4. Hilbert adjoints and some special operators in Hilbert space. Spectral decomposition of self-adjoint compact operator

In Linear Algebra basic courses two important operations on matrices are considered: the transposition $A \rightsquigarrow A^T$ and the adjoint $A \rightsquigarrow A^*$, where $A^* := A^T$ for real matrices, and in the complex case $A^* := \overline{(A^T)}$ (and $\overline{}$ means just the conjugation of the entries of matrix). Both are tightly related to some "operations on operators" represented (in an appropriate sense) by those matrices A, A^T/A^*. Moreover, these operations on operators can be defined not only in the finite dimensional case of "the basic course", but also for much more cases. Here we define the first (related to ^{T}) for normed spaces, and the second (related to *) for Hilbert spaces. They are called Banach*) and Hilbert adjoint operations, respectively.

---

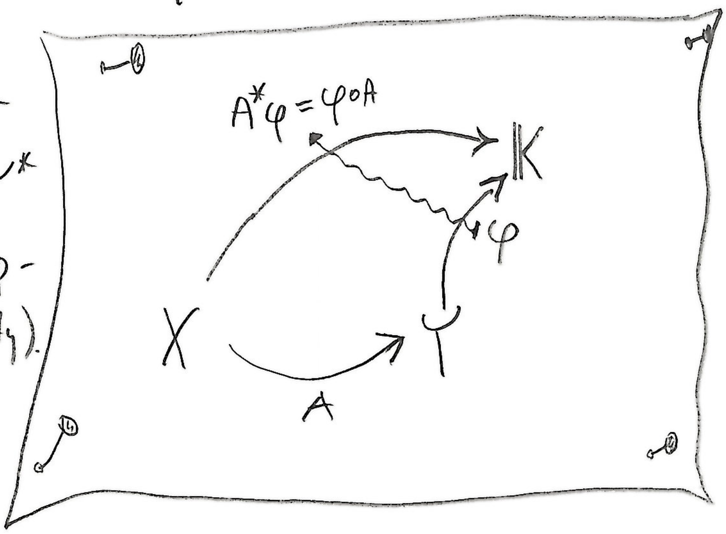
*) This suggests, that the spaces should be Banach, not normed only. And in fact, we use this adjoint usualy for the Banach case, but it is not necessary for the definition itself...

4.1. The Banach adjoint operator

Let X, Y be normed spaces and $A \in B(X,Y)$.
We define $A^*: Y^* \to X^*$ simply as follows:

$$A^*(\varphi) := \varphi \circ A \quad \text{for } \varphi \in Y^*. \qquad (1)$$

Surely, this is a proper definition, i.e. $\varphi \circ A \in X^*$ for $\varphi \in Y^*$ (the composition of two linear mappings is linear, similarly with continuity).
Moreover, $A^* \in \mathcal{L}(Y^*, X^*)$, obviously.
And even better: for any $\varphi \in Y^*$



$$\|A^*\varphi\|_{X^*} = \|\varphi \circ A\|_{X^*} \leq \|\varphi\|_{Y^*} \|A\|_{B(X,Y)},$$

hence $A^* \in B(Y^*, X^*)$ and $\|A^*\| \leq \|A\|$ (i.e., $\|A\|_{B(X,Y)}$).
The operator A^* defined in (1) is called **the Banach adjoint of A**.
And the estimate above can be improved...

$\boxed{\text{Fact}}$ ("On Banach adjoint")

Let X, Y, Z be normed spaces and $A \in B(X,Y)$. Then:

(i) $\|A^*\| = \|A\|$;

(ii) $*: B(X,Y) \to B(Y^*, X^*)$ is linear (and an isomorphism (isometry) onto its image);

(iii) If $B \in B(Y, Z)$ then
$$(BA)^* = A^* B^*;$$

(iii') If $X = Y$, then $I^* = I$;

(iv) If $X = Y \neq \{0\}$, then $\sigma(A^*) = \sigma(A)$.

$\boxed{\text{Proof}}$

(i)

By corollaries from the Hahn-Banach thm. we get:

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\| = \sup_{\|x\| \leq 1} \Big(\sup_{\substack{\|\varphi\| \leq 1 \\ (\varphi \in Y^*)}} |\varphi(Ax)| \Big) =$$

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

$$= \sup_{\|\varphi\| \leq 1} \|A^*\varphi\|_{X^*} = \|A^*\|.$$

(ii), (iii), (iii') are very easy to "calculate" from the definition of the Banach adjoint $\rightarrow \Delta$ And

(iv) follows just directly from (ii), (iii), (iii') by the definition of σ ($\rightarrow \Delta$). $\boxed{\checkmark}$

---

*) Note, that we use here the property of sup:
$$\sup\{\sup\{A_{i,j} : j \in J\} : i \in I\} = \sup\{\sup\{A_{i,j} : i \in I\} : j \in J\}$$
for any family of nonempty $\mathbb{R}$ subsets $\{A_{i,j}\}_{\substack{i \in I \\ j \in J}}$ with $I, J \neq \emptyset$. This is a rare example of possibility of interchanging (of) the order of two "passages to the limits" in Analysis without any extra strong assumptions. Prove this formula $\rightarrow \Delta$ (Hint $\boxed{\text{OTST} - 64}$: use another one: "$\sup(\bigcup_{s \in S} X_s) = \sup\{\sup X_s : s \in S\}$"!

4.2 The Hilbert-adjoint operator. Normal, self-adjoint and unitary operators

The Definition of Hilbert-adjoint

Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. We shall first make a simple observation – a corollary from the Riesz thm (the representation for Hilbert spaces).

Fact ("On the Hilbert-adjoint")

Suppose that $A \in B(\mathcal{H}_1, \mathcal{H}_2)$. Then there exists a unique map $B : \mathcal{H}_2 \longrightarrow \mathcal{H}_1$ such that

$$\forall_{\substack{x \in \mathcal{H}_1 \\ y \in \mathcal{H}_2}} \quad (Ax, y) = (x, B(y)); \qquad (1)$$

moreover a (any) map $B : \mathcal{H}_2 \to \mathcal{H}_1$ satisfying (1) belongs to $B(\mathcal{H}_2, \mathcal{H}_1)$, and $\|B\| \leq \|A\|$.

Proof

For any $y \in \mathcal{H}$ define $\varphi_y : \mathcal{H}_1 \longrightarrow \mathbb{K}$ by

$$\varphi_y(x) := (Ax, y) \quad \text{for } x \in \mathcal{H}_1.$$

Obviously, φ_y is a linear functional, and it is bounded by the Schwarz ineq. and by the boundedness of A $(\longrightarrow \triangle)$.
So $\varphi_y \in \mathcal{H}_1^*$, thus by the Riesz thm there exists a vector $z_y \in \mathcal{H}_1$ (a unique one...) such that

$$\forall_{x \in \mathcal{H}_1} \quad \varphi_y(x) = (x, z_y).$$

Hence, (1) holds, if we define $B(y) := z_y$.
The uniqueness follows from the above, too. And the uniqueness gives easily the linearity of B $(\longrightarrow \triangle)$. We also have:

$$\forall_{y \in \mathcal{H}_2} \quad \|B(y)\| = \|z_y\| = \|\varphi_y\| \leq \|A\| \cdot \|y\| \quad (\longrightarrow \triangle),$$

so $B \in B(\mathcal{H}_2, \mathcal{H}_1)$ and $\|B\| \leq \|A\|$. $\square$

This result allows now to define:

Definition

If $A \in B(\mathcal{H}_1, \mathcal{H}_2)$, then A^* is the operator from $B(\mathcal{H}_2, \mathcal{H}_1)$ such that

$$\forall_{\substack{x \in \mathcal{H}_1 \\ y \in \mathcal{H}_2}} \quad (Ax, y) = (x, A^* y) \qquad (1').$$

A^* is called <u>the Hilbert-adjoint to A</u>, but it is usually abbreviated to "<u>adjoint</u>", simply.

Remark

The notations used for Hilbert and Banach adjoints are the same "$*$" and the names are similar ... - this can make some formal problems and misunderstandings!
But fortunately they are different kind of objects - recall that the Banach-adjoint to A is an operator from $\mathcal{H}_2^*$ to $\mathcal{H}_1^*$, and not from $\mathcal{H}_2$ to $\mathcal{H}_1$, as the Hilbert adjoint.

And moreover they are tightly related:
denote here, for a moment, the Banach-adjoint of A by A^+, to distinguish it from the Hilbert-adjoint A^*.
If we use "the Riesz identifications" h_j, $j=1,2$
$h_j : \mathcal{H}_j \to \mathcal{H}_j^*$, $h_j(y) = \eta_{jy}$, where $\eta_{jy}(x) = (x,y)_{\mathcal{H}_j}$ for $x,y \in \mathcal{H}_j$, then one can easily check ($\to$ 4) that

$$A^* = h_1^{-1} \circ A^+ \circ h_2 . \qquad (2)$$

Several properties of A^* are similar to those of the Banach adjoint, however they can be proved without the use of (2).

$\boxed{\text{Theorem}}$ ("Properties of the Hilbert adjoint")

Suppose that $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ are Hilbert spaces and $A \in B(\mathcal{H}_1, \mathcal{H}_2)$.

1° $B(\mathcal{H}_1, \mathcal{H}_2) \ni T \rightsquigarrow T^* \in B(\mathcal{H}_2, \mathcal{H}_1)$ is a conjugate-linear map.

2° $(A^*)^* = A$

3° $\|A^*\| = \|A\|$

4° If $B \in B(\mathcal{H}_2, \mathcal{H}_3)$, then $(BA)^* = A^* B^*$.

If moreover $\mathcal{H}_1 = \mathcal{H}_2$, then

5° $I^* = I$

6° $\sigma(A^*) = \overline{\sigma(A)}$. *)

$\boxed{\text{Proof}}$

1°, 2°, 4° – are very simple to "calculate" from the definition (1') and Fact p. OTST-65. ($\longrightarrow$). Now 2° and $\|A^*\| \leq \|A\|$ (from the above fact) give 3°. 5° is obvious, and hence 6° follows from 1°, 4°, 5° ($\longrightarrow$).

---

*) Here "$\overline{}$" means the complex conjugated set, and not the closure...

Examples

Theorem ("The spectral decomposition of self-adjoint compact operators", "The Hilbert-Schmidt thm.")

If $A \in \mathcal{C}(\mathcal{H})$ and A is self-adjoint, then there exist an orthonormal base $\{x_i\}_{i\in\mathbb{I}}$ and a function $\{\lambda_i\}_{i\in\mathbb{I}}$ *) with $\lambda_i \in \sigma_p(A)$ for any $i\in\mathbb{I}$ such that

$$\forall_{i\in\mathbb{I}} \quad Ax_i = \lambda_i x_i \,. \tag{1}$$

I.e., the above result means that there exists an o.n. base consisting of the eigenvectors of such operator A!

Before we prove this thm. let us formulate some conclusions and remarks.

Remarks

(i) If the (base $\{x_i\}_{i\in\mathbb{I}}$ and "the eigenvalue-function" $\{\lambda_i\}_{i\in\mathbb{I}}$ on.) is as above, then for any $\mu \in \sigma_p(A)$

$$\operatorname{lin}\{x_i : i\in\mathbb{I},\ \lambda_i = \mu\} = \operatorname{Ker}(A-\mu I); \,^{**)} \tag{2}$$

in particular $\#\{i\in\mathbb{I} : \lambda_i = \mu\} = \dim \operatorname{Ker}(A-\mu I)$, and

---

*) We use the "sequence-type" notation for this function here for the finite matrix-analogy reason; formally, (this) $\{\lambda_i\}_{i\in\mathbb{I}}$ is just some $F : \mathbb{I} \to \sigma_p(A)$, with $F(i) := \lambda_i$ for each i.

**) And this is the eigenspace of A for the eigenvalue μ (i.e. $\{0\} \cup$ "the set of all the OTST - eigenvectors of A for μ").

a

$\mathbb{I}_* := \{ i \in \mathbb{I} : \lambda_i \neq 0 \}$ is at most countable.

(ii) If moreover (to the assumptions of the thm.) the above defined $\mathbb{I}_*$ is not finite, then the set of indices $\mathbb{I}$ can by chosen in such a way, that $\mathbb{I}_* = \mathbb{N}$, and then $\{\lambda_n\}_{n\in\mathbb{N}}$ converge to 0. *)

(iii) If moreover (again - to the assumptions of the thm., with no relations with (ii)) $\mathbb{I}$ is countable, then $\mathbb{I}$ can be chosen to be $\mathbb{N}$, and then $\lambda_n \longrightarrow 0$. The above holds in particular for any infinite dimensional separable $\mathcal{H}$.

$\boxed{\text{Proof}}$ (of the Rem.)

It is almost obvious from (1) with the use of the Corollary below, the fact that $\dim \operatorname{Ker}(A - \mu I) < +\infty$ for $\mu \in \sigma_p(A)$ and the fact that only 0 can be a limit point of $\sigma_p(A)$ for $A \in \mathcal{C}(\mathcal{H})$. $\boxed{\text{h}}$

$\boxed{\text{Corollary}}$

If $A \in \mathcal{C}(\mathcal{H})$ and A is self-adjoint, then: there exist an o.n. base $\{x_i\}_{i\in\mathbb{I}}$ and a $\overbrace{\text{function}}^{\sigma_p(A)\text{-valued}}$ $\{\lambda_i\}_{i\in\mathbb{I}}$ such that

$$Ax := \sum_{i\in\mathbb{I}} \lambda_i (x, x_i) x_i, \quad x \in \mathcal{H}; \qquad (3)$$

and the unitary transformation $\phi : \ell^2(\mathbb{I}) \longrightarrow \mathcal{H}$

---

*) Note that $\mathbb{I}_0 := \mathbb{I} \setminus \mathbb{I}_* \; (= \{ i \in \mathbb{I} : \lambda_i = 0 \})$ can be $\emptyset$ (but then $\dim \mathcal{H} < +\infty$, because $0 \in \sigma_p(A)$ if $\dim \mathcal{H} = +\infty$), it can be finite or not, including any cardinality of $\mathbb{I}_0$ (also uncountable).

OTST -

from Thm "On ℓ^2 representation of Hilbert space" (see (1) p. HS-65) transfers A to $M_{\{\lambda_i\}_{i\in I}}$ - the multiplication operator by $\{\lambda_i\}_{i\in I}$ in $\ell^2(I)$, in the sense, that

$$\Phi^{-1} A \,\Phi = M_{\{\lambda_i\}_{i\in I}}. \qquad (4)$$

$\boxed{\text{Proof}}$ (of the Coroll.)

We choose the same o.n. base and $\{\lambda_i\}_{i\in I}$ as in the Hilbert-Schmidt thm. Now, by the thm "On orthonormal base" (p. HS-58) for any $x\in\mathcal{H}$ we have

$$x=\sum_{i\in I}(x,x_i)\,x_i\,.$$

But $A\in B(\mathcal{H})$, hence (⟶ $\triangle$: please - explain the details below, see p. HS-41 -44)

$$Ax=\sum_{i\in I}(x,x_i)\,Ax_i=\sum_{i\in I}(x,x_i)\,\lambda_i x_i\,,$$

which gives (3). Now, to get (4) it suffices to check if for any $x=e_i$, $i\in I$ (see Example p. HS-64), because both the RHS and LHS of (4) are in $B(\ell^2(I))$ and $\{e_i\}_{i\in I}$ is linearly dense (, being an orthonormal base).
But for any e_i (4) is just obvious ... !

//

OTST-

c

Proof of the Hilbert-Schmidt thm

- For any $\mu \in \sigma_p(A)$ define $\mathcal{H}_\mu := \mathrm{Ker}(A - \mu I)$
  - it is a closed, nontrivial ($\neq \{0\}$) subspace of $\mathcal{H}$, since $\mu \in \sigma_p(A)$, hence $\mathcal{H}_\mu$ is a Hilbert space. Let $\{x_{(\mu,\alpha)}\}_{\alpha \in A_\mu}$ be an o.n. base for $\mathcal{H}_\mu$ (note, that A_μ is finite for any $\mu \neq 0$, however A_0 can be "large", if $0 \in \sigma_p(A)$). Now choose $\mathbb{I} := \{(\mu, \alpha) : \mu \in \sigma_p(A), \alpha \in A_\mu\}$ and for $i = (\mu, \alpha) \in \mathbb{I}$ define

$$x_i := x_{(\mu,\alpha)}. \quad *)$$

Observe that $\{x_i\}_{i \in \mathbb{I}}$ is an o.n. system, because $\mathcal{H}_\mu \perp \mathcal{H}_{\mu'}$ for $\mu, \mu' \in \sigma_p(A), \mu \neq \mu'$ by Lemma.

We obviously have

$$A x_i = A x_{(\mu,\alpha)} = \mu x_{(\mu,\alpha)} = \mu x_i \qquad \text{for } i = (\mu, \alpha) \in \mathbb{I},$$

because $x_i \in \mathcal{H}_\mu = \mathrm{Ker}(A - \mu I)$. Thus (1) holds, if we choose for $i = (\mu, \alpha) \in \mathbb{I}$,

$$\lambda_i := \mu.$$

- It suffices to prove that $\{x_i\}_{i \in \mathbb{I}}$ is linearly dense in $\mathcal{H}$, i.e that $Y^\perp = \{0\}$ for $Y := \mathrm{lin}\{x_i\}_{i \in \mathbb{I}}$ (see, e.g., Thm. "On o.n. base" p HS-58 + "A linear density

---

*) So we see that to make the proper "indexation" for the o.n. base consisting on eigenvectors of A we should use: 1) the eigenvalue μ 2) the properly chosen $\alpha \in A_\mu$ which is related to "the multiplicity" of μ as an eigenvalue.

OTST -

d

Criterion" p. HS-32).

- Suppose, that $Y^\perp \neq \{0\}$. Observe, that Y is an invariant space for A: – It is obvious that $\mathrm{lin}\{x_i\}_{i\in\mathbb{I}} =: Y_0$ is invariant, and hence, if $y \in Y$, then $y_n \to y$ for some $y_n \in Y_0$, $n \in \mathbb{N}$, so $Ay_n \to Ay$. But $Ay_n \in Y_0$, therefore $Ay \in \overline{Y_0} = Y$. By Fact "On invariant subspaces" $Y^\perp$ is also invariant for A.
- Thus $\tilde{A} := A_{|Y^\perp} : Y^\perp \to Y^\perp$, moreover $\tilde{A} \in \mathcal{C}(Y^\perp)$, because $A \in \mathcal{C}(\mathcal{H})$ (see Rem. "On independ. of precompactness" p. OTST-18; note that $Y^\perp$ is a closed space in $\mathcal{H}$).
- Now, by the Riesz - Schauder thm. (see p. OTST-), if $\lambda \in \sigma(\tilde{A})$ and $\lambda \neq 0$, then $\lambda \in \sigma_p(\tilde{A})$, so there exists $0 \neq x \in Y^\perp$ such that
$$Ax = \tilde{A}x = \lambda x, \quad \text{i.e} \quad x \in \mathrm{Ker}(A - \lambda I) \text{ and } \lambda \in \sigma_p(A).$$
But this means that $x \in \mathcal{H}_\lambda$, i.e., $x \in \mathrm{lin}\{x_{(\lambda,\alpha)}\}_{\alpha \in A_\lambda} \subset Y_0 \subset Y$. This is impossible, because $x \neq 0$ and $x \in Y \cap Y^\perp$. So, $\lambda = 0$, and $\sigma(\tilde{A}) \subset \{0\}$. Therefore $\|\tilde{A}\|_{sp} = 0$.
- But by Fact "On inv. subspaces" gives also that $\tilde{A}$ is selfadjoint. And thus $\|\tilde{A}\| = \|\tilde{A}\|_{sp} = 0$. Hence $\{0\} \neq Y^\perp = \mathrm{Ker}\,\tilde{A} \subset \mathrm{Ker}\,A$, but this means, that $0 \in \sigma_p(A)$, so $Y^\perp \subset \mathrm{Ker}\,A = \mathcal{H}_\mu \subset Y$ for $\mu = 0 \in \sigma_p(A)$. And so, $Y^\perp = Y^\perp \cap Y = \{0\}$ – a contradiction!

$\square$

e