11 Spectrum of compact operators

We begin with some general results on the spectrum of continuous operators (11.1–11.5), where we always assume that X is a Banach space over $\mathbb{C}$ (!), i.e. $\mathbb{K} = \mathbb{C}$, and that $T \in \mathscr{L}(X)$ (for the real case see 11.14). The main topic of this chapter is the Riesz-Schauder theory on the spectrum of compact operators (theorem 11.9).

11.1 Spectrum. We define the ***resolvent set*** of T by

$$\varrho(T) := \left\{ \lambda \in \mathbb{C} \; ; \; \mathscr{N}(\lambda \mathrm{Id} - T) = \{0\} \;\text{ and }\; \mathscr{R}(\lambda \mathrm{Id} - T) = X \right\}$$

and the ***spectrum*** of T by

$$\sigma(T) := \mathbb{C} \setminus \varrho(T) \,.$$

The spectrum can be decomposed into the ***point spectrum***

$$\sigma_p(T) := \left\{ \lambda \in \sigma(T) \; ; \; \mathscr{N}(\lambda \mathrm{Id} - T) \neq \{0\} \right\} ,$$

the ***continuous spectrum***

$$\sigma_c(T) := \left\{ \lambda \in \sigma(T) \; ; \; \mathscr{N}(\lambda \mathrm{Id} - T) = \{0\} \;\text{ and }\; \mathscr{R}(\lambda \mathrm{Id} - T) \neq X, \;\text{ but }\; \overline{\mathscr{R}(\lambda \mathrm{Id} - T)} = X \right\}$$

and the ***residual spectrum***

$$\sigma_r(T) := \left\{ \lambda \in \sigma(T) \; ; \; \mathscr{N}(\lambda \mathrm{Id} - T) = \{0\} \;\text{ and }\; \overline{\mathscr{R}(\lambda \mathrm{Id} - T)} \neq X \right\} .$$

11.2 Remarks.

(1) It holds that $\lambda \in \varrho(T)$ if and only if $\lambda \mathrm{Id} - T : X \to X$ is bijective. The inverse mapping theorem 7.8 yields that this is equivalent to the existence of

$$R(\lambda; T) := (\lambda \mathrm{Id} - T)^{-1} \in \mathscr{L}(X) \,.$$

The inverse $R(\lambda; T)$ is called the ***resolvent*** of T in λ and as a function of λ is called the ***resolvent function***.

(2) $\lambda \in \sigma_p(T)$ is equivalent to:

$$\text{There exists an } x \neq 0 \text{ with } Tx = \lambda x.$$

H.W. Alt, *Linear Functional Analysis*, Universitext,
DOI 10.1007/978-1-4471-7280-2_11

λ is then called an ***eigenvalue*** and x an ***eigenvector*** of T. If X is a function space, then x is also called an ***eigenfunction***. The subspace $\mathscr{N}(\lambda\mathrm{Id} - T)$ is the ***eigenspace*** of T corresponding to the eigenvalue λ. The eigenspace is a T-invariant subspace. (A subspace $Y \subset X$ is called ***T-invariant*** if $T(Y) \subset Y$.)

11.3 Theorem. The resolvent set $\varrho(T)$ is open and the resolvent function $\lambda \mapsto R(\lambda; T)$ is a complex analytic map from $\varrho(T)$ to $\mathscr{L}(X)$. It holds that

$$\|R(\lambda; T)\|^{-1} \le \operatorname{dist}(\lambda, \sigma(T)) \quad \text{for } \lambda \in \varrho(T)\,.$$

Remark: A map $F : D \to Y$, with $D \subset \mathbb{C}$ open and Y a Banach space, is called ***complex analytic*** if for every $\lambda_0 \in D$ there exists a ball $\mathrm{B}_{r_0}(\lambda_0) \subset D$ such that $F(\lambda)$ for $\lambda \in \mathrm{B}_{r_0}(\lambda_0)$ can be written as a power series in $\lambda - \lambda_0$ with coefficients in Y. Complex analytic maps are holomorphic (see A10.1).

Proof. Let $\lambda \in \varrho(T)$. Then we have for $\mu \in \mathbb{C}$ that

$$(\lambda - \mu)\mathrm{Id} - T = (\lambda\mathrm{Id} - T)\underbrace{\big(\mathrm{Id} - \mu R(\lambda; T)\big)}_{=:\, S(\mu)}\,.$$

It follows from 5.7 that $S(\mu)$ is invertible if

$$|\mu| \cdot \|R(\lambda; T)\| < 1\,,$$

and then $\lambda - \mu \in \varrho(T)$, with

$$R(\lambda - \mu; T) = S(\mu)^{-1} R(\lambda; T) = \sum_{k=0}^{\infty} \mu^k R(\lambda; T)^{k+1}\,.$$

Setting $d := \|R(\lambda; T)\|^{-1}$ yields that $\mathrm{B}_d(\lambda) \subset \varrho(T)$, i.e. $\operatorname{dist}(\lambda, \sigma(T)) \ge d$. □

11.4 Theorem. The spectrum $\sigma(T)$ is compact and nonempty (if $X \ne \{0\}$), with

$$\sup_{\lambda \in \sigma(T)} |\lambda| = \lim_{m \to \infty} \|T^m\|^{\frac{1}{m}} \le \|T\|\,.$$

This value is called the ***spectral radius*** of T.

Proof. Let $\lambda \ne 0$. We have from 5.7 that $\mathrm{Id} - \frac{T}{\lambda}$ is invertible if $\left\|\frac{T}{\lambda}\right\| < 1$, i.e. if $|\lambda| > \|T\|$, and then

$$R(\lambda; T) = \frac{1}{\lambda}\Big(\mathrm{Id} - \frac{T}{\lambda}\Big)^{-1} = \sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}}\,.$$

This shows that

$$r := \sup_{\lambda \in \sigma(T)} |\lambda| \leq \|T\| \,.$$

Since

$$\lambda^m \mathrm{Id} - T^m = (\lambda \mathrm{Id} - T) p_m(T) = p_m(T)(\lambda \mathrm{Id} - T)$$

with

$$p_m(T) := \sum_{i=0}^{m-1} \lambda^{m-1-i} T^i \,,$$

we conclude that

$$\begin{aligned} \lambda \in \sigma(T) \quad &\Longrightarrow \quad \lambda^m \in \sigma(T^m) \\ &\Longrightarrow \quad |\lambda^m| \leq \|T^m\| \qquad \text{(recall the bound established above)} \\ &\Longrightarrow \quad |\lambda| \leq \|T^m\|^{\frac{1}{m}} \,. \end{aligned}$$

This proves that

$$r \leq \liminf_{m \to \infty} \|T^m\|^{\frac{1}{m}} \,.$$

Next we show that

$$r \geq \limsup_{m \to \infty} \|T^m\|^{\frac{1}{m}} \,.$$

We recall from 11.3 that $R(\bullet; T)$ is a complex analytic map in $\mathbb{C} \setminus \overline{\mathrm{B}_r(0)}$ (if $\sigma(T)$ is empty, in $\mathbb{C}$). Hence, by Cauchy's integral theorem (see A10.1),

$$\frac{1}{2\pi \mathrm{i}} \int_{\partial \mathrm{B}_s(0)} \lambda^j R(\lambda; T) \,\mathrm{d}\lambda$$

is independent of s for $j \geq 0$ and $s > r$. However, if we choose $s > \|T\|$, then we obtain with the help of the representation of $R(\lambda; T)$ at the beginning of the proof that this integral is equal to

$$\begin{aligned} = \frac{1}{2\pi \mathrm{i}} \int_{\partial \mathrm{B}_s(0)} \sum_{k=0}^{\infty} \lambda^{j-k-1} T^k \,\mathrm{d}\lambda &= \frac{1}{2\pi} \sum_{k=0}^{\infty} s^{j-k} \Big(\int_0^{2\pi} e^{\mathrm{i}\theta(j-k)} \,\mathrm{d}\theta \Big) T^k \\ &= \sum_{k=0}^{\infty} s^{j-k} \delta_{j,k} T^k = T^j \,. \end{aligned}$$

Hence, for $j \geq 0$ and $s > r$,

$$\|T^j\| = \frac{1}{2\pi} \left\| \int_{\partial \mathrm{B}_s(0)} \lambda^j R(\lambda; T) \,\mathrm{d}\lambda \right\| \leq s^{j+1} \sup_{|\lambda| = s} \|R(\lambda; T)\| \,.$$

Consequently we obtain for $s > r$ and every subsequence $j \to \infty$ that

$$\|T^j\|^{\frac{1}{j}} \leq s \cdot \left(s \sup_{|\lambda| = s} \|R(\lambda; T)\| \right)^{\frac{1}{j}} \longrightarrow s \text{ or } 0 \,,$$

and hence

$$\limsup_{j\to\infty} \left\|T^j\right\|^{\frac{1}{j}} \le s\,.$$

As this holds for all $s > r$, we obtain the desired result on the spectral radius. In addition, if $\sigma(T)$ was empty, we would obtain for $j = 0$ and as $s \searrow 0$ that

$$\|\mathrm{Id}\| \le s \cdot \sup_{|\lambda|\le 1} \|R(\lambda;T)\| \longrightarrow 0\,,$$

i.e. $\mathrm{Id} = 0$, and so $X = \{0\}$. □

11.5 Remarks.

(1) If $\dim X < \infty$, then $\sigma(T) = \sigma_p(T)$.

(2) If $\dim X = \infty$ and $T \in \mathscr{K}(X)$, then $0 \in \sigma(T)$. But in general 0 is not an eigenvalue.

Proof (1). If $\lambda \in \sigma(T)$, then $\lambda\mathrm{Id} - T$ is not bijective, and so, as $\dim X < \infty$, it is also not injective, i.e. $\lambda \in \sigma_p(T)$. □

Proof (2). Let $T \in \mathscr{K}(X)$ and assume that $0 \in \varrho(T)$. Then (see 11.2(1)) $T^{-1} \in \mathscr{L}(X)$, and (see 10.3) so $\mathrm{Id} = T^{-1}T \in \mathscr{K}(X)$, which on recalling 4.10 implies that X is finite-dimensional.

Example without eigenvalue 0*:* The operator $T : C^0([0,1]) \to C^1([0,1])$ in 5.6(3) is injective. As an operator in $\mathscr{L}\big(C^0([0,1])\big)$ it is a compact operator in $\mathscr{K}\big(C^0([0,1])\big)$, by theorem 10.6. □

In the following we are interested in the point spectrum $\sigma_p(T)$ of an operator $T \in \mathscr{L}(X)$, i.e. we consider the ***eigenvalue problem*** corresponding to T: For a given $y \in X$ we look for all solutions $\lambda \in \mathbb{K}$ and $x \in X$ to

$$Tx - \lambda x = y\,.$$

If $\lambda \in \varrho(T)$, then there exists a uniquely determined solution x to this equation. If $\lambda \in \sigma_p(T)$, then the solution, if one exists, is not unique, i.e. on setting $A_\lambda := \lambda\mathrm{Id} - T$ we see that adding an element from $\mathscr{N}(A_\lambda)$ to a solution yields another solution (for $T \in \mathscr{K}(X)$ see also 11.11). On the other hand, the condition $y \in \mathscr{R}(A_\lambda)$ needs to be satisfied for a solution to the eigenvalue problem to exist at all. An important class of operators A_λ are those operators for which both the number of degrees of freedom for the solution x and the number of side conditions on y are finite:

11.6 Fredholm operators. A map $A \in \mathscr{L}(X;Y)$ is called a ***Fredholm operator*** if:

(1) $\dim \mathscr{N}(A) < \infty$,

(2) $\mathscr{R}(A)$ is closed,

(3) $\operatorname{codim} \mathscr{R}(A) < \infty$.

The ***index*** of a Fredholm operator is defined by

$$\operatorname{ind}(A) := \dim \mathscr{N}(A) - \operatorname{codim} \mathscr{R}(A)\,.$$

Remark: The ***codimension*** of the image of A being finite means that $Y = \mathscr{R}(A) \oplus Y_0$ for a finite-dimensional subspace $Y_0 \subset Y$. Then $\operatorname{codim} \mathscr{R}(A) := \dim Y_0$ is independent of the choice of Y_0: indeed, if $Y_1 \subset Y$ is a subspace with $\mathscr{R}(A) \cap Y_1 = \{0\}$, then Y_1 is finite-dimensional with $\dim Y_1 \leq \dim Y_0$, with equality if and only if $Y = \mathscr{R}(A) \oplus Y_1$.

Proof. We have from (2) and 4.9, respectively, that $Z := \mathscr{R}(A)$ and Y_0 are closed subspaces. Now let $P \in \mathscr{P}(Y)$ be the projection onto Y_0 with $Z = \mathscr{N}(P)$, as in 9.15. Then

$$S := P\big|_{Y_1} : Y_1 \to Y_0 \text{ is linear and injective,}$$

because if $y \in Y_1$ with $P(y) = 0$, then $y \in Z \cap Y_1 = \{0\}$. As Y_0 is finite-dimensional, it follows that Y_1 is also finite-dimensional, with $\dim Y_1 \leq \dim Y_0$.

If $Y = Z \oplus Y_1$, then it follows as above (interchange Y_0 and Y_1) that $\dim Y_0 \leq \dim Y_1$, and so $\dim Y_1 = \dim Y_0$. Conversely, if this holds, then S is bijective. For $x \in Y$ we then have that $y := S^{-1}Px \in Y_1$ with $Py = SS^{-1}Px = Px$, and so $x - y \in \mathscr{N}(P) = Z$, which proves that $Y = Z \oplus Y_1$. □

11.7 Example. Let $X = W^{1,2}(\Omega)$ and $Y = W^{1,2}(\Omega)'$. Then $A : W^{1,2}(\Omega) \to W^{1,2}(\Omega)'$, defined by

$$\langle v\,,\, Au\rangle_{W^{1,2}} := \int_\Omega \sum_{i,j} \partial_i v \cdot a_{ij} \partial_j u \,\mathrm{dL}^n \qquad \text{for } u, v \in W^{1,2}(\Omega)\,,$$

is a weak elliptic differential operator with Neumann boundary conditions. (We consider the homogeneous case in 6.5(2) with $h_i = 0$ and $b = 0$.) We have from 8.18(2) (where the symmetry $a_{ij} = a_{ji}$ was assumed) that:

The null space $\mathscr{N}(A)$ consists of the constant functions, and therefore $\dim \mathscr{N}(A) = 1$. The image of A is $\mathscr{R}(A) = \{F \in Y\,;\ \langle 1\,,\, F\rangle_{W^{1,2}} = 0\}$, and so it is closed, with $\operatorname{codim} \mathscr{R}(A) = 1$. It holds that

$$Y = \mathscr{R}(A) \oplus \operatorname{span}\{F_0\}\,, \quad \text{where} \quad \langle v\,,\, F_0\rangle_{W^{1,2}} := \int_\Omega v \,\mathrm{dL}^n\,.$$

Hence A is a Fredholm operator with index 0.

Observe: For the homogeneous Dirichlet problem (see 10.14(2)) the operator $A : W_0^{1,2}(\Omega) \to W_0^{1,2}(\Omega)'$ is an isomorphism.

A large class of Fredholm operators with $Y = X$ is given by compact perturbations of the identity (see also 12.8):

11.8 Theorem. Let $T \in \mathscr{K}(X)$. Then $A := \mathrm{Id} - T$ is a Fredholm operator with index 0. We prove this in several steps:

(1) $\dim \mathscr{N}(A) < \infty$,
(2) $\mathscr{R}(A)$ is closed,
(3) $\mathscr{N}(A) = \{0\} \implies \mathscr{R}(A) = X$,
(4) $\operatorname{codim} \mathscr{R}(A) \le \dim \mathscr{N}(A)$,
(5) $\dim \mathscr{N}(A) \le \operatorname{codim} \mathscr{R}(A)$.

Proof (1). On noting that $Ax = 0$ is equivalent to $x = Tx$, we have that $\mathrm{B}_1(0) \cap \mathscr{N}(A) \subset T\big(\mathrm{B}_1(0)\big)$, i.e. the unit ball in $\mathscr{N}(A)$ is precompact, and so 4.10 yields that $\mathscr{N}(A)$ is finite-dimensional. □

Proof (2). Let $x \in \overline{\mathscr{R}(A)}$ and let $Ax_n \to x$ as $n \to \infty$. We may assume without loss of generality that

$$\|x_n\| \le 2\, d_n \qquad \text{with } d_n := \operatorname{dist}(x_n, \mathscr{N}(A)),$$

because otherwise we choose $a_n \in \mathscr{N}(A)$ with $\|x_n - a_n\| \le 2 \operatorname{dist}(x_n, \mathscr{N}(A))$ and then proceed with $\widetilde{x}_n := x_n - a_n$, where

$$\operatorname{dist}(\widetilde{x}_n, \mathscr{N}(A)) = \operatorname{dist}(x_n, \mathscr{N}(A))\,.$$

First we assume that $d_n \to \infty$ for a subsequence $n \to \infty$. Setting

$$y_n := \frac{x_n}{d_n} \qquad \text{it holds that} \qquad Ay_n = \frac{Ax_n}{d_n} \to 0$$

as $n \to \infty$. Noting that the y_n are bounded and recalling that T is compact yields that there exists a subsequence such that $Ty_n \to y$ as $n \to \infty$. It follows that

$$y_n = Ay_n + Ty_n \to y\,,$$

and so, by the continuity of A,

$$Ay = \lim_{n\to\infty} Ay_n = 0\,.$$

Hence $y \in \mathscr{N}(A)$, which implies that

$$\|y_n - y\| \ge \operatorname{dist}(y_n, \mathscr{N}(A)) = \operatorname{dist}\Big(\frac{x_n}{d_n}, \mathscr{N}(A)\Big) = \frac{\operatorname{dist}(x_n, \mathscr{N}(A))}{d_n} = 1\,,$$

a contradiction. This shows that the d_n are bounded, and so are the x_n. For a subsequence we then have that $Tx_n \to z$ as $n \to \infty$, and so

$$x \longleftarrow Ax_n = A(Ax_n + Tx_n) \longrightarrow A(x+z)\,,$$

which means that $x \in \mathscr{R}(A)$. □

Proof (3). Assume that there exists an $x \in X \setminus \mathscr{R}(A)$. Then

$$A^n x \in \mathscr{R}(A^n) \setminus \mathscr{R}(A^{n+1}) \quad \text{for all } n \geq 0,$$

because otherwise $A^n x = A^{n+1} y$ for some y, then $A^n(x - Ay) = 0$, and from $\mathscr{N}(A) = \{0\}$ it then would follow (inductively) that $x - Ay = 0$, i.e $x \in \mathscr{R}(A)$, a contradiction. In addition $\mathscr{R}(A^{n+1})$ is closed, on noting that

$$A^{n+1} = (\mathrm{Id} - T)^{n+1} = \mathrm{Id} + \underbrace{\sum_{k=1}^{n+1} \binom{n+1}{k} (-T)^k}_{\in \mathscr{K}(X) \text{ on recalling 10.3}} ,$$

and so (2) yields that $\mathscr{R}(A^{n+1})$ is closed. Hence there exists an $a_{n+1} \in \mathscr{R}(A^{n+1})$ with

$$0 < \| A^n x - a_{n+1} \| \leq 2 \operatorname{dist}\big(A^n x, \mathscr{R}(A^{n+1})\big) .$$

Now consider

$$x_n := \frac{A^n x - a_{n+1}}{\| A^n x - a_{n+1} \|} \in \mathscr{R}(A^n) .$$

We have that

$$\operatorname{dist}\big(x_n, \mathscr{R}(A^{n+1})\big) \geq \tfrac{1}{2} , \tag{11-4}$$

because for $y \in \mathscr{R}(A^{n+1})$

$$\begin{aligned} \| x_n - y \| &= \frac{\| A^n x - (a_{n+1} + \| A^n x - a_{n+1} \| y) \|}{\| A^n x - a_{n+1} \|} \\ &\geq \frac{\operatorname{dist}\big(A^n x, \mathscr{R}(A^{n+1})\big)}{\| A^n x - a_{n+1} \|} \geq \frac{1}{2} . \end{aligned}$$

For $m > n$, we have $Ax_n + x_m - Ax_m \in \mathscr{R}(A^{n+1})$, and hence (11-4) implies that

$$\| Tx_n - Tx_m \| = \| x_n - (Ax_n + x_m - Ax_m) \| \geq \frac{1}{2} .$$

Hence $(Tx_n)_{n\in\mathbb{N}}$ contains no convergent subsequence, even though $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence. This is a contradiction to the compactness of T. □

Proof (4). By (1), the number $n := \dim \mathscr{N}(A)$ is finite. Let $\{x_1, \ldots, x_n\}$ be an arbitrary basis of $\mathscr{N}(A)$. If we assume that the claimed inequality is false, then there exist linearly independent vectors $y_1, \ldots, y_n$, such that $\operatorname{span}\{y_1, \ldots, y_n\} \oplus \mathscr{R}(A)$ is a proper subspace of X. Moreover, 9.16(1) yields the existence of $x'_1, \ldots, x'_n \in X'$ with

$$\langle x_l \, , \, x'_k \rangle = \delta_{k,l} \quad \text{for } k, l = 1, \ldots, n.$$

Setting

$$\widetilde{T}x := Tx + \sum_{k=1}^{n} \langle x\,,\,x_k'\rangle\,y_k$$

then defines an operator $\widetilde{T} \in \mathscr{K}(X)$, indeed, T is compact and $\widetilde{T} - T$ has a finite-dimensional image. In addition, $\mathscr{N}(\widetilde{A}) = \{0\}$, where $\widetilde{A} := \mathrm{Id} - \widetilde{T}$, because $\widetilde{A}x = 0$ implies, on recalling the choice of the y_k, that $Ax = 0$ and $\langle x\,,\,x_k'\rangle = 0$ for $k = 1,\ldots,n$. Therefore $x \in \mathscr{N}(A)$, and hence there exists a representation

$$x = \sum_{k=1}^{n} \alpha_k x_k\,, \quad \text{and so} \quad 0 = \langle x\,,\,x_l'\rangle = \sum_{k=1}^{n} \alpha_k\,\langle x_k\,,\,x_l'\rangle = \alpha_l$$

for $l = 1,\ldots,n$, which yields that $x = 0$. On applying (3) to the operator $\widetilde{A}$, it follows that $\mathscr{R}(\widetilde{A}) = X$. On noting that $\widetilde{A}x_l = -y_l$ for $l = 1,\ldots,n$ and that

$$\widetilde{A}\Big(x - \sum_{l=1}^{n} \langle x\,,\,x_l'\rangle\,x_l\Big) = Ax \qquad \text{for all } x \in X\,,$$

we conclude that $X = \mathscr{R}(\widetilde{A}) \subset \mathrm{span}\{y_1,\ldots,y_n\} \oplus \mathscr{R}(A)$, a contradiction to the above property. ☐

Proof (5). We have from (4) that $m := \mathrm{codim}\,\mathscr{R}(A) \le n := \dim \mathscr{N}(A)$.

First we reduce the claim to the case $m = 0$. To this end, choose $x_1,\ldots,x_n$ and $x_1',\ldots,x_n'$ as in the proof of (4) and $y_1,\ldots,y_m$ with

$$X = \mathrm{span}\{y_1,\ldots,y_m\} \oplus \mathscr{R}(A)\,.$$

As in the proof of (4), the operator

$$x \longmapsto \widetilde{T}x := Tx + \sum_{k=1}^{m} \langle x\,,\,x_k'\rangle\,y_k$$

is compact and $\widetilde{A} := \mathrm{Id} - \widetilde{T}$ is surjective with $\mathscr{N}(\widetilde{A}) = \mathrm{span}\{x_i\,;\ m < i \le n\}$. We need to show that $\mathscr{N}(\widetilde{A}) = \{0\}$. Hence the claim is reduced to the case $m = 0$.

In the case $m = 0$ it holds that $\mathscr{R}(A) = X$. We assume that there exists an $x_1 \in \mathscr{N}(A) \setminus \{0\}$. The surjectivity of A then yields that we can inductively choose $x_k \in X$, $k \ge 2$, with $Ax_k = x_{k-1}$. Then $x_k \in \mathscr{N}(A^k) \setminus \mathscr{N}(A^{k-1})$. It follows from the theorem on the almost orthogonal element that there exists a $z_k \in \mathscr{N}(A^k)$ with $\|z_k\| = 1$ and $\mathrm{dist}\big(z_k, \mathscr{N}(A^{k-1})\big) \ge \frac{1}{2}$. For $l < k$ this implies that $Az_k + z_l - Az_l \in \mathscr{N}(A^{k-1})$, and so the choice of z_k yields that

$$\|Tz_k - Tz_l\| = \|z_k - (Az_k + z_l - Az_l)\| \ge \tfrac{1}{2}\,.$$

This shows that $\{Tz_k\,;\ k \in \mathbb{N}\}$ contains no convergent subsequence. This is a contradiction to the sequence $\{z_k\,;\ k \in \mathbb{N}\}$ being bounded and the operator T being compact.

A second possible proof for $m = 0$ is as follows: We start with a decomposition $X = \widetilde{X} \oplus \mathscr{N}(A)$ with a closed subspace $\widetilde{X}$ (this follows from (1) and 9.16(2) for $Y = \{0\}$). Then $A : \widetilde{X} \to X$ is bijective, and so 7.8 yields that $\widetilde{A} := (A|_{\widetilde{X}})^{-1} : X \to \widetilde{X}$ is continuous. Now consider $\widetilde{A}$ as an element in $\mathscr{L}(X)$. Then $\widetilde{T} := \mathrm{Id} - \widetilde{A} \in \mathscr{K}(X)$, because if $\{x_k\,;\ k \in \mathbb{N}\}$ is bounded in X, then so is $\{\widetilde{A}x_k\,;\ k \in \mathbb{N}\}$, and hence there exists a subsequence with $T\widetilde{A}x_k \to x$ as $k \to \infty$. On the other hand,

$$T\widetilde{A}x_k = (\mathrm{Id} - A)\widetilde{A}x_k = \widetilde{A}x_k - x_k = -\widetilde{T}x_k\,.$$

Now (3) implies that $\mathscr{R}(\widetilde{A}) = X$, i.e. $\mathscr{N}(A) = \{0\}$.

A further possible proof of (5) will be given in 12.7. □

The fundamental theorem of this chapter is the

11.9 Spectral theorem for compact operators (Riesz-Schauder). For every operator $T \in \mathscr{K}(X)$ it holds that:

(1) The set $\sigma(T) \setminus \{0\}$ consists of countably (finitely or infinitely) many eigenvalues with 0 as the only possible cluster point. So if $\sigma(T)$ contains infinitely many elements, then $\overline{\sigma(T)} = \sigma_p(T) \cup \{0\}$, hence 0 is a cluster point of $\sigma(T)$.

(2) For $\lambda \in \sigma(T) \setminus \{0\}$

$$1 \le n_\lambda := \max\big\{\, n \in \mathbb{N}\ ;\ \mathscr{N}\big((\lambda\mathrm{Id} - T)^{n-1}\big) \neq \mathscr{N}\big((\lambda\mathrm{Id} - T)^{n}\big)\ \big\} < \infty\,.$$

The number $n_\lambda \in \mathbb{N}$ is called the ***order*** (or ***index***) of λ and $\dim \mathscr{N}(\lambda\mathrm{Id} - T)$ is called the ***multiplicity*** of λ.

(3) ***Riesz decomposition.*** For $\lambda \in \sigma(T) \setminus \{0\}$

$$X = \mathscr{N}\big((\lambda\mathrm{Id} - T)^{n_\lambda}\big) \oplus \mathscr{R}\big((\lambda\mathrm{Id} - T)^{n_\lambda}\big)\,.$$

Both subspaces are closed and T-invariant, and the ***characteristic subspace*** $\mathscr{N}\big((\lambda\mathrm{Id} - T)^{n_\lambda}\big)$ is finite-dimensional.

(4) For $\lambda \in \sigma(T) \setminus \{0\}$ it holds that $\sigma(T|_{\mathscr{R}((\lambda\mathrm{Id}-T)^{n_\lambda})}) = \sigma(T) \setminus \{\lambda\}\,.$

(5) If E_λ for $\lambda \in \sigma(T) \setminus \{0\}$ denotes the projection onto $\mathscr{N}\big((\lambda\mathrm{Id} - T)^{n_\lambda}\big)$ corresponding to the decomposition in (3), then

$$E_\lambda E_\mu = \delta_{\lambda,\mu} E_\lambda \quad \text{for } \lambda, \mu \in \sigma(T) \setminus \{0\}.$$

Proof (1). Let $0 \neq \lambda \notin \sigma_p(T)$. Then $\mathscr{N}(\mathrm{Id} - \frac{T}{\lambda}) = \{0\}$, and so $\mathscr{R}(\mathrm{Id} - \frac{T}{\lambda}) = X$ (recall 11.8(3)), i.e. $\lambda \in \varrho(T)$. This shows that

$$\sigma(T) \setminus \{0\} \subset \sigma_p(T)\,.$$

If $\sigma(T) \setminus \{0\}$ is not finite, then we choose $\lambda_n \in \sigma(T) \setminus \{0\}$, $n \in \mathbb{N}$, pairwise distinct and eigenvectors $e_n \neq 0$ to λ_n and define

$$X_n := \operatorname{span}\{e_1, \ldots, e_n\}\,.$$

The eigenvectors e_k, $k = 1, \ldots, n$, are linearly independent, because if there exists (this is an inductive proof) $1 < k \leq n$ with

$$e_k = \sum_{i=1}^{k-1} \alpha_i e_i$$

with already linearly independent vectors $e_1, \ldots, e_{k-1}$, then it follows that

$$0 = Te_k - \lambda_k e_k = \sum_{i=1}^{k-1} \alpha_i (Te_i - \lambda_k e_i) = \sum_{i=1}^{k-1} \alpha_i \underbrace{(\lambda_i - \lambda_k)}_{\neq 0} e_i\,,$$

and so $\alpha_i = 0$ for $i = 1, \ldots, k-1$, i.e. $e_k = 0$, a contradiction. This shows that X_{n-1} is a proper subspace of X_n. Hence the theorem on the almost orthogonal element (see 4.5) yields the existence of an $x_n \in X_n$ with

$$\|x_n\| = 1 \quad \text{and} \quad \operatorname{dist}(x_n, X_{n-1}) \geq \tfrac{1}{2}\,. \tag{11-5}$$

On noting that $x_n = \alpha_n e_n + \widetilde{x}_n$ with certain $\alpha_n \in \mathbb{C}$ and $\widetilde{x}_n \in X_{n-1}$, it follows from the T-invariance of the subspace X_{n-1} that $Tx_n - \lambda_n x_n = T\widetilde{x}_n - \lambda_n \widetilde{x}_n \in X_{n-1}$, and so it holds for $m < n$ that

$$\frac{1}{\lambda_n}(Tx_n - \lambda_n x_n) - \frac{1}{\lambda_m} Tx_m \in X_{n-1}\,.$$

Hence it follows from (11-5) that

$$\left\| T\Big(\frac{x_n}{\lambda_n}\Big) - T\Big(\frac{x_m}{\lambda_m}\Big) \right\| = \left\| x_n + \frac{1}{\lambda_n}(Tx_n - \lambda_n x_n) - \frac{1}{\lambda_m} Tx_m \right\| \geq \frac{1}{2}\,.$$

This shows that the sequence $\left(T\big(\frac{x_n}{\lambda_n}\big)\right)_{n\in\mathbb{N}}$ has no cluster point. As T is compact, this implies that $\left(\frac{x_n}{\lambda_n}\right)_{n\in\mathbb{N}}$ contains no bounded subsequences, which yields that

$$\frac{1}{|\lambda_n|} = \left\| \frac{x_n}{\lambda_n} \right\| \longrightarrow \infty \qquad \text{as } n \to \infty,$$

i.e. $\lambda_n \to 0$ as $n \to \infty$. Hence we have shown that 0 is the only cluster point of $\sigma(T) \setminus \{0\}$. In particular, it then holds that $\sigma(T) \setminus \mathrm{B}_r(0)$ is finite for every $r > 0$, and so $\sigma(T) \setminus \{0\}$ is countable. □

Proof (2). Let $A := \lambda\mathrm{Id} - T$. Then $\mathscr{N}(A^{n-1}) \subset \mathscr{N}(A^n)$ for all n. First we assume that:

$$\mathscr{N}(A^{n-1}) \text{ is a proper subset of } \mathscr{N}(A^n) \text{ for all } n \geq 1.$$

Similarly to the proof of (1), and on recalling the theorem on the almost orthogonal element, we choose an $x_n \in \mathscr{N}(A^n)$ with

$$\|x_n\| = 1 \quad \text{and} \quad \operatorname{dist}\bigl(x_n, \mathscr{N}(A^{n-1})\bigr) \geq \tfrac{1}{2}\,. \tag{11-6}$$

Then it follows for $m < n$ that

$$Ax_n + \lambda x_m - Ax_m \in \mathscr{N}(A^{n-1})\,,$$

and so with (11-6) that

$$\|Tx_n - Tx_m\| = \|\lambda x_n - (Ax_n + \lambda x_m - Ax_m)\| \geq \frac{|\lambda|}{2} > 0\,.$$

On the other hand, $\{x_n\,;\ n \in \mathbb{N}\}$ is a bounded sequence. This contradicts the compactness of T. Hence we can find an $n \in \mathbb{N}$ with $\mathscr{N}(A^{n-1}) = \mathscr{N}(A^n)$. This implies for $m > n$ that

$$\begin{aligned} x \in \mathscr{N}(A^m) \quad &\Longrightarrow \quad A^{m-n}x \in \mathscr{N}(A^n) = \mathscr{N}(A^{n-1}) \\ &\Longrightarrow \quad A^{n-1+m-n}x = 0 \\ &\Longrightarrow \quad x \in \mathscr{N}(A^{m-1})\,, \end{aligned}$$

and so $\mathscr{N}(A^m) = \mathscr{N}(A^{m-1})$, and it follows inductively that $\mathscr{N}(A^m) = \mathscr{N}(A^n)$ for all $m \geq n$. Hence we have shown that $n_\lambda < \infty$. Since $\mathscr{N}(A) \neq \{0\}$ it holds that $n_\lambda \geq 1$. □

Proof (3). Let $A := \lambda\mathrm{Id} - T$ as before. Then

$$\mathscr{N}(A^{n_\lambda}) \oplus \mathscr{R}(A^{n_\lambda}) \subset X\,,$$

because if $x \in \mathscr{N}(A^{n_\lambda}) \cap \mathscr{R}(A^{n_\lambda})$, then $A^{n_\lambda}x = 0$ and $x = A^{n_\lambda}y$ for a $y \in X$. Then $A^{2n_\lambda}y = 0$, and so $y \in \mathscr{N}(A^{2n_\lambda}) = \mathscr{N}(A^{n_\lambda})$ and hence $x = A^{n_\lambda}y = 0$. Now A^{n_λ} can be written as

$$A^{n_\lambda} = \lambda^{n_\lambda}\mathrm{Id} + \underbrace{\sum_{k=1}^{n_\lambda} \binom{n_\lambda}{k} \lambda^{n_\lambda - k}(-T)^k}_{\in\, \mathscr{K}(X) \text{ by } 10.3}\,. \tag{11-7}$$

Hence $\operatorname{codim} \mathscr{R}(A^{n_\lambda}) \leq \dim \mathscr{N}(A^{n_\lambda}) < \infty$ (recall 11.8(4) and 11.8(1)), which yields that

$$X = \mathscr{N}(A^{n_\lambda}) \oplus \mathscr{R}(A^{n_\lambda})\,.$$

As T commutes with A, i.e. $TA = AT$, T also commutes with A^{n_λ}, and so both subspaces are T-invariant. □

Proof (4). We denote by T_λ the restriction of T to $\mathscr{R}(A^{n_\lambda})$, where A^{n_λ} has been computed in (11-7). Then $T_\lambda \in \mathscr{K}\big(\mathscr{R}(A^{n_\lambda})\big)$, where $\mathscr{R}(A^{n_\lambda})$ is a closed subspace (recall 11.8(2)), and so a Banach space. Here we have used the fact that T and A^{n_λ} commute. Moreover, we have that

$$\mathscr{N}(\lambda\mathrm{Id} - T_\lambda) = \mathscr{N}(A) \cap \mathscr{R}(A^{n_\lambda}) = \{0\}\,,$$

and hence $\mathscr{R}(\lambda\mathrm{Id} - T_\lambda) = \mathscr{R}(A^{n_\lambda})$ (apply 11.8(3) to T_λ), which shows that $\lambda \in \varrho(T_\lambda)$. It remains to show that

$$\sigma(T_\lambda) \setminus \{\lambda\} = \sigma(T) \setminus \{\lambda\}\,.$$

Let $\mu \in \mathbb{C} \setminus \{\lambda\}$. We recall from above that $\mathscr{N}(A^{n_\lambda})$ is invariant under $\mu\mathrm{Id} - T$. Moreover, $\mu\mathrm{Id} - T$ is injective on this subspace. To see this, note that $x \in \mathscr{N}(\mu\mathrm{Id} - T)$ implies that $(\lambda - \mu)x = Ax$. If in addition $A^m x = 0$ for some $m \geq 1$, it follows that

$$(\lambda - \mu)A^{m-1}x = A^{m-1}((\lambda - \mu)x) = A^m x = 0\,,$$

and since $\lambda \neq \mu$ this means that $A^{m-1}x = 0$. Inductively (for decreasing m) this yields that $x = A^0 x = 0$. Hence we have shown that

$$\mathscr{N}(\mu\mathrm{Id} - T) \cap \mathscr{N}(A^m) = \{0\} \quad \text{for all } m \geq 1.$$

Setting $m = n_\lambda$ yields the injectivity of $\mu\mathrm{Id} - T$ on $\mathscr{N}(A^{n_\lambda})$. As this space is finite-dimensional, we have that $\mu\mathrm{Id} - T$ is also bijective on $\mathscr{N}(A^{n_\lambda})$. But this means that $\mu \in \varrho(T)$ if and only if $\mu \in \varrho(T_\lambda)$. This shows that by removing the (finite-dimensional) characteristic subspace corresponding to the eigenvalue λ we obtain a remaining operator T_λ for which $\sigma(T_\lambda) = \sigma(T) \setminus \{\lambda\}$. □

Proof (5). Let $\lambda, \mu \in \sigma(T) \setminus \{0\}$ be distinct, and let $A_\lambda := \lambda\mathrm{Id} - T$ and $A_\mu := \mu\mathrm{Id} - T$. Now every $x \in \mathscr{N}(A_\mu^{n_\mu})$, corresponding to the Riesz decomposition of X into $\mathscr{N}(A_\lambda^{n_\lambda}) \oplus \mathscr{R}(A_\lambda^{n_\lambda})$, has a representation $x = z + y$. As both subspaces are invariant under T, and hence also under A_μ, it follows that

$$0 = A_\mu^{n_\mu} x = \underbrace{A_\mu^{n_\mu} z}_{\in \mathscr{N}(A_\lambda^{n_\lambda})} + \underbrace{A_\mu^{n_\mu} y}_{\in \mathscr{R}(A_\lambda^{n_\lambda})}$$

and so $0 = A_\mu^{n_\mu} z$. On recalling from the above proof that A_μ is bijective on $\mathscr{N}(A_\lambda^{n_\lambda})$, and hence also $A_\mu^{n_\mu}$, it follows that $z = 0$, i.e. $x \in \mathscr{R}(A_\lambda^{n_\lambda})$. Therefore we have shown that

$$\mathscr{N}(A_\mu^{n_\mu}) \subset \mathscr{R}(A_\lambda^{n_\lambda})\,,$$

in other words

$$\mathscr{R}(E_\mu) \subset \mathscr{N}(E_\lambda)\,,$$

and hence $E_\lambda E_\mu = 0$. □

11.10 Corollary. If $T \in \mathscr{K}(X)$ and $\lambda \in \sigma(T) \setminus \{0\}$, then the resolvent function $\mu \mapsto R(\mu;T)$ has an ***(isolated) pole*** of ***order*** n_λ in λ, i.e. the function $\mu \mapsto (\mu-\lambda)^{n_\lambda} R(\mu;T)$ can be complex analytically extended to the point λ, and the value at the point λ is different from the null operator.

Proof. Consider the decomposition

$$X = \underbrace{\mathscr{N}\big((\lambda \mathrm{Id} - T)^{n_\lambda}\big)}_{=\mathscr{R}(E_\lambda)} \oplus \underbrace{\mathscr{R}\big((\lambda \mathrm{Id} - T)^{n_\lambda}\big)}_{=\mathscr{N}(E_\lambda)}$$

and the restrictions

$$T_0 := T \quad \text{to } \mathscr{R}(E_\lambda), \qquad T_1 := T \quad \text{to } \mathscr{N}(E_\lambda).$$

Since λ is an isolated point of $\sigma(T)$, there exists an $r > 0$ with $\mathrm{B}_r(\lambda) \setminus \{\lambda\} \subset \varrho(T)$. Then $\mathrm{B}_r(\lambda) \setminus \{\lambda\} \subset \varrho(T_0)$ and we have from 11.9(4) that $\mathrm{B}_r(\lambda) \subset \varrho(T_1)$, and it holds for $0 < |\mu| < r$ that

$$R(\lambda+\mu;T) = R(\lambda+\mu;T_0)E_\lambda + R(\lambda+\mu;T_1)(\mathrm{Id} - E_\lambda)\,.$$

It follows from 11.3 that $R(\lambda + \bullet\,; T_1)$ is complex analytic in $\mathrm{B}_r(0)$, and so it remains to show that $R(\lambda + \bullet\,; T_0)$ has a pole of order n_λ in 0. Consider

$$S(\mu) := \sum_{k=1}^{n_\lambda} \mu^{-k}(T_0 - \lambda \mathrm{Id})^{k-1} \qquad \text{for } \mu \neq 0.$$

It holds that

$$\begin{aligned} S(\mu)\big((\lambda+\mu)\mathrm{Id} - T_0\big) &= \sum_{k=1}^{n_\lambda} \mu^{1-k}(T_0 - \lambda \mathrm{Id})^{k-1} - \sum_{k=1}^{n_\lambda} \mu^{-k}(T_0 - \lambda \mathrm{Id})^{k} \\ &= \mathrm{Id} - \mu^{-n_\lambda}(T_0 - \lambda \mathrm{Id})^{n_\lambda} = \mathrm{Id} \end{aligned}$$

and similarly $\big((\lambda+\mu)\mathrm{Id} - T_0\big)S(\mu) = \mathrm{Id}$, i.e. $R(\lambda+\mu;T_0) = S(\mu)$. □

The assertion $\sigma(T) \setminus \{0\} \subset \sigma_p(T)$ in 11.9(1) can also be formulated as follows:

11.11 Fredholm alternative. If $T \in \mathscr{K}(X)$ and $\lambda \neq 0$, then it holds that:

Either the equation $Tx - \lambda x = y$ is uniquely solvable for every $y \in X$,

or the equation $Tx - \lambda x = 0$ has nontrivial solutions.

Note: See also theorem 12.8.

11.12 Finite-dimensional case. Let X be a finite-dimensional vector space over $\mathbb{C}$ and let $T : X \to X$ be linear. Then there exist pairwise distinct $\lambda_1, \dots, \lambda_m \in \mathbb{C}$, where $1 \le m \le \dim X$, such that

$$\sigma(T) = \sigma_p(T) = \{\lambda_1, \dots, \lambda_m\}\,,$$

and orders n_{λ_j} with the properties in 11.9(2) – 11.9(5), so that

$$X = \mathscr{N}\big((\lambda_1 \mathrm{Id} - T)^{n_{\lambda_1}}\big) \oplus \cdots \oplus \mathscr{N}\big((\lambda_m \mathrm{Id} - T)^{n_{\lambda_m}}\big).$$

Proof. We equip X with an arbitrary norm. Then $T \in \mathscr{K}(X)$ (see 10.2(3)), and similarly $T_\mu := T - \mu\mathrm{Id}$ for $\mu \in \mathbb{C}$. Now apply 11.9 to e.g. T_0 and T_1. □

11.13 Jordan normal form. Let $T \in \mathscr{K}(X)$ and let $\lambda \in \sigma_p(T)$ be as in 11.9 or 11.12, respectively. Set $A := \lambda\mathrm{Id} - T$. Then:

(1) For $n = 1, \dots, n_\lambda$ there exist subspaces E_n with $\mathscr{N}(A^{n-1}) \oplus E_n \subset \mathscr{N}(A^n)$ such that

$$\mathscr{N}\left(A^{n_\lambda}\right) = \bigoplus_{k=1}^{n_\lambda} N_k \ , \quad \text{where} \quad N_k := \bigoplus_{l=0}^{k-1} A^l(E_k) \ .$$

(2) The subspaces N_k, $k = 1, \dots, n_\lambda$, are T-invariant and the dimensions $d_k := \dim A^l(E_k)$ are independent of $l \in \{0, \dots, k-1\}$.

(3) If $\{e_{k,j} \ ; \ j = 1, \dots, d_k\}$ are bases of E_k, then

$$\{A^l e_{k,j} \ ; \ 0 \le l < k \le n_j, \ 1 \le j \le d_k\}$$

is a basis of $\mathscr{N}(A^{n_\lambda})$ and with

$$x = \sum_{k,j,l} \alpha_{k,j,l} \ A^l e_{k,j} \quad \text{and} \quad y = \sum_{k,j,l} \beta_{k,j,l} \ A^l e_{k,j}$$

it holds that $Tx = y$ is equivalent to

$$\begin{bmatrix} \beta_{k,j,0} \\ \vdots \\ \vdots \\ \beta_{k,j,k-1} \end{bmatrix} = \begin{bmatrix} \lambda & -1 & & 0 \\ 0 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & 0 & \lambda \end{bmatrix} \begin{bmatrix} \alpha_{k,j,0} \\ \vdots \\ \vdots \\ \alpha_{k,j,k-1} \end{bmatrix},$$

i.e. the matrix representing T with respect to this basis has a ***Jordan normal form***.

Proof. If E is a subspace with $\mathscr{N}(A^{n-1}) \oplus E \subset \mathscr{N}(A^n)$, then

$$\mathscr{N}(A^{n-l-1}) \oplus A^l(E) \subset \mathscr{N}(A^{n-l}) \quad \text{for } 0 \le l < n,$$

and A^l is injective on E. To see this, note that if $x \in E$ with $A^l x = 0$, then also $A^{n-1}x = 0$ because $l \le n-1$, and so $x \in \mathscr{N}(A^{n-1}) \cap E = \{0\}$. Based on this observation we inductively choose E_n for $n = n_\lambda, \dots, 1$ such that

$$\mathscr{N}(A^n) = \mathscr{N}(A^{n-1}) \oplus \bigoplus_{l=0}^{n_\lambda - n} A^l(E_{n+l}) \,.$$

This yields the desired results. □

11.14 Real case. If X is a Banach space over $\mathbb{R}$ and if $T \in \mathscr{K}(X)$, then the spectral theorem can be applied to their ***complexification***, i.e. let

$$\widetilde{X} := X \times X$$

and for $x = (x_1, x_2) \in \widetilde{X}$, $\alpha = a + \mathrm{i}b$ with $a, b \in \mathbb{R}$, let

$$\alpha x := (ax_1 - bx_2, ax_2 + bx_1)\,, \quad \overline{x} := (x_1, -x_2)\,.$$

With the above $\widetilde{X}$ becomes a vector space over $\mathbb{C}$. On setting

$$\|x\|_{\widetilde{X}} := \sup_{\theta' \in \mathbb{R}} \Big(\|\cos(\theta')x_1 - \sin(\theta')x_2\|_X^2 + \|\sin(\theta')x_1 + \cos(\theta')x_2\|_X^2 \Big)^{\frac{1}{2}}$$

it holds that $\left\|\mathrm{e}^{\mathrm{i}\theta}x\right\|_{\widetilde{X}} = \|x\|_{\widetilde{X}}$ for $x \in \widetilde{X}$ and $\theta \in \mathbb{R}$, and equipped with this norm $\widetilde{X}$ becomes a Banach space over $\mathbb{C}$. Then

$$\widetilde{T}x := (Tx_1, Tx_2)$$

defines the corresponding operator $\widetilde{T} \in \mathscr{K}(\widetilde{X})$, so that theorem 11.9 can now be applied.

Now if $\lambda \in \sigma_p(\widetilde{T})$ with eigenvector e, then

$$\widetilde{T}\overline{e} = \overline{\widetilde{T}e} = \overline{\lambda e} = \overline{\lambda}\,\overline{e}\,,$$

and so $\overline{\lambda} \in \sigma_p(\widetilde{T})$ with eigenvector $\overline{e}$. If $\lambda \in \mathbb{R}$, then the vectors $e_{k,j}$ in 11.13(3) can be chosen to satisfy $\overline{e_{k,j}} = e_{k,j}$. If $\lambda \notin \mathbb{R}$ and $e_{k,j}$ as in 11.13(3), then the vectors $\overline{e_{k,j}}$ have the properties in 11.13(3) with respect to $\overline{\lambda}$.
Remark: In the case when X is a Hilbert space, the above norm satisfies

$$\|x\|_{\widetilde{X}} = \Big(\|x_1\|_X^2 + \|x_2\|_X^2\Big)^{\frac{1}{2}}.$$