Bachelor’s Thesis

Stone’s Theorem and Applications

Sven Möller

September 27, 2010

Supervision:
Prof. Alexandru Aleman (Lunds Universitet)
Prof. Burkhard Kümmerer (TU Darmstadt)
Stone’s theorem establishes a bijection between self-adjoint operators on a Hilbert space and one-parameter groups of strongly continuous, unitary operators via a functional calculus. We will first introduce the necessary terminology and then prove this result. We will study many of the numerous applications of Stone’s theorem, especially those in quantum mechanics. We will also consider non-physics-related applications including an elegant proof of Bochner’s theorem on positive-definite functions.
Acknowledgements

I thank Profes. Alexandru Aleman and Burkhard Kümmerer for their support and guidance. I thank my fellow student Katharina Schade for helping me with the subtleties of the English language.
Contents

1 Introduction 5
   1.1 Overview ............................................. 5
   1.2 Motivation ......................................... 6

2 Stone's Theorem 9
   2.1 Unbounded, Self-adjoint Operators on Hilbert Spaces .................. 9
   2.2 Spectral Theorem and Functional Calculus .............................. 16
   2.3 Stone's Theorem ..................................... 20

3 Applications 25
   3.1 Bochner's Theorem .................................... 25
   3.2 Physical Background ................................... 30
   3.3 Examples I ........................................... 33
      3.3.1 Translation ....................................... 33
      3.3.2 Gauge Transformation ............................. 36
      3.3.3 Translation in Many Dimensions ................. 38
      3.3.4 Rotation in Cartesian Coordinates .............. 39
      3.3.5 Rotation in Polar Coordinates .................. 42
      3.3.6 Rotations in Three Dimensions .................. 43
      3.3.7 Dilation .......................................... 46
      3.3.8 Modified Translation ............................. 48
   3.4 Examples II: Time Evolution .................................. 52
      3.4.1 Free Particle ..................................... 53
      3.4.2 Free, Moving Particle ............................ 56
      3.4.3 General Wave Function ........................... 58
      3.4.4 Constant Potential ................................ 58

4 Conclusions 59

Bibliography 60
1 Introduction

“To those who do not know mathematics it is difficult to get across a real feeling as to the beauty, the deepest beauty, of nature... If you want to learn about nature, to appreciate nature, it is necessary to understand the language that she speaks in.” —RICHARD P. FEYNMAN

It is a postulate of quantum mechanics that physical observables are represented by self-adjoint operators on Hilbert spaces (see Section 3.2). In addition, most of the operators naturally occurring in the mathematical description of quantum mechanics are unbounded. The understanding and proper treatment of these operators is hence an important task in functional analysis. While the lack of the boundedness property makes the description of these operators extremely difficult, their self-adjointness offers new, effective ways of tackling them.

Stone’s theorem on strongly continuous unitary groups is one of these powerful tools, which make life a little bit easier. It gives us a way of writing certain families of unitary operators in terms of self-adjoint, possibly unbounded operators. Interpreted in physical terms, this yields statements like “momentum generates translation” or “angular momentum generates rotation”.

1.1 Overview

To begin with we will look at the statement of Stone’s theorem from a purely mathematical point of view. First, we have to introduce the necessary background to be able to work with unbounded and self-adjoint operators, which includes the important notion of the graph of an operator. Second, we must develop a functional calculus to give meaning to the expression $f(T)$, where $f$ is a function on the real line and $T$ is an operator. For this we will need the spectral theorem for unbounded, self-adjoint operators. Having those tools at hand, we can first look at the converse statement of Stone’s theorem and then at the theorem itself and prove both results.

We will then consider several applications of Stone’s theorem, many of these being relevant for quantum mechanics. We will also look at an elegant proof of Bochner’s theorem, which plays an important role in probability theory. An introduction to the theory of quantum mechanics and its mathematical formulation will be given. This mathematical framework can be stated with the help of a few postulates, from which the whole theory is deduced. The physical background will help us to understand the importance of Stone’s theorem and to interpret some results in terms of physical quantities.
1.2 Motivation

To familiarise ourselves with the statement of Stone’s theorem, let us first state and prove it for the special case of square matrices as operators on a finite dimensional \( \mathbb{C} \)-vector space. In this case the functional calculus can be defined via an ordinary power series and the statement of Stone’s theorem can be easily proved.

We begin by defining the exponential of a matrix via a power series.

**Definition 1.** Let \( A \in M_n(\mathbb{C}) = \mathbb{C}^{n \times n} \). We define the **matrix exponential** of \( A \) by

\[
\exp(A) := \expm(A) := \sum_{n=0}^{\infty} \frac{A^n}{n!}.
\]

(1.1)

**Remark 1.** This definition is well-defined since if we look at

\[
\sum_{n=0}^{\infty} \left\| \frac{A^n}{n!} \right\| \leq \sum_{n=0}^{\infty} \frac{\|A\|^n}{n!} = e^{\|A\|},
\]

(1.2)

we see that the exponential series is absolutely convergent and since \( \mathbb{C}^{n \times n} \) is complete, absolute convergence implies convergence. Thus the power series converges. (A submultiplicative norm was chosen, but on \( \mathbb{C}^{n \times n} \) all norms are equivalent.)

So, we have extended the function \( \exp : \mathbb{C} \to \mathbb{C} \) to a function \( \expm : M_n(\mathbb{C}) \to M_n(\mathbb{C}) \). In the following we will for simplicity always write \( \exp \) (and not \( \expm \)) or \( e^{(\cdot)} \) and it will be clear from the context what the domain of the function is. We can also extend many other functions \( f : \mathbb{C} \to \mathbb{C} \) (including all entire functions) to complex matrices using the same definition via a power series. This would then be called a functional calculus. For now, it is however sufficient to have an expression for the exponential of a matrix. We will deal with functional calculi in more detail when we define them for operators on a general (possibly infinite-dimensional) Hilbert space.

Having defined the matrix exponential, we can look at the expression \( e^{itA} \) for a Hermitian matrix \( A \) (i.e. \( A^* = A \)) and \( t \in \mathbb{R} \). In the following, let \( E \) denote the unit matrix.

**Proposition 1.** Let \( A \in M_n(\mathbb{C}) \) be Hermitian and define \( U(t) = e^{itA} \). Then:

1. For each \( t \in \mathbb{R} \), \( U(t) \) is a unitary matrix.
2. \( U(t)U(s) = U(t+s) \) for all \( t, s \in \mathbb{R} \).
3. The family \( \{U(t)\}_{t \in \mathbb{R}} \) forms an Abelian group under matrix multiplication.
4. \( \lim_{t \to 0} \frac{U(t)-E}{t} = iA \).

\(^{1}\)It is an extension in the sense that if we identify \( \mathbb{C} \) with the complex multiples of the unit matrix, then the function \( \expm \) restricted to the unit matrices becomes \( \exp \).
Proof. First note that

\[(U(t))^* = \left(\sum_{n=0}^{\infty} \frac{(itA)^n}{n!}\right)^* = \sum_{n=0}^{\infty} \frac{(-itA^*)^n}{n!} = \sum_{n=0}^{\infty} \frac{(-itA)^n}{n!} = U(-t).\]  \hspace{1cm} (1.3)

The matrix $A$ is Hermitian and hence diagonalizable by a unitary matrix $S$ ($S^*S = E$), i.e. $A = SDS^*$, where $D$ is diagonal. Thus

\[U(t) = \sum_{n=0}^{\infty} \frac{(itA)^n}{n!} = \sum_{n=0}^{\infty} \frac{(itSDS^*)^n}{n!} = \sum_{n=0}^{\infty} \frac{S(itD)^nS^*}{n!} = S\sum_{n=0}^{\infty} \frac{(itD)^n}{n!}S^* \hspace{1cm} (1.4)\]

and since for diagonal matrices matrix multiplication is “pointwise” on the diagonal elements, we get

\[U(t) = S\text{diag}(e^{-itd_1}, \ldots, e^{-itd_n})S^*, \hspace{1cm} (1.5)\]

where $\text{diag}(d_1, \ldots, d_n) = D$. For the ordinary exponential function we know that $e^{-x} = \frac{1}{e^x}$ and thus get

\[(U(t))^*U(t) = U(-t)U(t) = S\text{diag}(e^{-itd_1}, \ldots, e^{-itd_n})S^*S\text{diag}(e^{itd_1}, \ldots, e^{itd_n})S^* = S\text{diag}(e^{-itd_1}e^{itd_1}, \ldots, e^{-itd_n}e^{itd_n})S^* = SS^* = E. \hspace{1cm} (1.6)\]

So $U(t)$ is unitary and we showed (1). To prove (2) note that we can use the same diagonalisation argument as above to show that $U(s)U(t) = U(s+t)$ making use of the basic property of the exponential function that $e^x e^t = e^{x+t}$. Then the family $\{U(t)\}_{t \in \mathbb{R}}$ is also closed under multiplication and (2) gives commutativity. Furthermore, to every $U(t)$ the inverse is $U(-t)$, as we already found out, and $U(0) = E$. Together, this proves (3). Finally, to show (4), we write (note that all sums are absolutely convergent)

\[
\frac{U(t) - E}{t} = \sum_{n=0}^{\infty} \frac{(itA)^n}{n!} - E = \sum_{n=1}^{\infty} \frac{(iA)^{n-1}}{n!} = iA\sum_{n=0}^{\infty} \frac{(itA)^n}{(n+1)!},
\]

where the remaining sum gives $E$ in the limit $t \to 0$. This yields (4). \hfill \Box

This derivation was quite straightforward. The more interesting question is perhaps whether a converse statement is also true. Given a family of unitary matrices with the above properties, can we deduce that they have to be of the form $e^{itA}$, where $A$ is a Hermitian matrix? This is in fact true under certain premises as stated in the following theorem.

**Theorem 1.** Let $\{U(t)\}_{t \in \mathbb{R}}$ be a family of unitary matrices obeying

1. $U(s + t) = U(s)U(t)$ for all $s, t \in \mathbb{R}$,
2. $\lim_{s \to t} \|U(s) - U(t)\| = 0$.

Then there exists a unique Hermitian matrix $A$ such that

\[U(t) = e^{itA}. \hspace{1cm} (1.8)\]
Proof. Let $s,t \in \mathbb{R}$. Then $U(s)U(t) = U(s + t) = U(t + s) = U(t)U(s)$. Furthermore, $U(t)$ is unitary and hence diagonalisable for all $t \in \mathbb{R}$. So $\{U(t)\}_{t \in \mathbb{R}}$ is a family of diagonalisable, pairwise commuting matrices. That is the case if and only if they are simultaneously diagonalisable, i.e. there exists a matrix $S$ such that $U(t) = SD(t)S^{-1}$, where $D(t)$ is a diagonal matrix for all $t \in \mathbb{R}$.

$D(t)$ now fulfills the same functional equation $U(t)$ does since

$$D(t)D(s) = S^{-1}U(t)SS^{-1}U(s)S = S^{-1}U(t + s)S = D(t + s). \quad (1.9)$$

If we write $D(t) = \text{diag}(d_1(t), \ldots, d_n(t))$ the same is true for the diagonal entries $d_i(t)$ of $D(t)$. Condition (2) implies the continuity of $U(t)$ as function of $t$. This implies that $D(t)$ and consequently all the $d_i(t)$ are continuous. Hence we are given a functional equation

$$f(t + s) = f(t) \cdot f(s) \quad (1.10)$$

for all $s, t \in \mathbb{R}$, where $f$ is a complex-valued function and continuous. It is a well-known fact that such a function $f$ is necessarily of the form

$$f(t) = e^{ct}, \quad (1.11)$$

where $c$ is some complex number. So, we can write $d_i(t) = e^{ict}$. Since $U(t)$ is unitary, the eigenvalues $d_i(t)$ have unit modulus for all $t \in \mathbb{R}$. This can only be if the $c_i$ are real. So altogether this gives $D(t) = e^{iCt}$, where $C = \text{diag}(c_1, \ldots, c_1)$. So we have

$$U(t) = SD(t)S^{-1} = Se^{iCt}S^{-1} = e^{i\overbrace{SC}^A S^{-1}t} = e^{iAt}. \quad (1.12)$$

At last we want to verify the Hermiticity of $A$. For this note that $U(t)U(-t) = U(0) = E$. So $U(-t) = U(t)^*$ since $U(t)$ is unitary. Then

$$e^{-iAt} = (e^{iAt})^* = e^{-iA^*t}. \quad (1.13)$$

By (4) of Proposition 1 we can obtain $A$ from $U(t)$ by differentiation and hence $A = A^*$ for consistency. With the same arguments we get the uniqueness of $A$. □

The statement of Theorem 1 is exactly the statement of Stone’s theorem for the special case of complex matrices. In the following, we want to formulate this statement for operators on any Hilbert space. In order to do so, a considerable amount of preparatory work is necessary. In the next chapter we will thus familiarise ourselves with unbounded, self-adjoint operators on Hilbert spaces.
2 Stone’s Theorem

2.1 Unbounded, Self-adjoint Operators on Hilbert Spaces

The operators naturally occurring in the mathematical description of quantum mechanics are often unbounded and usually self-adjoint, corresponding to the fact that their eigenvalues representing physical observables are always real-valued. It is therefore important to know how such operators have to be treated. (An introduction to these concepts can be found in many textbooks including [RS72, BSU96].)

The first important observation is that an everywhere-defined operator on a Hilbert space \( H \) satisfying
\[
\langle A\varphi,\psi \rangle = \langle \varphi, A\psi \rangle
\]
\[(2.1)\]
for all \( \varphi,\psi \in H \) is necessarily bounded. This is a simple consequence of the closed graph theorem and is known under the name of Hellinger-Toeplitz theorem. The above condition \[(2.1)\] is exactly the definition of a symmetric or Hermitian \(^1\) bounded operator defined on all of \( H \). If we want to extend the notion of Hermiticity to unbounded operators, we must necessarily allow operators to be only defined on a subspace of \( H \). This leads to the concept of the domain of an operator.

**Definition 2.** An operator \( T \) on a Hilbert space \( H \) is a linear map from its domain \( D(T) \), a linear subspace of \( H \), into \( H \).

Usually the domain \( D(T) \) is dense in \( H \) since many properties such as self-adjointness can only be defined in that case. An operator is only well-defined if it is given together with its domain.

Before we define Hermiticity and similar properties for unbounded operators, let us introduce another important concept, namely that of the graph of an operator. It already appeared in the name of the afore mentioned closed graph theorem.

**Definition 3.** The graph of an operator \( T \) is the set of pairs
\[
\Gamma(T) := \{(\varphi,T\varphi) | \varphi \in D(T)\}.
\]
\[(2.2)\]

The graph \( \Gamma(T) \) is then a linear subspace of \( H \times H \). The latter is again a Hilbert space with the inner product
\[
\langle (\varphi_1,\varphi_2),(\psi_1,\psi_2) \rangle := \langle \varphi_1,\psi_1 \rangle + \langle \varphi_2,\psi_2 \rangle.
\]
\[(2.3)\]

\(^1\)While for matrices, being symmetric and Hermitian are two different (although similar) properties, for operators on a (complex) Hilbert space, one uses symmetric in the same sense as Hermitian.
The closed graph theorem states that the graph of an everywhere-defined operator on a Banach space is closed if and only if the operator is bounded. For general operators this statement does not hold but it is still useful to introduce the class of operators with closed graph; they will however not automatically be bounded.

**Definition 4.** An operator $T$ on a Hilbert space $H$ is called **closed** if its graph $\Gamma(T)$ is closed in $H \times H$.

Certainly not all operators are closed but it is often possible to extend them to a closed operator.

**Definition 5.** An operator $S$ on $H$ is said to be an **extension** of $T$ (write $S \supseteq T$) if $D(S) \supseteq D(T)$ and $S\varphi = T\varphi$ for all $\varphi \in D(T)$. Equivalently $\Gamma(S) \supseteq \Gamma(T)$.

**Definition 6.** An operator $T$ is called **closable** if it has a closed extension. If the operator is closable, then the smallest (w.r.t. inclusion of graphs) closed extension is called the **closure** of $T$ and denoted by $\overline{T}$.

**Remark 2.** Let $T$ be a closed operator on $H$ and let $x_n \rightarrow x$ and $Tx_n \rightarrow y$. Then $Tx = y$.

**Proof.** The pair $(x_n, Tx_n)$ converges to $(x, y)$ and since $\Gamma(T)$ is closed in $H \times H$, $(x, y) \in \Gamma(T)$. This can only be if $y = Tx$. \hfill $\square$

As we will see soon, a large class of operators, namely the symmetric operators, are always closable. There are however also non-closable operators, as the following example shows.

**Example 1.** We set $H = L^2([a, b])$ and define the operator $A$ on the domain $D(A) = C([a, b])$ by $(A\varphi)(x) \equiv \varphi(a)$ (constant function). It is possible to construct a sequence $(\varphi_n)_{n=1}^\infty$ of functions $\varphi_n \in D(A)$ such that $\varphi_n(a) = 1$ for all $n$ and $\|\varphi_n\|_{L^2} \rightarrow 0$ as $n \rightarrow \infty$. Then $A\varphi_n \equiv 1 \neq 0$ for all $n$ but $\varphi_n \rightarrow 0$. Hence, if the operator $A$ were closable, by Remark 2 we would get $0 = A0 = \lim_{n \rightarrow \infty} A\varphi_n = \lim_{n \rightarrow \infty}(A\varphi_n) \neq 0$, which is a contradiction.

The following proposition shows that we can naively take the closure of the graph of an operator and get the graph of the closure, at least if the closure exists.

**Proposition 2.** If $T$ is closable, then $\Gamma(\overline{T}) = \overline{\Gamma(T)}$.

**Proof.** Let $S$ be an arbitrary closed extension of $T$. Then $\Gamma(T) \subseteq \Gamma(S)$ and since $\Gamma(S)$ is closed, taking the closure, we get $\overline{\Gamma(T)} \subseteq \Gamma(S)$. That means that if $(0, \psi) \in \overline{\Gamma(T)}$, then it is also in $\Gamma(S)$ and thus $\psi = 0$ ($\square$). We set $D(R) = \{\psi|(\psi, \varphi) \in \Gamma(T)\}$ for some $\varphi$ and define $R$ by $R\psi = \varphi$, where $\varphi$ is the vector such that $(\psi, \varphi) \in \Gamma(T)$. $R$ is well-defined since $\varphi$ is unique, which we can see from the above observation $\square$. On the other hand we find a $\varphi$ for every $\psi \in D(T)$ by definition of $D(R)$. So we have $\Gamma(R) = \Gamma(T)$. Hence $R$ is another closed extension of $T$. But $R \subseteq S$, which was an arbitrary closed extension of $T$. So $R$ must be the smallest closed extension or the closure of $T$ and consequently $\Gamma(\overline{T}) = \Gamma(R) = \overline{\Gamma(T)}$. \hfill $\square$
Now, let us define what it means for two operators to be adjoint.

**Definition 7.** Two operators $T$ and $S$ on $\mathcal{H}$ are called **formally adjoint** if
\[ \langle T\psi, \varphi \rangle = \langle \psi, S\varphi \rangle \quad (2.4) \]
for all $\psi \in D(T)$ and all $\varphi \in D(S)$.

Note that for a given $T$, $S$ is not unique. $T$ may have many formal adjoints. Especially the trivial operator, which is only defined at the origin, is a formal adjoint of any operator.

**Proposition 3.** Let $T$ be a densely defined operator on $\mathcal{H}$. Then there exists a unique (w. r. t. inclusion of graphs) formally adjoint operator denoted by $T^*$, which we call the **adjoint** of $T$.

**Proof.** Let $D(T^*)$ bet the set of $\varphi \in \mathcal{H}$ for which there exists an $\eta \in \mathcal{H}$, such that
\[ \langle T\psi, \varphi \rangle = \langle \psi, \eta \rangle \quad (2.5) \]
for all $\psi \in D(T)$. For those $\varphi \in D(T^*)$ we define $T^* \varphi = \eta$. Since $D(T)$ is dense in $\mathcal{H}$, $\eta$ is uniquely determined. $T^*$ is by definition a formal adjoint of $T$. Since we chose $D(T^*)$ as large as possible, it holds that $\Gamma(S) \subseteq \Gamma(T^*)$ for all formal adjoints $S$ of $T$. Thus, $T^*$ is maximal. \hfill $\Box$

In general, it is impossible to say how large $T^*$ will be. $D(T^*)$ does not have to be dense and could even be just the origin (see Example 3). Note however, that the following statement holds.

**Remark 3.** If $S \subseteq T$ both densely defined, then $T^* \subseteq S^*$.

**Proof.** Since $T^*$ is the adjoint of $T$, it is also a formal adjoint of $T$. Then, by definition of the formal adjoint and since $S \subseteq T$, $T^*$ is also a formal adjoint of $S$. This immediately implies $T^* \subseteq S^*$ since $S^*$ is the maximal formal adjoint of $S$. \hfill $\Box$

**Example 2.** We want to find an example of an operator whose adjoint is only defined at the origin. Let us consider the Hilbert space $\mathcal{H} = L^2([a,b])$ and fix a sequence of functions $(\varphi_j)_{j=1}^\infty$ such that $\varphi_j \in C([a,b])$, $\sum_{j=1}^\infty |\varphi_j(x)|$ is uniformly convergent in $[a,b]$, and the $\varphi_j$ are total in $L^2([a,b])$, i.e. every function $f \in L^2([a,b])$ can be approximated arbitrarily well by a finite linear combination of the $\varphi_j$ w. r. t. the $L^2$-norm. We also fix a sequence $(a_j)_{j=1}^\infty$ of distinct points $a_j \in [a,b]$. We set $D(A) = C([a,b])$ and define the operator $A$ on $D(A)$ by
\[ (Af)(x) = \sum_{j=1}^\infty f(a_j)\varphi_j(x) \quad (2.6) \]
for all $f \in D(A)$. $A$ is well-defined in $L^2([a,b])$ since the series is uniformly convergent. Moreover, as $D(A)$ is dense in $L^2([a,b])$, the adjoint of $A$ exists. Let us now determine
the domain $D(A^*)$ of the adjoint. It is given by all $g \in \mathcal{H}$ such that there exists an $h =: A^*g$ so that
\[ \langle Af, g \rangle = \langle f, h \rangle \] (2.7)
for all $f \in D(A)$. In our case this means
\[ \sum_{j=1}^{\infty} f(a_j) \int_a^b \varphi_j(x)g(x)dx = \int_a^b f(x)\overline{h(x)}dx \] (2.8)
for any $f \in C([a,b])$. We then fix $j = j_0$ and consider a sequence of uniformly bounded functions $f_n \in C([a,b])$ such that $f_n(a_{j_0}) = 1$ for all $n$ and $f_n(x) = 0$ for $x \notin (a_{j_0} - \frac{1}{n}, a_{j_0} + \frac{1}{n})$. Inserting $f_n$ for $f$ in the above equation and taking the limit gives $\langle \varphi_{j_0}, g \rangle = 0$. Since $j_0$ was arbitrary and because the $\varphi_j$ are total in $L^2([a,b])$, the above implies $g = 0$. Hence the domain $D(A^*)$ of the adjoint of $A$ is only the origin.

The next example shows that even if the adjoint is defined on more than just the origin, it can be the zero operator. For this we do similar arguments as in the example before.

Example 3. We take our non-closable operator from Example 1, i.e. $\mathcal{H} = L^2([a,b])$, $D(A) = C([a,b])$, and $(A\varphi)(x) \equiv \varphi(a)$. Let us now determine $D(A^*)$. It is given by all $\psi \in L^2([a,b])$ such that there exists an $A^*\psi =: \eta \in L^2([a,b])$ with
\[ \varphi(a) \int_a^b \overline{\psi(x)}dx = \int_a^b \varphi(x)\overline{\eta(x)}dx \] (2.9)
for all $\varphi \in C([a,b])$. If we consider a sequence $\varphi_n \in C([a,b])$ with $\varphi_n(a) = 1$ for all $n$ and $\|\varphi_n\|_{L^2} \rightarrow 0$ as $n \rightarrow \infty$, then inserting $\varphi_n(x)$ for $\varphi(x)$ in the above equation and taking the limit yields $\langle 1, \psi \rangle = 0$. But then $\langle \varphi, \eta \rangle = 0$ already holds for all $\varphi \in C([a,b])$, which implies $\eta = 0$. So $A^*$ is just the zero operator on its domain.

The concepts of adjointness and closedness are related as shown by the following theorem.

Theorem 2. Let $T$ be a densely defined operator on a Hilbert space $\mathcal{H}$. Then:

1. $T^*$ is closed.
2. $T$ is closable if and only if $D(T^*)$ is dense in $\mathcal{H}$. In that case $\overline{T} = T^{**} := (T^*)^*$.
3. If $T$ is closable then $(\overline{T})^* = T^*$.

Proof. We define the operator $V$ on $\mathcal{H} \times \mathcal{H}$ by
\[ V(\varphi, \psi) = (-\psi, \varphi). \] (2.10)
$V$ is unitary since it is an isometry and bijective. Then we know that $V[E^\perp] = (V[E])^\perp$ for any subspace $E$ of $\mathcal{H} \times \mathcal{H}$. Now let $T$ be a densely defined linear operator on $\mathcal{H}$ and let $(\varphi, \eta) \in \mathcal{H} \times \mathcal{H}$. By definition $(\varphi, \eta) \in (V[\Gamma(T)])^\perp$ if and only
if \( \langle \varphi, \eta \rangle, (-T\psi, \psi) \rangle = 0 \) for all \( \psi \in D(T) \). The latter is the case if and only if \( \langle \varphi, T\psi \rangle = \langle \eta, \psi \rangle \) for all \( \psi \in D(T) \), which by definition is equivalent to \( \langle \varphi, \eta \rangle \in \Gamma(T^*) \). So \( \Gamma(T^*) = (V[\Gamma(T)])^\perp \), which is always a closed subspace of \( \mathcal{H} \times \mathcal{H} \). Hence, \( T^* \) is closed, which proves (1).

Next, we want to prove (2). \( \Gamma(T) \) is a linear subspace of \( \mathcal{H} \times \mathcal{H} \) and thus

\[
\overline{\Gamma(T)} = \left( \Gamma(T)^\perp \right)^\perp = \left( V^2 \left[ \Gamma(T)^\perp \right] \right)^\perp = \left( V \left[ (V[\Gamma(T)])^\perp \right] \right)^\perp = (V[\Gamma(T)])^\perp. \tag{2.11}
\]

Let us now assume \( T^* \) is densely defined. Then we can make the same arguments as above for \( T^* \), which results in \( (V[\Gamma(T)])^\perp = \Gamma(T^{**}) \) and thus \( \overline{\Gamma(T)} = \Gamma(T^{**}) \). So \( T \) is closable and \( \overline{T} = T^{**} \).

Let us conversely assume that \( T^* \) is not densely defined. Let \( \psi \in (D(T^*))^\perp \). Then \((\psi, 0) \in (\Gamma(T^*))^\perp \) and thus \((0, \psi) \in V \left( (\Gamma(T^*))^\perp \right) \). So the latter cannot be the graph of an operator. But since \( \overline{\Gamma(T)} = (V[\Gamma(T)])^\perp \), \( T \) cannot be closable by Proposition 2.

Finally, to prove (3), we write for \( T \), which is closable,

\[
T^* = \overline{T^2} = (T^*)^* = (T^{**})^* = (\overline{T})^*, \tag{2.12}
\]

which is the assertion.

Now, we are able to introduce the notions of Hermiticity and self-adjointness.

**Definition 8.** A densely defined operator on a Hilbert space \( \mathcal{H} \) is **symmetric** (or **Hermitian**) if

\[
\langle T\varphi, \psi \rangle = \langle \varphi, T\psi \rangle \tag{2.13}
\]

for all \( \psi, \varphi \in D(T) \).

**Remark 4.** An operator \( T \) is symmetric if and only if \( T \subseteq T^* \).

**Proof.** \( T \subseteq T^* \) by definition means \( D(T) \subseteq D(T^*) \) and \( T\varphi = T^*\varphi \) for all \( \varphi \in D(T) \). This means that for all \( \varphi, \psi \in D(T) \subseteq D(T^*) \) it holds that \( \langle T\varphi, \psi \rangle = \langle T^*\varphi, \psi \rangle = \langle \varphi, T\psi \rangle \), which means that \( T \) is symmetric.

If we assume on the other hand that \( T \) is symmetric, then we see that \( T \) is a formal adjoint of \( T \), which implies \( T \subseteq T^* \). \( \Box \)

**Definition 9.** An operator \( T \) is called **self-adjoint** if \( T = T^* \). An operator \( T \) is called **essentially self-adjoint** if its closure \( \overline{T} \) is self-adjoint.

We see that for unbounded operators Hermiticity and self-adjointness are two distinct properties, whereas for bounded operators the two notions coincide. We see immediately from the above Remark 4 that self-adjointness is the stronger property and it will be needed as premise in many of the following results. Hermiticity is usually a too weak assumption.

The following proposition will bring more structure into the abundance of definitions we have just made.
Proposition 4. A symmetric operator $T$ is always closable and the following holds:

1. $T$ is symmetric if and only if $T \subseteq T^{**} \subseteq T^*$.
2. $T$ is closed and symmetric if and only if $T = T^{**} \subseteq T^*$.
3. $T$ is self-adjoint if and only if $T = T^{**} = T^*$.
4. $T$ is essentially self-adjoint if and only if $T \subseteq T^{**} = T^*$.

**Proof.** $T$ is densely defined and since $D(T^*) \supseteq D(T)$ for any symmetric operator, $T^*$ is also densely defined. Then by (2) of Theorem 3, $T$ is closable with closure $T^{**}$. So $T^*$ is an extension of $T$ and $T^{**}$ is the smallest one, which gives (1). If $T$ is already closed, we get (2) and if $T = T^{**}$, (3) follows. (In all three cases the converse is trivial.) Finally let $T$ be essentially self-adjoint. The closure $T^{**} = T^{**}$ is self-adjoint. So, $(T^{**})^* = (T^*)^{**} = T^*$ if $T^{**} = T^*$. Conversely if $T^{**} = T^*$, the same arguments show that $T = T^{**}$ is self-adjoint.

In the following example, we will make use of the above Proposition 4 by inferring closability from Hermiticity.

**Example 4.** Let $\mathcal{H} = L^2(\mathbb{R})$. We define the two operators $T$ and $S$ with domains $D(T) = C_c^\infty(\mathbb{R})$ and $D(S) = C_0^1(\mathbb{R})$ by $T\psi := i\psi'$ for $\psi \in D(T)$ and $S\psi := i\psi'$ for $\psi \in D(S)$. Both operators are symmetric since

$$
\int_\mathbb{R} i\psi'(x)\overline{\varphi(x)} = -\int_\mathbb{R} i\psi(x)\overline{\varphi'(x)} = \int_\mathbb{R} \psi(x)i\overline{\varphi(x)'}
$$

(2.14)

for all $\psi, \varphi \in C_c^{\geq 1}(\mathbb{R})$. By the above Proposition 4 both operators are closable. $S$ is a proper extension of $T$, i.e. $S \supset T$. In the following, we want to show that $T = T^{**}$. For this, we introduce the approximate identity, a concept we will also use in the proof of Stone’s theorem (Theorem 7). We let $j(x)$ be a positive $C^\infty$-function supported in $(-1, 1)$ with $\int_\mathbb{R} j(x)dx = 1$. We then define $j_\varepsilon(x) = \frac{1}{\varepsilon}j(\frac{x}{\varepsilon})$ and for $\varphi \in D(S) = C_0^1(\mathbb{R})$ we set

$$
\varphi_\varepsilon(x) = \int_\mathbb{R} j_\varepsilon(x-t)\varphi(t)dt,
$$

(2.15)

i.e. we convolute $\varphi$ with the mollifier $j_\varepsilon$ resulting in a $\varphi_\varepsilon \in C_c^\infty$. Then

$$
|\varphi_\varepsilon(x) - \varphi(x)| \leq \int_\mathbb{R} j_\varepsilon(x-t)|\varphi(x) - \varphi(t)|dt \leq \left( \sup_{|x-t| \leq \varepsilon} |\varphi(x) - \varphi(t)| \right) \int_\mathbb{R} j_\varepsilon(x-t)dt
$$

$$
= \sup_{|x-t| \leq \varepsilon} |\varphi(x) - \varphi(t)|.
$$

(2.16)

Being continuous and compactly supported, $\varphi$ is uniformly continuous and hence $\varphi_\varepsilon \to \varphi$ uniformly and also $\varphi_\varepsilon \to \varphi$ in $L^2(\mathbb{R})$. In the same way

$$
\frac{d}{dx}\varphi_\varepsilon(x) = \int_\mathbb{R} \frac{d}{dx}j_\varepsilon(x-t)\varphi(t)dt = \int_\mathbb{R} j_\varepsilon(x-t)\frac{d}{dt}\varphi(t)dt
$$

(2.17)
Thus, $T\varphi$ is contained in the closure of the graph of $T$, which is the graph of the closure of $T$. Altogether we have $\overline{T} \supseteq S \supset T$. Since the closure is the smallest closed extension and both $S$ and $T$ are closable, we also get that $S = T$.

The next remark is very simple but quite useful observation and will be used many times throughout the text, so it is included here.

**Remark 5.** Given two operators $A$ and $B$ such that $A \subseteq B$. If $A$ is self-adjoint, then it already follows that $A = B$.

**Proof.** From $A \subseteq B$ it follows by Remark 3 that $B \subseteq B^* \subseteq A^* = A$ and hence $A = B$. 

Let us, before we go on to the next section, look at a basic criterion for the (essential) self-adjointness of operators. It will be very useful in the proof of Stone’s theorem.

**Theorem 3.** *(basic criterion for self-adjointness)* Let $T$ be a symmetric operator on a Hilbert space $H$. Then the following statements are equivalent:

1. $T$ is self-adjoint, i.e. $T = T^*$.
2. $T$ is closed and $\ker(T^* + i) = \{0\}$.
3. $\operatorname{Ran}(T \pm i) = H$.

**Proof.** (1 $\implies$ 2) Let $T$ be self-adjoint and suppose there is a $\varphi \in D(T^*) = D(T)$ with $T^*\varphi = i\varphi$. Then also $T\varphi = i\varphi$ and

$$
i \langle \varphi, \varphi \rangle = \langle i\varphi, \varphi \rangle = \langle T\varphi, \varphi \rangle = \langle \varphi, T^*\varphi \rangle = \langle \varphi, i\varphi \rangle = -i \langle \varphi, \varphi \rangle,
$$

which implies $\varphi = 0$. Similarly, $T^*\varphi = -i\varphi$ for some $\varphi \in D(T^*) = D(T)$ implies $\varphi = 0$. Since the $T^*$ is always closed and $T = T^*$, $T$ is closed.

(2 $\implies$ 3) If (2) holds, then $\{0\} = \ker(T^* + i) = \ker(T \mp i)^\perp = (\operatorname{Ran}(T \mp i))^\perp$, so $\operatorname{Ran}(T \pm i)$ is dense in $H$. It now suffices to show that $\operatorname{Ran}(T \pm i)$ is closed. Since $T$ is symmetric, it holds for $\varphi \in D(T)$

$$
\|T^* \pm i\varphi\|^2 = \|T\varphi\|^2 + \|\varphi\|^2 + \langle T\varphi, i\varphi \rangle + \langle i\varphi, T\varphi \rangle = \|T\varphi\|^2 + \|\varphi\|^2
$$

(2.19) since $\langle T\varphi, i\varphi \rangle = \langle \varphi, iT^* \varphi \rangle = \langle \varphi, iT\varphi \rangle = -i \langle \varphi, T\varphi \rangle$. Hence $\|\varphi\| \leq \|(T + i)\varphi\|$. So, given a sequence $(\varphi_n) \subseteq D(T)$ with $(T + i)\varphi_n \to \psi$, we have that $((T + i)\varphi_n)$ is a Cauchy sequence and thus $(\varphi_n)$ is itself a Cauchy sequence. So, $(\varphi_n)$ converges to some $\varphi \in H$ and since $D(T)$ is closed, $\varphi \in D(T)$ and $(T + i)\varphi = \psi$ by Remark 2 which means that $\psi \in \operatorname{Ran}(T + i)$. Hence $\operatorname{Ran}(T + i)$ is closed. Similarly $\operatorname{Ran}(T - i)$ is closed and we get $\operatorname{Ran}(T \pm i) = H$.

(3 $\implies$ 1) Suppose (3) holds. If $\psi \in D(T^*)$, then there exists a $\varphi \in D(T)$ such that $(T - i)\varphi = (T^* - i)\psi$. Now, since $T$ is symmetric, $D(T) \subseteq D(T^*)$ and thus $(T^* - i)(\varphi - \psi) = 0$. But since $\ker(T^* - i) = (\operatorname{Ran}(T + i))^\perp = H^\perp = \{0\}$, it follows that $\varphi = \psi$ and hence $\psi \in D(T)$. So we showed $D(T^*) \subseteq D(T)$, which implies $D(T^*) = D(T)$. Thus, $T = T^*$ and $T$ is self-adjoint.
A similar assertion can be made for essentially self-adjoint operators. It can be proved using a subset of the arguments used to prove the above theorem or stated as a corollary to it.

**Corollary 1.** *(basic criterion for essential self-adjointness)* Let $T$ be a symmetric operator on a Hilbert space $\mathcal{H}$. Then the following statements are equivalent:

1. $T$ is essentially self-adjoint.
2. $\ker(T^* \pm i) = \{0\}$.
3. $\text{ran}(T \pm i)$ are dense in $\mathcal{H}$.

Explicitly finding self-adjoint operators is not always an easy task. Often it is easier to find operators which are essentially self-adjoint and then just take the closure of them. Hence we will make use of Corollary 1 several times in this text.

### 2.2 Spectral Theorem and Functional Calculus

To formulate Stone’s theorem, we still have to give meaning to an expression like $e^{itA}$ for an unbounded operator $A$. In order to do so, it is necessary to develop a functional calculus. This is done with the help of the spectral theorem for (generally unbounded) self-adjoint operators. We will state (but not prove) the spectral theorem in several versions. But first, let us define an important concept, which will be used many times throughout this text.

**Definition 10.** A sequence of operators $(T_n)_{n \in \mathbb{N}}$ in $\mathcal{L}(X)$ (the bounded linear operators on $X$), where $X$ is a normed vector space, **converges strongly** (or **converges in the strong operator topology**) to $T \in \mathcal{L}(X)$ if

$$T_n x \to Tx \quad (2.20)$$

in norm for all $x \in X$. Similarly, an operator-valued function $T : \mathbb{R} \to \mathcal{L}(X)$ is **strongly continuous** at $t_0 \in \mathbb{R}$ if $t \mapsto T(t)x$ is norm-continuous for all $x \in X$.

Next, we state the spectral theorem in its first formulation.

**Theorem 4.** *(spectral theorem for unbounded self-adjoint operators)* Let $A$ be a self-adjoint operator on a separable Hilbert space $\mathcal{H}$ with domain $D(A)$. Then there exists a measure space $(M, \mu)$ with $\mu$ a finite measure, a unitary operator $U : \mathcal{H} \to L^2(M, \mu)$, and a real-valued function $f$ on $M$, which is finite almost everywhere, such that:

1. $\psi \in D(A)$ if and only if $f(\cdot)(U\psi)(\cdot) \in L^2(M, \mu)$.
2. If $\varphi \in U[D(A)]$, then $(UAU^{-1}\varphi)(m) = f(m)\varphi(m)$.

**Proof.** We omit the proof and refer to [RS72] for an outline of it. \qed

16
If we compare the statement of the theorem to the diagonalisation of a matrix in finite dimensions, we see that the unitary operator $U$ “diagonalises” the operator $A$ by transforming it into a real-valued function which acts on an element $\psi$ in $D(A)$ by pointwise multiplication of the corresponding function $\varphi = U\psi$ with $f$. This we can think of as a diagonal matrix acting on a vector.

Moreover, the theorem allows us to intuitively define the meaning of the formal expression $h(A)$, where $h$ is a bounded Borel function on $\mathbb{R}$. We simply set

$$
h(A) = U^{-1}T_{h(f)}U, \quad (2.21)
$$

where $T_{h(f)}$ is the multiplication operator on $L^2(M, \mu)$ which multiplies by the function $h \circ f$. To see why this definition makes sense, let us look at

$$
h(A)\psi = U^{-1}h(A)U^{-1}U\psi = U^{-1}(Uh(A)U^{-1}\varphi) = U^{-1}(g(\cdot)\varphi(\cdot)), \quad (2.22)
$$

where $g$ is the function representing the “diagonalised” $h(A)$. The only sensible choice is $g = h \circ f$. For this construction we restrict ourselves to bounded functions $h$ to ensure that $(h \circ f) \cdot \varphi$ is again in $L^2(M, \mu)$, i.e. $T_{h(f)}$ maps into $L^2(M, \mu)$. That of course means that $h(A)$ is a bounded operator and everywhere defined on $\mathcal{H}$.

Furthermore and more importantly, this definition is consistent with the definition of $h(A)$ for a bounded operator $A$ via a power series. To see this, we first note that the above spectral theorem also exists in a version for bounded self-adjoint operators. [RS72]

The statement then is that there exists a finite measure space $(M, \mu)$, a bounded real-valued function $f$ on $M$, and a unitary map $U : \mathcal{H} \to L^2(M, \mu)$, such that

$$
(UAU^{-1}\varphi)(m) = f(m)\varphi(m). \quad (2.23)
$$

Now, let us assume, we have a function $h$ on $\mathbb{R}$, which we can write as a power series with an infinite radius of convergence, i.e.

$$
h(x) = \sum_{n=0}^{\infty} a_n x^n. \quad (2.24)
$$

Then we define by analogy

$$
h(A) := \sum_{n=0}^{\infty} a_n A^n, \quad (2.25)
$$

which is well defined since $A$ is bounded and hence the sum converges absolutely. We get

$$
(A^n)\psi = U^{-1}UA^nU^{-1} U\psi = U^{-1}((UAU^{-1})^n)\varphi = U^{-1}((f(\cdot))^n \cdot \varphi(\cdot)) \quad (2.26)
$$

and thus

$$
h(A)\psi = \left(\sum_{n=0}^{\infty} a_n A^n\right)\psi = U^{-1}\left(\sum_{n=0}^{\infty} a_n (f(\cdot))^n \cdot \varphi(\cdot)\right) = U^{-1}((h \circ f) \cdot \varphi) = U^{-1}T_{h(f)}U\psi. \quad (2.27)
$$
The above arguments can only be made in the case of a bounded operator. This is for several reasons. First, if we simply apply a power series to an unbounded operator, we do not know whether this converges. This is guaranteed for a bounded operator by the absolute convergence which on a Banach space implies convergence. Second, an unbounded operator is not defined on the whole space. Thus, in order to apply it several times, we must restrict ourselves to a domain so that the function values always stay in that domain.

Implementing the idea of a functional calculus, the spectral theorem can be reformulated as follows.

**Theorem 5.** (spectral theorem - functional calculus form) Let $A$ be a self-adjoint operator on a Hilbert space $H$. Then there exists a unique map $\Phi_A$ from the bounded Borel functions on $\mathbb{R}$ into $L(H)$, such that:

1. $\Phi_A$ is an algebraic $*$-homomorphism, i.e.
   a) $\Phi_A(h \cdot g) = \Phi_A(h)\Phi_A(g)$,
   b) $\Phi_A(c \cdot h) = c \cdot \Phi_A(h)$,
   c) $\Phi_A(h + g) = \Phi_A(h) + \Phi_A(g)$,
   d) $\Phi_A(\overline{h}) = (\Phi_A(h))^*$.

2. $\|\Phi_A(h)\|_{L(H)} \leq \|h\|_\infty$, i.e. $\Phi_A$ is continuous.

3. If $h_n$ is a sequence of bounded Borel functions with $h_n \xrightarrow{n \to \infty} \text{Id}$ pointwise and $|h_n(x)| \leq |x|$ for all $x$ and $n$, then for any $\psi \in D(A)$, $\Phi_A(h_n)\psi \xrightarrow{n \to \infty} A\psi$.

4. If $h_n \xrightarrow{n \to \infty} h$ pointwise and the sequence $\|h_n\|_\infty$ is bounded then $\Phi_A(h_n) \xrightarrow{n \to \infty} \Phi_A(h)$ strongly.

5. If $A\psi = \lambda\psi$, then $\Phi_A(h)\psi = h(\lambda)\psi$.

**Proof.** Fix $A$. We set $\Phi(h) = h(A) = U^{-1}T_{h(f)}U$ with $U$ and $f$ from Theorem 4. Then the assertions follow. The full proof is omitted.

The map $\Phi_A$ can be viewed as the evaluation homomorphism of $h$ at $A$ since $\Phi_A(h) = h(A)$. Statement (3) of the theorem is saying \(\Phi_A(\text{Id}) = A\), which would be consistent with the above interpretation, the left-hand side is however not well-defined since $\text{Id}$ is not a bounded function.

We now have the means of evaluating any bounded Borel function at a self-adjoint operator and add or multiply these functions as we would do with them as functions on $\mathbb{R}$. This is what is called a functional calculus.

The above theorem gives us a very simple method of constructing orthogonal projections on $H$. If we let $\Omega \subseteq \mathbb{R}$ be a measurable set and denote by $1_\Omega$ its characteristic function, i.e.

$$1_\Omega(x) = \begin{cases} 1 & \text{if } x \in \Omega \\ 0 & \text{else} \end{cases} \quad (2.28)$$

\[2\text{Again, see } [\text{RS72}] \text{ for details.}\]
and let $P_A^\Omega := \Phi_A(1_\Omega) = 1_\Omega(A)$, then the following proposition holds.

**Proposition 5.** Let $\{P_A^\Omega\}$ be the family of all operators constructed as above, where $A$ is self-adjoint. Then:

1. Each $P_A^\Omega$ is an orthogonal projection, i.e.
   a) $P_A^\Omega P_A^\Omega = P_A^\Omega$,
   b) $P_A^\Omega$ is self-adjoint.
2. $P_A^\Omega = 0$ and $P_A^\emptyset = \text{Id}_H$.
3. If $\Omega = \bigcup_{n=1}^\infty \Omega_n$ with $\Omega_n \cap \Omega_m = \emptyset$ if $n \neq m$, then $\sum_{n=1}^N P_A^\Omega_n \xrightarrow{N \to \infty} P_A^\emptyset$ strongly.
4. $P_A^\emptyset P_A^\Omega = P_A^\emptyset$.

**Proof.** (1), (2), and (4) follow directly from the definition of the $P_A^\Omega$ and the functional calculus. (3) follows from (4) of Theorem 5. □

Such a family of operators is called a **projection-valued measure** (PVM) since for $\varphi \in \mathcal{H}$, $\langle \varphi, P_A^\Omega \varphi \rangle$ is a well-defined measure on $\mathbb{R}$, which we denote by $d \langle \varphi, P_A^\lambda \varphi \rangle$. Similarly $d \langle \varphi, P_A^\lambda \psi \rangle$ defines a complex measure for all $\varphi, \psi \in \mathcal{H}$. Now, for any bounded Borel function $h$ we can define $h(A)$ by

$$
\langle \varphi, h(A)\varphi \rangle = \int_{\mathbb{R}} h(\lambda) d \langle \varphi, P_A^\lambda \varphi \rangle. \quad (2.29)
$$

This definition of $h(A)$ then coincides with the one of Theorem 5. We write symbolically

$$
h(A) = \int_{\mathbb{R}} h(\lambda) d P_A^\lambda. \quad (2.30)
$$

Now suppose $h$ is an unbounded Borel function. Let

$$
D_h = \left\{ \varphi \left| \int_{\mathbb{R}} |h(\lambda)|^2 d \langle \varphi, P_A^\lambda \varphi \rangle < \infty \right. \right\}. \quad (2.31)
$$

Then $D_h$ is dense in $\mathcal{H}$ and we can define the operator $h(A)$ on $D_h$ by

$$
\langle \varphi, h(A)\varphi \rangle = \int_{\mathbb{R}} h(\lambda) d \langle \varphi, P_A^\lambda \varphi \rangle. \quad (2.32)
$$

In contrast to the case of a bounded $h$, $h(A)$ is again unbounded if $h$ is unbounded. For a more detailed treatment of the last few steps, see for example [RS72].
2.3 Stone’s Theorem

Having developed the functional calculus, we can now look at the operators of the form $U(t) = e^{itA}$, where $A$ is self-adjoint. Let us first study some properties of these operators.

**Theorem 6.** Let $A$ be a self-adjoint operator on a Hilbert space $\mathcal{H}$ and let $U(t) = e^{itA}$. Then:

1. $U(t)$ is a unitary operator for all $t \in \mathbb{R}$ and $U(t + s) = U(t)U(s)$ for all $s, t \in \mathbb{R}$. Furthermore, $\{U(t)\}_{t \in \mathbb{R}}$ forms an Abelian group under composition of operators.

2. $U(t)\varphi \to U(t_0)\varphi$ for all $\varphi \in \mathcal{H}$ as $t \to t_0$, i.e. $t \mapsto U(t)$ is continuous w.r.t. the strong operator topology.

3. $\frac{U(t)\psi - \psi}{t} \to iA\psi$ for all $\psi \in D(A)$ as $t \to 0$.

4. If $\lim_{t \to 0} \frac{U(t)\psi - \psi}{t}$ exists, then $\psi \in D(A)$.

**Proof.** The properties (1) follow directly from the functional calculus and the properties of the complex valued function $h_t(\lambda) := e^{it\lambda}$. It is convenient to use the language of the spectral theorem in its functional calculus version (Theorem 5) applied to the operator $A$. We write

$$U(t)^*U(t) = \Phi_A(h_t)^*\Phi_A(h_t) = \Phi_A(h_t^*h_t) = \Phi_A(h_t^*) = \Phi_A(1) = \text{Id}_\mathcal{H} \tag{2.33}$$

and hence $U(t)$ is unitary. Furthermore,

$$U(t)U(s) = \Phi_A(h_t)\Phi_A(h_s) = \Phi_A(h_t \cdot h_s) = \Phi_A(h_{t+s}) = U(t + s). \tag{2.34}$$

To see that the $U(t)$ form a group, we note that we already showed the closedness under composition. The associativity and commutativity also follow directly from the above observations. Moreover, it follows that $U(0)$ is the neutral element and $U(-t)$ is the inverse of $U(t)$.

To prove (2) we first note that if $t \mapsto U(t)$ is strongly continuous at $t = 0$, then it is strongly continuous everywhere since for $t \to t_0$ we can write

$$U(t)\varphi = U(t - t_0 + t_0)\varphi = U(t - t_0)U(t_0)\varphi = U(\tau)\psi \to U(0)\psi = U(t_0)\varphi \tag{2.35}$$

since $\tau \to 0$. Now, it is convenient to use the projection-valued measure formulation. We observe that

$$\|e^{itA}\varphi - \varphi\|^2 = \int_{\mathbb{R}} \left|e^{it\lambda} - 1\right|^2 \langle P_\lambda^A\varphi, \varphi \rangle \tag{2.36}$$

since for any function $h$

$$\|h(A)\varphi\|^2 = \langle h(A)\varphi, h(A)\varphi \rangle = \langle \varphi, h(A)^*h(A)\varphi \rangle = \langle \varphi, h(h(A)\varphi) \rangle = \left\langle \varphi, \left| h \right|^2 \cdot (A)\varphi \right\rangle. \tag{2.37}$$
The term $|e^{it\lambda} - 1|^2$ is dominated by the integrable function $g(\lambda) = 4$ (it is integrable since $\int_{\mathbb{R}} 4 \langle P_A^\lambda \varphi, \varphi \rangle = 4 \langle \varphi, \varphi \rangle$) and $|e^{it\lambda} - 1|^2 \to 0$ pointwise for all $\lambda \in \mathbb{R}$ as $t \to 0$. So $\|U(t)\varphi - \varphi\|^2 \to 0$ by Lebesgue’s dominated convergence theorem, which means that $t \to U(t)$ is strongly continuous at $t = 0$ and hence everywhere.

To prove (3) we use a similar technique. We again observe that

\[
\left\| \frac{U(t)\psi - \psi}{t} - iA\psi \right\|^2 = \int_{\mathbb{R}} \left| \frac{e^{it\lambda} - 1}{t} - i\lambda \right|^2 d \langle P_A^\lambda \varphi, \varphi \rangle.
\]  

(2.38)

In addition, for $x \in \mathbb{R}$,

\[
|e^{ix} - 1|^2 = \sin^2(x) + (\cos(x) - 1)^2 = \left(2 \sin \left(\frac{x}{2}\right)\right)^2 \leq x^2.
\]  

(2.39)

Thus

\[
\left| \frac{e^{it\lambda} - 1}{t} - i\lambda \right|^2 \leq \left( \left| \frac{\lambda}{t} \right| + |\lambda| \right)^2 = (2\lambda)^2,
\]  

(2.40)

which is integrable since

\[
\int_{\mathbb{R}} (2\lambda)^2 d \langle P_A\varphi, \varphi \rangle = \|2A\psi\|^2 < \infty
\]  

(2.41)

as $\psi \in D(A)$. So, we have found a dominating, measurable function and observing that $\left| \frac{e^{it\lambda} - 1}{t} - i\lambda \right|^2 \to 0$ pointwise for all $\lambda \in \mathbb{R}$ as $t \to 0$, we get (3) by the dominated convergence theorem.

Finally, we define

\[
D(B) := \left\{ \psi \left| \lim_{t \to 0} \frac{U(t)\psi - \psi}{t} \right. \text{exists} \right\}
\]  

(2.42)

and let $iB\psi := \lim_{t \to 0} \frac{U(t)\psi - \psi}{t}$. Then a simple computation gives that $B$ is symmetric, i.e. $B \subseteq B^*$. Namely, if we let $\varphi, \psi \in D(B)$, then

\[
\langle B\varphi, \psi \rangle = \lim_{t \to 0} \left\langle \varphi, iU(t)^* \left( \frac{U(t)\psi - \psi}{t} \right) \right\rangle = \lim_{t \to 0} \left\langle \varphi, \frac{iU(t)^* \psi - \psi}{t} \right\rangle
\]  

\[
= \lim_{t \to 0} \left\langle \varphi, iU(-t)\psi - \psi \right\rangle = \lim_{t \to 0} \left\langle \varphi, iU(t)\psi - \psi \right\rangle = \langle \varphi, B\psi \rangle.
\]  

(2.43)

By (3), $B \supseteq A$. $A$ is self-adjoint and hence $A = B$ by Remark 5.

**Definition 11.** Let \{\(U(t)\)\}_{t \in \mathbb{R}} be a family of unitary operators such that $U(t)U(s) = U(t+s)$ for all $t, s \in \mathbb{R}$. If in addition it holds that $U(t)\varphi \to U(t_0)\varphi$ for all $\varphi \in \mathcal{H}$ as $t \to t_0$, we call \{\(U(t)\)\}_{t \in \mathbb{R}} a **strongly continuous (one-parameter) unitary group**.

The name is justified since we showed in the proof of the above theorem that with these conditions \{\(U(t)\)\}_{t \in \mathbb{R}} forms a group.

Together with the above theorem, Stone’s theorem establishes a bijection between strongly continuous one-parameter unitary groups and self-adjoint operators on a Hilbert space. Stone’s Theorem is essentially the converse of the Theorem we just stated.
Theorem 7. (Stone’s theorem) Let \( \{U(t)\}_{t \in \mathbb{R}} \) be a strongly continuous one-parameter unitary group on a Hilbert space \( \mathcal{H} \). Then there exists a unique self-adjoint operator \( A \) on \( \mathcal{H} \) such that \( U(t) = e^{itA} \).

Definition 12. If \( \{U(t)\}_{t \in \mathbb{R}} \) is a strongly continuous one-parameter unitary group, then the self-adjoint operator \( A \) with \( U(t) = e^{itA} \) is called the infinitesimal generator of \( U(t) \).

Proof of theorem. The proof of this theorem is quite technical. By part (3) of the preceding Theorem 6 we know that we can obtain \( A \) by differentiating \( U(t) \) at \( t = 0 \). We will see that this can be done on a dense subset of \( \mathcal{H} \) consisting of “nice” vectors. This will give us an operator, which we will prove to be essentially self-adjoint by using the basic criterion (Corollary 1). In the end, we will see that \( U(t) \) is just the exponential of the closure of this operator.

To begin with, let \( f \in C_c^\infty(\mathbb{R}) \), i.e. \( f \) is infinitely often differentiable as has compact support. For each \( \varphi \in \mathcal{H} \) we define
\[
\varphi_f = \int_{-\infty}^{\infty} f(t)U(t)\varphi dt.
\]

This integral is Hilbert space-valued and defined as a Riemann integral, which is well-defined since \( U(t) \) is strongly continuous. Next, let \( D \) be the set of finite linear combinations of all such \( \varphi_f \) for \( \varphi \in \mathcal{H} \) and \( f \in C_c^\infty(\mathbb{R}) \). We again use the approximate identity \( j_\varepsilon(x) \). Recall, we let \( j(x) \) be any \( C_c^\infty(\mathbb{R}) \)-function with support contained in \((-1,1)\) and \( \int_{-\infty}^{\infty} j(x)dx = 1 \). We set \( j_\varepsilon(x) = \frac{1}{\varepsilon} j\left(\frac{x}{\varepsilon}\right) \). (Obviously, \( \int_{-\infty}^{\infty} j_\varepsilon(x)dx = 1 \).) Then
\[
\|\varphi_{j_\varepsilon} - \varphi\| = \left\| \int_{-\infty}^{\infty} j_\varepsilon(t)U(t)\varphi dt - \varphi \cdot 1 \right\| = \left\| \int_{-\infty}^{\infty} j_\varepsilon(t)U(t)\varphi dt - \varphi \int_{-\infty}^{\infty} j_\varepsilon(t)dt \right\|
\]
\[
= \left\| \int_{-\infty}^{\infty} j_\varepsilon(t)(U(t)\varphi - \varphi) dt \right\| \leq \int_{-\infty}^{\infty} j_\varepsilon(t) \|U(t)\varphi - \varphi\| dt
\]
\[
\leq \left( \int_{-\infty}^{\infty} j_\varepsilon(t) dt \right) \sup_{t \in [-\varepsilon,\varepsilon]} \|U(t)\varphi - \varphi\| = \sup_{t \in [-\varepsilon,\varepsilon]} \|U(t)\varphi - \varphi\|,
\]

where at “!” we used the inequality \( \| \int h(t)dt \| \leq \int \| h(t) \| dt \) for a Banach space-valued function \( h \). We can therefore conclude that \( D \) is dense in \( \mathcal{H} \) since if we let \( \varepsilon \to 0 \), then \( \varphi_{j_\varepsilon} \to \varphi \) because \( U(t) \) is strongly continuous.

---

3Since we defined the integral as a Riemann integral, one can prove this statement formally identically to the case of real- or complex-valued functions.
For a \( \varphi_f \in D \) we can compute
\[
\left( \frac{U(s) - \text{Id}}{s} \right) \varphi_f = \left( \frac{U(s) - \text{Id}}{s} \right) \int_{-\infty}^{\infty} f(t)U(t) \varphi dt = \int_{-\infty}^{\infty} f(t) \left( \frac{U(s + t) - U(t)}{s} \right) \varphi dt
\]
\[
= \int_{-\infty}^{\infty} f(t) \frac{U(s + t)}{s} \varphi dt - \int_{-\infty}^{\infty} f(t) \frac{U(t)}{s} \varphi dt
\]
\[
= \int_{-\infty}^{\infty} f(t-s) \frac{U(t)}{s} \varphi dt - \int_{-\infty}^{\infty} f(t) \frac{U(t)}{s} \varphi dt
\]
\[
= \int_{-\infty}^{\infty} f(t-s) - f(t) \frac{U(t)}{s} \varphi dt \xrightarrow{s \to 0} - \int_{-\infty}^{\infty} f'(t)U(t) \varphi dt = \varphi_{-f}
\]
(2.46)

since \( \frac{f(t-s) - f(t)}{s} \) converges uniformly to \( -f'(t) \) as \( s \to 0 \).

Now we can define the operator \( \tilde{A} \) on \( D \) by \( \tilde{A}\varphi_f = -i\varphi_{-f} \). Observe that \( U(t) : D \to D \) and \( A : D \to D \). The latter is clear by definition, the former can be shown by writing
\[
U(s)\varphi_f = U(s) \int_{-\infty}^{\infty} f(t)U(t) \varphi dt = \int_{-\infty}^{\infty} f(t)U(s + t) dt = \int_{-\infty}^{\infty} f(t)U(s) dt = \varphi_g.
\]
(2.47)

Furthermore \( U(t)\tilde{A}\varphi_f = \tilde{A}U(t)\varphi_f \) for \( \varphi_f \in D \), since (again using \( g(t) = f(t-s) \))
\[
\tilde{A}U(s)\varphi_f = \tilde{A}\varphi_g = -i\varphi_{-\varphi} = -iU(s)\varphi_{-f} = U(s)\tilde{A}\varphi_f,
\]
(2.48)

which simply reflects the fact that differentiation and translation commute.

Next, we want to show that \( \tilde{A} \) is symmetric. For this we write
\[
\langle \tilde{A}\varphi_f, \varphi_g \rangle = \lim_{s \to 0} \langle \left( \frac{U(s) - \text{Id}}{is} \right) \varphi_f, \varphi_g \rangle = \lim_{s \to 0} \langle \varphi_f, \left( \frac{\text{Id} - U(-s)}{is} \right) \varphi_g \rangle
\]
\[
= \langle \varphi_f, -i\varphi_{-g} \rangle = \langle \varphi_f, \tilde{A}\varphi_g \rangle.
\]
(2.49)

Now we can show that \( \tilde{A} \) is essentially self-adjoint using Corollary [I]. Suppose there is a \( \psi \in D(\tilde{A}^\ast) \) so that \( \tilde{A}^\ast\psi = i\psi \). Then for each \( \varphi \in D(\tilde{A}) = D \) we have
\[
\frac{d}{dt} \langle U(t)\varphi, \psi \rangle = \lim_{s \to 0} \left\langle \frac{U(s) + U(t) - U(t) - U(s)}{s} \varphi, \psi \right\rangle = \lim_{s \to 0} \left\langle \frac{U(s) - \text{Id}}{s} \varphi, \tilde{A}\psi \right\rangle
\]
\[
= \langle i\tilde{A}U(t)\varphi, \psi \rangle = i \langle U(t)\varphi, \tilde{A}^\ast\psi \rangle = i \langle U(t)\varphi, i\psi \rangle = \langle U(t)\varphi, \psi \rangle.
\]
(2.50)

So the complex-valued function \( f(t) = \langle U(t)\varphi, \psi \rangle \) satisfies the ordinary differential equation \( f' = f \), demanding an exponential solution \( f(t) = f(0)e^t \). On the other hand \( U(t) \) is unitary and thus has norm 1. So \( f(t) \) has to be bounded for positive and negative \( t \), which is only possible if \( f(0) = 0 = \langle \varphi, \psi \rangle \). Since \( D \) was dense in \( \mathcal{H} \) and \( \varphi \) was chosen arbitrarily, we conclude that \( \psi = 0 \). Similarly, we conclude that the equation \( \tilde{A}^\ast\psi = -i\psi \).
has no non-zero solutions. Then by Corollary 1 we know that \( \tilde{A} \) is essentially self-adjoint on \( D \), i.e. \( \tilde{A} := A \) is self-adjoint.

Then we define \( V(t) = e^{itA} \) and show that \( U(t) \) and \( V(t) \) coincide on \( D \). So let \( \varphi \in D \). Since \( \varphi \in D(A) \), \( V(t)\varphi \in D(A) \) and \( V'(t)\varphi = iAV(t)\varphi \) by (3) of Theorem 6. We already know that \( U(t)\varphi \in D \subseteq D(A) \) for all \( t \in \mathbb{R} \). If we now let \( w(t) = U(t)\varphi - V(t)\varphi \), then \( w(t) \) is a differentiable Hilbert space-valued function, since \( U(t) \) is strongly differentiable by assumption and \( V(t) \) because of Theorem 6. We get

\[
w'(t) = i\tilde{A}U(t)\varphi - iAV(t)\varphi = iAw(t).
\]

Hence,

\[
dt \|w(t)\|^2 = -i \langle Aw(t), w(t) \rangle + i \langle w(t), Aw(t) \rangle = 0
\] (2.52)

and so \( w(t) = 0 \) for all \( t \in \mathbb{R} \) since \( w(0) = 0 \) by definition. This means that \( U(t)\varphi = V(t)\varphi \) for all \( t \in \mathbb{R} \) and all \( \varphi \in D \). Since \( D \) is dense in \( \mathcal{H} \), we get \( U(t) = V(t) \) for all \( t \in \mathbb{R} \). Hence, we have found \( A \) to be the infinitesimal generator of \( U(t) \).

Finally, to see that the infinitesimal generator is unique, let us assume that there exists a self-adjoint operator \( B \) such that \( e^{itB} = U(t) = e^{itA} \). But then, by (3) of Theorem 6, \( A = B \).

**Remark 6.** We could have also introduced the concept of weakly continuous unitary group, where \( t \mapsto \langle U(t)\psi, \varphi \rangle \) is continuous for all \( \psi \) and \( \varphi \). However, the following computation shows that this already implies strong continuity:

\[
\|U(t)\varphi - \varphi\|^2 = \|U(t)\varphi\|^2 - \langle U(t)\varphi, \varphi \rangle - \langle \varphi, U(t)\varphi \rangle + \|\varphi\|^2 \to 2\|\varphi\|^2 - 2\|\varphi\|^2 = 0.
\] (2.53)
3 Applications

In this chapter we will study numerous applications of Stone’s theorem. Before we begin with the examples important for the mathematical formulation of quantum mechanics, let us look at one purely mathematical application, namely the proof of Bochner’s theorem.

3.1 Bochner’s Theorem

We want to state and prove Bochner’s theorem as a consequence of Stone’s theorem. The proof is, with some preliminary work done, much more elegant than Bochner’s original proof. The theorem relates positive-definite functions and positive measures via the Fourier transform and is an important result in probability theory concerning the characteristic function of a random variable.

Definition 13. A function \( f: \mathbb{R}^n \to \mathbb{C} \) is called positive-definite if for any given (not necessarily distinct) points \( \vec{x}_1, \ldots, \vec{x}_N \in \mathbb{R}^n \) the matrix \( A = (f(\vec{x}_i - \vec{x}_j))_{ij} \) is a Hermitian and positive semi-definite matrix, i.e.

\[
\vec{z}^* A \vec{z} \geq 0 
\]  

(3.1)

for all \( \vec{z} \in \mathbb{R}^N \).

Remark 7. It follows directly from the definition of positive-definite functions that

1. \( f(0) \geq 0 \),
2. \( f(\vec{x}) = \overline{f(-\vec{x})} \),
3. \( |f(\vec{x})| \leq f(0) \), thus \( f \) is bounded.

Proof. Set \( N = 1 \), \( \vec{x}_1 = 0 \), and \( \vec{z} = 1 \). Then (3.1) reads \( f(0) \geq 0 \), which proves (1).

With \( N = 2 \), \( \vec{x}_1 = \frac{1}{2} \vec{x} \), and \( \vec{x}_1 = -\frac{1}{2} \vec{x} \), (2) follows from the Hermiticity of \( A \). We get (3) if we choose \( \vec{z} = (1, -1)^T \) in the above situation and look at (3.1).

Now suppose, we have a finite positive measure \( \mu \) on \( \mathbb{R}^n \). We can then look at its Fourier transform defined by

\[
g(\vec{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i\vec{\xi} \cdot \vec{x}} d\mu(\vec{\xi})
\]

(3.2)
and let \( A = (g(\vec{x}_i - \vec{x}_j))_{i,j}, \vec{x}_1, \ldots, \vec{x}_N \in \mathbb{R}^n \), and \( \vec{z} \in \mathbb{R}^N \). Then

\[
\vec{z}^* A \vec{z} = \sum_{i,j=1}^{N} g(\vec{x}_i - \vec{x}_j) z_j z_i = \sum_{i,j=1}^{N} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\vec{\xi} \cdot (\vec{x}_i - \vec{x}_j)} d\mu(\vec{\xi}) z_j z_i = \sum_{i,j=1}^{N} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\vec{\xi} \cdot (\vec{x}_i - \vec{x}_j)} z_j z_i \phi(\vec{\xi}) d\mu(\vec{\xi})
\]

\[
= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left| \sum_{i=1}^{N} e^{-i\vec{\xi} \cdot \vec{x}_i} z_i \right|^2 d\mu(\vec{\xi}) \geq 0.
\]

(3.3)

This shows that the function \( g \) is positive-definite. Furthermore, \( g \) is continuous, which follows from the dominated convergence theorem and using that \( \mu \) is a finite measure.

One could now ask the question, if the converse is also true, i.e. whether for every positive-definite and continuous function \( g \) on \( \mathbb{R}^n \) there exists a positive finite measure \( \mu \) such that \( g \) is the Fourier transform of \( \mu \), i.e.

\[
g(\vec{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\vec{\xi} \cdot \vec{x}} d\mu(\vec{\xi}).
\]

(3.4)

Bochner’s theorem gives the answer.

**Theorem 8.** *(Bochner’s theorem)* A function \( g : \mathbb{R}^n \to \mathbb{C} \) is continuous and positive-definite if and only if it is the Fourier transform of a finite positive measure.

**Proof.** One direction we just proved, the other one remains to be shown. We will only prove the theorem in the one-dimensional case \( (n = 1) \). This is sufficient to see how Stone’s theorem comes into play. For the general, multidimensional case we would need to introduce the notion of commutativity of unbounded operators and prove some related results first. (We will however treat commutativity of unbounded operators to some extent in Section 3.3.6.)

Suppose \( g \) is a positive-definite and continuous function. Let \( K \) denote the space consisting of the complex-valued functions on \( \mathbb{R} \), which vanish except at a finite number of points. Then

\[
\langle \psi, \varphi \rangle_g := \sum_{x,y \in \mathbb{R}} g(x - y) \overline{\psi(x)} \varphi(y)
\]

fulfills all the properties required to form an inner product on \( K \), except that it is not positive-definite, but only positive semi-definite, i.e. we could have that \( \langle \varphi, \varphi \rangle_g = 0 \) for some \( \varphi \neq 0 \). It is however not difficult to construct a similar space with an inner product. For this let \( \mathcal{N} \) be the set of all \( \varphi \in K \) for which the semi-inner product vanishes. Then it is a well-known fact that the factor space \( K/\mathcal{N} \) of all equivalence classes modulo \( \mathcal{N} \) forms an inner product space (or pre-Hilbert space) with the induced inner product.

Now suppose that \( t \in \mathbb{R} \) and define \( U_t \) on \( K \) by

\[
(U_t(\varphi))(x) = \varphi(x - t)
\]

(3.6)
(right shift). $U_t$ preserves the form $\varphi \mapsto \langle \varphi, \varphi \rangle_{g}$ and thus maps equivalence classes to equivalence classes (mod $N$). Thus, it induces an isometry on $\mathcal{K}/N$. The same is true for $U_{-t}$ and so this isometry has dense range. Hence we can extend it to a unique unitary operator $\tilde{U}_t$ on the Hilbert space $\mathcal{H} := \mathcal{K}/N$. Furthermore $\tilde{U}_{t+s} = \tilde{U}_t \tilde{U}_s$ and $\tilde{U}_0 = \text{Id}_H$. Moreover, the continuity of $g$ implies that $t \mapsto \tilde{U}_t$ is strongly continuous, which we will show below. Thus, we are in the situation of Stone’s theorem and can write $\tilde{U}_t = e^{itA}$ for a self-adjoint operator $A$ on $\mathcal{H}$. Using the functional calculus in the projection valued-measure formulation, we get

$$\langle \varphi, \tilde{U}_t \psi \rangle_{g} = \int_{\mathbb{R}} e^{it\lambda} d \langle \varphi, P_{\lambda}^{A} \psi \rangle_{g}, \quad (3.7)$$

If we now let $\tilde{\varphi}_0$ be the equivalence class of $1\{0\}$, then

$$g(t) = \langle \tilde{U}_t \tilde{\varphi}_0, \tilde{\varphi}_0 \rangle_{g} = \langle \tilde{\varphi}_0, \tilde{U}_{-t} \tilde{\varphi}_0 \rangle_{g} = \int_{\mathbb{R}} e^{-it\lambda} d \langle \tilde{\varphi}_0, P_{\lambda}^{A} \tilde{\varphi}_0 \rangle_{g}, \quad (3.8)$$

which proves the assertion.

To complete the proof let us now prove the claim that $t \mapsto \tilde{U}_t$ is strongly continuous. We want to show that $t \mapsto \tilde{U}_t \psi$ is continuous for all $\psi \in \mathcal{H} = \mathcal{K}/N$. By Lemma 1 it is enough to show this for all $\psi \in \mathcal{K}/N$. If we denote by $[\psi]$ the equivalence class in $\mathcal{K}/N$ of $\psi \in \mathcal{K}$, then

$$\left\| \tilde{U}_t [\psi] - [\psi] \right\|_2^2 = \left\| [U_t \psi - \psi] \right\|_2^2 \leq \sum_{x,y \in \mathbb{R}} g(x-y)(\psi(x-t) - \psi(x))(\psi(y-t) - \psi(y)). \quad (3.9)$$

A simple calculation indeed shows that the latter term vanishes as $t \to 0$ provided that $g$ is continuous.

A result related to Bochner’s theorem is the following theorem on unitary dilation.

**Theorem 9. (unitary dilation theorem)** Let $f : \mathbb{R} \to \mathbb{C}$. Then the following are equivalent:

1. The function $f$ is positive-definite and continuous.

2. There exists a Hilbert space $\mathcal{H}$, a strongly continuous unitary group $\{U(t)\}_{t \in \mathbb{R}}$, and a $\xi \in \mathcal{H}$ such that $f(t) = \langle U_t \xi, \xi \rangle$ and $\xi$ is cyclic in $\mathcal{H}$, i.e. the linear span of $\{U_t \xi | t \in \mathbb{R}\}$ is dense in $\mathcal{H}$.

Moreover, $U_t$ is unique up to unitary equivalence, i.e. if there exists a second Hilbert space $\mathcal{H}'$, a strongly continuous unitary group $\{V_t\}_{t \in \mathbb{R}}$ on it, and a cyclic vector $\eta$ such that $f(t) = \langle V_t \eta, \eta \rangle$ then there exists a unitary map $S : \mathcal{H} \to \mathcal{H}'$ such that $U_t = S^{-1}V_t S$.

Together with Stone’s and Bochner’s theorem the above theorem forms a group of three theorems, where knowing two of them, one knows all three.
Proof. (1 ⇒ 2) In the proof of Bochner’s theorem we explicitly constructed a Hilbert space \( H \), a unitary operator \( \tilde{U}_t \) on it, and a vector \( \tilde{\varphi}_0 \in H \) such that \( f(t) = \langle \tilde{U}_t \tilde{\varphi}_0, \tilde{\varphi}_0 \rangle \). To complete this direction of the proof we have to show that this \( \tilde{\varphi}_0 \) is cyclic in \( H \). This follows directly from the construction since \( \tilde{U}_t \tilde{\varphi}_0 = [U_t \mathbb{1}_{\{0\}}] = [\mathbb{1}_{\{t\}}] \). The linear span of the \( \mathbb{1}_{\{t\}} \) is the space of all complex-valued functions on \( \mathbb{R} \) which vanish except at a finite number of point which was exactly the space \( K \). The linear span of the equivalence classes of the \( \mathbb{1}_{\{t\}} \) is then \( K/\mathcal{N} \) and this is dense in the Hilbert space \( H = K/\mathcal{N} \).

(2 ⇒ 1) We assume \( f(t) = \langle U_t \xi, \xi \rangle \). Using Stone’s theorem and the spectral measure formulation, this reads \( f(t) = \langle e^{it\lambda} \xi, \xi \rangle = \int_{\mathbb{R}} e^{it\lambda} d \langle \xi, P_\lambda \xi \rangle \). Using the arguments we made at the beginning of this section (3.2-3.3), we get that \( f \) is positive-definite and continuous.

To prove the uniqueness up to unitary equivalence, let us assume we have \( f(t) = \langle U_t \xi, \xi \rangle = \langle V_t \eta, \eta \rangle \) with \( \xi \in H \) and \( \eta \in H' \) both cyclic. We now want to define a unitary map \( S : H \rightarrow H' \) such that \( U_t = S^{-1}V_t S \). First, we map \( U_t \xi \mapsto V_t \eta \) for all \( t \in \mathbb{R} \). For linear combinations of the \( U_t \xi \) the map is defined by linearity, i.e.

\[
\sum_{i=1}^{n} \alpha_n U_{t_i} \xi \mapsto \sum_{i=1}^{n} \alpha_n V_{t_i} \eta.
\]

(3.10)

We have to show that this map is well-defined, i.e. that it is independent of the representation. So let us assume that a vector \( \varphi \in H \) can be written as two different linear combinations:

\[
\psi = \sum_{i=1}^{n} \alpha_n U_{t_i} \xi = \sum_{j=1}^{m} \beta_n U_{s_j} \xi.
\]

(3.11)

Multiplying by \( U_{-\tau} \) for some arbitrary \( \tau \in \mathbb{R} \) and rearranging gives

\[
\sum_{i=1}^{n} \alpha_n U_{t_i-\tau} \xi - \sum_{j=1}^{m} \beta_n U_{s_j-\tau} \xi = 0.
\]

(3.12)

Then, applying the scalar product,

\[
0 = \left\langle \sum_{i=1}^{n} \alpha_n U_{t_i-\tau} \xi - \sum_{j=1}^{m} \beta_n U_{s_j-\tau} \xi, \xi \right\rangle = \sum_{i=1}^{n} \alpha_n f(t_i - \tau) - \sum_{j=1}^{m} \beta_n f(s_j - \tau)
\]

\[
= \left\langle \sum_{i=1}^{n} \alpha_n V_{t_i-\tau} \eta - \sum_{j=1}^{m} \beta_n V_{s_j-\tau} \eta, \eta \right\rangle = \left\langle \sum_{i=1}^{n} \alpha_n V_{t_i} \eta - \sum_{j=1}^{m} \beta_n V_{s_j} \eta, V_\tau \eta \right\rangle.
\]

(3.13)

That means that if some \( \varphi \) is a linear combination of the \( \{V_\tau \eta | \tau \in \mathbb{R}\} \), then

\[
\left\langle \sum_{i=1}^{n} \alpha_n V_{t_i} \eta - \sum_{j=1}^{m} \beta_n V_{s_j} \eta, \varphi \right\rangle = 0
\]

(3.14)
and thus, since these linear combinations are dense in $\mathcal{H}'$ because $\eta$ is cyclic, it follows that

$$\sum_{i=1}^{n} \alpha_i V_{t_i}\eta - \sum_{j=1}^{m} \beta_j V_{s_j}\eta = 0,$$

which proves that the map is well defined. Now, since also $\xi$ is cyclic, we have defined a map on a dense subspace of $\mathcal{H}$ and since it is linear and bounded, it is uniquely extended to all of $\mathcal{H}$. We call this extension $S$.

So now we have a map $S \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$. Surjectivity of $S$ follows directly from construction and the cyclicity of $\eta$. Moreover $S$ is injective since by symmetry we could have defined a well-defined map $T : \mathcal{H}' \to \mathcal{H}$, which then would be the inverse of $S$. So $S$ is a bijection.

We have to check whether $S$ fulfills $U_t = S^{-1}V_tS$. It is enough to check this for finite linear combinations of the $U_t\xi$:

$$S^{-1}V_tS \sum_{i=1}^{n} \alpha_i U_{t_i}\xi = S^{-1}V_t \sum_{i=1}^{n} \alpha_i V_{t_i}\eta = S^{-1} \sum_{i=1}^{n} \alpha_i V_{t_i+t}\eta = \sum_{i=1}^{n} \alpha_i U_{t_i+t}\xi = U_t \sum_{i=1}^{n} \alpha_i U_{t_i}\xi.$$  

Finally, we want to see that $S$ is unitary. Knowing that $S$ is bijective, we still need to show that it is an isometry. Again it suffices to show this for all linear combinations of the $U_t\xi$:

$$\left\| S \sum_{i=1}^{n} \alpha_i U_{t_i}\xi \right\|^2 = \left\langle \sum_{i=1}^{n} \alpha_i V_{t_i}\eta, \sum_{j=1}^{n} \alpha_j V_{t_j}\eta \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \overline{\alpha_j} \langle V_{t_i-t_j}\eta, \eta \rangle = \left\| \sum_{i=1}^{n} \alpha_i U_{t_i}\xi \right\|^2$$

(3.17)

Altogether this proves the assertion of uniqueness up to unitary equivalence. \hfill $\square$

The Hilbert space, the unitary group, and the cyclic vector do not have to be found in the way described in the proof. Often, it is much simpler, as the following example shows. By the above theorem all of the representations are unitarily equivalent.

**Example 5.** Let $f(t) = e^{-\lambda|t|}$. We already know that

$$f(t) = \int_{\mathbb{R}} e^{-i\omega t} \frac{1}{\pi \alpha^2 + \omega^2} d\omega.$$  

(3.18)

Hence, if we choose $\mathcal{H} = L^2(\mathbb{R})$, $U_t = e^{-iTt}$, where $Tf$ is the multiplication operator with $f(\omega) = \omega$, and $\xi(\omega) = \sqrt{\frac{1}{\pi \alpha^2 + \omega^2}}$, then

$$f(t) = \langle U_t\xi, \xi \rangle_{\mathcal{H}}.$$  

(3.19)

Furthermore, $\xi$ is cyclic since the functions $\left\{ e^{i\omega t} \sqrt{\frac{1}{\pi \alpha^2 + \omega^2}} \mid t \in \mathbb{R} \right\}$ span a dense subspace of $L^2(\mathbb{R})$. 

29
3.2 Physical Background

“A philosopher once said, ‘It is necessary for the very existence of science that the same conditions always produce the same results.’ Well, they don’t!”
—RICHARD P. FEYNMAN

The theory of self-adjoint operators on Hilbert spaces is essential for the mathematical description of quantum mechanics. Quantum mechanics is a branch of modern physics describing the dynamics of particles at atomic and subatomic scales.

Newtonian physics are governed by Newton’s laws of motion, e.g. \( \vec{F} = m \vec{a} \) (the second law), where \( \vec{F} \) is the force on a particle of mass \( m \) and \( \vec{a} \) the resulting acceleration. Assuming one knew the position and the momentum of all particles in a system, one could calculate all their positions at any future time \( t \). The resulting universe would be deterministic.

We know however that this is not the case. While at length scales which matter in everyday life, the Newtonian description is (almost) exact, this breaks down at smaller length scales. In quantum mechanics, all particles are (represented by) wave functions. The absolute value squared of the amplitude of the wave function is the probability density of finding the particle at a certain position when measuring. So instead of knowing exactly where a particle will be at a future time, we can merely make a probability statement. In addition, the measurement causes the particle to actually be at the very position where it was measured to be; so the original wave function is changed. A crucial point in understanding quantum mechanics is that every measurement will interfere with the observed system.

Quantum mechanics was the first theory in physics which required abstract mathematical structures like infinite-dimensional Hilbert spaces and operators defined on them (see [Neu55, Mac04] for a good overview).

Any physical system is described by three basic ingredients: states, observables, and dynamics. In Newtonian physics, the state of a system is characterised by the position and the momentum of all particles, corresponding to a single point in phase space. Observables are measurable quantities like energy or angular momentum. Newton’s laws govern the dynamics of the system.

These three types of objects are also included in the mathematical formulation of quantum mechanics and are introduced by the postulates of quantum mechanics:

- Each physical system corresponds to a complex Hilbert space \( \mathcal{H} \) with the inner product of the two vectors \( |\varphi\rangle \) and \( |\psi\rangle \) denoted by \( \langle \varphi | \psi \rangle \). A state of the system is associated with a one-dimensional subspace of that Hilbert space or equivalently with the class of all vectors of unit length in that subspace. These all just differ by a complex phase factor.

\[1\] The question whether wave functions are only mathematical objects describing the physical reality or whether they are part of the reality themselves is one of the many open questions concerning the interpretation of quantum mechanics.
Physical observables are given by densely defined self-adjoint operators on the Hilbert space. The expected value of the observable $A$ for the system in a state represented by the unit vector $|\psi\rangle \in \mathcal{H}$ is given by $\langle \psi | A | \psi \rangle := \langle \psi | (A | \psi \rangle)$. One can show that the possible values for the observable must belong to the spectrum of the operator $A$. If the spectrum is discrete, the possible outcomes of measuring $A$ are exactly its eigenvalues. After the measurement, the system will be in the eigenstate of $A$ corresponding to the measured eigenvalue. (More generally, if the spectrum is not discrete, one can formulate a similar statement using the projection-valued measure.)

In the Schrödinger picture, the dynamics are given by the time evolution operator $U(t) := e^{-i t H}$, where $H$ is the self-adjoint Hamilton operator corresponding to the total energy of the system as observable. (In this text, we are using units in which $\hbar = 1$. In SI units the above would read $U(t) = e^{-i \frac{\hbar}{\mathcal{H}} t}$.) If $|\psi(t)\rangle$ is the state at time $t$, then the state at time $t + s$ is given by

$$|\psi(t + s)\rangle = U(s) |\psi(t)\rangle$$

(3.20)

for all $t, s \in \mathbb{R}$.

We already know the functional calculus, which gives meaning to the definition of the time-evolution operator. For this formulation to be true, one has to assume that $H$ does not explicitly depend on time. More generally, the Schrödinger equation holds:

$$i \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle.$$  

(3.21)

In this text, we will stick to the above operator formulation, which is valid in most isolated physical systems.

In any isolated system, the time-evolution is a unitary operation. This means that if the initial state is given by the unit vector $|\psi_0\rangle = |\psi(t = 0)\rangle$, then any future (or past) state also has unit length. Having unit length means that the probability of finding the particle anywhere is 1, which we demand from an isolated system. If, in addition, the physical description of the system is the same at all times (this corresponds to a Hamiltonian not explicitly depending on time), then we must demand that if we have an identical state at two different times (i.e. $|\psi(t)\rangle = |\psi(s)\rangle$), the time evolution should be the same for both (i.e. $|\psi(t + \tau)\rangle = |\psi(s + \tau)\rangle$). This means that the time evolution operator only depends on the time span $\tau$ but not on the initial time $t$ or $s$. Together with (3.21) this implies that $U(t)$ forms a unitary group. Stone’s theorem then guarantees the existence of a self-adjoint operator $H$, such that it generates the time evolution.

These are the fundamental mathematical postulates of quantum mechanics. To describe real physical systems, we must also specify which operators correspond to which observables. Let us for the sake of simplicity restrict ourselves to a system of one spinless particle. For such a particle the system is described by the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^3)$. 

31
A state $\langle \psi |$ corresponds to a function (i.e. its equivalence class in $L^2$) $\psi(\vec{x})$ with the normalization condition $\int |\psi(\vec{x})|^2 \, d^3x = 1$, where we consider two states to be equal if they only differ by a global (not $\vec{x}$-dependent) phase factor. As mentioned above, the physical interpretation of $\psi(\vec{x})$ is that the probability density of finding the particle at position $\vec{x}$ is given by $|\psi(\vec{x})|^2$.

If the system consists of several particles, then the Hilbert space is the tensor product of the state systems of the single particles. (We will discuss tensor products to some extent in Section 3.3.5.) Also, if we consider a particle which has spin, then its state space is the tensor product of the conventional $L^2$-space and the spin space. The spin operator has no corresponding observable in classical physics and hence cannot be described by just looking at the probability distribution in $L^2(\mathbb{R}^3)$. For all classical observables like energy, position, momentum, or angular momentum this is possible.

We will not include spin or many-particle systems in our considerations since for the purpose of this text, they are just another algebraic complication. So for us, the quantum mechanically relevant Hilbert space will always be the $L^2(\mathbb{R}^n)$ space with $n=3$, or $n=1$ or $n=2$ to simplify matters.

The two fundamental operators are the position and the momentum operator $Q$ and $P$. They have to satisfy the canonical commutation relation

$$PQ - QP = -iI. \quad (3.22)$$

One can show that from this commutation relation it follows that at least $P$ or $Q$ has to be unbounded and that means that the above relation is only a formal one. One has to specify domains for $P$ and $Q$. The standard representation in physics is the Schrödinger representation. Here, as we already assumed, $\mathcal{H} = L^2(\mathbb{R})$. The operator $P$ is taken to be the closure of $-i \frac{d}{dx}$ with domain $\mathcal{S}(\mathbb{R})$, the Schwartz space or the space of rapidly decreasing functions. This operator is essentially self-adjoint on $\mathcal{S}(\mathbb{R})$ (see Lemma 2) and hence $P$ is well-defined and self-adjoint. $Q$ is the maximally defined multiplication operator $T_x$, which multiplies by the function $f(x) = x$, and hence $Q$ is self-adjoint (compare to Section 3.3.2). So indeed, $P$ and $Q$ are both self-adjoint and, as a simple calculation shows, they fulfill the canonical commutation relation on $\mathcal{S}(\mathbb{R})$.

All other operators, which correspond to classical observables have to be built up from the fundamental operators $P$ and $Q$ as we just defined them. If we, for example, want to construct the Hamilton operator $H$ for a single non-relativistic particle, we must look at the Newtonian energy $E = \frac{p^2}{2m} + V(x)$, where $p$ is the momentum and $V$ a potential depending on $x$. By the correspondence rules one then defines

$$\tilde{H} = -\frac{1}{2m} \frac{d^2}{dx^2} + T_V(x) \quad (3.23)$$

on $\mathcal{S}(\mathbb{R})$, where $T_V(x)$ is the multiplication operator with $V(x)$. If the potential is “nice”, one can show that $\tilde{H}$ is essentially self-adjoint on $\mathcal{S}(\mathbb{R})$ and hence obtain the Hamilton operator $H$ as the closure of $\tilde{H}$.

In general, one of the major mathematical problems which arises in quantum mechanics is to prove essential self-adjointness for an operator written down formally using
the momentum and position operators, or if it is not, to find the “correct” self-adjoint
extension.

3.3 Examples I

Now that we have established the necessary theoretic background, we can look at many
interesting examples. Most of them are physics-related and whenever it is possible, we
will try to interpret our results in terms of the physical quantities discussed above.

3.3.1 Translation

The translation is a very simple example of a unitary map. If we can show that with
a certain parametrisation it forms a strongly continuous group, we can apply Stone’s
theorem. Let us first look at the one dimensional case, i.e. we look at the Hilbert space
\( \mathcal{H} = L^2(\mathbb{R}) \).

Definition 14. For \( \varphi \in L^2(\mathbb{R}) \) we define the translation operator by
\[
(U(a)\varphi)(x) = \varphi(x + a),
\]
(3.24)
i.e. \( U(a) \) shifts the function \( \varphi(x) \) to the left by \( a \).

It is clear from the definition that \( U(a) \) is an isometry since the Lebesgue integral is
translation-invariant and as translations are invertible, we have a unitary map for all
\( a \in \mathbb{R} \). Using
\[
(U(a)U(b)\varphi)(x) = (U(b)\varphi)(x + a) = \varphi(x + a + b) = (U(a + b)\varphi)(x),
\]
(3.25)
we note that \( \{U(a)\}_{a \in \mathbb{R}} \) forms a one-parameter group. In the following, we will show
that the translation operator group is also strongly continuous and hence we have a
strongly continuous one-parameter unitary group.

Lemma 1. Let \( X \) be a Banach space and let \( T \subset \mathcal{L}(X) \) be bounded, i.e. \( \sup_{T \in T} \|T\| = c < \infty \). Then in \( T \) are equivalent:

1. strong convergence,

2. strong convergence on a dense subspace \( M \) of \( X \).

Proof. The proof involves an \( \varepsilon \)-argument. We assume strong convergence on \( M \) and
want to show strong convergence on \( X \). So, let \( T_n \) be a sequence of operators in \( T \) such
that \( T_n \varphi \to T \varphi \) as \( n \to \infty \) for all \( \varphi \in M \). Now, let \( \psi \) be an arbitrary vector in \( X \). Then,
since \( M \) is dense in \( X \), we can find a \( \varphi \in M \) such that \( \|\psi - \varphi\| \leq \frac{\varepsilon}{3c} \). Moreover, if we
choose \( n \) sufficiently large, \( \|T_n \varphi - T \varphi\| \leq \frac{\varepsilon}{3} \). Together, this gives
\[
\|T_n \psi - T \psi\| \leq \|T_n \psi - T_n \varphi\| + \|T_n \varphi - T \varphi\| + \|T \varphi - T \psi\| \leq c\frac{\varepsilon}{3c} + \frac{\varepsilon}{3} + c\frac{\varepsilon}{3c} = \varepsilon,
\]
(3.26)
which proves the assertion. \( \square \)
Using the above Lemma\[11\] we can show the following theorem about the strong continuity of the translation group. We will show it in a more general setting, i.e. for \( L^p(\mathbb{R}^n) \) since this produces no further complication.

**Theorem 10.** The translation group on \( L^p(\mathbb{R}^n) \) \((1 \leq p < \infty)\) is strongly continuous, i.e. if we define
\[
(\tau_\vec{a} f)(\vec{x}) = f(\vec{x} + \vec{a}),
\]
then
\[
\lim_{|\vec{a}| \to 0} \tau_\vec{a} f = f
\]
in \( L^p(\mathbb{R}^n) \) for all \( f \in L^p(\mathbb{R}^n) \).

**Proof.** Let \( C_0(\mathbb{R}^n) \) := \( M \) be the continuous (complex-valued) functions on \( \mathbb{R}^n \) with compact support. Then \( C_0(\mathbb{R}^n) \) is dense in \( L^p(\mathbb{R}^n) \) =: \( X \) for \( 1 \leq p < \infty \).\[AE01\]

Moreover, \{\( \tau_\vec{a} | \vec{a} \in \mathbb{R}^n \} := \mathcal{T} \) is bounded by 1 (they are all isometries because of the translation invariance of the Lebesgue measure). Using Lemma\[11\] it is enough to show strong convergence on \( C_0(\mathbb{R}^n) \) to get strong convergence on all of \( L^p(\mathbb{R}^n) \). So, let \( g \in C_0(\mathbb{R}^n) \). Since \( g \) has compact support, we can find a compact set \( K \subset \mathbb{R}^n \) with \( \text{ supp}(\tau_\vec{a} g) \subseteq K \) for all \( |\vec{a}| \leq 1 \). \( g \) is also uniformly continuous and so there exists a \( \delta \in (0,1] \) with \( \|\tau_\vec{a} g - g\|_\infty \leq \varepsilon' \) for all \( \vec{a} \) with \( |\vec{a}| < \delta \). Since \( \text{ supp}(\tau_\vec{a} g - g) \subseteq K \), we get \( \|\tau_\vec{a} g - g\|_p \leq (\lambda^n(K))^\frac{1}{n} \|\tau_\vec{a} g - g\|_\infty \leq (\lambda^n(K))^\frac{1}{n} \varepsilon' \). Choosing \( \varepsilon' = \frac{\varepsilon}{(\lambda^n(K))^{\frac{1}{n}}} \), we get \( \|\tau_\vec{a} g - g\|_p \leq \varepsilon \). So \( \tau_\vec{a} \) is strongly continuous on \( C_0 \) and hence on \( L^p(\mathbb{R}^n) \). \( \square \)

Having shown that the translation forms a strongly continuous one-parameter unitary group, by Stone’s theorem there exists a self-adjoint infinitesimal generator \( A \), s.t. \( U(t) = e^{itA} \). By (3) and (4) of Theorem\[6\] we know that \( D(A) \) is exactly given by all the functions \( \psi \in L^2(\mathbb{R}) \) for which
\[
\lim_{t \to 0} \frac{U(t)\psi - \psi}{t} = \lim_{t \to 0} \frac{\psi(\cdot + t) - \psi(\cdot)}{t}
\]
exists, in which case
\[
A\psi = -i \lim_{t \to 0} \frac{U(t)\psi - \psi}{t}.
\]

If we look at the above limit pointwise, it is clear that for a differentiable function, this is just \(-i\) times the derivative. However, we have to consider the \( L^2 \)-limit. This means there has to exist a square-integrable function \( \varphi \) for which
\[
\lim_{t \to 0} \left\| \frac{U(t)\psi - \psi}{t} - \varphi \right\|_2 = \lim_{t \to 0} \int_\mathbb{R} \left| \frac{\psi(s + t) - \psi(s)}{t} - \varphi(s) \right|^2 ds = 0.
\]

This suggests that it is more convenient to look at the (global) weak derivative instead of the (local) ordinary derivative.

\[2\] Throughout the text we will make use of similar density statements. They are always of the form that some group of continuous, differentiable, compactly supported, or trigonometric functions is dense in \( L^p \).
**Proposition 6.** If the limit \( \lim_{t \to 0} \frac{U(t)\psi - \psi}{t} \) exists, then it is the weak derivative of \( \psi \), i.e. \( \langle \varphi, \eta \rangle = -\langle \psi, \eta' \rangle \) for all \( \eta \in C_0^\infty(\mathbb{R}) \).

**Proof.** Let \( \eta \in C_0^\infty(\mathbb{R}) \). Then

\[
\langle \varphi, \eta \rangle = \lim_{t \to 0} \left\langle \frac{U(t)\psi - \psi}{t}, \eta \right\rangle = \lim_{t \to 0} \left\langle \frac{U(t)\psi - \psi}{t}, \eta \right\rangle = \lim_{t \to 0} \left\langle \psi, \frac{U(-t)\eta - \eta}{t} \right\rangle = -\langle \psi, \eta' \rangle,
\]

(3.32)

where \( \eta' \) is the (ordinary) derivative of \( \eta \) (as a pointwise limit). The fact that \( \eta' \) is also the \( L^2 \)-limit of \( \frac{U(t)\eta - \eta}{t} \) as \( t \to 0 \) follows from the fact that the difference quotient \( \frac{\eta(x+t) - \eta(x)}{t} \) is bounded by \( \sup_x |\eta'(x)| \) (mean value theorem). Then, since \( \eta \) has compact support, we can find a positive \( L^2 \)-function which is a bound for \( \left|\frac{U(t)\eta - \eta}{t}\right| \) and by the dominated convergence theorem \( \eta' \) is also the \( L^2 \)-limit. Hence, by definition, \( \psi \) is weakly differentiable and \( \varphi \) is its weak derivative. \( \square \)

We now know that the domain of \( A \), which we can formally write as

\[
D(A) = \left\{ \psi \in L^2(\mathbb{R}) \left| \lim_{t \to 0} \frac{U(t)\psi - \psi}{t} \text{ exists and is in } L^2(\mathbb{R}) \right. \right\},
\]

(3.33)

is a subset of the weakly differentiable functions in \( L^2(\mathbb{R}) \). We can set \( D(D) := D(A) \) and define the operator \( D \) as the one mapping a function to its weak derivative. Then we get by Theorem 6

\[
A = -iD.
\]

(3.34)

This yields

\[
U(t) = e^{i(-1)tD} = e^{tD},
\]

(3.35)

which formally written as a power series corresponds to Taylor’s theorem.

The physical interpretation is also interesting. First, as a lemma, we proof what we have already stated in Section 3.2.

**Lemma 2.** The operator \( \tilde{P} = -i \frac{d}{dx} \) defined on \( S(\mathbb{R}) \) is essentially self-adjoint. (And hence the momentum operator \( P = \tilde{P} \) is self-adjoint.)

**Proof.** The Hermiticity of \( \tilde{P} \) follows from partial integration as we have seen before. We show essential self-adjointness by using the basic criterion (Corollary 1). We therefore want to show that \( \text{Ran}(\tilde{P} \pm i) \) is dense in \( L^2(\mathbb{R}) \). For this, let \( \varphi(x) \) be an arbitrary \( C_0^\infty(\mathbb{R}) \)-function. We are looking for a function \( \psi(x) \in S(\mathbb{R}) \) such that

\[
-i\psi'(x) \pm i\psi(x) = \varphi(x).
\]

(3.36)

Solving for \( \psi(x) \) gives

\[
\psi(x) = ie^{\pm x} \int_c^x e^{\mp y} \varphi(y)dy.
\]

(3.37)
If we can in both cases choose the integral such that $\psi \in \mathcal{S}(\mathbb{R})$, the assertion follows from the density of $C^\infty_0$ in $L^2$. Let us look at $\psi(x) = i e^{-x} \int e^{+y} \varphi(y) dy$; the other case can be treated symmetrically. The term in the integral is just another test function and hence vanishes outside a compact set $K$. We can choose the integral such that it is zero to the left of $K$ and takes a constant value to the right of $K$. Then, after multiplication with $e^{-x}$, we get a $C^\infty$-function which vanishes to the left of $K$ and decreases like $e^{-x}$ to the right of $K$. Hence, $\psi(x)$ is a rapidly decreasing function.

**Theorem 11.** The momentum operator $P$ is the infinitesimal generator of the translation group, i.e., $U(t) = e^{itP}$, where $U(t)$ is the translation operator.

**Proof.** We have seen that the momentum operator $P$ is the closure of $\tilde{P} := -i \frac{d}{dx}$ with $D(\tilde{P}) = \mathcal{S}(\mathbb{R})$. So $A$ and $P$ are certainly related. Both $P$ (see Lemma 2) and $A$ (as infinitesimal generator) are self-adjoint. Moreover, $\tilde{P}$ is certainly extended by $A$ since for Schwartz functions, the limit in (3.33) exists. Then by Remark 5, $A = P$. So, the operator $A$ we obtained as generator of the translation is identical to the momentum operator.

The result of the above theorem is often summarized as “momentum generates translation”, which is a somewhat trivial statement but gives an a posteriori justification for the name “infinitesimal generator”.

### 3.3.2 Gauge Transformation

The example discussed in the above Section 3.3.1 is perhaps the best-known application of Stone’s theorem in physics. The next example has also a physical interpretation, but describes a very different kind of unitary operation.

Let us stay in $L^2(\mathbb{R})^3$. We know that for a function $\psi$, $|\psi(x)|^2$ is interpreted as the probability density of finding the particle at the point $x$. The wave function $\psi$ itself is not a measurable quantity. This means we can redefine the phase locally without changing the physical reality. Such an operator is called a local gauge transformation.

**Definition 15.** Let $\chi$ be a measurable real-valued function on $\mathbb{R}$. Then we define $U(t)$ by

$$ (U(t)\psi)(x) = e^{it\chi(x)} \cdot \psi(x). $$

(3.38)

$U(t)$ is called local gauge transformation.

$U(t)$ is by definition unitary and fulfills $U(t + s) = U(t)U(s)$. Furthermore $\{U(t)\}_{t \in \mathbb{R}}$ is also strongly continuous as the following proposition shows.

3For this example, we could have easily chosen $L^2(\mathbb{R}^n)$ or $L^2(S)$, where $S$ is any set in $\mathbb{R}^n$. This would not have produced any additional difficulty.

4While a local gauge transformation does not change the physical interpretation of the momentary state of the system, the dynamics can change since they are determined by interference of the wave function and not just the probability density. It is however possible in some cases to replace the derivative in the equation of motion by a covariant derivative, which is simultaneously changed when making a gauge transformation, so that the equation of motion is overall invariant under this transformation. For this, the function $\chi$ is usually required to be differentiable, which we will not assume.
Proposition 7. The local gauge transformation is strongly continuous, i.e. the map \( t \mapsto U(t) \varphi \) is continuous for all \( \varphi \in \mathcal{H} \).

Proof. We write
\[
\|U(t)\psi - U(0)\psi\|^2 = \int_{\mathbb{R}} \left| e^{it\chi(x)} - 1 \right|^2 |\psi(x)|^2 \, dx = \int_{\mathbb{R}} 4 \sin^2 \left( \frac{t\chi(x)}{2} \right) |\psi(x)|^2 \, dx. \tag{3.39}
\]
Let us now consider a sequence \( t_n \xrightarrow{n \to \infty} 0 \). We then define (modulo zero sets) \( A_n := \{ x \in \mathbb{R} \left| |\chi(x)| \geq \frac{1}{\sqrt{t_n}} \right. \} \) and get
\[
\|U(t_n)\psi - \psi\|^2 \leq \int_{\mathbb{R} \setminus A_n} (t_n\chi(x))^2 |\psi(x)|^2 \, dx + 4 \int_{A_n} |\psi(x)|^2 \, dx \leq t_n \int_{\mathbb{R} \setminus A_n} |\psi(x)|^2 \, dx + 4 \int_{A_n} |\psi(x)|^2 \, dx. \tag{3.40}
\]
Now \( \mu(A) := \int_{A} |\psi|^2 \) defines a finite measure on \( \mathbb{R} \) and we get
\[
\|U(t_n)\psi - \psi\|^2 \leq t_n \mu(\mathbb{R} \setminus A_n) + 4\mu(A_n) \leq t_n \mu(\mathbb{R}) + 4\mu(A_n). \tag{3.41}
\]
For \( n \to \infty \) the left term goes to zero. For the right term we observe (assume w.l.o.g. that \( t_n \) is a monotonic sequence) that \( A_1 \supseteq A_2 \supseteq A_3 \ldots \) and \( \mu(A_1) \leq \mu(\mathbb{R}) < \infty \). Then by continuity from above
\[
\lim_{n \to \infty} \mu(A_n) = \mu \left( \bigcap_{n=1}^{\infty} A_n \right) = \mu(\emptyset) = 0. \tag{3.42}
\]
So \( t \to U(t) \) is strongly continuous in \( t = 0 \) and thus everywhere, which follows from the group properties as we have seen several times. \( \square \)

We can now apply Stone’s theorem and obtain the following result.

Theorem 12. The operator \( A \) defined by
\[
D(A) = \left\{ \psi \in L^2(\mathbb{R}) \left| \int_{\mathbb{R}} |\chi(x)|^2 |\psi(x)|^2 \, dx < \infty \right. \right\} \tag{3.43}
\]
and
\[
A\psi = T_\chi \psi = \chi(\cdot)\psi(\cdot). \tag{3.44}
\]
is the infinitesimal generator of the local gauge transformation
\[
(U(t)\psi)(x) = e^{it\chi(x)} \cdot \psi(x). \tag{3.45}
\]
This is of course what one would have expected naïvely.
Proof. Stone’s theorem states that there exists a self-adjoint operator \( A \), such that
\[
U(t) = e^{itA}.
\]
We again look at the set of \( \psi \in L^2(\mathbb{R}) \) for which
\[
-i \lim_{t \to 0} \frac{U(t)\psi - \psi}{t} = -i \lim_{t \to 0} \frac{e^{it\chi(\cdot) - 1}}{t} \psi(\cdot) = A\psi
\] (3.46)
exists in an \( L^2 \)-sense. This will be the domain \( D(A) \) of \( A \). It is clear that if the limit exists, it is just given by
\[
-i \lim_{t \to 0} \frac{U(t)\psi - \psi}{t} = \chi \cdot \psi.
\] (3.47)
Conversely, a simple computation shows that for every function \( \psi(x) \) such that \( \psi(x)\chi(x) \) is in \( L^2(\mathbb{R}) \), the limit does indeed exist. Hence we can read off \( A \) as being the maximally defined multiplication operator \( A = T_\chi \), which acts on a function by pointwise multiplying it with \( \chi \).

3.3.3 Translation in Many Dimensions

Now that we have familiarised ourselves with \( L^2(\mathbb{R}) \), we can extend our considerations to the \( n \)-dimensional case \( L^2(\mathbb{R}^n) \), where of course \( n = 3 \) will be most interesting from a physical point of view. Let us first do so with the translation operator.

Definition 16. Let \( U_i(t) \) be the \( i \)-th canonical translation operator which shifts a function \( \psi \in L^2(\mathbb{R}^n) \) by \( t \) in the direction of \( -\hat{e}_i \), where \( \hat{e}_i \) is the \( i \)-th canonical unit vector. So
\[
(U_i(t)\psi)(\vec{x}) = \psi(\vec{x} + t\hat{e}_i).
\] (3.48)

It can then be shown exactly as in the one-dimensional case that \( \{U_i(t)\}_{t \in \mathbb{R}} \) forms a strongly continuous one-parameter unitary group for all \( i \in \{1, 2, \ldots, n\} \), where we again make use of Theorem \ref{thm:strongly_continuous_group}. By Stone’s theorem we can then write
\[
U_i(t) = e^{itA_i},
\] (3.49)
for a self-adjoint operator \( A_i \) and as in the one-dimensional case we get that \( A_i = -iD_i \), where \( D_i \) maps a function in \( D(A_i) = D(D_i) \subseteq L^2(\mathbb{R}^n) \) to its weak partial derivative.

In the \( n \)-dimensional case we define the weak derivative as follows.

Definition 17. If \( \alpha = (\alpha_1, \ldots, \alpha_n) \) is a multi-index, then \( \varphi \) is called the \( \alpha \)-th weak derivative of \( \psi \) if
\[
\int_{\mathbb{R}^n} \varphi(\vec{x})\eta(\vec{x})d^n x = (-1)^{|\alpha|}\int_{\mathbb{R}^n} \psi(\vec{x})\partial^\alpha \eta(\vec{x})d^n x
\] (3.50)
for all test functions \( \eta \).

The domain \( D(A_i) \) is again given by all the functions in \( L^2(\mathbb{R}^n) \) for which the limit
\[
\lim_{t \to 0} \frac{U_i(t)\psi - \psi}{t}
\] (3.51)
exists, which is a subset of the weakly partially differentially functions in the direction of the \( i \)-th unit vector. So
\[
U_i(t) = e^{itA_i} = e^{itD_i}.
\]
(3.52)

Since the choice of a base of \( \mathbb{R}^n \) is arbitrary, we can describe translations into any given direction in the same way.

We can again interpret the result physically. Analogously to the one-dimensional case, one gets:

**Theorem 13.** The \( i \)-th momentum operator \( P_i \), which is the closure of \( \tilde{P}_i = -i\frac{\partial}{\partial x} \) defined on \( \mathcal{S}(\mathbb{R}^n) \), is the infinitesimal generator of the translation group in the direction of the \( i \)-th coordinate, i.e. \( U_i(t) = e^{tP_i} \).

### 3.3.4 Rotation in Cartesian Coordinates

Now that we have looked at translations in \( \mathbb{R}^n \), we can also consider linear maps which can be parametrised so that they form a unitary group. Rotations around a fixed axis come immediately into one’s mind. Let us for convenience first look at rotations in \( \mathbb{R}^2 \) since rotations (around a fixed axis) in higher dimensions can be reduced to that case by choosing an appropriate coordinate system.

**Definition 18.** We define the rotation matrix
\[
R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},
\]
(3.53)
which rotates a vector in \( \mathbb{R}^2 \) counterclockwise around the origin by the angle \( \theta \).

With this we define the rotation operator on \( L^2(\mathbb{R}^2) \).

**Definition 19.** The rotation operator \( U_\theta \) on \( L^2(\mathbb{R}^2) \) is defined by
\[
(U_\theta \psi)(\vec{x}) = \psi(R_\theta \vec{x}).
\]
(3.54)
It rotates the function \( \psi \) clockwise around the origin by an angle of \( \theta \).

**Proposition 8.** The rotation forms a strongly continuous one-parameter unitary group.

**Proof.** We note that the rotation operator is unitary for all \( \theta \) since it is an isometry, namely
\[
\|U_\theta \psi\|^2 = \int_{\mathbb{R}^2} |(U_\theta \psi)(\vec{x})|^2 \, d^2x = \int_{\mathbb{R}^2} |\psi(R_\theta \vec{x})|^2 \, d^2x \\
\overset{\vec{y} = R_\theta \vec{x}}{=} \int_{\mathbb{R}^2} |\psi(\vec{y})|^2 \, d^2y = \int_{\mathbb{R}^2} |\psi(\vec{y})|^2 \, d^2y = \|\psi\|^2
\]
(3.55)
and a bijection, namely
\[
(U_{-\theta}U_\theta \psi)(\vec{x}) = (U_\theta \psi)(R_{-\theta} \vec{x}) = \psi(R_\theta R_{-\theta} \vec{x}) = \psi(\vec{x}) = (U_\theta U_{-\theta} \psi)(\vec{x}).
\]
(3.56)
Furthermore,

\[(U_{\theta}U_{\phi}\psi)(\vec{x}) = (U_{\phi}\psi)(R_{\theta}\vec{x}) = \psi(R_{\phi}R_{\theta}\vec{x}) = (U_{\phi+\theta}\psi)(\vec{x}),\]

where we used the well known property that

\[R_{\phi}R_{\theta} = R_{\phi+\theta},\]

which can be easily shown using trigonometric addition theorems. So, if we can show that \(\theta \mapsto U_{\theta}\) is strongly continuous, we have shown that the \(U_{\theta}\) form a strongly continuous one-parameter unitary group.

We need to show that \(U_{\theta}\psi \rightarrow \psi\) in \(L^{2}(\mathbb{R}^{2})\) as \(\theta \rightarrow 0\) for all \(\psi \in C_{0}(\mathbb{R}^{2})\), i.e. for all continuous functions with compact support.

Let \(\psi \in C_{0}(\mathbb{R}^{2})\). Then \(\psi\) is uniformly continuous. Hence there is a \(\delta > 0\) such that \(||\psi(\vec{x}) - \psi(\vec{y})|| < \varepsilon'\) for all \(\vec{x}, \vec{y} \in \mathbb{R}^{2}\) with \(||\vec{x} - \vec{y}|| < \delta\). Also note that since \(\psi\) has compact support, there is an \(r > 0\) such that \(\text{supp}(\psi) \subseteq B_{r}(0)\). Thus, \(||R_{\theta}\vec{x} - \vec{x}|| \leq r\theta\).

So making \(\theta\) smaller than \(\frac{\delta}{r}\) we get \(||U_{\theta}\psi - \psi||_{\infty} \leq \varepsilon'\) and choosing \(\varepsilon' = \frac{\varepsilon}{(\lambda^{2}(K))^{\frac{1}{2}}}\) we get \(||U_{\theta}\psi - \psi||_{p} \leq (\lambda^{2}(K))^{\frac{1}{2}} ||U_{\theta}\psi - \psi||_{\infty} \leq \varepsilon\). Hence the rotation group is strongly continuous, which completes the proof.

We can then once again apply Stone’s theorem and determine the infinitesimal generator \(A\). Again \(D(A)\) is given by all the \(\psi \in L^{2}(\mathbb{R}^{2})\) for which

\[-i\lim_{\theta \rightarrow 0} \frac{U_{\theta}\psi - \psi}{\theta} = A\psi\]

exists. It is not directly apparent what the above limit means (if it exists).

**Proposition 9.** If the limit \(\lim_{\theta \rightarrow 0} \frac{U_{\theta}\psi - \psi}{\theta}\) exists, it is given by \(\varphi = x_{1} \cdot D_{2}\psi - x_{2} \cdot D_{1}\psi\), where \(D_{i}\psi\) is the weak partial derivative of \(\psi\) in the direction of the \(i\)-th unit vector.

**Proof.** To investigate this, let us assume the limit exists and call it \(\varphi\). Now let \(\eta \in C_{0}^{\infty}(\mathbb{R}^{2})\). Then

\[\langle \varphi, \eta \rangle = \lim_{\theta \rightarrow 0} \left\langle \frac{U_{\theta}\psi - \psi}{\theta}, \eta \right \rangle = \lim_{\theta \rightarrow 0} \left\langle \psi, \frac{U_{-\theta}\eta - \eta}{\theta} \right \rangle.\]

40
On the other hand since $\eta \in C_0^\infty(\mathbb{R}^2)$ we get pointwise

$$
\lim_{\theta \to 0} \frac{U_\theta \eta(\vec{x}) - \eta(\vec{x})}{\theta} = \lim_{\theta \to 0} \frac{\eta(R_\theta \vec{x}) - \eta(\vec{x})}{\theta}
$$

$$
= \lim_{\theta \to 0} \frac{\eta \left( \begin{array}{cc}
\cos(\theta) & \sin(\theta) \\
-\sin(\theta) & \cos(\theta)
\end{array} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \eta \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}{\theta}
$$

$$
= \lim_{\theta \to 0} \frac{\eta \left( \begin{array}{c}
\cos(\theta)x_1 + \sin(\theta)x_2 \\
-\sin(\theta)x_1 + \cos(\theta)x_2
\end{array} \right) - \eta \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}{\theta}
$$

$$
= \lim_{\theta \to 0} \frac{\eta \left( \begin{array}{c}
\cos(\theta)x_1 + \sin(\theta)x_2 \\
-\sin(\theta)x_1 + \cos(\theta)x_2
\end{array} \right) - \eta \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}{\theta}
$$

$$
= \lim_{\theta \to 0} \frac{\eta \left( \begin{array}{c}
x_1 \\
-\sin(\theta)x_1 + \cos(\theta)x_2
\end{array} \right) + \lim_{\theta \to 0} \frac{\eta \left( \begin{array}{c}
x_1 + \theta x_2 \\
-\theta x_1 + x_2
\end{array} \right)}{\theta} - \eta \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}{\theta}
$$

$$
= \lim_{\theta \to 0} \frac{\eta \left( \begin{array}{c}
x_1 + \theta x_2 \\
x_2
\end{array} \right) - \eta \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}{\theta x_2} + \lim_{\theta \to 0} \frac{\eta \left( \begin{array}{c}
x_1 + \theta x_2 \\
-\theta x_1 + x_2
\end{array} \right)}{\theta x_2} - \eta \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}
$$

$$
= x_2 \partial_1 \eta(\vec{x}) - x_1 \partial_2 \eta(\vec{x}).
$$

(3.61)

Since $\eta \in C_0^\infty(\mathbb{R}^2)$, the pointwise limit is also the $L^2$-limit and we have

$$
\langle \varphi, \eta \rangle = - \langle \psi, x_1 \partial_2 \eta - x_2 \partial_1 \eta \rangle.
$$

(3.62)

Hence, in the sense of weak derivatives

$$
\lim_{\theta \to 0} \frac{U_\theta \psi - \psi}{\theta} = \varphi = x_1 \cdot D_2 \psi - x_2 \cdot D_1 \psi,
$$

(3.63)

if it exists.

We have therefore identified the generator $A$ to be $-i(T_{x_1}D_2 - T_{x_2}D_1)$ with its domain being all functions for which the limit in (3.63) exists and lies again in $L^2(\mathbb{R}^2)$. As final result we can write

$$
e^{(x_1D_2 - x_2D_1)\theta} = U_\theta.
$$

(3.64)

We again find an intriguing physical interpretation of the above result. Formally, the generator $A = -i(T_{x_1}D_2 - T_{x_2}D_1)$ of the rotation corresponds to the observable $L_3 = xp_y - yp_x$ (making use of the correspondence rules for the momentum and position operator), which is the third component of the angular momentum $\vec{L} = \vec{x} \times \vec{p}$. Since the infinitesimal generator is per se self-adjoint, we can indeed view it as the operator corresponding to $L_3$. Hence, one says that “angular momentum generates rotation”.

41
3.3.5 Rotation in Polar Coordinates

As the reader may have noticed, the description of the rotation operator in Cartesian coordinates, i.e. in $L^2(\mathbb{R}^2)$, is a little bit laborious. Simpler and more efficient is the description in polar coordinates. We express the plane by $L^2(\mathbb{R}^+ \otimes L^2(D))$, where $\otimes$ denotes the tensor product and $D = [0, 2\pi)$ with addition defined modulo $2\pi$. (Hence, we identify 0 with $2\pi$, which means that the resulting space $D$ has no “endpoints” at 0 or $2\pi$.) $\mathbb{R}^+$ stands of course for the radius $r = \sqrt{x^2 + y^2}$ and $D$ for the angle $\phi = \arctan(y, x)$.

The rotation operator $U_\theta$ can now be defined much simpler.

**Definition 20.** The rotation operator in polar coordinates is defined by

$$(U_\theta \psi)(r, \phi) = \psi(r, \phi + \theta).$$

(3.65)

Hence, $U_\theta$ acts as identity on $L^2(\mathbb{R}^+)$ and as simple shift operator on $L^2(D)$.

We can write $U_\theta$ as a tensor product

$$U_\theta = \text{Id} \otimes V_\theta,$$

(3.66)

where we denote by $\text{Id}$ the identity map on $L^2(\mathbb{R}^+)$ and by $V_\theta$ the shift operator on $L^2(D)$. This tensor product is defined via its action on a multiplicative function $\psi$, which means that $\psi$ is the tensor product of two functions $\psi_1 \in L^2(\mathbb{R}^+)$ and $\psi_2 \in L^2(D)$, i.e.

$$\psi(r, \phi) = (\psi_1 \otimes \psi_2)(r, \phi) := \psi_1(r) \cdot \psi_2(\phi).$$

(3.67)

The action is then given by

$$\text{Id} \otimes V_\theta(\psi) = \text{Id} \otimes V_\theta(\psi_1 \otimes \psi_2) := (\text{Id} \psi_1) \cdot (V_\theta \psi_2) = U_\theta \psi.$$  (3.68)

That the family $\{U_\theta\}_{\theta \in \mathbb{R}}$ forms a strongly continuous unitary group is obvious from the treatment of the shift operator on $L^2(\mathbb{R})$.

We determine the generator $A$ of the group by looking at the limit

$$iA\psi = \lim_{\theta \to 0} \frac{U_\theta \psi - \psi}{\theta}$$

(3.69)

whenever it exists. If the function is differentiable with respect to $\phi$, this limit is simply the partial derivative $\frac{\partial \psi}{\partial \phi}$. To determine the generator exactly, we have to find a domain of essential self-adjointness for $-i \frac{\partial}{\partial \phi}$. We define $\tilde{A} = -i \frac{\partial}{\partial \phi}$ on the linear span of functions, which are tensor products of the form $\psi = \psi_1 \otimes \psi_2$, where $\psi_1 \in L^2(\mathbb{R}^+)$ and $\psi_2 \in C^\infty(D)$ (differentiability on $D$ includes differentiability at the point 0 = $2\pi$). The operator $\tilde{A}$ is symmetric (the boundary terms occurring when doing the partial integration vanish because of the periodicity) and hence we can show essential self-adjointness by using Corollary 1.

---

5More on tensor products in this context can be found in [RS72].
Proposition 10. The operator \( \tilde{A} \) defined as above is essentially self-adjoint on its domain.

Proof. We want to show that \( \text{Ran}(\tilde{A} \pm i) \) is dense in \( L^2(\mathbb{R}_+ ) \otimes L^2(\mathbb{D}) \). It is enough to show that all functions \( g \otimes f_n \) with \( f_n(\phi) = e^{in\phi} \) (\( n \in \mathbb{Z} \)) and \( g \in L^2(\mathbb{R}_+) \) are in \( \text{Ran}(\tilde{A} \pm i) \) since they span a space which is dense in \( L^2(\mathbb{R}_+) \otimes L^2(\mathbb{D}) \). A simple calculation yields

\[
(\tilde{A} \pm i) \left( \frac{1}{n \pm i} g \otimes e^{in\phi} \right) = g \otimes e^{in\phi}. \tag{3.70}
\]

Furthermore, \( \frac{1}{n \pm i} e^{in\phi} \) is in \( C^\infty(\mathbb{D}) \). Hence, \( \tilde{A} \) is essentially self-adjoint. \( \square \)

Theorem 14. The closure of \( \tilde{A} \) is the infinitesimal generator of the rotation in polar coordinates.

Proof. We know that \( \tilde{A} \) is self-adjoint. So, we have to show that this operator \( \tilde{A} \) is exactly the operator \( A \), we got as the generator of our unitary group. This is however simple, since (as we already said) for all \( C^\infty(\mathbb{D}) \)-functions, the limit (3.69) exists (because of the same arguments we made for the translation operator on \( \mathbb{R} \)). Hence \( A \supseteq \tilde{A} \) and since both sides are self-adjoint, we get \( A = \tilde{A} \) by Remark 5. \( \square \)

We interpret the result in physical terms. The chain rule gives that

\[
\frac{\partial}{\partial \phi} = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}, \tag{3.71}
\]

which we already identified as \( iL_3 \), when we worked in Cartesian coordinates. Hence, we get that \( e^{L_3 \theta} = U_\theta \), which again says that “angular momentum generates rotation”, only that we now worked in polar coordinates.

3.3.6 Rotations in Three Dimensions

As already mentioned, we actually want to describe rotations in \( \mathbb{R}^3 \). They are described by the special orthogonal group \( SO(3) \) of all orthogonal \( 3 \times 3 \)-matrices with unit determinant. They describe all proper (or physical) rotations. (If we allowed the determinant to be \(-1\), we include improper rotations or reflections.) It is a well known fact, that for every rotation in three dimensions one can find an axis through the origin left invariant by the rotation. Choosing the correct coordinate axes, the problem then reduces to the two-dimensional examples discussed above. In fact, for every axis through the origin, there is a one-parameter strongly continuous unitary group, describing all rotations around that axis. Moreover, if we use the three rotation groups around the Cartesian axes (or any other orthogonal coordinate system axes), we can write any rotation as a composition of these (Euler angles). It is important to note that while the rotations in one one-parameter unitary group commute, rotations from different subgroups do not commute in general, as the following example shows.
Example 6. Consider rotations around the $x$- and $y$-axes about an angle of $\frac{\pi}{2}$. Their rotation matrices are given by

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$  \hspace{1cm} (3.72)

and

$$R_y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$  \hspace{1cm} (3.73)

The products are

$$R_x R_y = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$  \hspace{1cm} (3.74)

and

$$R_x R_y = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}.$$  \hspace{1cm} (3.75)

Hence $R_x$ and $R_y$ do not commute. Consequently, the corresponding rotation operators do not commute.

The fact that the rotation operators around different axes do not commute has also important physical consequences. Let us assume, we are working with the unitary groups around the $x$-, $y$- and $z$-axis. They are generated by the infinitesimal generators, which we identified as the different components of the angular momentum ($L_x$, $L_y$, $L_z$), i.e. the self-adjoint operators corresponding to these.

Operators which commute, correspond to physical observables which can be measured simultaneously. These observables are called commensurable. In general, if the operators do not commute, the Heisenberg uncertainty principle\(^6\) dictates that one cannot measure both quantities arbitrarily exact at the same time.

From a mathematical point of view, the problem is the description of commutativity for unbounded operators. In contrast to bounded operators they do not form an algebra, nor even a linear space since they are all defined on different domains. Therefore, just writing down commutation relations like $AB - BA = 0$ (for a certain domain) is not the best approach. Hence, one has to come up with other ways of describing commutativity.

Indeed, for self-adjoint (and possibly unbounded) operators, it is possible to define a notion of commutativity [RS72]. One makes use of the projection-valued measure.

Definition 21. Two self-adjoint (possibly unbounded) operators are said to commute if and only if all the projections in their associated projection-valued measures (2.28) commute.

\(^6\)To be more exact: A generalisation of it due to [Rob29].
Remark 8. These projections are bounded operators on all of $\mathcal{H}$ and so commutativity is well-defined. Moreover, the spectral theorem for bounded self-adjoint operators gives that this definition is consistent with the usual definition of commutativity for bounded operators.

It follows directly from the spectral theorem that if $A$ and $B$ commute by the above definition, then all bounded Borel functions of $A$ and $B$ commute. In addition, the following holds.

**Theorem 15.** Two self-adjoint operators $A$ and $B$ commute (by the above definition) if and only if the unitary groups generated by them commute (in the ordinary sense).

**Proof.** The "$\Rightarrow$"-direction follows directly from the functional calculus as mentioned above. For the "$\Leftarrow$"-direction, we make use of some basic properties of the Fourier transform. Let $f \in S(\mathbb{R})$. Then, using Fubini’s theorem, we get

$$
\int_{\mathbb{R}} f(t) \langle e^{iAt} \varphi, \psi \rangle \, dt = \int_{\mathbb{R}} f(t) \left( \int_{\mathbb{R}} e^{-i\lambda t} d\langle P_A^\lambda \varphi, \psi \rangle \right) \, dt
= \sqrt{2\pi} \int_{\mathbb{R}} \hat{f}(\lambda) d\langle P_A^\lambda \varphi, \psi \rangle = \sqrt{2\pi} \langle \varphi, \hat{f}(A) \psi \rangle.
$$

(3.76)

Using the assumption $e^{iAt} e^{iBs} = e^{iBs} e^{iAt}$ and Fubini’s theorem again, this gives

$$
\langle \varphi, \hat{f}(A) \hat{g}(B) \psi \rangle = \int_{\mathbb{R}} \int_{\mathbb{R}} f(t) g(s) \langle \varphi, e^{-iAt} e^{-iBs} \rangle \, ds \, dt
= \langle \varphi, \hat{g}(B) \hat{f}(A) \psi \rangle.
$$

(3.77)

Hence, for all $f, g \in S(\mathbb{R})$, $\hat{f}(A) \hat{g}(B) = \hat{g}(B) \hat{f}(A)$. Using that the Fourier transform is a bijection from $S(\mathbb{R})$ to $S(\mathbb{R})$, we conclude that $f(A)g(B) = g(B)f(A)$ for all $f, g \in S(\mathbb{R})$. The characteristic function $1_M$ of a Borel set $M$ can be expressed as the pointwise limit of a sequence of uniformly bounded functions $f_n$ in $S(\mathbb{R})$. By the functional calculus, $f_n(A) \to P_M^A$ strongly. Similarly, we can find $g_n$ in $S(\mathbb{R})$ such that $g_n(B) \to P_N^B$ strongly for any Borel set $N$. Since the $f_n$ and $g_n$ are uniformly bounded and $f_n(A)g_n(B) = g_n(B)f_n(A)$ as $f_n, g_n \in S(\mathbb{R})$, we get that $P_M^A$ and $P_N^B$ commute.

We saw in Example 6 that the rotations around different axes do not commute and hence by the above considerations, the different components of the angular momentum do not commute. This is a well-known fact in physics. It is not possible to measure more than one component of the angular momentum at the same time.\footnote{It is however possible to measure one component and the magnitude of the angular momentum simultaneously.}

In fact, the angular momentum operators obey certain commutation relations. These are usually written in the form

$$
[L_i, L_j] = i\varepsilon_{ijk} L_k,
$$

(3.78)
where \([A, B] := AB - BA\) is the commutator of \(A\) and \(B\) and \(\varepsilon_{ijk}\) is the Levi-Civita symbol, which takes the value 1 if \((i, j, k)\) is an even permutation of \((1, 2, 3)\), the value -1 if it is an odd permutation of \((1, 2, 3)\), and 0 if any index is repeated. The Einstein summation convention is used, where one sums over doubly appearing indices. Generalizing our findings from the two-dimensional case, the \(L_i\) are given by

\[
L_i = \varepsilon_{ijk} x_j \frac{\partial}{\partial x_k}.
\] (3.79)

This commutation relation (3.78) is not a valid operator equation. It only becomes meaningful, if applied to a certain domain. In this case, if we apply the equation to a sufficiently differentiable function \(\psi\), we get the equation

\[
L_i(L_j \psi) - L_j(L_i \psi) = i\varepsilon_{ijk} L_k \psi,
\] (3.80)

which then is a good equation and its correctness can be easily verified using standard differentiation rules, though the calculation is quite lengthy.

### 3.3.7 Dilation

We have now considered translations, rotations, and phase transformations. Another simple operation one can think of is dilation. Of course it has to be defined in a way that it is unitary. Obviously, for every \(\lambda \neq 0\) the operator \(U\) defined by

\[
(U\psi)(x) = \sqrt[\lambda]{\psi}(\lambda x)
\] (3.81)

is an isometry from \(L^2(\mathbb{R})\) to \(L^2(\mathbb{R})\) and also a bijection and thus it is unitary.

**Definition 22.** We define the unitary **dilation operator** by

\[
U(\lambda)\psi(x) = e^{\frac{\lambda}{2}}\psi(e^{\lambda}x).
\] (3.82)

Then \(U(\lambda)U(\mu) = U(\lambda + \mu)\). So we are left with checking whether the map \(\lambda \mapsto U(\lambda)\) is strongly continuous.

**Proposition 11.** The dilation operator is strongly continuous.

**Proof.** The proof is similar to showing strong continuity for the rotation or translation operator. First note that

\[
\|U(\lambda)\psi - \psi\| = \left\|e^{\frac{\lambda}{2}}\psi(e^{\lambda}x) - \psi(x)\right\| \leq \left\|(e^{\frac{\lambda}{2}} - 1)\psi(e^{\lambda}x)\right\| + \left\|\psi(e^{\lambda}x) - \psi(x)\right\|
\]

\[
= \frac{e^{\frac{\lambda}{2}} - 1}{e^{\frac{\lambda}{2}}} \|\psi\| + \left\|\psi(e^{\lambda}x) - \psi(x)\right\|. \tag{3.83}
\]

The first term obviously goes to zero as \(\lambda \to 1\). So let us study the second term. As for translation and rotation is is enough that the term goes to zero as \(\lambda \to 0\) for all \(\psi \in C_0\) (using Lemma 1). For those functions we get (since they are uniformly continuous) that \(\left\|\psi(e^{\lambda}x) - \psi(x)\right\|_\infty\) goes to zero and thus \(\left\|\psi(e^{\lambda}x) - \psi(x)\right\|_2\) goes to zero as \(\psi\) has compact support. So indeed, the \(U(\lambda)\) form a unitary one-parameter strongly continuous group.

\[\square\]
Again, we can apply Stone’s theorem and determine the infinitesimal generator $A$ of $U(\lambda)$. We know $iA\psi$ is given by
\begin{equation}
\lim_{\lambda \to 0} \frac{U(\lambda)\psi - \psi}{\lambda} \tag{3.84}
\end{equation}
if it exists.

Now let us try to determine what this limit is. For this, let us first assume that $\psi \in C^1$, i.e. $\psi$ is continuously differentiable. Then the pointwise limit is for any $x$ given by
\begin{equation}
\lim_{\lambda \to 0} \frac{e^{\frac{\lambda}{2}}\psi(e^{\lambda}x) - \psi(x)}{\lambda} = \lim_{\lambda \to 0} \left( \frac{1}{2}e^{\frac{\lambda}{2}}\psi(e^{\lambda}x) + e^{\frac{\lambda}{2}}\psi'(e^{\lambda}x)e^{\lambda}x \right) = \frac{1}{2}\psi(x) + x \cdot \psi'(x), \tag{3.85}
\end{equation}
where we used l’Hôpital’s rule and the continuity of $\psi$ and $\psi'$. The next question is under which premises this convergence is also in $L^2$. For this we have to assume in addition that $\psi$ has compact support, so $\psi \in C^1_0$. Then:
\begin{equation}
\left| \frac{e^{\frac{\lambda}{2}}\psi(e^{\lambda}x) - \psi(x)}{\lambda} \right| = \left| \frac{e^{\frac{\lambda}{2}}\psi(x + (e^{\lambda} - 1)x) - \psi(x)}{\lambda} \right|
\leq \left| \frac{e^{\frac{\lambda}{2}} \left( \psi(x) + (e^{\lambda} - 1)x \| \psi' \|_\infty \right) - \psi(x)}{\lambda} \right|
= \left| \frac{e^{\frac{\lambda}{2}} - 1}{\lambda} \psi(x) + \frac{(e^{\lambda} - 1)\| \psi' \|_\infty}{\lambda} x \right|. \tag{3.86}
\end{equation}
If we assume that we have a sequence $(\lambda_n)$ converging to 0, then there is a compact set $K$, such that the left hand side of the equation is supported inside of it for all $\lambda_n$. On the other hand, the right hand side is monotonic in $\lambda$. So if we set $\Lambda = \sup_n (\lambda_n)$, then with
\begin{equation}
g(x) = \begin{cases}  \frac{e^{\frac{\lambda}{2}} - 1}{\Lambda} \psi(x) + \frac{(e^{\lambda} - 1)\| \psi' \|_\infty}{\Lambda} x & \text{if } x \in K \\ 0 & \text{else} \end{cases} \tag{3.87}
\end{equation}
we have found an $L^2$-function which dominates the left hand side and thus the left hand side also converges in $L^2$ to its pointwise limit by the dominated convergence theorem.

As we did in other examples, we could also look at the scalar product of any function $\psi$ for which the limit (3.84) exits with a $C^0_\infty$-function $\eta$. Then
\begin{equation}
\langle \psi, \eta \rangle = \left\langle \lim_{\lambda \to 0} \frac{U(\lambda)\psi - \psi}{\lambda}, \eta \right\rangle = \lim_{\lambda \to 0} \left\langle \frac{U(\lambda)\psi - \psi}{\lambda}, \eta \right\rangle = -\lim_{\lambda \to 0} \left\langle \psi, \frac{U(\lambda)\eta - \eta}{\lambda} \right\rangle
= -\left\langle \psi, \lim_{\lambda \to 0} \frac{U(\lambda)\eta - \eta}{\lambda} \right\rangle = -\left\langle \psi, \frac{1}{2}\eta + x \cdot \eta' \right\rangle, \tag{3.88}
\end{equation}
where the limit in the last step exists by the above calculations and since $C^1_0 \supset C^0_\infty$. 

47
If $\psi$ were sufficiently differentiable, integration by parts would give us
\[
-\left\langle \psi, \frac{1}{2} \eta + x \cdot \eta' \right\rangle = -\left\langle \psi, \frac{1}{2} \eta \right\rangle - \left\langle \psi \cdot x, \eta' \right\rangle = -\left\langle \frac{1}{2} \psi, \eta \right\rangle + \left\langle x \cdot \psi' + \psi, \eta \right\rangle
\]
which means that the limit $\lim_{\lambda \to 0} \frac{U(\lambda)}{\lambda} \psi - \psi$ is the “weak version” of
\[
\frac{1}{2} \psi + x \cdot \psi'
\]
if it exists.

### 3.3.8 Modified Translation

The unitary groups we have seen so far, were all more or less standard applications of Stone’s theorem. We calculated the infinitesimal generators and they were given by derivatives and/or multiplication operators. We were always able to determine the domain of the generator as those $\psi$ for which the limit $\lim_{t \to 0} U(t) \psi - \psi$ exists. In some cases we were able find another, more accessible characterisation of the domain. The next example should be understood as a warning that the domain of an operator really is crucial and that the unitary groups generated by two formally identical operators with different domains can look very different.

Let us look at an interesting modification of the translation operator. We define the operator $U_\delta(a)$ as the operator from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$ which translates the function by $a$ to the left (exactly as the standard translation operator in Section 3.3.1) but every part of the function moving over the origin is multiplied by a constant phase factor $e^{i\delta}$. For negative $a$, the part of the function moving over the origin in the other direction is multiplied by $e^{-i\delta}$. This operation is unitary since it is invertible and isometric. A proper definition is:

**Definition 23.** We define the **modified translation operator** by
\[
(U_\delta(a) \psi)(x) = \psi(x + a) e^{i\delta \mathbb{1}_{[-a,0]}(x)},
\]
where $\mathbb{1}_{[-a,0]}$ is the characteristic function on the interval $[-a,0]$ (or $[0,-a]$ if $a$ is negative).

Then, for $a, b \geq 0$
\[
(U_\delta(a)(U_\delta(b) \psi))(x) = (U_\delta(b) \psi)(x + a) e^{i\delta \mathbb{1}_{[-a,0]}(x)}
\]
\[
= \psi(x + a + b) e^{i\delta \mathbb{1}_{[-b,a]}(x+a)} e^{i\delta \mathbb{1}_{[-a,0]}(x)}
\]
\[
= \psi(x + a + b) e^{i\delta \mathbb{1}_{[-b-a,-a]}(x)} e^{i\delta \mathbb{1}_{[-a,0]}(x)}
\]
\[
= \psi(x + a + b) e^{i\delta \mathbb{1}_{[-(a+b),0]}(x)} = (U_\delta(a + b) \psi)(x).
\]

The same holds for other sign combinations of $a$ and $b$. This means, that the modified translation forms a group.
Proposition 12. The modified translation group is strongly continuous.

Proof. We proof the assertion using Lemma 1 and hence we just have to show that $U_\delta(a)\psi \to \psi$ as $a \to 0$ for all $\psi \in C_0(\mathbb{R})$. This we have already done for the ordinary translation operator (see Theorem 10), which we shall call $V(a)$ in this context. Then $\|U_\delta(a)\psi - \psi\| \leq \|U_\delta(a)\psi - V(a)\psi\| + \|V(a)\psi - \psi\|$, where the first term goes to zero as $a \to 0$ since $U_\delta(a)\psi$ and $V(a)\psi$ only differ on an interval of length $a$ and the second terms goes to zero as shown in Theorem 10. Lemma 1 gives the assertion.

This means we have a strongly continuous one-parameter unitary group. Now let us try to find out, for which $\psi$ in $L^2(\mathbb{R})$ the limit

$$\lim_{a \to 0} \frac{U_\delta(a)\psi - \psi}{a} =: iA_\delta \psi$$

exists and what it looks like in that case. $A_\delta$ is then the infinitesimal generator of $U_\delta(a)$.

For this let us assume that $\psi \in C^1$. Then for $x \neq 0$

$$\frac{(U_\delta(a)\psi)(x) - \psi(x)}{a} \to \psi'(x)$$

pointwise since the phase shift only occurs within a radius of $a$ around the origin. Thus by making $a$ sufficiently small the limit becomes the normal difference quotient, which converges to the derivative. Only at the point $x = 0$, the pointwise limit might not exist. One can easily show that for a differentiable function the pointwise limit in 0 exists if and only if $\psi(0) = 0$. On the other hand it does not matter, since the origin is only a zero set, so if the $L^2$-limit exists it will just be the derivative. The interesting question is of course under which circumstances the $L^2$-limit exists.

Proposition 13. Let $\psi \in C^1_0$. Then the $L^2$-limit $\lim_{a \to 0} \frac{U_\delta(a)\psi - \psi}{a}$ exists if and only if $\psi(0) = 0$.

Proof. To show this, assume $\psi(0) = 0$. Then

$$\left\| \frac{U_\delta(a)\psi - \psi}{a} \right\| \leq \left\| \frac{V(a)\psi - \psi}{a} \right\| + \left\| \frac{U_\delta(a)\psi - \psi}{a} - \frac{V(a)\psi - \psi}{a} \right\|,$$

where by $V(a)$ we now denote the ordinary translation operator as in previous examples.

Then we also know that the left term vanishes as $a \to 0$ (see Section 3.3.1). Let us now study the right term. We get

$$\left\| \frac{U_\delta(a)\psi - \psi}{a} - \frac{V(a)\psi - \psi}{a} \right\| = \frac{1}{a} \left\| U_\delta(a)\psi - V(a)\psi \right\| = \frac{1}{a} \left\| (e^{i\delta} - 1)\psi 1_{[0,a]} \right\|$$

$$= \frac{|e^{i\delta} - 1|}{a} \left( \int_0^a |\psi(x)|^2 \, dx \right)^{\frac{1}{2}}$$

$$\leq \frac{|e^{i\delta} - 1|}{a} \|\psi'\|_{L^\infty} \left( \int_0^a x^2 \, dx \right)^{\frac{1}{2}} = c \cdot a^{\frac{1}{2}} \xrightarrow{a \to 0} 0,$$
where in the second last step we used the mean value theorem and the fact that \( \psi(0) = 0 \). So the \( L^2 \)-limit exists indeed.

If we on the other hand assume that \( \psi(0) \neq 0 \), then

\[
\| \frac{U_\delta(a) \psi - \psi}{a} \| \geq \| \frac{V(a) \psi - \psi}{a} \| - \frac{\left| e^{i\delta} - 1 \right|}{a} \left( \int_0^a |\psi(x)|^2 \, dx \right) \frac{1}{2}.
\] (3.97)

Since the left term goes to zero, it suffices to show that the right terms diverges. As before, we get

\[
\| \frac{U_\delta(a) \psi - V(a) \psi}{a} \| \geq \frac{\left| e^{i\delta} - 1 \right|}{a} \left( \int_0^a |\psi(x)|^2 \, dx \right) \frac{1}{2} = c \cdot a^{-\frac{1}{2}} \to \infty,
\] (3.98)

which proves the claim.

Let us summarise what we have learned. We have seen that for a large class of functions namely the \( C^1_0 \)-functions, which vanish at the origin the limits \( \lim_{a \to 0} \frac{U_\delta(a) \psi - \psi}{a} = iA_\delta \psi \) and \( \lim_{a \to 0} \frac{V(a) \psi - \psi}{a} = : iB \psi \) exist and are identical, namely the derivative \( \psi' \) of \( \psi \). On the other hand a large class of functions lie outside the domain of \( A_\delta \), namely those in \( C^1_0 \) with non-vanishing value at the origin. We could be led to think that \( B \) is just an extension of \( A_\delta \) but since both \( A_\delta \) and \( B \) are self-adjoint, by Remark 5 that would already mean that \( A_\delta = B \). So in fact \( A_\delta \) and \( B \) are identical on a dense subset of \( L^2 \), but at the same time \( D(A_\delta) \setminus D(B) \) and \( D(B) \setminus D(A_\delta) \) are both non-empty. This demonstrates the fact that a symmetric operator may have more than one self-adjoint extension. In fact, since for every \( \delta \in [0, 2\pi) \) we get a distinct operator \( A_\delta \), there is an uncountable family of self-adjoint operators that are extensions of the operator which is defined on the \( C^1_0 \)-functions vanishing at the origin. Moreover \( B = A_0 \) by definition.

The situation is shown in Figure 3.1.

Finally, let us determine exactly what the operator \( A_\delta \) is, i.e. what its domain is and how it acts on a function in that domain. More precisely, we will try to find an essentially self-adjoint operator \( \tilde{A}_\delta \) whose closure is \( A_\delta \).

The problem at the origin arises from the phase shift in the operator. We solved this by restricting ourselves to functions that vanish at the origin. This is however a rather blunt approach. It is more subtle to compensate for the phase shift by looking at functions having a phase shift at the origin themselves. So, we define the space \( S_\delta(\mathbb{R}) \) to be the space of all Schwartz functions with the distinction that they have a phase shift of \( e^{-i\delta} \) at the origin (and hence have a discontinuity there, except if they are zero). They can be obtained exactly by taking any Schwartz function \( \psi(x) \) and multiplying with...
Figure 3.1: The drawing shows the domains of $A_\delta$ for different values of $\delta$ (black and grey ellipses).

$e^{-i\delta[0,\infty)(x)}$. To see that for those functions the limit exists, let us look at $\psi \in \mathcal{S}(\mathbb{R})$ and $\varphi(x) = \psi(x)e^{-i\delta[0,\infty)(x)}$ (so $\varphi \in \mathcal{S}_\delta(\mathbb{R})$). Then

$$
\frac{U_\delta(a)\varphi - \varphi}{a} = \frac{\varphi(\cdot + a)e^{i\delta[-a,0]} - \varphi}{a} = \frac{\psi(\cdot + a)e^{i\delta[0,\infty)}(\cdot+a)}{a} - \psi(x)e^{i\delta[0,\infty)}
$$

and hence, since $\psi \in \mathcal{S}(\mathbb{R})$, the limit exists in $L^2(\mathbb{R})$ since it is just the normal derivation operator on $\mathcal{S}(\mathbb{R})$ as we used it to define the momentum operator. We now define $\tilde{A}_\delta$ on $\mathcal{S}_\delta(\mathbb{R})$ by $\tilde{A}_\delta := -i \lim_{a \to 0} \frac{U_\delta(a) - \text{Id}}{a}$. As we have seen, $\tilde{A}_\delta$ acts on a function $\varphi \in \mathcal{S}_\delta(\mathbb{R})$ by removing the phase shift, acting like the normal momentum operator and restoring the phase shift. We want to show that $\tilde{A}_\delta$ is essentially self-adjoint on $\mathcal{S}_\delta(\mathbb{R})$. Then since $A_\delta \supset \tilde{A}_\delta$ we obtain $A_\delta$ as the closure of $\tilde{A}_\delta$.

So let us show that $\tilde{A}_\delta$ is essentially self-adjoint on $\mathcal{S}_\delta(\mathbb{R})$. We could do so by making a direct calculation as we did for the translation operator and the rotation operator (Lemma 5 and Proposition 10) or we can make use of the following useful theorem.

**Theorem 16.** Let $\{U(t)\}_{t \in \mathbb{R}}$ be a strongly continuous one-parameter unitary group and let $D$ be a dense subspace of $\mathcal{H}$ invariant under $U(t)$, i.e. $U(t)D \subset D$ for all $t \in \mathbb{R}$. Further, let $U(t)$ be strongly differentiable on $D$, i.e. the limit $\lim_{t \to 0} \frac{U(t)\psi - \psi}{t}$ exists for all $\psi \in D$. Then the operator defined on the domain $D$ by $-i$ times the strong derivative of $U(t)$ is essentially self-adjoint and its closure is the infinitesimal generator of $U(t)$. In other words: If $U(t) = e^{iAt}$, then $A|_D$ is essentially self-adjoint and $\overline{A|_D} = A$. 

51
Proof. The proof is analogous to the proof of essential self-adjointness in the proof of Stone’s theorem \cite{7}. We define $\tilde{A}$ on $D$ by $\tilde{A}\psi := -i\lim_{t \to 0} \frac{U(t)\psi - \psi}{t}$. Then $\tilde{A}$ is symmetric because if $\varphi, \psi \in D$, we have

$$\langle \tilde{A}\varphi, \psi \rangle = \lim_{t \to 0} \left( \frac{U(t)\varphi - \varphi}{it}, \psi \right) = \lim_{t \to 0} \left( \varphi, \frac{\psi - U(-t)\psi}{it} \right) = \langle \varphi, \tilde{A}\psi \rangle.$$

(3.101)

We want to show the essential self-adjointness using Corollary \ref{corollary}. Suppose there is a $\psi \in D(\tilde{A}^*)$ so that $\tilde{A}^*\psi = i\psi$. Then for each $\varphi \in D(\tilde{A}) = D$ we have

$$\frac{d}{dt} \langle U(t)\varphi, \psi \rangle = \lim_{s \to 0} \left( \frac{U(t+s) - U(t)}{s} \varphi, \psi \right) = \lim_{s \to 0} \left( \frac{U(s) - \Id}{s} \frac{\in D}{U(t)\varphi, \psi} \right)$$

(3.102)

So the complex-valued function $f(t) = \langle U(t)\varphi, \psi \rangle$ satisfies the ordinary differential equation $f' = f$ demanding an exponential solution $f(t) = f(0)e^{t}$. On the other hand $U(t)$ is unitary and thus has norm 1. So $f(t)$ has to be bounded for positive and negative $t$, which is only possible if $f(0) = \langle \varphi, \psi \rangle = 0$. Since $D$ was dense in $\mathcal{H}$ and $\varphi$ was chosen arbitrarily, we conclude that $\psi = 0$. Similarly, we conclude that the equation $\tilde{A}^*\psi = -i\psi$ has no non-zero solutions. Then by Corollary \ref{corollary} we know that $\tilde{A}$ is essentially self-adjoint on $D$. Moreover since $\tilde{A}$ is obviously an extension of $\tilde{A}$, we get that $\tilde{A} = A$. \qed

Now we can check essential self-adjointness in our example.

**Proposition 14.** The operator $\tilde{A}_\delta$ as defined above on its domain $\mathcal{S}_\delta(\mathbb{R})$ is essentially self-adjoint.

**Proof.** We want to use Theorem \ref{theorem} which we just stated. A simple calculation shows that $U_\delta(a)\mathcal{S}_\delta(\mathbb{R}) \subseteq \mathcal{S}_\delta(\mathbb{R})$. We just showed that $U_\delta(a)$ is strongly differentiable on $\mathcal{S}_\delta(\mathbb{R})$ (with $i\tilde{A}_\delta$ being the derivative) and $\mathcal{S}_\delta(\mathbb{R})$ is certainly dense in $L^2(\mathbb{R})$. Hence $\tilde{A}_\delta$ is essentially self-adjoint on $\mathcal{S}_\delta(\mathbb{R})$. \qed

Theorem \ref{theorem} also gives directly that the closure of $\tilde{A}_\delta$ is the infinitesimal generator of $U_\delta(a)$.

**Theorem 17.** The generator of the modified translation is exactly the closure of $\tilde{A}_\delta$ defined on $\mathcal{S}_\delta(\mathbb{R}) = \left\{ \psi e^{-i\delta \xi [0,\infty]} \big| \psi \in \mathcal{S}(\mathbb{R}) \right\}$.

### 3.4 Examples II: Time Evolution

So far we have started from a strongly continuous unitary group and have determined the infinitesimal generator of it. Let us now turn the tables and approach the problem from the other side, i.e., we start with a self-adjoint operator and want to construct the strongly continuous unitary group generated by it. The motivation for this lies in the time dependence of the wave function in quantum mechanics.

Before we investigate the time-evolution, let us state the following proposition, which gives an analogous result to Lemma \ref{lemma}. 

52
**Proposition 15.** The operator $\tilde{A} := \frac{d^2}{dx^2}$ defined on $\mathcal{S}(\mathbb{R})$ is essentially self-adjoint.

**Proof.** The procedure is the same as in the proof of Lemma 2, just a little bit more complicated. $\tilde{A}$ is certainly symmetric. We want to show essential self-adjointness by using the basic criterion (Corollary 1), i.e. by showing that $\text{Ran}(\tilde{A} \pm i)$ is dense in $L^2(\mathbb{R})$. For this, let $\varphi(x)$ be an arbitrary $C^0_\infty(\mathbb{R})$ function. We are looking for a function $\psi(x) \in \mathcal{S}(\mathbb{R})$ such that

$$\psi''(x) \pm i\psi(x) = \varphi(x). \quad (3.103)$$

Let us look at the case with the lower sign in $\pm$. The other case can be treated analogously. Using integrating factors one obtains

$$\varphi(x) = e^{ax} \left( e^{-2ax} (e^{ax} \psi(x))' \right)', \quad (3.104)$$

where $a = \frac{1+i}{\sqrt{2}}$. Solving for $\psi$ gives

$$\psi(x) = e^{-ax} \int_{c_1}^{x} e^{2ay} \int_{c_2}^{y} e^{-az} \varphi(z) dz dy \quad (3.105)$$

The term $e^{-az} \varphi(z)$ in the inner integral is just another test function and we can choose the integration constant $c_2$ such that the integral is constant to the left of a compact set $K$ and zero to the right of it. Multiplying by $e^{2ay}$ (note that $\Re(a) > 0$) gives a function, which is exponentially decreasing to the left of $K$ and zero to the right of it. Integrating gives a function, which is still exponentially decreasing on the left and constant on the right. Finally, multiplying by $e^{-ax}$ gives a function, which is exponentially decreasing on the left and on the right and hence a Schwartz function. \hfill $\square$

**3.4.1 Free Particle**

Recall, it is a postulate of quantum mechanics that the time evolution operator $U(t)$, which operates by

$$U(s)\psi_t = \psi_{s+t}, \quad (3.106)$$

i.e. it maps the wave function at time $t$ to the wave function at time $t + s$, is given by

$$U(t) = e^{-itH}, \quad (3.107)$$

where $H$ is the self-adjoint Hamilton operator representing the energy of the system. By Theorem 6 we then know that $\{U(t)\}_{t \in \mathbb{R}}$ forms a strongly continuous one-parameter unitary group. Depending on the described system, this group can take many forms.

The easiest example would be the case of a free particle. Free in this context means free of constraints, i.e. there are no potentials present. Then the energy is just given by the kinetic energy

$$E = \frac{p^2}{2m}. \quad (3.108)$$
where $m$ is the particle’s mass and $p$ its (here one-dimensional) momentum. Using the correspondence rules, we can write down the Hamilton operator. It must be a self-adjoint version of the second derivative. Again, we first define it for Schwartz functions:

$$\tilde{H} = \frac{1}{2m} \left( -i \frac{d}{dx} \right)^2 = -\frac{1}{2m} \frac{d^2}{dx^2},$$

(3.109)

In Proposition 15 we showed that $\tilde{H}$ is essentially self-adjoint on the domain $\mathcal{S}(\mathbb{R})$. Hence we define $H$ to be the self-adjoint closure of $\tilde{H}$.

There is however a more elegant and explicit way of writing $H$. It involves the Fourier transform on $L^2(\mathbb{R})$. It is a well-known fact that the Fourier transform $\mathcal{F}$ can be defined in a way that is is a unitary map from $L^2(\mathbb{R})$ into itself.\cite{Lax02}

**Definition 24.** For $\psi \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ we define the **unitary Fourier transform** by

$$\hat{\psi}(\xi) := (\mathcal{F}(\psi))(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix\xi} \psi(x)dx,$$

(3.110)

for other $L^2$-functions it is defined by density.

**Remark 9.** This unitary map $\mathcal{F} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ has some important properties. One is, that for a differentiable function whose derivative is again in $L^2(\mathbb{R})$, it holds:

$$\left( \mathcal{F} \left( \frac{d}{dx} \psi \right) \right)(\xi) = i\xi (\mathcal{F}(\psi)).$$

(3.111)

Analogously, if the second derivative is in $L^2(\mathbb{R})$,

$$\left( \mathcal{F} \left( \frac{d^2}{dx^2} \psi \right) \right)(\xi) = -\xi^2 (\mathcal{F}(\psi)).$$

(3.112)

We want to use the above observation to define some self-adjoint version of $\frac{d^2}{dx^2}$ via the Fourier transform.

**Definition 25.** We define the **Laplace operator** $\Delta$ by

$$\Delta := \mathcal{F}^{-1} T_{-x^2} \mathcal{F}$$

(3.113)

with the maximal possible domain

$$D(\Delta) = \{ \psi \in L^2(\mathbb{R}) \mid T_{-x^2} \mathcal{F} \psi \in L^2(\mathbb{R}) \}.$$  

(3.114)

By (3.112), the Laplace operator $\Delta$ applied to an $L^2$-function whose second derivative is again in $L^2$ is indeed exactly the second derivative.

**Proposition 16.** The Laplace operator $\Delta = \mathcal{F}^{-1} T_{-x^2} \mathcal{F}$ is self-adjoint.
Proof. First, we note that $\Delta$ is symmetric since for $\psi, \varphi \in D(\Delta)$ it holds:

$$
\langle \Delta \psi, \varphi \rangle = \langle F^{-1} T_{-x^2} F \psi, \varphi \rangle = \langle T_{-x^2} F \psi, F \varphi \rangle = \langle F \psi, T_{-x^2} F \varphi \rangle = \langle \psi, \Delta \varphi \rangle.
$$

(3.115)

Knowing that $\Delta$ is symmetric, we want to use the basic criterion (Theorem 3) to prove that $\Delta$ is also self-adjoint by showing that $\text{Ran}(\Delta \pm i) = L^2(\mathbb{R})$. We know

$$
\Delta \pm i = F^{-1} T_{-x^2} F \pm i = F^{-1} (T_{-x^2} \pm i) F = F^{-1} T_{-x^2 \pm i} F
$$

(3.116)

and since $F$ is a bijection from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$, it is enough to show that

$$
\text{Ran}(T_{-x^2 \pm i}) = L^2(\mathbb{R}).
$$

(3.117)

So let $\varphi \in L^2(\mathbb{R})$. Then define

$$
\psi(x) := \frac{\varphi(x)}{-x^2 \pm i},
$$

(3.118)

which is in $L^2(\mathbb{R})$ since $\left| \frac{1}{-x^2 \pm i} \right| \leq 1$. Then $T_{-x^2 \pm i} \psi = \varphi$, which proves that the range is all of $L^2(\mathbb{R})$ and hence $\Delta$ is self-adjoint.

Proposition 17. The Hamilton operator $H$ defined as closure of $\tilde{H} = -\frac{1}{2m} \frac{d^2}{dx^2}$ on $\mathcal{S}(\mathbb{R})$ is exactly a multiple of the Laplace operator, namely $H = -\frac{1}{2m} \Delta$.

Proof. To see this, recall that the domain of $\tilde{H}$ are the Schwartz functions. Applying $\Delta = F^{-1} T_{-x^2} F$ to a function $\psi \in \mathcal{S}$ just gives the second derivative. So $\tilde{H} \subseteq -\frac{1}{2m} \Delta$. Taking the closure, on gets $H \subseteq -\frac{1}{2m} \Delta$ and since both sides are self-adjoint operators, the assertion follows by Remark 5.

We have therefore gained a very useful identity, which will help when calculating the unitary group generated by $H$. This will give us the time evolution of a free particle.

Theorem 18. The time evolution of a single free non-relativistic particle is given by

$$
e^{-i t H} = F^{-1} T_g F
$$

(3.119)

with

$$
g(\xi) = e^{-i \frac{\xi^2}{2m}}.
$$

(3.120)

Proof. We have just learned how to “diagonalise” $H$, namely

$$
H = F^{-1} T_{-x^2} F.
$$

(3.121)

Hence, the assertion follows directly from the functional calculus.

Example 7. Let us study how the time evolution operator $e^{-itH}$ acts on a particular function. So let us assume that we start at $t = 0$ say with a normalised Gaussian peak of unit width

$$
\psi(x) = \frac{1}{\sqrt{\pi}} e^{-\frac{1}{2} x^2}.
$$

(3.122)
The Fourier transform is given by

\[ \hat{\psi}(\xi) = \frac{1}{\sqrt{\pi}} e^{-\frac{1}{2} \xi^2} \]  \hspace{1cm} (3.123)

and multiplying with \( e^{-i \frac{t}{2m} \xi^2} \) gives

\[ \frac{1}{\sqrt{\pi}} e^{-\frac{1+it}{2} \xi^2}. \]  \hspace{1cm} (3.124)

Finally, we take the inverse Fourier transform and get

\[ \psi_t(x) := (e^{-itH} \psi)(x) = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{1 + i \frac{t}{m}}} e^{-\frac{1}{2} \frac{x^2}{1 + i \frac{t}{m}}}, \]  \hspace{1cm} (3.125)

After a polar decomposition this reads

\[ \psi_t(x) = e^{i\phi(x,t)} \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{1 + \frac{x^2}{m^2}}} e^{-\frac{1}{2} \frac{x^2}{1 + \frac{x^2}{m^2}}}, \]  \hspace{1cm} (3.126)

where \( \phi \) is some real-valued function. Hence, we see that the initial peak becomes wider and flatter with increasing time \( t \). The physical interpretation is that we begin with a particle localised at \( x = 0 \) but over time the probability of measuring the particle further away from \( x = 0 \) increases (see Figure 3.2).

![Figure 3.2: Plot of \( |\psi_t(x)| \) for different values of \( t \).](image)

### 3.4.2 Free, Moving Particle

Let us further investigate the time evolution by making some variations. In the above example the initial peak became wider, but remained centered at the origin. This can
be explained by looking at the Fourier transform of the initial peak, which was given by

\[ \psi(\xi) = \frac{1}{\sqrt{\pi}} e^{-\frac{1}{2}\xi^2}. \]  

(3.127)

In physical terms the Fourier transform is a transform from the real to the momentum space. A Fourier transform symmetric around the origin means that there is no net momentum to the left or to the right and hence the centroid of the distribution remains at \( x = 0 \).

**Example 8.** Let us now consider the initial state

\[ \psi(x) = \frac{1}{\sqrt{\pi}} e^{-\frac{1}{2}x^2 + aix}, \]  

(3.128)

then the Fourier transform is given by

\[ \frac{1}{\sqrt{\pi}} e^{-\frac{1}{2}(a-\xi)^2}. \]  

(3.129)

This corresponds to a momentum distribution centred around \( a \). So it is reasonable to assume that the time evolved function will be a peak which becomes wider, but at the same time moves with speed \( \frac{a}{m} \) to the right. Multiplying with \( e^{-i\frac{1}{2m}\xi^2} \) and making the inverse Fourier transform, one obtains indeed

\[ |\psi_t(x)| = \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{1 + \frac{t^2}{m^2}}} e^{-\frac{1}{2}(x-\frac{at}{m})^2}, \]  

(3.130)

which is exactly the expected result (see Figure 3.3).

![Figure 3.3: Plot of $|\psi_t(x)|$ for different values of $t$.](image-url)
3.4.3 General Wave Function

In the above two examples we have seen how the time evolution acts on Gaussian peaks. It is intuitively clear that any given wave function can be approximated by a sum of Gaussian peaks of different amplitudes. This becomes exact in the limit where the sum becomes an integral. For the sake of completeness the according result shall be stated here, but without proof since this would lead to far away from the actual topic of this text.

**Theorem 19.** Let \( \psi_0 \in C_0^\infty \). Then

\[
\psi_t(x) = \frac{1}{\sqrt{2\pi i(1 + i)}} \int_{\mathbb{R}} \psi_0(y) \exp\left(\frac{i(x - y)^2}{4t}\right) dy \quad (3.131)
\]

for \( t \neq 0 \).

**Proof.** For a proof, see for example [Rau91]. \( \square \)

3.4.4 Constant Potential

So far, we have only looked at the case of a free particle. The next step would be to consider a potential \( V(x) \). The energy is given by

\[
E = T + V(x) = \frac{p^2}{2m} + V(x). \quad (3.132)
\]

The Hamilton-operator is then formally given by

\[
H = -\frac{1}{2m} \Delta + T_V(x). \quad (3.133)
\]

In general it is far from trivial to determine a domain of self-adjointness for this operator or to explicitly diagonalise it as we did with the Fourier transform for the pure \( \Delta \). Let us for the sake of argument consider the trivial potential, i.e. a constant potential. It is a well-known fact that we can add or subtract any constant to a potential without changing any calculation. This is because the choice of zero potential energy or zero height is completely arbitrary. So let us assume that \( V(x) = c \). Then

\[
\mathcal{F}(\frac{1}{2m} \Delta + T_c)\mathcal{F}^{-1} = \frac{1}{2m} \mathcal{F}(-\Delta)\mathcal{F}^{-1} + T_c = T_{\frac{1}{2m} x^2} + T_c = T_{\frac{1}{2m} x^2 + c}, \quad (3.134)
\]

where we used the linearity of the Fourier transform (i.e. \( \mathcal{F} T_c = T_c \mathcal{F} \)). The time evolution operator is then given by

\[
U(t) = \mathcal{F} T_g \mathcal{F}^{-1}, \quad (3.135)
\]

where

\[
g(\xi) = e^{-it(\frac{1}{2m} \xi^2 + c)} = e^{-it\frac{1}{2m} \xi^2} e^{-itc}. \quad (3.136)
\]

We can write this as

\[
(U(t)\psi)(x) = e^{-itc} \psi_t(x), \quad (3.137)
\]

where \( \psi_t \) denotes the time evolution of the free particle. So, we see that the only thing that changes by adding a constant potential is a global (time-dependent) phase factor, which has no physical relevance.
4 Conclusions

“A great deal more is known than has been proved.” —Richard P. Feynman

We built up the theory of unbounded operators on Hilbert spaces from scratch with a focus on important classes of operators like closed, symmetric, or self-adjoint ones. We learned how these properties are related and what they are good for. The goal of developing a functional calculus in mind, we had a closer look at self-adjoint operators and their spectral properties. The functional calculus in turn allowed us to give meaning to the exponential of an operator.

We then looked at strongly continuous groups of unitary operators. Using the functional calculus we were able to parametrise them in terms of self-adjoint operators, which made manipulating them and calculating with them easier. The corresponding result is Stone’s theorem, which is the main result of this text.

We realised that these unitary groups play an important role in quantum mechanics if their generator can be related to a physical observable. From this we deduced many remarkable features. In addition, we looked at an elegant proof of Bochner’s theorem.

The main conclusion one should draw is that putting the mathematical description of quantum mechanics on a firm theoretical setting is all but an easy task. We had to make use of sophisticated concepts and rely on high level results such as the spectral theorem for unbounded self-adjoint operators. The reader should also keep in mind that the level of difficulty when dealing with those unitary groups can increase tremendously, for example when looking at the time evolution of a particle in a more complicated potential.

As an extension of the strongly continuous unitary groups we discussed in this text, one could look at strongly continuous groups or strongly continuous semigroups (often abbreviated $C_0$-semigroups). For both, one can define an infinitesimal generator and especially the semigroups play an important role in modern mathematics. The corresponding calculations are more difficult than those in this text since one lacks some powerful tools such as the functional calculus for self-adjoint operators.
Bibliography


