8 Trace class operators

Throughout this section, we will work over the ground field $\mathbb{C}$.

8.1 Ordinary matrix trace and determinant

For $n \times n$ complex matrices T, the trace of T may be defined to be

$$\mathrm{Tr}(T) = \sum_{i=1}^{n} T_{ii}. \tag{8.1}$$

A little algebra shows the cyclicity property,

$$\mathrm{Tr}(T_1 T_2) = \mathrm{Tr}(T_2 T_1), \qquad T_1, T_2 \in \mathrm{M}_n(\mathbb{C}).$$

It follows that

$$\mathrm{Tr}(UTU^{-1}) = \mathrm{Tr}(T), \qquad T \in \mathrm{M}_n(\mathbb{C}),\ U \in \mathrm{GL}_n(\mathbb{C}),$$

so remarkably, trace is an invariant of T as a linear operator — the formula (8.1) is independent of basis choice.

There is a spectral theoretic reason for this. Namely, transforming T to its Jordan normal form, whose diagonal entries are the eigenvalues of T, we see that $\mathrm{Tr}(T)$ is the sum of the eigenvalues of T, counted with algebraic multiplicity. The latter sum can be taken to the definition of $\mathrm{Tr}(T)$, then one proves that the diagonal sum formula works.

Similarly, the determinant $\det_n(T)$ (which is also basis independent) is the product of the eigenvalues of T (with multiplicity). Furthermore, there is an identity

$$\det_n(\exp(T)) = e^{\mathrm{Tr}(T)}, \qquad T \in \mathrm{M}_n(\mathbb{C}). \tag{8.2}$$

Here, $\exp(T)$ is defined to be the linear operator obtained by the convergent (with respect to some matrix norm) power series

$$\exp(T) = 1 + T + \frac{T^2}{2} + \dots.$$

It is easiest to understand the series by transforming T into a direct sum of Jordan blocks. Let J be such a Jordan block, with diagonal part $\lambda \cdot 1_m$ (where λ is an eigenvalue of T), and N is the strictly upper triangular part. Then

$$\begin{aligned} \exp(J) &= 1 + (\lambda \cdot 1_m + N) + (\lambda \cdot 1_m + N)^2 + \cdots \\ &= \big(1 + \lambda + \lambda^2/2 + \cdots\big) \cdot 1_m + \big(p_1(\lambda) N + p_2(\lambda) N^2/2 + \cdots\big), \end{aligned}$$

where $p_i(\lambda)$ are some polynomials in λ. Since N is nilpotent, the sum involving powers of N is a finite sum, so we have $\exp(J) = \exp(\lambda)\cdot 1_m + S$, where S is some strictly upper-triangular matrix. The determinant of the upper triangular J equals that of its diagonal part, so

$$\det(\exp(J)) = (e^\lambda)^m = e^{m\lambda} = e^{\mathrm{Tr}(J)}.$$

As T is the direct sum of its Jordan blocks, we arrive immediately at (8.2).

8.2 Trace in infinite dimensions?

It would be futile to define a trace on all linear operators on X, if X is infinite-dimensional — what should $\mathrm{Tr}(\mathrm{id}_X)$ be? There are, roughly speaking, two types of difficulties:

1. (Sum along diagonal?) We have to pick a basis, and restrict to operators whose diagonal entries are summable. Absolute summability would be preferable, otherwise the order of the basis elements would matter. To get basis independence, we would need to restrict to basis changes which are "continuous enough" to preserve the diagonal summability. What notion of "basis" should be used?

2. (Sum of eigenvalues?) This would be basis independent, but we have to restrict to operators with at most countably many nonzero eigenvalues, each with finite multiplicity. Also, such a definition would be impractical since determining eigenvalues is extremely difficult in general.

So what we really want is a notion of trace on some fairly general class of linear operators, conveniently defined as a diagonal sum, then prove that this sum actually also sums up the eigenvalues. We will see how to do this for a certain class of operators on Hilbert spaces.

Continuous linear operators. Let X, Y be normed vector spaces. We write $\mathcal{L}(X, Y)$ for the vector space of continuous linear operators $T : X \to Y$. If $X = Y$, we write $\mathcal{L}(X) \equiv \mathcal{L}(X, X)$ for the *algebra*[13] of continuous linear operators on X.

[13] An algebra (over $\mathbb{C}$) is a complex vector space with a bilinear product. It is, in particular, a ring.

In a basic functional analysis course, one learns that $\mathcal{L}(X,Y)$ is precisely the normed vector space of *bounded* linear operators $T : X \to Y$, i.e., with finite *operator norm*,

$$\|T\|_{\mathrm{op}} = \sup_{\|x\|\leq 1} \|Tx\|.$$

The operator norm is *submultiplicative*,

$$\|T_1 T_2\|_{\mathrm{op}} \leq \|T_1\|_{\mathrm{op}} \|T_2\|_{\mathrm{op}}.$$

If Y is a Banach space, then so is $\mathcal{L}(X,Y)$, and thus $\mathcal{L}(Y)$ is a *Banach algebra*. Recall that a Banach algebra is an algebra with a submultiplicative norm making it a Banach space. The submultiplicativity is crucial for *spectral theory*, as inequalities like $\|T^n\|_{\mathrm{op}} \leq \|T\|_{\mathrm{op}}^n$ are used in analogy to $|z^n| = |z|^n$ to transfer complex function theory results to $\mathcal{L}(X)$-valued functions.

For any normed space X, there is a *dual space* X^* of continuous linear functionals $X \to \mathbb{C}$, which is always a Banach space under the operator norm. A special feature of Hilbert spaces $\mathcal{H}$ is the antilinear isometric isomorphism $\mathcal{H} \cong \mathcal{H}^*$ given by the *Riesz representation theorem*. Namely, $x \in \mathcal{H}$ corresponds to the continuous linear functional

$$\langle x|- \rangle : y \mapsto \langle x|y\rangle, \qquad y \in \mathcal{H},$$

and each continuous linear function $\mathcal{H} \to \mathbb{C}$ is of this form for a unique vector $\mathcal{H}$. From this duality, one readily proves that each continuous operator $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ has a corresponding *adjoint* operator $T^* \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$, defined by the equation

$$\langle T^* x|y\rangle_{\mathcal{H}_2} = \langle x|Ty\rangle_{\mathcal{H}_1}, \qquad \forall\, x \in \mathcal{H}_1, y \in \mathcal{H}_2.$$

Basic properties of the adjoint operation are: $T \mapsto T^*$ is antilinear, and

$$T^{**} = T, \qquad (T_1 T_2)^* = T_2^* T_1^*, \qquad \|T^*\|_{\mathrm{op}} = \|T\|.$$

Trace for positive operators on Hilbert spaces. An operator $T \in \mathcal{L}(\mathcal{H})$ is *positive*, denoted $T \geq 0$, if

$$\langle x|Tx\rangle \geq 0, \qquad \forall\, x \in \mathcal{H}.$$

Lemma 8.1. *Let T be a positive operator on a Hilbert space. Then*

$$\operatorname{Tr} T := \sum_{e\in B} \langle e|Te\rangle \ \in [0,\infty] \tag{8.3}$$

is independent of the choice of orthonormal basis[14] *B.*

(The ordering in the sum does not matter as all terms are non-negative.)

Proof. As explained in Section 8.3, there exists a self-adjoint $\sqrt{T} \in \mathcal{L}(\mathcal{H})$ such that $\sqrt{T}\sqrt{T} = T$. Then we can rewrite

$$\sum_{e\in B} \langle e|Te\rangle = \sum_{e\in B} \langle e|\sqrt{T}\sqrt{T}e\rangle = \sum_{e\in B} \langle \sqrt{T}e|\sqrt{T}e\rangle = \sum_{e\in B} \|\sqrt{T}e\|^2.$$

Expanding each $\sqrt{T}e$ with respect to any other orthonormal basis B', using the Parseval identity, rearranging sums (all terms are non-negative), and self-adjointness of $\sqrt{T}$,

$$\sum_{e\in B} \langle e|Te\rangle = \sum_{e\in B} \sum_{e'\in B'} |\langle e'|\sqrt{T}e\rangle|^2 = \sum_{e'\in B'} \sum_{e\in B} |\langle e|\sqrt{T}e'\rangle|^2 = \sum_{e'\in B} \langle e'|Te'\rangle.$$

□

Proposition 8.2. *Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces, and $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Then*

$$\operatorname{Tr}(T^*T) = \operatorname{Tr}(TT^*) \in [0,\infty].$$

Proof. Note that T^*T and TT^* are positive operators on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. By Lemma. 8.1, their traces are well defined, and we compute them with respect to any orthonormal bases B_i for $\mathcal{H}_i$, $i = 1, 2$,

$$\begin{aligned}
\operatorname{Tr}(T^*T) = \sum_{e\in B_1} \langle e|T^*Te\rangle = \sum_{e\in B_1} \|Te\|^2 &= \sum_{e\in B_1} \sum_{e'\in B_2} |\langle e'|Te\rangle|^2 \\
&= \sum_{e'\in B_2} \sum_{e\in B_1} |\langle e|T^*e'\rangle|^2 \\
&= \sum_{e'\in B_2} \|T^*e'\|^2 \\
&= \sum_{e'\in B_2} \langle e'|TT^*e'\rangle = \operatorname{Tr}(TT^*).
\end{aligned}$$

□

[14] Recall that an orthonormal basis is an orthonormal set whose span (finite linear combinations) is dense in the Hilbert space.

8.3 Square roots and polar decomposition

Given $T \geq 0$, it is a standard result[15] that there exists a unique positive (thus self-adjoint) operator $\sqrt{T} \geq 0$ which squares to T. This is the *square root lemma*[16]. We elaborate on the proof in Reed–Simon's *Methods of Modern Mathematical Physics, Vol. 1*, Theorem VI.9.

Construction of $\sqrt{T}$:

- Consider the power series $\sqrt{1-z} = 1 + c_1 z + c_2 z^2 + \cdots$, absolutely convergent on the open disk $|z| < 1$. Except for $c_0 = 1$, the coeffcients $c_n, n \geq 1$ are negative as the derivatives of $\sqrt{1-z}$ at $z = 0$ are negative. So

$$\sum_{n=1}^{N} |c_n| = -\sum_{n=1}^{N} c_n = 1 - \sum_{n=0}^{N} c_n = 1 - \lim_{r\to 1^-} \sum_{n=0}^{N} c_n r^n \leq 1 - \lim_{r\to 1^-} \sqrt{1-r} = 1.$$

  This holds for arbitrarily large N, so the c_n are absolutely summable. This means that the power series still converges absolutely for $|z| = 1$.

  If $0 \leq r \leq 1$, then $0 \leq 1 - r \leq 1$, and we have an absolutely convergent series

$$\sqrt{1-(1-r)} = 1 + c_1(1-r) + c_2(1-r)^2 + \ldots. \tag{8.4}$$

  The series can be squared to recover $(\sqrt{1-(1-r)})^2 = r$.

- (Exercise.) Using the polarization identity for the inner product, one shows that any positive operator is self-adjoint, and that the operator norm of self-adjoint operators T satisfies

$$\|T\|_{\text{op}} = \sup_{\|x\|\leq 1} |\langle x|Tx\rangle|, \qquad T = T^*. \tag{8.5}$$

  It follows that

$$\|T\|_{\text{op}}^2 = \sup_{\|x\|\leq 1} \|Tx\|^2 = \sup_{\|x\|\leq 1} |\langle Tx|Tx\rangle| = \sup_{\|x\|\leq 1} |\langle x|T^2x\rangle| = \|T^2\|_{\text{op}}. \tag{8.6}$$

[15] Another way to define $\sqrt{T}$ is to first develop the continuous functional calculus for self-adjoint operators.

[16] Another way to construct $\sqrt{T}$ is to invoke spectral theory and continuous functional calculus, but we are not assuming this here.

- Let T be a positive operator with $\|T\|_{\text{op}} \leq 1$. Then $1 - T$ is a positive (and self-adjoint) operator, and we use (8.5) to deduce

$$\|1 - T\|_{\text{op}} = \sup_{\|x\| \leq 1} \langle x|(1-T)x\rangle \leq 1 - \sup_{\|x\| \leq 1} \underbrace{\langle x|Tx\rangle}_{\geq 0} \leq 1.$$

  So we may apply the power series (8.4): there exists an operator

$$\sqrt{T} := 1 + c_1(1-T) + c_2(1-T)^2 + \cdots \in \mathcal{L}(\mathcal{H}),$$

  which squares to T.

  Furthermore,

$$\langle x|\sqrt{T}x\rangle = 1 + \sum_{n \geq 1} \underbrace{c_n}_{<0} \underbrace{\langle x|(1-T)^n x\rangle}_{\in [0,1]} \geq 1 + \underbrace{\sum_{n \geq 1} c_n}_{=-1} = 0,$$

  so $\sqrt{T}$ is a positive operator.

  Rescale to get $\sqrt{T} \geq 0$ for a general positive operator $T \geq 0$.

- For uniqueness, suppose $T \geq 0$ admits another positive square root $S \geq 0$ satisfying $S^2 = T$. Then $ST = S^3 = TS$, i.e., S commutes with T. The absolutely convergent series shows that $\sqrt{T}$ also commutes with S. Then

$$(\sqrt{T}-S)\sqrt{T}(\sqrt{T}-S)+(\sqrt{T}-S)S(\sqrt{T}-S) = (\sqrt{T}^2-S^2)(\sqrt{T}-S) = 0.$$

  As $\sqrt{T}$ and S are positive, they have square roots as well, and so each term summand on the left side of the above equation has the form A^*A and is thus positive. Therefore, they have to vanish individually. The difference of these summands is thus

$$(\sqrt{T} - S)^3 = 0,$$

  which implies $(\sqrt{T} - S)^4 = 0$. But $\sqrt{T} - S$ is self-adjoint, so we may use (8.6) to deduce that

$$\|\sqrt{T} - S\|_{\text{op}}^4 = \|(\sqrt{T} - S)^4\|_{\text{op}} = 0,$$

  Thus $\sqrt{T} - S = 0$.

The following is a useful Lemma exploiting the square root.

Lemma 8.3. *Any $S \in \mathcal{L}(\mathcal{H})$ is a linear combination of four unitaries.*

Proof. Let $C = C^* \in \mathcal{L}(\mathcal{H})$ be self-adjoint, with $\|C\|_{\text{op}} \leq 1$. Then $1 - C^2 \geq 0$, and the operators

$$C \pm i\sqrt{1 - C^2}$$

can be checked to be unitary. Then $C = \frac{1}{2}(C + i\sqrt{1 - C^2}) + \frac{1}{2}(C - i\sqrt{1 - C^2})$ is a linear combination of unitaries. By scaling, any self-adjoint element of $\mathcal{L}(\mathcal{H})$ is a linear combination of two unitaries. Then since any $S \in \mathcal{L}(\mathcal{H})$ is sum of two self-adjoint operators in $\mathcal{L}(\mathcal{H})$,

$$S = \frac{1}{2}(S + S^*) + \frac{1}{2i}(i(S - S^*)),$$

we deduce that S is a linear combination of four unitaries. □

Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces, and $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Since $T^*T \geq 0$, it has a unique positive square root, denoted

$$|T| = \sqrt{T^*T} \in \mathcal{L}(\mathcal{H}_1).$$

The notation $|T|$ generalizes $|z| = \sqrt{\bar{z}z}$ for complex numbers. We know that $z = u|z|$ for some phase rotation $u \in \mathrm{U}(1)$, where u is uniquely determined if $z \neq 0$. The analogue of the phase rotation u is an operator U which restricts to an isometry on $(\ker T)^\perp$.

Definition 5. Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. A *partial isometry* is an operator $U \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ which restricts to an isometry on $(\ker U)^\perp$.

Note that a partial isometry U with $\ker U = \{0\}$ is an isometry in the usual normed space sense ($\|Ux\| = \|x\|$ holds for all $x \in \mathcal{H}$).

A partial isometry U provides orthogonal decompositions,

$$\mathcal{H}_1 = \ker U \oplus (\ker U)^\perp, \qquad \mathcal{H}_2 = \operatorname{Range} U \oplus (\operatorname{Range} U)^\perp,$$

and restricts to a unitary map $(\ker U)^\perp \to \operatorname{Range} U$. It is readily checked that $U^* \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ is also a partial isometry, restricting to a unitary map $\operatorname{Range} U \to (\ker U)^\perp$. Then U^*U is the orthogonal projection onto $(\ker U)^\perp$.

Proposition 8.4 (Polar decomposition)**.** *Each* $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ *has a* polar decomposition

$$T = U|T|, \tag{8.7}$$

such that

$$\ker U = \ker T, \qquad \operatorname{Range} U = \overline{\operatorname{Range} T}. \tag{8.8}$$

Furthermore,

$$|T| = U^*U|T| = U^*T. \tag{8.9}$$

Proof. First, it is clear that U must restrict to the map

$$\begin{aligned} U : \operatorname{Range} |T| &\to \operatorname{Range} T \\ |T|x &\mapsto Tx. \end{aligned} \tag{8.10}$$

Eq. (8.10) is well-defined, since if $|T|x = |T|x'$, then

$$0 = \||T|(x - x')\|^2 = \langle x - x' | T^*T(x - x')\rangle = \|T(x - x')\|^2,$$

so $Tx = Tx'$. This calculation also shows that (8.10) is an isometry, and we extend it to an isometry $\overline{\operatorname{Range} |T|} \to \overline{\operatorname{Range} T}$. On the complementary subspace $\overline{\operatorname{Range} |T|}^{\perp} = (\operatorname{Range} |T|)^{\perp}$, we define U to be zero. We have constructed a partial isometry U on $\mathcal{H}$, satisfying (8.7) and $\operatorname{Range} U = \overline{\operatorname{Range} T}$.

As $|T|$ is self-adjoint, $(\operatorname{Range} |T|)^{\perp} = \ker |T|$ (Exercise). Furthermore,

$$\ker |T| = \ker T, \tag{8.11}$$

since

$$\big\||T|x\big\|^2 = \langle \sqrt{T^*T}x | \sqrt{T^*T}x\rangle = \langle x | T^*Tx\rangle = \|Tx\|^2.$$

Therefore, $\ker U = \ker T$, and (8.8) holds as well.

Finally, we have $U^*T = U^*U|T|$. But

$$U^*U = 0|_{\ker T} \oplus 1|_{\overline{\operatorname{Range} |T|}}.$$

So $|T| = U^*U|T| = U^*T$ holds. □

8.4 Normed $*$-ideal of trace class operators.

Definition 6. Let $\mathcal{H}$ be a Hilbert space. An operator $T \in \mathcal{L}(\mathcal{H})$ is *trace class* if for an(y) orthonormal basis B,

$$\|T\|_1 := \operatorname{Tr}|T| = \sum_{e\in B} \big\langle e\big||T|e\big\rangle < \infty. \tag{8.12}$$

The collection of trace class operators on $\mathcal{H}$ is denoted $\mathcal{L}^1(\mathcal{H})$.

Exercise 8.1. Use polar decomposition to show that for a trace class operator T,

$$\|T\|_{\text{op}} \leq \|T\|_1.$$

Proposition 8.5. *If T is trace class, then T^* is trace class, and $\|T^*\|_1 = \|T\|_1$.*

Proof. Write $T = U|T|$ for the polar decomposition of T, and recall from (8.9) that $U^*U|T| = |T|$. Then

$$|T^*|^2 = TT^* = U|T|^2U^* = U|T|U^*U|T|U^* = (U|T|U^*)^2.$$

So $|T^*|$ and $U|T|U^*$ are positive square roots of the same operator, thus $|T^*| = U|T|U^*$. Pick an orthonormal basis B comprising elements of $\ker T^*$ or $(\ker T^*)^\perp$. Recall that

$$\ker T^* = (\operatorname{Range} T)^\perp \overset{(8.8)}{=} (\operatorname{Range} U)^\perp = \ker U^*,$$

So $\{U^*e\}_{e\in B}$ remains an orthonormal set, which we can extend to an orthonormal basis B'. Thus

$$\begin{aligned}\|T^*\|_1 = \operatorname{Tr}|T^*| = \sum_{e\in B} \big\langle e\big|U|T|U^*e\big\rangle &= \sum_{e\in B} \big\langle U^*e\big||T|U^*e\big\rangle \\ &\leq \sum_{U^*e\in B'} \big\langle U^*e\big||T|U^*e\big\rangle \\ &= \operatorname{Tr}|T| = \|T\|_1.\end{aligned}$$

So T^* is trace class. By the same argument, we also get $\|T\|_1 = \|T^{**}\|_1 \leq \|T\|_1$. □

Definition 7. For a trace class operator T on a Hilbert space $\mathcal{H}$, its *trace* is

$$\operatorname{Tr}(T) := \sum_{e \in B} \langle e|Te\rangle \in \mathbb{C}, \tag{8.13}$$

where B is any orthonormal basis of $\mathcal{H}$.

Theorem 8.6. *The sum in* (8.13) *is well-defined and independent of the choice of orthonormal basis.*

Proof. As (8.13) is an unordered sum whose summands may not be positive, we need to show its absolute convergence.

Use polar decomposition to write

$$T = U|T| = U\sqrt{|T|}\sqrt{|T|}.$$

Then

$$\begin{aligned}
\sum_{e \in B} |\langle e|Te\rangle| &= \sum_{e \in B} \left|\left\langle \sqrt{|T|}U^* e \middle| \sqrt{|T|}e\right\rangle\right| \\
&\le \sum_{e \in B} \left\|\sqrt{|T|}U^* e\right\| \left\|\sqrt{|T|}e\right\| && \text{(Cauchy-Schwarz for } \mathcal{H}\text{)} \\
&\le \Big(\sum_{e \in B} \left\|\sqrt{|T|}U^* e\right\|^2\Big)^{1/2} \Big(\sum_{e \in B} \left\|\sqrt{|T|}e\right\|^2\Big)^{1/2} < \infty,
\end{aligned}$$

where the last line uses Cauchy–Schwarz for $\ell^2(B)$, justified provided the collections

$$\{\|\sqrt{|T|}e\|\}_{e \in B}, \qquad \{\|\sqrt{|T|}U^* e\|\}_{e \in B}$$

are square-summable. Indeed, the first collection has sum-of-squares

$$\sum_{e \in B} \left\|\sqrt{|T|}e\right\|^2 = \sum_{e \in B} \left\langle e \middle| |T| e\right\rangle = \operatorname{Tr}|T| < \infty,$$

while the second collection has sum-of-squares

$$\begin{aligned}
\sum_{e\in B} \big\| \sqrt{|T|} U^* e \big\|^2 &= \sum_{e\in B} \langle e | U \sqrt{|T|} \sqrt{|T|} U^* e \rangle \\
&= \sum_{e\in B} \langle e | \sqrt{|T|} U^* U \sqrt{|T|} e \rangle \qquad \text{(Prop. 8.2)} \\
&= \sum_{e\in B} \| U \sqrt{|T|} e \|^2 \\
&\le \sum_{e\in B} \| \sqrt{|T|} e \|^2 \qquad (\|U\|_{\mathrm{op}} \le 1) \\
&= \operatorname{Tr} |T| < \infty.
\end{aligned}$$

For later use, we record a useful inequality from the above calculations, regarding absolute summability of (8.13),

$$\sum_{e\in B} |\langle e | T e \rangle| \le (\operatorname{Tr}|T|)^{1/2} (\operatorname{Tr}|T|)^{1/2} = \operatorname{Tr}|T| = \|T\|_1. \tag{8.14}$$

Next, consider another orthonormal basis B'. We have

$$\begin{aligned}
\sum_{e\in B} \langle e | T e \rangle &= \sum_{e\in B} \big\langle \sqrt{|T|} U^* e \big| \sqrt{|T|} e \big\rangle \\
&= \sum_{e\in B} \sum_{e'\in B'} \big\langle \sqrt{|T|} U^* e \big| e' \big\rangle \big\langle e' \big| \sqrt{|T|} e \big\rangle. \qquad \text{(Expand in } B' \text{ basis).}
\end{aligned}$$

We may use Fubini's theorem to change the order of summation, because

$$\begin{aligned}
&\sum_{e\in B} \sum_{e'\in B'} \big| \big\langle \sqrt{|T|} U^* e \big| e' \big\rangle \big| \big| \big\langle e' \big| \sqrt{|T|} e \big\rangle \big| \\
&\quad \le \sum_{e\in B} \Big(\sum_{e'\in B'} \big| \big\langle \sqrt{|T|} U^* e \big| e' \big\rangle \big|^2 \Big)^{1/2} \Big(\sum_{e'\in B'} \big| \big\langle e' \big| \sqrt{|T|} e \big\rangle \big|^2 \Big)^{1/2} \\
&\quad = \sum_{e\in B} \big\| \sqrt{|T|} U^* e \big\| \big\| \sqrt{|T|} e \big\|
\end{aligned}$$

was checked to be finite by the $\ell^2(B)$ Cauchy–Schwarz argument. Then

$$\begin{aligned}
\sum_{e\in B}\langle e|Te\rangle &= \sum_{e'\in B'}\sum_{e\in B}\big\langle \sqrt{|T|}U^*e\big|e'\big\rangle\big\langle e'\big|\sqrt{|T|}e\big\rangle && \text{(Fubini)}\\
&= \sum_{e'\in B'}\sum_{e\in B}\big\langle \sqrt{|T|}e'\big|e\big\rangle\big\langle e\big|U\sqrt{|T|}e'\big\rangle && \text{(Adjoint)}\\
&= \sum_{e'\in B'}\big\langle \sqrt{|T|}e'\big|U\sqrt{|T|}e'\big\rangle && \text{(Contract in } B \text{ basis)}\\
&= \sum_{e'\in B'}\sum_{e\in B'}\big\langle \sqrt{|T|}e'\big|e\big\rangle\big\langle e\big|U\sqrt{|T|}e'\big\rangle && \text{(Expand in } B' \text{ basis)}\\
&= \sum_{e\in B'}\sum_{e'\in B'}\big\langle \sqrt{|T|}U^*e\big|e'\big\rangle\big\langle e'\big|\sqrt{|T|}e\big\rangle && \text{(Fubini, adjoint)}\\
&= \sum_{e\in B'}\langle\sqrt{|T|}U^*e\big|\sqrt{|T|}e\big\rangle && \text{(Contract in } B' \text{ basis)}\\
&= \sum_{e\in B'}\big\langle e\big|U\sqrt{|T|}\sqrt{|T|}e\big\rangle = \sum_{e\in B'}\langle e|Te\rangle.
\end{aligned}$$

□

Remark. If we take the spectral theorem/Borel functional calculus for bounded self-adjoint operators for granted, we could argue that $T \in \mathcal{L}^1(\mathcal{H})$ is a linear combination of four positive trace class operators, then use Lemma 8.1 to deduce the basis independence part of Theorem 8.6. But we want to avoid using big machinery unnecessarily.

Exercise 8.2. Let T be a trace class operator. Show that $\operatorname{Tr} T^* = \overline{\operatorname{Tr} T}$.

Proposition 8.7. *$\mathcal{L}^1(\mathcal{H})$ is a vector space, and $\|\cdot\|_1$ defines a norm on it.*

Proof. It is easy to see that scalar multiplication preserves the trace class property. Additivity is tricker. Let $T_1, T_2 \in \mathcal{L}^1(\mathcal{H})$ and consider the polar decompositions

$$T_1 = U_1|T_1|, \qquad T_2 = U_2|T_2|, \qquad (T_1+T_2) = V|T_1+T_2|.$$

For any finite orthonormal set F, we have

$$\begin{aligned}
\sum_{e\in F}\big\langle e\big||T_1+T_2|e\big\rangle &= \sum_{e\in F}\langle e|V^*(T_1+T_2)e\rangle \\
&\le \sum_{i=1}^{2}\sum_{e\in F}\big|\big\langle e\big|V^*U_i|T_i|e\big\rangle\big| \\
&= \sum_{i=1}^{2}\sum_{e\in F}\big|\big\langle \sqrt{|T_i|}U_i^*Ve\big|\sqrt{|T_i|}e\big\rangle\big| \\
&\le \sum_{i=1}^{2}\sum_{e\in F}\big\|\sqrt{|T_i|}U_i^*Ve\big\|\cdot\big\|\sqrt{T_i}e\big\| \\
&\le \sum_{i=1}^{2}\Big(\sum_{e\in F}\big\|\sqrt{|T_i|}U_i^*Ve\big\|^2\Big)^{1/2}\Big(\sum_{e\in F}\big\|\sqrt{|T_i|}e\big\|^2\Big)^{1/2} \\
&= \sum_{i=1}^{2}\Big(\sum_{e\in F}\big\langle e\big|V^*U_i|T_i|U_i^*Ve\big\rangle\Big)^{1/2}\Big(\sum_{e\in F}\big\langle e\big||T_i|e\big\rangle\Big)^{1/2} \\
&\le \sum_{i=1}^{2}\Big(\operatorname{Tr}V^*U_i|T_i|U_i^*V\Big)^{1/2}\big(\operatorname{Tr}|T_i|\big)^{1/2}.
\end{aligned} \tag{8.15}$$

Choose an orthonormal basis B comprising vectors either in $\ker V$ or $(\ker V)^\perp$. As V is a partial isometry, $\{Ve\}_{e\in B}$ is an orthonormal set, and

$$\begin{aligned}
\operatorname{Tr}V^*U_i|T_i|U_i^*V &= \sum_{e\in B}\big\langle e\big|V^*U_i|T_i|U_i^*Ve\big\rangle \\
&= \sum_{e\in B}\big\langle Ve\big|U_i|T_i|U_i^*Ve\big\rangle \\
&\le \operatorname{Tr}(U_i|T_i|U_i^*) \le \operatorname{Tr}|T_i|,
\end{aligned}$$

with the last inequality due to U also being a partial isometry. Putting this inequality into (8.15) yields

$$\sum_{e\in F}\big\langle e\big||T_1+T_2|e\big\rangle \le \sum_{i=1}^{2}\operatorname{Tr}|T_i|.$$

Since this is true for any finite subset F of an orthonormal basis B, we have

$$\|T_1+T_2\|_1 = \sum_{e\in B}\big\langle e\big||T_1+T_2|e\big\rangle \le \operatorname{Tr}|T_1|+\operatorname{Tr}|T_2| = \|T_1\|_1+\|T_2\|_1 < \infty, \tag{8.16}$$

i.e., $T_1 + T_2$ is trace class.

Finally, we check that $\|\cdot\|_1$ is indeed a norm. If $\|T\|_1 = 0$, then $\sum_{e\in B}\langle e||T|e\rangle = 0$, so each (non-negative) term $\langle e||T|e\rangle = \|\sqrt{|T|}e\|^2$ vanishes. It follows that $\sqrt{T} = 0$, thus $|T| = 0$, and also $T = U|T| = 0$. Homogeneity follows from $|\lambda T| = |\lambda||T|, \lambda \in \mathbb{C}$. The triangle inequality was already verified in (8.16). □

Proposition 8.8. *$\mathcal{L}^1(\mathcal{H})$ is a $*$-ideal in $\mathcal{L}(\mathcal{H})$, and the trace norm satisfies*

$$\|ST\|_1 \le \|S\|_{\mathrm{op}}\|T\|_1, \qquad \|TS\|_1 \le \|T\|_1\|S\|_{\mathrm{op}}, \qquad T \in \mathcal{L}^1(\mathcal{H}), S \in \mathcal{L}(\mathcal{H}).$$

In particular, $\|\cdot\|_1$ is a submultiplicative norm on the (non-unital) algebra $\mathcal{L}^1(\mathcal{H})$.

Proof. For the ideal property, let $S \in \mathcal{L}(\mathcal{H})$. By Lemma 8.3, S is a linear combination of four unitaries, so it will suffice to show that TW and WT are trace class whenever T is trace class and W is unitary. Certainly WT is trace class since $|WT| = \sqrt{T^*W^*WT} = \sqrt{T^*T} = |T|$. Also, $W^*|T|W$ is a positive operator squaring to $W^*|T|^2W = W^*T^*TW$, so

$$|TW| = W^*|T|W.$$

As $B' = \{We\}_{e\in B}$ is an orthonormal basis if B is, we have

$$\sum_{e\in B}\big\langle e\big||TW|e\big\rangle = \sum_{e\in B}\big\langle e\big|W^*|T|We\big\rangle = \sum_{We\in B'}\big\langle We\big||T|We\big\rangle = \operatorname{Tr}|T| < \infty,$$

thus TW is trace class as well.

Closure under adjoints was already checked in Prop. 8.5.

For the norm inequalities, we use polar decomposition to write $T = U\sqrt{|T|}\sqrt{|T|}$ as before. Let B be an orthonormal basis. We calculate

$$\begin{aligned}
|\operatorname{Tr}(ST)| &\le \sum_{e\in B}\big|\big\langle e\big|SU\sqrt{|T|}\sqrt{|T|}e\big\rangle\big| \\
&\le \sum_{e\in B}\Big|\big\langle \sqrt{|T|}U^*S^*e\big|\sqrt{|T|}e\big\rangle\Big| \\
&\le \sum_{e\in B}\big\|\sqrt{|T|}U^*S^*\big\|\big\|\sqrt{|T|}e\big\| \\
&\le \Big(\sum_{e\in B}\big\|\sqrt{|T|}U^*S^*\big\|^2\Big)^{1/2}\Big(\sum_{e\in B}\big\|\sqrt{|T|}e\big\|^2\Big)^{1/2},
\end{aligned}$$

using Cauchy–Schwarz in the last step. In the proof of Prop. 8.6 we saw that the collection $\{\|\sqrt{T}e\|\}_{e\in B}$ is square-summable to $\operatorname{Tr}|T|$, while the same method shows that

$$\sum_{e\in B}\|\sqrt{|T|}U^*S^*\|^2=\sum_{e\in B}\left\|SU\sqrt{|T|}\right\|^2\le\|S\|_{\mathrm{op}}^2\sum_{e\in B}\|\sqrt{|T|}\|^2=\|S\|_{\mathrm{op}}\|T\|_1.$$

Substituting into the earlier calculation, we get

$$|\operatorname{Tr}(ST)|\le\|S\|_{\mathrm{op}}\|T\|_1. \tag{8.17}$$

Now let $V|ST|$ be the polar decomposition of ST. Then

$$|ST|=V^*ST=V^*SU|T|,$$

therefore,

$$\begin{aligned}\|ST\|_1=\operatorname{Tr}|ST|&=\operatorname{Tr}(V^*SU|T|)\\&\le\|V^*SU\|_{\mathrm{op}}\|T\|_1 && \text{(Eq. (8.17))}\\&\le\|S\|_{\mathrm{op}}\|T\|_1 && (\|V^*\|_{\mathrm{op}}=\|V\|_{\mathrm{op}},\|U\|_{\mathrm{op}}\le 1)\end{aligned}$$

The other norm inequality follows from Prop. 8.5,

$$\|TS\|_1=\|(TS)^*\|_1=\|(S^*T^*)\|_1\le\|S^*\|_{\mathrm{op}}\|T^*\|_1=\|S\|_{\mathrm{op}}\|T\|_1.$$

Submultiplicativity of $\|\cdot\|_1$ follows readily: If $T_1,T_2\in\mathcal{L}^1(\mathcal{H})$, then

$$\|T_1T_2\|_1\le\|T_1\|_{\mathrm{op}}\|T_2\|_1\le\|T_1\|_1\|T_2\|_1,$$

with the last equality coming from Exercise 8.1. □

Corollary 8.9. *The trace* $\operatorname{Tr}:\mathcal{L}^1(\mathcal{H})\to\mathbb{C}$ *is a continuous linear functional satisfying the cyclicity property*

$$\operatorname{Tr}(ST)=\operatorname{Tr}(TS),\qquad T\in\mathcal{L}^1(\mathcal{H}),\,S\in\mathcal{L}(\mathcal{H}). \tag{8.18}$$

In particular, it is conjugation invariant,

$$\operatorname{Tr}T=\operatorname{Tr}(UTU^{-1}),\qquad T\in\mathcal{L}^1(\mathcal{H}),\,U\in\mathcal{L}(\mathcal{H})^\times. \tag{8.19}$$

Proof. Linearity of $\mathrm{Tr}(-)$ is obvious. Continuity, or equivalently, boundedness, follows from,

$$|\mathrm{Tr}\, T| = \Big| \sum_{e\in B} \langle e|Te\rangle \Big| \leq \sum_{e\in B} |\langle e|Te\rangle| \leq \|T\|_1,$$

where the last inequality is (8.14).

For cyclicity, it suffices by linearity and Lemma 8.3 to consider the case where S is unitary. In this case, $\{S^*e\}_{e\in B}$ is an orthonormal basis if B is one. Noting that ST and TS are trace class by the ideal property (Prop. 8.8), we have

$$\mathrm{Tr}(ST) = \sum_{e\in B} \langle e|STe\rangle = \sum_{e\in B} \langle S^*e|(TS)S^*e\rangle = \mathrm{Tr}(TS).$$

□

Remark 5. It is certainly possible for a product ST to be trace class without either S or T being trace class. So generalizing Eq. (8.18), we might ask: If $S, T \in \mathcal{L}(\mathcal{H})$ are such that ST and TS are trace class, is it still true that

$$\mathrm{Tr}(ST - TS) = \mathrm{Tr}(ST) - \mathrm{Tr}(TS) = 0?$$

The answer is yes, but it is not easy to show. Even more generally, an additive commutator $ST - TS$ could be trace class without either ST or TS being trace class. In this case, must we have

$$\mathrm{Tr}[S, T] = \mathrm{Tr}(ST - TS) = 0?$$

Remarkably, the answer is no, and many interesting examples come from quantum physics!

Remark 6. Eq. (8.19) shows that the trace enjoys basis independence beyond the class of orthonormal bases, indicating that it should have something to do with eigenvalues. However, note that if U is invertible but not unitary, then $\{Ue\}_{e\in B}$ is not an orthonormal basis when B is. (Rather, it is a kind of "Schauder basis".) We cannot compute $\mathrm{Tr}(T)$ by the inner product formula $\sum_{e\in B}\langle Ue|TUe\rangle$, since $(U^{-1})^* \neq U$.

8.5 Finite rank approximation of trace class

Definition 8. Let X, Y be normed vector spaces. A bounded operator $T \in \mathcal{L}(X, Y)$ is *finite rank* if $\operatorname{Range} T$ is finite dimensional. We write $\operatorname{Fin}(X, Y)$ for the set of finite rank operators, and $\operatorname{Fin}(X) = \operatorname{Fin}(X, X)$.

If $T \in \operatorname{Fin}(X, Y)$, then there are $\psi_1, \dots, \psi_N \in Y$ spanning $\operatorname{Range} T$, and coefficient functionals $\ell_1, \dots, \ell_N \in X^*$ such that

$$Tx = \sum_{n=1}^{N} \ell_n(x)\psi_n, \qquad x \in X. \tag{8.20}$$

Proposition 8.10. *Let $\mathcal{H}$ be a Hilbert space. Then $\operatorname{Fin}(\mathcal{H})$ is dense in $\mathcal{L}^1(\mathcal{H})$ (with respect to the trace norm).*

Proof. Let $T \in \operatorname{Fin}(\mathcal{H})$. By the Riesz representation theorem, the canonical form of T in (8.20) may be written as

$$Tx = \sum_{n=1}^{N} \langle \varphi_n | - \rangle \psi_n$$

for some vectors $\varphi_1, \dots, \varphi_n \in \mathcal{H}$. So $(\ker T)^\perp$ is finite-dimensional. Recall from (8.11) that $\ker |T| = \ker T$. We we may pick an orthonormal basis B for $\mathcal{H}$ comprising vectors from either $\ker |T|$ or $(\ker |T|)^\perp$. Then

$$\|T\|_1 = \operatorname{Tr} |T| = \sum_{e \in B} \big\langle e \big| |T| e \big\rangle$$

reduces to a finite sum. Thus $\operatorname{Fin}(\mathcal{H}) \subseteq \mathcal{L}^1(\mathcal{H})$.

Now let $T \in \mathcal{L}^1(\mathcal{H})$, so $\sum_{e \in B} \langle e||T|e\rangle < \infty$. Thus given $\epsilon > 0$, there is a finite subset $F \subseteq B$ such that

$$\sum_{e \in B \setminus F} \big\langle e \big| |T| e \big\rangle < \epsilon.$$

Write P for the orthogonal projection onto $\operatorname{span}\{e : e \in F\}$. Then TP is a finite rank operator, and we will show that TP is ϵ-close to T in trace norm. Write $T - TP = V|T - TP|$ and $T = U|T|$ for the polar decompositions. Noting that

$$(1 - P)e = \begin{cases} e, & e \in B \setminus F, \\ 0, & e \in F, \end{cases}$$

we have

$$\begin{aligned}\|T - TP\|_1 = \sum_{e\in B} \big\langle e\big||T - TP|e\big\rangle &= \sum_{e\in B} \big\langle e\big|V^*T(1-P)e\big\rangle \\ &= \sum_{e\in B\setminus F} \big\langle e\big|V^*U|T|e\big\rangle \\ &\le \sum_{e\in B\setminus F} \big\langle e\big||T|e\big\rangle < \epsilon,\end{aligned}$$

where we use a Cauchy–Schwarz argument over $\ell(B \setminus F)$ in the last line (as in the proof of Prop. 8.8). □