## 5 Compact linear operators

One of the most important results of Linear Algebra is that for every self-adjoint linear map A on a finite-dimensional space, there exists a basis consisting of eigenvectors. In particular, with respect to this basis the operator A can be represented by a diagonal matrix. The situation turns out to be much more complicated for operators on infinite dimensional spaces (in fact, self-adjoint operators may have no eigenvectors at all). Nonetheless, a satisfactory theory can be developed for so-called compact operators which we now introduce.

**Definition 5.1** (Compact linear operator). A linear operator  $T: X \to Y$  between normed spaces X and Y is called a **compact linear operator** if for every bounded sequence  $(x_n)_{n\geq 1}$  in X, the sequence  $(Tx_n)_{n\geq 1}$  has a convergent subsequence.

We note that every compact operator T is bounded. Indeed, if  $||T|| = \infty$ , then there exists a sequence  $(x_n)_{n\geq 1}$  such that  $||x_n|| \leq 1$  and  $||Tx_n|| \to \infty$ . Then  $(Tx_n)_{n\geq 1}$  cannot have a convergent subsequence. Hence,  $||T|| < \infty$ .

- **Example.** 1. Consider the linear operator  $T_N: \ell^2 \to \ell^2$  defined by  $T_N x = (x_1, x_2, \dots, x_N, 0, 0, \dots)$ . We claim that  $T_N$  is compact. Given any bounded sequence  $(x^{(n)})_{n\geq 1}$  in  $\ell^2$ , the sequence  $(T_N x^{(n)})_{n\geq 1}$  is bounded in  $\mathbb{F}^N \subset \ell^2$ . Since every bounded sequence in  $\mathbb{R}^N$  or  $\mathbb{C}^N$  has a convergent subsequence, it follows that  $T_N$  is compact.
  - 2. The identity operator  $I: \ell^2 \to \ell^2$  is not compact. To prove this, consider the bounded sequence  $(e_n)_{n\geq 1}$  where  $e_n = (0, \ldots, 0, 1, 0, \ldots)$ . Then for any  $n \neq m$ ,  $||Ie_n Ie_m||_2 = ||e_n e_m||_2 = \sqrt{2}$ . This implies that any subsequence of  $(Ie_n)_{n\geq 1}$  cannot be Cauchy and, hence, cannot converge. So that I is not compact.
  - 3. (integral operators) Let  $K \in C([0,1]^2)$  and  $T: L^2([0,1]) \to L^2([0,1])$  is defined by  $Tf(x) = \int_0^1 K(x,y)f(y)dy$ . Then the operator T is compact. We will not prove this in class, and it could be a possible topic for a level-M presentation.

We show that limits of compact operators is also compact.

**Theorem 5.2.** Let  $(T_n)_{n\geq 1}$  be a sequence of compact linear operators from a normed space X into a Banach spaces Y. If  $T_n \to T$  (that is,  $||T_n - T|| \to 0$ ), then the limit operator T is compact.

*Proof.* Since  $T_1$  is a compact operator, we know that the sequence  $(T_1(x_n))$  has a convergent (hence Cauchy) subsequence  $(T_1(x_{1,m}))$ , where  $(x_{1,m})$  is a subsequence of the original sequence  $(x_n)$ . The subsequence  $(x_{1,m})$  is bounded, so we can repeat the argument with  $T_2$  to produce a subsequence

 $(x_{2,m})$  of  $(x_{1,m})$  with the property that  $(T_2(x_{2,m}))$  converges. We continue in the same way, and then define a sequence  $(y_m) = (x_{m,m})$ . Notice that  $(y_m)$  is a subsequence of  $(x_n)$ , so it is bounded, say by  $||y_n|| \le c$ , and it has the property that for every fixed n, the sequence  $(T_n(y_m))$  is convergent, and hence Cauchy.

We claim that  $(T(y_m))$  is a Cauchy sequence in Y. Let  $\epsilon > 0$ . Since  $||T_n - T|| \to 0$ , there is some  $p \in \mathbb{N}$  such that  $||T_p - T|| < \frac{\epsilon}{3c}$ . Also, since  $(T_p(y_m))$  is Cauchy, there is some N > 0 such that  $||T_p(y_j) - T_p(y_k)|| < \frac{\epsilon}{3}$  whenever j, k > N. Therefore, for j, k > N, we have

$$||T(y_{j}) - T(y_{k})|| \le ||T(y_{j}) - T_{p}(y_{j}) + T_{p}(y_{j}) - T_{p}(y_{k}) + T_{p}(y_{k}) - T(y_{k})||$$

$$\le ||T(y_{j}) - T_{p}(y_{j})|| + ||T_{p}(y_{j}) - T_{p}(y_{k})|| + ||T_{p}(y_{k}) - T(y_{k})||$$

$$< ||T - T_{p}|| ||y_{j}|| + \frac{\epsilon}{3} + ||T - T_{p}|| ||y_{k}||$$

$$< \frac{\epsilon}{3c} c + \frac{\epsilon}{3} + \frac{\epsilon}{3c} c = \epsilon,$$

which proves that  $(T(y_m))$  is Cauchy. Since Y is a Banach space, it is by definition complete, so  $(T(y_m))$  converges. We have thus produced, for an arbitrary bounded sequence  $(x_n) \subset X$ , a convergent subsequence of its image under T. Therefore, T is compact.

**Example.** Let  $T: \ell^2 \to \ell^2$  be an operator defined by  $Tx = (\lambda_n x_n)_{n \geq 1}$  for a sequence  $\lambda_n \to 0$ . We claim that T is compact. To show this, we approximate T by compact operators  $T_N$  such that  $T_N x = (\lambda_1 x_1, \ldots, \lambda_N x_N, 0, \ldots)$ . As in the previous example we observe that  $T_N$  is a compact operator. For  $x \in \ell^2$ ,

$$||T_N x - T x||_2 = \left(\sum_{n>N} |\lambda_n x_n|^2\right)^{1/2} \le \left(\sup_{n>N} |\lambda_n|\right) \left(\sum_{n>N} |x_n|^2\right)^{1/2}$$

$$\le \left(\sup_{n>N} |\lambda_n|\right) ||x||_2.$$

This shows that  $||T_N - T|| \le \sup_{n>N} |\lambda_n|$ . Since  $\lambda_n \to 0$ , it follows that  $||T_N - T|| \to 0$ . Hence, T is compact by the previous theorem.

We note that for the operator T in the previous example there exists a basis of eigenvectors  $e_n$  with eigenvalues  $\lambda_n$ . Remarkably, every self-adjoint compact operator on a Hilbert space is of this form: it can be diagonalised with respect to suitable orthonormal set. The following is the main result of this lecture:

**Theorem 5.3** (spectral theorem). Let H be a Hilbert space and  $T: H \to H$  a compact self-adjoint operator. Then there exists an orthonormal set

 $(e_n)_{n\geq 1}$  consisting of eigenvectors of T with eigenvalues  $\lambda_n$  such that

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n, \quad x \in H.$$

The proof of this theorem require several auxiliary results.

**Theorem 5.4.** Let T be a self-adjoint operator on a Hilbert space H. Then

$$||T|| = \sup\{|\langle Tx, x \rangle| : ||x|| = 1\}.$$

*Proof.* Let  $m = \sup\{|\langle Tx, x \rangle| : ||x|| = 1\}$ . For  $x \in H$  with ||x|| = 1,

$$|\langle Tx, x \rangle| \le ||Tx|| ||x|| \le ||T|| ||x||^2 = ||T||.$$

This shows that m < ||T||.

For any  $x, y \in H$ ,

$$\operatorname{Re} \langle Tx, y \rangle = \frac{1}{4} (\langle T(x+y), x+y \rangle - \langle T(x-y), x-y \rangle).$$

This formula can be checked by expanding the right hand side. We use that  $|\langle Tu, u \rangle| \le m||u||^2$  for all  $u \in H$ . Then from the above formula, we obtain

Re 
$$\langle Tx, y \rangle \le \frac{1}{4} (|\langle T(x+y, x+y) | + |\langle T(x-y), x-y \rangle |)$$
  
 $\le \frac{1}{4} (m||x+y||^2 + m||x-y||^2) = \frac{m}{2} (||x||^2 + ||y||^2),$ 

where we used the parallelogram identity. Replacing x by  $\lambda x$  where  $|\lambda| = 1$  and  $\lambda$  is chosen so that  $\lambda \langle Tx, y \rangle = |\langle Tx, y \rangle|$  is real and non-negative, we obtain

$$|\langle Tx, y \rangle| \le \frac{m}{2} (||x||^2 + ||y||^2).$$

Suppose that  $||Tx|| \neq 0$  and take  $y = ||x|| \frac{Tx}{||Tx||}$ . This gives  $||x|| ||Tx|| \leq m||x||^2$ . Hence,  $||Tx|| \leq m||x||$ . This inequality also obviously holds when Tx = 0. We conclude that  $||T|| \leq m$ .

**Theorem 5.5.** Let T be a compact self-adjoint operator on a Hilbert space H. Then either ||T|| or -||T|| is an eigenvalue.

*Proof.* We assume that that  $T \neq 0$ .

By the previous theorem there exists  $x_n \in H$  with  $||x_n|| = 1$  such that  $|\langle Tx_n, x_n \rangle| \to ||T||$ . Since  $T^* = T$ ,  $\overline{\langle Tx_n, x_n \rangle} = \langle x_n, Tx_n \rangle = \langle Tx_n, x_n \rangle$ , so that  $\langle Tx_n, x_n \rangle$  is real. Passing to subsequence we may assume that  $\langle Tx_n, x_n \rangle \to \lambda$  where  $\lambda = ||T||$  or  $\lambda = -||T||$ . Then

$$||Tx_n - \lambda x_n||^2 = ||Tx_n||^2 - 2\lambda \langle Tx_n, x_n \rangle + \lambda^2 ||x_n||^2 \le 2\lambda^2 - 2\lambda \langle Tx_n, x_n \rangle \to 0.$$

This shows that  $Tx_n - \lambda x_n \to 0$ . Since T is compact, passing to a subsequence we may assume that  $Tx_n \to y$  for some  $y \in H$ . Then it follows that  $\lambda x_n \to y$ . Since T is continuous,  $\lambda Tx_n \to Ty$ , but  $\lambda Tx_n \to \lambda y$ . Hence, we conclude that  $Ty = \lambda y$ .

We claim that  $y \neq 0$ . Indeed, by the triangle inequality,

$$||Tx_n|| \ge ||\lambda x_n|| - ||Tx_n - \lambda x_n|| = |\lambda| - ||Tx_n - \lambda x_n|| \to |\lambda| = ||T|| > 0.$$

Hence,  $||y|| = \lim ||Tx_n|| > 0$ . We have proved that  $\lambda$  is an eigenvalue of T.

**Theorem 5.6.** Let T be a compact self-adjoint operator on a Hilbert space H and  $\lambda_n$ 's are eigenvalues of T with linearly independent eigenvectors  $x_n$ . Then  $\lambda_n$ 's are real, and for every c > 0 there are only finitely many n's such that  $|\lambda_n| \geq c$ .

*Proof.* We have

$$\lambda_n ||x_n||^2 = \langle Tx_n, x_n \rangle = \langle x_n, Tx_n \rangle = \bar{\lambda}_n ||x_n||^2.$$

Since  $x_n \neq 0$ , it follows that  $\lambda_n = \bar{\lambda}_n$  and  $\lambda_n$  is real.

Suppose that  $\lambda_n \neq \lambda_m$ . Then

$$\lambda_n \langle x_n, x_m \rangle = \langle Tx_n, x_m \rangle = \langle x_n, Tx_m \rangle = \lambda_m \langle x_n, x_m \rangle$$
.

This implies that  $\langle x_n, x_m \rangle = 0$ , that is  $x_n$  and  $x_m$  are orthogonal. If  $\lambda_n$ 's are repeating, then consider  $N_{\lambda} = \{n : \lambda_n = \lambda\}$ . We note that linear combinations of  $x_n$ 's with  $n \in N_{\lambda}$  are also eigenvectors with eigenvalue  $\lambda$ . Using the Gram-Schmidt orthogonalisation algorithm, we can construct orthonormal eigenvectors for  $\lambda_n$  with  $n \in N_{\lambda}$ .

Now we can assume that  $x_n$ 's are orthonormal. Suppose that for infinitely many n's  $|\lambda_n| \geq c$ . Then for such indices  $n \neq m$ ,

$$||Tx_n - Tx_m||^2 = ||\lambda_n x_n - \lambda_m x_m||^2 = |\lambda_n|^2 + |\lambda_m|^2 \ge 2c^2.$$

This gives a sequence of the form  $(Tx_n)$  which does not contain any Cauchy subsequence. This contradicts compactness of the operator T.

We note that Theorem 5.6 implies that if we order the eigenvalues of T as  $|\lambda_1| \ge |\lambda_2| \ge \cdots$ , then  $|\lambda_n| \to 0$ .

**Theorem 5.7.** Let T be a self-adjoint operator on a Hilbert space H, and U is a closed subspace of H such that  $T(U) \subset U$ . Then  $T(U^{\perp}) \subset U^{\perp}$ .

*Proof.* Let 
$$x \in U^{\perp}$$
. Then for any  $y \in U$ ,  $Ty \in U$ , and  $\langle Tx, y \rangle = \langle x, Ty \rangle = 0$ . Hence,  $Tx \in U^{\perp}$ .

Now we are ready for the proof of the main Theorem 5.3. The idea is to apply Theorem 5.5 inductively.

Proof of Theorem 5.3. By Theorem 5.5, there exists an eigenvector  $e_1$  with eigenvalue  $\lambda_1 = \pm \|T\|$ . Note that if  $Tx = \lambda x$  with  $x \in 0$ , then  $\|Tx\| = \|\lambda\| \|x\| \le \|T\| \|x\|$ , so that  $\|\lambda\| \le \|T\|$ . Hence,  $\|\lambda_1\|$  is maximal among eigenvalues. After rescaling we can assume that  $\|e_1\| = 1$ . The subspace  $U_1 = \operatorname{span}(e_1)$  is closed and T-invariant. By Theorem 5.7,  $U_1^{\perp}$  is also T-invariant. Now we can apply the same construction to the operator  $T_2: U_1^{\perp} \to U_1^{\perp}$ , which is the restriction of the operator T to  $U_1^{\perp}$ , to construct a unit eigenvector  $e_2 \in U_1^{\perp}$  with eigenvalue  $\lambda_2$  such that  $|\lambda_2| = \|T_2\|$ . Since  $\|T_2\| \le \|T\|$ , we have  $|\lambda_2| \le |\lambda_1|$ . Next, we consider  $T_3: U_2^{\perp} \to U_2^{\perp}$  where  $U_2 = \langle e_1, e_2 \rangle$ , and so on ...

After n steps we produce an orthonormal set  $\{e_1, \ldots e_n\}$  consisting of eigenvectors with eigenvalues  $\lambda_1, \ldots, \lambda_n$  such that  $|\lambda_1| \geq \cdots \geq |\lambda_n|$ . Setting  $U_n = \langle e_1, \ldots e_n \rangle$ , we have the orthogonal decomposition  $H = U_n \oplus U_n^{\perp}$ . For a vector  $x \in H$ , we set  $y_n = x - \sum_{i=1}^n \langle x, e_i \rangle e_i$ . Since  $e_i$ 's are orthonormal, it is easy to check that  $y_n \in U_n^{\perp}$ . Hence, we have the orthogonal sum  $x = \sum_{i=1}^n \langle x, e_i \rangle e_i + y_n$ , and by the Pythagoras' theorem,  $||x||^2 = \sum_{i=1}^n |\langle x, e_i \rangle|^2 + ||y_n||^2$ . In particular,  $||y_n|| \leq ||x||$ . We observe that

$$\left\| Tx - \sum_{i=1}^{n} \lambda_i \langle x, e_i \rangle e_i \right\| \le \left\| T \left( x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right) \right\|$$
$$= \left\| T_{n+1} y_n \right\| \le \left\| T_{n+1} \right\| \|y_n\| \le |\lambda_{n+1}| \|x\|.$$

Since by Theorem 5.6,  $|\lambda_{n+1}| \to 0$ , this proves that the operator T can be represented as  $Tx = \sum_{i=1}^{\infty} \lambda_i \langle x, e_i \rangle e_i$ ,  $x \in H$ .