 d\mu_{f, f}(u) = \lim_{n \rightarrow \infty} \frac{1}{\pi} \int_{E_n} u^2 d\mu_{f, f}(u) \\ &= \lim_{n \rightarrow \infty} \frac{1}{\pi} \int_{\mathbb{R}} u^2 d\mu_{\mu(E_n)f, \mu(E_n)f}(u) \\ &= \lim_{n \rightarrow \infty} \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{L\mu(E_n)f, L\mu(E_n)f}(u) \text{ by (viii), since } \mu(E_n)f \in D_c \subseteq D, \\ &= \lim_{n \rightarrow \infty} \|\mu(E_n)f\|^2. \end{aligned}$$

The idea is good, but short of showing that  $f \in D$ . We use the same calculation in estimating a Cauchy tail:

$$\begin{aligned} \|L\mu(E_{n+m})f - L\mu(E_n)f\|^2 &= \|L\mu(E_{n+m} \setminus E_n)f\|^2 \\ &= \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{L\mu(E_{n+m} \setminus E_n)f, L\mu(E_{n+m} \setminus E_n)f}(u) \\ &= \frac{1}{\pi} \int_{E_{n+m} \setminus E_n} u^2 d\mu_{f, f}(u) \\ &\rightarrow 0 \text{ as } n \rightarrow \infty \text{ by dominated convergence.} \end{aligned}$$

We have shown the existence of  $h = \lim_{n \rightarrow \infty} L\mu(E_n)f$ . Since  $L$  is closed and  $\mu(E_n)f \rightarrow f$ , we have that  $f \in D$  and  $h = Lf$ .

(xi). We compute:

$$\frac{1}{\pi} \int_{\mathbb{R}} u^2 d\mu_{m(L)f, m(L)f}(u) = \frac{1}{\pi} \int_{\mathbb{R}} u^2 |m(u)|^2 d\mu_{f, f}(u) = \int_{\mathbb{R}} |m'(u)|^2 d\mu_{f, f}(u) < \infty,$$

hence, by (x),  $m(L)f \in D$  if  $f \in H$ . We can now compute

$$\begin{aligned} \langle f, m'(L)g \rangle &= \frac{1}{\pi} \int_{\mathbb{R}} um(u) d\mu_{f, g}(u) = \frac{1}{\pi} \int_{\mathbb{R}} u d\mu_{f, m(L)g}(u) \\ &= \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{f, Lm(L)g}(u) \text{ since } m(L)g \in D \\ &= \langle f, Lm(L)g \rangle. \end{aligned}$$

We have shown that  $m'(L) = Lm(L)$ .

In the other direction,

$$\begin{aligned}\langle f, m'(L)g \rangle &= \frac{1}{\pi} \int_{\mathbb{R}} um(u) d\mu_{f,g}(u) \\ &= \frac{1}{\pi} \int_{\mathbb{R}} m(u) d\mu_{f,Lg}(u) \text{ since } g \in D \\ &= \frac{1}{\pi} \int_{\mathbb{R}} d\mu_{f,m(L)Lg}(u) = \langle f, m(L)Lg \rangle,\end{aligned}$$

hence,  $m'(L) = m(L)L$  on  $D$ . □

### 7.7. The spectral theorem: multiplication form

**7.7.1. The multiplicative form of the spectral theorem.** We proceed like in the case of the unitary operators. Let  $D \xrightarrow{L} H$  be self-adjoint. A subset  $V$  of  $H$  is *invariant for  $L$  in the sense of the measurable calculus* if  $m(L)(V) \subseteq V$  for all  $m \in L_0^\infty(\mathbb{R})$ .

LEMMA 7.9. *Let  $D \xrightarrow{L} H$  be a self-adjoint operator.*

- (i) *If  $V$  is invariant for  $L$ , then  $V$  is linear and closed.*
- (ii) *If  $V$  is invariant for  $L$ , then  $V^\perp$  is invariant, and  $\pi|_V m(L) = m(L)\pi|_V$ .*
- (iii) *If  $V$  is invariant for  $L$ , then  $D \cap V$  is dense in  $V$  and  $L(D \cap V) \subseteq V$  is closed and symmetric.*
- (iv) *If  $V$  is invariant for  $L$ , then  $D \cap V \xrightarrow{L} V$  is self-adjoint.*
- (v) *There is an index set  $A$  and invariant subspaces  $V_\alpha$ ,  $\alpha \in A$ , such that*

$$H = \bigoplus_{\alpha \in A} V_\alpha.$$

*If  $H$  is separable, the index set is at most countable.*

PROOF. (i) is clear. (ii) If  $g \in V^\perp$  and  $f \in V$ , then

$$\langle m(L)g, f \rangle = \langle g, m(L)^* f \rangle = \langle g, \overline{m}(L)f \rangle = 0,$$

thus  $m(L)g \in V^\perp$ . For  $h \in H$ ,

$$m(L)h = m(L)\pi_V h + m(L)\pi_{V^\perp} h \in V \oplus V^\perp,$$

hence,  $\pi_V(m(L)h) = m(L)\pi_V h$ , as wished.

(iii) At first sight it may not seem obvious that  $L|_V$  is densely defined in  $V$ . Let  $\mu$  be the projection valued measure associated with  $L$ , and recall that  $D_c = \cup_N > 0 \mu[-N, N]$  is dense in  $D$ . If  $h \in V$ , then  $\mu[-N, N]h \in V \cap D_c$  because  $V$  is invariant. Knowing this, the proof that  $D_c \cap V$  is dense in  $V$  is wholly similar to that that  $D_c$  is dense in  $H$ .

For  $h = \mu[-N, N]f \in V \cap D_c$  ( $f \in V$ ), we have that  $Lh = m(L)f$  with  $m(t) = t\chi_{[-N, N]}(t)$ , hence  $Lh \in V$  because  $V$  is invariant. This shows that  $L|_V$  maps  $D_c \cap V$  into  $V$ . Since  $L$  maps  $D_c$  into  $D_c$ ,  $L$  maps  $D_c \cap V$  into  $D_c \cap V$ .

We now verify that  $L(D \cap V) \subseteq V$ . Suppose  $g \in D \cap V \subseteq D$ . Since  $L$  is self-adjoint, there is  $f$  in  $H$ , unique, such that  $g = R(i)f$ . Then,

$$D \cap V \ni g = R(i)f = R(i)\pi_V f + R(i)\pi_{V^\perp} f,$$

with  $R(i)\pi_V f \in V$  and  $R(i)\pi_{V^\perp} f \in V^\perp$  because  $V$  and  $V^\perp$  are both invariant. Thus  $R(i)\pi_{V^\perp} f = 0$ , which in turns implies that  $\pi_{V^\perp} f = 0$ , since  $R(i)$  is a bijection. Hence,  $f \in V$ . We have shown that  $g = R(i)f$  with  $f \in V$ . Then,

$$Lg = (L - i)R(i)f + iR(i)f = f + R(i)f \in V,$$

as we had to prove, because  $V$  is invariant.

The symmetry of  $L|_V : D \cap V \rightarrow V$  follows immediately from what we have proved and the fact that  $L$  is symmetric. Finally,

$$\Gamma(L|_V) = \Gamma(L) \cap V \times V,$$

which is closed because  $\Gamma(L)$  and  $V$  are closed.

(iv) To show that  $L|_D : D \cap V \rightarrow H$  is self-adjoint, it suffices to verify that  $\pm i \in \rho(L|_V)$ , i.e. that  $L|_V \pm i : D \cap V \rightarrow V$  is invertible; something we did while proving (iii).

(v) The proof is step-by-step identical to the one for the unitary operator case.  $\square$

. We now proceed exactly like in §7.2.5. A vector  $f \in H$  is *cyclic* for the self-adjoint operator  $D \xrightarrow{L} H$  if  $\{m(L)f : m \in L_0^\infty(\mathbb{R})\}$  is dense in  $H$ .

**PROPOSITION 7.12.** *Suppose  $L$  has a cyclic vector  $f$ . Then, the map  $\Theta : L^\infty(\mathbb{R}) \rightarrow H$ ,  $\Theta(f) = f(L)x$ , extends to a unitary map from  $L^2(\mathbb{R}, \mu_{f,f})$  to  $H$ , such that*

$$(7.7.1) \quad ((\Theta^{-1}L\Theta)m)(t) = tm(t).$$

That is,  $L$  is unitarily equivalent to multiplication times  $t$  on  $L^2(\mathbb{R}, \mu_{f,f})$ ,

$$\begin{array}{ccc} L^2(\mathbb{R}, \mu_{f,f}) & \xrightarrow{\Theta} & H \\ M_t \downarrow & & \downarrow L \\ L^2(\mathbb{R}, \mu_{f,f}) & \xrightarrow{\Theta} & H \end{array}$$

PROOF. The map  $\Theta$  is an  $L^2(\mu_{f,f})$  isometry on  $L_0^\infty(\mathbb{R})$ ,

$$\|m(L)f\|^2 = \langle f, m(L)^* m(L)f \rangle = \langle f, |m|^2(L)f \rangle = \int_{\mathbb{R}} |m|^2 d\mu_{f,f}.$$

As such,  $\Theta$  extends to an isometry of  $\overline{L_0^\infty(\mathbb{R})}^{L^2(\mu_{f,f})}$ , the closure of  $L_0^\infty(\mathbb{R})$  in the  $L^2(\mu_{f,f})$  norm. The hypothesis that  $f$  is cyclic means that  $\Theta(L_0^\infty(\mathbb{R}))$  is dense in  $H$ , hence that  $\Theta$  is a surjective isometry.

Suppose now that  $m \in L_0^\infty(\mathbb{R})$  is such that  $xm(x) = m'(x)$  defines a bounded function. Such  $m$ 's are dense in  $L^2(\mu_{f,f})$ , and, by theorem 7.14 (ix),

$$L\Theta(m) = Lm(L)f = m'(L)f = \Theta(M_x m)f.$$

The relation extends by the unitary equivalence to the domain on  $M_x$ .  $\square$

**THEOREM 7.15.** [*Spectral theorem for self-adjoint operators: multiplicative form*] Let  $D \xrightarrow{L} H$  be a self-adjoint operator on a separable Hilbert space  $H$ . Then, there exists a locally compact space  $X$ , a finite Borel measure  $\nu$  on  $X$ , and a continuous function  $\varphi : X \rightarrow \mathbb{R}$ , and a unitary map  $\Theta : L^2(\nu) \rightarrow H$ , such that

$$(7.7.2) \quad (\Theta^{-1}U\Theta)m = \varphi m,$$

is multiplication times  $\varphi$ .

The proof follows from proposition 7.12 and lemma 7.9 (v) (with an index set  $A$  which is at most countable, in the separable case), line by line as in the corresponding statement for the unitary operators. In fact, the proof produces a rather concrete model for  $X$  and  $\varphi$ .

- (i)  $X = \mathbb{R} \times A$ , with the product topology ( $A$  is considered discrete);
- (ii)  $d\nu(x, n) = d\mu_{f_n, f_n}(x)$ , where  $f_n$  is the cyclic vector for the  $n^{\text{th}}$  summand in the decomposition of  $H$  into invariant vectors;
- (iii)  $\varphi(x, n) = x$ :  $\varphi$  multiplies times the coordinate  $x$  on each copy  $\mathbb{R} \times \{n\}$  of the real line.

A picture of  $\varphi$  can be obtained by changing coordinates in such a way  $\mathbb{R} \times \{n\}$  is mapped onto  $(-1/2 + n, 1/2 + n)$ .

### 7.7.2. Some consequences.

7.7.2.1. *The spectral theory of a multiplication operator.* We translate here in the world of self-adjoint operators what we saw in §7.2.6 for the unitary operators. We omit the proofs, which are essentially the same.

The multiplicative form of the spectral theorem tells us that any self-adjoint operator is unitarily equivalent to a multiplication operator  $M_\eta$  of the following form, for which we can explicitly write down the spectrum, the spectral measures, and so on.

- (a) *We have a locally compact space  $X$  and a finite, positive Borel measure  $\nu$ , which we can assume to be finite, and a surjective local homeomorphism  $\eta : X \rightarrow \mathbb{R}$ , such that  $M_\eta : L^2(\nu) \rightarrow L^2(\nu)$  is the operator  $M_\eta f = \eta f$ . Actually, we should specify the domain of  $M_\eta$ ,*

$$D(M_\eta) = \{f : X \rightarrow \mathbb{C} \text{ such that } \int_X (1 + \eta^2) |f|^2 d\nu < \infty\}.$$

- (b)  $\sigma(M_\eta) = \overline{\eta(\text{supp}(\nu))} \subseteq \mathbb{T}$  *is the spectrum of  $M_\eta$ .*

- (c) *Let  $f \in L^2(\nu)$ . The spectral measure  $\mu_{f,f}$  is such that, for any bounded function  $m : \mathbb{R} \rightarrow \mathbb{C}$ ,*

$$\int_X m(\eta) |f|^2 d\nu = \langle f, m(M_\eta) f \rangle_{L^2(\nu)} = \int_{\mathbb{R}} m d\mu_{f,f}.$$

*That is,  $d\mu_{f,f} = \eta_*(|f|^2 d\nu)$ . Similarly,  $d\mu_{f,g} = \eta_*(\bar{g} f d\nu)$ .*

- (d) *The p.v.m.  $\mu$  maps Borel measurable sets  $E$  in  $\mathbb{R}$  to projections of  $L^2(\nu)$  in the following fashion:*

$$\int_X \bar{f} [\mu(E)g] d\nu = \langle f, \mu(E)g \rangle_{L^2(\nu)} = \mu_{f,g}(E) = \int_X \chi_E(\eta) \bar{f} g d\nu,$$

*i.e.  $(\chi_E(\eta) = \chi_E \circ \eta)$*

$$(7.7.3) \quad \mu(E)g = \chi_E(\eta)g, \quad \mu(E) = M_{\chi_E(\eta)}.$$

- (e) *A Borel set  $E \subseteq \mathbb{R}$  is a null-set for  $\mu$  if and only if  $0 = \chi_E \circ \eta = \chi_{\eta^{-1}(E)} \nu$  - a.e., i.e. if and only if  $\nu(\eta^{-1}(E)) = 0$ .*

- (f) *By (e) and the topological properties of  $\eta$ ,  $\overline{\eta(\text{supp}(\nu))} = \text{supp}(\mu)$ .*

- (g) *By (b) and (f),*

$$(7.7.4) \quad \text{supp}(\mu) = \sigma(M_\eta).$$

7.7.2.2. *Some applications of the spectral theorem in its multiplicative form.*

THEOREM 7.16. *Let  $L : D \rightarrow H$  be self-adjoint,  $\{\mu_{x,x} : x \in H\}$  be the family of the spectral measures for  $L$ , and  $\mu$  be the projection valued measure associated with  $L$ . We have*

$$(7.7.5) \quad \text{supp}(\mu) = \sigma(L).$$

PROOF OF THEOREM 7.16. By the multiplicative form of the spectral theorem, it suffices to show it for multiplication operators. This is done in (7.7.4).  $\square$

COROLLARY 7.4. *With the same notation of theorem 7.16, we also have the following.*

- (i) *For a function  $f \in L^\infty(\mu)$ ,  $f(L) = 0$  if and only if  $f$  vanishes  $\mu$ -a.e on  $\sigma(L)$ . Thus, the map  $f \mapsto f\chi_{\sigma(L)}$  is the identity of  $L^\infty(\mu)$ . To highlight this fact, we also write  $L^\infty(\mu) = L^\infty(\sigma(L))$ , and for  $\varphi : \sigma(L) \rightarrow \mathbb{C}$  measurable and bounded, we write  $\varphi(L) = f(L)$ , where  $f$  is any measurable extension of  $\varphi$  to  $\mathbb{R}$ .*
- (ii) *If  $\{f_n : n \geq 1\}$  is a sequence in  $L^\infty(\sigma(L))$ , if  $\|f_n\|_{L^\infty(\sigma)} \leq C$  is uniformly bounded, and  $f_n \rightarrow f$   $\mu$ -a.e., then  $f_n(L) \rightarrow f(L)$  in the strong operator topology.*

Another application is a relationship between unitary and self-adjoint operators. We will see some deeper result when we discuss Stone's theorem of strongly continuous, one parameter groups of unitary operators.

THEOREM 7.17. (i) *Let  $H \xrightarrow{U} H$  be a unitary operator. Then there exist self-adjoint operators  $L$ , some of which are bounded, such that  $U = e^{iL}$ .*

(ii) *Viceversa, if  $D \xrightarrow{L} H$  is self-adjoint, then  $e^{itL}$  is a unitary operator for all  $t \in \mathbb{R}$ .*

PROOF. (i) See proposition 7.3. (ii) By the measurable calculus,  $L$  is unitarily equivalent to  $M_\eta$  on some measure space, with  $\eta$  measurable and real valued. It is easy to see, by the uniqueness of the measurable calculus, that  $e^{itM_\eta} = M_{e^{it\eta}}$ , and the latter is a unitary operator on  $L^2(\nu)$  for all  $t \in \mathbb{R}$ .  $\square$



### 7.8. One parameter groups of unitary operators and Stone's theorem

A *one parameter group of unitary operators* on a Hilbert space  $H$  is a map  $t \mapsto U_t$  from  $\mathbb{R}$  to the group  $\mathcal{U}(H)$  of the unitary operators on  $H$  such that

$$(7.8.1) \quad U_{s+t} = U_s U_t, \text{ and } U_0 = I.$$

The group is *strongly continuous* if the map is continuous with respect to the strong operator topology,

$$(7.8.2) \quad \lim_{t \rightarrow 0} U_t h = h \text{ in } H$$

for all  $h \in H$ .

PROPOSITION 7.13. Let  $D \xrightarrow{L} H$  be self-adjoint, and for  $t \in \mathbb{R}$  let  $U_t := e^{itL}$ . Then,

(i)  $t \mapsto U_t$  is a strongly continuous group of unitary operators;

(ii) for all  $f \in D$ ,

$$(7.8.3) \quad \lim_{t \rightarrow 0} \frac{U_t f - f}{t} = iL f$$

in  $H$ ;

(iii) if  $f \in H \setminus D$ , then the limit in 7.8.3 does not exist.

The proposition might be read as providing a solution for the abstract, general version of the *Schroedinger equation*. The equation

$$(7.8.4) \quad i^{-1} \frac{d}{dt} u(t) = Lu(t),$$

where  $u : \mathbb{R} \rightarrow H$ , has a solution  $u$  satisfying  $u(0) = f \in D$ , which is provided by

$$(7.8.5) \quad u(t) = e^{itL} f.$$

It is an important fact that the function  $u$  in (7.8.5) makes perfect sense, and it is strongly continuous, even when  $f \in H \setminus D$ .

PROOF. By the spectral theorem in multiplicative form, it suffices to prove the proposition for an operator of the form  $L = M_g$  on  $H = L^2(\nu)$ , where  $\nu$  is a finite Radon measure on a locally compact metric space  $X$ , and  $g : X \rightarrow \mathbb{R}$  is continuous. In this framework,  $U_t = M_{e^{itg}}$  is the operator of multiplication times  $e^{itg}$ . Since  $|e^{itg}| = 1$ , the operator is unitary, as we remarked earlier, and  $U_t^*$  is multiplication times  $e^{-itg}$ .

(i) We only have to verify continuity.

$$\|U_{t+s}f - U_tf\|_{L^2(\nu)}^2 = \int_X |e^{itg} - 1|^2 |f|^2 d\nu \rightarrow 0 \text{ as } s \rightarrow 0,$$

by dominated convergence.

(ii) We have

$$\frac{U_tf - f}{t} - iLf = \left( \frac{e^{itg} - 1}{t} - ig \right) f \rightarrow 0 \text{ as } t \rightarrow 0 \text{ } \nu - a.e. \text{ on } X,$$

and

$$\left| \frac{e^{itg} - 1}{t} \right| = |g| \left| \frac{e^{itg} - 1}{tg} \right| \leq C|g|$$

because  $|(e^{it} - 1)/t|$  is bounded on  $\mathbb{R} \setminus \{0\}$ . We have then that

$$\left| \left( \frac{e^{itg} - 1}{t} - ig \right) f \right| \leq (C + 1)|fg| \in L^2(\nu),$$

if  $f \in D$ , the domain of  $M_g$ . Then

$$\lim_{t \rightarrow 0} \left\| \left( \frac{e^{itg} - 1}{t} - ig \right) f \right\|_{L^2(\nu)} = 0,$$

by dominated convergence.

(iii) Viceversa, if there is  $h \in L^2(\nu)$  such that

$$\lim_{t \rightarrow 0} \left\| \frac{e^{itg} - 1}{t} f - h \right\|_{L^2(\nu)} = 0,$$

then, for some sequence  $t_n \rightarrow 0$ ,

$$ifg - h = \lim_{n \rightarrow \infty} \frac{e^{it_n g} - 1}{t_n} f - h = 0 \text{ } \nu - a.e.,$$

hence,  $L^2(\nu) \ni fg, f$ , which is the condition for membership of  $f$  in  $D$ .  $\square$

This simple proposition has a converse.

**THEOREM 7.18.** [*Stone's theorem*] Let  $\{U_t\}_{t \in \mathbb{R}}$  be a strongly continuous one parameter group of unitary operators on  $H$ . Then, there exists a unique self-adjoint operator  $D \xrightarrow{L} H$  such that  $U_t = e^{itL}$ .

I present here two proofs. The first one, from Reed and Simon, does not require the Lebesgue-Hausdorff theorem. The second one, from [Tao's notes](#), makes use of it.

**PROOF FROM [16].** For  $\varphi \in C_c^\infty(\mathbb{R})$  and  $f \in H$ , let

$$f_\varphi = \int_{\mathbb{R}} \varphi(t) U_t f dt,$$

which you might think of as a "smoothing" of the vector  $f$ , and let

$$D = \text{span}\{f_\varphi : \varphi \in C_c^\infty, f \in H\}.$$

Let  $\eta_\epsilon(t) = 1/\epsilon \eta(t/\epsilon)$  and approximation of the identity,

$$\eta \in C_c^\infty(\mathbb{R}), \quad \eta(t) = 0 \text{ if } |t| > 1, \quad \eta \geq 0, \quad \int_{\mathbb{R}} \eta(t) dx = 1.$$

We have then:

$$\begin{aligned} \|f_{\eta_\epsilon} - f\| &= \left\| \int_{\mathbb{R}} \eta_\epsilon(t) [U_t f - f] dt \right\| \leq \int_{\mathbb{R}} \eta_\epsilon(t) \sup_{|t| \leq \epsilon} \|U_t f - f\| dt \\ &\rightarrow 0 \text{ as } \epsilon \rightarrow 0. \end{aligned}$$

This shows that  $D$  is dense in  $H$ .

For  $f_\varphi \in D$ ,

$$\begin{aligned} \frac{U_s - I}{s} f_\varphi &= \int_{\mathbb{R}} \varphi(t) \frac{U_{s+t} - U_t}{s} f dt \\ &= \int_{\mathbb{R}} \frac{\varphi(\tau - s) - \varphi(\tau)}{s} U_\tau f d\tau \\ &\rightarrow - \int_{\mathbb{R}} \varphi'(\tau) U_\tau f d\tau \text{ in } H \text{ as } s \rightarrow 0, \end{aligned}$$

since

$$\lim_{s \rightarrow 0} \frac{\varphi(\tau - s) - \varphi(\tau)}{s} = -\varphi'(\tau),$$

uniformly with respect to  $\tau$ .

For  $f_\varphi \in D$ , it is now natural defining

$$L f_\varphi := i^{-1} f_{-\varphi'},$$

so that  $L : D \rightarrow D$ . Also, it is easily verified that  $U_s : D \rightarrow D$ , and in fact that  $U_s[f_\varphi] = [U_sf]_\varphi$ . We now verify that

$$(7.8.6) \quad U_s L f_\varphi = L U_s f_\varphi.$$

In fact,

$$\begin{aligned} U_s L f_\varphi &= U_s i^{-1} f_{-\varphi'} = i^{-1} [U_sf]_{-\varphi'} = L [U_sf]_\varphi \\ &= L U_s f_\varphi. \end{aligned}$$

Also,  $L$  is symmetric on  $D$ ,

$$\langle f_\varphi L g_\psi \rangle = \lim_{s \rightarrow 0} \left\langle f_\varphi, \frac{u_s - I}{is} g_\psi \right\rangle = \lim_{s \rightarrow 0} \left\langle \frac{u_{-s} - I}{-is} f_\varphi, g_\psi \right\rangle = \langle L f_\varphi g_\psi \rangle.$$

We have to prove that  $D \xrightarrow{L} H$  is essentially self-adjoint, i.e. that  $\ker(L^* - i) = 0$ . Suppose that  $u \in \ker(L^* - i)$ . Then, for all  $f \in D$ ,

$$\begin{aligned} \frac{d}{dt} \langle u, U_t f \rangle &= \langle u, i L U_t f \rangle = i \langle L^* u, U_t f \rangle \\ &= i \langle i u, U_t f \rangle = \langle u, U_t f \rangle. \end{aligned}$$

i.e.  $\alpha(t) = \langle u, U_t f \rangle$  satisfies  $\alpha'(t) = \alpha(t)$ ,  $\alpha(0) = \langle u, f \rangle$ . It must be

$$\alpha(t) = e^t \langle u, f \rangle.$$

On the other hand, it must be  $|\alpha(t)| = |\langle u, U_t f \rangle| \leq \|u\| \cdot \|f\|$ , hence  $0 = \alpha(0) = \langle u, f \rangle$ . Since  $D$  is dense in  $H$ ,  $u = 0$ .  $\square$

The second proof is more spectral. We can not use directly the spectral theorem in its multiplicative form, because we do not know at this stage if the operators  $U_t$  can be simultaneously "diagonalized". The proof has to go back to spectral measures, which have to be reconstructed from scratch to fit the present case, where we have a continuum of operators to work with simultaneously.

PROOF FROM [22]. **Step I.** By proposition 7.13, the domain of  $L$  and its action are determined by the group,

$$D = \left\{ f \in H : \lim_{t \rightarrow 0} \frac{U_t f - f}{t} \text{ exists in } H \right\}, \quad Lf = i^{-1} \lim_{t \rightarrow 0} \frac{U_t f - f}{t} f.$$

**Step II.** We show that (simultaneous diagonalization) for each  $f$  in  $H$  there is a finite, positive Radon measure  $\mu_{f,f}$  on  $\mathbb{R}$  such that:

$$\mu_{f,f}(\mathbb{R}) = \|f\|^2, \text{ and } \langle f, U_t f \rangle = \int_{\mathbb{R}} e^{itx} d\mu_{f,f}(x).$$

The function  $\varphi(t) = \langle f, U_t f \rangle$  is continuous on  $\mathbb{R}$ ,

$$|\varphi(t) - \varphi(s)| \leq \|f\| \cdot \|[U_t - U_s]f\| = \|f\| \cdot \|U_{t-s}f - f\| \rightarrow 0 \text{ as } t \rightarrow s,$$

uniformly, by the hypothesis of continuity. Also,  $\varphi$  is positive definite,

$$\begin{aligned} \sum_{l,m=1}^n \varphi(t_m - t_l) c_m \overline{c_l} &= \sum_{l,m=1}^n \langle c_l U_{t_l} f, c_m U_{t_m} f \rangle \\ &= \sum_{m=1}^n \|c_m U_{t_m} f\|^2 \geq 0. \end{aligned}$$

By Bochner's theorem a measure with the desired properties exists and it is unique. Also,

$$\mu_{f,f}(\mathbb{R}) = \int_{\mathbb{R}} e^{i0x} d\mu_{f,f}(x) = \langle f, I f \rangle = \|f\|^2.$$

**Step III.** By polarization, for all  $f, g \in H$  we can construct Radon, finite, complex measures  $\mu_{f,g}$  such that

$$\langle f, U_t g \rangle = \int_{\mathbb{R}} e^{itx} d\mu_{f,g}(x),$$

and the map  $(f, g) \mapsto \mu_{f,g}$  is sesquilinear. Moreover,

$$\begin{aligned} \int_{\mathbb{R}} e^{itx} d\mu_{f,g}(x) &= \langle f, U_t g \rangle = \langle U_{-t} f, g \rangle = \overline{\langle g, U_{-t} f \rangle} \\ &= \int_{\mathbb{R}} e^{itx} d\overline{\mu_{g,f}}(x), \end{aligned}$$

i.e.  $\mu_{f,g} = \overline{\mu_{g,f}}$ , by the injectivity of the Fourier transform on the space of the complex measures.

**Step IV.** For each  $m \in L_0^\infty(\mathbb{R})$  there is a unique operator  $m(L)$  in  $\mathcal{L}(H)$  such that

$$\langle f, m(L)g \rangle = \int_{\mathbb{R}} m(x) d\mu_{f,g}(x).$$

Clearly, the map  $m \mapsto m(L)$  is linear. Also,

$$(7.8.7) \quad \|m(L)\| \leq \|m\|_{L_0^\infty}.$$

Moreover, If  $m_n \rightarrow m$  pointwise,  $m_n, m \in L_0^\infty(\mathbb{R})$ , and the  $m_n$ 's are uniformly bounded, then

$$(7.8.8) \quad \lim_{n \rightarrow \infty} m_n(L) = m(L)$$

in the weak operator topology. Elementary estimates, the definition of  $\mu_{f,g}$ , and the fact that  $\mu_{f,f}(\mathbb{R}) = \|f\|^2$ , imply that

$$\left| \int_{\mathbb{R}} m(x) d\mu_{f,g}(x) \right| \leq C \|m\|_{L_0^\infty} (\|f\| + \|g\|).$$

Passing to sup for  $\|f\|, \|g\| \leq 1$ , we see that the sesquilinear form

$$(f, g) \mapsto \int_{\mathbb{R}} m(x) d\mu_{f,g}(x)$$

is bounded, hence that a bounded operator  $m(L) \in \mathcal{L}$  with the desired properties exists, and the estimate on its norm improves to  $\|m(L)\| \leq \|m\|_{L_0^\infty}$ , by estimating the integral.

The limit in (7.8.8) follows from dominated convergence.

**Step V.** We have

$$d\mu_{f,m(L)g}(x) = m(x) d\mu_{f,g}(x).$$

We start with  $m(x) = e^{isx}$ , which corresponds to, we will see,  $m(L) = U_s$ . For all  $t$ ,

$$(7.8.9) \quad \int_{\mathbb{R}} e^{itx} e^{isx} d\mu_{f,g}(x) = \int_{\mathbb{R}} e^{i(t+s)x} d\mu_{f,g}(x) = \langle f, U_{t+s}g \rangle$$

$$(7.8.10) \quad = \langle f, U_t(U_s g) \rangle = \int_{\mathbb{R}} e^{itx} d\mu_{U_s f, g}(x),$$

hence,  $e^{isx} d\mu_{f,g}(x) = d\mu_{U_s f, g}(x)$ . As a consequence,  $\mu_{f, U_t g} = \mu_{U_{-t} f, g}$ , since

$$\int_{\mathbb{R}} e^{isx} d\mu_{f, U_t g}(x) = \langle f, U_s U_t g \rangle = \langle U_{-t} f, U_s g \rangle = \int_{\mathbb{R}} e^{isx} d\mu_{U_{-t} f, g}(x)$$

and the Fourier transform is injective on the class of the finite Borel measures. Then, for  $m \in L_0^\infty$  and  $t \in \mathbb{R}$ ,

$$\begin{aligned} \int_{\mathbb{R}} e^{itx} m(x) d\mu_{f,g}(x) &= \int_{\mathbb{R}} m(x) d\mu_{f, U_t g}(x) = \int_{\mathbb{R}} m(x) d\mu_{U_{-t} f, g}(x) \\ &= \langle U_{-t} f, m(L)g \rangle = \langle f, U_t m(L)g \rangle \\ &= \int_{\mathbb{R}} e^{itx} d\mu_{f, m(L)g}(x), \end{aligned}$$

which, again by injectivity of the Fourier transform, implies  $m(x) d\mu_{f,g}(x) = d\mu_{f, m(L)g}(x)$ .<sup>3</sup>

---

<sup>3</sup>I thank Terence Tao for providing me this argument, which shortened a more intricate one in a previous version of the notes.

**Step VI.** *The map  $m \mapsto m(L)$  is a  $*$ -homomorphism from  $L_0^\infty(\mathbb{R})$  to  $\mathcal{L}(H)$ . We proceed exactly as in the case of the unitary operators. Linearity is clear from Step IV, and the multiplicative property follows from Step V,*

$$\begin{aligned} \langle f, (mn)(L)g \rangle &= \int_{\mathbb{R}} mnd\mu_{x,y} = \int_{\mathbb{R}} md\mu_{x,n(L)y} \\ &= \int_{\mathbb{R}} d\mu_{x,m(L)n(L)y} = \langle f, m(L)n(L)g \rangle, \end{aligned}$$

if  $f, g \in H$  and  $m, n \in L_0^\infty(\mathbb{R})$ . The fact that  $m(L)^* = \overline{m}(L)$  is left as an exercise.

**Step VII.** *For each Borel measurable  $E$  in  $\mathbb{R}$ , define  $\mu(E) \in \mathcal{L}(H)$  by*

$$(7.8.11) \quad \langle f, \mu(E)g \rangle = \int_E d\mu_{f,g}.$$

*Then,  $\mu : \mathbf{B}(\mathbb{R}) \rightarrow \mathcal{L}(H)$  is a projection valued measure.* The proof of this fact is wholly similar to the one in section 7.2.4.

**Step VIII.** *There exists a self-adjoint operator  $L$  having  $\mu$  as its spectral measure.* We first define  $L$  as a densely defined, symmetric operator, and show that it is essentially self-adjoint. Let

$$D_c = \cup_{N \geq 1} \mu[-N, N],$$

which is dense in  $H$  (the proof is identical to that we have seen when proving the spectral theorem). For  $g \in D_c$  define  $Lg$  by

$$\langle f, Lg \rangle := \int_{\mathbb{R}} u d\mu_{f,g}(u).$$

The integral converges absolutely because, for some  $N \geq 1$ ,  $d\mu_{f,g}(u) = d\mu_{f, \mu[-N, N]g}(u) = \chi_{[-N, N]}(u) d\mu_{f,g}(u)$ .

Since the function  $u \mapsto u$  is real valued,  $L$  is symmetric. Moreover,  $L(\mu[-N, N](H)) \subseteq \mu[-N, N](H)$ . From this, it is easy to deduce that  $(L \pm i)(H)$  is dense in  $H$ . In fact, on  $\mu[-N, N](H)$ ,

$$\langle f, (L \pm i)g \rangle = \int_{-N}^N (u \pm i) d\mu_{f,g}(u),$$

and we have  $1 \leq |u \pm i|^2 \leq N^2 + 1$ , hence  $L \pm i : \mu[-N, N](H) \rightarrow \mu[-N, N](H)$  is invertible with bounded inverse. Hence,  $L$  is essentially self-adjoint.

Next, we have to show that  $\mu$  is the p.v.m. associated to  $L$ , or, equivalently, that  $\mu_{f,g}$  are the spectral measures associated with  $L$ , let's call them  $\nu_{f,g}$ . If  $z \in \mathbb{C} \setminus \mathbb{R}$ , then

$$\langle f, (L - z)^{-1}g \rangle = \int_{\mathbb{R}} \frac{d\nu_{f,g}(u)}{u - z}.$$

If we verify that

$$(7.8.12) \quad \langle f, (L - z)^{-1}g \rangle = \int_{\mathbb{R}} \frac{d\mu_{f,g}(u)}{u - z},$$

we have  $\mu_{f,g} = \nu_{f,g}$  by the uniqueness of the Herglotz representation, and we are done.

Let  $m(u) = \frac{1}{u-z}$ , so that  $\langle f, m(L)g \rangle = \int_{\mathbb{R}} \frac{d\mu_{f,g}(u)}{u-z}$ . Then,

$$\langle f, (L - z)m(L)g \rangle = \int_{\mathbb{R}} (u - z)d\mu_{f,m(L)}(u) = \langle f, (u - z)m(z)d\mu_{f,g}(u) \rangle = \langle f, g \rangle.$$

Hence  $m(L) = (L - z)^{-1}$ , and (7.8.12) holds.

**Step IX. Conclusion.** By the multiplicative form of the spectral theorem, after a unitary map our operator  $L$  has the form  $L\varphi = \eta\varphi$  on  $L^2(X, \lambda)$  for some measure space  $(X, \lambda)$  (and  $\eta$  is measurable and real valued), and the operators  $U_t$  have the form  $U_t\varphi = e^{it\eta}\varphi$ . Stone's theorem now follows from its converse, proposition 7.13.  $\square$

## 7.9. Some applications to bounded and compact self-adjoint operators, and to the spectrum

Having at your disposal various versions of the spectral theorem for self-adjoint operators, especially the multiplicative one, by means of pictures, intuition, and measure theory, a number of conclusions can be drawn. We make a list of some of them, without striving for the greatest generality. The hints should be enough for you to come up with a proof.

### 7.9.1. Bounded operators, continuous calculus, spectrum.

- (i) *A self-adjoint operator  $L$  is bounded if and only if the corresponding p.v.m. measure has bounded support in  $\mathbb{R}$ . If  $\mu$  has support in  $[-N, N]$ , then  $L = L\mu[-N, N]$  is bounded. Viceversa, if the support of  $\mu$  is unbounded you can find elements  $f_n \in H$  with  $\|f_n\| = 1$  and  $\|Lf_n\| \geq n$ .*
- (ii) **Continuous calculus.** *Let  $L$  be a self-adjoint operator, and for  $\varphi \in C(\sigma(L))$ , consider the operators  $\varphi(L)$ , as we did in the measurable calculus. Then,  $\varphi \mapsto \varphi(L)$  is a homeomorphic isometry from  $C(\sigma)$  endowed with the sup norm and a subalgebra of  $\mathcal{L}(H)$  with respect to the operator norm.*
- (iii) **The spectral mapping theorem.** *Let  $D \xrightarrow{L} H$  be a self-adjoint operator (possibly unbounded), and let  $m \in L^\infty(\mu)$ , where  $\mu$  is the*



associated p.v.m. Then,  $\sigma(m(L)) = \text{ess-ran}(m)$ , where the essential range is with respect to  $\mu$ . If  $m \in C(\sigma(L))$ , then  $\sigma(m(L)) = m(\sigma(L))$ .

- (iv) An **eigenvalue**  $\lambda$  of  $D \xrightarrow{L} H$  is a scalar such that  $Lf = \lambda f$  for some  $0 \neq f \in D$ . Any such  $f$  is an **eigenvector** relative to  $\lambda$ , and the subspace of the eigenvectors relative to  $\lambda$  is the **eigenspace** relative to  $\lambda$ . It is clear that  $\lambda \in \sigma(L)$ , hence  $\lambda \in \mathbb{R}$  if  $L$  is self-adjoint. In this case, if  $\mu \neq \lambda$  are two eigenvalues, the corresponding eigenspaces are orthogonal. The set of the eigenvalues,  $\sigma_d(L)$ , is the **discrete spectrum** of  $L$ . To see this compute  $\langle f, Lg \rangle = \langle Lf, g \rangle$  when  $f, g$  are eigenvectors relative to  $\lambda, \mu$ , respectively.

- (v) If  $f$  is an eigenvector relative to  $\lambda$  and  $m \in C_c(\sigma(L))$ , then  $m(\lambda)$  is an eigenvalue, having  $f$  as eigenvector. If  $\|f\| = 1$ , then

$$m(\lambda) = \langle f, m(L)f \rangle = \int_{\mathbb{R}} m(x) d\mu_{f,f}(x)$$

for all  $m \in C_c(\mathbb{R})$ ; which correspond to the fact that  $\mu_{f,f}$  is a Dirac delta at  $\lambda$ .

- (vi) If  $\lambda \in \sigma(L)$  is not an eigenvalue, then  $L - \lambda$  has dense range in  $H$ . A simple way to prove this is by verifying the statement when  $L = M_\eta$  is a multiplication operator on  $L^2(X, \nu)$ , where  $X$  is locally compact and  $\nu$  is a finite Radon measure. There, you can show that  $\lambda$  is an eigenvalue if and only if  $\nu$  "contains" a positive multiple of Dirac delta  $a\delta_\zeta$  (i.e.  $\nu \geq a\delta_\zeta$  and  $\nu - a\delta_\zeta \perp \delta_\zeta$ ) and  $\lambda = \eta(\zeta)$ . If  $\lambda$  is not an eigenvalue, then,  $\eta^{-1}(\lambda) \subseteq X$  is a set which does not support point masses of  $\nu$ .

- (vii) Let  $\mu$  be the p.v.m. associated to  $D \xrightarrow{L} H$ , self-adjoint. Then,  $\lambda \in \mathbb{R} \setminus \{0\}$  is an eigenvalue of  $L$  if and only if  $\mu(\{\lambda\}) \neq 0$ . Let  $E = \mu(\{\lambda\})(H) \ni f$ . Then

$$\langle g, Lf \rangle = \langle g, L\mu(\{\lambda\})f \rangle = \int_{\{\lambda\}} x d\mu_{g,f}(x) = \lambda \int_{\{\lambda\}} d\mu_{g,f}(x) = \lambda \langle g, f \rangle.$$

Thus, if  $E \neq 0$ , then  $\lambda$  is an eigenvalue.

Viceversa, if  $\lambda$  is an eigenvalue with eigenvector  $f$ , then for all  $m \in C_c(\mathbb{R})$

$$\lambda \int_{\mathbb{R}} m(x) d\mu_{f,f}(x) = \lambda \langle f, m(L)f \rangle = \langle f, Lm(L)f \rangle = \int_{\mathbb{R}} xm(x) d\mu_{f,f}(x),$$

This implies that  $\mu_{f,f}$  is supported on  $\{\lambda\}$ , and

$$\|\mu(\{\lambda\})f\|^2 = \int_{\{\lambda\}} d\mu_{f,f}(x) = \int_{\mathbb{R}} d\mu_{f,f}(x) = \|f\|^2,$$

which shows that  $f = \mu(\{\lambda\})f \in E$ .

(viii) **The Cayley map**  $\Gamma : \mathbb{R} \rightarrow \mathbb{T}$ ,

$$\Gamma(x) = \frac{x - i}{x + i},$$

induces a bijection

$$L \mapsto U = \Gamma(L) = (L - i)(L + i)^{-1}$$

from the family of the self-adjoint operators onto the family of the unitary operators. If  $\mu_{f,g}$  are the spectral measures for  $U$  and  $\nu_{f,g}$  are those for  $L$ , and  $m \in L_0^\infty(\mathbb{T})$ , then

$$\begin{aligned} \langle f, m(U)g \rangle &= \int_{\mathbb{T}} m(e^{it}) d\mu_{f,g}(e^{it}) \\ &= \int_{\mathbb{R}} m(\Gamma(x)) d\mu_{f,g}(\Gamma(x)) \\ &= \int_{\mathbb{R}} (\Gamma^* m)(x) d([\Gamma^{-1}]_* \mu_{f,g})(x) \\ &= \langle f, m(\Gamma(L))g \rangle. \end{aligned}$$

Recall that the *pull-back* of  $\mathbb{T} \xrightarrow{m} \mathbb{C}$  by  $\Gamma$  is  $\mathbb{R} \xrightarrow{\Gamma^* m = m \circ \Gamma} \mathbb{C}$ , while the push-forward of a measure  $\mu$  on  $\mathbb{T}$  by  $\Phi = \Gamma^{-1}$  is defined by

$$(\Phi_* \mu)(E) = \mu(\Phi^{-1}(E)).$$

What we said can be rephrased as  $\nu_{f,g} = [\Gamma^{-1}]_* \mu_{f,g}$ .

**7.9.2. Compact self-adjoint operators and their inverses.** A bounded operator  $L : H \rightarrow H$  is *compact* if the image  $L(B_1)$  of the unit ball of  $H$  has compact closure in  $H$ . Here we always assume  $L$  to be self-adjoint.

- (i) *Let  $L$  be self-adjoint, bounded, and let  $\mu$  be the corresponding p.v.m. Then,  $L$  is compact if and only if  $\mu|_{\mathbb{R} \setminus (-\epsilon, \epsilon)}$  reduces to a finite number of point masses.* The if part is easy. For the only if, suppose there is  $\epsilon > 0$  for which the requirement fails. Then we can find infinitely many, disjoint, measurable subsets  $E_n$  in  $\mathbb{R} \setminus (-\epsilon, \epsilon)$  with  $\mu(E_n) \neq 0$ . Pick a unit vector  $f_n$  in each  $[\mu(E_n)](H)$ :  $\{f_n\}$  is bounded, but  $\{Lf_n\}$  does not have convergent subsequences.

- (ii) From (i) it is easy proving that self-adjoint  $L$  is compact if and only if there are  $L_n$  self-adjoint with finite rank such that  $\|L - L_n\| \rightarrow 0$  as  $n \rightarrow \infty$ .
- (iii) **Hilbert-Schmidt Theorem.** If  $L$  is self-adjoint, then  $L$  is compact if and only if (when  $H$  is separable and infinite dimensional) (a)  $\sigma(L) = \{\lambda_n\} \cup \{0\}$ , with finitely many  $\lambda_n$ 's, or  $\lambda_n \rightarrow 0$ ; (b) each  $\lambda_n$  is an eigenvalue, and its eigenspace is finite dimensional.
- (iv) An self-adjoint operator  $D \xrightarrow{\Delta} H$  has a compact inverse (separable, infinite dimensional case) if and only if: (a)  $\sigma(M) = \{\mu_n\} \subset \mathbb{R} \setminus \{0\}$  with  $|\mu_n| \rightarrow 0$ ; (b) each  $\mu_n$  is an eigenvalue, and its eigenspace is finite dimensional.
- (v) With  $M$  as in (iv), and  $M$  **positive**, the eigenvalues  $0 < \mu_1 \leq \dots \leq \mu_n \leq \dots$  can be computed using the **Rayleigh quotients**.

$$\mu_1 = \min_{0 \neq x \in H} \frac{\langle x, Mx \rangle}{\|x\|^2};$$

pick  $x_1$  such that the minimum is achieved

$$\mu_2 = \min_{0 \neq x \perp x_1} \frac{\langle x, Mx \rangle}{\|x\|^2} \geq \mu_1;$$

pick  $x_2$  such that the minimum is achieved

...

$$\mu_n = \min_{0 \neq x \perp x_1, \dots, x_{n-1}} \frac{\langle x, Mx \rangle}{\|x\|^2} \geq \mu_{n-1};$$

pick  $x_n$  such that the minimum is achieved

...



## Index

- Banach algebra, 60, 135
- Banach space, 5
  - dual, 60
  - of bounded operators, 60
  - reflexive, 62
  - valued functions: integral, 79
  - weak topology, 65
  - weak\* topology, 66
- Borel
  - $\sigma$ -algebra, 6
  - complex measure, 8
  - measurable function, 6
  - measure, 6, 33
- continuous calculus
  - unitary operator, 177
- cyclic vector, 184, 217
- densely defined
  - operator, 191
- equicontinuous, 107
- essentially self-adjoint operator,
  - 195, 200, 224, 227
- Fourier
  - coefficients, 15
- Fourier analysis
  - in Hilbert spaces, 46
  - trigonometric system, 51
- Fourier synthesis
  - in Hilbert spaces, 46
  - trigonometric system, 51
- Fourier transform, 13
  - of a measure, 13
- Gram-Schmidt algorithm, 47
- Hahn-Banach theorem
  - complex, 59
  - real, 58
- Hamel basis, 56
- harmonic function, 16
- Hilbert space, 40
  - dimension, 49
- holomorphic
  - Banach space valued function,
    - 74
  - calculus in a Banach algebra,
    - 144
- inner product, 35, 39
- invariant subspace, 216
  - unitary operator, 185
- lemma
  - of Riesz in Banach spaces, 93
  - resolvent identity, 147
- measurable calculus
  - self-adjoint operator, 209, 210
  - unitary operator, 179
- momentum operator, 13, 201
- Neumann series, 97, 138
- normed linear spaces, 1

- operator
  - adjoint, 45, 64
  - adjoint of a densely defined operator, 193
  - closed, 192
  - compact, 94
  - compact vs. finite rank, 95
  - compact, self-adjoint, 99
  - Hilbert-Schmidt, 104, 108
  - integral kernel, 106
  - positive, 197
  - self-adjoint, 221
  - symmetric, 192
  - trace, 106
  - unitary, 175, 221
- operator topology
  - strong, 61, 182, 189, 211, 220, 221
  - uniform, 61
  - weak, 61, 179, 180, 226
- orthogonal system, 45
- orthonormal basis, 45
- orthonormal system, 45
- Plancherel formula
  - abstract, 46
  - Fourier transform, 13
  - trigonometric system, 52
- polarization identity, 40
- sesquilinear forms, 81
- position operator, 196
- positive definite function, 33
- projection, 42
  - lemma, 41
- projection valued measure
  - self-adjoint operator, 211, 219
- projection valued measures
  - unitary operator, 182, 188, 189
  - unitary operator: nullsets, 183
- Radon measure (finite): finite
  - Borel measure, 6
- resolvent
  - function, 138, 196
  - set, 138, 196
- Schroedinger equation, 172, 221
- self-adjoint
  - densely defined operator, 194
- sesquilinear form, 81
- spectral measures
  - self-adjoint operator, 206
  - unitary operator, 179, 187
- spectrum
  - self-adjoint operator, 219
  - symmetric, densely defined operator, 195
  - unitary operator, 175, 187, 189
- strongly continuous one parameter
  - group of unitary operators, 221
- Sturm-Liouville theory, 110
- theorem
  - Ascoli-Arzelà, 107
  - Baire, 70
  - Banach-Alaoglu: sequential form, 69
  - Banach-Alaoglu: topological form, 68
  - Banach-Steinhaus, or uniform boundedness principle, 71
  - Bessel inequality, 45
  - Bochner, 33
  - closed graph, 78
  - Fredholm, 96
  - Fredholm alternative, 98
  - Hellinger-Toeplitz, 83
  - Hilbert-Schmidt, 99
  - inverse mapping, 78
  - Lax-Milgram, 83
  - Lebesgue-Hausdorff, 9, 180
  - of Herglotz for harmonic functions, 31

- of Herglotz for holomorphic functions, 32
- open mapping, 78
- orthogonal decomposition, 42
- predual of separable is separable, 63
- Riesz-Markov-Kakutani for complex measures, 8
- Riesz-Markov-Kakutani for positive measures, 7, 178
- Riesz-Schauder, 98
- Riesz representation in Hilbert spaces, 44
- spectral in multiplicative form
  - for self-adjoint operators, 218
- spectral in multiplicative form
  - for unitary operators, 185
- Stone, 174, 223
- Tychonoff, 67
- trigonometric system, 14
  - completeness, 22, 50
- weak topology, 94
  - and compact operators, 94
- Zorn lemma, 53





## Bibliography

- [1] N. Aronszajn, [Theory of reproducing kernels](#), Trans. AMS (1950)
- [2] Eric Bédos, [An introduction to Sturm-Liouville theory](#), lecture notes
- [3] Arne Beurling, [Ensembles exceptionnels](#), Acta Math. 72 (1940), 1–13.
- [4] Garrett Birkhoff and Erwin Kreyszig [The Establishment of Functional Analysis](#) (Historia Mathematica 11(3) 1984, p. 258-321)
- [5] Guy Bonneau, Jacques Faraut, Galliano Valent, [Self-adjoint extensions of operators and the teaching of quantum mechanics](#) (2001), [arXiv:quant-ph/0103153](#)
- [6] Gerald B. Folland, [Real Analysis: Modern Techniques and Their Applications](#), Wiley 2nd Edition (1999)
- [7] Markus Haase, [The Functional Calculus Approach to the Spectral Theorem](#) (2020) [arXiv:2003.06130](#).
- [8] Paul R. Halmos [Introduction to Hilbert Space and the Theory of Spectral Multiplicity](#), New York, Chelsea Pub. Co. 1957
- [9] Alexander Kegeles [Lecture notes on operator algebras and their application in physics](#) (2022) [arXiv:2208.10151](#)
- [10] Paul Koosis, [Introduction to  \$H^p\$  Spaces](#), 2<sup>nd</sup> edition (1999) Cambridge University Press.
- [11] Emmanuel Kowalski [Spectral theory in Hilbert spaces](#) (2009), ETH Zürich lecture notes,
- [12] Peter D. Lax, Functional analysis, Wiley-Interscience, 2002
- [13] É. Matheron [Three Proofs of Tychonoff's Theorem](#) (2020), lecture notes
- [14] Vern I. Paulsen, Mrinal Raghupathi, [An Introduction to the Theory of Reproducing Kernel Hilbert Spaces](#) (2016) Cambridge University Press
- [15] Michael Reed, Barry Simon, [Methods of Modern Mathematical Physics: Functional analysis](#), Academic Press, 1972
- [16] M. Reed, B. Simon [Methods of modern mathematical physics: II Fourier Analysis, Self-Adjointness](#) (1975), Academic Press
- [17] Walter Rudin [Fourier Analysis on Groups](#) (1962), Wiley-Interscience
- [18] S.M. Srivastava [A Course on Borel Sets](#), Springer 1998.
- [19] S.M. Srivastava [How did Cantor Discover Set Theory and Topology?](#) [Resonance, Vol. 19, p. 977-999, (2014)]

- [20] Leonard Susskind, Art Friedman, Quantum Mechanics: The Theoretical Minimum, Penguin UK
- [21] Leonard Susskind, [The theoretical minimum: quantum mechanics](#), video lectures
- [22] Terence Tao, [The spectral theorem and its converses for unbounded symmetric operators \(2011\)](#), from Tao's blog.
- [23] Michael Taylor, [The Spectral Theorem for Self-Adjoint and Unitary Operators \(2018\)](#), lecture notes.
- [24] Michael Taylor, [Spectral Theory](#), lecture notes
- [25] Gerald Teschl [Mathematical Methods in Quantum Mechanics With Applications to Schrödinger Operators \(2014\)](#) (Graduate Studies in Mathematics, Volume 157, Amer. Math. Soc.).
- [26] [Wiener algebra](#), Wikipedia
- [27] Wenjing Wu, [The Bochner integral](#), lecture notes