Universiteit Utrecht

Faculteit Bètawetenschappen

The spectral theorem for unbounded self-adjoint operators and Nelson's theorem

BACHELOR THESIS

Kevin Zwart
(5533082)

Wis- en Natuurkunde (TWIN)

Supervisor:

Prof. Dr. E.P. van den Ban
Mathematical Institute

June 18, 2018

Abstract

In this thesis, we will introduce the notion of unbounded operators on a Hilbert space. We will discuss the definition of the adjoint of an operator, and what it means for an operator to be self-adjoint. After that, we will restrict ourselves to bounded operators and prove the Spectral Theorem for normal bounded operators. The notion of a spectral measure will be introduced as well. After that, we return to unbounded operators and we will consider the Cayley-transform. With that and the Spectral Theorem for normal bounded operators, we prove the Spectral Theorem for unbounded self-adjoint operators. Then we look into the situations that two unbounded self-adjoint operators commute on a common domain. We then prove a theorem of Nelson that gives some criteria which imply that the spectral measures of the two operators commute. And finally we will consider two examples of unbounded operators that play a role in Quantum Mechanics.

Contents

1 Introduction 1

2 Unbounded operators 2
2.1 Unbounded operators and some basic properties 2
2.2 The adjoint of an operator 5
2.3 Inverse of an operator and the spectrum 8
2.4 Symmetric, normal and self-adjoint operators 9

3 Spectral theorem for bounded normal operators 13
3.1 C^*-algebras and representations 13
3.2 Spectral measures 15
3.3 Spectral measure and representation of bounded measurable functions 21
3.4 The spectral theorem for bounded normal operators 26

4 Spectral Theorem for unbounded self-adjoint operators 31
4.1 The Cayley transform 31
4.2 The spectral theorem for unbounded self-adjoint operators 33

5 Theorem of Nelson 40
5.1 The theorem of Nelson 40

6 Position and momentum operators 48
6.1 Tempered distributions 49
6.2 Fourier transform 50
6.3 Position and momentum operators 53

References I

1 Introduction

A theorem often discussed in an introduction course in Linear Algebra, is that a symmetric matrix can be diagonalized with the eigenvalues on the diagonal. This theorem is often called the Spectral Theorem. In an introduction course in Functional Analysis, the same theorem is proven for a compact, normal operator on a Hilbert space. In both these cases, both operators were bounded and the spectrum of the operators are countable. So if we were to allow the operator to be unbounded on a Hilbert space, and the spectrum be uncountable, does there still exist a spectral theorem for certain operators?

The Spectral Theorem for unbounded self-adjoint operators answers that question with yes. If A is a (possibly unbounded) operator and A is self-adjoint, then A can be written as a 'sum' over the elements of the spectrum times a projection operator. In fact, this 'sum' will be an integral with respect to some projection valued measure. This also involves the notion of a spectral measure.

As we will see, spectral measures have interesting properties. If we can show that two spectral measures commute, then any operator written as an integral over the first spectral measure commutes with any operator that can be written as an integral over the second spectral measure. So we wish to investigate the commutation of two spectral measures. The theorem of Nelson will give us a way to characterize some situations in which two spectral measures, corresponding to two self-adjoint operators, commute.

We will begin by studying unbounded operators on a Hilbert space, and adjoints of these operators. In particular we will look into the notion of self-adjointness for unbounded operators. Then we will study spectral measures and the Spectral Theorem for normal bounded operators. Then we will prove the Spectral Theorem for unbounded, self-adjoint operators. Next we will look at the theorem of Nelson for self-adjoint operators, and finally we will consider an application in Quantum Mechanics.

We assume the reader has at least a Bachelor level of understanding of Linear Algebra, Topology, real Analysis, Measure Theory and Functional Analysis. We will also use some Distribution Theory, but that will be explained in short.

Most of the definitions and results are based on the results in the book of John B. Conway ([1]). Most of the time we will follow the proofs given in the book. However, sometimes we will diverge from the results given in the book.

Finally, I would like to thank my supervisor prof. dr. E. P. van den Ban for guiding me in writing this thesis. Even though he is very busy, he always had a moment each week to discuss the progress and helped me out when it was needed. It was a great experience! I also would like to thank my family, girlfriend and friends for supporting me when it was necessary.

2 Unbounded operators

In this thesis, we wish to investigate unbounded operators on a Hilbert space, and their spectral resolutions in particular. The theorem of Nelson tells us that if we start with two self-adjoint operators A and B, and $A^2 + B^2$ is essentially self-adjoint, then the spectral resolutions of A and B commute. In order to discuss this theorem, we must first investigate unbounded operators and self-adjointness. Unbounded linear operators are no longer continuous, so some theorems based on continuity do not hold anymore. Although, as we continue down the line, a lot of the properties of bounded operators still hold for arbitrary operators.

In this section we will prove some elementary properties of (unbounded) operators. We will discuss the domain and closures of operators, symmetric, normal and self-adjoint operators and essentially self-adjointness. The reader who is already familiar with these concepts, may read onto the next section, Section 3.

2.1 Unbounded operators and some basic properties

Before we are ready to define what an unbounded operator is, we need to broaden our definition of a linear operator. In this way, we do not have to speak of unbounded operators and if we want to address them, we can do so without having the problem of a not well-defined operator on all of $\mathscr{H}$. For the rest of this thesis, all Hilbert spaces are assumed to be defined over $\mathbb{C}$. So the inner product is a sesquilinear mapping into $\mathbb{C}$. Also, we assume that every Hilbert space is separable.

Definition 2.1.1. If $\mathscr{H}, \mathscr{K}$ are Hilbert spaces, we define a *linear operator* $A : \mathscr{H} \to \mathscr{K}$ as a function whose domain of definition is a linear subspace (not necessarily closed), $\mathrm{Dom}(A)$, in $\mathscr{H}$ and such that for any $\lambda, \mu \in \mathbb{C}$ and $x, y \in \mathrm{Dom}(A)$ we find $A(\lambda x + \mu y) = \lambda A(x) + \mu A(y)$. We call A a *bounded operator* if there exists a $c > 0$ such that $\|Ax\| \leq c\|x\|$ for all $x \in \mathrm{Dom}(A)$. We write $\|A\| := \sup\{\|Ax\| | \, x \in \mathrm{Dom}(A), \|x\| \leq 1\}$

Remark 2.1.2. Before we continue, we want to note three things. First of all, whenever we say $A : \mathscr{H} \to \mathscr{K}$ is a linear operator, we mean that $A : \mathrm{Dom}(A) \to \mathscr{K}$ is a linear operator with $\mathrm{Dom}(A) \subseteq \mathscr{H}$. It might suggest that A is everywhere defined, but note the implicit assumption of a domain which might not be the whole Hilbert space.

Secondly, we note that the notion of boundedness in the new definition of a linear operator, is equivalent to the notion of boundedness for linear operators everywhere defined. Of course, if A is bounded in the old definition, then A is defined on all of $\mathscr{H}$ and $\|A\| = \sup\{\|Ax\| | \, \|x\| \leq 1\} = c$. Then it also is a bounded linear operator in the new definition, because $\mathrm{Dom}(A) = \mathscr{H}$. On the other hand, if A is a bounded operator defined on $\mathrm{Dom}(A)$, then we can find a bounded operator $\tilde{A}$ such that $\|\tilde{A}x\| \leq c\|x\|$ for all $x \in \mathscr{H}$.

Lemma 2.1.3. *Let* $A : \mathscr{H} \to \mathscr{K}$ *be bounded. Then there exists a unique bounded operator* A' *with* $\mathrm{Dom}(A') = \overline{\mathrm{Dom}(A)}$ *such that* $A'x = Ax$ *if* $x \in \mathrm{Dom}(A)$

Proof. Let $x \in \overline{\mathrm{Dom}(A)}$. Then there exists a sequence $(x_n)_n \subseteq \mathrm{Dom}(A)$ such that $x_n \to x$ as $n \to \infty$. So $(x_n)_n$ is also a Cauchy sequence. Because $x_n \in \mathrm{Dom}(A)$, we have $\|Ax_n\| \leq c\|x_n\|$ for any n. Therefore $(Ax_n)_n$ is also a Cauchy sequence. Because a Hilbert space is complete, $\lim_{n\to\infty} Ax_n$ exists. Define

$$A'x := \lim_{n\to\infty} Ax_n$$

We will show that this definition of $A'x$ does not depend on the choice of the sequence $(x_n)_n$. For let $(y_n)_n \subseteq \mathrm{Dom}(A)$ be another sequence converging to x. Then $\|Ax_n - Ay_n\| \leq c\|x_n - y_n\| \to 0$ for both sequences converge to x. Therefore $\|Ax_n - Ay_n\| \to 0$ and so they converge to the same value. Therefore is $A'x$ well-defined. We also see that $\mathrm{Dom}(A') = \overline{\mathrm{Dom}(A)}$ and $A' : \mathscr{H} \to \mathscr{K}$ is a linear operator. Finally, we see that $\|A'\| = \|A\|$ □

Unless otherwise specified we will always identify a bounded operator $A : \mathscr{H} \to \mathscr{K}$ with domain $\mathrm{Dom}(A)$ with the unique bounded extension $\tilde{A} : \mathscr{H} \to \mathscr{K}$ such that $\mathrm{Dom}(\tilde{A}) = \mathscr{H}$ and $\tilde{A} = 0$ on $\mathrm{Dom}(A')^{\perp}$, where A' is as in previous Lemma. Note that $\|\tilde{A}\| = \|A\|$.

On the other hand, if A is an unbounded operator in the old definition, we assumed A to be defined everywhere and $\sup\{\|Ax\| \mid \|x\| \leq 1\}$ tends to go to infinity. In other words, we can find $x \in \mathscr{H}$ such that $\|Ax\|$ is arbitrary large. Therefore there must exist $x \in \mathscr{H}$ such that Ax cannot be well-defined. So the notion of a domain is needed in order to work with unbounded operators.

Example 2.1.4. Consider the Hilbert space $L^2(\mathbb{R})$ with the inner product $\langle f, g\rangle = \int_{\mathbb{R}} f(x)\overline{g(x)}dx$. Then consider the linear mapping $\hat{x} : f \mapsto xf$. It is not guaranteed that for every $f \in L^2(\mathbb{R})$ we have $xf \in L^2$. In fact, define the function $f : \mathbb{R} \to \mathbb{R}$ as

$$f(x) = \begin{cases} \frac{1}{x^{3/2}} & x \geq 1 \\ 0 & else \end{cases}$$

Then we find that

$$\int_{\mathbb{R}} |f(x)|^2 dx = \int_1^{\infty} \frac{1}{x^3} dx < \infty$$

So $f \in L^2(\mathbb{R})$. Therefore, one might expect that $\hat{x}(f) \in L^2(\mathbb{R})$. However, we have

$$\int_{\mathbb{R}} |xf(x)|^2 dx = \int_1^{\infty} \frac{1}{x} dx = \infty$$

Thus $xf \notin L^2(\mathbb{R})$. Therefore we have that the operator $\hat{x}$ cannot be defined on the whole space $L^2(\mathbb{R})$. We need to specify a suitable domain for $\hat{x}$. Define $\mathrm{Dom}(\hat{x}) = \{f \in L^2(\mathbb{R}) | xf \in L^2(\mathbb{R})\}$. This operator we defined is called the *position operator*, and plays an important role in Quantum Mechanics. We will look at this specific operator in Section 6 ⊘

Definition 2.1.5. If $\mathscr{H}, \mathscr{K}$ are normed vector spaces, we denote by $\mathscr{B}(\mathscr{H}, \mathscr{K})$ the set of bounded operators from $\mathscr{H}$ into $\mathscr{K}$. We define $\mathscr{B}(\mathscr{H})$ to be the set of bounded operators on $\mathscr{H}$, and we write $\mathscr{H}'$ for the set of bounded linear functionals $A : \mathscr{H} \to \mathbb{C}$. Note that $\mathscr{B}(\mathscr{H}, \mathscr{K})$ is a linear space.

Definition 2.1.6. We say that A is *densely defined*, if $\mathrm{Dom}(A)$ lies dense in $\mathscr{H}$.

Note that if A is a linear operator from $\mathscr{H}$ to $\mathscr{K}$, then A is also a linear operator from the Hilbert space $\overline{\mathrm{Dom}(A)}$ to $\mathscr{K}$. Thus if we replace the Hilbert space $\mathscr{H}$ with $\overline{\mathrm{Dom}(A)}$, then we can arrange that A is densely defined. If A is densely defined with respect to $\mathscr{H}$, then we can approximate any $x \in \mathscr{H}$ by an element in $\mathrm{Dom}(A)$. This fact will be key to some of the definitions later.

Remark 2.1.7. With these definitions it is easy to see that, if A, B are linear operators from $\mathscr{H}$ into $\mathscr{K}$, then $A + B$ is defined on $\mathrm{Dom}(A + B) = \mathrm{Dom}(A) \cap \mathrm{Dom}(B)$. Additionally, if $\mathscr{H}, \mathscr{K}, \mathscr{L}$ are Hilbert spaces and $A : \mathscr{H} \to \mathscr{K}$, $B : \mathscr{K} \to \mathscr{L}$ are linear operators, then $\mathrm{Dom}(BA) = A^{-1}(\mathrm{Dom}(B))$.

Definition 2.1.8. If $\mathscr{H}, \mathscr{K}$ are Hilbert spaces and A, B are linear operators from $\mathscr{H}$ into $\mathscr{K}$, we say B is an *extension* of A if

1. $\mathrm{Dom}(A) \subseteq \mathrm{Dom}(B)$,
2. if $x \in \mathrm{Dom}(A)$ then $Ax = Bx$.

We write $A \subseteq B$ if B is an extension of A.

We note that if $A \in \mathscr{B}(\mathscr{H})$, then $\mathrm{Dom}(A) = \mathscr{H}$, and so the only extension of A is A itself. Therefore the notion of extensions is only relevant for unbounded operators.

We remind ourselves of the definition of a graph. Now that the domain is not necessarily the whole Hilbert space anymore, we need to refine the definition.

Definition 2.1.9. If $A : \mathscr{H} \to \mathscr{K}$ is a linear operator, the *graph* of A is defined as

$$\mathrm{gra}(A) := \{(x, Ax) \in \mathscr{H} \times \mathscr{K} | x \in \mathrm{Dom}(A)\}$$

Lemma 2.1.10. *Let $A : \mathscr{H} \to \mathscr{K}$ be a linear operator. Define the mapping $\|\cdot\|_{gr} : \mathrm{Dom}(A) \to \mathbb{R}$ as*

$$\|x\|_{gr} = \sqrt{\|x\|_{\mathscr{H}}^2 + \|Ax\|_{\mathscr{K}}^2}$$

This mapping is a norm on $\mathrm{Dom}(A)$. This norm is called the graph norm

Together with the previous definition, it is easy to see that $B \subseteq A$ if and only if $\mathrm{gra}(B) \subseteq \mathrm{gra}(A)$.

One of the important theorems about graphs, is the Closed Graph Theorem. For completeness, we will give the theorem here in the context of a Hilbert space and for a general linear mapping. A proof of this theorem can be found in any Functional Analysis book, for example $[7, p.123]$.

Theorem 2.1.11. *(Closed Graph Theorem) Let $\mathscr{H}, \mathscr{K}$ be Hilbert spaces, and T a linear operator from $\mathscr{H}$ into $\mathscr{K}$ with $\mathrm{Dom}(T) = \mathscr{H}$. If $\mathrm{gra}(T)$ is closed, then $T \in \mathscr{B}(\mathscr{H}, \mathscr{K})$*

Definition 2.1.12. An operator $A : \mathscr{H} \to \mathscr{K}$ is called *closed* if its graph $\mathrm{gra}(A)$ is closed in $\mathscr{H} \times \mathscr{K}$. We call A *closable* if there exists a closed extension of A. We denote by $\mathscr{C}(\mathscr{H}, \mathscr{K})$ the set of all closed, densely defined operators from $\mathscr{H}$ into $\mathscr{K}$. We denote $\mathscr{C}(\mathscr{H}) = \mathscr{C}(\mathscr{H}, \mathscr{H})$.

Lemma 2.1.13. *[1, Prop X.1.4, p.304] Let $\mathscr{H}, \mathscr{K}$ be Hilbert spaces. A linear operator $A : \mathscr{H} \to \mathscr{K}$ is closable if and only if $\overline{\mathrm{gra}(A)}$ is the graph of a linear operator.*

Proof. Let $\overline{\mathrm{gra}(A)}$ be the graph of a linear operator. Then by definition there exists a linear operator $B : \mathscr{H} \to \mathscr{K}$ with $\mathrm{gra}(B) = \overline{\mathrm{gra}(A)}$. Since $\mathrm{gra}(A) \subseteq \overline{\mathrm{gra}(A)}$, we have that B is an extension of A, and thus A is closable.

Now, let A be a closable operator. In other words, A has a closed extension $B : \mathscr{H} \to \mathscr{K}$. Let $(0, x) \in \overline{\text{gra}(A)}$. Because $\text{gra}(A) \subseteq \text{gra}(B)$ and $\text{gra}(B)$ is closed, we have $(0, x) \in \text{gra}(B)$ and thus $x = B(0) = 0$. Define

$$\mathscr{D} = \{y \in \mathscr{H} | \exists z \in \mathscr{K} : (y, z) \in \overline{\text{gra}(A)}\}$$

If $x \in \mathscr{D}$ and $y_1, y_2 \in \mathscr{K}$ such that $(x, y_2), (x, y_2) \in \overline{\text{gra}(A)}$ then $(0, y_1 - y_2) \in \overline{\text{gra}(A)}$. Thus by the same argument, $y_1 - y_2 = B(0) = 0$ and so $y_1 = y_2$. So we have for every $x \in \mathscr{D}$ a unique $y \in \mathscr{K}$ such that $(x, y) \in \overline{\text{gra}(A)}$. Define $T : \mathscr{H} \to \mathscr{K}$ with $\text{Dom}(T) = \mathscr{D}$ as $Tx = y$ where y is such that $(x, y) \in \overline{\text{gra}(A)}$. We only need to show that this operator is a linear operator and $\text{gra}(T) = \overline{\text{gra}(A)}$. It is easy to check that T is a linear operator, by using the fact that $\overline{\text{gra}(A)}$ is a linear subspace. By construction we find $\text{gra}(T) \subseteq \overline{\text{gra}(A)}$. On the other hand, if $(x, y) \in \overline{\text{gra}(A)}$ then $Tx = y$ by definition, and so $(x, y) = (x, Tx) \in \text{gra}(T)$. So $\overline{\text{gra}(A)} = \text{gra}(T)$, and so $\overline{\text{gra}(A)}$ is the graph of a linear operator. □

Definition 2.1.14. Let A be a closable operator as in Prop. 2.1.13. The operator whose graph is $\overline{\text{gra}(A)}$ is called the *closure of A*. It is denoted by $\overline{A}$.

2.2 The adjoint of an operator

For bounded operators, we know that the adjoint of an operator is defined as the unique operator such that $\langle Ax, y\rangle = \langle x, A^*y\rangle$. For arbitrary linear operators, we cannot use this definition anymore. It might be that Ax is not defined, or the mapping $x \mapsto \langle Ax, y\rangle$ is not bounded for certain $y \in \mathscr{H}$. The latter was needed in the proof of uniqueness of the vector A^*y, so if it is not bounded anymore, uniqueness might not occur. Therefore, another definition is needed.

Definition 2.2.1. Let $\mathscr{H}, \mathscr{K}$ be Hilbert spaces. If $A : \mathscr{H} \to \mathscr{K}$ is densely defined, define the set:

$$\text{Dom}(A^*) = \{y \in \mathscr{K} | x \mapsto \langle Ax, y\rangle_{\mathscr{K}} \text{is a bounded linear functional on } \text{Dom}(A)\}$$

By $\langle \cdot, \cdot\rangle_{\mathscr{K}}$ we mean the inner product defined on $\mathscr{K}$.

Remark 2.2.2. In order to introduce the adjoint of an operator, we consider $y \in \text{Dom}(A^*)$. Then the mapping $f : x \mapsto \langle Ax, y\rangle_{\mathscr{K}}$ is a bounded linear functional on $\text{Dom}(A)$. Because $\text{Dom}(A)$ lies dense in $\mathscr{H}$, by Remark 2.1.2 it has a unique extension $\tilde{f}$ that is defined on all of $\mathscr{H}$. Because it is a bounded linear functional, we can use the Riesz Representation Theorem to conclude that there exists a unique $z \in \mathscr{H}$ such that $f(x) = \langle x, z\rangle_{\mathscr{H}}$ for every $x \in \mathscr{H}$. Thus we find the equation for $x \in \text{Dom}(A)$

$$\langle Ax, y\rangle_{\mathscr{K}} = \langle x, z\rangle_{\mathscr{H}} \tag{1}$$

Definition 2.2.3. Let A be a linear operator. For $y \in \text{Dom}(A^*)$ we define A^*y to be the unique element $z \in \mathscr{H}$ determined by Remark 2.2.2. By this definition, we find a linear operator $A^* : \mathscr{K} \to \mathscr{H}$ with domain $\text{Dom}(A^*)$, where A^*y is defined in such a way that

$$\langle Ax, y\rangle = \langle x, A^*y\rangle \qquad (x \in \text{Dom}(A)) \tag{2}$$

Remark 2.2.4. We observe that this definition is an implicit definition of A^*. It is in general very hard to find an explicit formula for the domain, or for A^* itself. In some cases, properties of A will make an explicit formula possible (self-adjointness for example), but for most cases calculating A^* is hard to almost impossible.

Proposition 2.2.5. *[1, Prop. X.1.6, p.305] Let $\mathscr{H}, \mathscr{K}$ be Hilbert spaces, and $A : \mathscr{H} \to \mathscr{K}$ be a densely defined operator. Then:*

1. *A^* is a closed linear operator.*
2. *A^* is densely defined if and only if A is closable.*
3. *if A is closable, then $\overline{A} = (A^*)^* := A^{**}$*

Before we prove this lemma, we introduce another lemma to aid us in the proof of the above lemma.

Lemma 2.2.6. *If $A : \mathscr{H} \to \mathscr{K}$ be a densely defined operator, and $J : \mathscr{H} \times \mathscr{K} \to \mathscr{K} \times \mathscr{H}$ is defined as $J(x, y) = (-y, x)$, then J is an isometric isomorphism and*

$$\mathrm{gra}(A^*) = [J(\mathrm{gra}(A))]^\perp \tag{3}$$

Proof. It should be clear that J is an isometric isomorphism. Thus the only real interesting part of the lemma is the second part. Remember that the inner product $\langle \cdot, \cdot \rangle_{\mathscr{H} \times \mathscr{K}}$ on $\mathscr{H} \times \mathscr{K}$ is given by $\langle (x, X), (y, Y) \rangle_{\mathscr{H} \times \mathscr{K}} = \langle x, y \rangle_{\mathscr{H}} + \langle X, Y \rangle_{\mathscr{K}}$. If $x \in \mathrm{Dom}(A)$ and $x' \in \mathrm{Dom}(A^*)$, we have:

$$\langle (J(x, Ax)), (x', A^*x') \rangle_{\mathscr{H} \times \mathscr{K}} = \langle (-Ax, x), (x', A^*x') \rangle_{\mathscr{H} \times \mathscr{K}} = -\langle Ax, x' \rangle_{\mathscr{H}} + \langle x, A^*x' \rangle_{\mathscr{K}} = 0$$

by Equation (1). Therefore $\mathrm{gra}(A^*) \subseteq [J(\mathrm{Dom}(A))]^\perp$. On the other hand, if $(x, y) \in [J(\mathrm{gra}(A))]^\perp$, then for any $z \in \mathrm{Dom}(A)$ we have $0 = \langle (x, y), J(z, Az) \rangle_{\mathscr{H} \times \mathscr{K}} = -\langle x, Az \rangle_{\mathscr{H}} + \langle y, z \rangle_{\mathscr{K}}$. So $\langle Az, x \rangle = \langle z, y \rangle$. So by definition we find $x \in \mathrm{Dom}(A^*)$ and $A^*x = y$ □

Proof of Proposition 2.2.5. 1) Since $\mathrm{gra}(A^*) = [J(\mathrm{Dom}(A))]^\perp$, we have $\mathrm{gra}(A^*)$ is a closed set. Therefore A^* is a closed operator.

2) First assume $\mathrm{Dom}(A^*)$ is dense in $\mathscr{K}$. Then $(A^*)^* = A^{**}$ is defined. Then by 1), A^{**} is a closed operator. We need to show $A \subseteq A^{**}$. Let $x \in \mathrm{Dom}(A)$. Define $f : \mathrm{Dom}(A^*) \to \mathbb{C}$ by $f(y) = \langle A^*y, x \rangle$. We see that $|f(y)| = |\langle A^*y, x \rangle| = |\langle y, Ax \rangle| \leq \|Ax\|\|y\|$ by Remark 2.2.2. Therefore we see that f is a bounded operator on $\mathrm{Dom}(A^*)$, and thus $x \in \mathrm{Dom}(A^{**})$. Additionally we see that for any $x \in \mathrm{Dom}(A), y \in \mathrm{Dom}(A^*)$:

$$\langle Ax, y \rangle_{\mathscr{K}} = \langle x, A^*y \rangle_{\mathscr{H}} = \langle A^{**}x, y \rangle_{\mathscr{K}}$$

So $\langle (A - A^{**})x, y \rangle = 0$. This is true for any $y \in \mathrm{Dom}(A^*)$. Since $\mathrm{Dom}(A^*)$ lies dense in $\mathscr{K}$, we must have $Ax = A^{**}x$ for any $x \in \mathrm{Dom}(A)$. We conclude that $A \subseteq A^{**}$ and so A is closable.

On the other hand, let A be a closable operator. Consider $x \in [\mathrm{Dom}(A^*)]^\perp$. We wish to show $x = 0$, because this shows that $[\mathrm{Dom}(A^*)]^\perp = (0)$, and so $\mathrm{Dom}(A^*)$ lies dense in $\mathscr{K}$. Since $x \in (\mathrm{Dom}(A^*))^\perp$ we have that

$$(x, 0) \in (\mathrm{gra}(A^*))^\perp = \left[[J(\mathrm{gra}(A))]^\perp\right]^\perp = \overline{J(\mathrm{gra}(A))}$$

Because J is is an isomorphism on $\mathscr{H} \times \mathscr{K}$, we have $\overline{J(\mathrm{gra}(A))} = J(\overline{\mathrm{gra}(A)})$. Therefore $(x, 0) \in J(\overline{\mathrm{gra}(A)})$. So there exists a $(y, z) \in \overline{\mathrm{gra}(A)}$ such that $J(y, z) = (-z, y) = (x, 0)$. So $-z = x$ and $y = 0$. Thus $(0, -x) \in \overline{\mathrm{gra}(A)}$. But we know that A is closable, so $\overline{\mathrm{gra}(A)}$ is a graph. So $x = 0$. We conclude that A^* is densely defined.

3) Let A be closable. We know already by 2) that A^{**} is a closed extension. We only need to prove that $\mathrm{gra}(A^{**}) = \overline{\mathrm{gra}(A)}$. Define $J' : \mathscr{K} \times \mathscr{H} \to \mathscr{H} \times \mathscr{K}$ by $J'(x, y) = (-y, x)$. Note that we only switched $\mathscr{H}$ and $\mathscr{K}$. By 2), $A^* : \mathscr{K} \to \mathscr{H}$ is densely defined and so by going through the same proof as in Lemma 2.2.6 we find $\mathrm{gra}(A^{**}) = [J'(\mathrm{gra}(A^*))]^\perp$. But we note that $J' \circ J(x, y) = J'(-y, x) = -(x, y)$ for any $x \in \mathscr{H}$ and $y \in \mathscr{K}$. Therefore, $J' \circ J = -I$. By the same calculations, $J \circ J' = -I$. Hence $J' = -J^{-1}$. We know that J is an isometric isomorphism, thus $J^{-1} = J^*$. Therefore we conclude

$$J'(\mathrm{gra} A^*) = J^{-1}(\mathrm{gra}(A)) = -J^*(\mathrm{gra}(A)) = J^*(\mathrm{gra}(A))$$

Hence we can conclude $\mathrm{gra}(A^{**}) = [J^*(\mathrm{gra}(A^*))]^\perp$. Therefore

$$\mathrm{gra}(A^{**}) = [J^*(\mathrm{gra}(A^*))]^\perp = \left[J^*\left(J(\mathrm{gra}(A))^\perp\right)\right]^\perp$$

Because J is an isometric isomorphism on $\mathscr{H} \times \mathscr{K}$, $J^* = J^{-1}$ and J^{-1} is continuous. Hence

$$\mathrm{gra}(A^{**}) = \left[J^{-1}\left(J(\mathrm{gra}(A))^\perp\right)\right]^\perp = \left[J^{-1} \circ J(\mathrm{gra}(A)^\perp)\right]^\perp = \mathrm{gra}(A)^{\perp\perp} = \overline{\mathrm{gra}(A)}$$

So $A^{**} = \overline{A}$. □

Corollary 2.2.7. *[1, Cor. X.1.8, p.305] Let $\mathscr{H}, \mathscr{K}$ be Hilbert spaces, and $A \in \mathscr{C}(\mathscr{H}, \mathscr{K})$. Then $A^* \in \mathscr{C}(\mathscr{K}, \mathscr{H})$ and $A = A^{**}$. Here is $\mathscr{C}(\mathscr{H}, \mathscr{K})$ defined as in Definition 2.1.12.*

This corollary looks a lot like the case of bounded operators. Of course, if A is bounded, then we can define it on all of $\mathscr{H}$, and then we can use the previous corollary to conclude that $A^{**} = A$, which is known for bounded operators.

Corollary 2.2.8. *[1, Prop. X.1.13, p.307] If $A : \mathscr{H} \to \mathscr{K}$ is a densely defined linear operator, then*

$$(\mathrm{ran}(A))^\perp = \ker(A^*) \tag{4}$$

If A is also closed, then

$$(\mathrm{ran}(A^*))^\perp = \ker(A) \tag{5}$$

Proof. 1) if $x \in (\mathrm{ran}(A))^\perp$, then for any $y \in \mathrm{Dom}(A)$ we have $0 = \langle Ay, x\rangle$. So $x \in \mathrm{Dom}(A^*)$ and $A^*x = 0$. The other inclusion is clear.

2) if A is closed, $A \in \mathscr{C}(\mathscr{H}, \mathscr{K})$. Thus by Corollary 2.2.7 we have $A^{**} = A$, and we use Equation (4) to conclude Equation (5). □

2.3 Inverse of an operator and the spectrum

In order to define an inverse of a linear operator, we remember Remark 2.1.7. If $A : \mathscr{H} \to \mathscr{K}$ is a linear operator, and $B : \mathscr{K} \to \mathscr{H}$ is a bounded linear operator, we see that $\mathrm{Dom}(AB) = B^{-1}(\mathrm{Dom}(A))$ and $\mathrm{Dom}(BA) = \mathrm{Dom}(A)$ because B is defined on all of $\mathscr{K}$.

Definition 2.3.1. Let $\mathscr{H}, \mathscr{K}$ be Hilbert spaces, and $A : \mathscr{H} \to \mathscr{K}$ be a linear operator with domain $\mathrm{Dom}(A)$. We say A is *boundedly invertible* if there exists a **bounded** linear operator $B : \mathscr{K} \to \mathscr{H}$ such that $AB = I_{\mathscr{K}}$ and $BA \subseteq I_{\mathscr{H}}$. We call B a *(bounded) inverse* of A. In light of the following proposition part 2, we denote $B = A^{-1}$.

Note that if A is boundedly invertible, then $BA \subseteq I_{\mathscr{K}}$, and therefore is BA bounded on its domain. Therefore it is possible to extend is, as noted in Remark 2.1.2.

For any bounded operator it was enough to be bijective, in order to have a bounded inverse (by the Open Mapping Theorem). For any arbitrary operator, we do not have the advantage of having $\mathscr{H}$ as a domain. However, the following proposition ensures that we can find a bounded inverse of A if $A : \mathrm{Dom}(A) \to \mathscr{K}$ is bijective, and the graph of A is closed.

Proposition 2.3.2. *[1, Prop. X.1.14, p. 307] Let $\mathscr{H}, \mathscr{K}$ be Hilbert spaces, and $A : \mathscr{H} \to \mathscr{K}$ be a linear operator. Then*

1. *A is bounded invertible if and only if $\ker(A) = (0)$, $\mathrm{ran}(A) = \mathscr{K}$ and the graph of A is closed.*

2. *If A is boundedly invertible, its inverse is uniquely defined.*

Proof. 1) First let A be boundedly invertible. Let $B \in \mathscr{B}(\mathscr{K}, \mathscr{H})$ be an inverse of A. Then $\mathrm{Dom}(B) = \mathscr{K}$. Let $x \in \ker(A)$. Since $BA \subseteq I_{\mathscr{H}}$, we have $0 = B(0) = BA(x) = I_{\mathscr{H}}(x) = x$. So $\ker(A) = (0)$. Additionally, if $y \in \mathscr{K}$, we see that $y = I_{\mathscr{K}}(y) = AB(y) = A(B(y))$. So $y \in \mathrm{ran}(A)$. So $\mathrm{ran}(A) = \mathscr{K}$. Also, note

$$\mathrm{gra}(A) = \{(x, Ax) \in \mathscr{H} \times \mathscr{K} | x \in \mathrm{Dom}(A)\} = \{(Bx, x) \in \mathscr{H} \times \mathscr{K} | x \in \mathscr{K}\}$$

Because B is a bounded operator, $\mathrm{gra}(A)$ is closed.

Now let $\ker(A) = (0)$, $\mathrm{ran}(A) = \mathscr{K}$ and assume $\mathrm{gra}(A)$ is closed. Because of the first two properties, A is a bijective operator on its domain. Therefore $Bx := A^{-1}x$ is well defined for any $x \in \mathscr{K}$. Because $\mathrm{gra}(A)$ is closed, the same holds for $\mathrm{gra}(B)$. By The Closed Graph Theorem, Theorem 2.1.11 we find $B \in \mathscr{B}(\mathscr{K}, \mathscr{H})$.

2). Let A be boundedly invertible, and assume B_1 and B_2 are bounded inverses of A. Thus $AB_1 = AB_2 = I_{\mathscr{K}}$ and so $A(B_1 - B_2)x = 0$ for any $x \in \mathscr{K}$. So $A(B_1x - B_2x) = 0$, and so $B_1x - B_2x \in \ker(A)$. But because A is boundedly invertible, $\ker(A) = (0)$. So $B_1x - B_2x = 0$ and so $B_1x = B_2x$ for any $x \in \mathscr{K}$. Therefore $B_1 = B_2$. □

Definition 2.3.3. Let $\mathscr{H}$ be a Hilbert space, and $A : \mathscr{H} \to \mathscr{H}$ be a linear operator. Define the *resolvent set* $\rho(A)$ by $\rho(A) := \{\lambda \in \mathbb{C} | A - \lambda I \text{ is boundedly invertible}\}$. The *spectrum of* A is defined to be the set $\sigma(A) := \mathbb{C} \setminus \rho(A)$. (Most of the times, we will omit the I, and just write $A - \lambda$). We also define the *point spectrum* as $\sigma_p(A) := \{\lambda \in \sigma(A) | A - \lambda \text{ is not injective}\}$.

This definition is exactly the same as the definition for the spectrum for any bounded operator. The only subtlety about this definition, lies with the domain of A. It should therefore be no surprise that the following proposition holds for an arbitrary operator. For a proof, see for example $[10, Prop.\ 2.7,\ p.\ 29]$.

Proposition 2.3.4. *Let $A : \mathscr{H} \to \mathscr{H}$ be a linear operator.*

1. *If $\lambda \in \mathbb{C}$, then $\operatorname{gra}(A)$ is closed if and only if $\operatorname{gra}(A - \lambda)$ is closed.*
2. *If $A \in \mathscr{C}(\mathscr{H})$, then $\sigma(A^*) = \{\overline{\lambda} | \lambda \in \sigma(A)\}$. Additionally if $\lambda \in \rho(A)$ we have*

$$[(A - \lambda)^*]^{-1} = \left[(A - \lambda)^{-1}\right]^*$$

2.4 Symmetric, normal and self-adjoint operators

Now that we have defined what the adjoint of an arbitrary linear operator is, we can continue to define what symmetric operators and self-adjoint operators are. These are the operators we will be considering most of the times in this thesis.

In linear algebra, a linear transformation $M : \mathbb{R}^n \to \mathbb{R}^n$ is called *symmetric* whenever $\langle Mx, y\rangle = \langle x, My\rangle$ for any $x, y \in \mathbb{R}^n$. In functional analysis, when we have a bounded operator $A : \mathscr{H} \to \mathscr{H}$, we call A *self-adjoint* whenever $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in \mathscr{H}$. In conclusion, these two names seem to represent the same operation in some sense. However, for unbounded operators there is an important difference between symmetric and self-adjoint operators.

Definition 2.4.1. If $\mathscr{H}$ is a Hilbert space, and $A : \mathscr{H} \to \mathscr{H}$ a linear operator, then we say A is *symmetric* if A is densely defined, and for all $x, y \in \operatorname{Dom}(A)$ we have:

$$\langle Ax, y\rangle = \langle x, Ay\rangle$$

Definition 2.4.2. Let $\mathscr{H}$ be a Hilbert space. A densely defined operator $A : \mathscr{H} \to \mathscr{H}$ is said to be *self-adjoint* if $A = A^*$.

Remark 2.4.3. We note that Definition 2.4.2 implicitely claims that A is symmetric, and $\operatorname{Dom}(A) = \operatorname{Dom}(A^*)$. So in other words, we cannot define the adjoint on any other vector other than the vectors in $\operatorname{Dom}(A)$. For bounded operators, $\operatorname{Dom}(A) = \mathscr{H}$, so automatically if A is symmetric, it is self-adjoint.

However, for unbounded operators these terms are not quite the same. There exist symmetric operators on $L^2(\mathbb{R})$, that are not self-adjoint. See for example $[1, p.\ 306]$.

Remark 2.4.4. Proposition 2.2.5 tells us that any self-adjoint operator A must have a closed graph.

Proposition 2.4.5. *If $A : \mathscr{H} \to \mathscr{H}$ is densely defined, then the following statements are equivalent:*

1. *A is symmetric.*
2. *for any $x \in \operatorname{Dom}(A)$ we have $\langle Ax, x\rangle \in \mathbb{R}$.*

3. $A \subseteq A^*$.

Proof. 1) $\Rightarrow$ 2) : Let A be symmetric. Then for any $x, y \in \mathrm{Dom}(A)$ we find $\langle Ax, y\rangle = \langle x, Ay\rangle$. So especially $\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$ for any $x \in \mathrm{Dom}(A)$. So $\langle Ax, x\rangle \in \mathbb{R}$.

2) $\Rightarrow$ 1) : Let $x \in \mathrm{Dom}(A)$, and consider $\langle Ax, x\rangle \in \mathbb{R}$. Then $\langle Ax, x\rangle = \overline{\langle x, Ax\rangle} = \langle x, Ax\rangle$. Therefore, let $x, y \in \mathrm{Dom}(A)$. We find then

$$\langle A(x \pm y), x \pm y\rangle - \langle A(x \mp y), x \mp y\rangle = \pm 2\langle Ax, y\rangle \pm 2\langle Ay, x\rangle$$

Therefore it follows that

$$\begin{aligned}&\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle + i\langle A(x+iy), x+iy\rangle - i\langle A(x-iy), x-iy\rangle \\ &= 2\langle Ax, y\rangle + 2\langle Ay, x\rangle + i(-i\langle Ax, y\rangle + i\langle Ay, x\rangle - i\langle Ax, y\rangle + i\langle Ay, x\rangle) = 4\langle Ax, y\rangle\end{aligned}$$

Because $\langle Az, z\rangle = \langle z, Az\rangle$ for any $z \in \mathrm{Dom}(A)$, we find

$$\begin{aligned}\langle Ax, y\rangle &= \frac{1}{4}(\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle \\ &\quad + i\langle A(x+iy), x+iy\rangle - i\langle A(x-iy), x-iy\rangle) \\ &= \frac{1}{4}(\langle x+y, A(x+y)\rangle - \langle x-y, A(x-y)\rangle \\ &\quad + i\langle x+iy, A(x+iy)\rangle - i\langle x-iy, A(x-iy)\rangle) = \langle x, Ay\rangle\end{aligned}$$

Therefore, we see that A is symmetric.

1) $\Rightarrow$ 3) : Let A be symmetric, and consider $x \in \mathrm{Dom}(A)$. Then for any $y \in \mathrm{Dom}(A)$ we have $|\langle Ay, x\rangle| = |\langle y, Ax\rangle| \leq \|y\|\|Ax\|$, so $y \mapsto \langle Ay, x\rangle$ is bounded, and thus $x \in \mathrm{Dom}(A^*)$, and we see $A^*x = Ax$ for any $x \in \mathrm{Dom}(A)$. So $A \subseteq A^*$.

3) $\Rightarrow$ 1) : If $A \subseteq A^*$, we have $A^*x = Ax$ for $x \in \mathrm{Dom}(A)$. Hence, for any $y \in \mathrm{Dom}(A)$ we see that $\langle Ay, x\rangle = \langle y, A^*x\rangle = \langle y, Ax\rangle$. So A is symmetric. □

Lemma 2.4.6. *Let $A : \mathscr{H} \to \mathscr{H}$ be a symmetric linear operator. If B is a symmetric extension of A, then*

$$A \subseteq B \subseteq B^* \subseteq A^* \tag{6}$$

Proof. Let B be a symmetric extension of A. The first inclusion is by definition, and the second by Proposition 2.4.5. For the third inclusion, consider $y \in \mathrm{Dom}(B^*)$. We then have $\langle Bx, y\rangle = \langle Ax, y\rangle$ for any $x \in \mathrm{Dom}(A)$. Therefore $y \in \mathrm{Dom}(A^*)$ and we find $B^*y = A^*y$. So A^* is an extension of B^*. □

Another type of operator which is important in Operator Theory, is a *normal* linear operator. For a bounded operators A, we say A is normal if $A^*A = AA^*$, so A and A^* commute. For arbitrary operators, we have a similar definition, with the subtlety of defining the domains in the proper way.

Definition 2.4.7. Let $\mathscr{H}$ be a Hilbert space. A linear operator A on $\mathscr{H}$ is *normal* if A is a closed, densely defined operator and $A^*A = AA^*$.

Remark 2.4.8. Note that the equation $A^*A = AA^*$ implicitly carries the condition $\mathrm{Dom}(A^*A) = \mathrm{Dom}(AA^*)$, where we define $\mathrm{Dom}(AA^*)$ as in Remark 2.1.7. It does not say anything about the domain of A or A^* itself. Apart from that, it is the same definition as for bounded operators on a Hilbert space.

Also note, that if A is a self-adjoint operator, then A is a normal operator.

In light of the theorem of Nelson, we need to introduce another definition, which is almost the same as self-adjointness. In short, we only need to take the closure in order to find an self-adjoint operator.

Definition 2.4.9. Let $\mathscr{H}$ be a Hilbert space. A linear operator $A : \mathscr{H} \to \mathscr{H}$ is called *essentially self-adjoint*, if A is closable and its closure $\overline{A}$ is self-adjoint. In other words $\overline{A} = \overline{A}^*$.

For the rest of this chapter, we will have a look at the spectral properties of symmetric operators. We will need them in order to prove the Spectral Theorem, which will be used in the theorem of Nelson. A lot of properties of the spectrum for bounded self-adjoint operators still hold, especially the lemma which states that the spectrum of a self-adjoint operator is a subset of the real line.

Additionally, we will see that there is an interesting way of concluding that a symmetric operator is self-adjoint. We will use this characterization later on to prove that some operators are self-adjoint.

Lemma 2.4.10. *Let A be a symmetric operator on a Hilbert space $\mathscr{H}$, and let $\lambda \in \mathbb{C}$ be given as $\lambda = \alpha + i\beta$ with $\alpha, \beta \in \mathbb{R}$. Then*

1. *For every $x \in \mathrm{Dom}(A)$ we have $\|(A-\lambda)x\|^2 = \|(A-\alpha)x\|^2 + \beta^2\|x\|^2$*
2. *If $\beta \neq 0$ we have $\ker(A-\lambda) = (0)$*
3. *If A is closed and $\beta \neq 0$, then $\mathrm{ran}(A-\lambda)$ is closed*

Proof. 1): We note that for any $x \in \mathrm{Dom}(A)$

$$\begin{aligned}\|(A-\lambda)x\|^2 &= \|(A-\alpha)x\|^2 + \|i\beta x\|^2 - 2\mathrm{Re}(\langle (A-\alpha)x, i\beta x\rangle)\\ &= \|(A-\alpha)x\|^2 + \beta^2\|x\|^2 + 2\mathrm{Re}(i\langle (A-\alpha)x, \beta x\rangle)\\ &= \|(A-\alpha)x\|^2 + \beta^2\|x\|^2 + 2\mathrm{Re}(i(\beta\langle Ax, x\rangle + \alpha\beta\langle x, x\rangle))\end{aligned}$$

Since A is symmetric, $\langle Ax, x\rangle \in \mathbb{R}$ and so part 1) follows.

2): Since $\|(A-\lambda)x\|^2 = \|(A-\alpha)x\|^2 + \beta^2\|x\|^2 \geq \beta^2\|x\|^2$, if $\beta \neq 0$, we have $\ker(A-\lambda) = (0)$.

3): Let A be closed and $\beta \neq 0$. Take a sequence $(y_n)_n \subseteq \mathrm{ran}(A-\lambda)$ such that $y_n \to y$ as $n \to \infty$. We wish to show $y \in \mathrm{ran}(A-\lambda)$. Because $y_n \in \mathrm{ran}(A-\lambda)$, there exists $x_n \in \mathrm{Dom}(A-\lambda)$ such that $y_n = (A-\lambda)x_n$. Notice that $\|x_n\| \leq \|(A-\lambda)x_n\|$ and so $(x_n)_n$ is a Cauchy sequence, and thus has a limit in $\mathscr{H}$. Define $x := \lim_{n\to\infty} x_n$. Since $(x_n, y_n) \in \mathrm{gra}(A-\lambda)$, we have $(x_n, y_n) \to (x, y)$. Because A is a closed operator we have $\mathrm{gra}(A-\lambda)$ is closed. Thus $(x, y) \in \mathrm{gra}(A-\lambda)$. Therefore $y \in \mathrm{ran}(A-\lambda)$. □

Theorem 2.4.11. *[1, Cor. X.2.9, p. 311] If A is a closed, symmetric linear operator on $\mathscr{H}$, then the following statements are equivalent.*

1. *A is self-adjoint*
2. $\sigma(A) \subseteq \mathbb{R}$
3. $\ker(A^* - i) = \ker(A^* + i) = 0$

Proof. 1) $\Rightarrow$ 2): Let A be a self-adjoint operator. If $x \in \ker(A - \lambda)$ then $Ax = \lambda x$. Then $\lambda\|x\|^2 = \langle \lambda x, x\rangle = \langle Ax, x\rangle$ which is a real number. Therefore $\lambda \in \mathbb{R}$. Let $\operatorname{Im}(\lambda) \neq 0$. Then we can conclude $\ker(A - \lambda) = \ker(A^* - \lambda) = (0)$. So $A - \lambda$ is injective. It is easy to see that $\operatorname{Dom}(A) = \operatorname{Dom}(A - \lambda)$, so $A - \lambda$ is densely defined. Additionally $(A - \lambda)^* = A^* - \overline{\lambda} = A - \overline{\lambda}$. Then by Corollary 2.2.8 we find

$$[\operatorname{ran}(A - \lambda)]^{\perp} = \ker((A - \lambda)^*) = \ker(A - \overline{\lambda}) = (0)$$

Therefore $\operatorname{ran}(A - \lambda)$ is dense. By Lemma 2.4.10, $A - \lambda$ has closed range, and so $A - \lambda$ is surjective. Therefore by Proposition 2.3.2 we have $A - \lambda$ is boundedly invertible, and so $\lambda \in \rho(A)$. Therefore $\sigma(A) \subseteq \mathbb{R}$.

2) $\Rightarrow$ 3): By Corollary 2.2.8, we have $\ker(A^* \pm i) = [\operatorname{ran}(A \mp i)]^{\perp}$. Since $\sigma(A) \subseteq \mathbb{R}$, $A \pm i$ is boundedly invertible, and so $[\operatorname{ran}(A \mp i)]^{\perp} = \mathscr{H}^{\perp} = (0)$. So 3) follows.

3) $\Rightarrow$ 1): If 3) holds, we have by Corollary 2.2.8 that $\operatorname{ran}(A + i)$ is dense, and by Lemma 2.4.10 we have $A + i$ is surjective. So if $x \in \operatorname{Dom}(A^*)$, there exists $y \in \operatorname{Dom}(A)$ such that $(A + i)y = (A^* + i)x$. But since A is symmetric, $A^*y = Ay$ and so $(A^* + i)y = (A^* + i)x$ and thus $y = x \in \operatorname{Dom}(A)$. So $A = A^*$. □

As it seems, the dimension of $\ker(A \pm i)$ seem to play an important role in showing whether A is self-adjoint or not. It is convenient to give them names, for we will need them again.

Definition 2.4.12. Suppose that A is a closed symmetric operator on a Hilbert space $\mathscr{H}$. Let

$$\mathscr{L}_+ := \ker(A^* - i) = [\operatorname{ran}(A + i)]^{\perp} \quad (7)$$

$$\mathscr{L}_- := \ker(A^* + i) = [\operatorname{ran}(A - i)]^{\perp} \quad (8)$$

$\mathscr{L}_+$ and $\mathscr{L}_-$ are called the *deficiency subspaces* of A. The pair of numbers $n_{\pm} := \dim(\mathscr{L}_{\pm})$ are called the *deficiency indices* of A.

Remark 2.4.13. We remark that for an arbitrary linear operator A, it is possible for the deficiency indices to be any pair of nonnegative integers. It is also possible that n_+ or n_- (or both) are ∞. See for example $[6, p.\ 138]$.

Definition 2.4.14. A *partial isometry* is a linear operator W on a Hilbert space $\mathscr{H}$ with $\operatorname{Dom}(W) = \mathscr{H}$ such that for $x \in [\ker(W)]^{\perp}$, $\|Wx\| = \|x\|$. We define $[\ker(W)]^{\perp}$ the *initial space* of W, and ran(W) is defined as the *final space* of W.

The following theorem gives a one-to-one correspondence between closed symmetric extensions of a symmetric operator, and partial isometries. We will not prove this theorem. The reader who is interested in the proof, may read $[1, p.\ 313 - 315]$ or $[6, p.\ 138 - 140]$.

Theorem 2.4.15. *[1, Thm. X.2.17, p. 314] Let A be a closed symmetric operator on a Hilbert space $\mathscr{H}$. If W is a partial isometry with initial space in $\mathscr{L}_+$ and final space $\mathscr{L}_-$, define*

$$\mathscr{D}_W := \{x + y + Wy | x \in \text{Dom}(A), y \in initial(W)\}$$

and define A_W on $\mathscr{D}_W$ by

$$A_W(x + y + Wy) = Ax + iy - iWy$$

Then A_W is a closed symmetric extension of A. Conversely, if B is a closed symmetric extension of A, then there exists a unique partial isometry W such that $B = A_W$ with A_W defined as above.

The following lemma is just a recap of the results of some of the theorems, but gives an efficient way of showing whether an operator is self-adjoint or has self-adjoint extensions.

Lemma 2.4.16. *Let A be a closed symmetric operator with deficiency indices $n_\pm$. Then*

1. *A is self-adjoint if and only if $n_+ = n_- = 0$*

2. *A has a self-adjoint extension if and only if $n_+ = n_-$. In this case the set of self-adjoint extensions is in a natural correspondence with the set of unitary isomorphisms of $\mathscr{L}_+$ onto $\mathscr{L}_-$*

Proof. The proof of $1)$ is just a rephrasing of $2.4.11$.

For $2)$, we first note that $\mathscr{L}_\pm$ are closed linear subspaces, and thus Hilbert spaces. Additionally note that $n_+ = n_-$ means that $\dim(\mathscr{L}_+) = \dim(\mathscr{L}_-)$. We assumed that all of our Hilbert spaces are separable, so the equality $n_+ = n_-$ is true if and only if $\mathscr{L}_+$ and $\mathscr{L}_-$ are isomorphic. But saying that the two are isomorphic, is equivalent to saying there is a partial isometry W with initial space $\mathscr{L}_+$ and final space $\mathscr{L}_-$. □

3 Spectral theorem for bounded normal operators

If we want to understand the theorem of Nelson, we need to understand what a spectral measure is. A spectral measure is a function that sends every measurable subset of a set X to a projection operator in a Hilbert space $\mathscr{H}$. It follows that we can 'diagonalize' any unbounded self-adjoint operator A as an 'integral' over the spectrum of A. This is known as the spectral theorem for unbounded self-adjoint operators. In order to prove the spectral Theorem for unbounded self-adjoint operators, we need the spectral theorem for *bounded normal* operators. Thus we will study this first. Readers already familiar with the Spectral Theorem for bounded operators, may skip Section 4.

3.1 C^*-algebras and representations

The spectral theorem for bounded normal operators is a corollary of a more general theorem, which states that there exists an one-to-one correspondence between spectral measures and representations of certain C^*-algebras. In order to state this more general theorem, we need to know what spectral measures and representations of C^*-algebras are. In this section, we will shortly introduce C^*-algebras and representations. We assume $\mathbb{K}$ as either one of the fields $\mathbb{R}$ or $\mathbb{C}$.

Definition 3.1.1. An *algebra* $\mathscr{A}$ over $\mathbb{K}$, is a vector space $\mathscr{A}$ over $\mathbb{K}$ that also has a multiplication defined on it that makes it into a ring such that if $\lambda \in \mathbb{K}$ and $x, y \in \mathscr{A}$ then $\lambda(xy) = (\lambda x)y = x(\lambda y)$.

Remark 3.1.2. Note that having an identity element in the algebra is not included in the definition. It is therefore not needed for any algebra to have an identity.

Definition 3.1.3. We define a *Banach algebra* as an algebra $\mathscr{A}$ over $\mathbb{K}$ with a norm $\|\cdot\|$ such that $\mathscr{A}$, equipped with this norm, is a Banach space and such that for any $x, y \in \mathscr{A}$,

$$\|xy\| \leq \|x\|\|y\|$$

If $\mathscr{A}$ has an identity element 1, it is assumed $\|1\| = 1$

Example 3.1.4. Let X be a Banach space, and put $\mathscr{A} = \mathscr{B}(X)$. If multiplication is defined as composition, $\mathscr{A}$ becomes a Banach algebra with identity. ⊘

Now that we have a notion of an algebra, we can define what a C^*-algebra is. In short, we add an additional operation on our algebra, that looks a lot like the conjugation operation in $\mathbb{C}$. But in more general setiings, the conjugation operation is not necessarily abelian. In other words, if $x, y \in \mathscr{A}$ we might have $(xy)^* \neq (yx)^*$.

Definition 3.1.5. We define an *involution* on a Banach algebra $\mathscr{A}$ as the map $x \mapsto x^*$ of $\mathscr{A}$ into $\mathscr{A}$ in such a way that the following properties hold for any $x, y \in \mathscr{A}$ and $\lambda \in \mathbb{K}$

- $(x^*)^* = x$
- $(xy)^* = y^*x^*$
- $(\lambda x + y)^* = \overline{\lambda}x^* + y^*$

Definition 3.1.6. A *C^*-algebra* over $\mathbb{K}$ is a Banach algebra $\mathscr{A}$ over $\mathbb{K}$ with an involution such that for every $x \in \mathscr{A}$ we have

$$\|x^*x\| = \|x\|^2$$

Example 3.1.7. $\mathbb{C}$ is a C^*-algebra, if we consider the conjugation defined on $\mathbb{C}$ as the involution. Also $M_n(\mathbb{C})$, the space of complex-valued $n \times n$ matrices, is a C^*-algebra if we consider the complex conjugate as the involution ⊘

Example 3.1.8. Let $\mathscr{H}$ be a Hilbert space. We already know that $\mathscr{B}(\mathscr{H})$ is a Banach algebra. If we define the involution operation * as A^*= the operator's adjoint, then this defines an involution, and so $\mathscr{B}(\mathscr{H})$ is a C^*-algebra. ⊘

In Group Theory, one often encounters homomorphisms and isomorphism. Such mappings conserve the structure of the group. For algebras, we have a similar notion of homomorphisms and isomorphisms.

Definition 3.1.9. Let $\mathscr{A}_1, \mathscr{A}_2$ be C^*-algebras over $\mathbb{K}$, and $h : \mathscr{A}_1 \to \mathscr{A}_2$. We call h a *homomorphism of a C^*-algebra* if $h(xy) = h(x)h(y)$ and $h(x+\lambda y) = h(x)+\lambda h(y)$ for every $x, y \in \mathscr{A}_1$ and $\lambda \in \mathbb{K}$. We call h an *isomorphism* if h is bijective. We call h a **-homomorphism* if h is a homomorphism, and in addition $h(x^*) = h(x)^*$ is true for any $x \in \mathscr{A}_1$. We call h a **-isomorphism* if h is an isomorphism and a *-homomorphism.

Definition 3.1.10. Let $\mathscr{A}$ be a C^*-algebra over $\mathbb{K}$. A *representation of a C^*-algebra* is a pair $(\mathscr{H}, \pi)$ where $\mathscr{H}$ is a Hilbert space and $\pi : \mathscr{A} \to \mathscr{B}(\mathscr{H})$ a *-homomorphism, where $\mathscr{B}(\mathscr{H})$ is as usual the space of bounded operators on $\mathscr{H}$. If $\mathscr{A}$ has an identity, it is assumed $\pi(1) = I$, the identity operator. Often the mention of $\mathscr{H}$ is omitted and say π is a representation of the C^*-algebra.

3.2 Spectral measures

Next, we wish to define a spectral measure. This will turn out to be a projection operator on a Hilbert space $\mathscr{H}$ which is dependent upon the Borel subsets of $\mathbb{C}$. In this way, it will give a sort of 'weight' to each Borel subset of $\mathbb{C}$, and therefore we can integrate over this spectral measure. We will look into all of these items in this section.

Definition 3.2.1. Let $\mathscr{H}$ be a Hilbert space. An *idempotent* on $\mathscr{H}$ is a bounded, linear operator $E : \mathscr{H} \to \mathscr{H}$ such that $E^2 = E$. A *projection* is an idempotent E such that $\ker(E) = \operatorname{ran}(E)^\perp$.

Lemma 3.2.2. *Let $\mathscr{H}$ be a Hilbert space. If E is an idempotent on $\mathscr{H}$, then* $\ker(E) = \operatorname{ran}(I - E)$ *and* $\ker(I - E) = \operatorname{ran}(E)$.

Proof. If $x \in \ker(E)$ then $Ex = 0$, and so $(I - E)x = x - Ex = x$, so $x \in \operatorname{ran}(I - E)$. On the other hand if $x \in \operatorname{ran}(I - E)$ then there exists a $y \in \mathscr{H}$ such that $(I - E)y = x$. So

$$(I - E)x = (I - E)^2 y = (I - 2E + E^2)y = (I - E)y = x$$

Thus $x - Ex = x$ and we conclude $x \in \ker(E)$. So the first part is proven. The proof for the second equality goes in a similar way, so we will omit it. □

Lemma 3.2.3. *[1, Thm. II.3.3, p. 37] Let $\mathscr{H}$ be a Hilbert space, E an idempotent on $\mathscr{H}$ and $E \neq 0$. Then the following statements are equivalent:*

1. *E is a projection.*
2. *E is a self-adjoint operator.*
3. *E is a normal operator.*
4. $\|E\| = 1$
5. $\langle Ex, x\rangle \geq 0$ *for any* $x \in \mathscr{H}$.

Proof. $1) \Rightarrow 2)$: Assume E to be a projection. Let $x, y \in \mathscr{H}$. Then $x = x_1 + x_2$ and $y = y_1 + y_2$ with $x_1, y_1 \in \ker(E) = (\operatorname{ran}(E))^\perp$ and $x_2, y_2 \in (\ker(E))^\perp = \overline{\operatorname{ran}(E)}$. Therefore if $x_2 \in \ker(E)^\perp$, we can find a sequence $(x_n)_n \subseteq \mathscr{H}$ such that $Ex_n \to x_2$ as $n \to \infty$. Then we find that

$$\langle Ex, y\rangle = \langle Ex_2, y\rangle = \lim_{n\to\infty} \langle EEx_n, y\rangle = \lim_{n\to\infty} \langle Ex_n, y\rangle = \lim_{n\to\infty} (\langle Ex_n, y_1\rangle + \langle Ex_n, y_2\rangle)$$

Because $y_1 \in \operatorname{ran}(E)^\perp$, we find that $\langle Ex_n, y_1\rangle = 0$. Therefore we conclude

$$\langle Ex, y\rangle = \lim_{n\to\infty} \langle Ex_n, y_2\rangle = \langle x_2, y_2\rangle$$

We could do the same reasoning for y, and we see that $\langle x, Ey\rangle = \langle x_2, y_2\rangle$. So we find $\langle Ex, y\rangle = \langle x, Ey\rangle$ for $x, y \in \mathscr{H}$. Therefore, E is self-adjoint.

$2) \Rightarrow 3)$: There should be no surprises here, it is true by definition.

$3) \Rightarrow 1)$: Let E be a normal operator. Note that for any $x \in \mathscr{H}$ we have

$$\|Ex\|^2 = \langle Ex, Ex\rangle = \langle E^*Ex, x\rangle = \langle EE^*x, x\rangle = \langle E^*x, E^*x\rangle = \|E^*x\|^2$$

So $\ker(E) = \ker(E^*)$. By Corollary 2.2.8 we have $\ker(E^*) = \operatorname{ran}(E)^\perp$, and so E is a projection.

$1) \Rightarrow 4)$: If $x \in \mathscr{H}$ we have $\|Ex\|^2 = \langle Ex, Ex\rangle = \langle Ex, x\rangle \leq \|Ex\|\|x\|$ because E is self-adjoint. So $\|Ex\| \leq \|x\|$ for $x \in \mathscr{H}$. On the other hand, if $y \in \operatorname{ran}(E)$, then there exists a $z \in \mathscr{H}$ such that $y = Ez$. Then

$$\|Ey\| = \|E(Ez)\| = \|Ez\| = \|y\|$$

So this concludes that $\|E\| = 1$.

$4) \Rightarrow 1)$: Let $x \in (\ker(E))^\perp$. Because $\ker(E) = \operatorname{ran}(I - E)$, we have $x - Ex \in \ker(E)$. This means that $0 = \langle x - Ex, x\rangle = \|x\|^2 - \langle Ex, x\rangle$ and thus $\|x\|^2 = \langle Ex, x\rangle \leq \|Ex\|\|x\| \leq \|x\|^2$. Therefore the inequality signs must be an equality sign. So we conclude that for any $x \in (\ker(E))^\perp$ we have $\|Ex\|^2 = \|x\|^2 = \langle Ex, x\rangle$. But then we can conclude that

$$\|x - Ex\|^2 = \|x\|^2 + \|Ex\|^2 - 2\mathrm{Re}(\langle Ex, x\rangle) = 0$$

Thus we find $x \in \ker(I - E) = \operatorname{ran}(E)$ and so $(\ker(E))^\perp \subseteq \operatorname{ran}(E)$.

On the other hand, let $y \in \operatorname{ran}(E)$. So we can write $y = y_1 + y_2$ with $y_1 \in \ker(E)$ and $y_2 \in (\ker(E))^\perp$. Then $y_2 \in \operatorname{ran}(E)$, and so $Ey_2 = y_2$. So $y = Ey = Ey_2 = y_2$ and so $y \in (\ker(E))^\perp$. So E is a projection.

$1) \Rightarrow 5)$: Let E be a projection. Then by previous statements $E = E^*$, and thus is

$$0 \leq \|Ex\|^2 = \langle Ex, Ex\rangle = \langle E^2x, x\rangle = \langle Ex, x\rangle$$

for any $x \in \mathscr{H}$. So $\langle Ex, x\rangle \geq 0$

$5) \Rightarrow 1)$: Let $x \in \operatorname{ran}(E)$ and $y \in \ker(E)$. Then $0 \leq \langle E(x + y), x + y\rangle = \langle x, x\rangle + \langle x, y\rangle$. Thus $-\|x\|^2 \leq \langle x, y\rangle$. Assume there exists x and y such that $\langle x, y\rangle = \lambda \neq 0$. Then certainly if $z = \mu y$ for $\mu \in \mathbb{C} \setminus \{0\}$ gives $\langle x, z\rangle = \overline{\mu}^{-1}\langle x, y\rangle \neq 0$. Thus, define $\mu = -2\overline{\lambda^{-1}}\|x\|^2$, we find $\langle x, y\rangle = -2\lambda^{-1}\|x\|^2\lambda = -2\|x\|^2$, and thus we find $\|x\|^2 \leq -2\|x\|^2$ which is a contradiction. Thus $\langle x, y\rangle = 0$ for any $x \in \operatorname{ran}(E)$ and $y \in \ker(E)$. Hence E is a projection.

□

For our definition of a spectral measure, we first need to introduce some topologies defined on $\mathscr{B}(\mathscr{H})$.

Definition 3.2.4. If $\mathscr{H}$ is a Hilbert space, we define the *weak operator topology* (WOT) on $\mathscr{B}(\mathscr{H})$ as the locally convex topology given by the seminorms $\{p_{x,y} | x, y \in \mathscr{H}\}$ where $p_{x,y}(A) = |\langle Ax, y\rangle|$. The *strong operator topology* (SOT) is the topology defined on $\mathscr{B}(\mathscr{H})$ by the family of seminorms $\{p_x | x \in \mathscr{H}\}$, where $p_x(A) = \|Ax\|$.

The next proposition gives some properties of these topologies. Because they do not give a lot of insight in the problem we are working with, we will omit the proof.

Proposition 3.2.5. *Let $\mathscr{H}$ be a Hilbert space, and $(A_j)_j$ a sequence in $\mathscr{B}(\mathscr{H})$. Then*

1. *$A_j \to A$ (WOT) if and only if $\langle A_j x, y\rangle \to \langle Ax, y\rangle$ for any $x, y \in \mathscr{H}$*

2. *If $\sup\{\|A_j\|\} < \infty$ and $\mathscr{A} \subseteq \mathscr{H}$ a subset such that $\overline{span(\mathscr{A})} = \mathscr{H}$, then $A_j \to A$ (WOT) if and only if $\langle A_j x, y\rangle \to \langle Ax, y\rangle$ for any $x, y \in \mathscr{A}$*

3. *$A_j \to A$ (SOT) if and only if $\|A_j x - Ax\| \to 0$ for any $x \in \mathscr{H}$*

4. *If $\sup\{\|A_j\|\} < \infty$ and $\mathscr{A} \subseteq \mathscr{H}$ such that $\overline{span(\mathscr{A})} = \mathscr{H}$, then $A_j \to A$ (SOT) if and only if $\|A_j x - Ax\| \to 0$ for any $x \in \mathscr{A}$*

5. *If $\mathscr{H}$ is separable, then the WOT and SOT are metrizable on bounded subsets of $\mathscr{B}(\mathscr{H})$*

With all the previous definitions, we can finally define a spectral measure. Remember that a σ-algebra $\mathscr{A}$ of a set X, is a collection of subsets of X such that $\emptyset \in \mathscr{A}$, if $A \in \mathscr{A}$ then $X \setminus A \in \mathscr{A}$, and finally if $(A_j)_{j\in I} \subseteq \mathscr{A}$ is a countable sequence of sets, then $\bigcup_{j\in I} A_j \in \mathscr{A}$. If X is a topological space, then the *Borel σ-algebra* is defined as the σ-algebra generated by all open subsets of X.

Definition 3.2.6. Let X be a set, $\mathscr{A}$ a σ-algebra of X, and $\mathscr{H}$ a Hilbert space. We define a *spectral measure* for $(X, \mathscr{A}, \mathscr{H})$ as a function $E : \mathscr{A} \to \mathscr{B}(\mathscr{H})$ such that:

1. For any set $A \in \mathscr{A}$ we have $E(A)$ is a projection.

2. $E(\emptyset) = 0$ and $E(X) = I$.

3. If $A, B \in \mathscr{A}$ then $E(A \cap B) = E(A)E(B)$.

4. If $(A_j)_j$ is a pairwise disjoint sequence of sets in $\mathscr{A}$, then

$$E\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} E(A_j)$$

Remark 3.2.7. We need to say a few things about the last part of the definition. It is not *a priori* clear that the sum of part 4 of the definition converges at all. We will show that the sum does converge with respect to the strong operator. First we note that if A, B are disjoint, we have that $A \cap B = \emptyset$ and so by part 2 and 3, we see that $0 = E(A)E(B) = E(A \cap B) = E(B)E(A)$, and so the projections have orthogonal ranges. With the following claim we can conclude that part 4 of the definition is unambiguous, if $(E(A_n))_n$ is a sequence of projections with orthogonal ranges.

Claim 1. For any sequence of projections $(E_j)_j$ with pairwise orthogonal ranges, we have that $\sum_{j=1}^{n} E_n \to E$ (SOT) where E is the projection with $\operatorname{ran}(E) = \overline{\oplus_{j\in\mathbb{N}} \operatorname{ran}(E_j)}$. We write $E = \sum_{j=1}^{\infty} E_j$.

Proof. Let $0 \neq x \in \mathscr{H}$ be given. First we show that the sum converges (SOT). We do this by constructing a orthonormal set, and then use the identity of Parseval. Define

$$A := \left\{ \frac{E_j x}{\|E_j x\|} \mid j \in \mathbb{N} \text{ and } \|E_j x\| \neq 0 \right\}$$

We claim A is an orthonormal set. For take any $y, z \in A$, then there exists $j, k \in \mathbb{N}$ such that $\langle y, z\rangle = \frac{1}{\|E_j x\| \|E_k x\|}\langle E_j x, E_k x\rangle$. Since E_j is a projection, it is self adjoint. But E_j and E_k have orthogonal ranges, so $\langle E_j x, E_k x\rangle = 0$. So A is an orthogonal set. It should be clear that any vector in A has norm 1. Therefore A is an orthonormal system, and thus we can extend A such that it becomes a orthonormal basis, call it A'. Then by Parseval's Identity, we see

$$\|x\|^2 = \sum_{y\in A'} |\langle x, y\rangle|^2 \geq \sum_{y\in A} |\langle x, y\rangle|^2 = \sum_{A} \frac{1}{\|E_j x\|^2} |\langle x, E_j x\rangle|^2 = \sum_{j=1}^{\infty} \|E_j x\|^2$$

by using that E_j is self-adjoint and idempotent. So we see that $\sum_{j=1}^{\infty} \|E_j x\|^2 \leq \|x\|^2$.

We note that $\| \sum_{j=1}^{n} E_j x\|^2 = \sum_{j=1}^{n} \|E_j x\|^2$ because of the orthogonality of $\operatorname{ran}(E_j)$. Therefore we see that $\| \sum_{j=1}^{\infty} E_j x\|^2 = \sum_{j=1}^{\infty} \|E_j x\|^2 \leq \|x\|^2$ and so the series $\sum_j E_j$ converges (SOT) to a unique operator. We also proved that $\sum_{j=1}^{\infty} E_j$ is a bounded operator, and linear. Note that the range of this operator is $\overline{\oplus \operatorname{ran}(E_j)}$. Also $E_j E_k = E_k E_j = 0$ for any $j \neq k$, and therefore we conclude $(\sum_{j=1}^{n} E_j)^2 = \sum_{j=1}^{n} E_j^2 = \sum_{j=1}^{n} E_j$, and so $\sum_{j=1}^{\infty} E_j$ is idempotent. Finally, since $\sum_{j=1}^{n} E_j$ is self adjoint, and $\langle \cdot, \cdot \rangle$ is continuous, we have that $\sum_{j=1}^{\infty} E_j$ is self adjoint, and thus a projection. □

Example 3.2.8. Let X be a compact space, $\mathscr{A}$ the Borel σ-algebra, and $\mathscr{H} = L^2(X, \mu)$ where μ is a Borel measure. Then define $E : \mathscr{A} \to \mathscr{B}(\mathscr{H})$ by $E(A) = \chi_A$, the characteristic function on A. Then E is a spectral measure on $(X, \mathscr{A}, L^2(X, \mu))$. ⊘

Example 3.2.9. Let X be any set, $\mathscr{A} = \mathscr{P}(X)$: the power set of X, and $\mathscr{H}$ is a separable Hilbert space. Because $\mathscr{H}$ is separable, we can find a countable orthonormal basis $(e_j)_j$. Fix a sequence $(x_n)_n$ in X. For any $A \in \mathscr{A}$, define $E(A)$ as the projection upon $\overline{span\{e_j | x_j \in A\}}$. Then E is a spectral measure. ⊘

Definition 3.2.10. Let X be any set, and $\mathscr{A}$ be a σ-algebra on X. If μ is a measure on $(X, \mathscr{A})$ and $A \in \mathscr{A}$, we define the *variation of* μ, denoted $|\mu|$, by

$$|\mu|(A) = \sup \left\{ \sum_{j=1}^{m} |\mu(E_j)| \mid (E_j)_j^m \text{is a measurable partition of } A \right\} \tag{9}$$

We define the *total variation of* μ as $\|\mu\| = |\mu|(X)$

The name spectral measure suggests that this operator is some sort of measure. Indeed, the spectral measure gives rise to a measure, as the following lemma shows.

Lemma 3.2.11. *If E is a spectral measure for $(X, \mathscr{A}, \mathscr{H})$ and $x, y \in \mathscr{H}$, then*

$$E_{x,y}(A) := \langle E(A)x, y\rangle \tag{10}$$

defines a (complex-valued) measure on $\mathscr{A}$. Additionally, $\|E_{x,y}\| \leq \|x\|\|y\|$.

Proof. Of course, $E_{x,y}(\emptyset) = \langle 0, y \rangle = 0$. Next, take a countable sequence $(A_j)_{j \in I} \subseteq \mathscr{A}$ of pairwise disjoint sets. Then we have

$$E_{x,y}\left(\bigcup_{j=1}^{\infty} A_j\right) = \left\langle E\left(\bigcup_{j=1}^{\infty} A_j\right) x, y \right\rangle = \left\langle \sum_{j=1}^{\infty} E(A_j)x, y \right\rangle = \sum_{j=1}^{\infty} \langle E(A_j)x, y \rangle = \sum_{j=1}^{\infty} E_{x,y}(A_j)$$

So $E_{x,y}$ is a measure. Next, if $A_1, A_2, \ldots, A_n$ are pairwise disjoint sets in $\mathscr{A}$, let $a_j \in \mathbb{C}$ such that $|a_j| = 1$ and $|\langle E(A_j)x, y \rangle| = a_j \langle E(A_j)x, y \rangle$. Then

$$\sum_{j=1}^{n} |E_{x,y}(A_j)| = \sum_{j=1}^{n} a_j \langle E(A_j)x, y \rangle = \left\langle \sum_{j=1}^{n} E(A_j)a_j x, y \right\rangle \leq \left\| \sum_{j=1}^{n} E(A_j)a_j x \right\| \|y\|$$

But note that the set $\{E(A_j)a_j x | \ 1 \leq j \leq n\}$ is a finite set of orthogonal vectors, because $(E(A_j))$ is a sequence of pairwise orthogonal projections. Thus

$$\left\| \sum_j E(A_j)a_j x \right\|^2 = \sum_j \|E(A_j)x\|^2 = \left\| \sum_j E(A_j)x \right\|^2 = \left\| E(\bigcup_{j=1}^{\infty} A_j)x \right\|^2 \leq \|x\|^2$$

Hence $\sum_j |E_{x,y}(A_j)| \leq \|x\| \|y\|$. So $\|E_{x,y}\| \leq \|x\| \|y\|$. □

Now that we know that there exists a measure, we can use the tools of Measure Theory. Especially, we can integrate functions with respect to this measure. This will help us define the integral over the spectral measure itself. It turns out that if we integrate a bounded function with respect to this measure, we can find a bounded linear operator such that the inner product of this bounded linear operator can be described as the integral over this function. In order to prove this, we need a lemma.

Definition 3.2.12. If $\mathscr{H}, \mathscr{K}$ are Hilbert spaces, a function $f : \mathscr{H} \times \mathscr{K} \to \mathbb{K}$ is called a *sesquilinear form* if for $x, y \in \mathscr{H}$, $\phi, \psi \in \mathscr{K}$ and $\alpha, \beta \in \mathbb{K}$,

1. $f(\alpha x + \beta y, \phi) = \alpha f(x, \phi) + \beta f(y, \phi)$
2. $f(x, \alpha\phi + \beta\phi) = \overline{\alpha} f(x, \phi) + \overline{\beta} f(x, \psi)$

We say f is *bounded*, if $\sup_{\|x\|, \|\phi\| \leq 1} \{\|f(x, \phi)\|\} = M < \infty$. The constant M is called a *bound of* f

Lemma 3.2.13. *Let $\mathscr{H}, \mathscr{K}$ be Hilbert spaces. If $f : \mathscr{H} \times \mathscr{K} \to \mathbb{K}$ is a bounded sesquilinear form with bound M, then there exist unique operators $A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ and $B \in \mathscr{B}(\mathscr{K}, \mathscr{H})$ such that*

$$f(x, \phi) = \langle Ax, \phi \rangle = \langle x, B\phi \rangle$$

for all $x \in \mathscr{H}$ and $\phi \in \mathscr{K}$. Additionally $\|A\|, \|B\| \leq M$.

Proof. We will only prove the existence of A. The existence of A is showed in a similar way. Let $\phi \in \mathscr{K}$. Then $f_\phi : \mathscr{H} \to \mathbb{K}, x \mapsto f(x, \phi)$ is a bounded linear functional. By the Riesz Representation Theorem, there exists a unique $y \in \mathscr{H}$ such that $f_\phi(x) = \langle x, y \rangle$. Define the operator $B : \mathscr{K} \to \mathscr{H}$ by $B\phi = y$. So $f_\phi(x) = \langle x, B\phi\rangle$. By uniqueness of y, we have a well-defined operator. Additionally, by uniqueness of the Riesz Representation Theorem, B is a linear operator. Finally, the Riesz Representation Theorem also tells us that $\|f_\phi\| = \|y\|$ and thus:

$$\|B\| = \sup_{\|\phi\|\leq 1} \|B\phi\| = \sup_{\|\phi\|\leq 1} \|f_\phi\| = \sup_{\|x\|,\|\phi\|\leq 1} \|f_\phi(x)\| = M$$

So $B \in \mathscr{B}(\mathscr{K}, \mathscr{H})$ and $\|B\| \leq M$.

Next, assume there is another $B_2 \in \mathscr{B}(\mathscr{K}, \mathscr{H})$ which has these properties. Then $f(x, \phi) = \langle x, B\phi\rangle = \langle x, B_2\phi\rangle$, so $\langle x, (B - B_2)\phi\rangle = 0$. This is true for any $x \in \mathscr{H}$ so $(B - B_2)\phi = 0$. This again is true for any $\phi \in \mathscr{K}$ and so $B = B_2$. So B is unique. □

Proposition 3.2.14. *If E is a spectral measure for $(X, \mathscr{A}, \mathscr{H})$ and $\phi : X \to \mathbb{C}$ is a bounded $\mathscr{A}$-measurable function, then there is a unique operator $I(\phi) \in \mathscr{B}(\mathscr{H})$ such that if $\epsilon > 0$ and $\{A_1, \ldots, A_n\}$ is an $\mathscr{A}$-partition of X with $\sup\{|\phi(x) - \phi(x')|\ |x, x' \in A_k\} < \epsilon$ for $1 \leq k \leq n$, then for any $x_k \in A_k$ we have*

$$\|I(\phi) - \sum_{k=1}^{n} \phi(x_k)E(A_k)\| < \epsilon$$

Proof. Define $B(x, y) = \int \phi(z)d\langle E(z)x, y\rangle$ for $x, y \in \mathscr{H}$. By Lemma 3.2.11, we can see that B is a sesquilinear form with $|B(x, y)| \leq \|\phi\|_\infty\|x\|\|y\|$. Thus by Lemma 3.2.13 there exists a unique operator $I(\phi)$ such that $B(x, y) = \langle I(\phi)x, y\rangle$.

Next, let $\{A_1, \ldots, A_n\}$ be an $\mathscr{A}$-partition satisfying the requirements. If $y, z \in \mathscr{H}$ and $x_k \in A_k$ for $1 \leq k \leq n$ then

$$\begin{aligned} |\langle I(\phi)y, z\rangle - \sum_{k=1}^{n} \phi(x_k)\langle E(A_k)y, z\rangle| &= |\sum_{k=1}^{n} \int_{A_k} (\phi(x) - \phi(x_k))\, d\langle E(x)y, z\rangle| \\ &\leq \sum_{k=1}^{n} \int_{A_k} |\phi(x) - \phi(x_k)|\, d|\langle E(x)y, z\rangle| \\ &\leq \epsilon \sum_{k=1}^{n} \int_{A_k} d|\langle E(x)y, z\rangle| \leq \epsilon\|y\|\|z\| \end{aligned}$$

This is true for any $y, z \in \mathscr{H}$, thus we see that $\|I(\phi) - \sum_{k=1}^{n} \phi(x_k)E(A_k)\| < \epsilon$. □

The operator $I(\phi)$ obtained in the previous proposition, is called the *integral of ϕ with respect to the spectral measure E* and is denoted as $I(\phi) = \int_X \phi dE$. The given proof given of Proposition 3.2.14 also implies for $y, z \in \mathscr{H}$ and ϕ a bounded $\mathscr{A}$-measurable function, that

$$\langle\left(\int_X \phi dE\right) y, z\rangle = \int_X \phi(x)\ d\langle E(x)y, z\rangle = \int_X \phi(x)\ dE_{y,z}(x) \tag{11}$$

3.3 Spectral measure and representation of bounded measurable functions

Now that we introduced integrating with respect to a spectral measure, we prove that there is a 1-1 corespondence between spectral measures and representations of the set of bounded measurable functions. And as it turns out, the spectral theorem for bounded, normal operators is a consequence of this correspondence.

Definition 3.3.1. Let X and Y be two sets, and $\mathscr{A}$, $\mathscr{B}$ be two σ-algebras on X, Y respectively. If $f : X \to Y$ is a function, we say f is *measurable* if $f^{-1}(B) \in \mathscr{A}$ for every $B \in \mathscr{B}$. We call a function $f : X \to \mathbb{C}$ $\mathscr{A}$*-measurable* if f is measurable with respect to the spaces X and $\mathbb{C}$, equipped with $\mathscr{A}$ and the Borel σ-algebra respectively.

Definition 3.3.2. Let $\mathscr{A}$ be a σ-algebra defined on some set X. We define $B(X, \mathscr{A})$ as the set of *bounded* $\mathscr{A}$*-measurable* functions $f : X \to \mathbb{C}$.

Remark 3.3.3. If we equip $B(X, \mathscr{A})$ with the norm $\|f\|_\infty = \sup\{|f(x)|\ |x \in X\}$, then $B(X, \mathscr{A})$ becomes a Banach algebra. If we additionally define the mapping $f^*(x) := \overline{f(x)}$ then $B(X, \mathscr{A})$ becomes a C^*-algebra.

Note that the integral with respect to a spectral measure E is a mapping from $B(X, \mathscr{A})$ into $\mathscr{B}(\mathscr{H})$. This raises the question whether the integral over a spectral measure is a representation. And indeed, it is, as we will prove in the following theorem:

Theorem 3.3.4. *[1, Prop. IX.1.12, p. 258] If E is a spectral measure for $(X, \mathscr{A}, \mathscr{H})$ and $\rho : B(X, \mathscr{A}) \to \mathscr{B}(\mathscr{H})$ is defined as $\rho(\phi) = \int \phi dE$, then ρ is a representation of the C^*- algebra $B(X, \mathscr{A})$ and $\rho(\phi)$ is a bounded normal operator for every $\phi \in \mathscr{B}(X, \mathscr{A})$.*

Proof. First of all, we see that ρ is linear, for let $\phi, \psi \in B(X, \mathscr{A})$, then $\rho(\phi + \psi) = \int_X (\phi + \psi) dE$. Let $\lambda \in \mathbb{C}$ and $x, y \in \mathscr{H}$. Then

$$\begin{aligned}\langle \left(\int_X (\phi + \lambda\psi) dE\right) x, y\rangle &= \int_X (\phi(z) + \lambda\psi(z))\, d\langle E(z)x, y\rangle \\ &= \int_X \phi(z) d\langle E(z)x, y\rangle + \lambda \int_X \psi(z) d\langle E(z)x, y\rangle \\ &= \langle \rho(\phi)x, y\rangle + \langle \lambda\rho(\psi)x, y\rangle\end{aligned}$$

This is true for all $x, y \in \mathscr{H}$, and so we can conclude that ρ is linear.

Next, let $\phi \in B(X, \mathscr{A})$. Then for $x, y \in \mathscr{H}$ we find

$$\langle \rho(\phi)^* x, y\rangle = \langle x, \rho(\phi)y\rangle = \overline{\langle \rho(\phi)y, x\rangle} = \overline{\int_X \phi(z) d\langle E(z)y, x\rangle} = \int_X \overline{\phi(z)} d\langle E(z)x, y\rangle = \langle \rho(\overline{\phi})x, y\rangle$$

because E is a self-adjoint operator. This again is true for any $x, y \in \mathscr{H}$, so $\rho(\phi)^* = \rho(\overline{\phi})$.

Finally, we need to prove that ρ is multiplicative. Let $\phi, \psi \in B(X, \mathscr{A})$. Let $\epsilon > 0$ and choose an $\mathscr{A}$-partition $\{A_1, \dots, A_n\}$ of X such that $\sup\{|f(x) - f(x')|\ |x, x' \in A_k\} < \epsilon$ for $1 \leq k \leq n$ and $f = \phi, \psi$ or $\phi\psi$.

We claim that such a partition always exists. Because ϕ is bounded, we can find $r > 0$ such that $\phi(X) \subseteq \overline{B(0, r)} := \{z \in \mathbb{C} |\ |z| \leq r\}$. Consider the collection of sets $\{B(x, \epsilon) \mid x \in \phi(X)\}$.

This is an open cover of $\overline{B(0,r)}$. Because $\overline{B(0,r)}$ is compact, there exists a finite subcover of this collection. So there exists finitely many points $x_1, \dots, x_k$ such that $\overline{B(0,r)} \subseteq \cup_{j=1}^{k} B(x_j, \epsilon)$. By construction of this cover, it holds that if $\phi(y_1), \phi(y_2) \in B(x_j, \epsilon)$ then $|\phi(y_1) - \phi(y_2)| < \epsilon$. Define the set

$$B_j := B(x_j, \epsilon) \setminus \left(\bigcup_{i=1, i \neq j}^{k} B(x_i, \epsilon) \right)$$

Then we note that B_j is a Borel set, in other words B_j is an element of the Borel σ-algebra. Additionally, define the sets

$$C_{ij} = B(x_i, \epsilon) \cap B(x_j, \epsilon) \qquad i \neq j$$

Then C_{ij} is a Borel set as well, and so we see that the collection

$$\mathscr{C}_\phi := \{B_1, \dots, B_k, C_{11}, \dots C_{1k}, \dots, C_{(k-1)k}\}$$

is a partition of Borel sets of $\overline{B(0,r)}$. Rename all the sets such that $\mathscr{C}_\phi = \{U_1, \dots, U_m\}$ for some $m \in \mathbb{N}$. Because ϕ is $\mathscr{A}$-measurable, $\phi^{-1}(U_i) \in \mathscr{A}$ for $1 \leq i \leq m$. Then for any $x, y \in \phi^{-1}(U_i)$ we find $|\phi(x) - \phi(y)| < \epsilon$. Also note that the collection $\mathscr{B}_\phi := \{\phi^{-1}(U_1), \dots, \phi^{-1}(U_m)\}$ is an $\mathscr{A}$-partition of X. So we found our partition. We can do this for ψ and $\phi\psi$ as well, and choose the common refinement of the associated partitions $\mathscr{B}_\phi$, $\mathscr{B}_\psi$ and $\mathscr{B}_{\phi\psi}$.

Hence if $x \in A_k$ for $1 \leq k \leq n$, by Proposition 3.2.14:

$$\left\| \int_X f \, dE - \sum_{j=1}^{n} f(x_k) E(A_k) \right\| < \epsilon$$

for $f = \phi, \psi$ or $\phi\psi$. Thus, using the triangle inequality, we find:

$$\begin{aligned} \left\| \int_X \phi\psi dE - \left(\int_X \phi dE \right) \left(\int_X \psi dE \right) \right\| &\leq \left\| \int_X \phi\psi dE - \sum_{k=1}^{n} \phi(x_k)\psi(x_k)E(A_k) \right\| \\ &+ \left\| \sum_{k=1}^{n} \phi(x_k)\psi(x_k)E(A_k) - \left(\sum_{i=1}^{n} \phi(x_i)E(A_i) \right) \left(\sum_{j=1}^{n} \psi(x_j)E(A_j) \right) \right\| \\ &+ \left\| \left(\sum_{i=1}^{n} \phi(x_i)E(A_i) \right) \left(\sum_{j=1}^{n} \psi(x_j)E(A_j) \right) - \left(\int_X \phi \, dE \right) \left(\int_X \psi \, dE \right) \right\| \end{aligned}$$

Note that the first term on the right is smaller than ϵ. Additionally, $E(A_i)E(A_j) = E(A_i \cap A_j)$ and $\{A_1, \dots, A_n\}$ is a partition, so $A_i \cap A_j = \emptyset$. Therefore, the middle term is zero, and we are left with

$$\begin{aligned}
&\left\| \int_X \phi\psi dE - \left(\int_X \phi \, dE\right)\left(\int_X \psi \, dE\right) \right\| \\
&< \epsilon + \left\| \left(\sum_{i=1}^n \phi(x_i)E(A_i)\right)\left(\sum_{j=1}^n \psi(x_j)E(A_j)\right) - \left(\int_X \phi \, dE\right)\left(\int_X \psi \, dE\right) \right\| \\
&\leq \epsilon + \left\| \left(\sum_{i=1}^n \phi(x_i)E(A_i)\right)\left(\sum_{j=1}^n \psi(x_j)E(A_j) - \int_X \psi dE\right) \right\| \\
&\quad + \left\| \left(\sum_{i=1}^n \phi(x_i)E(A_i) - \int_X \phi dE\right)\left(\int_X \psi dE\right) \right\| \leq \epsilon(1 + \|\phi\| + \|\psi\|) \leq \epsilon M
\end{aligned}$$

Here $M = (1 + \|\phi\| + \|\psi\|)$. Since ϵ is arbitrary, we find that

$$\rho(\phi\psi) = \int \phi\psi dE = \left(\int \phi dE\right)\left(\int \psi dE\right) = \rho(\phi)\rho(\psi)$$

We conclude that ρ is a representation of $B(X, \mathscr{A})$. The fact that $\rho(\phi)$ is a normal operator immediately follows from earlier computations. □

Corollary 3.3.5. *If X is a compact Hausdorff space and E is a spectral measure defined on the Borel subsets of X, then $\rho : C(X) \to \mathscr{B}(\mathscr{H})$ defined as $\rho(u) = \int u dE$ is a representation of the C^*-algebra $C(X)$.*

So if we have a spectral measure and the set of continuous functions on X, we know that integration with respect to this spectral measure will give us a representation. Is the converse also true? If we start with a representation of the continuous functions, that this will give us a spectral measure? The answer is yes, and we will use the rest of this section to prove this theorem. But before we do this, we need a version of the Riesz Representation Theorem. We will not prove this version, for it will be a long proof that would not give us any insights in how spectral measures work.

Definition 3.3.6. Let X be a locally compact space equipped with the Borel σ-algebra $\mathscr{A}$. A *regular Borel measure* is a mapping $\mu : \mathscr{A} \to \mathbb{C}$ such that:

1. The mapping μ is a complex-valued measure.
2. Let $\epsilon > 0$ and let $A \in \mathscr{A}$. Then there exists a compact set $K \subseteq A$ and an open set $U \supseteq A$ such that $|\mu|(U \setminus K) < \epsilon$.

We write $M(X)$ for the space of all regular Borel measures on X.

We note that if we equip $M(X)$ with the norm $\|\mu\| = |\mu|(X)$ then $M(X)$ becomes a normed space.

Theorem 3.3.7. *(Variant of the Riesz-Representation Theorem)[1, Thm. C.18, p. 383] Let X be a locally compact Hausdorff space, and $\mu \in M(X)$. Define $F_\mu : C_0(X) \to \mathbb{C}$ by*

$$F_\mu(f) = \int f \, d\mu$$

Then $F_\mu \in C_0(X)'$ and the mapping $\mu \mapsto F_\mu$ is an isometric isomorphism of $M(X)$ onto $C_0(X)'$.

In this theorem, $C_0(X)$ is the space of continuous functions $f : X \to \mathbb{C}$ such that f vanishes at infinty. In other words, the space $K = \{x \in X \mid |f(x)| \geq \epsilon\}$ is compact for every $\epsilon > 0$. We also note that $C_0'(X)$ is the dual space of $C_0(X)$.

Definition 3.3.8. If V is a normed space, we define the *weak** *topology* on V' as the topology defined by the seminorms $\{p_x | x \in V\}$ where

$$p_x(x') = |x'(x)|$$

Here V' is the dual space of V

Lemma 3.3.9. *[1, Prop. V.4.1, p. 131], Prop.V.4.1, pg.131) If V is a normed space, then the unit ball in V is weak* dense in the unit ball in V''.*

For a proof, we refer to $[1, Prop.\ V.4.1, p.\ 131]$. Finally we can state, and prove, the other inclusion we already mentioned.

Theorem 3.3.10. *[1, Thm. IX.1.14, p. 259] Let X be a compact Hausdorff space. If $\rho : C(X) \to \mathscr{B}(\mathscr{H})$ is a representation, there exists a unique spectral measure E defined on the Borel subsets of X such that for all $x, y \in \mathscr{H}$ the measure $\langle E(\cdot)x, y\rangle = E_{x,y}$ is a regular measure and*

$$\rho(u) = \int u \, dE \tag{12}$$

for every $u \in C(X)$.

Proof. First we note that X is compact. Therefore $C(X) \equiv C_0(X)$. Let $x, y \in \mathscr{H}$. Then the mapping $u \mapsto \langle \rho(u)x, y\rangle$ is a bounded linear operator on $C(X)$ with norm $\leq \|x\|\|y\|$. Thus by Theorem 3.3.7, there exists a unique measure $\mu_{x,y} \in M(X)$ such that

$$\langle \rho(u)x, y\rangle = \int_X u \, d\mu_{x,y} \tag{13}$$

for all $u \in C_0(X)$. Because $\mu_{x,y}$ is uniquely defined, we see that the mapping $(x, y) \mapsto \mu_{x,y}$ is a sesquilinear mapping. Also $\|\mu_{x,y}\| \leq \|x\|\|y\|$. Let $\mathscr{A}$ be the Borel σ-algebra of X, and let $\phi \in B(X, \mathscr{A})$ be fixed. Define the map $[x, y] := \int \phi \, d\mu_{x,y}$. Then $[\cdot, \cdot]$ is a sesquilinear mapping and

$$|[x, y]| \leq \|\phi\|_\infty \|\mu_{x,y}\| \leq \|\phi\|_\infty \|x\|\|y\|$$

Hence by Lemma 3.2.13 there exists a unique operator A such that $[x, y] = \langle Ax, y\rangle$ with $\|A\| \leq \|\phi\|_\infty$. We write $A = \tilde{\rho}(\phi)$. Then $\tilde{\rho} : B(X, \mathscr{A}) \to \mathscr{B}(\mathscr{H})$ is a well-defined function with $\|\tilde{\rho}(\phi)\| \leq \|\phi\|_\infty$. Thus, by definition of $[\cdot, \cdot]$ we have for $x, y \in \mathscr{H}$

$$\langle \tilde{\rho}(\phi)x, y\rangle = \int_X \phi \, d\mu_{x,y} \tag{14}$$

Next, we want to prove that this $\tilde{\rho}$ is a representation of $B(X, \mathscr{A})$. First we note that $\tilde{\rho}|_{C(X)} = \rho$ by comparing Equations (13) and (14). Next, it is easy to see that $\tilde{\rho}$ is linear, by Equation (14). Next, if $\phi \in B(X, \mathscr{A})$, we can consider $\phi \in M(X)'(= C''(X))$, with the correspondence $\mu \mapsto \int \phi \, d\mu$. By Lemma 3.3.9, the set $\{u \in C(X) \mid \|u\| \leq \|\phi\|\}$ lies weak* dense in the set

$\{L \in (M(X))' \mid \|L\| \leq \|\phi\|\}$. Thus there exists a sequence $(u_i)_i$ in $C(X)$ such that $\|u_i\| \leq \|\phi\|$ for all u_i and $\int u_i \, d\mu \to \int \phi \, d\mu$ for every $\mu \in M(X)$. If $\psi \in B(X, \mathscr{A})$ and $\mu \in M(X)$ then $\psi\mu \in M(X)$. So $\int u_i\psi \, d\mu \to \int \phi\psi \, d\mu$ for any $\psi \in B(X, \mathscr{A})$ and $\mu \in M(X)$. Thus, by Lemma 3.2.5 we have $\int u_i\psi \, d\mu \to \int \phi\psi \, d\mu$ (WOT) for all $\psi \in B(X, \mathscr{A})$. In particular, for $\phi \in B(X, \mathscr{A})$ and $\psi \in C(X)$ we find

$$\begin{aligned}\tilde{\rho}(\phi\psi) &= \lim_{WOT, j\to\infty} \tilde{\rho}(u_i\psi) = \lim_{WOT, j\to\infty} \rho(u_i\psi) \\ &= \lim_{WOT, j\to\infty} \rho(u_i)\rho(\psi) = \lim_{WOT, j\to\infty} \tilde{\rho}(u_i)\rho(\psi) = \tilde{\rho}(\phi)\rho(\psi)\end{aligned}$$

Hence $\tilde{\rho}(u_i\psi) = \rho(u_i)\tilde{\rho}(\psi)$ for any $\psi \in B(X, \mathscr{A})$ and $u_i \in C(X)$. Since $\tilde{\rho}(u_i) \to \tilde{\rho}(\phi)$ (WOT) and $\tilde{\rho}(u_i\psi) \to \tilde{\rho}(\phi\psi)$ (WOT), it implies

$$\tilde{\rho}(\phi\psi) = \tilde{\rho}(\phi)\tilde{\rho}(\psi)$$

for $\phi, \psi \in B(X, \mathscr{A})$.

To prove that $\tilde{\rho}(\phi)^* = \tilde{\rho}(\overline{\phi})$, consider $\phi \in B(X, \mathscr{A})$ and let $(u_i)_i$ be the sequence obtained in previous paragraph. If $\mu \in M(X)$ we define the measure $\overline{\mu}$ as $\overline{\mu}(A) := \overline{\mu(A)}$. Because $\rho(u_i) \to \tilde{\rho}(\phi)$ (WOT) we find $\rho(u_i)^* \to \tilde{\rho}(\phi)^*$ (WOT). So

$$\int \overline{u_i} \, d\mu = \overline{\int u_i \, d\overline{\mu}} \to \overline{\int \phi d\overline{\mu}} = \int \overline{\phi} d\mu$$

is true for any measure μ. Therefore we can conclude that $\rho(\overline{u_i}) \to \tilde{\rho}(\overline{\phi})$. We know that ρ is a representation, so $\rho(\overline{u_i}) = \rho(u_i)^*$. Hence $\rho(u_i)^* = \rho(\overline{u_i}) \to \tilde{\rho}(\overline{\phi})$. Because the weak operator topology is Hausdorff, limits are unique. So $\tilde{\rho}(\phi)^* = \tilde{\rho}(\overline{\phi})$. So $\tilde{\rho}$ is a representation.

For any Borel subset A of X define the operator $E(A) := \tilde{\rho}(\chi_A)$, where χ_A is the characteristic function on A. We show that E is a spectral measure. Let A be some Borel subset of X. Then $E(A)^2 = \tilde{\rho}(\chi_A)\tilde{\rho}(\chi_A) = \tilde{\rho}(\chi_A) = E(A)$. So $E(A)$ is a idempotent on $\mathscr{H}$. Additionally, $E(A)^* = \tilde{\rho}(\chi_A)^* = \tilde{\rho}(\chi_A) = E(A)$, so $E(A)$ is self-adjoint. So by Lemma 3.2.3 $E(A)$ is a projection for A a Borel subset of X.

Additionally, because $\chi_\emptyset = 0$ and $\chi_X = I$ we have $E(\emptyset) = 0$ and $E(X) = I$. Next let A, B be Borel subsets. Then $E(A \cap B) = \tilde{\rho}(\chi_{A\cap B}) = \tilde{\rho}(\chi_A\chi_B) = E(A)E(B)$. And finally, let $(A_i)_i$ be a pairwise disjoint sequence of Borel sets and set $\Lambda_n = \bigcup_{k=n+1}^{\infty} A_k$. Because $\tilde{\rho}$ is a representation, we use induction to see that E is finitely additive. Thus, for $x \in \mathscr{H}$, we find

$$\begin{aligned}\left\| E(\bigcup_{k=1}^{\infty} A_k)x - \sum_{k=1}^{n} E(A_k)x \right\|^2 &= \langle E(\Lambda_n)x, E(\Lambda_n)x \rangle \\ &= \langle E(\Lambda_n)x, x \rangle \\ &= \langle \tilde{\rho}(\chi_{\Lambda_n})x, x \rangle \\ &= \int \chi_{\Lambda_n} \, d\mu_{x,x} \\ &= \sum_{k=n+1}^{\infty} \mu_{x,x}(\Lambda_n)\end{aligned}$$

The last sum clearly goes to 0 as $n \to \infty$. Therefore, E is a spectral measure.

Now that we have constructed a spectral measure, we need to prove $\rho(u) = \int u\ dE$. If $\tilde{\rho}(\phi) = \int \phi\ dE$ for any $\phi \in B(X, \mathscr{A})$, then surely it is true for $u \in C(X)$. Fix $\phi \in B(X, \mathscr{A})$ and $\epsilon > 0$. If $\{A_1, \ldots, A_n\}$ is any Borel partition of X such that $\sup\{|\phi(x) - \phi(x')|\ |x, x' \in A_k\} < \epsilon$ for $1 \leq k \leq n$. Then we find $\|\phi - \sum_{k=1}^n \phi(x_k)\chi_{A_k}\|_\infty < \epsilon$ for any choice of $x \in A_k$. Because $\|\tilde{\rho}\| = 1$, we have $\|\tilde{\rho}(\phi) - \sum_{k=1}^n \phi(x_k)E(A_k)\| < \epsilon$. We use Proposition 3.2.14 to conclude $\tilde{\rho}(\phi) = \int \phi\ dE$ for any $\phi \in B(X, \mathscr{A})$.

The only thing left, is to show that E is the unique measure such that $\rho(u) = \int u\ dE$ for $\phi \in C(X)$. First we proof the uniqueness of E for $\tilde{\rho}$. Assume there exists another spectral measure F such that $\tilde{\rho}(\phi) = \int \phi\ dF$. Then $\int \phi\ dE = \int \phi\ dF$. Let A be some Borel subset. Then $E(A) = \int \chi_A\ dE = \int \chi_A\ dF = F(A)$. So $E(A) = F(A)$ for any Borel subset, so $E = F$.

Now consider any spectral measure G such that $\rho(u) = \int u\ dG$ for $u \in C(X)$. Consider a Borel set A. By previous paragraph, there exists a sequence $(u_i)_i$ in $C(X)$ such that $\int u_i\ d\mu \to \int \chi_A\ d\mu$ for every $\mu \in M(X)$. So surely for the measures $E_{x,y}$ and $G_{x,y}$ for $x, y \in \mathscr{H}$. This results in

$$\begin{aligned}\langle E(A)x, y\rangle &= \langle \left(\int_X \chi_A dE\right) x, y\rangle = \int_X \chi_A(z)\ d\langle E(z)x, y\rangle = \lim_{i\to\infty} \int_X u_i(z)\ d\langle E(z)x, y\rangle \\ &= \lim_{i\to\infty} \int_X u_i\ d\langle G(z)x, y\rangle = \int_X \chi_A\ d\langle G(z)x, y\rangle = \langle \left(\int_X \chi_A\ dG\right) x, y\rangle = \langle G(A)x, y\rangle\end{aligned}$$

This is true for any $x, y \in \mathscr{H}$ and thus $E(A) = G(A)$. So it is uniquely determined. □

3.4 The spectral theorem for bounded normal operators

With Theorem 3.3.10 we can give the proof of the spectral theorem in the bounded case. We consider a specific C^*-algebra with a specific representation and apply the previous theorem, and we find the spectral theorem.

Before we prove the spectral theorem, we give some theorems regarding C^*-algebras. These theorems are needed for the proof of the spectral theorem. We will not prove these theorems however, because they do not give much insight in the proof of the spectral theorem itself.

Definition 3.4.1. If $\mathscr{A}$ be a C^*-algebra over $\mathbb{K}$ and $a \in \mathscr{A}$, we define the C^**-algebra generated by* a as

$$C^*(a) := \overline{\{p(a, a^*)|p(z, \overline{z}) \text{ is a polynomial}\}}$$

Example 3.4.2. Let $\mathscr{A}$ be a C^*-algebra, and let $a \in \mathscr{A}$ a normal element (in other words, $aa^* = a^*a$). Then $C^*(a)$ is an abelian C^*-algebra. For example, if $\mathscr{A} = \mathscr{B}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$, and $N \in \mathscr{B}(\mathscr{H})$ is a normal operator, then $C^*(N)$ is an abelian C^*-algebra. ⊘

Definition 3.4.3. Let $\mathscr{A}$ be a C^*-algebra over $\mathbb{K}$ with identity and $a \in \mathscr{A}$. We say a is invertible, if there exists $x, y \in \mathscr{A}$ such that $xa = 1 = ay$.

We also define the *spectrum of* a, denoted $\sigma_{\mathscr{A}}(a)$, as

$$\sigma_{\mathscr{A}}(a) := \{\lambda \in \mathbb{K}|(a - \lambda 1) \text{ is not invertible in } \mathscr{A}\}$$

Proposition 3.4.4. *[1, Prop. VIII.1.4, p. 235] Let* $\mathscr{A}$ *and* $\mathscr{B}$ *be* C^**-algebras over* $\mathbb{K}$ *with a common identity and norm such that* $\mathscr{B} \subseteq \mathscr{A}$*. If* $a \in \mathscr{B}$ *then* $\sigma_{\mathscr{A}}(a) = \sigma_{\mathscr{B}}(a)$.

Theorem 3.4.5. *(The Stone-Weierstraß Theorem)[1, Thm. V.8.1, p. 145] If X is compact and $\mathscr{A}$ is a closed subalgebra of $C(X)$ such that*

1. $\mathrm{Id} \in \mathscr{A}$,
2. *If $x, y \in X$ and $x \neq y$, then there is an $f \in \mathscr{A}$ such that $f(x) \neq f(y)$,*
3. *If $f \in \mathscr{A}$ then $\overline{f} \in \mathscr{A}$,*

then $\mathscr{A} = C(X)$.

We also need another tool called the functional calculus. In order to introduce the functional calculus, we introduce the notion of a maximal ideal space.

Definition 3.4.6. Let $\mathscr{A}$ be an abelian C^*-algebra over $\mathbb{K}$. Let

$$\Sigma = \{h : \mathscr{A} \to \mathbb{C} \mid h \text{ is a non-zero homomorphism}\}$$

If we equip Σ with the smallest topology such that for every $b \in \mathscr{A}$ the mapping $\mathrm{ev}_b : \Sigma \to \mathbb{C}$, $h \mapsto h(b)$ is continuous, we call Σ the *maximal ideal space*.

Maximal ideal spaces have some interesting properties. One of these properties is the following theorem.

Theorem 3.4.7. *[1, Thm. VIII.2.1, p. 236] Let $\mathscr{A}$ is a C^*-algebra, and let $a \in \mathscr{A}$ be a normal element (so $aa^* = a^*a$). Consider $C^*(a)$, and Σ the maximal ideal space of $C^*(a)$. Then the mapping $\gamma : C^*(a) \to C(\Sigma)$ given by $x \mapsto \mathrm{ev}_x$ is an isometric *-isomorphism.*

So we see that $C^*(a)$ is isomorphic to $C(\Sigma)$. Additionally, the have another theorem which tells us that we can construct a homeomorphism between Σ and $\sigma_{C^*(a)}(a)$.

Proposition 3.4.8. *[1, Prop. VIII.2.3, p. 237] Let $\mathscr{A}$ be a C^*-algebra, and let $a \in \mathscr{A}$ be a normal element. If $\mathscr{B} := C^*(a)$, and Σ the maximal ideal space of $C^*(a)$, then the map $\mathrm{ev}_a : \Sigma \to \sigma_{\mathscr{B}}(a)$ given by $\mathrm{ev}_a(h) = h(a)$ is a homeomorphism. Additionally, if $p(z, \overline{z})$ is a polynomial in z and $\overline{z}$ and $\gamma : \mathscr{B} \to C(\Sigma)$ is as in Theorem 3.4.7, then $\gamma(p(a, a^*)) = p \circ \mathrm{ev}_a$.*

Now, we note that if $\mathrm{ev}_a : \Sigma \to \sigma_{\mathscr{A}}(a)$ is defined as in Proposition 3.4.8, we can define $(\mathrm{ev}_a)^* : C(\sigma_{\mathscr{A}}(a)) \to C(\Sigma)$ by $(\mathrm{ev}_a)^*(f) = f \circ \mathrm{ev}_a$. Then this is an isometric *-isomorphism because ev_a is a homeomorphism. Therefore, by the last part of previous proposition, $\gamma(p(a, a^*)) = (\mathrm{ev}_a)^*(p)$. Of course, any polynomial is just a function on $\sigma(a)$, and therefore we can define the mapping $\rho : C(\sigma_{\mathscr{A}}(a)) \to C^*(a)$ by $\rho = \gamma^{-1} \circ (\mathrm{ev}_a)^*$, such that the following diagram commutes.

$$\begin{array}{ccc} C(\sigma_{\mathscr{A}}(a)) & \xrightarrow{\gamma} & C(\Sigma) \\ \rho\big\downarrow & \nearrow_{(\mathrm{ev}_a)^*} & \\ C^*(a) & & \end{array}$$

Note that if $p \in C(\sigma_{C^*(a)}(a))$ is a polynomial in z and $\overline{z}$, then $\rho(p(z, \overline{z})) = p(a, a^*)$. In particular, $\rho(z) = a$.

Definition 3.4.9. Let $\mathscr{A}$ be a C^*-algebra over $\mathbb{K}$ with identity, and let $a \in \mathscr{A}$ be a normal element. Define $\rho : C(\sigma(a) \to C^*(a)$ as in the previous diagram. If $f \in C(\sigma(a))$ we define

$$f(a) := \rho(f)$$

The mapping $f \mapsto f(a)$ of $C(\sigma_{\mathscr{A}}(a)) \to \mathscr{A}$ is called the *functional calculus for a*.

Theorem 3.4.10. *[1, Thm. VIII.2.6, p. 238] Let ρ be the functional calculus of a. Then $\rho : C(\sigma_{C^*(a)}(a)) \to C^*(a)$ is an isometric *-isomorphism.*

Finally, we can state and prove the spectral theorem for bounded, normal operators.

Theorem 3.4.11 (The spectral theorem for bounded, normal operators)**.** *[1, Thm IX.2.2, p. 263] Let $\mathscr{H}$ be a Hilbert space, and $N \in \mathscr{B}(\mathscr{H})$ be a normal operator. Then there exists a unique regular spectral measure E on the Borel subsets of $\sigma(N)$ (here $\sigma(N)$ is defined as in Definition 2.3.3) such that:*

1. *$N = \int_{\sigma(N)} z \, dE(z)$.*
2. *If A is a nonempty relatively open subset of $\sigma(N)$, $E(A) \neq 0$.*
3. *If $T \in \mathscr{B}(\mathscr{H})$, then $TN = NT$ and $TN^* = N^*T$ if and only if $TE(A) = E(A)T$ for every Borel subset A of $\sigma(N)$.*

Proof. 1): We consider the set $\mathscr{A} = C^*(N)$, the C^*-algebra generated by N. In other words, $\mathscr{A}$ is the closure of the set of all polynomials in N and N^*. Then, by Theorem 3.4.9 there exists an isometric *-isomorphism $\rho : C(\sigma_{\mathscr{A}}(N)) \to \mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ 'given by $\rho(u) = u(N)$. Here $\sigma(N) = \{\lambda \in \mathbb{C} | (N - \lambda)$is not invertible in $C^*(N)\}$. So this is a representation of $C(\sigma_{\mathscr{A}}(N))$. Because $I \in C^*(N)$ and $I \in \mathscr{B}(\mathscr{H})$, we use Proposition 3.4.4 to conclude that $\sigma_{\mathscr{A}}(N) = \sigma(N)$ where $\sigma(N)$ is the spectrum of N as defined in Definition 2.3.3. So $\rho : C(\sigma(N)) \to \mathscr{A}$ is a representation.

Since N is bounded, $\sigma(N)$ is compact, and so we conclude by Theorem 3.3.10, that there exists a unique regular spectral measure E defined on the Borel subsets of $\sigma(N)$ such that $\rho(u) = \int_{\sigma(N)} u \, dE$ for all $u \in C(\sigma(N))$. In particular, it holds for $u(z) = z$. Then

$$N = \int_{\sigma(N)} z \, dE(z)$$

.

2): Next, let A be some nonempty relatively open subset of $\sigma(N)$. Then we can find a nonzero continuous function u on $\sigma(N)$ such that $0 \leq u \leq \chi_A$. It follows then that $\|E(A)\| = \|\tilde{\rho}(\chi_A)\| \geq \|\tilde{\rho}(u)\| > 0$. The last inequality is because ρ is a *-isomorphism. So $E(A) \neq 0$.

3) $\Rightarrow$: Let $T \in \mathscr{B}(\mathscr{H})$ such that $TN = NT$ and $TN^* = N^*T$. Consider the set

$$\mathscr{C} := \{u \in C(\sigma(N)) \, | T\rho(u) = \rho(u)T\}$$

Then it easily follows that $Tp(N, N^*) = p(N, N^*)T$ for any polynomial $p \in \mathbb{C}[z, \overline{z}]$. So the set of polynomials is a subset of $\mathscr{C}$. Because the polynomials are point-separating, we find by

Theorem 3.4.5 that $\overline{\{p(z,\overline{z})\}} = C(\sigma(N))$. But we just concluded $\{p(z,\overline{z})\} \subseteq \mathscr{C}$, so $C(\sigma(N)) = \overline{\{p(z,\overline{z})\}} \subseteq \mathscr{C} \subseteq C(\sigma(N))$. Therefore $\mathscr{C} = C(\sigma(N))$. So we conclude that $T\rho(u) = \rho(u)T$ for any $u \in C(\sigma(N))$. In other words $Tu(N) = u(N)T$.

Define the following set

$$\mathscr{D} := \{A \subseteq \sigma(N) |\ A \text{ is a Borel set and } TE(A) = E(A)T\}$$

It follows quite easily from the definition of a spectral measure that $\mathscr{A}$ is a σ-algebra.

If G is an open set in $\sigma(N)$, then there exists a sequence $(u_n)_n$ of positive continuous functions on $\sigma(N)$ such that $u_n(z) \uparrow \chi_G(z)$ as $n \to \infty$ pointwise for all $z \in \sigma(N)$. This means that for $x, y \in \mathscr{H}$:

$$\begin{aligned}
\langle TE(G)x, y\rangle = \langle E(G)x, T^*y\rangle &= \int_{\sigma(N)} \chi_G(z)\ d\langle E(z)x, T^*y\rangle \\
&= \lim_{n\to\infty} \int_{\sigma(N)} u_n(z)\ d\langle E(z)x, T^*y\rangle \\
&= \lim_{n\to\infty} \langle u_n(N)x, T^*y\rangle = \lim_{n\to\infty} \langle Tu_n(N)x, y\rangle = \lim_{n\to\infty} \langle u_n(N)Tx, y\rangle \\
&= \lim_{n\to\infty} \int_{\sigma(N)} u_n(z)\ d\langle E(z)Tx, y\rangle = \langle E(G)Tx, y\rangle
\end{aligned}$$

Here we used the Lebesgue Dominated Convergence Theorem, to move the limit out of the integral. This is true for any $x, y \in \mathscr{H}$, and so $G \in \mathscr{D}$. But this means that every open subset is in $\mathscr{D}$, and thus $\mathscr{A}$ is the Borel σ-algebra. Therefore if A is a Borel subset, $E(A)T = TE(A)$.

3) $\Leftarrow$: Assume that for every Borel subset A of $\sigma(N)$, we have $E(A)T = TE(A)$. Then we have $\langle E(A)Tx, y\rangle = \langle TE(A)x, y\rangle = \langle E(A)x, T^*y\rangle$. In other words we have that

$$\langle NTx, y\rangle = \int_{\sigma(N)} z\ d\langle E(z)Tx, y\rangle = \int_{\sigma(N)} z\ d\langle E(z)x, T^*y\rangle = \langle Nx, T^*y\rangle = \langle TNx, y\rangle$$

This is true for any $x, y \in \mathscr{H}$, so $NT = TN$. We can do the same trick to see $N^*T = TN^*$. This concludes part 3. □

The spectral measure E obtained by the spectral theorem, is called the *spectral measure for* N. It is also common to say: Let $N = \int_{\sigma(N)} \lambda\ dE(\lambda)$ be the *spectral decomposition* of N. We indirectly use Theorem 3.4.11 to find this E.

Remark 3.4.12. If we look at Theorem 3.4.11, it seems like this theorem tells us that N can be 'diagonalized', in a sense that N is an 'infinite dimensional matrix' with the 'eigenvalues' on its diagonal. Intuitively it feels correct, in the finite dimensional case the spectral theorem reduces to this exact case.

Example 3.4.13. Consider a finite dimensional Hilbert space $\mathscr{H}$, say $\dim(\mathscr{H}) = n$. Let $N \in \mathscr{B}(\mathscr{H})$ be a normal operator, so with respect to some orthonormal basis is N some matrix with complex coefficients such that $N^*N = NN^*$. Because $\mathscr{H}$ is of dimension n, we have that $|\sigma(N)| \leq n$. Therefore, we can enumerate the eigenvalues of A, say $\sigma(N) = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$

with $k := |\sigma(N)| \leq n$. We can use Theorem 3.4.11 to gain the spectral measure E of N. Therefore, we find that for any $x, y \in \mathscr{H}$:

$$\langle Nx, y\rangle = \int_{\sigma(N)} \lambda \, d\langle E(\lambda)x, y\rangle = \sum_{i=1}^{k} \lambda_i \, \langle E(\{\lambda_i\})x, y\rangle = \langle \sum_{i=1}^{k} \lambda_i E(\{\lambda_i\})x, y\rangle \qquad (x, y \in \mathscr{H})$$

Therefore $N = \sum_{i=1}^{k} \lambda_i E(\{\lambda_i\})$, with $E(\{\lambda_i\})$ some projection operator.

But what is $E(\{\lambda_i\})$ exactly? We know it is a projection operator. Since N is normal, we have $NN^* = N^*N$, and so by part 3 of Theorem 3.4.11, $E(\{\lambda_i\})N = NE(\{\lambda_i\})$. Next, consider $x \in \mathscr{H}$. Then

$$NE(\{\lambda_i\})x = E(\{\lambda_i\})Nx = E(\{\lambda_i\}) \left(\sum_{j=1}^{k} \lambda_j E(\{\lambda_j\}) \right) x = E(\{\lambda_i\})\lambda_i E(\{\lambda_i\})x = \lambda_i E(\{\lambda_i\})x$$

because E is a spectral measure. So we see that $\operatorname{ran}(E(\{\lambda_i\})) \subseteq E_{\lambda_i}$, where E_{λ_i} is the eigenspace of the eigenvalue λ_i. On the other hand, if $x \in E_{\lambda_i}$, then $Nx = \lambda_i x$. Thus for $1 \leq j \leq n$

$$\begin{aligned} E(\{\lambda_j\})Nx &= E(\{\lambda_j\}) \left(\sum_{i=1}^{k} \lambda_i E(\{\lambda_i\}) \right) x = \lambda_j E(\{\lambda_j\})x \\ &= \lambda_i E(\{\lambda_j\})x \end{aligned}$$

So if $\lambda_j \neq \lambda_i$, then we must have $E(\{\lambda_i\})x = 0$. If $\lambda_j = \lambda_i$, then we see that $Nx = \lambda_i E(\{\lambda_i\})x = \lambda_i x$, so $x \in \operatorname{ran}(E(\{\lambda_i\}))$. Therefore we see that $E(\lambda_i)$ is the projection onto the eigenspace E_{λ_i}.

So in short, we see that N is diagonalized into projections onto the eigenspace of each eigenvectors. We knew already that this was possible from linear algebra, but it is good to see that the spectral theorem gives the same result as we would have expect it would give. ⊘

The next theorem and corollary seem innocent, yet are important. We will consider the consequences of this corollary in Section 5.

Theorem 3.4.14 (Theorem of Fuglede)**.** *[3, p. 35] Let $A, B \in \mathscr{B}(\mathscr{H})$ be two normal operators. If $AB = BA$ then $BA^* = A^*B$.*

Corollary 3.4.15. *(Corollary of Theorem 3.4.11) Let $A, B \in \mathscr{B}(\mathscr{H})$ be two normal operators, and let E_A, E_B be the spectral measures for A, B respectively. If $AB = BA$, then for any two Borel sets U, V we have $E_A(U)E_B(V) = E_B(V)E_A(U)$.*

Proof. Let $A, B \in \mathscr{B}(\mathscr{H})$. By previous Theorem, we see that if $AB = BA$, then $A^*B = BA^*$ and $AB^* = B^*A$. Also $A^*B^* = B^*A^*$. Because $AB = BA$ and $A^*B = BA^*$, we use part 3) of Theorem 3.4.11 to conclude that for any Borel set U we have $E_A(U)B = BE_A(U)$. Additionally, $E_A(U)$ is a projection, thus self-adjoint. So

$$E_A(U)B^* = E_A(U)^*B^* = (BE_A(U))^* = (E_A(U)B)^* = B^*E_A(U)^* = B^*E_A(U)$$

By the same theorem, we see that for any Borel set V, we have $E_A(U)E_B(V) = E_B(V)E_A(U)$. □

4 Spectral Theorem for unbounded self-adjoint operators

We return to general operators. For self-adjoint operators, a similar spectral theorem exists for unbounded operators. We will prove this in this section. We start with a useful tool to prove the spectral theorem for self-adjoint operators: the Cayley transform. After that, we will prove the spectral theorem for unbounded, self-adjoint operators. The reader who is already familiar with the spectral theorem for self-adjoint operators, can read on to the theorem of Nelson which will be discussed in Section 5.

4.1 The Cayley transform

To introduce the Cayley transform, we define the function $f : \mathbb{R} \to \mathbb{C}$ by

$$f(z) = \frac{z-i}{z+i} \tag{15}$$

It is clear $f(0) = -1$ and $f(1) = -i$. In fact, $f(\mathbb{R}) = \mathbb{S} \setminus \{1\}$ where $\mathbb{S} := \{z \in \mathbb{C} |\ |z| = 1\}$. The inverse of f is given by $g : \mathbb{S} \setminus \{1\} \to \mathbb{R}$

$$g(z) = i\frac{z+1}{z-1} \tag{16}$$

So, if A is a self-adjoint operator we have $\sigma(A) \subseteq \mathbb{R}$, so $f(\sigma(A)) \subseteq \mathbb{S} \setminus \{1\}$. We know that a bounded, unitary operator U has the property that $\sigma(U) \subseteq \mathbb{S}$. So this raises the question, is $f(A)$ a bounded, unitary operator? The answer is yes, and $f(A)$ is called the Cayley transform.

Theorem 4.1.1 (The Cayley transform)**.** *[1, Thm. X.3.1, p. 317]*

1. *If A is a closed densely defined symmetric operator with deficiency subspaces $\mathscr{L}_\pm$, and if $U : \mathscr{H} \to \mathscr{H}$ is defined by letting $U = 0$ on $\mathscr{L}_+$ and*

$$U := (A-i)(A+i)^{-1} \tag{17}$$

on $\mathscr{L}_+^\perp$, then U is a partial isometry with initial space $\mathscr{L}_+^\perp$, final space $\mathscr{L}_-^\perp$ and such that $(I-U)(\mathscr{L}_+^\perp)$ is dense in $\mathscr{H}$

2. *If U is a partial isometry with initial space $\mathscr{M}$ and final space $\mathscr{N}$, and such that $(I-U)(\mathscr{M})$ is dense in $\mathscr{H}$, then*

$$A := i(I+U)(I-U)^{-1} \tag{18}$$

is a densely defined closed symmetric operator with deficiency subspaces $\mathscr{L}_+ = \mathscr{M}^\perp$ and $\mathscr{L}_-^\perp = \mathscr{N}^\perp$.

3. *If A is given as in* 1) *and U is defined by Equation* (17)*, then A and U satisfy Equation* (18)*. If U is given as in* 2) *and A is defined by Equation* (18)*, then A and U satisfy Equation* (17)*.*

Proof. 1): Let A as in the theorem. By Lemma 2.4.10 part 3), we have $\operatorname{ran}(A \pm i)$ is closed. Hence $\mathscr{L}_\pm^\perp = \operatorname{ran}(A \pm i)$. By the same Lemma, we also have $\ker(A+i) = (0)$, and therefore $(A+i)^{-1}$ is well defined on $\mathscr{L}_+^\perp$. Note that $(A+i)^{-1}(\mathscr{L}_+) \subseteq \operatorname{Dom}(A)$. In this way, the operator in Equation (17) makes sense and is well-defined. If $y \in \mathscr{L}_+^\perp$, then there exists a unique $x \in \operatorname{Dom}(A)$ such that $y = (A+i)x$. Then we get, by Lemma 2.4.10:

$$\|Uy\|^2 = \|(A-i)x\|^2 = \|Ax\|^2 + \|x\|^2 = \|(A+i)x\|^2 = \|y\|^2$$

Therefore we conclude that U is a partial isometry, with initial space $\mathscr{L}_+^\perp$ and final space

$$\operatorname{ran}(U) = \operatorname{ran}(A-i) = \mathscr{L}_-^\perp$$

The only thing we need to show, is that $(I-U)(\mathscr{L}_+^\perp)$ is dense in $\mathscr{H}$. If $y \in \mathscr{L}_+^\perp$, there exists a unique $x \in \operatorname{Dom}(A)$ such that $y = (A+i)x$. Hence

$$(I-U)y = y - (A-i)x = (A+i)x - (A-i)x = 2ix$$

So $(I-U)(\mathscr{L}_+^\perp) = \operatorname{Dom}(A)$. Because A is densely defined, $(I-U)(\mathscr{L}_+^\perp)$ is dense in $\mathscr{H}$.

2): Now assume U is the partial isometry as in the theorem. Let $x \in \ker(I-U)$, then $Ux = x$ and so $\|Ux\| = \|x\|$. Hence $x \in \mathscr{M} = (\ker(U))^\perp$. It is easy to see that U^*U is a projection upon the initial space of U, in this case $\mathscr{M}$. Therefore, we see that

$$x = U^*Ux = U^*x$$

Therefore we have $x \in \ker(I-U^*)$. So we see $x \in \ker(I-U^*) = [\operatorname{ran}(I-U)]^\perp \subseteq [(I-U)(\mathscr{M})]^\perp = (0)$ by assumption. Therefore $x = 0$ and $I-U$ is injective.

Define $\mathscr{D} := (I-U)(\mathscr{M})$. Then $I-U$ is bijective on $\mathscr{D}$ and so $(I-U)^{-1}$ can be defined on $\mathscr{D}$. Because $I-U$ is bounded on $\mathscr{M}$, we have $\operatorname{gra}((I-U)|_{\mathscr{M}})$ is closed and so $\operatorname{gra}((I-U)^{-1})$ is closed. Hence, if we define the linear operator $A := i(I+U)(I-U)^{-1}$ we find A is closed and densely defined, with domain $\mathscr{D}$.

To prove that A is symmetric, consider $\phi, \psi \in \mathscr{D}$. Then there exist $x, y \in \mathscr{M}$ such that $\phi = (I-U)x$ and $\psi = (I-U)y$. Hence we find

$$\langle A\phi, \psi\rangle = \langle i(I+U)x, (I-U)y\rangle = i\left(\langle x,y\rangle + \langle Ux,y\rangle - \langle x,Uy\rangle - \langle Ux,Uy\rangle\right)$$

Next we note that $x, y \in \mathscr{M}$, so $\langle Ux, Uy\rangle = \langle x,y\rangle$. So we find $\langle A\phi,\psi\rangle = i(\langle Ux,y\rangle - \langle x,Uy\rangle)$. If we consider $\langle \phi, A\psi\rangle$ we get

$$\langle \phi, A\psi\rangle = \langle (I-U)x, i(I+U)y\rangle = -i\left(\langle x,y\rangle - \langle Ux,y\rangle + \langle x,Uy\rangle - \langle Ux,Uy\rangle\right) = \langle A\phi,\psi\rangle$$

So A is symmetric.

Finally, we need to prove that $\mathscr{L}_+ = \mathscr{M}^\perp$ and $\mathscr{L}_- = \mathscr{N}^\perp$. It is sufficient to show $\mathscr{M} = \operatorname{ran}(A+i)$ and $\mathscr{N} = \operatorname{ran}(A-i)$ because A is closed and symmetric. Consider $x \in \mathscr{M}$, and define $y = (I-U)x$. Then $(A+i)y = Ay + iy = i(I+U)x + i(I-U)x = 2ix$. Therefore $x \in \operatorname{ran}(A+i)$. On the other hand if $\phi \in \operatorname{ran}(A+i)$, then there exists $\psi \in \mathscr{D}$ such that $\phi = (A+i)\psi$. Since $\psi \in \mathscr{D}$ there exists a $\varphi \in \mathscr{M}$ such that $\psi = (I-U)\varphi$. This results in

$\phi = A\psi + i\psi = i(I+U)\varphi + i(I-U)\varphi = 2i\varphi$. So $\phi \in \mathscr{M}$. Thus $\mathscr{M} = \operatorname{ran}(A+i)$ By changing the plus into a minus at $A+i$, we see that $\operatorname{ran}(A-i) = \operatorname{ran}(U) = \mathscr{N}$.

3): If A is given as in 1) and U as in Equation (17). Let $y \in (I-U)(\mathscr{L}_+^\perp)$. It is equivalent to say that there exists $x \in \mathscr{L}_+^\perp$ such that $y = (I-U)x$. Because $x \in \mathscr{L}_+^\perp = \operatorname{ran}(A+i)$ it is equivalent to say that there exists $z \in \operatorname{Dom}(A)$ such that $x = (A+i)z$. Therefore, we find

$$y = x - Ux = (A+i)z - (A-i)z = 2iz$$

Hence $z = \frac{1}{2i}y$. Therefore $\operatorname{Dom}(A) = \operatorname{Dom}(i(I+U)(I-U)^{-1})$. Next we see that

$$\begin{aligned} i(I+U)(I-U)^{-1}y &= i(I+U)x = ix + iUx \\ &= i(A+i)z + i(A-i)z = 2iAz = Ay \end{aligned}$$

So $A = i(I+U)(I-U)^{-1}$.

Now assume U is given as in 2) and A as in Equation (18). If $\phi \in \operatorname{ran}(A+i)$ there exists a $\psi \in \operatorname{Dom}(A)$ such that $\phi = (A+i)\psi$. Since $\operatorname{Dom}(A) = (I-U)(\mathscr{M})$, it is equivalent to say that there exists $\varphi \in \mathscr{M}$ such that $\psi = (I-U)\varphi$. Therefore

$$\phi = (A+i)\psi = A\psi + i\psi = i(I+U)\varphi + i(I-U)\varphi = 2i\varphi$$

Thus we see $\phi = 2i\varphi$. So we find for $\phi \in \operatorname{ran}(A+i)$:

$$(A-i)(A+i)^{-1}\phi = (A+i)\psi = A\psi + i\psi = (i(I+U)\varphi) + i(I-U)\varphi = 2iU\varphi = U\phi$$

Hence $U = (A-i)(A+i)^{-1}$. □

Definition 4.1.2. If A is a densely defined closed symmetric operator, the partial isometry U defined in Theorem 4.1.1 is called the *Cayley transform* of A.

Corollary 4.1.3. *If A is a self-adjoint operator and U is its Cayley transform, then U is a unitary operator with* $\ker(I-U) = (0)$ *and* $(I-U)(\mathscr{H})$ *lies dense in* $\mathscr{H}$. *Conversely, if U is a unitary operator such that* $1 \notin \sigma_p(U)$ *and* $(I-U)(\mathscr{H})$ *lies dense in* $\mathscr{H}$, *then the operator A defined by Theorem* 4.1.1 *is self-adjoint.*

Proof. Note that A is self-adjoint if and only if $\mathscr{L}_\pm = 0$. A partial isometry is a unitary operator if and only if its initial space and its final space are $\mathscr{H}$. Therefore if A is self-adjoint, then U is a unitary operator such that $(I-U)(\mathscr{H})$ lies dense in $\mathscr{H}$. By the first part of the proof of part 2) of Theorem 4.1.1 we find that $\ker(I-U) = (0)$.

If U is a unitary operator such that $\ker(I-U) = (0)$ and $(I-U)(\mathscr{H})$ lies dense, we can immediately use the second part of Theorem 4.1.1 to conclude the second part. □

4.2 The spectral theorem for unbounded self-adjoint operators

The main goal of this section, is to give the spectral theorem for self-adjoint operators. In short, we know that if A is a self-adjoint operator, and therefore we can transform it to a bounded, unitary operator U by the Cayley transform. Since U is a bounded, normal operator, we know that there exists a unique spectral measure E_U such that $U = \int \lambda \, dE_U(\lambda)$. Then we transform back to find the unique spectral measure for A.

But before we give the spectral theorem for self-adjoint operators, we wish to know how to integrate any measurable function with respect to a spectral measure. The reason why we wish to know this, is that the function $f : \sigma(A) \to \mathbb{C}$ given by $f(z) = z$ may not be bounded anymore, and so Proposition 3.2.14 does not hold anymore.

Proposition 4.2.1. *[4, Prop. 10.1, p. 202] Let E be a spectral measure for $(X, \mathscr{A}, \mathscr{H})$ and let $f : X \to \mathbb{C}$ be an $\mathscr{A}$-measurable function (so not necessarily bounded). Define the subspace $W_f \subset \mathscr{H}$ by*

$$W_f := \left\{ x \in \mathscr{H} | \int_X |f(\lambda)|^2 d\langle E(\lambda)x, x\rangle < \infty \right\} \tag{19}$$

Then there exists a unique (not necessarily bounded) operator A on $\mathscr{H}$ with domain W_f, with the property that

$$\langle Ax, x\rangle = \int_X f(\lambda)\ d\langle E(\lambda)x, x\rangle \tag{20}$$

for all $x \in W_f$. This unique operator will be denoted as $\int_X f\ dE$ and has domain W_f. Sometimes we will write $W_f = \mathrm{Dom}(\int_X f\ dE)$. This operator also satisfies the following equation for all $x \in W_f$

$$\left\| \left(\int_X f\ dE \right) x \right\|^2 = \int_X |f|^2(\lambda)\ d\langle E(\lambda)x, x\rangle \tag{21}$$

Remark 4.2.2. It should be noted that if f is a bounded function, $W_f = \mathscr{H}$ and this coincides with our definition of $\int f\ dE$ and so the only case we need to consider is when f is not bounded.

In order to prove Proposition 4.2.1, we consider another proposition. After this proposition is proved, the previous Proposition immediately follows.

Proposition 4.2.3. *[4, Prop. 10.2, p. 203] Let E be a spectral measure on $(X, \mathscr{A}, \mathscr{H})$, $f : X \to \mathbb{C}$ be an $\mathscr{A}$-measurable function and W_f as in Equation 19. Then*

1. *The space W_f is a dense linear subspace of $\mathscr{H}$, and the mapping $Q_f : W_f \to \mathbb{C}$ given by*

$$Q_f(x) = \int_X f(\lambda)\ d\langle E(\lambda)x, x\rangle$$

is a mapping such that $Q_f(\lambda x) = |\lambda|^2 Q_f(x)$ for $x \in W_f$ and $\lambda \in \mathbb{C}$, and such that the mapping $L_f : W_f \times W_f \to \mathbb{C}$ defined by

$$L_f(y, x) := \frac{1}{2}(\overline{Q_f(x+y)} - \overline{Q_f(x)} - \overline{Q_f(y)}) + \frac{i}{2}(\overline{Q(x+iy)} - \overline{Q(x)} - \overline{Q(iy)})$$

is a sesquilinear form.

2. *If $x, y \in W_f$ we have*

$$|L_f(x, y)| \leq \|y\| \sqrt{\int_X |f(\lambda)|^2\ d\langle E(\lambda)x, x\rangle} \tag{22}$$

3. *For each $x \in W_f$ there is a unique $z \in \mathscr{H}$ such that $L_f(y,x) = \langle y, z\rangle$ for all $y \in W_f$. Additionally, the mapping $x \mapsto z$ is linear and for all $y \in W_f$ we have*

$$\|z\|^2 = \int_X |f(\lambda)|^2 \, d\langle E(\lambda)x, x\rangle$$

Proof. 1): We note that if f is a bounded measurable function, then $W_f = \mathscr{H}$ because of Theorem 3.3.4 and therefore $\int f \, dE$ is a bounded operator. Then the rest of part 1) is automatically true.

Next let f be unbounded. First we prove that W_f is a dense linear subspace of $\mathscr{H}$. It is clear that if $x \in W_f$, then for $\mu \in \mathbb{C}$ we find $\int_X |f(\lambda)|^2 \, d\langle E(\lambda)\mu x, \mu x\rangle = |\mu|^2 \int_X |f(\lambda)|^2 \, d\langle E(\lambda)x, x\rangle < \infty$. So $\mu x \in W_f$. Next, if $x, y \in W_f$ then we get that, since for any set $A \in \mathscr{A}$ we have $E(A)$ is a projection. So we find

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

So the measure $\langle E(\cdot)(x+y), (x+y)\rangle$ can be bounded by the other two measures. Therefore we see

$$\int_X |f(\lambda)|^2 \, d\langle E(\lambda)(x+y), x+y\rangle \leq 2\int_X |f(\lambda)|^2 \, d\langle E(\lambda)x, x\rangle + 2\int_X |f(\lambda)|^2 \, d\langle E(\lambda)y, y\rangle$$

So $x + y \in W_f$, and so W_f is a linear subspace of $\mathscr{H}$.

Next we prove that W_f lies dense in $\mathscr{H}$. Consider the sets $A_n = \{x \in X| \, |f(x)| \leq n\}$. Then if $x \in \operatorname{ran}(E(A_n))$ there exists $y \in \mathscr{H}$ such that $x = E(A_n)y$. It follows then that $0 = E(\emptyset)y = E((X \setminus A_n) \cap A_n)y = E(X \setminus A_n)x$ and thus is $\langle E(X \setminus A_n)x, x\rangle = 0$. Therefore we find

$$\int_X |f(z)|^2 \, d\langle E(z)x, x\rangle = \int_{A_n} |f(z)|^2 \, d\langle E(z)x, x\rangle \leq n^2\langle E(A_n)x, x\rangle < \infty \tag{23}$$

So $x \in W_f$. Thus $\operatorname{ran}(E(A_n)) \subseteq W_f$. Because $\cup_{n\in\mathbb{N}} A_n = X$, we see that the union of the ranges of $E(A_n)$ are dense in $\mathscr{H}$, and for each A_n the range is contained in W_f. So W_f lies dense in $\mathscr{H}$.

Define the function $f_n := \chi_{A_n} f$ where χ_{A_n} is the characteristic function on A_n. Then f_n is a bounded function on X, and $f_{n+1} \geq f_n$. Therefore, using Lebesgue dominated convergence theorem

$$\lim_{n\to\infty} \int_X |f_n(z)|^2 \, d\langle E(z)x, x\rangle = \int_X \lim_{n\to\infty} |f_n(z)|^2 \, d\langle E(z)x, x\rangle = \int_X |f(z)|^2 \, d\langle E(z)x, x\rangle$$

or in other words $\lim_{n\to\infty} Q_{f_n}(x) = Q_f(x)$. Therefore

$$Q_f(\lambda x) = \lim_{n\to\infty} Q_{f_n}(\lambda x) = |\lambda|^2 \lim_{n\to\infty} Q_{f_n}(x) = |\lambda|^2 Q_f(x)$$

The other part of 1) is analogous.

2): First, let f be bounded. Then we know for $x, y \in \mathscr{H}$ that $L_f(y,x) = \langle y, \left(\int_X f(\lambda) \, dE\right) x\rangle$ and thus

$$|L_f(x,y)| \leq \|y\| \left\| \left(\int_X f(\lambda) \, dE(\lambda)\right) x \right\| = \|y\| \cdot \sqrt{\int_X |f(\lambda)|^2 \, d\langle E(\lambda)x, x\rangle}$$

Next, let f be unbounded and $x, y \in W_f$. We note that $L_f(y,x) = \lim_{n\to\infty} L_{f_n}(y,x)$ because $Q_f(z) = \lim_{n\to\infty} Q_{f_n}(z)$ for all $z \in W_f$. Therefore we see, again using Lebesgue dominated convergence theorem:

$$\begin{aligned}
|L_f(y,x)| = \lim_{n\to\infty} |L_{f_n}(y,x)| = \lim_{n\to\infty} \left|\left\langle y, \left(\int_X f_n \, dE\right) x \right\rangle\right| &\leq \lim_{n\to\infty} \left\| \left(\int_X f_n(\lambda) \, dE(\lambda)\right) x \right\| \|y\| \\
&= \lim_{n\to\infty} \|y\| \cdot \sqrt{\int_X |f_n(\lambda)|^2 \, d\langle E(\lambda)x, x\rangle} \\
&= \|y\| \sqrt{\int_X |f(\lambda)|^2 \, d\langle E(\lambda)x, x\rangle}
\end{aligned}$$

For the last inequality sign is because of Equation (23) and continuity of the square root. So we see that part 2 is also true.

3): First, let f be bounded. Then we already found a unique operator which does exactly this, see Proposition 3.2.14. Therefore the only interesting part is f unbounded. Because Equation 22 is true for any $x \in W_f$, we conclude that if we consider $x \in W_f$ to be fixed, we have that the mapping $L_x : y \mapsto L(y,x)$ is bounded and linear. By Remark 2.1.2, we can extend this operator to a bounded operator $\tilde{L}_x(y)$ such that $\tilde{L}_x(y) = L(y,x)$ for $y \in W_f$. Because $\tilde{L}_x$ is a bounded functional, we find by the Riesz Representation Theorem that there exists a unique $z \in \mathscr{H}$ such that $\tilde{L}_x = \langle \cdot, z\rangle$. If $y \in W_f$ we find:

$$L(y,x) = \tilde{L}_x(y) = \langle y, z\rangle$$

We note that f_n is bounded, so $\int_X f_n \, dE$ is a bounded operator. Therefore we find for $n, m \in \mathbb{N}$:

$$\begin{aligned}
\left\| \int_X f_n \, dEx - \int_X f_m \, dEx \right\|^2 &= \left\| \int_X f_n - f_m \, dEx \right\| \\
&= \int_X |f_n(z) - f_m(z)| \, d\langle E(z)x, x\rangle
\end{aligned}$$

Because $\int_X |f(z)|^2 \, d\langle E(z)x, x\rangle < \infty$ we get by Lebesgue dominated theorem

$$\lim_{n\to\infty} \int_X |f_n(z)|^2 \, d\langle E(z)x, x\rangle = \int_X |f(z)| \, d\langle E(z)x, x\rangle$$

and so $(\int_X f_n \ dEx)_n$ is a Cauchy sequence. Therefore the series converges. Define the vector $\xi := \lim_{n\to\infty} \int_X f_n \, dEx$. W find then

$$\langle y, \xi\rangle = \lim_{n\to\infty} \langle y, \int_X f_n \, dEx\rangle = \lim_{n\to\infty} L_{f_n}(y,x) = L_f(y,x) = \langle y, z\rangle$$

Because $y \in W_f$, and W_f lies dense in $\mathscr{H}$, we find that $z = \lim_{n\to\infty} \int_X f_n \, dEx$. Then we find

$$\begin{aligned}
\|z\|^2 &= \lim_{n\to\infty} \| \int_X f_n \, dEx\|^2 = \lim_{n\to\infty} \int_X |f_n(z)|^2 \, d\langle E(z)x, x\rangle \\
&= \int_X |f(z)| \, d\langle E(z)x, x\rangle
\end{aligned}$$

Therefore $\|z\|^2 = \int_X |f(z)| \, d\langle E(z)x, x\rangle$. □

By construction, we see that we get the same result as for the bounded case, only with a domain we need to take into account. For if $f : X \to \mathbb{C}$ and $x \in W_f$ and $y \in \mathscr{H}$, we find

$$\left\langle \left(\int_X f \, dE \right) x, y \right\rangle = \int_X f(z) \, d\left\langle E(z)x, y \right\rangle$$

Theorem 4.2.4. *[4, Thm. 10.30, p. 223] Let g as in Equation 16. If A is a self-adjoint operator on $\mathscr{H}$, and U is its Cayley transform, define for the set $U \subseteq \sigma(A)$ the operator*

$$E_A(U) := E_U(g^{-1}(U))$$

Then E_A is a spectral measure and

$$A = \int_{\sigma(A)} \lambda \, dE_A$$

Here E_U is the spectral measure of the bounded operator U.

Proof. Define for any Borel set B of $\sigma(A)$ the mapping $E_A(B) := E_U(g^{-1}(B))$. We must show that this is indeed a spectral measure. Let B be a Borel set of $\sigma(A)$. Observe $g : \mathbb{S} \setminus \{1\} \to \mathbb{R}$ is a homeomorphism. Hence, g preserves the Borel sets. Therefore, g is a measurable function and $g^{-1}(B)$ is a Borel set of $\sigma(U)$. Therefore $E_A(B)$ is a spectral measure, because $E_U(g^{-1}(B))$ is a spectral measure.

For any $x \in \mathrm{Dom}(A)$ and $y \in \mathscr{H}$, let B_A be a Borel set in $\sigma(A)$ and B_U be a Borel set in $\sigma(U)$. Then define the measures

$$E^A_{x,y}(B_A) := \langle E_A(B_A)x, y \rangle = \langle E_U(g^{-1}(B_A))x, y \rangle \qquad E^U_{x,y}(B_U) := \langle E_U(B_U)x, y \rangle$$

We see that $E^A_{x,y}(B_A) = E^U_{x,y}(g^{-1}(B_A))$. By the abstract change of coordinates, we find for any measurable function $u : \sigma(A) \to \mathbb{C}$

$$\int_{\sigma(A)} u \, dE^A_{x,y} = \int_{\sigma(U)} u \circ g \, dE^U_{x,y} \tag{24}$$

For a proof of the abstract change of coordinates, we refer to $[8, p.\ 154]$. We know that

$$\langle Ax, y \rangle = \langle g(U)x, y \rangle = \int_{\sigma(U)} g(\lambda) \, dE^U_{x,y}(\lambda) = \int_{\sigma(U)} (\mathrm{Id} \circ g)(\lambda) \, dE^U_{x,y}(\lambda)$$

and so using Equation (24) we find that

$$\langle Ax, y \rangle = \int_{\sigma(A)} \lambda \, dE^A_{x,y}(\lambda) = \int_{\sigma(A)} \lambda \, d\langle E_A(\lambda)x, y \rangle = \langle \left(\int_{\sigma(A)} \lambda \, dE_A(\lambda) \right) x, y \rangle$$

This again is true for any $y \in \mathscr{H}$, and so we find

$$Ax = \left(\int_{\sigma(A)} \lambda \, dE_A(\lambda) \right) x \qquad (x \in \mathrm{Dom}(A)) \tag{25}$$

Thus the equality holds. ☐

Theorem 4.2.5 (The spectral theorem for unbounded self-adjoint operators)**.** *([4, Thm. 10.4, p. 205] Let A be a self-adjoint operator on $\mathscr{H}$. Then there exists a unique spectral measure E_A on the Borel subsets of $\sigma(A)$ such that:*

1. *$A = \int_{\sigma(A)} \lambda \, dE_A(\lambda)$*
2. *If B is a Borel set of $\mathbb{R}$, and $B \cap \sigma(A) = \emptyset$ then $E_A(B) = \emptyset$. Additionally if B is an open subset of $\mathbb{R}$ and $B \cap \sigma(A) \neq \emptyset$ then $E_A(B) \neq 0$*
3. *Let $T \in \mathscr{B}(\mathscr{H})$. If $TA \subseteq AT$ then $T(\int \phi \, dE_A) \subseteq (\int \phi \, dE_A)T$ for any Borel measurable function ϕ.*

Proof. 1): By Theorem 4.2.4 we found a measure such that $A = \int \lambda \, dE_A(\lambda)$. The only thing we need to prove, is uniqueness. Let E_A be the spectral measure found in Theorem 4.2.4 and E_U the spectral measure of the Cayley transform of A, denoted U. Let P be another spectral measure such that $A = \int_{\sigma(A)} \lambda \, dP(\lambda)$. Then we define $P_U(B) := P(f^{-1}(B))$ for B a Borel set in $\mathbb{S} \setminus \{1\}$, where f is defined as in Equation (15). Then by the same argument as in previous theorem, P_U is a spectral measure on $\mathbb{S} \setminus \{1\}$. Also, we see that

$$U = f(A) = \int_{\sigma(A)} f(\lambda) \, dP(\lambda)$$

Let $x, y \in \mathscr{H}$, B_A a Borel set of $\sigma(A)$ and B_U a Borel set of $\sigma(U)$. Define the following measures

$$P^A_{x,y}(B_A) = \langle P(B_A)x, y\rangle \qquad P^U_{x,y}(B_U) = \langle P_U(B_U)x, y\rangle = \langle P(f^{-1}(B_U))x, y\rangle$$

We note again that $P^U_{x,y}(B_U) = P^A_{x,y}(f^{-1}(B_U))$, and so again, we find for any measurable function $u : \sigma(U) \to \mathbb{C}$:

$$\int_{\sigma(U)} u(\lambda) \, dP^U_{x,y} = \int_{\sigma(A)} u \circ f \, dP^A_{x,y} \tag{26}$$

And so, by the same reasoning as in Theorem 4.2.4, we conclude that $U = \int_{\sigma(U)} z \, dP_U(z)$. But we know, by Theorem 3.4.11, that there exists a unique spectral measure such that $U = \int z \, dE_U$. Therefore $P_U = E_U$. Transforming back gives us that $E_A = P$. So E_A is the unique spectral measure.

2): Let B be a Borel set of $\mathbb{R}$ such that $B \cap \sigma(A) = \emptyset$. Then $g^{-1}(B) \cap \sigma(U) = g^{-1}(B \cap \sigma(A)) = g^{-1}(\emptyset) = \emptyset$. Therefore we can conclude

$$E_A(B) = E_U(g^{-1}(B)) = E_U(g^{-1}(\sigma(U) \cap B)) = E_U(\emptyset) = 0$$

On the other hand, if B is open and $B \cap \sigma(A) \neq \emptyset$, then $g^{-1}(B)$ is open and $g^{-1}(B) \cap \sigma(U) \neq \emptyset$. Therefore $E_A(B) = E_U(g^{-1}(B)) \neq 0$ by Theorem 3.4.11.

3) Let $T \in \mathscr{B}(\mathscr{H})$. If $TA \subseteq AT$ then $TA = AT$ on $\mathrm{Dom}(A)$. Therefore we have $(A \pm i)T = T(A \pm i)$ on $\mathrm{Dom}(A)$. Thus we can conclude $(A+i)^{-1}T = T(A+i)^{-1}$, and thus we have $TU = UT$ where U is given by Equation (17). Because $U^* = U^{-1} = (A+i)(A-i)^{-1}$, we also conclude $TU^* = U^*T$ by the same argument. By Theorem 3.4.11 we have $E_U(B)T = TE_U(B)$ for any

Borel subset of $\sigma(U)$. Thus $E_A(C)T = TE_A(C)$ for any Borel subset C of $\sigma(A)$, by construction of E_A. Therefore we get, for any $x, y \in \mathscr{H}$:

$$\langle E_A(C)Tx, y\rangle = \langle TE_A(C)x, y\rangle = \langle E_A(C)x, T^*y\rangle$$

Consider any measurable function $\phi : \sigma(A) \to \mathbb{C}$, and define the operator $Z_\phi := \int \phi \; dE_A$. This operator is well defined by Proposition 4.2.1 with domain W_ϕ. Therefore consider $x \in W_\phi$ and $y \in \mathscr{H}$. We find

$$\begin{aligned}\langle Z_\phi Tx, y\rangle &= \int_{\sigma(A)} \phi(z) \; d\langle E_A(z)Tx, y\rangle \\ &= \int_{\sigma(A)} \phi(z) \; d\langle E_A(z)x, T^*y\rangle = \langle Z_\phi x, T^*y\rangle = \langle TZ_\phi x, y\rangle\end{aligned}$$

This is true for any $y \in \mathscr{H}$, and therefore we find $Z_\phi Tx = TZ_\phi x$ for $x \in W_f$. We conclude that $T(\int \phi \; dE_A) \subseteq (\int \phi \; dE_A)T$. □

If we combine the spectral theorem and Theorem 4.2.1, we find that we can integrate any measurable function over the spectral measure of A. Therefore, if $\phi : \sigma(A) \to \mathbb{C}$ is a measurable function, and A is a self-adjoint operator with spectral measure E_A, we define

$$\phi(A) := \int_{\sigma(A)} \phi(\lambda) \; dE_A(\lambda) \tag{27}$$

In fact, the mapping $\phi \mapsto \int \phi \; dE_A$ has a lot of properties we already know for the bounded variant. The following lemma sums them up. Because the proof is just some computations, and would not give a lot of insight, we omit the proof. For a proof, see for example $[10, p.\ 78]$.

Lemma 4.2.6. *[1, Thm. 4.10, p. 323] If $(X, \mathscr{A})$ is a measurable space, $\mathscr{H}$ a Hilbert space, and E is a spectral measure for $(X, \mathscr{A}, \mathscr{H})$, let $\Phi(X, \mathscr{A})$ be the algebra of all $\mathscr{A}$-measurable functions $\phi : X \to \mathbb{C}$. Then for $\phi, \psi \in \Phi(X, \mathscr{A})$:*

1. *the operator $\int_X \phi \; dE$ is closed,*
2. $\left(\int_X \phi \; dE\right)^* = \int_X \overline{\phi} \; dE$,
3. $\left(\int_X \phi \; dE\right)\left(\int \psi \; dE\right) \subseteq \int_X \phi\psi \; dE$ *and* $\mathrm{Dom}(\left(\int_X \phi \; dE\right)\left(\int \psi \; dE\right)) = W_\phi \cap W_\psi$ *with* W_ϕ,W_ψ *as in Equation* 19,
4. *If ψ is bounded,* $\left(\int_X \phi \; dE\right)\left(\int_X \psi \; dE\right) = \left(\int_X \psi \; dE\right)\left(\int_X \phi \; dE\right) = \int_X \phi\psi \; dE$,
5. $\left(\int_X \phi \; dE\right)^* \left(\int_X \phi \; dE\right) = \int_X |\phi|^2 \; dE$.

5 Theorem of Nelson

5.1 The theorem of Nelson

Now we have the tools to start working on the theorem of Nelson. The theorem of Nelson is especially interesting, because it tells us that if two self-adjoint operators A and B commute on a special dense linear subset, and when some operator is essentially self-adjoint, then for any measurable functions f and g we have $f(A)$ and $g(B)$ also commute. In Nelson's paper, Nelson was inspired by analytic vectors and Lie-theory to prove this theorem [5, Cor. 9.2, p. 603]. We will however prove it by using elementary analysis. But first we need a definition of commuting.

Definition 5.1.1. Let A, B be two linear operators on a Hilbert space $\mathscr{H}$. We say A and B *commute* if $ABx = BAx$ for every $x \in \mathrm{Dom}(AB) \cap \mathrm{Dom}(BA)$. If A and B are self-adjoint, we say A and B *commute strongly* if the spectral measures E_A and E_B, of A and B respectively, commute.

It appears commutation of the spectral measures seem to imply something, for the term *strongly* commuting suggests that this form of commutation is stronger than normal commutation. And in fact, if the spectral measures commute, then all possible operators created by integrating two measurable functions with respect to these measures also commute. The following proposition and its proof are my own work and have not been taken from the literature.

Proposition 5.1.2. *Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be two measurable spaces, and $\mathscr{H}$ a Hilbert space. Let E_X be a spectral measure for $(X, \mathscr{A}, \mathscr{H})$ and E_Y a spectral measure for $(Y, \mathscr{B}, \mathscr{H})$.*

If $E_X(U)$ and $E_Y(V)$ commute for any $U \in \mathscr{A}$ and $V \in \mathscr{B}$, then for any $\mathscr{A}$-measurable function $f : X \to \mathbb{C}$ and any $\mathscr{B}$-measurable function $g : Y \to \mathbb{C}$ we have

$$\left(\int_X f\ dE_X\right)\left(\int_Y g\ dE_Y\right)x = \left(\int_Y g\ dE_Y\right)\left(\int_X f\ dE_X\right)x$$

for any $x \in \mathrm{Dom}(\int_X f\ dE_X \int_Y g\ dE_Y) \cap \mathrm{Dom}(\int_Y g\ dE_Y \int_X f\ dE_X)$. In other words, $\int_X f\ dE_X$ and $\int_Y g\ dE_Y$ commute.

Proof. We start with a claim.

Claim 1. $\int_Y g\ dE_Y$ commutes with $E_X(U)$ for any $U \in \mathscr{A}$ on $\mathrm{Dom}(\int_Y g\ dE_Y)$.

Proof. Take $x \in \mathrm{Dom}(\int_Y g\ dE_Y)$. Then we find that for any $V \in \mathscr{B}$ and $y \in \mathscr{H}$:

$$\langle E_Y(V)E_X(U)x, y\rangle = \langle E_X(U)E_Y(V)x, y\rangle = \langle E_Y(V)x, E_X(U)y\rangle$$

Therefore we immediately find that

$$\begin{aligned}
\langle E_X(U)\left(\int_Y g\ dE_Y\right)x, y\rangle &= \langle\left(\int_Y g\ dE_Y\right)x, E_X(U)y\rangle \\
&= \int_Y g(z)\ d\langle E_Y(z)x, E_X(U)y\rangle \\
&= \int_Y g\ d\langle E_Y(z)E_X(U)x, y\rangle \\
&= \langle\left(\int_Y g\ dE_Y\right)E_X(U)x, y\rangle
\end{aligned}$$

Since this is true for any $y \in \mathscr{H}$, we have that $\left(\int_Y g\, dE_Y\right) E_X(U) = E_X(U) \left(\int_Y g\, dE_Y\right)$ on $\mathrm{Dom}(\int_Y g\, dE_Y)$. □

Next, we want to do this for f and g. We work in a similar way as in the previous claim. Let $x \in \mathrm{Dom}(\int_X f\, dE_X \int_Y g\, dE_Y) \cap\ \mathrm{Dom}(\int_Y g\, dE_Y \int_X f\, dE_X)$, and $y \in \mathrm{Dom}((\int g\, dE_Y)^*)$. Then for any $U \in \mathscr{B}$ we find the equation

$$\begin{aligned}\left\langle E_X(U)\left(\int_Y g\, dE_Y\right)x, y\right\rangle &= \left\langle \left(\int_Y g\, dE_Y\right) E_X(U)x, y\right\rangle \\ &= \left\langle E_X(U)x, \left(\int_Y g\, dE_Y\right)^* y\right\rangle\end{aligned}$$

And so we find that the following holds

$$\begin{aligned}\left\langle \left(\int_X f\, dE_X\right)\left(\int_Y g\, dE_Y\right)x, y\right\rangle &= \int_X f(z)\, d\left\langle E_X(z)\left(\int_Y g\, dE_Y\right)x, y\right\rangle \\ &= \int_X f\, d\left\langle E_X(z)x, \left(\int_Y g\, dE_Y\right)^* y\right\rangle \\ &= \left\langle \left(\int_X f\, dE_X\right)x, \left(\int_Y g\, dE_Y\right)^* y\right\rangle \\ &= \left\langle \left(\int_Y g\, dE_Y\right)\left(\int_X f\, dE_X\right)x, y\right\rangle\end{aligned}$$

This is only true for any $y \in \mathrm{Dom}((\int_Y g\, dE_Y)^*)$. Because E_Y is a spectral measure, by Lemma 4.2.6 we have that $(\int_Y g\, dE_Y)^* = \int_Y \overline{g}\, dE_Y$. By Proposition 4.2.3 we find that the domain of $\int_Y \overline{g}\, dE_Y$ lies dense in $\mathscr{H}$.

So let $y \in \mathscr{H}$. Then there exists a sequence $(y_n)_n \subseteq \mathrm{Dom}(\int_Y \overline{g}\, dE_Y)$ such that $y_n \to y$. Because the inner product is continuous, we find:

$$\begin{aligned}\left\langle \left(\int_X f\, dE_X\right)\left(\int_Y g\, dE_Y\right)x, y\right\rangle &= \lim_{n\to\infty}\left\langle \left(\int_X f\, dE_X\right)\left(\int_Y g\, dE_Y\right)x, y_n\right\rangle \\ &= \lim_{n\to\infty}\left\langle \left(\int_Y g\, dE_Y\right)\left(\int_X f\, dE_X\right)x, y_n\right\rangle \\ &= \left\langle \left(\int_Y g\, dE_Y\right)\left(\int_X f\, dE_X\right)x, y\right\rangle\end{aligned}$$

Again, this is true for any $y \in \mathscr{H}$. And therefore we conclude that

$$\left(\int_X f\, dE_X\right)\left(\int_Y g\, dE_Y\right)x = \left(\int_Y g\, dE_Y\right)\left(\int_X f\, dE_X\right)x$$

for any $x \in \mathrm{Dom}(\int_X f\, dE_X \int_X g\, dE_Y) \cap \mathrm{Dom}(\int_Y g\, dE_Y \int f\, dE_X)$ □

This result looks promising. If we consider two strongly commuting self-adjoint operators A and B, then not only do these operators commute, we also find that A^n and B^n commute for any $n \in \mathbb{N}$, and the operators e^A and $B^2 - B$ commute. So if the spectral measures commute, we see that a lot of operators commute. Therefore, we wish to investigate whether two self-adjoint operators strongly commute.

Because of Corollary 3.4.15 and the spectral theorem for bounded operators (Theorem 3.4.11), we see that the spectral measures of two bounded self-adjoint operators A and B commute if and only if A and B commute. Therefore, if A and B commute, then they automatically commute strongly.

Additionally, if A is a self-adjoint operator, and B is a bounded self-adjoint operator, they strongly commute as well.

Lemma 5.1.3. *Let A be a self-adjoint operator and let B be a bounded self-adjoint operator. If A and B commute on $\mathrm{Dom}(A)$, then A and B strongly commute*

Proof. Since A and B commute on $\mathrm{Dom}(A)$, then by the spectral theorem for unbounded operators (Theorem 4.2.5) we find that B and $(\int \phi \ dE_A)$ commute. Taking $\phi = \chi_U$ for any Borel subset $U \subseteq \sigma(A)$ we see this immediately tells us that $E_A(U)B = BE_A(U)$. Then using the spectral theorem for bounded, normal operators, we see that the spectral measures commute. In other words, A and B commute strongly. □

But when both A and B are unbounded, this correspondence does not need to hold anymore. However, Nelson's theorem will give us one criterium to show that two self-adjoint operators strongly commute. The rest of this section will be dedicated to stating the theorem of Nelson and proving it.

Definition 5.1.4. Let A be a linear operator on $\mathscr{H}$, and let $\mathscr{D}$ be a subset of $\mathrm{Dom}(A)$. We say $\mathscr{D}$ is a *core* of A if $\mathscr{D}$ lies dense in $\mathrm{Dom}(A)$ with respect to the graph norm, where the graph norm is given as in Lemma 2.1.10.

Remark 5.1.5. Note that if A is closable, we find that if $\mathscr{D}$ is a core of A, then $\overline{A|_{\mathscr{D}}} = \overline{A}$. If A is closed, then it tells us $\overline{A|_{\mathscr{D}}} = A$. If A is self-adjoint, then $A|_{\mathscr{D}}$ is essentially self-adjoint.

Cores can be useful, because some of the properties of an operator can be seen if we only look at a core. So we can find some properties of the operator, without using the whole domain. We will be using cores as well to prove the theorem of Nelson. First we note an interesting result that characterizes whenever two self-adjoint operators strongly commute. The proof is also not based on literature.

Lemma 5.1.6. *Let $\mathscr{H}$ be a Hilbert space, and A, B be self-adjoint operators on $\mathscr{H}$. Then the following assertions are equivalent:*

1. *A and B strongly commute,*
2. *For all $z \in \mathbb{C} \setminus \mathbb{R}$ the operator $(A + z)^{-1}$ commutes with B,*
3. *The operators $(A + i)^{-1}$ and $(A - i)^{-1}$ commute with B.*

Proof. 1) $\Rightarrow$ 2): Assume that A and B commute strongly. We note that $(A+z)^{-1} = f(A)$ with $f : \sigma(A) \to \mathbb{C}$ is given by $f(x) = \frac{1}{x+z}$. Since $\sigma(A) \subset \mathbb{R}$ and $z \in \mathbb{C} \setminus \mathbb{R}$, we have that $f(x)$ is a continuous function, and therefore measurable. We use Proposition 5.1.2 to conclude that $(A+z)^{-1}$ and B commutes for $z \in \mathbb{C} \setminus \mathbb{R}$.

2) $\Rightarrow$ 3): If $(A+z)^{-1}$ and B commute for any $z \in \mathbb{C} \setminus \mathbb{R}$, then it is surely true for $z = \pm i$.

3) $\Rightarrow$ 1): Let $(A \pm i)^{-1}$ and B commute. So $(A \pm i)^{-1}B \subseteq B(A \pm i)^{-1}$. Because A is a self-adjoint operator, we have $(A \pm i)^{-1}$ is a bounded operator. Therefore, by the spectral theorem for unbounded operators, Theorem 4.2.5, we find that $(A \pm i)^{-1}\left(\int \phi \, dE_B\right) \subseteq \left(\int \phi \, dE_B\right)(A \pm i)^{-1}$ for any Borel measurable function ϕ. We note that if $V \subseteq \mathbb{C}$ is a Borel set, then

$$E_B(V)(A \pm i)^{-1} = \left(\int \chi_V \, dE_B\right)(A \pm i)^{-1} = (A \pm i)^{-1}(\left(\int \chi_V \, dE_B\right)) = (A \pm i)^{-1}E_B(V)$$

Therefore we see that $E_B(V)x = (A-i)^{-1}E_B(V)(A-i)x$ for any $x \in \mathrm{Dom}(A)$. Hence we can conclude

$$\begin{aligned}(A-i)(A+i)^{-1}E_B(V) = (A-i)E_B(V)(A+i)^{-1} &= (A-i)E_B(V)(A-i)^{-1}(A-i)(A+i)^{-1} \\ &= (A-i)(A-i)^{-1}E_B(V)(A-i)(A+i)^{-1} \\ &= E_B(V)(A-i)(A+i)^{-1}\end{aligned}$$

Hence we can conclude that $E_B(V)$ commutes with U; the Cayley transform of A. We can do the same trick to prove that $E_B(V)$ commutes with U^*. Because U is a bounded operator, we use the spectral theorem for bounded operators, Theorem 3.4.11, to conclude that for any Borel set $W \subseteq \mathbb{C}$ that $E_B(V)E_U(W) = E_U(W)E_B(V)$ where E_U is the spectral measure for U. Because the spectral measure of A is defined as $E_A(W) = E_U(g^{-1}(W))$, we immediately see $E_A(W)E_B(V) = E_B(V)E_A(W)$. So A and B commute strongly. □

Next, we will prove two theorems. These theorems are the core of the proof of Nelson's theorem, and if we have proven these two, then the theorem of Nelson will follow directly.

Remark 5.1.7. We want to adress the notion that the two upcoming lemmas are based on [9, Lm. 1, p. 365], but is slightly changed. In our version, we do not need $\mathrm{Dom}(B) \subseteq \mathrm{Dom}(A)$, but instead assume that the linear subspace $\mathscr{D}$ has the additional property that $\mathscr{D} \subseteq \mathrm{Dom}(A) \cap \mathrm{Dom}(B)$. The proof of these lemmas will also differ slightly.

Lemma 5.1.8. *[9, Lm. 1, p. 365] Let A be a self-adjoint operator, and B a symmetric operator. Suppose that there exists a linear subspace $\mathscr{D} \subset \mathrm{Dom}(AB) \cap \mathrm{Dom}(BA) \cap \mathrm{Dom}(A) \cap \mathrm{Dom}(B)$ such that*

1. *$ABx = BAx$ for any $x \in \mathscr{D}$,*
2. *$\mathscr{D}$ is a core for A and B,*
3. *$\|Bx\| \leq \lambda \|(A+i)x\|$ for all $x \in \mathscr{D}$ and some $\lambda > 0$.*

Then B is essentially self-adjoint and $\overline{B}$ and A strongly commute.

Proof. We start the proof with a claim:

Claim 1. There exists a $C \in \mathscr{B}(\mathscr{H})$ such that $Bx = C(A+i)x$ for $x \in \mathscr{D}$.

Proof. We note that $i \notin \sigma(A)$, so $A+i$ is boundedly invertible. Define the subspace $\mathscr{H}_0 := (A+i)(\mathscr{D})$. Thus for any $x \in \mathscr{D}$ there exists a $y \in \mathscr{H}_0$ such that $x = (A+i)^{-1}y$. Thus we have

$$\|Bx\| = \|B(A+i)^{-1}y\| \leq \lambda\|(A+i)(A+i)^{-1}y\| = \lambda\|y\|$$

because of the assumption. Define $C := B(A+i)^{-1}$ on $\mathscr{H}_0$, then $\|Cy\| \leq \lambda\|y\|$ for all $y \in \mathscr{H}_0$. So C is a bounded operator on $\mathscr{H}_0$. By Remark 2.1.2 we can extend C in such a way that C becomes a bounded operator on $\mathscr{H}$ with $C = B(A+i)^{-1}$ on $\mathscr{H}_0$. We conclude that if $x \in \mathscr{D}$ then $Bx = B(A+i)^{-1}(A+i)x = C(A+i)x$. □

So we have $Bx = C(A+i)x$ for any $x \in \mathscr{D}$. Because $\mathscr{D}$ is a core of B, we have $B^* = (B|_{\mathscr{D}})^*$ where $B|_{\mathscr{D}}$ denotes the operator B with domain $\mathscr{D}$. We then find

$$B^* = (B|_{\mathscr{D}})^* = (C(A+i)|_{\mathscr{D}})^* = (C(A+i))^* = (A-i)C^*$$

Define the operator $|A-i| := ((A-i)^*(A-i))^{1/2}$. In other words,

$$|A-i| = \int_{\sigma(A)} \sqrt{(\lambda+i)(\lambda-i)}\, dE(\lambda) = \int \sqrt{\lambda^2+1}\, dE(\lambda)$$

where E is the spectral measure of A. We note that $\mathrm{Dom}(\int_{\sigma(A)} \frac{1}{\sqrt{\lambda^2+1}}\, dE) = \mathscr{H}$. Because $\sqrt{\lambda^2+1} > 1$ for any $\lambda \in \mathbb{R}$, we find for $x \in \mathscr{H}$,

$$\int_{\sigma(A)} \left|\frac{1}{\lambda^2+1}\right| d\langle E(\lambda)x, x\rangle \leq \int_{\sigma(A)} 1\, d\langle E(\lambda)x, x\rangle = \langle x, x\rangle$$

Therefore $\mathrm{Dom}(\int \frac{1}{\sqrt{\lambda^2+1}}\, dE) = \mathscr{H}$. Additionally, we find

$$\left\|\left(\int_{\sigma(A)} \frac{1}{\sqrt{\lambda^2+1}}\, dE\right) x\right\|^2 = \int_{\sigma(A)} \frac{1}{\lambda^2+1}\, d\langle E(\lambda)x, x\rangle \leq \|x\|^2$$

Therefore we see that $|A-i|^{-1} := \int_{\sigma(A)} \frac{1}{\sqrt{\lambda^2+1}}\, dE$ is a bounded operator defined on $\mathscr{H}$ with norm $\||A-i|^{-1}\| \leq 1$. Also note by Lemma 4.2.6 that

$$\left(|A-i|^{-1}\right)^* = \left(\int_{\sigma(A)} \frac{1}{\sqrt{\lambda^2+1}}\, dE\right)^* = \int_{\sigma(A)} \frac{1}{\sqrt{\lambda^2+1}}\, dE = |A-i|^{-1}$$

So $|A-i|^{-1}$ is a bounded, self-adjoint operator.

Next consider $(A-i)|A-i|^{-1}$. Then we find for $x \in \mathscr{H}$

$$\begin{aligned}\|(A-i)|A-i|^{-1}x\|^2 &= \left\|\left(\int_{\sigma(A)} \frac{\lambda-i}{\sqrt{\lambda^2+1}}\, dE(\lambda)\right) x\right\|^2 \\ &= \int_{\sigma(A)} \frac{\lambda-i}{\sqrt{\lambda^2+1}} \cdot \frac{\lambda+i}{\sqrt{\lambda^2+1}}\, d\langle E(\lambda)x, x\rangle = \int_{\sigma(A)} 1\, d\langle E(\lambda)x, x\rangle = \|x\|^2\end{aligned}$$

Thus $\|(A-i)|A-i|^{-1}\| = 1$ and so $(A-i)|A-i|^{-1}$ is a bounded operator.

Define $F := (A-i)|A-i|^{-1}C^*$. Then note that F is a bounded operator. Additionally, by Lemma 4.2.6 we see that $(A \pm i) = \int_{\sigma(A)} \lambda \pm i \; dE(\lambda)$ and $|A-i|^{-1}$ commute. Therefore we see

$$\begin{aligned} |A-i|F &= |A-i|(A-i)|A-i|^{-1}C^* = |A-i||A-i|^{-1}(A-i)C^* \\ &= (A-i)C^* = B^* \end{aligned}$$

Because B is a symmetric operator, we have $B \subseteq B^*$, and so for any $x \in \mathrm{Dom}(B)$ we have

$$Bx = B^*x = |A-i|Fx$$

Therefore, if $x \in \mathscr{D}$ we find:

$$(A+i)Bx = (A+i)|A-i|Fx = |A-i||A-i|^{-1}(A+i)|A-i|Fx = |A-i|(A+i)Fx \quad (28)$$

We assumed $Abx = BAx$ for $x \in \mathscr{D}$. So we find for $x \in \mathscr{D}$

$$(A+i)Bx = B(A+i)x = |A-i|F(A+i)x \quad (29)$$

Comparing (28) and (29), we see that $(A+i)Fx = F(A+i)x$ for $x \in \mathscr{D}$. Since $\mathscr{D}$ is a core for A, $(A+i)(\mathscr{D})$ lies dense in $\mathscr{H}$. Therefore, if $y \in \mathscr{H}$, there exists a sequence $(y_n)_n \subseteq \mathscr{D}$ such that $(A+i)y_n \to y$. Then $y_n = (A+i)^{-1}(A+i)y_n \to (A+i)^{-1}y$. And so we find

$$\begin{aligned} (A+i)^{-1}Fy &= \lim_{n\to\infty}(A+i)^{-1}F(A+i)y_n = \lim_{n\to\infty}(A+i)^{-1}(A+i)Fy_n \\ &= \lim_{n\to\infty} Fy_n = F(A+i)^{-1}y \end{aligned}$$

Therefore we find that F and $(A+i)^{-1}$ commute. Taking adjoints, we see that F^* and $(A-i)^{-1}$ commute. Hence F^* commute with $|A-i|$.

Since $F = (A-i)|A-i|^{-1}C^*$, we find $F^* = C((A-i)|A-i|^{-1})^* = C(A+i)|A-i|^{-1}$. Remember $B = C(A+i)$ on $\mathscr{D}$ and B is symmetric. Therefore if $x \in \mathscr{D}$, we see that the following holds

$$\begin{aligned} Bx &= C(A+i)|A-i|^{-1}|A-i|x = F^*|A-i|x = |A-i|F^*x \\ &= B^*x = |A-i|Fx \end{aligned}$$

So $F^* = F$ on $\mathscr{D}$. Since $\mathscr{D}$ is dense in $\mathscr{H}$ and F is bounded, we find $F^* = F$.

Now that we know that F is self-adjoint, we prove that $\overline{B}$ is self-adjoint. It is sufficient to show that the deficiency indices are 0. Suppose that $B^*x = zx$ with $z = \pm i$ for some $x \in \mathrm{Dom}(B^*)$. Because $B^* = |A-i|F$, we find $Fx = z|A-i|^{-1}x$. Then

$$\langle Fx, x\rangle = z\langle |A-i|^{-1}x, x\rangle = z\||A-i|^{-1/2}x\|$$

Because F is self-adjoint, the left side of the equation is real. Because $z = \pm i$ we conclude $|A-i|^{-1/2}x = 0$ and so $x = 0$. Therefore $n_+ = n_- = 0$. So B^* is self-adjoint, or equivalently $B^* = B^{**} = \overline{B}$. So B is essentially self-adjoint.

Because $\overline{B} = B^* = |A-i|F$ and F commutes with $(A+i)^{-1}$, we find

$$(A+i)^{-1}\overline{B} = (A+i)^{-1}|A-i|F = |A-i|F(A+i)^{-1} = \overline{B}(A+i)^{-1}$$

on $\mathrm{Dom}(B^*) = \mathrm{Dom}(\overline{B})$. We also know that F^* and $(A-i)^{-1}$ commutes, and $F = F^*$. Therefore, F and $(A-i)^{-1}$ commute, and thus by the same calculation we conclude that $(A-i)^{-1}$ and $\overline{B}$ commute on $\mathrm{Dom}(\overline{B})$. By Lemma 5.1.6 we have that A and $\overline{B}$ are strongly commuting. □

Lemma 5.1.9. *[9, Lm. 2, p. 366] Let A be a self-adjoint operator, and assume B_1, B_2 are symmetric operators. Suppose that*

1. *There exist linear subspaces $\mathscr{D}_i \subseteq \mathrm{Dom}(AB_i) \cap \mathrm{Dom}(B_iA) \cap \mathrm{Dom}(A) \cap \mathrm{Dom}(B_i)$ for $i = 1,2$ such that $AB_ix = B_iAx$ for all $x \in \mathscr{D}_i$ and such that $\mathscr{D}_i$ is a core of A and B_i for $i = 1,2$.*

2. *There exists a linear subspace $\mathscr{D}_{12} \subseteq \mathrm{Dom}(B_1B_2) \cap \mathrm{Dom}(B_2B_1)$ with $\mathscr{D}_{12} \subseteq \mathscr{D}_1 \cap \mathscr{D}_2$ such that $B_1B_2x = B_2B_1x$ for all $x \in \mathscr{D}_{12}$, such that $\mathscr{D}_{12}$ is a core for A and such that $B_1(\mathscr{D}_{12}) \subseteq \mathscr{D}_2$ and $B_2(\mathscr{D}_{12}) \subseteq \mathscr{D}_1$.*

3. *$\|B_ix\| \leq \lambda\|(A+i)x\|$ for some $\lambda > 0$ and for all $x \in \mathscr{D}_i$ for $i = 1,2$.*

Then $\overline{B_1}$ and $\overline{B_2}$ are strongly commuting self-adjoint operators.

Proof. The fact that $\overline{B_i}$ are self-adjoint, is a consequence of Lemma 5.1.8. Thus we only need to prove $\overline{B_1}$ and $\overline{B_2}$ commute strongly. We use the same notation as in previous proof.

By Lemma 5.1.8, we know that $B_i = |A-i|F_i$ on $\mathscr{D}_i$ with $i = 1,2$. Then we get for $x \in \mathscr{D}_{12}$

$$\begin{aligned} B_1B_2x &= |A-i|F_1|A-i|F_2x = |A-i|F_1F_2|A-i|x \\ &= B_2B_1x = |A-i|F_2|A-i|F_1x = |A-i|F_2F_1|A-i|x \end{aligned}$$

where we used that $F = F^*$ commutes with $|A-i|$. Because $\ker|A-i| = (0)$, the previous calculations implies $F_1F_2|A-i|x = F_2F_1|A-i|x$ for any $x \in \mathscr{D}_{12}$. Since $\mathscr{D}_{12}$ is a core of A, $|A-i|(\mathscr{D}_{12})$ lies dense in $\mathscr{H}$. Thus we can conclude that $F_1F_2 = F_2F_1$. Then we consider $\overline{B_1}$. Because $\overline{B_1} = |A-i|F_1$ we get for $y \in \mathrm{Dom}(\overline{B_1})$:

$$F_2\overline{B_1}y = F_2|A-i|F_1y = |A-i|F_1F_2y = \overline{B_1}F_2y$$

because, again $F_2 = F_2^*$ commutes with $|A-i|$. Therefore F_2 and $\overline{B_1}$ commute on $\mathrm{Dom}(\overline{B_1})$. Because $F_2 \in \mathscr{B}(\mathscr{H})$, F_2 commutes with $(\overline{B_1}+i)^{-1}$. Therefore, it holds that

$$\overline{B_2}(\overline{B_1}+i)^{-1} = |A-i|F_2(\overline{B_1}+i)^{-1} = (\overline{B_1}+i)^{-1}|A-i|F_2 = (\overline{B_1}+i)^{-1}\overline{B_2}$$

By Lemma 5.1.6 we have that $\overline{B_1}$ and $\overline{B_2}$ strongly commute.

□

Now we are ready for the theorem of Nelson. The theorem we are discussing here is slightly adapted from the one Nelson proved. He proved it for symmetric operators with an dense linear subset contained in a lot of different domains. We will only be considering self-adjoint operators, and a slightly different set.

Theorem 5.1.10 (Theorem of Nelson)**.** *[5, Cor. 9.2, p. 603] Let B_1, B_2 be two self-adjoint operators on a Hilbert space $\mathscr{H}$, and let $\mathscr{D}$ be a dense linear subspace of $\mathscr{H}$ such that:*

1. *$\mathscr{D}$ is contained in $\mathrm{Dom}(B_i)$ for $i = 1, 2$,*
2. *$B_i(\mathscr{D}) \subseteq \mathscr{D}$ for $i = 1, 2$,*
3. *$\mathscr{D}$ is a core of B_1 and B_2*
4. *$B_1 B_2 x = B_2 B_1 x$ for all $x \in \mathscr{D}$.*

If $(B_1^2 + B_2^2)|_{\mathscr{D}}$ is essentially self-adjoint, then B_1 and B_2 strongly commute.

Proof. We wish to use Lemma 5.1.9 to prove this theorem. Let B_1 and B_2 be as in Lemma 5.1.9. Define $A := \overline{B_1^2 + B_2^2|_{\mathscr{D}}}$. Then A is self-adjoint, by assumption. To show that the first requirement holds, note that $\mathscr{D} \subseteq \mathrm{Dom}(B_1)$ and $B_1(\mathscr{D}) \subseteq \mathscr{D}$. Therefore $\mathscr{D} \subseteq \mathrm{Dom}(B_1^2)$. By the same reasoning we find $\mathscr{D} \subseteq \mathrm{Dom}(B_2^2)$. Hence $\mathscr{D} \subseteq \mathrm{Dom}(B_1^2 + B_2^2) \subseteq \mathrm{Dom}(A)$. Additionally, we see that $A(\mathscr{D}) = (B_1^2 + B_2^2)(\mathscr{D}) \subseteq \mathscr{D}$, and therefore $\mathscr{D} \subseteq \mathrm{Dom}(B_j A)$ and $\mathscr{D} \subseteq \mathrm{Dom}(AB_j)$ for $j = 1, 2$. Therefore we see that

$$\mathscr{D} \subseteq \mathrm{Dom}(AB_j) \cap \mathrm{Dom}(B_j A) \cap \mathrm{Dom}(A) \cap \mathrm{Dom}(B_j)$$

for $j = 1, 2$. By assumption we also know that $B_1 B_2 x = B_2 B_1 x$, and so

$$AB_1 x = (B_1^2 + B_2^2) B_1 x = B_1 (B_1^2 + B_2^2) x = B_1 A x$$

Finally, because $(B_1^2 + B_2^2)|_{\mathscr{D}}$ is essentially self-adjoint, $\mathscr{D}$ is a core of $B_1^2 + B_2^2$ and so it is a core of A. By assumption, $\mathscr{D}$ is a core of B_1 and B_2. Therefore the first requirement of Lemma 5.1.9 holds for $\mathscr{D}_1 = \mathscr{D}_2 = \mathscr{D}$.

Because $B_1(\mathscr{D}) \subseteq \mathscr{D}$ and $\mathscr{D} \subseteq \mathrm{Dom}(B_2)$, we find $\mathscr{D} \subseteq \mathrm{Dom}(B_1 B_2)$. By the same reasoning, $\mathscr{D} \subseteq \mathrm{Dom}(B_2 B_1)$. Therefore $\mathscr{D} \subseteq \mathrm{Dom}(B_1 B_2) \cap \mathrm{Dom}(B_2 B_1)$. We conclude that if we define $\mathscr{D}_{12} := \mathscr{D}$, we also see that the second requirement of Lemma 5.1.9 holds.

Finally, we need to show that $\|B_i x\| \leq \|(A + i)x\|$ for $i = 1, 2$ and for all $x \in \mathscr{D}$. If we can show this to be true, then we can use Lemma 5.1.9 and we are done. We will prove the inequality for B_1. We note that if $x \in \mathscr{D}$ we get

$$\begin{aligned}
\|(A + i)x\|^2 &= \|(B_1^2 + B_2^2 + i)x\|^2 \\
&= \langle \left(B_1^2 + B_2^2 + i\right) x, \left(B_1^2 + B_2^2 + i\right) x \rangle \\
&= \|(B_1^2 + i)x\|^2 + \|B_2^2 x\|^2 + 2\mathrm{Re}(\langle (B_1^2 + i)x, B_2^2 x \rangle) \\
&\geq \|(B_1^2 + i)x\|^2 + 2\mathrm{Re}(\langle B_1^2 x, B_2^2 x \rangle + i \langle x, B_2^2 x \rangle)
\end{aligned}$$

Because $B_2(\mathscr{D}) \subseteq \mathscr{D}$ we have $B_1 B_2^2 x = B_2^2 B_1 x$, and thus we can conclude

$$\begin{aligned}
\|(A + i)x\|^2 &= \|(B_1^2 + i)x\|^2 + 2\mathrm{Re}(\langle B_1 x, B_1 B_2^2 x \rangle + i \langle B_2 x, B_2 x \rangle) \\
&= \|(B_1^2 + i)x\|^2 + 2\mathrm{Re}(\langle B_2 B_1 x, B_2 B_1 x \rangle + i \langle B_2 x, B_2 x \rangle) \\
&= \|(B_1^2 + i)x\|^2 + 2\mathrm{Re}(\|B_2 B_1 x\|^2 + i \|B_2 x\|^2) \\
&= \|(B_1^2 + i)x\|^2 + \|B_2 B_1 x\|^2 \geq \|(B_1^2 + i)x\|^2
\end{aligned}$$

So we found $\|(A+i)x\| \geq \|(B_1^2+i)x\|$ for $x \in \mathscr{D}$. If we can prove $\|(B_1^2+i)x\| \geq \|B_1x\|$ for $x \in \mathscr{D}$, we are done. Note that B_1 is a self-adjoint operator, so $B_1 = \int_{\mathbb{R}} \lambda \, dE_1$ where E_1 is the spectral measure of B_1. Thus

$$B_1 = B_{[-1,1]} + B_\infty$$

where $B_{[-1,1]} := \int_{[-1,1]} \lambda \, dE_1$ and $B_\infty = \int_{\sigma(B_1)\setminus[-1,1]} \lambda \, dE_1$. Because $[-1,1]$ and $\sigma(B_1)\setminus[-1,1]$ are disjoint sets, we get that for any Borel set $U \subseteq [-1,1]$ and any Borel set $V \subset \sigma(B_1)\setminus[-1,1]$ that $E_1(U)E_1(V) = E_1(V)E_1(U) = 0$. Therefore is $B_{[-1,1]}B_\infty = B_\infty B_{[-1,1]} = 0$. Hence

$$B_1^2 = B_{[-1,1]}^2 + B_\infty^2$$

Additionally, we note that for any set $K \subseteq \mathbb{R}$ we have $\int_K \lambda \, dE(\lambda) = \int_{\sigma(B_1)} \chi_K \lambda \, dE(\lambda)$. Therefore we see that

$$B_{[-1,1]}^* = \int_{\sigma(B_1)} \overline{\chi_{[-1,1]}\lambda} \, dE(\lambda) = \int_{\sigma(B_1)} \chi_{[-1,1]}\lambda \, dE(\lambda) = B_{[-1,1]}$$

$$B_\infty^* = \int_{\sigma(B_1)} \overline{\chi_{\sigma(B_1)\setminus[-1,1]}\lambda} \, dE = \int_{\sigma(B_1)} \chi_{\sigma(B_1)\setminus[-1,1]}\lambda \, dE = B_\infty$$

Thus $B_{[-1,1]}$ and B_∞ are self-adjoint. Therefore we get for $x \in \mathscr{D}$

$$\begin{aligned}
\|B_1x\|^2 = \|B_{[-1,1]}x\|^2 + \|B_\infty x\|^2 &= \int_{[-1,1]} \lambda^2 \, d\langle E_1(\lambda)x, x\rangle + \int_{\mathbb{R}\setminus[-1,1]} |\lambda|^2 \, d\langle E_1(\lambda)x, x\rangle \\
&\leq \sup_{\lambda\in[-1,1]} \{\lambda\} \cdot \langle E_1([-1,1])x, x\rangle + \int_{\mathbb{R}\setminus[-1,1]} |\lambda|^4 \langle E_1(\lambda)x, x\rangle \\
&= \langle E_1([-1,1])x, x\rangle + \|B_\infty^2 x\|^2 \leq \|E_1([-1,1])\|\|x\|^2 + \|B_\infty^2 x\|^2 \\
&= \|x\|^2 + \|B_\infty^2 x\|^2 \leq \|x\|^2 + \|B_\infty^2 x\|^2 + \|B_{[-1,1]}^2 x\|^2 = \|(B_{[-1,1]}^2 + B_\infty^2 + i)x\|^2
\end{aligned}$$

The last equality sign is because $B_1^2 = B_{[-1,1]}^2 + B_\infty^2$ is symmetric, so $\|(B_1^2+i)x\|^2 = \|B_1^2x\|^2 + i^2\|x\|^2$. If we read the whole equation, we see that $\|B_1x\|^2 \leq \|(B_1^2+i)x\|^2$ for $x \in \mathscr{D}$. Same holds for B_2 if we swap the 1 for a 2. So we can use Lemma 5.1.9 to conclude that B_1 and B_2 strongly commute. □

In conclusion, if we can find that we can apply Nelson's theorem for any two self-adjoint operators A and B, we know by Proposition 5.1.2 that for any two Borel measurable functions $f : \sigma(A) \to \mathbb{C}$ and $g : \sigma(B) \to \mathbb{C}$ we have that $f(A)$ and $g(B)$ commute. We will consider two examples to apply this to.

6 Position and momentum operators

In this section, we consider two examples. The reader might be familiar with these concepts, for these operators are being used in Quantum Mechanics.

6.1 Tempered distributions

Before we formulate the examples, we give the notion of a distribution. Distributions turn out to be necessary for these examples, because we wish to differentiate functions which are not differentiable by definition. Because the focus of this thesis is not distribution theory, we will prove only a few theorems. The reference to the proofs is within brackets after the theorem for those who are interested. The reader who are already familiar with distribution theory and Fourier analysis on distributions, can continue reading Section 6.3.

Definition 6.1.1. Let $X \subseteq \mathbb{R}^n$. We define $\mathscr{D}(X) := C_0^\infty(X)$; the space of infinitely differentiable $\phi : X \to \mathbb{C}$ with compact support. A sequence $(\phi_j)_j \subseteq \mathscr{D}(X)$ is said to *converges to ϕ in $\mathscr{D}(X)$*, if:

1. there exists a compact $K \subseteq X$ such that $\operatorname{supp}(\phi_j) \subset K$ for every j
2. for every multi-index $\alpha \in \mathbb{N}^n$ the sequence $(\partial^\alpha \phi_j)_j$ converges uniformly on X to $\partial^\alpha \phi$

Definition 6.1.2. Let $\mathscr{D}(X)$ as in Defintion 6.1.1. We define a distribution on X as a linear map $u : \mathscr{D}(X) \to \mathbb{C}$, such that if $\lim_{j\to\infty} \phi_j = \phi$ in $\mathscr{D}(X)$ then $\lim_{j\to\infty} u(\phi) = u(\phi)$. We write the space of distributions as $\mathscr{D}'(X)$.

Lemma 6.1.3. *[2, p. 37] Let $f \in L_{loc}(X)$. So $f : X \to \mathbb{C}$ is locally integrable. Then f can be interpreted as a distribution via the mapping* $\mathrm{test}\cdot : L_{loc}(X) \to \mathscr{D}'(X)$ *defined by*

$$(\mathrm{test} f)(\phi) = \int_{\mathbb{R}^n} f(x)\phi(x)dx$$

Most of the times we will omit the test, and say f is a distribution.

Since $L^2(\mathbb{R}^n) \subseteq L_{loc}(\mathbb{R}^n)$, we know that if $u \in L^2(\mathbb{R}^n)$ then $u \in \mathscr{D}'(\mathbb{R}^n)$. But does a distribution come from a function?

Lemma 6.1.4. *Let $u \in \mathscr{D}'(\mathbb{R}^n)$. Then $u \in L^2(\mathbb{R}^n) \iff$ there exists a $C > 0$ such that $|u(\phi)| \leq C\|\phi\|_{L^2(\mathbb{R}^n)}$ for any $\phi \in \mathscr{D}(\mathbb{R}^n)$*

Proof. Let $u \in L^2(\mathbb{R}^n)$. So $\int_{\mathbb{R}^n} |u(x)|^2 dx < \infty$. Let $\phi \in \mathscr{D}(\mathbb{R}^n)$, then there exists a compact set $K \subseteq \mathbb{R}^n$ such that $\operatorname{supp}(\phi) \subseteq K$. We claim that $\phi \in L^2(\mathbb{R}^n)$. We see that

$$\int_{\mathbb{R}^n} |\phi(x)|^2 dx = \int_K |\phi(x)|^2 dx \leq \sup_{x\in K}\{|\phi(x)|^2\} \int_K dx < \infty$$

because $|\phi|^2$ is a continuous function on K, so it is bounded on K. So $\phi \in L^2(\mathbb{R}^n)$. Then we use Hölder's inequality to conclude

$$|u(\phi)| = \left|\int_{\mathbb{R}^n} u(x)\phi(x)dx\right| \leq \int_{\mathbb{R}^n} |u(x)\phi(x)|dx = \|u\phi\|_{L^1(\mathbb{R}^n)} \leq \|u\|_{L^2(\mathbb{R}^n)}\|\phi\|_{L^2(\mathbb{R}^n)}$$

This is true for any $\phi \in \mathscr{D}'(\mathbb{R}^n)$, so the first part is proven.

Next, let $|u(\phi)| \leq C\|\phi\|_{L^2(\mathbb{R}^n)}$ for all $\phi \in \mathscr{D}(\mathbb{R}^n)$. Because $\mathscr{D}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, it is possible to extend u to a bounded functional $\tilde{u}$ on $L^2(\mathbb{R}^n)$. Hence by the Riesz Representation Theorem, there exists a unique $g \in L^2(\mathbb{R}^n)$ such that $\tilde{u} = \langle \cdot, g\rangle$. So for $\phi \in \mathscr{D}(\mathbb{R}^n)$ we have

$$\tilde{u}(\phi) = u(\phi) = \langle \phi, g\rangle = \int_{\mathbb{R}^n} \phi(x)\overline{g(x)}dx = (\text{test } \overline{g})(\phi)$$

So $u = \text{test } \overline{g}$, and thus $u \in L^2(\mathbb{R}^n)$. □

If f is a continuously differentiable function on $X \subseteq \mathbb{R}^n$, then

$$(\text{test}\partial_j f)(\phi) = \int_X \frac{\partial f}{\partial x_i}\phi \, dx = -\int_X f\frac{\partial \phi}{\partial x_i}dx = -(\text{test}f)(\partial_j\phi)$$

because the boundary term is absent, as $\phi(x) = 0$ for sufficiently large x. Motivated by this, we define a differentiation for distributions:

Definition 6.1.5. If $X \subseteq \mathbb{R}^n$, we define for an arbitrary distribution u on X:

$$\partial_j u(\phi) = -u(\partial_j\phi) \qquad (1 \leq j \leq n, \phi \in \mathscr{D}(X))$$

Here $\partial_j := \frac{\partial}{\partial x_j}$, the differential operator with respect to the jth component.

With this notion any distribution, and thus any locally integrable function, is infinitely differentiable. This definition of differentiable is consistent with our regular definition of differentiation on functions; in other words, if f is continuously differentiable, then $\text{test}(\partial_j f) = \partial_j(\text{test}f)$ for $1 \leq j \leq n$.

6.2 Fourier transform

Now that we have introduced the basics of distribution theory, we can talk about Fourier transforms. In order to do so, we first introduce the space on which Fourier transform acts.

Definition 6.2.1. A function ϕ on $\mathbb{R}^n$ is said to be *rapidly decreasing* if for every multi-index $\beta \in \mathbb{N}^n$ the function $x \mapsto x^\beta\phi(x)$ is bounded on $\mathbb{R}^n$. We define $\mathscr{S} = \mathscr{S}(\mathbb{R}^n)$ as the space of all $\phi \in C^\infty(\mathbb{R}^n)$ such that $\partial^\alpha\phi$ is rapidly decreasing for every multi-index α. This space is called the space of *Schwartz-functions*.

If $(\phi_j)_j$ is a sequence in $\mathscr{S}$ and $\phi \in \mathscr{S}$, then ϕ_j is said to *converge to ϕ in $\mathscr{S}$* if for all multi-indices α, β the sequence of functions $\left(x^\beta\partial^\alpha\phi_j\right)_j$ converges uniformly on all of $\mathbb{R}^n$ to $x^\beta\partial^\alpha\phi$.

Definition 6.2.2. We define a *tempered distribution* as a linear mapping $u : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ in such a way that if $\lim_{n\to\infty}\phi_n = \phi$ in $\mathscr{S}(\mathbb{R}^n)$, then $\lim_{n\to\infty} u(\phi_n) = u(\phi)$. The space of tempered distributions is denoted by $\mathscr{S}'(\mathbb{R}^n) = \mathscr{S}'$.

We say that a sequence $(u_j)_j \subseteq \mathscr{S}'(\mathbb{R}^n)$ *converges to u in $\mathscr{S}'(\mathbb{R}^n)$*, denoted $\lim_{j\to\infty} u_j = u$, if $u \in \mathscr{S}'(\mathbb{R}^n)$ and $\lim_{j\to\infty} u_j(\phi) = u(\phi)$ for every $\phi \in \mathscr{S}(\mathbb{R}^n)$. In other words, if $(u_j)_j$ converges pointwise to $u \in \mathscr{S}'(\mathbb{R}^n)$.

With this new definition, we have $\mathscr{D}(\mathbb{R}^n) \subseteq \mathscr{S}(\mathbb{R}^n) \subseteq C^\infty(\mathbb{R}^n)$. Since $\mathscr{D}(\mathbb{R})$ lies dense in $C^\infty(\mathbb{R}^n)$, we have that $\mathscr{S}(\mathbb{R}^n)$ lies dense in $C^\infty(\mathbb{R}^n)$. Also note that the identity mapping $\mathrm{Id} : \mathscr{D}(\mathbb{R}^n) \to \mathscr{S}$ is continuous. Additionally, we have the continuous inclusion $\mathscr{S}' \subset \mathscr{D}'(\mathbb{R}^n)$ (see $[2, p.\ 189]$ for example).

Definition 6.2.3. On $\mathscr{S}$ we define the Fourier transform $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$, by

$$(\mathscr{F}\phi)(\xi) = \int_{\mathbb{R}^n} e^{-i\langle \xi, x\rangle_{\mathbb{R}^n}} \phi(x)dx$$

where $\langle \xi, x\rangle_{\mathbb{R}^n} = \sum_{i=1}^n \xi_i x_i$, the standard inner product on $\mathbb{R}^n$.

Since ϕ is a rapidly decreasing function, this integral is well defined for any $\xi \in \mathbb{R}^n$. One of the more promising features of the Fourier transform is that differential operators are transformed into polynomial operators and the other way around, as the following lemma shows. We define:

$$D_j := -i\partial_j \quad (1 \leq j \leq n) \tag{30}$$

Then we can use this definition to state the lemma.

Lemma 6.2.4. *[2, p. 183] The Fourier transform $\mathscr{F} : u \mapsto \mathscr{F}u$ defines a continuous linear mapping from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}(\mathbb{R}^n)$. For every $1 \leq j \leq n$, $\phi \in \mathscr{S}(\mathbb{R}^n)$ and $\xi, a \in \mathbb{R}^n$ we have*

$$\mathscr{F}(D_j\phi)(\xi) = \xi_j \mathscr{F}(\phi)(\xi) \qquad\qquad \mathscr{F}(x_j\phi)(\xi) = -D_j\mathscr{F}(\phi)(\xi)$$

Example 6.2.5. For $a \in \mathbb{C}$ with $\mathrm{Re}(a) > 0$ we define $u_a(x) = e^{-ax^2/2}$ for $x \in \mathbb{R}$. Then surely, $u_a \in \mathscr{S}(\mathbb{R})$. Thus $\mathscr{F}u_a \in \mathscr{S}(\mathbb{R})$. Note that $Du_a = iaxu_a$, and so taking the Fourier transform on both sides gives:

$$D\mathscr{F}u_a = \mathscr{F}(-xu_a) = \mathscr{F}(i\frac{1}{a}Du_a) = i\frac{1}{a}\xi\mathscr{F}u_a$$

This is an ordinary differential equation, and thus we get $\mathscr{F}u_a = A(a)u_{\frac{1}{a}}$ with $A(a) = (\mathscr{F}u_a)(0) = \int_{\mathbb{R}} u_a(x)dx = \sqrt{\frac{2\pi}{a}}$. So in conclusion $\mathscr{F}u_a(\xi) = \sqrt{\frac{2\pi}{a}}e^{-\frac{\xi^2}{2a}}$.

By this, we see that, if we apply $\mathscr{F}$ again, we get

$$\mathscr{F}(\mathscr{F}u_a) = \mathscr{F}(A(a)u_{\frac{1}{a}}) = A(a)\mathscr{F}u_{\frac{1}{a}} = A(a)B(a)u_a$$

with $B(a) = \int_{\mathbb{R}} u_{\frac{1}{a}}(x)dx = \int_{\mathbb{R}} e^{-\frac{x^2}{2a}} = \sqrt{2\pi a}$. So we see that $\mathscr{F}\mathscr{F}u_a = 2\pi u_a$ $\oslash$

Note that for this example, we got the inverse of this particular Fourier transform. But the example is not unique, as the following theorem states

Theorem 6.2.6. *[2, Thm. 14.13, p. 185] The Fourier transform $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is bijective, with inverse $\mathscr{F}^{-1} = (2\pi)^{-n}S \circ \mathscr{F} = (2\pi)^{-n}\mathscr{F} \circ S$, where $S : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$, $S(\phi)(x) = \phi(-x)$. This can be written as, for $\phi \in \mathscr{S}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$:*

$$\phi(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\langle \xi, x\rangle_{\mathbb{R}^n}} \mathscr{F}\phi(\xi)d\xi \tag{31}$$

Now that we have some theorems about the Fourier transforms on $\mathscr{S}$, we continue to define the Fourier transform on $\mathscr{S}'$.

Definition 6.2.7. If $u \in \mathscr{S}'$, we define its Fourier transform $\mathscr{F}u \in \mathscr{S}'$ as, for $\phi \in \mathscr{S}$,

$$\mathscr{F}u(\phi) = u(\mathscr{F}\phi)$$

Note that this definition coincides with the definition we gave for $\phi \in \mathscr{S} \subset \mathscr{S}'$ in Definition 6.2.3. For if we consider $\phi \in \mathscr{S}$ as a distribution, we have

$$\begin{aligned}(\mathscr{F}\phi)(\psi) = \phi(\mathscr{F}\psi) &= \int_{\mathbb{R}^n} \phi(\xi)\mathscr{F}\psi(\xi)d\xi \\ &= \int_{\mathbb{R}^n} \phi(\xi) \int_{\mathbb{R}^n} e^{-i\langle \xi,x\rangle_{\mathbb{R}^n}}\psi(x)dxd\xi \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i\langle \xi,x\rangle_{\mathbb{R}^n}}\phi(\xi)\psi(x)d\xi dx = \int_{\mathbb{R}^n} (\mathscr{F}\phi)(x)\psi(x)dx\end{aligned}$$

As it happens a lot in the theory of distributions, many theorems that hold for functions, also apply (although slightly adjusted) for distributions. The most important in case of Fourier theory is the next theorem.

Theorem 6.2.8. *[2, Thm. 14.14, p. 191] For every $u \in \mathscr{S}(\mathbb{R}^n)$ we have $\mathscr{F}u \in \mathscr{S}'(\mathbb{R}^n)$. The mapping $\mathscr{F} : u \mapsto \mathscr{F}u$ is a continuous linear mapping from $\mathscr{S}'(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$. Also, for every $u \in \mathscr{S}'(\mathbb{R}^n)$ and $1 \leq j \leq n$ we have*

$$\mathscr{F}(D_j u) = \xi_j \mathscr{F}u$$

$$\mathscr{F}(x_j u) = -D_j \mathscr{F}u$$

Finally, $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(R^n)$ is bijective with inverse equal to $(2\pi)^{-n}\mathscr{F} \circ S' = (2\pi)^{-n}S' \circ \mathscr{F}$ where $S' : \mathscr{S}' \to \mathscr{S}'$ by $S'(u) = u \circ S$ with S as in Lemma 6.2.6.

Lemma 6.2.9. *[2, p. 190] We have the following inclusions:*

$$\mathscr{D}(\mathbb{R}^n) \subseteq \mathscr{S}(\mathbb{R}^n) \subseteq L^2(\mathbb{R}^n) \subseteq \mathscr{S}'(\mathbb{R}^n)$$

These inclusions are also continuous, or in other words, convergence in one space includes convergence in the bigger space.

Fourier Theory turns out to be useful to study $L^2(\mathbb{R}^n)$. In fact, with the notion of distributions we get that the Fourier transform is a unitary isomorphism.

Theorem 6.2.10. *[2, Thm. 14.32, p. 196] If u belongs to $L^2(\mathbb{R}^n)$, then $\mathscr{F}u \in L^2(\mathbb{R}^n)$. Also, if $\phi, \psi \in L^2(\mathbb{R}^n)$ then*

$$\langle \mathscr{F}\phi, \mathscr{F}\psi\rangle_{L^2(\mathbb{R}^n)} = \langle \phi, \psi\rangle_{L^2(\mathbb{R}^n)}$$

It follows that the restriction of $\tilde{\mathscr{F}} := (2\pi)^{n/2}\mathscr{F}$ to $L^2(\mathbb{R}^n)$ defines a unitary isomorphism on $L^2(\mathbb{R}^n)$.

Corollary 6.2.11. *Let $u \in L^2(\mathbb{R}^n)$.*

1. *If* $x_j u \in L^2(\mathbb{R}^n)$*, then* $D_j(\mathscr{F}(u)) \in L^2(\mathbb{R}^n)$.
2. *If* $D_j u \in L^2(\mathbb{R}^n)$*, then* $x_j \mathscr{F}(u) \in L^2(\mathbb{R})$.

Proof. 1): Let $\iota : L^2(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ be the inclusion mapping. Then

$$\iota(\mathscr{F}(x_j u)) = \mathscr{F}\left(\iota(x_j u)\right) = \mathscr{F}(x_j \iota(u)) = -D_j \mathscr{F}(\iota(u)) = -D_j \iota(\mathscr{F}(u)) = \iota(-D_j \mathscr{F}(u))$$

Therefore we can conclude $\iota(\mathscr{F}(x_j u)) = \iota(-D_j \mathscr{F}(u))$ in $\mathscr{S}(\mathbb{R}^n)$. Therefore, $\mathscr{F}(x_j u) = -D\mathscr{F}(u)$ in L^2. So $D_j \mathscr{F}(u) \in L^2(\mathbb{R})$.

2): Let ι be the inclusion map again. Then we see

$$\iota(\mathscr{F}(D_j u)) = \mathscr{F}(\iota(D_j u)) = \mathscr{F}(D_j(\iota(u))) = x_j \mathscr{F}(\iota(u)) = x_j \iota(\mathscr{F}(u)) = \iota(x_j \mathscr{F}(u))$$

So $\iota(\mathscr{F}(D_j u)) = \iota(x_j \mathscr{F}(u))$ in $\mathscr{S}'(\mathbb{R}^n)$. Therefore, $\mathscr{F}(D_j u) = x_j \mathscr{F}(u)$ and $x_j \mathscr{F}(u) \in L^2(\mathbb{R}^n)$. □

6.3 Position and momentum operators

After quite some distribution theory, we finally consider the two examples. These operators are important in the quantum mechanics.

First, we consider the Hilbert space $\mathscr{H} = L^2(\mathbb{R}^n)$. Consider $f \in \mathscr{H}$. In the classical sense, the derivative of f may or may not exist. For example, let $B := B(0, r)$ be an open ball of $\mathbb{R}^n$ of radius r centered around the origin. Then $\chi_B \in L^2(\mathbb{R}^n)$ but the derivative is not well defined everywhere. Nevertheless, the derivative does exist if we consider χ_B in a distributional sense, by Definition 6.1.5. The only problem with this way of doing it, is that the derivative might not be a function itself, but rather a distribution. And if it is a function, this function might not be square integrable. Thus the operator $\partial_j : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ has a certain domain on which it is defined. We are now ready for the definition of the momentum and position operators in $L^2(\mathbb{R}^n)$. First we consider $n = 1$:

Definition 6.3.1. Let $\mathscr{H} = L^2(\mathbb{R})$, and define the linear operator $\hat{p} : \mathscr{H} \to \mathscr{H}$ by $\hat{p}(u) = i\frac{d}{dx}u$. The domain on which this operator is defined is

$$\mathrm{Dom}(\hat{p}) = \{f \in L^2(\mathbb{R}) |\, i\frac{d}{dx}f \in L^2(\mathbb{R}) \text{ in the distributional sense}\}$$

This operator is called the *momentum operator*.

By Example 2.1.4 we also need a domain for the operator that is defined as $\hat{x}$.

Definition 6.3.2. Let $\mathscr{H} = L^2(\mathbb{R})$. Define the linear operator $\hat{x} : \mathscr{H} \to \mathscr{H}$ by $\hat{x}(u) = xu$. The domain on which this operator is defined is $\mathrm{Dom}(\hat{x}) = \{f \in L^2(\mathbb{R}) | xf \in L^2(\mathbb{R})\}$.

Our first goal is to show that both $\hat{x}$ and $\hat{p}$ are self-adjoint. We start with the operator $\hat{x}$.

Lemma 6.3.3. *The operator* $\hat{x}$ *is a self-adjoint operator*

Proof. First, we must show that $\mathrm{Dom}(\hat{x})$ lies dense in $L^2(\mathbb{R})$. Note that $\mathscr{D}(\mathbb{R}) \subseteq \mathrm{Dom}(\hat{x}) \subseteq L^2(\mathbb{R})$. Since $\mathscr{D}(\mathbb{R})$ lies dense in $L^2(\mathbb{R})$ we have that $\mathrm{Dom}(\hat{x})$ lies dense in $L^2(\mathbb{R})$.

Next, we note that if $f, g \in \mathrm{Dom}(\hat{x})$, we have

$$\langle \hat{x}(f), g \rangle = \int_{\mathbb{R}} x f(x)\overline{g(x)}dx = \int_{\mathbb{R}} f(x)\overline{xg(x)}dx = \langle f, \hat{x}(g) \rangle$$

So $\hat{x}$ is symmetric. Thus the only thing left to show is $\mathrm{Dom}(\hat{x}) = \mathrm{Dom}(\hat{x}^*)$. Since $\hat{x}$ is symmetric, we know by Lemma 2.4.5 that $\hat{x} \subseteq \hat{x}^*$. Therefore $\mathrm{Dom}(\hat{x}) \subseteq \mathrm{Dom}(\hat{x}^*)$.

On the other hand, given $f \in \mathrm{Dom}(\hat{x}^*)$, we have that the mapping $\phi \mapsto \langle \hat{x}(\phi), f \rangle$ is a bounded linear functional op $\mathrm{Dom}(\hat{x})$. Thus, there exists a $C > 0$ such that for any $\phi \in \mathrm{Dom}(\hat{x})$ we have $|\langle \hat{x}(\phi), f \rangle| \leq C\|\phi\|_{L^2(\mathbb{R})}$. This is surely true for $\phi \in \mathscr{D}(\mathbb{R})$. Also note that $f \in L^2(\mathbb{R})$, so $f, \overline{f} \in L_{loc}(\mathbb{R})$, and thus $\overline{xf} \in L_{loc}(\mathbb{R})$. And so the distribution $\mathrm{test}(\overline{xf}) = \overline{xf}$ is well defined. Thus

$$\overline{xf}(\phi) = \int_{\mathbb{R}} \phi(x)(\overline{xf})(x)dx = \int_{\mathbb{R}} x\phi(x)\overline{f}(x)dx = \langle \hat{x}(\phi), f \rangle$$

Therefore, $|\overline{xf}(\phi)| = |\langle \hat{x}(\phi), f \rangle| \leq C\|\phi\|_{L^2}$. Then, by Lemma 6.1.4 we find that $xf \in L^2(\mathbb{R})$. But this means that $f \in \mathrm{Dom}(\hat{x})$. So $\mathrm{Dom}(\hat{x}) = \mathrm{Dom}(\hat{x}^*)$, and thus is $\hat{x}$ self-adjoint □

Next, we look at the linear operator $\hat{p}$.

Lemma 6.3.4. *$\hat{p}$ is self-adjoint.*

Proof. We consider the following claim.

Claim 1. $\mathscr{F}(\mathrm{Dom}(\hat{x})) = \mathrm{Dom}(\hat{p})$.

Proof. Consider $f \in \mathrm{Dom}(\hat{x})$. Then $xf \in L^2(\mathbb{R})$ and thus $\mathscr{F}(xf) \in L^2(\mathbb{R})$. By Lemma 6.2.4 we see $\mathscr{F}(xf) = -D\mathscr{F}(f)$. So $-i\frac{d}{dx}\mathscr{F}(f) \in L^2(\mathbb{R})$ and so $\mathscr{F}(f) \in \mathrm{Dom}(\hat{p})$. For the other inclusion, we follow the same reasoning, only now backwards. □

First we prove that $\hat{p}$ is symmetric. Consider $f, g \in \mathrm{Dom}(\hat{p})$. Then there exists $h \in \mathrm{Dom}(\hat{x})$ such that $\mathscr{F}(h) = f$. Then, by Corollary 6.2.11,

$$\begin{aligned}
\langle \hat{p}(f), g \rangle &= \langle \hat{p}(\mathscr{F}h), g \rangle = \langle \mathscr{F}(\hat{x}(h)), g \rangle \\
&= \langle \hat{x}(h), \mathscr{F}^{-1}g \rangle = \langle h, \hat{x}(\mathscr{F}^{-1}g) \rangle \\
&= \langle h, \hat{x} \circ S \circ \mathscr{F}g) \rangle = \langle h, -S \circ \hat{x} \circ \mathscr{F}(g) \rangle \\
&= \langle h, \mathscr{F}^{-1}(\hat{p}(g)) \rangle = \langle \mathscr{F}(h), \hat{p}(g) \rangle = \langle f, \hat{p}(g) \rangle
\end{aligned}$$

So $\hat{p}$ is symmetric. Next, let $f \in \mathrm{Dom}(\hat{p}^*)$. Then the mapping $g \mapsto \langle \hat{p}(g), f \rangle$ is bounded for $g \in \mathrm{Dom}(\hat{p})$. Because $g \in \mathrm{Dom}(\hat{p})$, there is a $h \in \mathrm{Dom}(\hat{x})$ such that $\mathscr{F}(h) = g$. Therefore it is equivalent to say that $h \mapsto \langle \hat{p}(\mathscr{F}h), f \rangle$ is bounded for $h \in \mathrm{Dom}(\hat{x})$. But note that

$$\langle \hat{p}(\mathscr{F}h), f \rangle = \langle \mathscr{F}(\hat{x}(h)), f \rangle = \langle \hat{x}(h), \mathscr{F}^{-1}f \rangle$$

So the mapping $h \mapsto \langle \hat{x}(h), \mathscr{F}^{-1}f \rangle$ is bounded on $\mathrm{Dom}(\hat{x})$. Therefore we can conclude $\mathscr{F}^{-1}f \in \mathrm{Dom}(\hat{x}^*) = \mathrm{Dom}(\hat{x})$. But now, $f \in \mathscr{F}(\mathrm{Dom}(\hat{x})) = \mathrm{Dom}(\hat{p})$. So $\hat{p}$ is self-adjoint.

□

One might ask whether we can generalize this to $\mathbb{R}^n$. In fact, we can. Before we do so, we generalize the operators we are discussing to $L^2(\mathbb{R}^n)$.

Definition 6.3.5. Let $\mathscr{H} = L^2(\mathbb{R}^n)$, and define the linear operator $\hat{p}_j : \mathrm{Dom}(\hat{p}_j) \to L^2(\mathbb{R}^n)$ by $\hat{p}_j(u) = i\partial_j u$ for $j = 1, \dots, n$. The domain on which this operator is defined is

$$\mathrm{Dom}(\hat{p}_j) = \{f \in L^2 | \; i\partial_j f \in L^2 \text{ in the distributional sense}\}$$

This operator is called *the momentum operator in the j-th direction*.

Also define $\hat{x}_j : \mathrm{Dom}(\hat{x}_j) \to L^2(\mathbb{R}^n)$ by $\hat{x}_j(u) = x_j u$ where $x_j : \mathbb{R}^n \to \mathbb{R}, (x_1, \dots, x_n) \mapsto x_j$. The domain on which this operator is defined, is given by

$$\mathrm{Dom}(\hat{x}_j) = \{f \in L^2(\mathbb{R}^n) | \; x_j f \in L^2\}$$

This operator is called *the position operator in the j-th direction*.

Lemma 6.3.6. *The operator $\hat{x}_j$ is self-adjoint for $j = 1, \dots, n$.*

Proof. Change $\mathbb{R}$ to $\mathbb{R}^n$ in the proof of Lemma 6.3.3. □

Corollary 6.3.7. *The operator $\hat{p}_j$ is a self-adjoint operator for $j = 1, \dots, n$.*

Proof. By changing $\mathbb{R}$ to $\mathbb{R}^n$ in Lemma 6.3.4 and using Lemma 6.2.4, we see that it easily follows from these two lemmas. □

Now that we know of $\hat{x}_j$ and $\hat{p}_j$ are self-adjoint, we know they have a unique spectral measure. In fact, we can easily see what the spectral measure for $\hat{x}$ is.

Lemma 6.3.8. *Equip $\mathbb{R}$ with the Borel σ-algebra, denoted $\mathscr{A}$. Define the operator $E : \mathscr{A} \to \mathscr{B}(L^2(\mathbb{R}))$ by $E(A)(f) = \chi_A f$ where χ_A is the characteristic function of the set A. Then E is a spectral measure, and*

$$\int_{\sigma(\mathbb{R})} \lambda \, dE(\lambda) = \hat{x}$$

Proof. We first prove that E is a spectral measure. Let $A \subseteq \mathbb{R}$ be a Borel set. If $f \in L^2(\mathbb{R}^n)$ we have $E(A)E(A)f = \chi_A \chi_A f = \chi_A f = E(A)f$. Thus $E(A)^2 = E(A)$. Next we note that for any $f \in L^2$ we have $\langle E(A)f, f \rangle = \int \chi_A(x)|f(x)|^2 dx = \int_A |f(x)|^2 dx \geq 0$. By Lemma 3.2.3 we have that $E(A)$ is a projection for any Borel set $A \subseteq \mathbb{R}$.

Next, $E(\emptyset) = 0$ and $E(\mathbb{R}) = I$. Next, if A, B are Borel sets, then $E(A \cap B)f = \chi_{A \cap B} f = \chi_A \chi_B f = E(A)E(B)f$ for any $f \in L^2(\mathbb{R})$. And finally, if $\{A_n\}_n$ is a collection of pairwise disjoint sets, then $E(\cup_n A_n)f = \chi_{\cup_n A_n} f = (\chi_{A_1} + \chi_{A_2} + \cdots)f = \sum_{i=1}^{\infty} \chi_{A_i} f$ because the A_i are pairwise disjoint. Therefore, E is a spectral measure.

Next, note that for any Borel set A we have $\langle E(A)f, g \rangle = \int_A f(x)\overline{g(x)} dx$. Therefore, we see that if $f \in \mathrm{Dom}(\hat{x})$ and $g \in L^2(\mathbb{R})$ we have

$$\langle \hat{x}(f), g \rangle = \int_{\mathbb{R}} x f(x) \overline{g(x)} \, dx = \int_{\mathbb{R}} x \; d\langle E(x)f, g \rangle$$

Because $\hat{x}$ is self-adjoint, there exists only one spectral measure such that this is true. Thus we see that E is the spectral measure for $\hat{x}$. □

We see that if we consider $L^2(\mathbb{R}^n)$, the spectral measure of $\hat{x}_j$ is $E_j(A)f = \chi_{\mathbb{R}^{j-1}\times A\times\mathbb{R}^{n-j}}f$. It is easy to check that these spectral measures commute. Therefore $\hat{x}_i$ and $\hat{x}_j$ strongly commute for $1 \leq i, j \leq n$. However, we wish to show this using Nelson's theorem.

Corollary 6.3.9. *Consider $L^2(\mathbb{R}^n)$. Then the operators $\hat{x}_j$ and $\hat{x}_k$ commute strongly for all j, k.*

Proof. Let i, j be given. We start with a claim

Claim 1. $\hat{x}_j^2 + \hat{x}_k^2$ is self-adjoint on $L^2(\mathbb{R}^n)$.

Proof. The proof is the exact same proof as the proof of the self-adjointness of $\hat{x}$ in Lemma 6.3.3, where we replace $\hat{x}$ by $\hat{x}_j^2 + \hat{x}_k^2$, and $\mathbb{R}$ by $\mathbb{R}^n$. □

Next, we note that $\mathscr{D}(\mathbb{R}^n)$ lies dense in $L^2(\mathbb{R}^n)$, and $\mathscr{D}(\mathbb{R}^n) \subseteq \mathrm{Dom}(\hat{x}_j^2 + \hat{x}_k^2)$. Hence $\mathscr{D}(\mathbb{R}^n)$ lies dense in $\mathrm{Dom}(\hat{x}_j^2 + \hat{x}_k^2)$. Therefore we have that $\mathscr{D}(\mathbb{R}^n)$ is a core of $\hat{x}_j^2 + \hat{x}_k^2$. Also note that $\hat{x}_l(\mathscr{D}(\mathbb{R}^n)) \subseteq \mathscr{D}(\mathbb{R}^n)$ for $l = j, k$. On $\mathscr{D}(\mathbb{R}^n)$, we have that

$$\hat{x}_j\hat{x}_k\phi(x) = x_jx_k\phi(x) = \hat{x}_k\hat{x}_j\phi(x) \qquad (\phi \in \mathscr{D}(\mathbb{R}^n))$$

Because $\hat{x}_j^2 + \hat{x}_k^2$ is self-adjoint, we have that $(\hat{x}_j^2 + \hat{x}_k^2)|_{\mathscr{D}(\mathbb{R}^n)}$ is essentially self-adjoint. Therefore we can use the theorem of Nelson to conclude that $\hat{x}_j$ and $\hat{x}_k$ strongly commute. □

Finally, we consider $\hat{p}_j$, and try to find if the same holds as well.

Corollary 6.3.10. *Consider $L^2(\mathbb{R}^n)$. Then $\hat{p}_j$ and $\hat{p}_k$ commute strongly for all i, j.*

Proof. First we prove that $\hat{p}_j^2 + \hat{p}_k^2$ is self-adjoint. This proof is analogous to the proof of Lemma 6.3.4, so we will omit it. We replace $\hat{x}$ and $\hat{p}$ with $\hat{x}_j^2 + \hat{x}_k^2$ and $\hat{p}_j^2 + \hat{p}_k^2$, and note that the Fourier transform is linear.

Next, we use the same reasoning as in previous corollary to conclude that $\hat{p}_j$ and $\hat{p}_k$ strongly commute. □

Remember that strongly commuting of two operators resulted in the commutation of two measurable functions f and g as in Proposition 5.1.2.

Example 6.3.11. Define the functions $f : \sigma(\hat{x}_1) \to \mathbb{C}$ by $f(x) = e^x$ and $g : \sigma(\hat{x}_2) \to \mathbb{C}$ by $g(y) = |x|$. Because $\hat{x}_1$ and $\hat{x}_2$ commute strongly, we find that $f(\hat{x}_1) = e^{\hat{x}_1}$ and $g(\hat{x}_2) = |\hat{x}_2|$ commute. ⊘

Example 6.3.12. The same can be done for $\hat{p}_1$ and $\hat{p}_2$, because they strongly commute as well. So for example, the operators $e^{i\partial_1}$ and $\sin(i\partial_2)$ commute. ⊘

References

[1] J. B. Conway. *A course in Functional Analysis*. Springer, New York, 2 edition, 1990. ISBN 0387972455.

[2] J. J. Duistermaat and J. A. C. Kolk. *Distributions: Theory and Applications*. Springer, New York, 2010. ISBN 9780817646721.

[3] B. Fuglede. *A Commutativity Theorem for Normal Operators*, volume 36. National Academy of Sciences, 1950.

[4] B. C. Hall. *Quantum Theory for Mathematicians*. Springer, New York, 2013. ISBN 9781461471165.

[5] E. Nelson. *Analytic Vectors*, volume 70. Annals of Mathematics, 1959.

[6] M. Reed and B. Simon. *Methods of Modern Mathematical Physics Volume II: Fourier Analysis, Self-adjointness*. Acadamic Press Inc., New York, 1975. ISBN 0125850026.

[7] B. P. Rynne and M. A. Youngson. *Linear Functional Analysis*. Springer, New York, 2008. ISBN 9781852332570.

[8] R. L. Schilling. *Measures, Integrals and Martingales*. Cambridge University Press, Cambridge, 2 edition, 2017. ISBN 9781316620243.

[9] K. Schmüdgen. *Strongly Commuting Selfadjoint Operators And Commutants of Unbounded Operator Algebras*, volume 102. Proceedings of the American Mathematical Society, 1988.

[10] K. Schmüdgen. *Unbounded Self-adjoint Operators on Hilbert Space*. Springer Netherlands, 2012. ISBN 9789400747524.