3. Automorphisms and spectrum

This part can be treated as a "very initial introduction" to Spectral Theory – a very important topic of analysis, one of the main pillars of Operator Theory.
We shall introduce here only some basic spectral notions: the spectrum, the resolvent set and the resolvent operator function.
This is nearly related to automorphisms of Banach spaces.

3.1. Invertibility of bounded operators

We shall study here the problem of invertibility of $T \in B(X)$ for X being a Banach space. The main question we ask and try to answer is:

" Is a »small perturbation« of an invertible operator also invertible?"

We start from the simplest invertible operator.

◆ Perturbations of I

We shall show that the answer is "YES" when we perturb the identity operator I, even if we do not treat the above "smallness" very strictly ...

If $A \in \mathcal{L}(X)$, then we denote $A^0 := I$, $A^{n+1} := AA^n$, $n \geqslant 0$ as usual, and $A^n \in \mathcal{L}(X)$ for any n; moreover $A^n \in B(X)$ if $A \in B(X)$. Observe, that the definition $A^0 = I$ is valid for any A, in particular for $A = 0$ ("$0^0 = I$").

Recall the general function terminology: $F: X \to Y$ is invertible iff F is "onto" Y and injective (i.e., together, it is bijective)*). For F-invertible (as $F: X \to \underline{\underline{Y}}$) F^{-1} denotes the inverse function. We have preserved this terminology for linear operators (see p. OF-11).

Let us recall now a well-known "geometric sequence formula" for numbers:

$$\frac{1}{1-a} = \sum_{n=0}^{+\infty} a^n, \quad |a| < 1 \qquad (a^0 := 1 \text{ also for } a = 0).$$

This may be surprising, but it turns out, that the number a can be "replaced" by any bounded operator in a Banach space!

$\boxed{\text{Lemma}}$ ("On $(I-A)^{-1}$")

If X is a Banach space and $A \in B(X)$, $\|A\| < 1$, then $(I-A)$ is invertible, $\sum_{n=0}^{+\infty} A^n$ is convergent and absolutely convergent in $B(X)$ and

$$B(X) \ni (I-A)^{-1} = \sum_{n=0}^{\infty} A^n, \quad \|(I-A)^{-1}\| \leqslant \frac{1}{1-\|A\|}. \qquad (1)$$

$\boxed{\text{Proof}}$

Consider first the scalar series $\sum_{n=0}^{+\infty} \|A^n\|$. We have

---

*) However, some mathematicians use the other terminology, where "invertible" means injective only, and then "inverse" is the appropriate function from $F(X)$ onto X also when $F(X) \neq Y$. But here we do not use this extended meaning!

$\|A^n\| \leq \|A\|^n$ for $n \geqslant 1$ (see e.g. Fact p. OF-44) and also $\|A^0\| = \|I\| \leq \|A\|^0 = 1$ *). But $q := \|A\| < 1$, so with $\|A^n\| \leq q^n$ we get the convergence of $\sum\limits_{n=0}^{\infty} \|A^n\|$. Hence $\sum\limits_{n=0}^{+\infty} A^n$ is absolutely convergent, and it is convergent, by the completeness of $B(X)$ (see thm. "On completness of $B(X,Y)$", p. OF-46). So, to finish the proof, it is sufficient to check that for $S := \sum\limits_{n=0}^{+\infty} A^n$:

$$(I - A) S = S (I - A) = I \qquad (2)$$

(we should use then also $\|S\| \leq \sum\limits_{n=0}^{+\infty} \|A\|^n = \frac{1}{1-\|A\|}$). But we have

$$S(I-A) = \sum_{n=0}^{+\infty} \left(A^n - A^{n+1}\right) = (I-A)S \qquad (3)$$

– it is an obvious corollary from the fact "on the product continuity", which we formulate and prove below; note, that (3) is in fact just the commuting (interchanging of the order) of the operations of "$\sum\limits_{n=0}^{\infty}$" and of left / right multiplication by $(I-A)$.

But we have:

$$\sum_{n=0}^{+\infty} (A^n - A^{n+1}) = \sum_{n=0}^{+\infty} A^n - \sum_{n=0}^{+\infty} A^{n+1} \overset{**)}{=} S - (S - I) = I,$$

and (2) follows. $\boxed{4}$

---

*) We adopt here the agreement, that $0^0 = 1$, i.e., $a^0 = 1$ for any scalar a. Note also, that "$\|A^0\| = \|I\| = 1$" can be not true... – we have $\|I\| = 0$ if (and only if...) $X = \{0\}$.

**) We use here some "trivial" properties of series in the norm space. – contin. on the next page...

Now formulate the result used above in the proof of (3).

$\boxed{\text{Fact}}$ ("On the product continuity")

If X, Y, Z are normed spaces, then the multiplication $\cdot$ (i.e., the composition $\circ$) is a bilinear continuous operation from $B(Y,Z) \times B(X,Y)$ into $B(X,Z)$.

$\boxed{\text{Proof}}$

The bilinearity is obvious. To prove the continuity suppose that $A_n, A \in B(Y,Z)$, $B_n, B \in B(X,Y)$ and $A_n \to A$, $B_n \to B$. Then

$$A_n B_n - AB = (A_n - A) B_n + A(B_n - B),$$

so, by "submultiplicativity" of the operator <u>norm</u>, (see Fact p. OF-44)

$$0 \leq \|A_n B_n - AB\| \leq \|A_n - A\| \|B_n\| + \|A\| \|B_n - B\| \to 0 + 0 = 0,$$

because $\|B_n\| \to \|B\|$ (the continuity of the norm in <u>any</u> normed spaces...), hence $\{\|B_n\|\}$ is bounded. Therefore $A_n B_n \to AB$. $\boxed{\square}$

---

the continuation of **) from the previous page:

They seem so trivial, that we can forget that we use any property, at all... — But be careful and try to <u>formulate</u> them, and <u>prove</u> them (for any normed space) $\to$ ⚠.

Automorphisms

If X is a norm space, then isomorphisms (the linear continuous ones) of X onto X are called automorphisms. The set of all the automorphisms of X is sometimes denoted by $Aut(X)$, but we shall rather use here

$$B_*(X)$$

for short (note that $B_*(X) \subset B(X)$, but it is not a subspace of $B(X)$, excluding the case of $X = \{0\}$ ...).

Generally speaking, to be an authomorphism is a stronger condition than to be an invertible function from $B(X)$ – we need also the continuity of the inverse operator. We recall however that for Banach spaces the problem is easier.

Remark

Let X be a Banach space, $A \in B(X)$

Then TFCAE:

(i) $A \in B_*(X)$

(ii) $Ker(A) = \{0\}$ and $Ran(A) = X$

(iii) A is invertible (as a function $A : X \longrightarrow X$).

Proof

(ii)$\Leftrightarrow$(iii) is known (and obvious) for linear operators. (i)$\Rightarrow$(iii) – obvious and (iii)$\Rightarrow$(i) follows from The inverse mapping thm.

(and is the only place, where the completness of X is important). $\boxed{\text{KL}}$

$\boxed{\text{Theorem}}$ ("On $B_*(X)$")

Let X be a Banach space. Then $B_*(X)$ is a nonempty open subset of $B(X)$ and the inversion operation $^{-1}: B_*(X) \longrightarrow B_*(X)$ is a homeomorphism of $B_*(X)$ (onto $B_*(X)$). Moreover, if $A \in B_*(X)$ and $r_A := \|A^{-1}\|^{-1}$ *), and $\Theta_A := \{T \in B(X): \; \|A^{-1}(A-T)\| < 1\}$, then Θ_A is open,

$$K(A, r_A) \subset \Theta_A \subset B_*(X),$$

and for $T \in \Theta_A$

$$\left.\begin{aligned} T^{-1} &= \left(\sum_{n=0}^{+\infty} (A^{-1}(A-T))^n\right) A^{-1}, \\ \|T^{1} - A^{-1}\| &\leq \|A^{-1}\| \frac{\|A^{-1}(A-T)\|}{1 - \|A^{-1}(A-T)\|}; \end{aligned}\right\} \quad (4)$$

analogously, if $\Theta'_A := \{T \in B(X): \; \|(A-T)A^{-1}\| < 1\}$, then Θ'_A is open, $K(A, r_A) \subset \Theta'_A \subset B_*(A)$, and for $T \in \Theta'_A$

$$\left.\begin{aligned} T^{-1} &= A^{-1} \sum_{n=0}^{+\infty} ((A-T)A^{-1})^n, \\ \|T^{-1} - A^{-1}\| &\leq \|A^{-1}\| \frac{\|(A-T)A^{-1}\|}{1 - \|(A-T)A^{-1}\|}. \end{aligned}\right\} \quad (4')$$

---

*) When $\|A^{-1}\| = 0$ (i.e., only when $X = \{0\}$ ...) then "0^{-1}" means $+\infty$, here.

$\boxed{\text{Proof}}$

Suppose that $A \in B_*(X)$, then for any $T \in B(X)$

$$T = A - (A-T) = A\left(I - A^{-1}(A-T)\right) \qquad (5)$$
$$= \left(I - (A-T)A^{-1}\right)A. \qquad (5')$$

Consider $R := A^{-1}(A-T)$, $R' := (A-T)A^{-1}$. Using Lemma "On $(I-A)^{-1}$" to R, R' instead of A, we get $\Theta_A, \Theta'_A \subset B_*(X)$, because if $C \in B_*(X)$, then $AC, CA \in B_*(X)$. Moreover, for such C, $(AC)^{-1} = C^{-1}A^{-1}$ and $((CA)^{-1} = A^{-1}C^{-1}$, hence the formula (1) of the Lemma gives (4), (4').

Now using $\|A^{-1}(A-T)\| \leq \|A^{-1}\| \, \|A-T\|$ we get $\Theta_A \subset B_*(X)$ and analogically for Θ'_A. So $B_*(X)$ is open; $B_*(X) \neq \emptyset$ since $I \in B_*(X)$ (also when $X = \{0\}$, but then $I = 0$ and $B(X) = B_*(X) = \{0\}$).

Observe, that the second part of (4) (and of (4'), too) shows, that the (the) inversion $^{-1}$ is Lipschitz on $K(A, r_A/2)$ for $A \in B_*(X)$ – This proves the continuity of $^{-1}$, but $^{-1}$ is onto $B_*(X)$, because for any $A \in B_*(X)$ $A = (A^{-1})^{-1}$. This also shows, that -1 is inverse to itself – hence -1 is a homeomorphism. $\boxed{\text{(4)}}$

3.2 The resolvent set and the spectrum

On the definition of the spectrum

We shall define here the main object for our topic — the spectrum, denoted by $\sigma(A)$, for the operator A.

Recall first the definition of the spectrum of the matrix from "Linear Algebra", i.e, for a square $d \times d$ matrix A:

"$\sigma(A)$ is the set of all $\lambda \in \mathbb{C}$, such that $\det(A - \lambda I) = 0$; this condition is equivalent to

$$\operatorname{Ker}(A - \lambda I) \neq \{0\}. \qquad (1)$$

and also to *)

$$\operatorname{Ran}(A - \lambda I) \neq \mathbb{C}^d. \text{"} \qquad (1')$$

In fact, both conditions (1), (1') (each of them) is equivalent to

$$(A - \lambda I) \text{ is not invertible} \qquad (2)$$

in the case of X being $\mathbb{K}^d$. But this is <u>not</u> true for the infinite dimensional spaces! Moreover for such X each of the sets of λ-s defined by (1), (1'), (2) can be different! (surely, there are inclusions ...).

$\boxed{\text{Definition}}$ **)

Let X be a linear space and $A \in \mathcal{L}(X)$. Then

---

*) In (1) and (1') the operator is identified with its matrix (for a fixed base)

**) We define $\sigma(A)$ and $\rho(A)$ here for <u>any linear operator in any linear space X</u>, however only for normed X, those notions are formulated often and the definition is ... →

cont. on next p.

$$\sigma(A) := \{\lambda \in \mathbb{K} : (2) \text{ holds}\}$$

is the spectrum of A. The set

$$\varrho(A) := \mathbb{K} \setminus \sigma(A)$$

is called the resolvent (or the resolvent set) of A.

And

$$\sigma_p(A) := \{\lambda \in \mathbb{K} : (1) \text{ holds}\}$$

is called the point spectrum of A or the set of eigenvalues of A. Each $\lambda \in \sigma_p(A)$ is called eigenvalue of A, and any $x \in X \setminus \{0\}$ satisfying the equation

$$Ax = \lambda x \qquad (3)$$

is an eigenvector of A (for λ) (and (3) itself is called "eigenequation"). The subspace

$$\operatorname{Ker}(A - \lambda I)$$

is called the eigenspace for A and λ ("of A for λ", too...).

Note, that $x \in \operatorname{Ker}(A - \lambda I)$ is equivalent to (3), but the eigenspace for A and λ and the set of all the eigenvectors of A for λ is not the same! Those two sets differs by the zero vector – this is not an eigenvector...

→ cont. of $**$) from the previous page: .. different very often. The requirement, that there is no any continuous inverse to $(A - \lambda I)$, seems to be more popular. But the problem is, that such a definition has no sense if X is only a linear space, with no fixed topology. Fortunately, those two definitions meet, when the normed space X is a Banach space and $A \in B(X)$.

Our considerations before the definition can be now formulated as follows.

Corollary

If $\dim X < +\infty$, then $\sigma_p(A) = \sigma(A)$ for any $A \in \mathcal{L}(X)$.

We add here also one more definition – of the resolvent operator:

$R_A : \rho(A) \longrightarrow \mathcal{L}(X)$

given by $R_A(\lambda) := (A - \lambda I)^{-1}$ is called the resolvent operator function, and each $R_A(\lambda)$ is the resolvent (or the resolvent operator) for (/of) A and λ.

Remark

If X is a Banach space and $A \in B(X)$ then $R_A(\lambda) \in B(X)$ for any $\lambda \in \rho(A)$, i.e. $R_A : \rho(A) \to B(X)$.

The above follows directly from the Inverse mapping thm. (compare to Remark p. OTST-38)

For spectral studies the choice of $\mathbb{K}$ starts to be very important, more important than before in our AFI course (it was sometimes important in some Hilbert Spaces parts till now, but still $\mathbb{K} = \mathbb{R}$ was "acceptable"...).

The non-emptiness problem and some basic properties of spectrum

Example

Consider the <u>real</u> matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

We have $\det(A - \lambda I) = \det\begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1$.

But $\lambda^2 + 1 = 0$ has no real roots... And it has two complex roots i and $-i$. Thus $\sigma(A) = \emptyset$, if we treat A as an operator in $\mathbb{R}^2$ with $\mathbb{K} = \mathbb{R}$ and $\sigma(A) = \{i, -i\}$ if we treat it as a $\mathbb{C}^2$ operator <u>with</u> $\mathbb{K} = \mathbb{C}$ (note, that for $\mathbb{C}^2$ we have still a choice of $\mathbb{K}$, however the dimension would be 4 not 2 for $\mathbb{K} = \mathbb{R}$, so the matrix would be "to small"...).

The above problem of empty spectrum for $\mathbb{K} = \mathbb{R}$ is not the only one – generally $\mathbb{R}$ is "not convenient" for spectral studies – the theory would be "very poor". So, we shall often assume, that $\mathbb{K} = \mathbb{C}$ here. Note, that this does not mean, that we cannot do anything for $\mathbb{K} = \mathbb{C}$! This paradoxically means, that the $\mathbb{R}$ theory is much complicated, i.e., complex (nomen omen...)! But many "real" problems ("real" for $\mathbb{R}$...) can be solved by the general idea of making "complexification" first – i.e. by finding the "complex" objects corresponding to the real ones; and by solving the complex problems for those complex objects. The last step is then to find the proper way back — from "complex" results to the "real" ones (which could be not so easy to get ...). Below, we see some poits where $\mathbb{K}=\mathbb{C}$ starts to be important...

Theorem ("On spectrum")

If $X\neq\{0\}$ is a Banach space and $A\in B(X)$, then

(i) $\sigma(A)$ is a compact subset of $\mathbb{K}$ and $\forall_{\lambda\in\sigma(A)}\ |\lambda|\leq\|A\|$,

If, moreover, $\mathbb{K}=\mathbb{C}$, then:

(iii) $\sigma(A)\neq\emptyset$,

(ii) $\rho(A)$ is an open, $R_A:\rho(A)\to B(X)$ is continuous and

$$\forall_{\lambda\in\mathbb{K}}\ \left(|\lambda|>\|A\| \Rightarrow \|R_A(\lambda)\|<\frac{1}{|\lambda|-\|A\|}\right) \quad (1)$$

**)

$$R_A(\lambda_1)-R_A(\lambda_2)=(\lambda_1-\lambda_2)R_A(\lambda_1)R_A(\lambda_2),\quad \lambda_1,\lambda_2\in\rho(A). \quad (2)$$

(iv) R_A is a ~~weakly~~ vector analytic*) function, i.e., for any $\varphi\in(B(X))^*$ $\varphi\circ R_A$ is analytic*) (as a function from $\rho(A)$ into $\mathbb{C}$).

---

*) analytic = holomorphic here. There are many kinds of "analyticity" of operator-(e.g. weak-operator, vector-operator...) valued, and generally, vector-valued functions of complex variables. (note, that $R_A(\lambda)$ is a bounded operator, so it is a vector in $B(X)$...) but they are all <u>the same</u> in the case of Banach space valued functions!

**) This formula is often called "<u>the (1-st) resolvent formula</u>".

$\boxed{\text{Proof}}$

If $\lambda_0 \in \varrho(A)$, then $(A-\lambda_0 I)$ is invertible, and $(A-\lambda_0 I) \in B_*(X)$ by Fact.

Let $r_0 := \|(A-\lambda_0 I)^{-1}\|^{-1}$, then for any $C \in K(0, r_0)$ also $(A-\lambda_0 I)+C \in B_*(X)$, by thm. "on $B_*(X)$". In particular for any $\lambda \in \mathbb{K}$ such that $|\lambda-\lambda_0|<r$ $A-\lambda I = A-\lambda_0 I + (\lambda_0-\lambda)I \in B_*(X)$, so $A-\lambda I$ is invertible. Thus $\varrho(A)$ is open and $\sigma(A) = \mathbb{K}\setminus\varrho(A)$ is a closed set. Moreover, if $0 \neq \lambda \in \sigma(A)$, then $A-\lambda I \notin B_*(X)$, hence also $I - \lambda^{-1}A = = (-\lambda^{-1})(A-\lambda I) \notin B_*(X)$. From Lemma "On $(I-A)^{-1}$" we now conclude, that $\|\lambda^{-1}A\| \geq 1$, so $|\lambda| \leq \|A\|$ – and this finishes the proof of (i) ($\sigma(A)$ is closed and bounded subset of $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$ $\Rightarrow$ $\sigma(A)$ – compact).

Let now $\lambda_1, \lambda_2 \in \varrho(A)$. We have:

$$R_A(\lambda_1) - R_A(\lambda_2) = (A-\lambda_1 I)^{-1} - (A-\lambda_2 I)^{-1} = (A-\lambda_1 I)^{-1} I - I(A-\lambda_2 I)^{-1} =$$
$$= \underline{(A-\lambda_1 I)^{-1}}(A-\lambda_2 I)(A-\lambda_2 I)^{-1} - \underline{(A-\lambda_1 I)^{-1}}(A-\lambda_1 I)(A-\lambda_2 I)^{-1} =$$
$$= R_A(\lambda_1)\left[A-\lambda_2 I - (A-\lambda_1 I)\right] R_A(\lambda_2) =$$
$$= R_A(\lambda_1)\cdot(\lambda_1-\lambda_2)I\cdot R_A(\lambda_2) = (\lambda_1-\lambda_2)R_A(\lambda_1)R_A(\lambda_2),$$

i.e. (2) holds.

Observe that R_A can be written in the form $h \circ p_A$, where $h : B_*(X) \to B_*(X)$ is the $^{-1}$ homeomorphism of $B_*(X)$ (see thm "on $B^*(X)$") and $p_A : \varrho(A) \to B_*(X)$ is given by $p_A(\lambda) := A - \lambda I$ for $\lambda \in \varrho(A)$. Obviously p_A is a Lipschitz function, thus it is continuous, so $R_A = h\circ p_A$ is also continuous. When $|\lambda| > \|A\|$, then $\lambda \in \varrho(A)$ by (i), and $A-\lambda I = (-\lambda)(I-\frac{1}{\lambda}A)$, hence

$\|R_A(\lambda)\| = \|(A-\lambda I)^{-1}\| = |\lambda^{-1}| \frac{1}{1-\|\frac{1}{\lambda}A\|} = \frac{1}{|\lambda| - \|A\|}$, by Lemma "On $(I-A)^{-1}$" again.
This gives (ii). In particular $\|R_A(\lambda)\| \le \frac{2}{|\lambda|}$ for $|\lambda| > 2\|A\|$.

Let now $\lambda, \lambda_0 \in \rho(A)$ and $\lambda \ne \lambda_0$. Then by (2):

$$\frac{R_A(\lambda) - R_A(\lambda_0)}{\lambda - \lambda_0} \overset{*)}{=} R_A(\lambda) R_A(\lambda_0).$$

Hence, if $\lambda_n \in \rho(A) \setminus \{\lambda_0\}$, $\lambda_n \to \lambda_0$, then $R_A(\lambda_n) \xrightarrow[n]{} R_A(\lambda_0)$ by the continuity of R_A, and by the continuity of the product (see Fact p. OTST – 37)

$$\frac{R_A(\lambda_n) - R_A(\lambda_0)}{\lambda_n - \lambda_0} \xrightarrow[n]{} (R_A(\lambda_0))^2$$

i.e, there exists the limit (in λ_0) of the λ_0-difference quotient for the operator function: $\lambda \mapsto R_A(\lambda)$, and the limit is $(R_A(\lambda_0))^2$. **)

Assume now that $\mathbb{K} = \mathbb{C}$ and let $\varphi \in (B(X))^*$ Denote by f_φ the scalar function given by $\varphi \circ R_A$. Hence $f_\varphi : \rho(A) \longrightarrow \mathbb{C}$ and $\rho(A)$ is an open subset of $\mathbb{C}$. Moreover, using again any $\{\lambda_n\}$ as above, we get (we use also the linearity and the continuity of φ):

$$\frac{f_\varphi(\lambda_n) - f_\varphi(\lambda_0)}{\lambda_n - \lambda_0} \xrightarrow[n]{} \varphi\left(R_A(\lambda_0)^2\right) \qquad (2)$$

Thus f_φ is complex-differentiable, i.e. - holomorphic (analytic), and (iv) holds. Suppose, that $\sigma(A) = \emptyset$, i.e, $\rho(A) = \mathbb{C}$. Thus R_A is bounded, since it is continuous and (1) holds. Hence, if $\varphi \in X^*$, then $f_\varphi : \mathbb{C} \to \mathbb{C}$ is bounded and holomorpic. So, f_φ is constant by the Liouville thm, which implies $f_\varphi'(\lambda) = 0$ for any $\lambda \in \mathbb{C}$. But (2) gives $0 = f_\varphi'(\lambda) = \varphi((R_A(\lambda))^2)$ for <u>any</u> $\varphi \in (B(X))^*$.

---

*) We use here (and in some other cases) $\frac{x}{\lambda} := \lambda^{-1}x$ for some vector x and $\lambda \in \mathbb{K} \setminus \{0\}$...

**) We can say that R_A is <u>differentiable</u> and $R_A'(\lambda_0) = (R_A(\lambda_0))^2$ for $\lambda_0 \in \rho(A)$.

Therefore using the conclusions from the Hahn-Banach theorem (applied to the normed space $B(X)$) (e.g. Thm.1 "A continuous functional for a fixed vector p. LF-19) we get $(R_A(\lambda))^2 = 0$ for any $\lambda \in \mathbb{C}$ (e.g. for $\lambda = 0$). Thus $I = (A-\lambda I)^2 (R_A(\lambda))^2 = 0$ – which means, that $X = \{0\}$ – a contradiction. $\boxed{\text{\Lightning}}$

Below we formulate several remarks concerning the above results and their proofs, but first we need a definition.

Definition

Let $A \in \mathcal{L}(X)$. The spectral norm *) or the spectral radius of A is

$$\|A\|_{sp} \overset{*)}{:=} \begin{cases} \sup\{|\lambda| : \lambda \in \sigma(A)\} & \text{if } \sigma(A) \neq \emptyset, \\ 0 & \text{if } \sigma(A) = \emptyset. \end{cases}$$

Remarks

1. Part (i) shows in particular, that $\|A\|_{sp} \leq \|A\|$ for $A \in B(X)$ and X – a Banach space. The example p. OTST-44 shows that if $\mathbb{K} = \mathbb{R}$, then we can obtain $\|A\|_{sp} < \|A\|$. But also for $\mathbb{K} = \mathbb{C}$ the inequality can be sharp... In fact, there is a famous

---

*) But be carefull! Both the name "spectral norm" and the notation $\|\cdot\|_{sp}$ are confusing – $\| \ \|_{sp}$ is NOT any norm "for operators"...

"spectral radius formula" (see ...) which expresses the $\|A\|_{sp}$ in terms of $\|A^n\|$ for all n-s.
For some special classes of operators, however, the equality $\|A\|_{sp} = \|A\|$ holds (see ...).

2. We have $R_A(\lambda_1)\,R_A(\lambda_2) = R_A(\lambda_2)\,R_A(\lambda_1)$, i.e each two of the resolvent operators for A commute.

3. Part (iii) was formulated only for $\mathbb{K}=\mathbb{C}$, but if $\mathbb{K}=\mathbb{R}$ then the proof above shows the differenciability of $\varphi \circ R_A$ too (but we cant say "holomorphic" then, since $\rho(A)$ is open in $\mathbb{R}$ only (and) not in $\mathbb{C}$...).

Some further properties of spectrum, spectral radius formula

We collect here some extra facts related to spectrum.

Fact ("On eigenvectors independence") *)

If X is a linear space ($A \in \mathcal{L}(X)$) and x_i is an eigenvector for A and $\lambda_i \in \sigma_p(A)$ for $i = 1, \dots n$ ($n \geq 1$), where $\lambda_i \neq \lambda_j$ for $i \neq j$, then $\{x_i\}_{i=1,\dots n}$ is linearly independent.

Proof

Note, that we have obviously $x_i \neq x_j$ for $i \neq j$ (since if $x_i = x_j$ and $i \neq j$ then $\lambda_i x_i = Ax_i = Ax_j = \lambda_j x_j = \lambda_j x_i \Rightarrow x_i = 0$, because $\lambda_i \neq \lambda_j$, but x_i is an eigenvector...). Thus $\#\{x_1, \dots, x_n\} = n$.

*) This result can be known from "Linear Algebra I"...

Since $W := \{x_1, \dots x_n\}$ is finite, there exist a minimal subset W' of W such that $\operatorname{lin} W' = \operatorname{lin} W$ $(\longrightarrow \triangle \dots)$.
Obviously such W' is linearly independent, and non-empty, because $W \neq \{0\}$, by those minimality. If $W' \subsetneq W$ then for some $\emptyset \neq I \subsetneq \{1, \dots n\}$ we have $W' = \{x_j : j \in I\}$. If $k_0 \in \{1, \dots, n\} \setminus I$, then $x_{k_0} = \sum_{i \in I} \alpha_i x_i$ for some $\alpha_i \in \mathbb{K}$, so taking $\tilde{I} := \{i \in I : \alpha_i \neq 0\}$ we have $\tilde{W} := \{x_i : i \in \tilde{I}\}$ – lin. independent too, and $x_{k_0} = \sum_{i \in \tilde{I}} \alpha_i x_i$; moreover $\tilde{I} \neq \emptyset$, because $x_{k_0} \neq 0$. So

$$\lambda_{k_0} x_{k_0} = A x_{k_0} = \sum_{i \in \tilde{I}} \alpha_i A x_i = \sum_{i \in \tilde{I}} \lambda_i \alpha_i x_i .$$

1° case: $\lambda_{k_0} = 0$. Then $\sum_{i \in \tilde{I}} \lambda_i \alpha_i x_i = 0 \Rightarrow \forall_{i \in \tilde{I}} \lambda_i \alpha_i = 0$, by the lin. indep. Thus $\forall_{i \in \tilde{I}} \lambda_i = 0$ – but $\tilde{I} \neq \emptyset$ and $k_0 \notin \tilde{I}$ and this is a contradiction since for any $i_0 \in \tilde{I}$ $i_0 \neq k_0$, but $\lambda_{k_0} = 0 = \lambda_{i_0}$.

2°. case: $\lambda_{k_0} \neq 0$. Then $\sum_{i \in \tilde{I}} \alpha_i x_i = x_{k_0} = \sum_{i \in \tilde{I}} \frac{\lambda_i}{\lambda_{k_0}} \alpha_i x_i$, i.e. $\forall_{i \in \tilde{I}} \frac{\lambda_i}{\lambda_{k_0}} = 1$. Thus again $\lambda_{i_0} = \lambda_{k_0}$ for some $\tilde{I} \ni i_0 \neq k_0$ – – a contradiction.

Hence, finally, $W' = W$, i.e, W is linearly independent. $\boxed{4}$

Now, for X – linear and $A \in \mathcal{L}(X)$ and for f – a $\mathbb{K}$-coefficients polynomial of the form

$$f(s) := \sum_{i=0}^{m} a_i s^i \tag{1}$$

where $m \in \mathbb{N}$, $a_0, \dots, a_m \in \mathbb{K}$, we define

$$f(A) := \sum_{i=0}^{m} a_i A^i \in \mathcal{L}(X). \tag{1'}$$

E.g., when $f(s) \equiv s^n$, then $f(A) = A^n$. Moreover, if $A \in B(X)$ when

X is a normed space, then $f(A) \in B(X)$.

One can easily "calculate" ($\longrightarrow$ △), that the operation $f \rightsquigarrow f(A)$ *) has the following properties:

$\boxed{\text{Fact}}$ ("On polynomial-functional calculus")

The mapping $Pol(\mathbb{K}) \ni f \rightsquigarrow f(A) \in \mathcal{L}(X)$ is a homeomorphism of the algebras $Pol(\mathbb{K})$ - of all the $\mathbb{K}$-coefficients polynomials and $\mathcal{L}(X)$, i.e.:

(i) it is a linear mapping

(ii) $\mathbb{1}(A) = I$, where $\mathbb{1}$ is the constant 1 polynomial

(iii) $\forall_{f, g \in Pol(\mathbb{K})} \quad (f \cdot g)(A) = f(A) \cdot g(A)$.

Observe that for any $\lambda \in \sigma_p(A)$ and any $f \in Pol(\mathbb{K})$ we have $f(\lambda) \in \sigma_p(f(A))$, because if $X \setminus \{0\} \ni x$ and

$$Ax = \lambda x$$

then for f given by (1)

$$f(A)x = \sum_{i=0}^{m} a_i A^i x = \sum_{i=0}^{m} a_i \lambda^i x = \left(\sum_{i=0}^{m} a_i \lambda^i\right) x = f(\lambda) x.$$

It turns out, that a related result holds for "the full" spectrum...

$\boxed{\text{Fact}}$ ("On the polynomial - spectrum calculus")

If $X \neq \{0\}$ is a $\mathbb{C}$-linear space and $A \in \mathcal{L}(X)$, then $\sigma(f(A)) = f(\sigma(A))$ and $f(\sigma_p(A)) \subset \sigma_p(f(A))$ for any $f \in Pol(\mathbb{C})$.

---

*) We call $f(A)$ the polynomial (or function) f of A.

Proof

The proof of the inclusion " $\subset$ " for σ_p was just made (and works also for $\mathbb{K}=\mathbb{R}$ ...).

Observe first, that for $f = c\mathbb{1}$, $c \in \mathbb{C}$ the assertion

$$\sigma(f(A)) = f(\sigma(A)) \qquad (2)$$

is obvious. Let $f \in \mathrm{Pol}(\mathbb{C})$, $\deg f \geq 1$ and let $\mu_0 \in \mathbb{C}$.

Denote $g := f - \mu_0 \mathbb{1}$. Hence $m := \deg g \geq 1$, too and by the Fundamental Algebra Theorem g has the form:

$$g = c\, g_1 \cdot \ldots \cdot g_m$$

where $0 \neq c \in \mathbb{C}$ and $g_j(s) = s - \lambda_j$ for any $j = 1, \ldots m$, with some $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ (depending on f and μ_0).

By Fact "On polynomial functional calculus" we have:

$$g(A) = c\, g_1(A) \cdot \ldots \cdot g_m(A)$$

and moreover the operators $g_1(A), \ldots g_m(A)$ commute. – Using this commuting one easily see that $g_1(A) \cdots g_m(A)$ is invertible iff each of $g_1(A), \ldots g_m(A)$ is invertible! ($\longrightarrow \triangle$). Hence $g(A)$ is not invertible iff for some $j = 1, \ldots n$ $g_j(A)$ is not invertible – which means that $\lambda_j \in \sigma(A)$, because $g_j(A) = A - \lambda_j I$. So, $\mu_0 \in \sigma(f(A))$ iff $g(A) = f(A) - \mu_0 I$ is not invertible iff $\exists_{\lambda \in \{z \in \mathbb{C} : g(z) = 0\}}\ \lambda \in \sigma(A)$ iff $\{z \in \mathbb{C} : f(z) = \mu_0\} \cap \sigma(A) \neq \emptyset$ iff $\mu_0 \in f(\sigma(A))$. ∎

$\boxed{\text{Theorem}}$ ("On the spectrum of compact operator", "The Riesz - Schauder Theorem")

Let $\{0\} \neq X$ be a Banach space and $K \in \mathcal{C}(X)$. Then:

(1) If $\dim X = +\infty$, then $0 \in \sigma(K)$.

(2) If $\lambda \in \sigma(K) \setminus \{0\}$, then $\lambda \in \sigma_p(K)$ and $\dim \operatorname{Ker}(K - \lambda I) < +\infty$.

(3) $\forall_{\varepsilon > 0}$ $\sigma(K) \cap \{\lambda \in \mathbb{C} : |\lambda| > \varepsilon\}$ is a finite set; in particular, if $\lambda \in \mathbb{C}$ is an accumulation point (limit point) of $\sigma_p(K)$, then $\lambda = 0$.

$\boxed{\text{Proof}}$

(1) Suppose, that $\dim X = +\infty$. If $0 \notin \sigma(K)$, then $\operatorname{Ran} K = X$, hence by thm "on properties of $\mathcal{C}(X,Y)$" part (iv) (p. OTST-2,1) $\operatorname{Ran} K$ is not closed, but X is closed! - Thus $0 \in \sigma(K)$.

(2) Suppose, that $\lambda \in \sigma(K) \setminus \{0\}$. If $\lambda \notin \sigma_p(K)$, then $\operatorname{Ker}(K - \lambda I) = \{0\}$. But $K - \lambda I = -\lambda(I - \frac{1}{\lambda}K)$ and $-\frac{1}{\lambda}K$ is also compact, $\operatorname{Ker}(K - \lambda I) = \operatorname{Ker}(I - \frac{1}{\lambda}K) = \{0\}$. Hence, by

$\boxed{\text{OTST} -\quad}$

Thm "On compact pert. ..." p. OTST-31, part 3 we have codim $\mathrm{Ran}(I - \frac{1}{\lambda}K) = 0$ $(= \dim\{0\})$. Hence (by the def. of codim) $\mathrm{Ran}(I - \frac{1}{\lambda}K) = X$, but $\mathrm{Ran}(K - \lambda I) = \mathrm{Ran}(I - \frac{1}{\lambda}K) = X$. Thus $K - \lambda I$ is invertible, and $\lambda \in \varrho(K)$, which contradicts $\lambda \in \sigma(K)$. Therefore $\lambda \in \sigma_p(K)$. Now, the eigenspace $\mathrm{Ker}(K - \lambda I) = \mathrm{Ker}(I - \frac{1}{\lambda}K)$, hence $\dim \mathrm{Ker}(K - \lambda I) < +\infty$ by theorem "On compact..." p OTST-31, part 1.

(3) . . . .

OTST-

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

$$\forall_{i\in\mathbb{I}} \; Ax_i = \lambda_i x_i . \qquad (1)$$

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

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

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 $\mathcal{H}$.

Proof

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 $A \in \mathcal{C}(\mathcal{H})$. [h]

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 $\sigma_p(A)$-valued function $\{\lambda_i\}_{i \in \mathbb{I}}$ 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 $\mathbb{I}_0 := \mathbb{I} \setminus \mathbb{I}_* \left(= \{ i \in \mathbb{I} : \lambda_i = 0 \} \right)$ 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 $M_{\{\lambda_i\}_{i\in\mathbb{I}}}$ – the multiplication operator by $\{\lambda_i\}_{i\in\mathbb{I}}$ in $\ell^2(\mathbb{I})$, in the sense, that

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

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

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

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

please – $\triangle$!: explain the details below, see p. HS-41 -44

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

$$A x = \sum_{i\in\mathbb{I}} (x, x_i) A x_i = \sum_{i=\mathbb{I}} (x, x_i)\lambda_i x_i ,$$

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

$\boxed{\text{OTST-}}$

Proof of the Hilbert-Schmidt thm

- For any $\mu \in \sigma_p(A)$ define $\mathcal{H}_\mu := \operatorname{Ker}(A-\mu I)$ – 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 $\{x_{(\mu,\alpha)}\}_{\alpha \in \mathcal{A}_\mu}$ be an o.n. base for $\mathcal{H}_\mu$ (note, that $\mathcal{A}_\mu$ is finite for any $\mu \neq 0$, however $\mathcal{A}_0$ can be "large", if $0 \in \sigma_p(A)$). Now choose $\mathbb{I} := \{(\mu,\alpha) : \mu \in \sigma_p(A), \alpha \in \mathcal{A}_\mu\}$ 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 $x_i \in \mathcal{H}_\mu = \operatorname{Ker}(A-\mu I)$. 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 $Y := \operatorname{lin}\{x_i\}_{i\in\mathbb{I}}$ (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 which is related to "the multiplicity" of μ as an eigenvalue. 2) the properly chosen $\alpha \in \mathcal{A}_\mu$

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

$$Ax = \tilde{A}x = \lambda x \text{ , i.e } x \in \mathrm{Ker}(A - \lambda I) \text{ and } \lambda \in \sigma_p(A).$$

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! $\boxed{\text{c.b.}}$

e