FROM THE DEFINITION OF A HILBERT SPACE TO THE PROOF OF SPECTRAL THEOREMS

Maciej Horowski

Institute of Physics, Warsaw University Division
15-424 Bialystok, Lipowa 41, Poland

1 Hilbert spaces

1.1

Let H be a complex vector space equipped with a scalar product denoted by $(\cdot|\cdot)$, so $(H, (\cdot|\cdot))$ is a unitary space. Every scalar product $(\cdot|\cdot)$ defines a norm $\|\cdot\|$ by

$$\|x\|^2 \stackrel{\text{df}}{=} (x|x) \quad x \in H.$$

Definition 1.1.1 *A complete unitary space is called a Hilbert space.*

It will be usefull to recall the two following inequalities:
a) Schwarz inequality

$$\forall x, y \in H \quad |(x|y)|^2 \leq (x|x)(y|y);$$

b) Minkowski inequality

$$\forall x, y \in H \quad \|x + y\| \leq \|x\| + \|y\|.$$

1.2 Examples of Hilbert spaces

a) $\mathbf{C}^n$ with the standard scalar product:

$$(z|w) \stackrel{\text{df}}{=} \sum_{i=1}^{n} \bar{z}_i w_i \quad \forall z = (z_1, \ldots, z_n), w = (w_1, \ldots, w_n) \in \mathbf{C}^n;$$

b)

$$l^2 \stackrel{\text{df}}{=} \left\{ \{a_n\} \subset \mathbf{C}, n \in \mathbf{N} : \sum_n |a_n|^2 < \infty \right\},$$

with the scalar product given by:

$$(\{a_n\}|\{b_n\}) \stackrel{\text{df}}{=} \sum_n \bar{a}_n b_n;$$

c) $L^2(\mathbf{R}^n, d\lambda)$, where $d\lambda$ is the Lebesgue measure, with the standard scalar product:

$$(f|g) \stackrel{\text{df}}{=} \int_{\mathbf{R}^n} \overline{f(x)} g(x) d\lambda(x), \quad f, g \in L^2(\mathbf{R}^n, d\lambda).$$

1.3

Definition 1.3.1 *a) A Hilbert space is called separable if there exists a countable set which is dense in this space or, equivalently, this space has a countable basis.*
b) Two vectors $x, y \in H$ are called orthogonal if $(x|y) = 0$; we denote it by $x \perp y$.
c) A set in a Hilbert space is called orthogonal if any two elements of this set are orthogonal.
d) If the norm of any element of an orthogonal set is one, this set is called orthonormal.
e) An orthonormal (o.n.) basis of a Hilbert space is any complete orthonormal set.

1.4

Proposition 1.4.1 *If H_1 and H_2 are Hilbert spaces, we have a Hilbert space structure on $H_1 \times H_2$ given by:*

$$((x,y)|(x',y'))_{H_1 \times H_2} \stackrel{\text{df}}{=} (x|x')_{H_1} + (y|y')_{H_2} \quad \forall (x,y),(x',y') \in H_1 \times H_2.$$

□

Definition 1.4.1 *The Hilbert space $H_1 \times H_2$ with the scalar product defined above is called a direct sum of Hilbert spaces and we denote it by $H_1 \oplus H_2$.*

Proposition 1.4.2 *A scalar product as a function $(\cdot|\cdot) : H \times H \to \mathbf{C}$ is continuous.*

Proof
If $x_n \to x$ and $y_n \to y$, then $\{x_n\}$, as a convergent sequence is bounded.

$$\begin{aligned} &|(x_n|y_n) - (x|y) - (x_n|y) + (x_n|y)| \leq \\ &\leq |(x_n|y_n - y) + (x_n - x|y)| \leq \\ &\leq M\|y_n - y\| + \|y\|\|x_n - x\|. \end{aligned}$$

□

1.5

Proposition 1.5.1 *If $\{e_n\}_{n \in \mathbf{N}}$ is an o.n. basis and $\{t_n\} \in l^2$, the series $\sum_n t_n e_n$ is convergent and*

$$\|\sum_n t_n e_n\| = \left(\sum_n |t_n|^2\right)^{\frac{1}{2}}. \tag{1}$$

Proof
Let $S_N \stackrel{\text{df}}{=} \sum_{n=1}^N t_n e_n$, then

$$\|S_N - S_M\|^2 = \|\sum_{n=M+1}^N t_n e_n\|^2 =$$

$$= (t_{M+1}e_{M+1} + ... + t_N e_N | t_{M+1}e_{M+1} + ... + t_N e_N) = \sum_{n=M+1}^N |t_n|^2.$$

Since the series $\sum_n |t_n|^2$ is convergent, the sequence $\{S_N\}$ is Couchy, so it is convergent.
Since the scalar product is continuous, we have

$$\left(\sum_n t_n e_n | \sum_m t_m e_m\right) = \sum_{n,m} \bar{t}_n t_m (e_n|e_m) = \sum_n |t_n|^2.$$

□

Proposition 1.5.2 *If $\{e_n\}$ is an o.n. basis, $x \in H$ and $t_n \stackrel{\mathrm{df}}{=} (e_n|x)$, then*
a) $\{t_n\} \in l^2$;
b) $x = \sum_n e_n t_n$;
c) $\|x\|^2 = \sum_n |t_n|^2$.

□

1.6 Example. Fourier series

Let us consider $L^2([0, 2\pi])$ with a scalar product given by

$$(f|g) \stackrel{\mathrm{df}}{=} \frac{1}{2\pi} \int_0^{2\pi} \bar{f(\lambda)} g(\lambda) d\lambda.$$

The family of functions defined by

$$e_k(\lambda) \stackrel{\mathrm{df}}{=} e^{i\lambda k}, \quad k \in \mathbf{Z},$$

form an orthonormal set and we shall show that it is an o.n. basis:
Let as assume that there exists $f \in L^2([0, 2\pi])$ such that $(f|f) = 1$ and $(e_k|f) = 0 \ \ \forall k \in \mathbf{Z}$. Then we have

$$0 = (e_k|f) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\lambda} f(\lambda) d\lambda \Rightarrow$$

$$\Rightarrow \int_0^{2\pi} \sum_{k=-N}^{N} a_k e^{-ik\lambda} f(\lambda) d\lambda = 0 \quad \forall N \in \mathbf{N}, \forall a_k \in \mathbf{C}.$$

Inasmuch as the family of functions of the form $\sum_{k=-N}^{N} a_k e^{-i\lambda k}$ fulfil the assumptions of the Stone–Weierstrass theorem on a circle, so by this theorem $f(\lambda) \equiv 0$, which is contradiction. □

Now let as consider the consequence of the Proposition 1.5.2:

$$f \in L^2([0, 2\pi]) \ \Rightarrow \ t_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\lambda} f(\lambda) d\lambda,$$

moreover

$$\frac{1}{2\pi} \int_0^{2\pi} |f(\lambda)|^2 d\lambda = \sum_{k=-\infty}^{\infty} |t_k|^2,$$

finally we obtain the Fourier series

$$f(\lambda) = \sum_{k=-\infty}^{\infty} t_k e^{ik\lambda},$$

which means that

$$\frac{1}{2\pi} \int_0^{2\pi} |f(\lambda) - \sum_{k=-N}^{N} t_k e^{ik\lambda}|^2 d\lambda \longrightarrow 0$$

if $N \to \infty$.

2 Subspaces

2.1

Definition 2.1.1 *A vector subspace F of a Hilbert space H is closed if F is closed in the topology generated by the scalar product in H.*

2.2

Definition 2.2.1 *Let G be any subset of H, then*

$$G^{\perp} \stackrel{\text{df}}{=} \{x \in H : \forall y \in G \ \ (x|y) = 0\}.$$

Proposition 2.2.1 *Let G be any subset of H, then*
a) $G^{\perp}$ is a closed subspace;
b) $G \subset G^{\perp\perp}$.

□

2.3

Theorem 2.3.1 (on the orthogonal projection (Beppo-Levi)) *If F is a closed subspace of H and $x \in H$, then*
a) $x = x_{\parallel} + x_{\perp}$, where $x_{\parallel} \in F$, $x_{\perp} \in F^{\perp}$;
b) this decomposition is unique.

<u>Proof</u>
a) Let $\{e_n\}$ be an o.n. basis in F and let

$$x_{\parallel} \stackrel{\text{df}}{=} \sum_n (e_n|x)e_n \quad \text{and} \quad x_{\perp} \stackrel{\text{df}}{=} x - x_{\parallel}.$$

It is enough to show that $x_{\perp} \in F^{\perp}$:

$$(e_m|x_{\perp}) = (e_m|x) - (e_m|\sum_n (e_n|x)e_n) = 0.$$

b) Let $x = x_{\parallel} + x_{\perp}$ and $x = x'_{\parallel} + x'_{\perp}$, then

$$F \ni x'_{\parallel} - x_{\parallel} = x_{\perp} - x'_{\perp} \in F^{\perp} \quad \Rightarrow \quad x'_{\parallel} - x_{\parallel} = 0.$$

□

Corollary 2.3.1 *If F is a closed subspace, then $F^{\perp\perp} = F$.*

<u>Proof</u>
It is enough to show that $F^{\perp\perp} \subset F$. Let $x \in F^{\perp\perp}$, then by the Theorem 2.3.1 we have

$$x = x_{\parallel} + x_{\perp} \quad x_{\parallel} \in F, \ x_{\perp} \in F^{\perp},$$

therefore

$$0 = (x|x_{\perp}) = (x_{\parallel}|x_{\perp}) + (x_{\perp}|x_{\perp}),$$

which gives

$$(x_{\perp}|x_{\perp}) = 0 \quad \Rightarrow \quad x_{\perp} = 0 \quad \Rightarrow \quad x \in F.$$

□

2.4

Proposition 2.4.1 *a) $G^{\perp\perp}$ is the smallest closed subspace containing the set G.*
b) Let F be a subspace, then F is dense if and only if $F^{\perp} = \{0\}$.

□

3 Linear functionals

3.1

Definition 3.1.1 *A linear (antilinear) functional is any linear (antilinear) map* $l : H \to \mathbf{C}$.

Proposition 3.1.1 *A linear functional* l *is continuous if and only if*

$$\exists c > 0 \ \ \forall x \in H \quad |l(x)| \leq c\|x\|.$$

□

3.2

Theorem 3.2.1 (Frechet, Riesz) *For every continuous linear functional* l *there exists a vector* $y \in H$ *such that*

$$l(x) = (y|x) \quad \forall x \in H.$$

<u>Proof</u>
Let $\{e_n\}$ be an o.n. basis in H. First we shall show, that the series

$$\sum_n \overline{l(e_n)} e_n$$

is convergent:

$$\sum_n |l(e_n)|^2 = \sum_n \overline{l(e_n)} l(e_n) = l(\sum_n \overline{l(e_n)} e_n) \leq$$

$$\leq c\| \sum_n \overline{l(e_n)} e_n \| \leq c \left(\sum_n |l(e_n)|^2 \right)^{\frac{1}{2}},$$

hence

$$\sum_n |l(e_n)|^2 \leq c^2.$$

Now we can put

$$y \stackrel{\text{df}}{=} \sum_n \overline{l(e_n)} e_n$$

and calculate

$$(y|e_n) = (\sum_m \overline{l(e_m)} e_m | e_n) = l(e_n).$$

□

Remark 3.2.1 *It is very easy to formulate an analog of this theorem for continuous antilinear functionals.*

4 Barrelled spaces

4.1

Definition 4.1.1 *A subset* $Q \subset H$ *is called absolutely convex if*

$$\forall x, y \in Q \ \ \forall \alpha, \beta \in \mathbf{C} \ \ \left(|\alpha|^2 + |\beta|^2 \leq 1\right) \Rightarrow (\alpha x + \beta y \in Q)\,.$$

Definition 4.1.2 $Q \subset H$ *is called absorbing if* $\bigcup_{n=1}^{\infty} nQ = H$.

Definition 4.1.3 *A closed, absolutely convex, and absorbing set is called a barrel.*

Definition 4.1.4 *A topological vector space is called barrelled if every barrel is a neighbourhood of zero.*

Proposition 4.1.1 *A Hilbert space is barrelled.*

□

4.2

Proposition 4.2.1 *Let* R *be a subset in* H *such that*

$$\forall x \in H \ \ \exists M > 0 \ \ \forall y \in R \ \ |(x|y)| < M.$$

Then

$$\exists N > 0 \ \ \forall y \in R \ \ \|y\| < N.$$

Proof
Let $R^o \stackrel{\text{df}}{=} \{z \in H : \forall y \in R \ \ |(z|y)| \leq 1\}$. It is easy to see that R^o is closed and absolutely convex. We shall show that it is absorbing:
If $x \in H$, then $\frac{1}{M}x \in R^o$. Thus $x \in MR^o \subset nR^o$ for some $n \in \mathbf{N}$.
Hence R^o is a barrel, so it is a neighbourhood of zero i.e. there exists $r > 0$ such that the closed ball $K(0, r) \subset R^o$.
Let $y \in R$. Since $\|\frac{ry}{\|y\|}\| = r$, $\frac{ry}{\|y\|} \in R^o$, which gives

$$|(\frac{ry}{\|y\|}|y)| = r\|y\| \leq 1,$$

therefore

$$\|y\| \leq \frac{1}{r}.$$

□

5 Bounded operators

5.1

Definition 5.1.1 *A linear map* $A : H_1 \to H_2$ *between two Hilbert spaces is called a bounded operator if*

$$\exists M \geq 0 \ \ \forall x \in H_1 \quad \|Ax\|_2 \leq M\|x\|_1. \tag{2}$$

The space of all bounded operators beetwen H_1 *and* H_2 *will be denoted by* $B(H_1, H_2)$*. Moreover, we define* $B(H) \stackrel{\text{df}}{=} B(H, H)$.

Proposition 5.1.1 *a) If* $\dim H_1 < \infty$*, then every linear map is bounded.*
b) Every bounded operator is continuous.

□

5.2

Of course, $B(H_1, H_2)$ is a complex vector space and, in addition, it has a natural metric structure:

$$\|A\| \stackrel{\text{df}}{=} \sup_{\|x\|\le 1} \frac{\|Ax\|_2}{\|x\|_1},$$

so $\|A\|$ is the smalest M which satisfies the condition (2). It is not difficult to show that in the topology generated by this norm $B(H_1, H_2)$ is complete (i.e. it is a Banach space), moreover, for $B(H)$ we have the following inequality for the superposition of operators:

$$\|A \cdot B\| \le \|A\| \cdot \|B\|.$$

Definition 5.2.1 *Let $\{A_n\}$ be a sequence in $B(H_1, H_2)$. Then we say:*
a) $\{A_n\}$ is convergent if $\{\|A_n\|\}$ is convergent;
b) $\{A_n\}$ is strongly convergent if for all $x \in H$ $\{\|A_n x\|\}$ is convergent;
c) $\{A_n\}$ is weakly convergent if for all $x, y \in H$ $\{(x|A_n y)\}$ is convergent.

5.3

Proposition 5.3.1 *For every $A \in B(H_1, H_2)$ there exists exactly one $A^* \in B(H_2, H_1)$ such that*

$$\forall x \in H_1,\ \forall y \in H_2 \quad (y|Ax)_2 = (A^*y|x)_1.$$

Proof
For $y \in H_2$ we have the following linear map:

$$H_1 \ni x \longmapsto (y|Ax) \in \mathbf{C}. \tag{3}$$

Because of

$$|(y|Ax)| \le \|y\|\|Ax\| \le \|y\|\|A\|\|x\|,$$

the functional (3) is continuous, so by the Theorem 3.2.1 there exists $z_y \in H_1$ such that $(y|Ax) = (z_y|x)$ for all $x \in B(H_1)$.
Now we can define a linear map

$$H_2 \ni y \longmapsto A^*y \stackrel{\text{df}}{=} z_y \in H_1.$$

A^* is linear:

$$\begin{aligned}(A^*(\alpha y_1 + \beta y_2)|x) &= (\alpha y_1 + \beta y_2|Ax) = \\ &= \bar{\alpha}(y_1|Ax) + \bar{\beta}(y_2|Ax) = \\ &= \bar{\alpha}(A^*y_1|x) + \bar{\beta}(A^*y_2|x) = \\ &= (\alpha A^*y_1 + \beta A^*y_2|x),\end{aligned}$$

A^* is continuous:

$$\begin{aligned}\|A^*y\|^2 = (A^*y|A^*y) = (y|AA^*y) \le \\ \le \|y\|\|A\|\|A^*y\|,\end{aligned}$$

therefore

$$\|A^*y\| \le \|y\|\|A\|,$$

hence

$$\|A^*\| \le \|A\|.$$

□

Definition 5.3.1 *The operator A^* is called an adjoint operator of A.*

5.4

The following proposition show the elementary properties of the map

$$*: B(H_1, H_2) \to B(H_2, H_1).$$

Proposition 5.4.1 *a)* $A^{**} = A$, *so* $*$ *is an involution;*
b) $\|A^*\| = \|A\|$*;*
c) $(\lambda A + \mu B)^* = \bar{\lambda} A^* + \bar{\mu} B^*$, *so* $*$ *is antilinear;*
d) $(AB)^* = B^* A^*$*;*
e) $\|A^* A\| = \|A\|^2$.

Proof
a) $(y|A^{**}x) = (A^*y|x) = (y|Ax)$,
b) from the proof of the Proposition 5.3.1 we know that $\|A^*\| \leq \|A\|$, but $\|A\| = \|A^{**}\| \leq \|A^*\|$,
e) first, we have

$$\|A^* A\| \leq \|A^*\| \|A\| = \|A\|^2,$$

second, we have

$$\|A\|^2 = \sup_{\|x\| \leq 1} \|Ax\|^2 = \sup_{\|x\| \leq 1} (Ax|Ax) = \sup_{\|x\| \leq 1} (x|A^* Ax) \leq$$

$$\leq \sup_{\|x\| \leq 1} \|x\| \|A^* Ax\| \leq \sup_{\|x\| \leq 1} \|x\| \|A^* A\| \|x\| \leq \|A^* A\|.$$

□

5.5

Definition 5.5.1 *In the case* $H_1 = H_2 = H$ *we define*
a) if $A = A^*$, *then* A *will be called a Hermitian operator;*
b) if for all $x \in H$ $(x|Ax) \geq 0$, *then* A *will be called a positive operator;*
c) if $AA^* = A^* A$, *then* A *will be called a normal operator.*

Proposition 5.5.1 *Every positive operator is Hermitian.*

□

5.6 Projection operators

Let as consider a closed subspace $F \subset H$. Then by the Theorem 2.3.1, for all $x \in H$ we have $x = x_\parallel + x_\perp$, where $x_\parallel \in F$ and $x_\perp \in F^\perp$. Because this decomposition is unique, we may define the following linear operator:

$$H \ni x \mapsto P_F x \stackrel{\mathrm{df}}{=} x_\parallel \in H.$$

Proposition 5.6.1 *a)* $P_F \in B(H)$, *i.e.* P_F *is continuous;*
b) $P_F^2 = P_F$, *i.e.* P_F *is idempotent;*
c) $P_F^* = P_F$, *i.e.* P_F *is Hermitian.*

Proof
a) $\|P_F x\| = \|x_\parallel\| \leq \|x_\parallel + x_\perp\| = \|x\|$,
c) $(y|P_F x) = (y_\parallel + y_\perp | P_F(x_\parallel + x_\perp)) = (y_\parallel + y_\perp | x_\parallel) = (y_\parallel | x_\parallel) =$
$= (y_\parallel | x_\parallel + x_\perp) = (P_F(y_\parallel + y_\perp) | x_\parallel + x_\perp) = (P_F y | x)$. □

Proposition 5.6.2 *If* $P \in B(H)$ *and* $P^2 = P = P^*$, *then there exists a closed subspace* $F \subset H$ *such that* $P = P_F$.

Proof
We define a closed subspace in H:

$$F \stackrel{\text{df}}{=} \{x \in H,\ Px = x\}.$$

For every $x \in H$ we have $x = Px + (I - P)x$, where I is the identity operator. We show that $(I - P)x \in F^{\perp}$:
Let $f \in F$ then

$$(f|(I-P)x) = ((I-P)^* f|x) = ((I-P)f|x) = (f - Pf|x) = (0|x) = 0.$$

□

Proposition 5.6.3 *If $F, G, F + G$ are projection operators, then $FG = GF = 0$.*

Proof

$$F + G = (F+G)^2 = F^2 + FG + GF + G^2 = F + G + FG + FG,$$

which gives that $FG + GF = 0$. Multiplying this by F we obtain:

$$0 = F(FG + GF)F = FGF + FGF = 2FGF = 2FG^2F = 2(GF)^* GF.$$

Because $\|A^*A\| = \|A\|^2$, we conclude that $FG = 0$. □

Proposition 5.6.4 *Let $E_1, ..., E_n \in B(H)$ be a family of projection operators such that $E_i \neq 0$, $E_iE_j = 0$ for $i \neq j$ and $\sum_{i=1}^n E_i = I$. Then for $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$*

$$\|\sum_{i=1}^n \lambda_i E_i\| = \max\{|\lambda_1|, \ldots, |\lambda_n|\}.$$

Proof
For all $x \in H$, $x = \sum_{i=1}^n E_i x$ and every component of this sum is orthogonal to each other. Moreover, if $\|x\| \leq 1$ then $\sum_{i=1}^n \|E_i x\| \leq 1$. Now using the notation $a_i \stackrel{\text{df}}{=} \|E_i x\|^2$ we may calculate:

$$\|\sum_{i=1}^n \lambda_i E_i\|^2 = \sup_{\|x\| \leq 1} \|\sum_{i=1}^n \lambda_i E_i x\|^2 = \sup_{\|x\| \leq 1} \sum_{i=1}^n |\lambda_i|^2 \|E_i x\|^2 =$$

$$= \sup_{a_i \geq 0, \sum a_i \leq 1} \sum_{i=1}^n |\lambda_i|^2 a_i = \max\{|\lambda_1|^2, \ldots, |\lambda_n|^2\}.$$

□

5.7 Isometric and unitary operators

Definition 5.7.1 *A linear operator $U : H_1 \to H_2$ is called isometric if*

$$\forall x, y \in H_1 \quad (x|y)_1 = (Ux|Uy)_2,$$

if additionally U is "onto," then we call it a unitary operator.

Proposition 5.7.1 *For every unitary operator $U : H_1 \to H_2$ we have:*
a) $U \in B(H_1, H_2)$ and $\|U\| = 1$;
b) $U^*U = I_1$ – identity operator in H_1;
c) $UU^* = I_2$ – identity operator in H_2.

□

Proposition 5.7.2 *If $U \in B(H_1, H_2)$, $U^*U = I_1$ and $UU^* = I_2$ then U is unitary.*

Proof
Obviously U is an isometric operator. To show that for all $z \in H_2$ there exists $x \in H_1$ such that $Ux = z$ we put $x = U^*z$. □

Corollary 5.7.1 *Operator U is unitary if and only if $U^* = U^{-1}$.*

□

6 Unbounded operators

6.1

The operator $Af \stackrel{\text{df}}{=} \frac{df}{dt}$ acting in $L^2([0,1], d\lambda)$ is not an operator in the sense of the Definition 5.1.1, therefore we need more general definition:

Definition 6.1.1 *A linear map T from H_1 to H_2 defined on a linear subspace $D_T \subset H_1$ ($T : H_1 \supset D_T \to H_2$) is called an operator. D_T is called a domain of the operator T.*

Definition 6.1.2 *If T, S are operators from H_1 to H_2 such that $D_T \subset D_S$ and for all $x \in D_T$ $Tx = Sx$, then we say that operator S is the extension of T and we write $T \subset S$.*

6.2

Definition 6.2.1 *An operator T is closed if for all sequences $\{x_n\} \subset D_T$ we have*

$$(x_n \to x, \ \ Tx_n \to y) => (x \in D_T, \ \ Tx = y) \, .$$

Definition 6.2.2 *An operator T is closable if for all sequences $\{x_n\} \subset D_T$ we have*

$$(x_n \to 0 \ \ Tx_n \to y) => (y = 0).$$

From every closable operator we can obtain a closed operator:

Definition 6.2.3 *If T is a closable operator from H_1 to H_2, then we define:*

$$D_{\overline{T}} \stackrel{\text{df}}{=} \{x \in H_1 : \exists \{x_n\} \subset D_T, \{x_n\} \ \& \ \{Tx_n\} \text{ are convergent}\} \, ,$$

$$\overline{T}x \stackrel{\text{df}}{=} \lim_{n \to \infty} Tx_n .$$

$\overline{T}$ is called a closure of T. Of course $T \subset \overline{T}$.

6.3

Definition 6.3.1 *For every operator T from H_1 to H_2 we define a subspace*

$$G_T \stackrel{\text{df}}{=} \{(x, Tx) : x \in D_T\} \subset H_1 \oplus H_2,$$

which is called a graph of the operator T.

Proposition 6.3.1 *A subset $G \subset H_1 \oplus H_2$ is a graph of some operator T if and only if G is a vector subspace and $G \cap (H_1 \oplus \{0\}) = \{0\}$.*

Sketch of the proof

$$D_T \stackrel{\text{df}}{=} \{x \in H_1 : \exists y \in H_2 \ \ (x,y) \in G\},$$

$$Tx \stackrel{\text{df}}{=} y.$$

□

Proposition 6.3.2 *T is closed if and only if G_T is closed.*

□

Proposition 6.3.3 *T is closable if and only if $\overline{G_T}$ (the closure of G_T) is a graph of some operator.*

Proof
=>

$$T \subset \overline{T} <=> G_T \subset G_{\overline{T}} => \overline{G_T} \subset G_{\overline{T}} => \overline{G_T} \cap (H_1 \oplus \{0\}) = \{0\}.$$

<=
Let as consider a sequence $\{x_n\} \subset D_T$ such that $x_n \to 0$ and $Tx_n \to y$, then $(0,y) \in \overline{G_T}$, hence $y = 0$ which means that T is closable. □

6.4

Analogically to bounded operators we want to introduce the notion of an adjoint operator. For the operator $T : H_1 \supset D_T \to H_2$ we want to define an operator T^* which satisfies conditions:

$$T^* : H_2 \supset D_{T^*} \to H_1,$$

$$\forall x \in D_T \ \forall y \in D_{T^*} \quad (x|T^*y)_1 = (Tx|y)_2. \tag{4}$$

Since these two conditions are satisfied by $T^* = 0$, $D_{T^*} = \{0\}$, we need some kind of expansion.

Definition 6.4.1 *Let $T : H_1 \supset D_T \to H_2$ and D_T be dense in H_1 i.e. $\overline{D_T} = H_1$. Then the operator T^* adjoint to T may be defined by:*

$$D_{T^*} \stackrel{\text{df}}{=} \{y \in H_2 : \exists x \in H_1 \ \forall x' \in D_T \ (x'|x)_1 = (Tx'|y)_2 \ \},$$

$$T^*y \stackrel{\text{df}}{=} x.$$

Remark 6.4.1 *If D_T is not dense, then the equations $(x'|x)_1 = (Tx'|y)_2$ do not determine x unique, so T^* cannot exists.*

From now on we shall consider only operators with dense domains.

Definition 6.4.2 *Let $V : H_1 \oplus H_2 \to H_2 \oplus H_1$ be given by:*

$$V(x,y) \stackrel{\text{df}}{=} (-y, x).$$

Remark 6.4.2 *V is a unitary operator.*

Proposition 6.4.1 *The condition (4) is equivalent to the inclusion:*

$$G_{T^*} \subset VG_T^{\perp}.$$

Proof
It is enough to show that $V^{-1}G_{T^*}$ is orthogonal to G_T:
For $x \in D_T$ and $y \in D_{T^*}$ we have $V^{-1}(x, T^*y) = (T^*y, -x)$. Thus, the condition of orthogonality has the form $(x|T^*y)_1 + (Tx|-y)_2 = 0$, which is exactly (4). □

Proposition 6.4.2 *Let T be an operator. D_T is dense in H_1 if and only if $VG_T^\perp$ is a graph.*

Proof
$(0,x) \in VG_T^\perp \quad <=> \quad (x,0) \in G_T^\perp \quad <=> \quad x \perp D_T \quad => \quad x = 0$ because D_T is dense. □

Corollary 6.4.1 *We have an equivalent definition of the adjoint operator T^*:*
T^* is represented by the graph $VG_T^\perp$.

□

Corollary 6.4.2 *For every operator T such that $\overline{D_T} = H_1$ the operator T^* is closed.*

□

Proposition 6.4.3 *a) If the operator T is closable, then D_{T^*} is dense;*
b) if D_{T^*} is dense, then T is closable and $T^{**} = \overline{T}$.

Proof
a) Let $y \perp D_{T^*}$, so $(y,0) \perp G_{T^*}$. Because V is unitary,

$$(y,0) \in G_{T^*}^\perp = VG_T^{\perp\perp} = V\overline{G_T}, \tag{5}$$

therefore $(0,y) \in \overline{G_T}$. Because T is closable, then $\overline{G_T}$ is a graph, hence $y = 0$.
b) from the formula (5) we obtain

$$(0,y) \in \overline{G_T} <=> y \perp D_{T^*}.$$

Since D_{T^*} is dense, $y = 0$. Thus, $\overline{G_T}$ is a graph.

$$G_{T^{**}} = V^{-1}G_{T^*}^\perp = V^{-1}(VG_T^\perp)^\perp = V^{-1}VG_T^{\perp\perp} = \overline{G_T} = G_{\overline{T}}.$$

□

Corollary 6.4.3 *If the operator T is closed, then $T = T^{**}$.*

□

Definition 6.4.3 *Let $H_1 = H_2$. Then:*
a) if $T \subset T^*$, T is called symmetric;
b) if $T = T^*$, T is called self-adjoint;
c) if $\overline{T} = T^*$, T is called essentially self-adjoint.

Corollary 6.4.4 *T is symmetric if and only if for all $x, y \in D_T$*

$$(Tx|y) = (x|Ty).$$

□

6.5

The following theorem shows why the domain of an unbounded operator cannot be an entire Hilbert space:

Theorem 6.5.1 *If T is a closed operator from H_1 to H_2 and $D_T = H_1$ then T is bounded.*

Proof
Let P_i, $i = 1, 2$, denote the orthogonal projection in $H_1 \oplus H_2$ onto the i-th component. Of course $G_T = \{(x, Tx) : \ x \in H_1\}$ is a Hilbert space and $P_1 \mid_{G_T}: G_T \to H_1$ is an isomorphism of Hilbert spaces. Thus

$$x \overset{(P_1|_{G_T})^{-1}}{\longmapsto} (x, Tx) \overset{P_2}{\longmapsto} Tx$$

being the superposition of continuous maps, is continuous. □

6.6

We denote by $\mathcal{C}(H_1, H_2)$ a set of all closed operators with the dense domain. If $H_1 = H_2 = H$, we write $\mathcal{C}(H)$.
Note that $\mathcal{C}(H_1, H_2)$ is not a vector space, because, for example, $T + (-T) \neq 0$, but $T + (-T) = 0 \mid_{D_T}$.

Theorem 6.6.1 *If $T \in \mathcal{C}(H)$, then T^*T is self-adjoint.*

Proof
Since T is closed, G_T is a Hilbert subspace of $H \oplus H$. So we have the orthogonal decomposition:

$$H \oplus H = G_T \oplus G_T^{\perp} = G_T \oplus V^{-1} G_{T^*}. \tag{6}$$

Thus, for every $(x, 0) \in H \oplus H$ there exists $u \in D_T$ and $v \in D_{T^*}$ such that

$$(x, 0) = (u, Tu) + V^{-1}(v, T^*v) = (u - T^*v, Tu + v)$$

and therefore

$$x = u - T^*v \quad \text{and} \quad v = -Tu,$$

hence $Tu \in D_{T^*}$ and $u \in D_{T^*T}$, which allows us to write $-T^*v = T^*Tu$. So we proved that

$$\forall x \in H \ \ \exists u \in D_{T^*T} \quad x = u + T^*Tu. \tag{7}$$

Now we shall show that D_{T^*T} is dense in H:
Let $x \perp D_{T^*T}$. Then, by (7), we have

$$0 = (x|u) = (u|u) + (T^*Tu|u) = (u|u) + (Tu|Tu) \geq 0,$$

which implies $u = 0$, therefore $x = 0$.
It is easy to show that T^*T is symmetric:

$$(x|T^*Ty) = (Tx|Ty) = (T^*Tx|y).$$

Finally, we shall show that $D_{(T^*T)^*} \subset D_{T^*T}$:
Let $x \in D_{(T^*T)^*}$ and $y \overset{\text{df}}{=} (T^*T)^*x$. Then for each $h \in D_{T^*T}$

$$(y|h) = ((T^*T)^*x|h) = (x|T^*Th). \tag{8}$$

From (7) we know that there exists $u \in D_{T^*T}$ such that $x + y = u + T^*Tu$.

$$(x - u|h + T^*Th) = (x|h) - (u|h) + (x|T^*Th) - (u|T^*Th) =$$

$$= (x|h) - (u|h) + (y|h) - (T^*Tu|h) = (x + y|h) - (u + T^*Tu|h) = 0,$$

but $h + T^*Th$ runs over whole H (formula (7)) so $x - u = 0$. This means that $x \in D_{T^*T}$. □

7 Spectral measures and integrals

7.1 σ–algebras

Let X be any set and $\mathcal{B}$ a family of subsets of X.

Definition 7.1.1 *$\mathcal{B} \neq \emptyset$ is called a σ–algebra of sets if:*
a) if $A \in \mathcal{B}$, then $X \setminus A \in \mathcal{B}$;
b) if $A_i \in \mathcal{B}$, $i \in \mathbf{N}$, then $\cup_i A_i \in \mathcal{B}$.

Definition 7.1.2 *The σ–algebra of subsets generated by $\mathcal{B}_0$ is the smallest σ–algebra of sets which contains $\mathcal{B}_0$.*

Definition 7.1.3 *Let X be a topological space. Then*
a) the σ–algebra of Borel sets is the σ–algebra generated by the family of all open sets;
b) the σ–algebra of Baire sets is the σ–algebra generated by the family of all compact sets of the G_δ type (countable intersection of open sets).

Remark 7.1.1 *a) If X is a metric space, then every compact set is of the G_δ type;*
b) (σ–algebra of Borel sets) $\subset$ (σ–algebra of Baire sets);
c) if $X = \cup_{n=1}^{\infty} K_n$, where K_n is compact, then
(σ–algebra of Borel sets) $=$ (σ–algebra of Baire sets).

□

7.2 Measures

Definition 7.2.1 *Let $(X, \mathcal{B})$ be a set with a σ–algebra of sets. A measure is any map*

$$\mu : \mathcal{B} \longrightarrow \overline{\mathbf{R}_+} \stackrel{\text{df}}{=} [0, \infty],$$

which satisfies the condition:

$$(A_i \in \mathcal{B},\ i \in \mathbf{N}\ A_i \cap A_j = \emptyset,\ i \neq j) => \left(\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i) \right).$$

If additionally $\mu(X) = 1$, then μ is called a probabilistic measure.

Definition 7.2.2 *Let as consider $(X, \mathcal{B})$. The function $f : X \to \mathbf{R}$ is called $\mathcal{B}$–measurable if*

$$\forall c \in \mathbf{R} \quad f^{-1}(] - \infty, c]) \in \mathcal{B}.$$

Proposition 7.2.1 *The measurable functions have the following properties:*
a) they form an algebra (with ordinary pointwise addition and multiplication);
b) if $\{f_n\}$ is a sequence of measurable functions and for all $x \in X$, $f_n(x) \to f(x)$, then f is measurable.

□

Definition 7.2.3 *$Baire(X)$ is, by definition, the set of all functions which are measurable with respect to the σ–algebra of Baire sets.*

Proposition 7.2.2 *$C_0(X) \subset Baire(X)$, or, more precisely $Baire(X)$ is the smallest family of functions which contains $C_0(X)$ and is closed with respect to pointwise limits.*

□

7.3 Integrals

Let as consider a set X with a σ–algebra of sets $\mathcal{B}$ and a measure μ.

Definition 7.3.1 *Let χ_A be the characteristic function of the set $A \in \mathcal{B}$. If $\mu(A) < \infty$, then we define:*

$$\int_X \chi_A(x)\mu(dx) \stackrel{\mathrm{df}}{=} \mu(A).$$

Definition 7.3.2 *A function f is called a $\mathcal{B}$–measurable stair function if $\mu(\{x : f(x) \neq 0\}) < \infty$ and the value set of f is finite.*

We see that f is a linear combination of characteristic functions:

$$f = \sum_{i=1}^{n} \alpha_i \chi_{A_i},$$

thus, we can define the integral

$$\int_X f(x)\mu(dx) \stackrel{\mathrm{df}}{=} \sum_{i=1}^{n} \alpha_i \mu(A_i).$$

The set of all $\mathcal{B}$–measurable stair functions forms a vector structure and the integral is a positive functional on this space, so we can apply the Stone procedure to expand the integral on $\mathcal{L}^p(\mu)$ and to construct $L^p(\mu)$ etc. Let as recall the following theorem from the general theory of the integral:

Theorem 7.3.1 *Let X be a locally compact topological space and let $\phi : C_0(X) \to \mathbf{R}$ be a linear map which preserves positivity (i.e. if for all $x \in X$ $f(x) \geq 0$ then $\phi(f) \geq 0$). Then there exists a measure μ on the σ–algebra of Baire sets such that*

$$\phi(f) = \int_X f(x)\mu(dx).$$

□

Below we shall consider measures not necessarily positive i.e. the measures of the form $\mu_1 - \mu_2$ or $\mu_1 + i\mu_2$ (complex), where μ_1, μ_2 are positive.

7.4 Spectral measures

Definition 7.4.1 *Let $(X, \mathcal{B})$ be a set with a σ–algebra of sets and let $B(H)$ be the space of all bounded operators in a Hilbert space H. A spectral measure is any map $E : \mathcal{B} \to B(H)$ which satisfies the following conditions:*
a) for all $A \in \mathcal{B}$, $E(A)$ is a projection operator;
b) if $A_i \in \mathcal{B}$, $A_i \cap A_j = \emptyset$, $i, j \in \mathbf{N}$, $i \neq j$, $x \in H$, then

$$E\left(\cup_{i=1}^{\infty} A_i\right) x = \sum_{i=1}^{\infty} E(A_i)x;$$

c) $E(X) = I$ is the identity operator.

Remark 7.4.1 *Note that in the condition b) we have a strong convergence, not an operator convergence.*

Proposition 7.4.1 *Let E be a spectral measure and let $A, B \in \mathcal{B}$. Then*

$$E(A)E(B) = E(B)E(A) = E(A \cap B).$$

Proof

Let $A' = A - B$, $B' = B - A$ and $C' = A \cap B$. We have

$$E(A) = E(A') + E(C'), \quad E(B) = E(B') + E(C').$$

Now, by the Proposition 5.6.3 we can calculate:

$$E(A)E(B) = E^2(C') = E(C') = E(A \cap B).$$

□

Remark 7.4.2 *If $E : \mathcal{B} \to B(H)$ is a spectral measure and $x \in H$, then the map $\mathcal{B} \ni A \mapsto (x|E(A)x) \in \overline{\mathbf{R}_+}$ is a measure, and we denote it by $(x|Ex)$.*

7.5 Spectral integrals

Definition 7.5.1 *Let $(X, \mathcal{B}, E)$ be a set with a σ–algebra of sets and a spectral measure. For the given stair function $f(x) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(x)$, we define the spectral integral*

$$\int_X f(x)E(dx) \stackrel{\text{df}}{=} \sum_{i=1}^{n} \alpha_i E(A_i).$$

Proposition 7.5.1 *The spectral integral has the following properties:*
a) if $f \equiv 1$ then $\int_X f(x)E(dx) = I$;
b) for all $\alpha, \beta \in \mathbf{C}$ and f_1, f_2 being stair functions

$$\int_X (\alpha f_1 + \beta f_2)(x)E(dx) = \alpha \int_X f_1(x)E(dx) + \beta \int_X f_2(x)E(dx);$$

c) for all f_1, f_2 being stair functions

$$\int_X f_1(x)f_2(x)E(dx) = \left(\int_X f_1(x)E(dx)\right)\left(\int_X f_2(x)E(dx)\right);$$

d) $\int_X \overline{f(x)}E(dx) = \left(\int_X f(x)E(dx)\right)^*$;
e) $\| \int_X f(x)E(dx)\| \leq \sup_{x \in X} |f(x)|$.

Proof

b) Taking into account the definition and the fact that the integral does not depend on the representation of the stair function, i.e.

$$\left(\sum_{i=1}^{n} \alpha_i \chi_{A_i} = \sum_{j=1}^{m} \beta_j \chi_{B_j}\right) => \left(\sum_{i=1}^{n} \alpha_i E(A_i) = \sum_{j=1}^{m} \beta_j E(B_j)\right),$$

the proof is obvious. c) For $f_1 = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$ and $f_2 = \sum_{j=1}^{m} \beta_i \chi_{B_j}$ we have

$$f_1 f_2 = \sum_{i,j} \alpha_i \beta_j \chi_{A_i} \chi_{B_j} = \sum_{i,j} \alpha_i \beta_j \chi_{A_i \cap B_j},$$

which gives

$$\int_X f_1(x)f_2(x)E(dx) = \sum_{i,j} \alpha_i \beta_j E(A_i \cap B_j) = \sum_{i,j} \alpha_i \beta_j E(A_i)E(B_j).$$

e) The proof follows directly from the Proposition 5.6.4. □

Corollary 7.5.1 *If $\{f_n\}$ is a sequence of stair functions which is uniformly convergent to the function f then it follows from e) that the sequence $\int_X f_n(x)E(dx)$ is convergent and its limit depends only on f.*

□

Definition 7.5.2 *If $\{f_n\}$ is a sequence of stair functions which is uniformly convergent to the function f, then we define*

$$\int_X f(x)E(dx) \stackrel{\text{df}}{=} \lim_{n\to\infty} \int_X f_n(x)E(dx).$$

Proposition 7.5.2 *Every bounded $\mathcal{B}$–measurable function f may be conformly aproximated with stair functions.*

Proof
Since f is bounded we may cover the set $f(X)$ by $\cup_{i=1}^n \Delta_i$, where $|\Delta_i| < \varepsilon$. Now, defining:

$$A_i \stackrel{\text{df}}{=} \{x \in X : f(x) \in \Delta_i\},$$

and

$$f_\varepsilon \stackrel{\text{df}}{=} \sum_{i=1}^n \chi_{A_i},$$

we have

$$\sup_{x\in X} |f(x) - f_\varepsilon(x)| \leq \varepsilon.$$

□

Theorem 7.5.1 *Let X be a compact space and let $\phi : C(X) \to B(H)$ be a map which satisfies the following conditions:*
a) ϕ is linear;
b) ϕ is multiplicative i.e. $\phi(f_1 f_2) = \phi(f_1)\phi(f_2)$;
c) ϕ is symmetric i.e. $\phi(\bar{f}) = (\phi(f))^*$;
d) $\phi(1) = I$.
Then there exists a spectral measure E on the σ–algebra of Baire sets such that for all $f \in C(X)$

$$\phi(f) = \int_X f(x)E(dx).$$

Proof
Step 1.
Let $M \stackrel{\text{df}}{=} \sup_{x\in X} |f(x)|$. Then $M^2 - \bar{f}f \geq 0$, so there exists a nonnegative function g such that

$$M^2 = \bar{f}f + \bar{g}g.$$

Acting on this equality by ϕ, we obtain

$$M^2 I = \phi(f)^*\phi(f) + \phi(g)^*\phi(g)$$

and taking the average value over $v \in H$ we have

$$M^2(v|v) = (v|\phi(f)^*\phi(f)v) + (v|\phi(g)^*\phi(g)v) =$$

$$= (\phi(f)v|\phi(f)v) + (\phi(g)v|\phi(g)v) = \|\phi(f)v\|^2 + \|\phi(g)v\|^2.$$

Finally we obtain

$$\|\phi(f)\| \leq \sup_{x \in X} |f(x)|.$$

Step 2.
If $f \geq 0$ then $\phi(f) \geq 0$. Since because $f = \bar{f}^{\frac{1}{2}} f^{\frac{1}{2}}$, $\phi(f) = \phi(f^{\frac{1}{2}})^* \phi(f^{\frac{1}{2}})$.
Step 3.
For each $u \in H$ we define the map

$$C(X) \ni f \mapsto (u|\phi(f)u) \in \mathbf{C}.$$

This map assigns positive numbers to positive functions, so by the Theorem 7.3.1 there exists a measure μ_u such that

$$(u|\phi(f)u) = \int_X f(x)\mu_u(dx).$$

Using the formula

$$(u|Av) = \frac{1}{4}\sum_{k=0}^{3} i^k (i^k u + v|A(i^k u + v)),$$

we obtain a complex measure $\mu_{u,v}$, which satisfies

$$(u|\phi(f)v) = \int_X f(x)\mu_{u,v}(dx).$$

Step 4.
Now we may extend ϕ on a whole family of bounded Baire functions, denoted by $Baire_b(X)$. Because X is compact, $1 \equiv f \in L^1(\mu_{u,v})$, so for all $f \in Baire_b(X)$ we define

$$(u|\tilde{\phi}(f)v) \stackrel{\text{df}}{=} \int_X f(x)\mu_{u,v}(dx).$$

As an exercise we propose to prove that this definition is good i.e.
i) the right-hand side of this definition is antilinear in u and continous in v,
ii) $\tilde{\phi}$ fulfil the conditions a)–d). For example, b) follows from the Lebesgue theorem on majority convergence.
Step 5.
We define the map

$$\mathcal{B} \ni A \mapsto E(A) \stackrel{\text{df}}{=} \tilde{\phi}(\chi_A) \in B(H).$$

We prove that this is a spectral measure:
i) $E(A)$ is a projection operator:

$$E(A)^2 = \tilde{\phi}(\chi_A)\tilde{\phi}(\chi_A) = \tilde{\phi}(\chi_A^2) = \tilde{\phi}(\chi_A) = E(A),$$

$$E(A)^* = \tilde{\phi}(\chi_A)^* = \tilde{\phi}(\bar{\chi_A}) = \tilde{\phi}(\chi_A) = E(A),$$

ii) $E(X) = \tilde{\phi}(1) = \phi(1) = I$,
iii) countable additivity of E in strong topology:
Let $A = \cup_{i=1}^{\infty} A_i$ where $A_i \cap A_j = \emptyset$ for $i \neq j$. For each $u \in H$ we define a sequence $\{w_N\}$ by

$$w_N \stackrel{\text{df}}{=} \sum_{i=1}^{N} E(A_i)u.$$

To show that this sequence is convergent, it should be noted that the series

$$\sum_{i=1}^{\infty} \|E(A_i)u\|^2$$

is convergent. It follows from the fact that partial sums for this series are bounded by $\|u\|^2$. Since the vectors $E(A_i)u$ are orthogonal, we have

$$\|w_N - w_M\|^2 = \|\sum_{i=N+1}^{M} E(A_i)u\|^2 = \sum_{i=N+1}^{M} \|E(A_i)u\|^2,$$

so $\{w_N\}$ is convergent.
Finally, from the fact that for all $x \in X$

$$\chi_A(x) = \lim_{N\to\infty} \sum_{i=1}^{N} \chi_{A_i}(x),$$

we have

$$(u|E(A)v) = \sum_{i=1}^{\infty} (u|E(A_i)v),$$

which gives

$$\lim_{N\to\infty} w_N = \sum_{i=1}^{\infty} E(A_i)u = E(\cup_{i=1}^{\infty} A_i)u.$$

Step 6.
Finally, we have that for every stair function f

$$\tilde{\phi}(f) = \int_X f(x)E(dx) \tag{9}$$

and, by the continuity, this is true for every $f \in Baire_b(X)$. □

Definition 7.5.3 (Spectral integral for arbitrary Baire function.) *Let $f \in Baire(X)$ and $\{f_n\}$ be a sequence in $Baire_b(X)$ such that for all $x \in X$*

$$|f_n(x)| \le |f(x)| \quad f_n(x) \to f(x).$$

We define an unbounded operator by

$$\int_X f(x)E(dx)u \stackrel{\mathrm{df}}{=} \lim_{n\to\infty} \int_X f_n(x)E(dx)u,$$

with a domain consisting of those $u \in H$ for which there exists the limit on the right side of this definition.

7.6

Directly from the definitions of spectral measure and integral we obtain:

Proposition 7.6.1 *If $\phi : X_1 \to X_2$ is a measurable mapping and E is a spectral measure on X_1, then*

$$F(\Delta) \stackrel{\mathrm{df}}{=} E(\phi^{-1}(\Delta))$$

is spectral measure on X_2 and

$$\int_{X_2} f(x_2)F(dx_2) = \int_{X_1} f(\phi(x_1))E(dx_1).$$

□

8 Commutative C^*–algebras

8.1

Definition 8.1.1 *$\mathcal{A}$ is a C^*–algebra if:*
a) $\mathcal{A}$ is a vector space over $\mathbf{C}$*;*
b) there is in $\mathcal{A}$ a bilinear and associative multiplication

$$\mathcal{A} \times \mathcal{A} \ni (a, b) \mapsto ab \in \mathcal{A},$$

(if additionally there exists $1 \in \mathcal{A}$ such that for all $a \in \mathcal{A}$ $1a = a1 = a$, we call $\mathcal{A}$ the algebra with identity);
c) there exists an antylinear map $\mathcal{A} \ni a \mapsto a^* \in \mathcal{A}$ such that for all $a, b \in \mathcal{A}$

$$a^{**} = a, \quad (ab)^* = b^* a^*,$$

(it follows from this that $1^* = 1$);
d) $\mathcal{A}$ is a normed space such that for all $a, b \in \mathcal{A}$

$$\|ab\| \leq \|a\|\|b\|, \quad \|a^*\| = \|a\|;$$

e) $\mathcal{A}$ is complete in this norm;
f) for all $a \in \mathcal{A}$

$$\|a^* a\| = \|a\|^2;$$

g) if additionally for all $a \in \mathcal{A}$, $ab = ba$, then $\mathcal{A}$ is called a commutative C^*–algebra.

Definition 8.1.2 *Element $a \in \mathcal{A}$ is called invertible if there exists $a^{-1} \in \mathcal{A}$, such that $aa^{-1} = a^{-1}a = 1$. In opposite case, a is called noninvertible.*

8.2 Examples

a) the set of bounded operators in a Hilbert space is a C^*–algebra;
b) if $A \in H$ and $A^* = A$, then the closure of the space of all polinomials of A is a commutative C^*–algebra;
c) if Λ is a compact topological space, then $C(\Lambda)$ with

$$f^* \stackrel{\mathrm{df}}{=} \bar{f}, \quad \|f\| \stackrel{\mathrm{df}}{=} \sup_{\lambda \in \Lambda} |f(\lambda)|,$$

is a commutative C^*–algebra. It turns aut that every commutative C^*–algebra is of this form, more precisly, we have the following theorem which will be proved in this section:

Theorem 8.2.1 (Gelfand) *Let $\mathcal{A}$ be a commutative C^*–algebra. There exists a compact space Λ and a map $\mathcal{A} \ni a \mapsto \hat{a} \in C(\Lambda)$ such that for all $a, b \in \mathcal{A}$, $r, s \in \mathbf{C}$ and $\lambda \in \Lambda$:*
a) $\widehat{(ra + sb)}(\lambda) = r\hat{a}(\lambda) + s\hat{b}(\lambda)$;
b) $\widehat{ab}(\lambda) = \hat{a}(\lambda)\hat{b}(\lambda)$;
c) $\widehat{a^*}(\lambda) = \overline{\hat{a}(\lambda)}$;
d) $\|a\| = \sup_{\lambda \in \Lambda} |\hat{a}(\lambda)|$;
e) $\{\hat{a} : a \in \mathcal{A}\} = C(\Lambda)$.
Shortly, we have an isomorphism of C^*–algebras.

Definition 8.2.1 *A character of the C^*–algebra $\mathcal{A}$ is any map $\mathcal{A} \ni a \mapsto \chi(a) \in \mathbf{C}$ such that for all $a, b \in \mathcal{A}$ and $r, s \in \mathbf{C}$:*
a) $\chi(ra + sb) = r\chi(a) + s\chi(b)$;
b) $\chi(ab) = \chi(a)\chi(b)$;
c) $\chi(a^*) = \overline{\chi(a)}$;
d) $\chi(1) = 1$.

Remark 8.2.1 *The notion of character is important because we shall define* Λ *as the space of all characters of* $\mathcal{A}$ *equipped in some topology and we shall put:*

$$\hat{a}(\chi) \stackrel{\text{df}}{=} \chi(a).$$

From now on $\mathcal{A}$ will denote a commutative C^*–algebra with identity.

8.3 Ideals

Definition 8.3.1 *A subset* $\mathcal{I} \subset \mathcal{A}$ *is called an ideal if:*
a) $\mathcal{I}$ *is a vector subspace in* $\mathcal{A}$*;*
b) for all $a \in \mathcal{I}$ *and* $b \in \mathcal{A}$ *we have* $ab \in \mathcal{I}$*;*
c) $1 \notin \mathcal{I}$ *(i.e.* $\mathcal{I} \neq \mathcal{A}$*).*

Corollary 8.3.1 *If* $\mathcal{I}$ *is an ideal, then* $\mathcal{A}/\mathcal{I}$ *is an algebra, where*

$$(a + \mathcal{I}) + (b + \mathcal{I}) \stackrel{\text{df}}{=} a + b + \mathcal{I},$$

$$(a + \mathcal{I})(b + \mathcal{I}) \stackrel{\text{df}}{=} ab + \mathcal{I}.$$

□

Proposition 8.3.1 *Every element of each ideal is noninvertible and every noninvertible element in* $\mathcal{A}$ *is in some ideal.*

Proof
Let $a, a^{-1} \in \mathcal{I}$, then $1 \in \mathcal{I}$, which is a contradiction.
For noninvertible $a \in \mathcal{A}$ we define the ideal

$$\mathcal{I} \stackrel{\text{df}}{=} \{ab : \ b \in \mathcal{A}\}.$$

□

8.4 Maximal ideals

Definition 8.4.1 $\mathcal{M}$ *is called a maximal ideal if:*
a) $\mathcal{M}$ *is an ideal;*
b) if there exists an ideal $\mathcal{I}$*, such that* $\mathcal{M} \subset \mathcal{I}$*, then* $\mathcal{M} = \mathcal{I}$*.*

Proposition 8.4.1 *Each ideal is included in some maximal ideal.*

Proof
Follows from the Kuratowski – Zorn lema. □

Corollary 8.4.1 *If* $a \in \mathcal{A}$ *is noninvertible, then there exists a maximal ideal* $\mathcal{M}$ *such that* $a \in \mathcal{M}$*.*

□

Proposition 8.4.2 *If* $\mathcal{M}$ *is a maximal ideal, then* $\mathcal{A}/\mathcal{M} \simeq \mathbf{C}$*.*

Proof
Follows directly from the Gelfand–Mazur theorem, which will be proved later. □

Corollary 8.4.2 *The map* $\mathcal{A} \ni a \mapsto a + \mathcal{M} \in \mathcal{A}/\mathcal{M} \simeq \mathbf{C}$ *is a character.*

□

8.5

Proposition 8.5.1 *If $\|1-a\|<1$, then a is invertible.*

Proof
Let as consider the series $\sum_{n=0}^{\infty}(1-a)^n$. It is convergent, so by putting $a=1-(1-a)$ we see that this series is the inversion of a. □

Proposition 8.5.2 *The closure of an ideal is an ideal too.*

Proof
Let as consider an ideal $\mathcal{I}$, then its closure $\bar{\mathcal{I}}$ is a vector subspace in $\mathcal{A}$.
Let $a=\lim a_n \in \bar{\mathcal{I}}$, $a_n \in \mathcal{I}$, then for all $b \in \mathcal{A}$ we have $ba = \lim ba_n \in \bar{\mathcal{I}}$.
$1 \notin \bar{\mathcal{I}}$ follows from the Proposition 8.5.1. □

Corollary 8.5.1 *Every maximal ideal is closed.*

□

8.6

Definition 8.6.1 *a) The spectrum of an element $a \in \mathcal{A}$ is the set*

$$Sp\, a \overset{\text{df}}{=} \{\lambda \in \mathbf{C} : a-\lambda 1 \text{ is noninvertible}\};$$

b) the resolvent set of a:

$$\varrho(a) \overset{\text{df}}{=} \mathbf{C} \setminus Sp\, a.$$

Proposition 8.6.1 *a) If $\lambda \in \mathbf{C}$ is such that $|\lambda| > \|a\|$, then $\lambda \in \varrho(a)$;*
b) $\varrho(a)$ is open in $\mathbf{C}$;
c) the map

$$\varrho(a) \ni \lambda \mapsto R(a,\lambda) \overset{\text{df}}{=} (a-\lambda 1)^{-1} \in \mathcal{A}$$

is holomorphic. $R(a,\lambda)$ is called the resolvent of a;
d) $Sp\, a \neq \emptyset$.

Proof
a) Since $|\lambda| > \|a\|$, the series

$$-\frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{a^n}{\lambda^n} \tag{10}$$

is convergent. Moreover

$$(a-\lambda 1)\left(-\frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{a^n}{\lambda^n}\right) = -\sum_{n-0}^{\infty}\frac{a^{n+1}}{\lambda^{n+1}} + \sum_{n=0}^{\infty}\frac{a^n}{\lambda^n} = 1.$$

Thus we have

$$R(a,\lambda) = (a-\lambda 1)^{-1} = -\frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{a^n}{\lambda^n}.$$

b) Let $\lambda_0 \in \varrho(a)$, e.i. $(a-\lambda_0 1)^{-1}$ exists. Let as consider λ from some neighbourhood of λ_0. Then

$$(a-\lambda 1)^{-1} = ((a-\lambda_0 1) + (\lambda_0-\lambda)1)^{-1} =$$

$$= \left\{(a-\lambda_0 1)\left[1-(\lambda-\lambda_0)(a-\lambda_0 1)^{-1}\right]\right\}^{-1} =$$

$$= (a - \lambda_0 1)^{-1} \sum_{n=0}^{\infty} \left[(a - \lambda_0 1)^{-1} (\lambda - \lambda_0) \right]^n .$$

This series is convergent if $|\lambda - \lambda_0| < \|a - \lambda_0 1\|^{-1}$. Finally we see that each point in $\varrho(a)$ has a neighbourhood in this set.
c) The proof is obvious because $R(a, \lambda)$ is given by the series (10).
d) Assuming that $Sp\, a = \emptyset$, we have $R(a, \lambda)$ defined on the whole $\mathbf{C}$ as a holomorphic function, which by (10) satisfies

$$R(a, \lambda) \to 0 \quad \text{if } |\lambda| \to \infty.$$

Thus $R(a, \lambda) \equiv 0$ which is contradiction. □

Corollary 8.6.1 *a) $Sp\, a$ is compact;*
b) we have the following equality

$$\sup_{\lambda \in Sp\, a} |\lambda| = \|a\|.$$

<u>Proof</u>
b) Because the series (10) is convergent, then

$$\lim_{n \to \infty} \sqrt[n]{\frac{\|a^n\|}{|\lambda|^n}} < 1,$$

which implies that

$$\sup_{\lambda \in Sp\, a} |\lambda| = \lim_{n \to \infty} \sqrt[n]{\|a^n\|}$$

is the radius of the minimal circle containing $Sp\, a$.
But $\|a^* a\| = \|a\|^2$, so

$$\|a^2\|^2 = \|(a^*)^2 a^2\| = \|a^* a a^* a\| = \|a^* a\|^2 = \|a\|^4,$$

therefore $\|a^2\| = \|a\|^2$ and

$$\|a^{2^k}\| = \|a\|^{2^k}.$$

□

Another corollary from the Proposition 8.6.1 is the following theorem

Theorem 8.6.1 (Gelfand, Mazur) *If every $a \in \mathcal{A} \setminus \{0\}$ is invertible, then*

$$\mathcal{A} = \{\lambda 1 : \lambda \in \mathbf{C}\}.$$

<u>Proof</u>
For $a \in \mathcal{A} \setminus \{0\}$ there exists $\lambda \in Sp\ a$ such that $a - \lambda 1$ is noninvertible, so $a - \lambda 1 = 0$. □

8.7

Let f be an entire function on $\mathbf{C}$ i.e. $f(z) = \sum_{n=0}^{\infty} c_n z^n$ and this series is convergent for all $z \in \mathbf{C}$. For all $a \in \mathcal{A}$ we define

$$f(a) \overset{\text{df}}{=} \sum_{i=0}^{\infty} c_n a^n.$$

It is easy to show that

$$f(a) = -\frac{1}{2\pi i} \oint f(\zeta)(a - \zeta 1)^{-1} d\zeta,$$

where the integral is taken over the contour including $Sp\, a$, moreover

$$Sp\, f(a) = f(Sp\, a).$$

Proposition 8.7.1 *If $a = a^*$, then $Sp\, a \subset \mathbf{R}$.*

Proof
Let as consider

$$b \stackrel{\text{df}}{=} e^{ia} = \sum_{n=0}^{\infty} \frac{(ia)^n}{n!}.$$

We have $b^* b = e^{-ia} e^{ia} = e^0 = 1$, so $\|b\| = 1$, which gives

$$Sp\, b \subset \{\lambda \in \mathbf{C} : |\lambda| \leq 1\}.$$

But $Sp\, b = e^{iSp\, a}$, so

$$(z \in Sp\, a) => (Imz \geq 0),$$

moreover $Sp\, b^* = e^{-iSp\, a}$, so

$$(z \in Sp\, a) => (Imz \leq 0).$$

□

8.8

Proposition 8.8.1 *If $a \in \mathcal{A}$ and $\mathcal{M}$ is a maximal ideal in $\mathcal{A}$, then there exists exactly one complex number $\lambda_{\mathcal{M}}$ such that*

$$a - \lambda_{\mathcal{M}} 1 \in \mathcal{M}.$$

Proof
Let as consider the C^*–algebra $\mathcal{A}/\mathcal{M}$ with the norm given by

$$\|a + \mathcal{M}\| \stackrel{\text{df}}{=} \inf_{a' \in a + \mathcal{M}} \|a'\|.$$

Because $\mathcal{M}$ is maximal, then $\mathcal{A}/\mathcal{M}$ satisfies the conditions of the Gelfand – Mazur theorem, thus

$$\mathcal{A}/\mathcal{M} \simeq \mathbf{C}.$$

□

Definition 8.8.1 *Let $\mathbf{M}$ denote the set of all maximal ideals in $\mathcal{A}$. We define the map called the Gelfand transformation*

$$\mathcal{A} \ni a \mapsto \hat{a} \in \{\textit{complex functions on } \mathbf{M}\},$$

given by

$$\hat{a}(\mathcal{M}) \stackrel{\text{df}}{=} \lambda_{\mathcal{M}}.$$

Theorem 8.8.1 *For all $a, b \in \mathcal{A}$:*
a) $\widehat{(a + b)} = \hat{a} + \hat{b}$;
b) $\widehat{ab} = \hat{a}\hat{b}$;
c) $\hat{1} = 1$ – constant function on $\mathbf{M}$;
d) $\hat{a}(\mathbf{M}) = Sp\, a$;
e) defining

$$\|\hat{a}\| \stackrel{\text{df}}{=} \sup_{\mathcal{M} \in \mathbf{M}} |\hat{a}(\mathcal{M}|,$$

we have that $\|a\| = \|\hat{a}\|$;
f) $\widehat{a^*} = \bar{\hat{a}}$;
g) $\hat{a}(\mathcal{M}) = 0$ if and only if $a \in \mathcal{M}$.

Proof
b) $ab - \hat{a}(\mathcal{M})\hat{b}(\mathcal{M})1 = (a - \hat{a}(\mathcal{M})1)b + \hat{a}(\mathcal{M})(b - \hat{b}(\mathcal{M})1) \in \mathcal{M}$.
d) if $\lambda \in Sp\, a$, then there exists a maximal ideal $\mathcal{M}$ such that $a - \lambda 1 \in \mathcal{M}$, so $\lambda = \hat{a}(\mathcal{M})$. Thus $Sp\, a \subset \hat{a}(\mathbf{M})$.
Now let $\lambda \in \hat{a}(\mathbf{M})$ which means that there exists a maximal ideal $\mathcal{M}$ such that $\lambda = \hat{a}(\mathcal{M})$, so $a - \lambda 1$ is noninvertible i.e. $\lambda \in Sp\, a$. Thus $\hat{a}(\mathbf{M}) \subset Sp\, a$. e) we have

$$\|\hat{a}\| = \sup_{\mathcal{M}\in\mathbf{M}} |\hat{a}(\mathcal{M})| = \sup_{\lambda \in Sp\, a} |\lambda| = \|a\|.$$

f) the proof follows from the observation that

$$b = b^* \quad => \quad \hat{b}(\mathcal{M}) \in \mathbf{R}$$

and that $a = \frac{1}{2}(a + a^*) + i\frac{1}{2i}(a - a^*)$ and $a^* = \frac{1}{2}(a + a^*) - i\frac{1}{2i}(a - a^*)$. □

8.9 Topology on M

To prove the spectral theorem, it is enough to limit our considerations to the case when $\mathcal{A}$ is generated by a finite number of elements $a_1, ..., a_n \in \mathcal{A}$ i.e. $\mathcal{A}$ is the closure of the space of all polinomials of a_i, $i = 1, \ldots, n$.

Proposition 8.9.1 *If $\mathcal{M}_1, \mathcal{M}_2 \in \mathbf{M}$ and $\hat{a}_i(\mathcal{M}_1) = a_i(\mathcal{M}_2)$, $i = 1, ..., n$, then $\mathcal{M}_1 = \mathcal{M}_2$.*

Proof
Follows from the fact that

$$\mathcal{M} = \{a \in \mathcal{A} : \hat{a}(\mathcal{M}) = 0\}.$$

□

Corollary 8.9.1 *We have the injection*

$$\mathbf{M} \ni \mathcal{M} \mapsto (a_1(\mathcal{M}), \ldots, a_n(\mathcal{M})) \in \mathbf{C}^n. \tag{11}$$

□

Definition 8.9.1 *The topology on $\mathbf{M}$ is defined as a pull-back, by the map (11), of the usual topology from $\mathbf{C}^n$.*

Theorem 8.9.1 *a) $\mathbf{M}$ is compact;*
b) for all $a \in \mathcal{A}$ the function $\hat{a}$ is continuous;
c) each continous function on $\mathbf{M}$ is of the form $\hat{a}$, for some $a \in \mathcal{A}$.

Proof
a) $\mathbf{M}$ is bounded because $|\hat{a}_i(\mathcal{M})| \leq \|a_i\|$.
We shall show that $\mathbf{M}$ is closed:
Let $\{\mathcal{M}_k\}$ be a Couchy sequence in $\mathbf{M}$, so there exists $(\lambda_1, \ldots, \lambda_n) \in \mathbf{C}^n$ such that for all $i = 1, \ldots, n$,

$$\lim_{k\to\infty} \hat{a}_i(\mathcal{M}_k) = \lambda_i.$$

We must show that there exists the maximal ideal $\mathcal{M}_\infty$ such that $\hat{a}_i(\mathcal{M}_\infty) = \lambda_i$ for all i. It follows from the proposition 8.9.1 that for all $a \in \mathcal{A}$ there exists $\lambda_a \in \mathbf{C}$ such that

$$\lambda_a = \lim_{k\to\infty} \hat{a}(\mathcal{M}_k).$$

Thus we have the multiplicative and linear map

$$\mathcal{A} \ni a \mapsto \lambda_a \in \mathbf{C},$$

which leads to the definition:

$$\mathcal{M}_\infty \stackrel{\mathrm{df}}{=} \{a \in \mathcal{A} : \lambda_a = 0\}.$$

It is easy to see that $\mathcal{M}_\infty$ is an ideal. Because *codim* $\mathcal{M}_\infty = 1$, we obtain that $\mathcal{M}_\infty$ is a maximal ideal.
Let $z_i \stackrel{\mathrm{df}}{=} \hat{a}_i(\mathcal{M}_\infty)$, then $a_i - z_i 1 \in \mathcal{M}_\infty$ i.e.

$$0 = \lambda_{a_i - z_i 1} = \lambda_i - z_i.$$

Finally we have

$$\lim_{k\to\infty} \mathcal{M}_k = \mathcal{M}_\infty \in \mathbf{M}.$$

b) All $\hat{a}_i$ are continuous, so all its polinomials and limits of polinomials (because convergence in $\mathcal{A}$ implies uniform convergence of the Gelfand transformations).
c) The proof follows directly from the Stone–Weierstrass theorem. □

Remark 8.9.1 *In this way we completed the proof of the Gelfand theorem 8.2.1.*

9 Spectral theorems

9.1 Spectral theorem for Hermitian operators

Let as assume that $\mathcal{A}$ is a C^*–subalgebra in $B(H)$ generated by a Hermitian operator A.

Theorem 9.1.1 *If $A \in B(H)$ is a Hermitian operator, then there exists a spectral measure E on $Sp\,A$, such that*

$$A = \int_{Sp\,A} \lambda E(d\lambda).$$

Proof
The inverse Gelfnad trasformation is an example of an isomorphism between $C(\mathbf{M})$ and $\mathcal{A}$ which satisfies the conditions of the Theorem 7.5.1. This implies that there exists a spectral measure E' on $\mathbf{M}$ such that for all $B \in \mathcal{A}$ we have

$$B = \int_{\mathbf{M}} \hat{B}(\lambda)E'(d\lambda),$$

where $\hat{B}$ is the Gelfand transformation of B. Moreover, due to the Theorem 8.8.1 d) and the Proposition 8.9.1, we may identify $\mathbf{M}$ with $Sp\,A \subset \mathbf{R}$.
Taking into account the Proposition 7.6.1, we see that there exists a spectral measure E on $Sp\,A$ such that

$$B = \int_{Sp\,A} \hat{B}(\hat{A}_{-1}(\lambda))E(d\lambda), \tag{12}$$

where $\hat{A}_{-1}$ is the inverse map to the Gelfand transformation of the generator of $\mathcal{A}$, i.e. $\hat{A}_{-1} : Sp\,A \to \mathbf{M}$. Putting A in (12) we complete the proof. □

This proof show that we may also formulate the spectral theorem for unitary operators or, more generally for normal operators. If an operator A is not normal, then we cannot imbed $\mathcal{A}$ into some commutative algebra.

9.2 Cayley transformation

Now let T be an unbounded operator in a Hilbert space H and let D_T denote the domain of T.

Proposition 9.2.1

$$(T = T^*) => ((T \pm iI)D_T = H).$$

Proof
Step 1.
We shall prove that $(T \pm iI)D_T$ is dense in H:
Let as assume that there exist $0 \neq u \perp (T \pm iI)D_T$, which means that for all $v \in D_T$ we have $(u|(T \pm iI)v) = 0$, so $(u|Tv) = (\pm iu|v)$, which implies that $u \in D_{T^*}$ and $T^*u = iu$. But $T^* = T$, so $Tu = iu$, which iscontradiction.
Step 2.
We shall prove that $(T \pm iI)D_T$ is closed in H:
Because $T = T^*$, then for each $v \in D_T$ we can directly calculate that

$$\|(T \pm iI)v\|^2 = \|Tv\|^2 + \|v\|^2. \tag{13}$$

Now let $h \in H$, then there exists a sequence $\{v_n\}$ such that

$$(T + iI)v_n \to h \quad \text{if} \quad n \to \infty.$$

Thus, by (13), $\{v_n\}$ and $\{Tv_n\}$ are Couchy sequences, so there exist $v, w \in H$ such that

$$v_n \to v \quad and \quad Tv_n \to w.$$

Moreover, T is closed (because $T = T^*$), then $Tv = w$.
Finally $(T + iI)v = h$. □

Corollary 9.2.1 *$T \pm iI$ is invertible.*

Proof
It follows from (13) that $ker(T \pm iI) = \{0\}$. □

Definition 9.2.1 *The Cayley transformation of an operator $T = T^*$ is the operator*

$$U_T \stackrel{\text{df}}{=} (T - iI)(T + iI)^{-1}.$$

Proposition 9.2.2 *U_T has the following properties:*
a) the domain of U_T is the whole H;
b) $U_T(H) = H$;
c) for all $h \in H$ we have $\|U_T h\| = \|h\|$.

□

Corollary 9.2.2 *U_T is a unitary operator.*

□

9.3 Idea of a proof of the spectral theorem for self-adjoint operators

In particular, we have the spectral representation of U_T:

$$U_T = \int_{|\lambda|=1} \lambda E_{U_T}(d\lambda).$$

The main idea is as follows:
Using the spectral representation of U_T we want to obtain the representation of T due to the inverse Cayley transformation:

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

Thus we propose

$$T = \int_{|\lambda|=1} -i\frac{\lambda+1}{\lambda-1} E_{U_T}(d\lambda),$$

where, of course

$$-i\frac{\lambda+1}{\lambda-1} \in \mathbf{R}.$$

It is easy to check that the mapping

$$S^1 - \{1\} \ni \lambda \mapsto C^{-1}(\lambda) \stackrel{\text{df}}{=} -i\frac{\lambda+1}{\lambda-1} \in \mathbf{R}$$

is "onto." So we may define the spectral measure on $\mathbf{R}$:

$$E_T(\Delta) \stackrel{\text{df}}{=} E_{U_T}(C(\Delta)).$$

Now, by the Proposition 7.6.1 we obtain:

$$T = \int_{S^1} C^{-1}(\lambda) E_{U_T}(d\lambda) = \int_{\mathbf{R}} \mu E_{U_T}(C(d\mu)) =$$

$$= \int_{\mathbf{R}} \mu E_T(d\mu),$$

where $\mu = C^{-1}(\lambda)$.

9.4 The spectral theorem for self-adjoint operators

Theorem 9.4.1 *If T is a self-adjoint operator in a Hilbert space H, then there exists a spectral measure E_T on the real line $\mathbf{R}$ such that:*

$$T = \int_{-\infty}^{+\infty} \mu E_T(d\mu).$$

Proof
For $u \in D_T$ we have (from the Cayley transformation) the following formula

$$U_T(T + iI)u = (T - iI)u \tag{14}$$

or, in another form

$$(U_T - I)Tu = -i(U_T + I)u. \tag{15}$$

Let as consider the equation $(U_T - I)v = 0$. By the Proposition 9.2.1 let $x \in D_T$ be such that $v = (T + iI)x$. Using (14) we obtain $U_T v = (T - iI)x$. But our equation gives $U_T v = v$. Therefore

$(T+iI)x=(T-iI)x$, which means $ix=-ix$ or $x=0$, so $v=0$. Concluding, the equation (15) determines Tu in a unique manner.

It will be convenient to define

$$C_\varepsilon \stackrel{\text{df}}{=} \{\lambda \in S^1 : |\arg\lambda| \geq \varepsilon\}$$

and

$$H_\varepsilon \stackrel{\text{df}}{=} E_{U_T}(C_\varepsilon)H.$$

For $u \in H_\varepsilon \cap D_T$ we have a good defined bounded operator:

$$Tu = \int_{C_\varepsilon} -i\frac{\lambda+1}{\lambda-1}E_{U_T}(d\lambda)u,$$

which, for our u, we may write in the following form:

$$Tu = \int_{S^1} -i\frac{\lambda+1}{\lambda-1}E_{U_T}(d\lambda)u,$$

(of course, $S^1 = C_0$).

We shall check that this formula satisfies the equation (15):

$$(U_T - I)\int_{S^1} -i\frac{\lambda+1}{\lambda-1}E_{U_T}(d\lambda)u =$$

$$= \int_{S^1}(\lambda-1)E_{U_T}(d\lambda)\int_{S^1} -i\frac{\lambda+1}{\lambda-1}E_{U_T}(d\lambda)u =$$

$$= -i\int_{S^1}(\lambda+1)E_{U_T}(d\lambda)u = -i(U_T+I)u.$$

This allows us to put for $u \in D_T$:

$$Tu \stackrel{\text{df}}{=} \lim_{\varepsilon\to 0}\int_{C_\varepsilon} -i\frac{\lambda+1}{\lambda-1}E_{U_T}(d\lambda)u. \tag{16}$$

First of all we check that this Tu satisfies (15) too:
Since the operator $U_T - I$ is continous, it commutes with the operation of taking limit

$$(U_T - I)\lim_{\varepsilon\to 0}\int_{C_\varepsilon} -i\frac{\lambda+1}{\lambda-1}E_{U_T}(d\lambda)u =$$

$$= \lim_{\varepsilon\to 0}\int_{C_\varepsilon} -i(\lambda+1)E_{U_T}(d\lambda)u \stackrel{(*)}{=}$$

$$= -i\int_{C_0}(\lambda+1)E_{U_T}(d\lambda)u = -i(U_T+I)u,$$

where the equality (*) follows from the fact that:

$$\lim_{\varepsilon\to 0}\|\int_{C_0\setminus C_\varepsilon}(\lambda+1)E_{U_T}(d\lambda)u\|^2 = (u|E_{U_T}(\{1\})u) = 0,$$

because $E_{U_T}(\{1\}) = 0$ (i.e. there does not exist a vector $0 \neq v \in H$, such that $U_T v = v$).

To prove that the limit in (16) exists, it should be noted that we may rewrite (15) in the following form:

$$\int_{C_0}(\lambda-1)E_{U_T}(d\lambda)Tu = -i\int_{C_0}(\lambda+1)E_{U_T}(d\lambda)u.$$

Now, because $\int_{C_\varepsilon} (\lambda - 1)^{-1} E_{U_T}(d\lambda)$ a bounded operator,

$$\int_{C_\varepsilon} 1 E_{U_T}(d\lambda) T u = \int_{C_\varepsilon} (\lambda - 1)^{-1} E_{U_T}(d\lambda) \int_{C_0} (\lambda - 1) E_{U_T}(d\lambda) T u =$$

$$= \int_{C_\varepsilon} -i \frac{\lambda + 1}{\lambda - 1} E_{U_T}(d\lambda) u.$$

Thua the limit exists and by the Proposition 7.6.1 we have for all $u \in D_T$:

$$Tu = \int_{-\infty}^{+\infty} \mu E_T(d\mu) u,$$

where $\mu = -i\frac{\lambda+1}{\lambda-1}$. Thus we have

$$T \subset \int_{-\infty}^{+\infty} \mu E_T(d\mu).$$

It is easy to see that $\int_{-\infty}^{+\infty} \mu E_T(d\mu)$ is symmetric, then, because $T = T^*$, we have:

$$T = \int_{-\infty}^{+\infty} \mu E_T(d\mu).$$

$\square$