3.3. Compact Operators

Synopsis. Equivalent characterizations of compact operators. The spectral theorem for normal, compact operators. Atkinson's theorem. Fredholm operators and index. Invariance properties of the index. Exercises.

3.3.1. An operator T on an infinite-dimensional Hilbert space $\mathfrak{H}$ has *finite rank* if $T(\mathfrak{H})$ is a finite-dimensional subspace of $\mathfrak{H}$ (hence closed, cf. 2.1.9). The set $\mathbf{B}_f(\mathfrak{H})$ of operators in $\mathbf{B}(\mathfrak{H})$ with finite rank is clearly a subspace, and it is easily verified that $\mathbf{B}_f(\mathfrak{H})$ is not only a subalgebra, but even an ideal in $\mathbf{B}(\mathfrak{H})$ (cf. 4.1.2).

If $T \in \mathbf{B}_f(\mathfrak{H})$, we may use 3.2.5 to obtain an orthogonal decomposition $\mathfrak{H} = T(\mathfrak{H}) \oplus \ker T^*$, which shows that $T^*(\mathfrak{H}) = T^*T(\mathfrak{H})$, so that T^* has finite rank. Thus, $\mathbf{B}_f(\mathfrak{H})$ is a *self-adjoint* ideal in $\mathbf{B}(\mathfrak{H})$ [i.e. $(\mathbf{B}_f(\mathfrak{H}))^* = \mathbf{B}_f(\mathfrak{H})$ as a set].

The class $\mathbf{B}_f(\mathfrak{H})$ bears much the same relation to $\mathbf{B}(\mathfrak{H})$ as the class $C_c(X)$ in relation to $C_b(X)$ (when X is a locally compact Hausdorff space; see 1.7.6 and 2.1.14). These classes describe local phenomena on $\mathfrak{H}$ and on X. Passing to a limit in norm may destroy the exact "locality," but enough structure is preserved to make these "quasilocal" operators and functions very attractive. We shall study the closure of $\mathbf{B}_f(\mathfrak{H})$ in this section as a noncommutative analogue of $C_0(X)$ in function theory.

3.3.2. Lemma. *There is a net $(P_\lambda)_{\lambda \in \Lambda}$ of projections in $\mathbf{B}_f(\mathfrak{H})$ such that $\|P_\lambda x - x\| \to 0$ for each x in $\mathfrak{H}$.*

PROOF. Take any orthonormal basis $\{e_j \mid j \in J\}$ for $\mathfrak{H}$ (cf. 3.1.12), and let Λ be the net of finite subsets of J, ordered under inclusion. For each λ in Λ let P_λ denote the projection of $\mathfrak{H}$ on the subspace $\operatorname{span}\{e_j \mid j \in \lambda\}$, so that $(P_\lambda)_{\lambda \in \Lambda}$ is indeed a net in $\mathbf{B}_f(\mathfrak{H})$. If $x \in \mathfrak{H}$, we have $x = \sum \alpha_j e_j$, whence $\|P_\lambda x - x\|^2 = \sum |\alpha_j|^2$, the summation being over all $j \notin \lambda$; and this tends to zero by Parseval's identity (3.1.11). □

3.3.3. Theorem. *Let $\mathfrak{B}$ denote the closed unit ball in a Hilbert space $\mathfrak{H}$. Then the following conditions on an operator T in $\mathbf{B}(\mathfrak{H})$ are equivalent:*

(i) $T \in (\mathbf{B}_f(\mathfrak{H}))^=$.
(ii) $T|\mathfrak{B}$ *is a weak–norm continuous function from $\mathfrak{B}$ into $\mathfrak{H}$.*
(iii) $T(\mathfrak{B})$ *is compact in $\mathfrak{H}$.*
(iv) $(T(\mathfrak{B}))^=$ *is compact in $\mathfrak{H}$.*
(v) *Each net in $\mathfrak{B}$ has a subnet whose image under T converges in $\mathfrak{H}$.*

PROOF. (i) ⇒ (ii). Let $(x_\lambda)_{\lambda \in \Lambda}$ be a weakly convergent net in $\mathfrak{B}$ with limit x. Given $\varepsilon > 0$ there is by assumption an S in $\mathbf{B}_f(\mathfrak{H})$ with $\|S - T\| < \varepsilon/3$, whence

$$\|Tx_\lambda - Tx\| \le 2\|T - S\| + \|Sx_\lambda - Sx\|$$
$$\le \tfrac{2}{3}\varepsilon + \|Sx_\lambda - Sx\|.$$

Since every operator in $\mathbf{B}(\mathfrak{H})$ is weak–weak continuous (cf. 3.1.10), we have $Sx_\lambda \to Sx$, weakly. However, on the finite-dimensional subspace $S(\mathfrak{H})$ all vector space topologies coincide by 2.1.9, so that $Sx_\lambda \to Sx$ in norm. Eventually, therefore, $\|Tx_\lambda - Tx\| < \varepsilon$. Since ε was arbitrary, T is weak–norm continuous by 1.4.3.

(ii) ⇒ (iii) Since $\mathfrak{B}$ is weakly compact (3.1.10), the image $T(\mathfrak{B})$ is norm(!) compact by 1.6.7.

(iii) ⇒ (iv) is trivial, since $T(\mathfrak{B})$ is closed by 1.6.5.

(iv) ⇒ (v) It follows from 1.6.2(v) that if $T(\mathfrak{B})$ is relatively compact, then every net has a convergent subnet.

(v) ⇒ (i) Take $(P_\lambda)_{\lambda \in \Lambda}$ as in 3.3.2. Then $P_\lambda T \in \mathbf{B}_f(\mathfrak{H})$ for every λ, and we claim that $P_\lambda T \to T$. If not, there is an $\varepsilon > 0$ and (passing if necessary to a subnet of Λ) for every λ a unit vector x_λ with $\|(P_\lambda T - T)x_\lambda\| \geq \varepsilon$. By assumption we may assume that the net $(Tx_\lambda)_{\lambda \in \Lambda}$ is norm(!) convergent in $\mathfrak{H}$ with a limit y. But then by 3.3.2

$$\begin{aligned}\varepsilon \leq \|(I - P_\lambda)Tx_\lambda\| &\leq \|(I - P_\lambda)(Tx_\lambda - y)\| + \|(I - P_\lambda)y\| \\ &\leq \|Tx_\lambda - y\| + \|(I - P_\lambda)y\| \to 0,\end{aligned}$$

a contradiction. Thus, $\|P_\lambda T - T\| \to 0$, as desired. □

3.3.4. The class of operators satisfying 3.3.3 is called the *compact operators* [after condition (iii)] and is denoted by $\mathbf{B}_0(\mathfrak{H})$ to signify that these operators "vanish at infinity." Unfortunately, this notation is not standard, and the reader will more often find the letters $\mathbf{K}$ or $\mathbf{C}$ [sometimes $\mathbf{K}(\mathfrak{H})$ or $\mathbf{C}(\mathfrak{H})$] used. We see from condition (i) that $\mathbf{B}_0(\mathfrak{H})$ is a norm closed, self-adjoint ideal in $\mathbf{B}(\mathfrak{H})$ (and actually the smallest such; for separable Hilbert spaces even the *only* closed ideal). Note that $I \notin \mathbf{B}_0(\mathfrak{H})$ when $\mathfrak{H}$ is infinite-dimensional, but that $\mathbf{B}_0(\mathfrak{H})$ has an approximate unit consisting of projections of finite rank [cf. the proof of the implication (v) ⇒ (i)].

3.3.5. Lemma. *A diagonalizable operator T in $\mathbf{B}(\mathfrak{H})$ is compact* iff *its eigenvalues $\{\lambda_j | j \in J\}$ corresponding to an orthonormal basis $\{e_j | j \in J\}$ belongs to $c_0(J)$.*

Proof. We have $Tx = \sum \lambda_j (x|e_j)e_j$ for every x in $\mathfrak{H}$, cf. (∗) in 3.2.14. If $T \in \mathbf{B}_0(\mathfrak{H})$ and $\varepsilon > 0$ is given, we let

$$J_\varepsilon = \{j \in J \,|\, |\lambda_j| \geq \varepsilon\}.$$

If J_ε is infinite, the net $(e_j)_{j \in J_\varepsilon}$ will converge weakly to zero for any well-ordering of J_ε, because $(e_j|x) \to 0$ by Parseval's identity 3.1.11. Since $\|Te_j\| = |\lambda_j| \geq \varepsilon$ for j in J_ε, this contradict condition (ii) in 3.3.3. Thus, J_ε is finite for each $\varepsilon > 0$, which means that the λ_j's vanish at infinity.

Conversely, if J_ε is finite for each $\varepsilon > 0$, we let

$$T_\varepsilon = \sum_{j \in J_\varepsilon} \lambda_j(\cdot|e_j)e_j.$$

Then T_ε has finite rank and

$$\|(T - T_\varepsilon)x\|^2 = \left\| \sum_{j \notin J_\varepsilon} \lambda_j (x|e_j) e_j \right\|^2$$
$$= \sum_{j \notin J_\varepsilon} |\lambda_j|^2 |(x|e_j)|^2 \leq \varepsilon^2 \|x\|^2.$$

Thus $\|T - T_\varepsilon\| \leq \varepsilon$, whence $T \in \mathbf{B}_0(\mathfrak{H})$ by 3.3.3(i). □

Lemma 3.3.6. *If x is an eigenvector for a normal operator T in $\mathbf{B}(\mathfrak{H})$, corresponding to the eigenvalue λ, then x is an eigenvector for T^*, corresponding to the eigenvalue $\bar{\lambda}$. Eigenvectors for T corresponding to different eigenvalues are orthogonal.*

PROOF. The operator $T - \lambda I$ is normal and its adjoint is $T^* - \bar{\lambda} I$. Consequently, we have

$$\|(T^* - \bar{\lambda} I)x\| = \|(T - \lambda I)x\| = 0;$$

cf. 3.2.7. If $Ty = \mu y$ with $\lambda \neq \mu$, we may assume that $\lambda \neq 0$, whence

$$(x|y) = \lambda^{-1}(Tx|y) = \lambda^{-1}(x|T^* y) = \lambda^{-1}(x|\bar{\mu} y) = \lambda^{-1}\mu(x|y),$$

so that $(x|y) = 0$. □

3.3.7. Lemma. *Every normal, compact operator T on a complex Hilbert space $\mathfrak{H}$ has an eigenvalue λ with $|\lambda| = \|T\|$.*

PROOF. With $\mathfrak{B}$ as the unit ball of $\mathfrak{H}$ we know from 3.3.3 that $T: \mathfrak{B} \to \mathfrak{H}$ is weak–norm continuous. If therefore $x_i \to x$ weakly in $\mathfrak{B}$, then

$$|(Tx_i|x_i) - (Tx|x)| = |(T(x_i - x)|x_i) + (Tx|x_i - x)|$$
$$\leq \|T(x_i - x)\| + |(Tx|x_i - x)| \to 0.$$

This shows that the function $x \to |(Tx|x)|$ is weakly continuous on $\mathfrak{B}$. Since $\mathfrak{B}$ is weakly compact, the function attains it maximum (1.6.7), and by 3.2.25 that maximum is $\|T\|$. Thus,

$$|(Tx|x)| = \|T\|$$

for some x in $\mathfrak{B}$. Now

$$\|T\| = |(Tx|x)| \leq \|Tx\| \|x\| \leq \|T\|,$$

so that, in fact, $|(Tx|x)| = \|Tx\| \|x\|$. But as we saw in 3.1.3, equality holds in the Cauchy–Schwarz inequality only when the vectors are proportional, and, therefore, $Tx = \lambda x$ for some λ. Evidently $|\lambda| = \|T\|$. □

3.3.8. Theorem. *Every normal, compact operator T on a complex Hilbert space $\mathfrak{H}$ is diagonalizable and its eigenvalues (counted with multiplicity) vanish at infinity. Conversely, each such operator is normal and compact.*

Proof. We need only show that every normal, compact operator T is diagonalizable, since the rest of the statement in 3.3.8 is contained in 3.3.5. Toward this end, consider the family of orthonormal systems in $\mathfrak{H}$ consisting entirely of eigenvectors for T. Clearly this family is inductively ordered under inclusion, so by Zorn's lemma (1.1.3) it has a maximal element $\{e_j | j \in J\}$. Let $\{\lambda_j | j \in J\}$ denote the corresponding system of eigenvalues, and let P denote the projection on the closed subspace spanned by the e_j's [so that these form a basis for $P(\mathfrak{H})$]. For each x in $\mathfrak{H}$ we then have, by 3.3.6, that

$$\begin{aligned} TPx &= T\sum (x|e_j)e_j = \sum (x|e_j)\lambda_j e_j \\ &= \sum (x|\bar{\lambda}_j e_j)e_j = \sum (x|T^* e_j)e_j \\ &= \sum (Tx|e_j)e_j = PTx. \end{aligned}$$

It follows that T and P commute, so that the operator $(I - P)T$ is normal and compact. If $P \neq I$, we either have $(I - P)T = 0$, and then every unit vector e_0 in $(I - P)\mathfrak{H}$ is an eigenvector for T, or else $(I - P)T \neq 0$, in which case by 3.3.7 there is a unit vector e_0 in $(I - P)\mathfrak{H}$ with $Te_0 = \lambda e_0$ and $|\lambda| = \|(I - P)T\|$. Both cases contradict the maximality of the system $\{e_j | j \in J\}$, and therefore $P = I$. □

3.3.9. It will be convenient, especially for the next section, to introduce the notation $x \odot y$ for the rank one operator in $\mathbf{B}(\mathfrak{H})$ determined by the vectors x and y in $\mathfrak{H}$ by the formula

$$(x \odot y)z = (z|y)x.$$

Note that the map $x, y \to x \odot y$ is a sesquilinear map of $\mathfrak{H} \times \mathfrak{H}$ into $\mathbf{B}_f(\mathfrak{H})$. If $\|e\| = 1$, then $e \odot e$ is the one-dimensional projection of $\mathfrak{H}$ on $\mathbb{C}e$. Every normal, compact operator on $\mathfrak{H}$ can now by 3.3.8 be written in the form

$$T = \sum \lambda_j e_j \odot e_j$$

for a suitable orthonormal basis $\{e_j | j \in J\}$. The sum converges in norm, because either the set $J_0 = \{j \in J | \lambda_j \neq 0\}$ is finite [so that $T \in \mathbf{B}_f(\mathfrak{H})$] or else is countably infinite, in which case the sequence $\{\lambda_j | j \in J_0\}$ converges to zero. We say that the compact set

$$\operatorname{sp}(T) = \{\lambda_j | j \in J_0\} \cup \{0\}$$

is the *spectrum* of T.

For every continuous function f on $\operatorname{sp}(T)$ we define

$$f(T) = \sum f(\lambda_j) e_j \odot e_j.$$

Then $f(T)$ is compact iff $f(0) = 0$, and the map $f \to f(T)$ is an isometric $*$-preserving homomorphism of $C(\operatorname{sp}(T))$ into $\mathbf{B}(\mathfrak{H})$. Moreover, if $f(z) = \sum \alpha_{nm} z^n \bar{z}^m$, a polynomial in the two commuting variables z and $\bar{z}$, then $f(T) = \sum \alpha_{nm} T^n T^{*m}$. This result is the *spectral (mapping) theorem* for normal, compact operators. In the next chapter we shall show a generalized version of the spectral mapping theorem, valid for every normal operator.