Unbounded operators in a Hilbert space and the Trotter product formula

Jordan Bell
jordan.bell@gmail.com
Department of Mathematics, University of Toronto

August 25, 2015

1 Unbounded operators

Let H be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$. We do not assume that H is separable. By an **operator in** H we mean a linear subspace $\mathscr{D}(T)$ of H and a linear map $T : \mathscr{D}(T) \to H$. We define

$$\mathscr{R}(T) = \{Tx : x \in \mathscr{D}(T)\}.$$

If $\mathscr{D}(T)$ is dense in H we say that T **is densely defined**.

Write

$$\mathscr{G}(T) = \{(x, y) \in H \times H : x \in \mathscr{D}(T), y = Tx\}.$$

When $\mathscr{G}(T) \subset \mathscr{G}(S)$, we write

$$T \subset S,$$

and say that S **is an extension of** T. If $\mathscr{G}(T)$ is a closed linear subspace of $H \times H$, we say that T **is closed**.

We say that an operator T in H is **closable** if there is a closed operator S in H such that $T \subset S$. If T is closable, one proves that there is a unique closed operator $\overline{T}$ in H with $T \subset \overline{T}$ and such that if S is a closed operator satisfying $T \subset S$ then $\overline{T} \subset S$.

Suppose that T is a densely defined operator in H. We define $\mathscr{D}(T^*)$ to be the set of those $y \in H$ for which

$$x \mapsto \langle Tx, y \rangle, \qquad x \in \mathscr{D}(T),$$

is continuous. For $y \in \mathscr{D}(T^*)$, by the Hahn-Banach theorem there is some $\lambda_y \in H^*$ such that

$$\lambda_y x = \langle Tx, y \rangle, \qquad x \in \mathscr{D}(T).$$

Next, by the Riesz representation theorem, there is a unique $x_y \in H$ such that

$$\lambda_y x = \langle x, x_y \rangle, \qquad x \in H,$$

and hence

$$\langle x, x_y \rangle = \langle Tx, y \rangle, \qquad x \in \mathscr{D}(T).$$

If $v \in H$ satisfies

$$\langle x, v \rangle = \langle Tx, y \rangle, \qquad x \in \mathscr{D}(T),$$

then

$$\langle x, v \rangle = \langle x, x_y \rangle, \qquad x \in \mathscr{D}(T),$$

and because $\mathscr{D}(T)$ is dense in H this implies that $v = x_y$. We define $T^* : \mathscr{D}(T^*) \to H$ by $T^* y = x_y$, which satisfies

$$\langle Tx, y \rangle = \langle x, T^* y \rangle, \qquad x \in \mathscr{D}(T).$$

T^* is called **the adjoint of** T. One checks that $\mathscr{D}(T^*)$ is a linear subspace of H and that $T^* : \mathscr{D}(T^*) \to H$ is a linear map. We say that T **is self-adjoint** when $T = T^*$.

For operators S and T in H we define

$$\mathscr{D}(S + T) = \mathscr{D}(S) \cap \mathscr{D}(T)$$

and

$$\mathscr{D}(ST) = \{x \in \mathscr{D}(T) : Tx \in \mathscr{D}(S)\}.$$

One checks that

$$(R + S) + T = R + (S + T), \qquad (RS)T = R(ST),$$

and

$$RT + ST = (R + S)T, \qquad TR + TS \subset T(R + S).$$

We now determine the adjoint of products of densely defined operators.[1]

Theorem 1. *If S, T, and ST are densely defined operators in H, then*

$$T^* S^* \subset (ST)^*.$$

If $S \in \mathscr{B}(H)$, then

$$T^* S^* = (ST)^*.$$

Proof. Let $y \in \mathscr{D}(T^* S^*)$ and let $x \in \mathscr{D}(ST)$. Then $S^* y \in \mathscr{D}(T^*)$ and $x \in \mathscr{D}(T)$, so

$$\langle Tx, S^* y \rangle = \langle x, T^* S^* y \rangle.$$

On the other hand, $y \in \mathscr{D}(S^*)$, so

$$\langle STx, y \rangle = \langle Tx, S^* y \rangle.$$

Hence

$$\langle STx, y \rangle = \langle x, T^* S^* y \rangle,$$

[1] Walter Rudin, *Functional Analysis*, second ed., p. 348, Theorem 13.2.

which implies that $(ST)^*y = T^*S^*y$ for each $y \in \mathscr{D}(T^*S^*)$, that is, $T^*S^* \subset (ST)^*$.

Suppose that $S \in \mathscr{B}(H)$, hence $S^* \in \mathscr{B}(H)$, for which $\mathscr{D}(S^*) = H$. Let $y \in \mathscr{D}((ST)^*)$. For $x \in \mathscr{D}(ST)$,

$$\langle Tx, S^*y \rangle = \langle STx, y \rangle = \langle x, (ST)^*y \rangle .$$

This implies that $S^*y \in \mathscr{D}(T^*)$ and hence $y \in \mathscr{D}(T^*S^*)$, showing

$$\mathscr{D}((ST)^*) \subset \mathscr{D}(T^*S^*).$$

□

If T is an operator in H, we say that T **is symmetric** if

$$\langle Tx, y \rangle = \langle x, Ty \rangle , \qquad x, y \in \mathscr{D}(T).$$

Theorem 2. *Let T be a densely defined operator in H. T is symmetric if and only if $T \subset T^*$.*

Proof. Suppose that T is symmetric and let $(y, Ty) \in \mathscr{G}(T)$. For $x \in \mathscr{D}(T)$,

$$| \langle Tx, y \rangle | = | \langle x, Ty \rangle | \leq \|x\| \, \|Ty\| ,$$

hence $x \mapsto \langle Tx, y \rangle$ is continuous on $\mathscr{D}(T)$, i.e. $y \in \mathscr{D}(T^*)$. For $x \in \mathscr{D}(T)$, on the one hand,

$$\langle Tx, y \rangle = \langle x, T^*y \rangle ,$$

and on the other hand,

$$\langle Tx, y \rangle = \langle x, Ty \rangle .$$

Therefore $\langle x, T^*y \rangle = \langle x, Ty \rangle$ for all $x \in \mathscr{D}(T)$, and because $\mathscr{D}(T)$ is dense in H we get that $T^*y = Ty$, i.e. $(y, Ty) \in \mathscr{G}(T^*)$. Therefore $\mathscr{G}(T) \subset \mathscr{G}(T^*)$.

Suppose that $\mathscr{G}(T) \subset \mathscr{G}(T^*)$. Let $x, y \in \mathscr{D}(T)$. We have $(y, Ty) \in \mathscr{G}(T^*)$, i.e. $y \in \mathscr{D}(T^*)$ and $T^*y = Ty$. Hence

$$\langle Tx, y \rangle = \langle x, T^*y \rangle = \langle x, Ty \rangle ,$$

showing that T is symmetric. □

One proves that if T is a symmetric operator in H then T is closable and $\overline{T}$ is symmetric. An operator T in H is said to be **essentially self-adjoint** when T is densely defined, symmetric, and $\overline{T}$ (which is densely defined) is self-adjoint.

2 Graphs

For $(a, b), (c, d) \in H \times H$, we define

$$\langle (a, b), (c, d) \rangle = \langle a, c \rangle + \langle b, d \rangle .$$

This is an inner product on $H \times H$ with which $H \times H$ is a Hilbert space. We define $V : H \times H \to H \times H$ by

$$V(a, b) = (-b, a), \qquad (a, b) \in H \times H,$$

which belongs to $\mathscr{B}(H \times H)$. It is immediate that $VV^* = I$ and $V^*V = I$, namely, V **is unitary**. As well, $V^2 = -I$, whence if M is a linear subspace of $H \times H$ then $V^2 M = M$. The following theorem relates the graphs of a densely defined operator and its adjoint.[2]

Theorem 3. *Suppose that T is a densely defined operator in H. It holds that*

$$\mathscr{G}(T^*) = (V\mathscr{G}(T))^{\perp}.$$

Theorem 4. *If T is a densely defined operator in H, then T^* is a closed operator.*

Proof. $V\mathscr{G}(T)$ is a linear subspace of $H \times H$. The orthogonal complement of a linear subspace of a Hilbert space is a closed linear subspace of the Hilbert space, and thus Theorem 3 tells us that $\mathscr{G}(T^*)$ is a closed linear subspace of $H \times H$, namely, T^* is a closed operator. □

Let T be a densely defined operator in H. If T is self-adjoint, then the above theorem tells us that T is itself a closed operator.

Theorem 5. *Suppose that T is a closed densely defined operator in H. Then*

$$H \times H = V\mathscr{G}(T) \oplus \mathscr{G}(T^*)$$

is an orthogonal direct sum.

Proof. Generally, if M is a linear subspace of $H \times H$,

$$H \times H = \overline{M} \oplus M^{\perp} = \overline{M} \oplus (\overline{M})^{\perp}$$

is an orthogonal direct sum. For $M = V\mathscr{G}(T)$, because $\mathscr{G}(T)$ is a closed linear subspace of $H \times H$, so is M. Thus

$$H \times H = V\mathscr{G}(T) \oplus (V\mathscr{G}(T))^{\perp}.$$

By Theorem 3, this is

$$H \times H = V\mathscr{G}(T) \oplus \mathscr{G}(T^*),$$

proving the claim. □

[2] Walter Rudin, *Functional Analysis*, second ed., p. 352, Theorem 13.8.

If T is an operator in H that is one-to-one, we define $\mathscr{D}(T^{-1}) = \mathscr{R}(T)$, and T^{-1} is a densely defined operator with domain $\mathscr{D}(T^{-1})$.

The following theorem establishes several properties of symmetric densely defined operators.[3] We remind ourselves that if T is an operator in H, the statement $\mathscr{D}(T) = H$ means that T is a linear map $H \to H$, from which it does not follow that T is continuous.

Theorem 6. *Suppose that T is a densely defined symmetric operator in H. Then the following statements are true:*

1. *If $\mathscr{D}(T) = H$ then T is self-adjoint and $T \in \mathscr{B}(H)$.*
2. *If T is self-adjoint and one-to-one, then $\mathscr{R}(T)$ is dense in H and T^{-1} is densely defined and self-adjoint.*
3. *If $\mathscr{R}(T)$ is dense in H, then T is one-to-one.*
4. *If $\mathscr{R}(T) = H$, then T is self-adjoint and $T^{-1} \in \mathscr{B}(H)$.*

If $T \in \mathscr{B}(H)$ then $T^{**} = T$. The following theorem says that this is true for closed densely defined operators.[4]

Theorem 7. *If T is a closed densely defined operator in H, then $\mathscr{D}(T^*)$ is dense in H and $T^{**} = T$.*

The following theorem gives statements about $I + T^*T$ when T is a closed densely defined operator.[5]

Theorem 8. *Suppose that T is a closed densely defined operator in H and let $Q = I + T^*T$, with*

$$\mathscr{D}(Q) = \mathscr{D}(T^*T) = \{x \in \mathscr{D}(T) : Tx \in \mathscr{D}(T^*)\}.$$

The following statements are true:

1. *$Q : \mathscr{D}(Q) \to H$ is a bijection, and there are $B, C \in \mathscr{B}(H)$ with $\|B\| \leq 1$, $B \geq 0$, $\|C\| \leq 1$, $C = TB$, and*

   $$B(I + T^*T) \subset (I + T^*T)B = I.$$

   T^*T is self-adjoint.

2. *Let T_0 be the restriction of T to $\mathscr{D}(T^*T)$. Then $\mathscr{G}(T_0)$ is dense in $\mathscr{G}(T)$.*

Let T be a symmetric operator in H. We say that T is **maximally symmetric** if $T \subset S$ and S being symmetric imply that $S = T$. One proves that a self-adjoint operator is maximally symmetric.[6]

The following theorem is about $T + iI$ when T is a symmetric operator in H.[7]

[3] Walter Rudin, *Functional Analysis*, second ed., p. 353, Theorem 13.11.
[4] Walter Rudin, *Functional Analysis*, second ed., p. 354, Theorem 13.12.
[5] Walter Rudin, *Functional Analysis*, second ed., p. 354, Theorem 13.13.
[6] Walter Rudin, *Functional Analysis*, second ed., p. 356, Theorem 13.15.
[7] Walter Rudin, *Functional Analysis*, second ed., p. 356, Theorem 13.16.

Theorem 9. *Suppose that T is a symmetric operator in H and let j be i or $-i$. Then:*

1. $\|Tx + jx\|^2 = \|x\|^2 + \|Tx\|^2$ *for* $x \in \mathscr{D}(T)$.
2. *T is closed if and only if $\mathscr{R}(T + jI)$ is a closed subset of H.*
3. *$T + jI$ is one-to-one.*
4. *If $\mathscr{R}(T + jI) = H$ then T is maximally symmetric.*

3 The Cayley transform

Let T be a symmetric operator in H and define

$$\mathscr{D}(U) = \mathscr{R}(T + iI).$$

Theorem 9 tells us that $T + iI$ is one-to-one. Because

$$\mathscr{D}(T - iI) = \mathscr{D}(T) = \mathscr{D}(T - iI)$$

and $\mathscr{D}((T + iI)^{-1}) = \mathscr{R}(T + iI)$,

$$\begin{aligned}\mathscr{D}((T - iI)(T + iI)^{-1}) &= \{x \in \mathscr{R}(T + iI) : (T + iI)^{-1}x \in \mathscr{D}(T)\} \\ &= \{x \in \mathscr{R}(T + iI) : (T + iI)^{-1}x \in \mathscr{D}(T + iI)\} \\ &= \mathscr{R}(T + iI) \\ &= \mathscr{D}(U).\end{aligned}$$

We define

$$U = (T - iI)(T + iI)^{-1}.$$

U is called the **Cayley transform of** T.

We have

$$\mathscr{R}(U) = U\mathscr{D}(U) = U\mathscr{R}(T+iI) = (T-iI)(T+iI)^{-1}\mathscr{R}(T+iI) = (T-iI)\mathscr{D}(T+iI),$$

and $\mathscr{D}(T + iI) = \mathscr{D}(T) = \mathscr{D}(T - iI)$ so

$$\mathscr{R}(U) = (T - iI)\mathscr{D}(T - iI) = \mathscr{R}(T - iI).$$

Also, for $x \in \mathscr{D}(T)$, Theorem 9 tells us

$$\|(T + iI)x\|^2 = \|Tx + ix\|^2 = \|x\|^2 + \|Tx\|^2 = \|Tx - ix\|^2 = \|(T - iI)x\|^2,$$

hence for $x \in \mathscr{D}(U)$, for which $(T + iI)^{-1}x \in \mathscr{D}(T + iI) = \mathscr{D}(T)$,

$$\|Ux\| = \left\|(T - iI)(T + iI)^{-1}x\right\| = \left\|(T + iI)(T + iI)^{-1}x\right\| = \|x\|,$$

showing that U is an **isometry in** H.

The Cayley transform of a symmetric operator in H (which we do not presume to be densely defined) has the following properties.[8]

[8] Walter Rudin, *Functional Analysis*, second ed., p. 385, Theorem 13.19.

Theorem 10. *Suppose that T is a symmetric operator in H. Then:*

1. *U is closed if and only if T is closed.*
2. *$\mathscr{R}(I-U) = \mathscr{D}(T)$, $I-U$ is one-to-one, and*

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

3. *U is unitary if and only if T is self-adjoint.*

If V is an operator in H that is an isometry and $I-V$ is one-to-one, then there is a symmetric operator S in H such that V is the Cayley transform of S.

4 Resolvents

Let T be an operator in H. The **resolvent set of** T, denoted $\rho(T)$, is the set of those $\lambda \in \mathbb{C}$ such that $T-\lambda I : \mathscr{D}(T) \to H$ is a bijection and $(T-\lambda I)^{-1} \in \mathscr{B}(H)$. That is, $\lambda \in \rho(T)$ if and only if there is some $S \in \mathscr{B}(H)$ such that

$$S(T-\lambda I) \subset (T-\lambda I)S = I.$$

We call $R : \rho(T) \to \mathscr{B}(H)$ defined by

$$R(\lambda) = (T-\lambda I)^{-1}$$

the **resolvent of** T. The **spectrum of** T is $\sigma(T) = \mathbb{C} \setminus \rho(T)$. It is a fact that $\rho(T)$ is open, that $\sigma(T)$ is closed, and that if $\sigma(T) \neq \mathbb{C}$ then T is a closed operator, that

$$R(z) - R(w) = (z-w)R(z)R(w), \qquad z, w \in \rho(T),$$

and

$$\frac{d^n R}{dz^n}(z) = n! R^{n+1}(z), \qquad z \in \rho(T).$$

If T is a self-adjoint operator in H, one proves that $\sigma(T) \subset \mathbb{R}$.

5 Resolutions of the identity

Let $(\Omega, \mathscr{S})$ be a measurable space. A **resolution of the identity** is a function

$$E : \mathscr{S} \to \mathscr{B}(H)$$

satisfying:

1. $E(\emptyset) = 0$, $E(\Omega) = I$.
2. For each $a \in \mathscr{S}$, $E(a)$ is a self-adjoint projection.
3. $E(a \cap b) = E(a)E(b)$.

4. If $a \cap b = \emptyset$, then $E(a \cup b) = E(a) + E(b)$.
5. For each $x, y \in H$, the function $E_{x,y} : \mathscr{S} \to \mathbb{C}$ defined by

$$E_{x,y}(a) = \langle E(a)x, y \rangle, \qquad a \in \mathscr{S},$$

is a complex measure on $\mathscr{S}$.

We check that if $a_n \in \mathscr{S}$ and $E(a_n) = 0$ for each $n = 1, 2, \ldots$, then for $a = \bigcup_{n=1}^{\infty} a_n$, $E(a) = 0$.

Let $\{D_i\}$ be a countable collection of open discs that is a base for the topology of $\mathbb{C}$, i.e., $\bigcup D_i = \mathbb{C}$ and for each i, j and for $z \in D_i \cap D_j$, there is some k such that $x \in D_k \subset D_i \cap D_j$. Let $f : (\Omega, \mathscr{S}) \to (\mathbb{C}, \mathscr{B}_{\mathbb{C}})$ be a measurable function and let V be the union of those D_i for which $E(f^{-1}(D_i)) = 0$. Then $E(f^{-1}(V)) = 0$. The **essential range of** f is $\mathbb{C} \setminus V$, and we say that f is **essentially bounded** if the essential range of f is a bounded subset of $\mathbb{C}$. We define the **essential supremum of** f to be

$$\|f\|_\infty = \sup\{|\lambda| : \lambda \in \mathbb{C} \setminus V\}.$$

Now define B to be the collection of bounded measurable functions $(\Omega, \mathscr{S}) \to (\mathbb{C}, \mathscr{B}_{\mathbb{C}})$, which is a Banach algebra with the norm

$$\sup\{|f(\omega) : \omega \in \Omega\},$$

for which

$$N = \{f \in B : \|f\|_\infty = 0\}$$

is a closed ideal. Then B/N is a Banach algebra, denoted $L^\infty(E)$, with the norm

$$\|f + N\|_\infty = \|f\|_\infty.$$

The unity of $L^\infty(E)$ is $1 + N$. Because $L^\infty(E)$ is a Banach algebra, it makes sense to speak about the spectrum of an element of $L^\infty(E)$. For $f + N \in L^\infty(E)$, the spectrum of $f + N$ is the set of those $\lambda \in \mathbb{C}$ for which there is no $g + N \in L^\infty(E)$ satisfying $(g + N)(f + N - \lambda(1 + N)) = 1 + N$. Check that the spectrum of $f + N$ is equal to the essential range of g, for any $g \in f + N$.

A subset A of $\mathscr{B}(H)$ is said to be **normal** when $ST = TS$ for all $S, T \in A$ and $T \in A$ implies that $T^* \in A$.[9] (To say that $T \in \mathscr{B}(H)$ is normal means that $TT^* = T^*T$, and this is equivalent to the statement that the set $\{T, T^*\}$ is normal.)

Theorem 11. *If $(\Omega, \mathscr{S})$ is a measurable space and $E : \mathscr{S} \to H$ is a resolution of the identity, then there is a closed normal subalgebra A of $\mathscr{B}(H)$ and a unique isometric *-isomorphism $\Psi : L^\infty(E) \to A$ such that*

$$\langle \Psi(f)x, y \rangle = \int_\Omega f dE_{x,y}, \qquad f \in L^\infty(E), \quad x, y \in H.$$

[9] Walter Rudin, *Functional Analysis*, second ed., p. 319, Theorem 12.21.

Furthermore,

$$\|\Psi(f)x\|^2 = \int_\Omega |f|^2 dE_{x,x}, \qquad f \in L^\infty(E), \quad x \in H.$$

For $f \in L^\infty(E)$, we define

$$\int_\Omega f dE = \Psi(f).$$

For $L^\infty(E)$, $\sigma(\Psi(f))$ is equal to the essential range of f.[10]

6 The spectral theorem

The following is the spectral theorem for self-adjoint operators.[11]

Theorem 12. *If T is a self-adjoint operator in H, then there is a unique resolution of the identity*

$$E : \mathscr{B}_\mathbb{R} \to \mathscr{B}(H)$$

such that

$$\langle Tx, y\rangle = \int_\mathbb{R} \lambda dE_{x,y}(\lambda), \qquad x \in \mathscr{D}(T), \quad y \in H.$$

This resolution of the identity satisfies $E(\sigma(T)) = I$.

If T is a self-adjoint operator in H applying the spectral theorem and then Theorem 11, we get that there is a closed normal subalgebra A of $\mathscr{B}(H)$ and a unique isometric *-isomorphism $\Psi : L^\infty(E) \to A$ such that

$$\langle \Psi(f)x, y\rangle = \int_{\sigma(T)} f(\lambda) dE_{x,y}(\lambda), \qquad f \in L^\infty(E), \quad x, y \in H.$$

For $t \in \mathbb{R}$ and $f_t : \sigma(T) \to \mathbb{C}$ defined by $f_t(\lambda) = e^{it\lambda}$, this defines

$$e^{itT} = \Psi(f_t) = \int_{\sigma(T)} e^{it\lambda} dE(\lambda).$$

Because Ψ is a *-homomorphism, for $t \in \mathbb{R}$ we have

$$\Psi(f_t)^*\Psi(f_t) = \Psi(\overline{f_t})\Psi(f_t) = \Psi(f_{-t})\Psi(f_t) = \Psi(f_{-t}f_t) = \Psi(f_0) = I,$$

and likewise $\Psi(f_t)\Psi(f_t)^* = I$, showing that $e^{itT} = \Psi(f_t)$ is unitary. We denote by $\mathscr{U}(H)$ the collection of unitary elements of $\mathscr{B}(H)$. $\mathscr{U}(H)$ is a subgroup of the group of invertible elements of $\mathscr{B}(H)$.

[10]Walter Rudin, *Functional Analysis*, second ed., p. 366, Theorem 13.27.
[11]Walter Rudin, *Functional Analysis*, second ed., p. 368, Theorem 13.30.

Furthermore, because Ψ is a *-homomorphism, for $t \in \mathbb{R}$ we have

$$I = \Psi(f_0) = \Psi(f_t f_{-t}) = \Psi(f_t)\Psi(f_{-t}) = e^{itT} e^{i(-t)T},$$

and for $s, t \in \mathbb{R}$ we have

$$e^{isT} e^{itT} = \Psi(f_s)\Psi(f_t) = \Psi(f_s f_t) = \Psi(f_{s+t}) = e^{i(s+t)T},$$

showing that $t \mapsto e^{itT}$ is a one-parameter group $\mathbb{R} \to \mathscr{B}(H)$.

For $t \in \mathbb{R}$ and $x \in H$, by Theorem 11 we have

$$\|\Psi_t x - x\|^2 = \|\Psi(f_t - 1)x\|^2 = \int_{\sigma(T)} |f_t - 1|^2 dE_{x,x} = \int_{\sigma(T)} |e^{it\lambda} - 1|^2 dE_{x,x}(\lambda).$$

For each $\lambda \in \sigma(T)$, $|e^{it\lambda} - 1|^2 \to 0$ as $t \to 0$, and thus we get by the dominated convergence theorem

$$\int_{\sigma(T)} |e^{it\lambda} - 1|^2 dE_{x,x}(\lambda) \to 0, \qquad t \to 0.$$

That is, for each $x \in H$,

$$\left\| e^{itT} x - x \right\| \to 0$$

as $t \to 0$, showing that $t \mapsto e^{itT}$ is **strongly continuous**, i.e. $t \mapsto e^{itT}$ is continuous $\mathbb{R} \to \mathscr{B}(H)$ where $\mathscr{B}(H)$ has the strong operator topology.

Conversely, **Stone's theorem on one-parameter unitary groups**[12] states that if $\{U_t : t \in \mathbb{R}\}$ is a strongly continuous one-parameter group of bounded unitary operators on H, then there is a unique self-adjoint operator A in H such that $U_t = e^{itA}$ for each $t \in \mathbb{R}$.

For $t \neq 0$, define $g_t : \sigma(T) \to \mathbb{C}$ by $g_t(\lambda) = \frac{e^{it\lambda} - 1}{t}$. By Theorem 12, for $x \in \mathscr{D}(T)$ and $y \in H$,

$$\langle iTx, y \rangle = i \langle Tx, y \rangle = i \int_{\mathbb{R}} \lambda dE_{x,y}(\lambda)$$

and by Theorem 11,

$$\langle \Psi(g_t)x, y \rangle = \int_{\sigma(T)} g_t dE_{x,y} = \int_{\sigma(T)} \frac{e^{it\lambda} - 1}{t} dE_{x,y}(\lambda),$$

so

$$\langle \Psi(g_t)x - iTx, y \rangle = \int_{\sigma(T)} \left(\frac{e^{it\lambda} - 1}{t} - i\lambda \right) dE_{x,y}(\lambda).$$

For each $\lambda \in \sigma(T)$, $\frac{e^{it\lambda} - 1}{t} - i\lambda \to 0$ as $t \to 0$, and for each t,

$$\left| \frac{e^{it\lambda} - 1}{t} - i\lambda \right| \leq \left| \frac{e^{it\lambda} - 1}{t} \right| + |\lambda| \leq 2|\lambda|,$$

[12] cf. Walter Rudin, *Functional Analysis*, second ed., p. 382, Theorem 38.

and as $x \in \mathscr{D}(T)$, by Theorem 12 we have that $\lambda \mapsto |\lambda|$ belongs to $L^1(E_{x,y})$. Thus by the dominated convergence theorem,

$$\langle \Psi(g_t)x - iTx, y\rangle = \int_{\sigma(T)} \left(\frac{e^{it\lambda} - 1}{t} - i\lambda \right) dE_{x,y}(\lambda) \to 0$$

as $t \to 0$. In particular,

$$\|\Psi(g_t)x - iTx\|^2 \to 0$$

as $t \to 0$. That is, for each $x \in \mathscr{D}(T)$,

$$\frac{e^{itT}x - x}{t} \to iTx$$

as $t \to 0$. In other words, iT is the **infinitesimal generator** of the one-parameter group e^{itT}.[13] We remark that because $T^* = T$, the adjoint of iT is $(iT)^* = \bar{i}T^* = -iT^* = -iT = -(iT)$.

7 Trotter product formula

We remind ourselves that for an operator T in H to be closed means that $\mathscr{G}(T)$ is a closed linear subspace of $H \times H$.

Theorem 13. *Let T be an operator in H. T is closed if and only if the linear space $\mathscr{D}(T)$ with the norm*

$$\|x\|_T = \|x\| + \|Tx\|.$$

is a Banach space.

The following is the **Trotter product formula**, which shows that if A, B, and $A+B$ are self-adjoint operators in a Hilbert space, then for each t, $(e^{itA/n}e^{itB/n})^n$ converges strongly to $e^{it(A+B)}$ as $n \to \infty$.[14]

Theorem 14. *Let H be a Hilbert space, not necessarily separable. If A and B are self-adjoint operators in H such that $A+B$ is a self-adjoint operator in H, then for each $t \in \mathbb{R}$ and for each $\psi \in H$,*

$$e^{it(A+B)}\psi = \lim_{n\to\infty} \left((e^{itA/n}e^{itb/n})^n \psi \right).$$

Proof. The claim is immediate for $t = 0$, and we prove the claim for $t > 0$; it is straightforward to obtain the claim for $t < 0$ using the truth of the claim for $t > 0$. Let $D = \mathscr{D}(A+B) = \mathscr{D}(A) \cap \mathscr{D}(B)$. Because $A+B$ is self-adjoint, $A+B$ is closed (Theorem 4), so by Theorem 13, the linear space D with the norm $\|\phi\|_{A+B} = \|\phi\| + \|(A+B)\phi\|$ is a Banach space. Because D is a Banach

[13] cf. Walter Rudin, *Functional Analysis*, second ed., p. 376, Theorem 13.35.

[14] Barry Simon, *Functional Integration and Quantum Physics*, p. 4, Theorem 1.1; Konrad Schmüdgen, *Unbounded Self-adjoint Operators on Hilbert Space*, p. 122, Theorem 6.4.

space, the uniform boundedness principle[15] tells us that if Γ is a collection of bounded linear maps $D \to H$ and if for each $\phi \in D$ the set $\{\gamma\phi : \gamma \in \Gamma\}$ is bounded in H, then the set $\{\|\gamma\| : \gamma \in \Gamma\}$ is bounded, i.e. there is some C such that $\|\gamma\phi\| \leq C \|\phi\|_{A+B}$ for all $\gamma \in \Gamma$ and all $\phi \in D$.

For $s \in \mathbb{R}$, let $S_s = e^{is(A+B)}$, $V_s = e^{isA}$, $W_s = e^{isB}$, $U_s = V_s W_s$, which each belong to $\mathscr{B}(H)$. For $n \geq 1$,

$$\sum_{j=0}^{n-1} U_{t/n}^j (S_{t/n} - U_{t/n}) S_{t/n}^{n-j-1} = U_{t/n}^n - S_{t/n}^n = U_{t/n}^n - S_t,$$

so, because a product of unitary operators is a unitary operator and a unitary operator has operator norm 1 and also using the fact that $S_{t/n}^{n-j-1} = S_{t-\frac{j+1}{n}}$, for $\xi \in H$ we have

$$\begin{aligned}
\left\|(S_t - U_{t/n}^n)\xi\right\| &= \left\|\sum_{j=0}^{n-1} U_{t/n}^j (S_{t/n} - U_{t/n}) S_{t/n}^{n-j-1}\xi\right\| \\
&\leq \sum_{j=0}^{n-1} \left\|(S_{t/n} - U_{t/n}) S_{t/n}^{n-j-1}\xi\right\| \\
&= \sum_{j=0}^{n-1} \left\|(S_{t/n} - U_{t/n}) S_{t-\frac{j+1}{n}}\xi\right\| \\
&\leq \sum_{j=0}^{n-1} \sup_{0 \leq s \leq t} \left\|(S_{t/n} - U_{t/n}) S_s\xi\right\|.
\end{aligned}$$

That is,

$$\left\|(S_t - U_{t/n}^n)\xi\right\| \leq n \sup_{0 \leq s \leq t} \left\|(S_{t/n} - U_{t/n}) S_s\xi\right\|, \qquad \xi \in H, \quad n \geq 1. \tag{1}$$

Let $\phi \in D$. On the one hand, because $i(A+B)$ is the infinitesimal generator of $\{S_s : s \in \mathbb{R}\}$, we have

$$\frac{S_s - I}{s}\phi \to i(A+B)\phi, \qquad s \downarrow 0. \tag{2}$$

On the other hand, for $s \neq 0$ we have, because an infinitesimal generator of a one-parameter group commutes with each element of the one-parameter group,

$$V_s(iB\phi) + V_s\left(\frac{W_s - I}{s} - iB\right)\phi + \frac{V_s - I}{s}\phi = \frac{U_s - I}{s}\phi,$$

and as V_s converges strongly to I as $s \downarrow 0$ and as iB is the infinitesimal generator of the one-parameter group $\{W_s : s \in \mathbb{R}\}$ and iA is the infinitesimal generator

[15] Walter Rudin, *Functional Analysis*, second ed., p. 45, Theorem 2.6.

of the one-parameter group $\{V_s : s \in \mathbb{R}\}$,

$$V_s(iB\phi) + V_s\left(\frac{W_s - I}{s} - iB\right)\phi + \frac{V_s - I}{s}\phi \to iB\phi + iA\phi$$

as $s \downarrow 0$, i.e.

$$\frac{U_s - I}{s}\phi \to i(A+B)\phi, \qquad s \downarrow 0. \tag{3}$$

Using (2) and (3), we get that for each $\phi \in D$,

$$\frac{S_s - U_s}{s}\phi \to 0, \qquad s \downarrow 0.$$

Therefore, for each $\phi \in D$, with $s = t/n$ we have

$$\frac{n}{t}(S_{t/n} - U_{t/n})\phi \to 0, \qquad n \to \infty,$$

equivalently (t is fixed for this whole theorem),

$$\lim_{n\to\infty} \left\| n(S_{t/n} - U_{t/n})\phi \right\| = 0, \qquad \phi \in D. \tag{4}$$

For each $n \geq 1$, define $\gamma_n : D \to H$ by $\gamma_n = n(S_{t/n} - U_{t/n})$. Each γ_n is a linear map, and for $\phi \in D$,

$$\|\gamma_n\phi\| \leq n\left\|S_{t/n}\phi\right\| + n\left\|U_{t/n}\phi\right\| \leq n\|\phi\| + n\|\phi\| \leq 2n\|\phi\|_{A+B},$$

showing that each γ_n is a bounded linear map $D \to H$, where D is a Banach space with the norm $\|\phi\|_{A+B} = \|\phi\| + \|(A+B)\phi\|$. Moreover, (4) shows that for each $\phi \in D$, there is some C_ϕ such that

$$\|\gamma_n\phi\| \leq C_\phi, \qquad n \geq 1.$$

Then applying the uniform boundedness principle, we get that there is some $C > 0$ such that for all $n \geq 1$ and for all $\phi \in D$,

$$\|\gamma_n\phi\| \leq C\|\phi\|_{A+B},$$

i.e.

$$\left\|n(S_{t/n} - U_{t/n})\phi\right\| \leq C\|\phi\|_{A+B}, \qquad n \geq 1, \quad \phi \in D. \tag{5}$$

Let K be a compact subset of D, where D is a Banach space with the norm $\|\phi\|_{A+B} = \|\phi\| + \|(A+B)\phi\|$. Then K is totally bounded, so for any $\epsilon > 0$, there are $\phi_1, \ldots, \phi_M \in K$ such that $K \subset \bigcup_{m=1}^M B_{\epsilon/C}(\phi_m)$. By (4), for each m, $1 \leq m \leq M$, there is some n_m such that when $n \geq n_m$,

$$\left\|n(S_{t/n} - U_{t/n})\phi_m\right\| \leq \epsilon.$$

Let $N = \max\{n_1, \dots, n_M\}$. For $n \geq N$ and for $\phi \in D$, there is some m for which $\|\phi - \phi_m\|_{A+B} < \frac{\epsilon}{C}$, and using (5), as $\phi - \phi_m \in D$, we get

$$\begin{aligned}\left\|n(S_{t/n} - U_{t/n})\phi\right\| &\leq \left\|n(S_{t/n} - U_{t/n})(\phi - \phi_m)\right\| + \left\|n(S_{t/n} - U_{t/n})\phi_m\right\| \\ &\leq C \|\phi - \phi_m\|_{A+B} + \epsilon \\ &< \epsilon + \epsilon.\end{aligned}$$

This shows that any compact subset K of D and $\epsilon > 0$, there is some n_ϵ such that if $n \geq n_\epsilon$ and $\phi \in K$, then

$$\left\|n(S_{t/n} - U_{t/n})\phi\right\| < \epsilon. \tag{6}$$

Let $\phi \in D$, let $s_0 \in \mathbb{R}$, and let $\epsilon > 0$. Because $s \mapsto S_s$ is strongly continuous $\mathbb{R} \to \mathscr{B}(H)$, there is some $\delta_1 > 0$ such that when $|s - s_0| < \delta_1$, $\|S_s\phi - S_{s_0}\phi\| < \epsilon$, and there is some $\delta_2 > 0$ such that when $|s - s_0| < \delta_2$, $\|S_s(A+B)\phi - S_{s_0}(A+B)\phi\| < \epsilon$, and hence with $\delta = \min\{\delta_1, \delta_2\}$, when $|s - s_0| < \delta$ we have

$$\begin{aligned}\|S_s\phi - S_{s_0}\phi\|_{A+B} &= \|S_s\phi - S_{s_0}\phi\| + \|(A+B)(S_s\phi - S_{s_0}\phi)\| \\ &= \|S_s\phi - S_{s_0}\phi\| + \|S_s(A+B)\phi - S_{s_0}(A+B)\phi)\| \\ &< \epsilon + \epsilon,\end{aligned}$$

showing that $s \mapsto S_s\phi$ is continuous $\mathbb{R} \to D$. Therefore $\{S_s\phi : 0 \leq s \leq t\}$ is a compact subset of D, so applying (6) we get that for any $\epsilon > 0$, there is some n_ϵ such that if $n \geq n_\epsilon$ and $0 \leq s \leq t$, then

$$\left\|n(S_{t/n} - U_{t/n})S_s\phi\right\| < \epsilon,$$

and therefore if $n \geq n_\epsilon$ then

$$\sup_{0 \leq s \leq t} \left\|n(S_{t/n} - U_{t/n})S_s\phi\right\| \leq \epsilon. \tag{7}$$

Finally, let $\epsilon > 0$. The statement that $A+B$ is self-adjoint in H entails the statement that D is dense in H, so there is some $\phi \in D$ such that $\|\phi - \psi\| < \epsilon$. For $n \geq 1$,

$$\begin{aligned}\left\|(S_t - U_{t/n}^n)\psi\right\| &\leq \left\|(S_t - U_{t/n}^n)(\psi - \phi)\right\| + \left\|(S_t - U_{t/n}^n)\phi\right\| \\ &\leq 2\|\psi - \phi\| + \left\|(S_t - U_{t/n}^n)\phi\right\| \\ &< \epsilon + \left\|(S_t - U_{t/n}^n)\phi\right\|.\end{aligned}$$

Using (1) with $\xi = \phi$ and then using (7), there is some n_ϵ such that when $n \geq n_\epsilon$,

$$\left\|(S_t - U_{t/n}^n)\phi\right\| \leq n \sup_{0 \leq s \leq t} \left\|(S_{t/n} - U_{t/n})S_s\phi\right\| \leq \epsilon.$$

Therefore for $n \geq n_\epsilon$,

$$\left\|(S_t - U_{t/n}^n)\psi\right\| < 2\epsilon,$$

proving the claim. □