1 The spectral theorem for compact and selfadjoint operators

Let H be a separable Hilbert space, and let $T = T^* \in B(H)$ be a selfadjoint, compact operator. This means that given any bounded sequence $\{x_n\}_{n\geq 1}$, one can always find a convergent subsequence for $\{Tx_n\}_{n\geq 1}$.

Theorem 1. *There exists an orthonormal basis in H, $\{\psi_j\}_{j\geq 1}$, and a sequence of real numbers $\{\lambda_j\}_{j\geq 1}$ accumulating at 0 and satisfying $||T|| = |\lambda_1| \geq |\lambda_2| \geq \dots$, such that for every $f \in H$ we have:*

$$Tf = \sum_{j\geq 1} \lambda_j \psi_j \langle f, \psi_j\rangle. \tag{1.1}$$

1.1 Proof of Theorem 1

Lemma 2. *Let $z \in \mathbb{C}$. We have $\mathrm{null}(T - z) = \{\mathrm{range}(T - \overline{z})\}^\perp$, and $H = \mathrm{null}(T - z) \oplus \overline{\{\mathrm{range}(T - \overline{z})\}}$.*

Proof. Let us prove the first equality. We know that T is symmetric, hence $\langle (T - z)f, g\rangle = \langle f, (T - \overline{z})g\rangle$ for all vectors $f, g \in H$. If $f \in \mathrm{null}(T - z)$, then $0 = \langle f, (T - \overline{z})g\rangle$ for all g, thus $f \in \{\mathrm{range}(T-\overline{z})\}^\perp$. If $f \in \{\mathrm{range}(T-\overline{z})\}^\perp$, then $\langle (T-z)f, g\rangle = 0$ for all $g \in H$, thus $(T-z)f = 0$ and $f \in \mathrm{null}(T - z)$.

Let us prove the second equality. We know that for any linear subspace M we have $\{M^\perp\}^\perp = \overline{M}$. Thus:

$$H = \mathrm{null}(T - z) \oplus \{\mathrm{null}(T - z)\}^\perp = \mathrm{null}(T - z) \oplus \overline{\mathrm{range}(T - \overline{z})}. \tag{1.2}$$

□

Lemma 3. *Let $z = x + iy$. Then $||(T - z)f|| \geq |y|\ ||f||$ for every $f \in H$. In particular, $\mathrm{null}(T - z) = \{0\}$ and $T - z$ is injective.*

Proof. It is an easy consequence of the fact that $\langle Tf, f\rangle$ is real and:

$$||f||\ ||(T - z)f|| \geq |\langle (T - z)f, f\rangle| = |\langle (T - x)f, f\rangle - iy||f||^2| \geq |y|\ ||f||^2.$$

□

Lemma 4. *Assume that for a given z, there exists $\delta > 0$ such that*

$$||(T - z)f|| \geq \delta||f||, \quad \forall f \in H. \tag{1.3}$$

Then $T - z$ is injective and surjective, thus $z \in \rho(T)$.

Proof. Let us write $z = x + iy$. Clearly, $T - z$ is injective. Our goal is to prove that $\mathrm{range}(T - z) = H$.

If $y \neq 0$, then (1.3) is a consequence of Lemma 3. Thus we can also assume that

$$||(T - \overline{z})f|| \geq \delta||f||, \quad \forall f \in H. \tag{1.4}$$

When $y = 0$, (1.4) contains no additional information.

In both cases, (1.4) implies that $T - \overline{z}$ is injective, thus $\mathrm{null}(T - \overline{z}) = \{0\}$. Using (1.2) with z replaced by $\overline{z}$ we obtain that the range of $T - z$ is dense in H:

$$\overline{\mathrm{range}(T - z)} = H. \tag{1.5}$$

The only remaining thing in the proof is to show that $\mathrm{range}(T - z)$ is a closed set, which together with (1.5) would show the surjectivity of $T - z$.

Let us do that. Assume that $\{y_n\}_{n\geq 1} \subset \text{range}(T-z)$ converges to $y_\infty \in H$. We have to show that $y_\infty \in \text{range}(T-z)$. There exists $\{x_n\}_{n\geq 1} \subset H$ such that $y_n = (T-z)x_n$. Using (1.3) we can write:

$$||x_{n+k} - x_n|| \leq \frac{1}{\delta}||(T-z)(x_{n+k} - x_n)|| = \frac{1}{\delta}||y_{n+k} - y_n||, \quad \forall n, k \geq 1. \tag{1.6}$$

Since $\{y_n\}_{n\geq 1}$ is Cauchy, (1.6) implies the same thing for $\{x_n\}_{n\geq 1}$. Thus there exists $x_\infty \in H$ such that $\lim_{n\to\infty} x_n = x_\infty$. Using this in the equality $Tx_n = zx_n + y_n$ together with the continuity of T, we obtain $Tx_\infty = zx_\infty + y_\infty$ and:

$$y_\infty = (T-z)x_\infty \in \text{range}(T-z).$$

□

Remark 1. The previous lemma shows that if T is a selfadjoint operator, then $z \in \rho(T)$ if

$$||(T-z)x|| \geq \delta > 0 \quad \forall x \in B_1(0) \tag{1.7}$$

Thus $\lambda \in \sigma(T)$ if

$$\inf_{||x||=1} ||(T-\lambda)x|| = 0,$$

or more precisely, if there exists a sequence $\{x_n\}_{n\geq 1}$ with $||x_n|| = 1$ such that

$$\lim_{n\to\infty} (T-z)x_n = 0. \tag{1.8}$$

Remark 2. Lemma 3 and Lemma 4 prove that $\sigma(T) \subset \mathbb{R}$. Moreover, if $|z| > ||T||$ we can write

$$(T-z)^{-1} = -\sum_{n\geq 0} \frac{1}{z^{n+1}} T^n, \tag{1.9}$$

thus $\sigma(T) \subset [-||T||, ||T||]$.

Let us now characterize the structure and nature of the spectrum of T.

Lemma 5. *If $\lambda \in \sigma(T)$ and $\lambda \neq 0$, then there exists at least one eigenvector $f \neq 0$ such that $Tf = \lambda f$.*

Proof. Because $\lambda \in \sigma(T)$, we have the bounded sequence $\{x_n\}_{n\geq 1}$ from (1.8). Since T is compact, we can find a subsequence $\{x_{n_k}\}_{k\geq 1}$ such that $\{Tx_{n_k}\}_{k\geq 1}$ is convergent to some y_∞. We can write:

$$x_{n_k} = \frac{1}{\lambda} T x_{n_k} - \frac{1}{\lambda}(T-\lambda)x_{n_k},$$

and since the r.h.s. converges to $\frac{1}{\lambda}y_\infty$ we conclude that $\lim_{k\to\infty} x_{n_k} = \frac{1}{\lambda}y_\infty$. The continuity of T implies that $\lim_{k\to\infty} Tx_{n_k} = \frac{1}{\lambda}Ty_\infty$. Hence:

$$0 = \lim_{k\to\infty}(Tx_{n_k} - \lambda x_{n_k}) = \frac{1}{\lambda}Ty_\infty - y_\infty.$$

Moreover, since $||x_{n_k}|| = 1$ implies that $||y_\infty|| = 1$, and we can choose our eigenvector $f = y_\infty$.

□

Lemma 6. *If $\lambda_1 \neq \lambda_2$ belong to the spectrum, and if f_1 and f_2 are two corresponding eigenvectors, then $\langle f_1, f_2\rangle = 0$.*

Proof. Use the symmetry of T and write:

$$0 = \langle Tf_1, f_2\rangle - \langle f_1, Tf_2\rangle = (\lambda_1 - \lambda_2)\langle f_1, f_2\rangle.$$

□

Lemma 7. *The spectrum of T cannot have other accumulation points outside 0. In other words, $\sigma(T) \setminus \{0\}$ is a discrete set consisting from isolated points.*

Proof. Assume that $\lambda \neq 0$ is an accumulation point of $\sigma(T)$. It means that we can find a sequence of points $\{\lambda_n\}_{n\geq 1} \subset \sigma(T)$, all distinct and not zero, such that

$$\lim_{n\to\infty} \lambda_n = \lambda.$$

From Lemma 5 we obtain at least an eigenvector x_n, $||x_n|| = 1$, such that $Tx_n = \lambda_n x_n$, or $x_n = \frac{1}{\lambda_n} T x_n$. Since T is compact, there exists a subsequence x_{n_k} such that Tx_{n_k} converges to some y. Thus

$$\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty} \frac{1}{\lambda_{n_k}} T x_{n_k} = \frac{1}{\lambda} y.$$

Thus we have just constructed a convergent subsequence of $\{x_n\}_{n\geq 1}$. But since each x_n corresponds to a different λ_n, Lemma 6 tells us that $||x_j - x_k|| = \sqrt{2}$ if $j \neq k$, therefore this sequence cannot have Cauchy subsequences. We arrived to a contradiction.

□

Lemma 8. *Assume that $\lambda \in \sigma(T) \setminus \{0\}$. Then the dimension of $\mathrm{null}(T - \lambda)$ is finite.*

Proof. Assume the contrary, i.e. the existence of infinitely many linearly independent vectors in $\mathrm{null}(T - \lambda)$. Up to a Gramm-Schmidt procedure, we can consider them to be orthogonal and normalized to one. If $\{x_n\}_{n\geq 1}$ is such a list, then again $||x_j - x_k|| = \sqrt{2}$, thus it cannot have any convergent subsequences. But since $x_n = \frac{1}{\lambda} T x_n$, the compactness of T would generate a convergent subsequence for $\{x_n\}_{n\geq 1}$, and we arrive to a contradiction.

□

Until now we know that the spectrum of T is contained in the interval $[-||T||, ||T||]$, it consists from isolated points outside 0, and the nullspace associated to each of its nonzero points is finite dimensional. Thus the nonzero spectrum is only composed from eigenvalues with finite geometric multiplicity, and they can only accumulate to 0.

Lemma 9. *At least one of the numbers $\pm||T||$ is an eigenvalue for T.*

Proof. From the definition of the norm, we have $||T|| = \sup_{||x||=1} ||Tx||$. Thus there exists a sequence $\{x_n\}_{n\geq 1}$, $||x_n|| = 1$ such that $\lim_{n\to\infty} ||Tx_n|| = ||T||$. Since T is compact, we can find a subsequence x_{n_k} such that $\lim_{k\to\infty} T x_{n_k} = y$, thus $||y|| = ||T||$. In order to simplify notation, denote $||T||$ by λ. Then we have:

$$\begin{aligned}
&\lim_{k\to\infty} ||(T^2 - \lambda^2) x_{n_k}||^2 = \lim_{k\to\infty} \langle (T^2 - \lambda^2) x_{n_k}, (T^2 - \lambda^2) x_{n_k} \rangle \\
&= \lim_{k\to\infty} \{ \langle T^2 x_{n_k}, T^2 x_{n_k} \rangle - 2\lambda^2 \langle T^2 x_{n_k}, x_{n_k} \rangle + \lambda^4 ||x_{n_k}||^2 \} = \langle Ty, Ty \rangle - 2\lambda^2 \langle y, y \rangle + \lambda^4 \\
&= \langle Ty, Ty \rangle - \lambda^4 \geq 0.
\end{aligned} \tag{1.10}$$

Thus we get $||Ty|| \geq \lambda^2$. Moreover:

$$\begin{aligned}
0 &\leq ||(T - \lambda)(T + \lambda) y||^2 = ||(T^2 - \lambda^2) y||^2 = \langle (T^2 - \lambda^2) y, (T^2 - \lambda^2) y \rangle \\
&= \langle T^2 y, T^2 y \rangle - 2\lambda^2 \langle T^2 y, y \rangle + \lambda^4 ||y||^2 = \langle T^2 y, T^2 y \rangle - 2\lambda^2 \langle Ty, Ty \rangle + \lambda^4 ||y||^2 \\
&\leq ||T^2 y||^2 - \lambda^6 \leq 0.
\end{aligned} \tag{1.11}$$

In the last line above we used (1.10). Thus (1.11) implies $(T-\lambda)(T+\lambda)y = 0$. Now if $(T+\lambda)y = 0$, it means that $-\lambda$ is an eigenvalue. If $f = (T + \lambda)y \neq 0$, then necessarily $(T - \lambda)f = 0$ which means that λ is an eigenvalue.

□

The previous result together with Lemma 8 imply the existence of a finite number of eigenvectors of T which span the subspace $M_\lambda := \text{null}(T-\lambda)$ where $\lambda = \pm||T||$. Denote by $\{\psi_j(\lambda)\}_{j=1}^{\dim(M_\lambda)}$ an orthonormal basis of M_λ, consisting of eigenvectors of T. Denote by P_λ the orthogonal projection associated to $\text{null}(T-\lambda)$:

$$P_\lambda f := \sum_{j=1}^{\dim(M_\lambda)} \langle f, \psi_j(\lambda)\rangle \psi_j(\lambda). \tag{1.12}$$

By direct computation, one can show that $P_\lambda^* = P_\lambda = P_\lambda^2$.

By convention, if λ is not in the spectrum of T, then $M_\lambda = \{0\}$ and $P_\lambda = 0$. Denote by

$$M_1 := M_{+||T||} \cup M_{-||T||}. \tag{1.13}$$

Lemma 10. *The subspace M_1 is a finite dimensional, closed linear subspace, which is left invariant by T (that is $TM_1 \subset M_1$). The same is true for $M_1^\perp$.*

Proof. Every $f \in M_1$ can be written as a finite linear combination of the type $f = \sum_j \langle f, \psi_j\rangle \psi_j$. Since all ψ_j's are eigenvectors of T, then $Tf \in M_1$.

Now let us prove that $M_1^\perp$ is invariant under T. Let $g \in M_1^\perp$. Then for every $f \in M_1$ we have:

$$\langle Tg, f\rangle = \langle g, Tf\rangle = 0,$$

since $Tf \in M_1$. Hence $Tg \in M_1^\perp$.

□

Now consider the decomposition $H = M_1 \oplus M_1^\perp$. The previous invariance result allows us to write our operator T as a direct sum $T = (||T||P_{+||T||} - ||T||P_{-||T||}) \oplus T_1$, where T_1 is simply the restriction of T to $M_1^\perp$. The next technical result is the following:

Lemma 11. *The restriction T_1 is also compact and selfadjoint. Moreover, $||T_1|| < ||T||$.*

Proof. The fact that T is compact and selfadjoint follows from

$$T_1 = T(1 - P_{+||T||} - P_{-||T||}) = (1 - P_{+||T||} - P_{-||T||})T.$$

Now let us prove that $||T_1|| < ||T||$. Clearly, $||T_1|| \le ||T||$, so we only need to prove that the two norms cannot be equal. Assume that they are equal. Then applying Lemma 9 to T_1, it would provide an eigenvector $\phi \in M_1^\perp$, $||\phi|| = 1$, for T_1. But ϕ would also be an eigenvector for T corresponding to $||T||$ or $-||T||$, thus $\phi \in M_1$, contradicting $\phi \neq 0$.

□

Remark 3. We have the inclusion $\text{null}(T) \subset M_1^\perp$; indeed, let $f \in \text{null}(T)$ and let ψ_j one eigenvector of T from M_1 corresponding to the eigenvalue $\lambda \neq 0$. Then

$$0 = \frac{1}{\lambda}\langle Tf, \psi_j\rangle = \frac{1}{\lambda}\langle f, T\psi_j\rangle = \langle f, \psi_j\rangle.$$

Thus f is orthogonal to any linear combination of ψ_j's, thus $f \in M_1^\perp$.

The proof of Theorem 1 is now almost over. If $M_1^\perp = \text{null}(T)$, then we have $H = M_1 \oplus \text{null}(T)$ and $T = (||T||P_{+||T||} - ||T||P_{-||T||}) \oplus 0$.

Otherwise, define M_2 as the subspace of $M_1^\perp$ corresponding to the union of $\text{null}(T_1 + ||T_1||)$ with $\text{null}(T_1 - ||T_1||)$ and decompose $H = M_1 \oplus (M_2 \oplus M_2^\perp)$. Here T_1 decomposes as

$$T_1 = (||T_1||P_{+||T_1||} - ||T_1||P_{-||T_1||}) \oplus T_2.$$

By induction, we obtain the decomposition

$$H = M_1 \oplus M_2 \cdots \oplus (M_n \oplus M_n^{\perp})$$

and

$$T = \oplus_{j=0}^{n-1}(||T_j||P_{+||T_j||} - ||T_j||P_{-||T_j||}) \oplus T_n,$$

where T_n is the restriction of T_{n-1} to $M_n^{\perp}$. By convention, $T_0 = T$. Reasoning as in the proof of Remark 3, we get that $\mathrm{null}(T) \subseteq M_n^{\perp}$. If they are equal, then we stop. Otherwise, we continue the reduction procedure.

Now assume that we never get $\mathrm{null}(T) = M_n^{\perp}$. It follows that $T_n \neq 0$, and also $\lim_{n\to\infty} ||T_n|| = 0$ because Lemma 7 forbids the accumulation of eigenvalues outside 0.

Lemma 12. *We have* $\oplus_{j\geq 0} M_j = \overline{\mathrm{range}(T)}$.

Proof. Fix $f \in H$. The vector $\sum_{j=0}^{n-1}(||T_j||P_{+||T_j||}f - ||T_j||P_{-||T_j||}f)$ can be seen as an element of $\oplus_{j\geq 0} M_j$, where all components with an index larger than $n+1$ are zero. We know that $T_n f = Tf - \sum_{j=0}^{n-1}(||T_j||P_{+||T_j||}f - ||T_j||P_{-||T_j||}f)$, and $||T_n f|| \to 0$ when n grows. Thus we can approximate Tf arbitrarily well with elements of $\oplus_{j\geq 0} M_j$. □

Corollary 13. *We have the decomposition* $H = \{\oplus_{j\geq 0} M_j\} \oplus \mathrm{null}(T)$.

Proof. Put $z = 0$ in (1.2) and use Lemma 12.

□

We can now conclude the proof of Theorem 1. The orthonormal basis consists from the eigenvectors of T corresponding to non-zero eigenvalues, put together with an arbitrary basis in $\mathrm{null}(T)$. The numbers λ_j's are either the eigenvalues of T or zero.

2 The singular value decomposition of a compact operator

Theorem 14. *Let H be a separable Hilbert space, and let A be a compact operator. Then there exist two orthonormal basis of H, $\{e_j\}_{j\geq 1}$ and $\{f_j\}_{j\geq 1}$, and a nonincreasing sequence of non-negative numbers $s_j \geq 0$ accumulating at zero such that for every $f \in H$ we have:*

$$Af = \sum_{j\geq 1} s_j \langle f, e_j\rangle f_j.$$

Proof. Let $T := A^*A$. We see that T is compact, selfadjoint and positive. Moreover, $\mathrm{null}(A) = \mathrm{null}(T)$; indeed, if $x \in \mathrm{null}(A)$ then $Tx = A^*(Ax) = 0$, thus $x \in \mathrm{null}(T)$. If $x \in \mathrm{null}(T)$, then $0 = \langle x, Tx\rangle = ||Ax||^2$ thus $Ax = 0$ and $x \in \mathrm{null}(A)$.

According to Theorem 1, there exists an orthonormal basis $\{e_j\}_{j\geq 1}$ consisting of eigenvectors of T, and let λ_j be their corresponding (non-zero) eigenvalues. We have

$$Af = \sum_{j\geq 1} \langle f, e_j\rangle Ae_j \tag{2.1}$$

In the above sum, only those e_j's appear which are not spanning the null space of T. Denote by $f_j := \frac{1}{||Ae_j||} Ae_j$, if $Ae_j \neq 0$. Clearly, from (2.1) it follows that the f_j's span the closure of the range of A. Now let us prove that the f_j's are orthogonal on each other. If $j \neq k$ we have

$$\langle f_j, f_k\rangle = \frac{1}{||Ae_j||\,||Ae_k||}\langle e_j, A^*Ae_k\rangle = 0.$$

We can extend the f_j basis in an arbitrary way to $\mathrm{range}(A)^{\perp}$. Finally, let us denote by $s_j := ||Ae_j|| = \sqrt{\langle e_j, A^*Ae_j\rangle} = \sqrt{\lambda_j}$. From (2.1) and the definition of f_j's and s_j's, the theorem is proved.

□