U is well-defined, that is, if $|A|\psi = |A|\phi$ then $A\psi = A\phi$. U is isometric and so extends to an isometry of $\overline{\text{Ran}\,|A|}$ to $\overline{\text{Ran}\,A}$. Extend U to all of $\mathscr{H}$ by defining it to be zero on $(\text{Ran}\,|A|)^\perp$. Since $|A|$ is self-adjoint, $(\text{Ran}\,|A|)^\perp =$ Ker $|A|$. Furthermore, $|A|\psi = 0$ if and only if $A\psi = 0$ so that Ker $|A| =$ Ker A. Thus Ker $U =$ Ker A. Uniqueness is left to the reader. ∎

In Problem 20 of Chapter VII, the reader will prove that U is a strong limit of polynomials in A and A^* so that U is in the "von Neumann algebra" generated by A.

VI.5 Compact operators†

Many problems in classical mathematical physics can be handled by reformulating them in terms of integral equations. A famous example is the Dirichlet problem discussed at the end of this section. Consider the simple operator K, defined in $C[0, 1]$ by

$$(K\varphi)(x) = \int_0^1 K(x, y)\varphi(y)\,dy \tag{VI.4}$$

where the function $K(x, y)$ is continuous on the square $0 \le x, y \le 1$. $K(x, y)$ is called the **kernel** of the **integral operator** K. Since

$$|(K\varphi)(x)| \le \left(\sup_{0 \le x, y \le 1} |K(x, y)|\right)\left(\sup_{0 \le y \le 1} |\varphi(y)|\right)$$

we see that

$$\|K\varphi\|_\infty \le \left(\sup_{0 \le x, y < 1} |K(x, y)|\right)\|\varphi\|_\infty$$

so K is a bounded operator on $C[0, 1]$. K has another property which is very important. Let B_M denote the functions φ in $C[0, 1]$ such that $\|\varphi\|_\infty \le M$. Since $K(x, y)$ is continuous on the square $0 \le x, y \le 1$ and since the square is compact, $K(x, y)$ is uniformly continuous. Thus, given an $\varepsilon > 0$, we can find $\delta > 0$ such that $|x - x'| < \delta$ implies $|K(x, y) - K(x', y)| < \varepsilon$ for all $y \in [0, 1]$. Thus, if $\varphi \in B_M$

$$\begin{aligned}|(K\varphi)(x) - (K\varphi)(x')| &\le \left(\sup_{y \in [0, 1]} |K(x, y) - K(x', y)|\right)\|\varphi\|_\infty \\ &\le \varepsilon M\end{aligned}$$

† A supplement to this section begins on p. 368.

Therefore the functions $K[B_M]$ are equicontinuous. Since they are also uniformly bounded by $\|K\|M$, we can use the Ascoli theorem (Theorem I.28) to conclude that for every sequence $\varphi_n \in B_M$, the sequence $K\varphi_n$ has a convergent subsequence (the limit may not be in $K[B_M]$). Another way of saying this is that the set $K[B_M]$ is **precompact**; that is, its closure is compact in $C[0, 1]$. It is clear that the choice of M was not important so what we have shown is that K takes bounded sets into precompact sets. It is this property which makes the so called "Fredholm alternative" hold for nice integral equations like (VI.4). This section is devoted to studying such operators.

Definition Let X and Y be Banach spaces. An operator $T \in \mathscr{L}(X, Y)$ is called **compact** (or **completely continuous**) if T takes bounded sets in X into precompact sets in Y. Equivalently, T is compact if and only if for every bounded sequence $\{x_n\} \subset X$, $\{Tx_n\}$ has a subsequence convergent in Y.

The integral operator (VI.4) is one example of a compact operator. Another class of examples is:

Example (finite rank operators) Suppose that the range of T is finite dimensional. That is, every vector in the range of T can be written $Tx = \sum_{i=1}^N \alpha_i y_i$, for some fixed family $\{y_i\}_{i=1}^N$ in Y. If x_n is any bounded sequence in X, the corresponding α_i^n are bounded since T is bounded. The usual subsequence trick allows one to extract a convergent subsequence from $\{Tx_n\}$ which proves that T is compact.

An important property of compact operators is given by (compare Problem 34):

Theorem VI.11 A compact operator maps weakly convergent *sequences* into norm convergent sequences.

Proof Suppose $x_n \xrightarrow{w} x$. By the uniform boundedness theorem, the $\|x_n\|$ are bounded. Let $y_n = Tx_n$. Then $\ell(y_n) - \ell(y) = (T'\ell)(x_n - x)$ for any $\ell \in Y^*$. Thus, y_n converges weakly to $y = Tx$ in Y. Suppose that y_n does not converge to y in norm. Then, there is an $\varepsilon > 0$ and a subsequence $\{y_{n_k}\}$ of $\{y_n\}$ so that $\|y_{n_k} - y\| \ge \varepsilon$. Since the sequence $\{x_{n_k}\}$ is bounded and T is compact $\{y_{n_k}\}$ has a subsequence which converges to a $\tilde{y} \ne y$. This subsequence must then also converge weakly to $\tilde{y}$, but this is impossible since y_n converges weakly to y. Thus y_n converges to y in norm. ∎

We note that if X is reflexive then the converse of Theorem VI.11 holds (Problem 20). The following theorem is important since one can use it to prove that an operator is compact by exhibiting it as a norm limit of compact operators or as an adjoint of a compact operator.

Theorem VI.12 Let X and Y be Banach spaces, $T \in \mathscr{L}(X, Y)$.

(a) If $\{T_n\}$ are compact and $T_n \to T$ in the norm topology, then T is compact.
(b) T is compact if and only if T' is compact.
(c) If $S \in \mathscr{L}(Y, Z)$ with Z a Banach space and if T *or* S is compact, then ST is compact.

Proof (a) Let $\{x_m\}$ be a sequence in the unit ball of X. Since T_n is compact for each n, we can use the diagonalization trick of I.5 to find a subsequence of $\{x_m\}$, call it $\{x_{m_k}\}$, so that $T_n x_{m_k} \to y_n$ for each n as $k \to \infty$. Since $\|x_{m_k}\| \le 1$ and $\|T_n - T\| \to 0$, an $\varepsilon/3$-argument shows that the sequence $\{y_n\}$ is Cauchy, so $y_n \to y$. It is not difficult to show using an $\varepsilon/3$ argument that $Tx_{m_k} \to y$. Thus T is compact.
(b) See the Notes and Problem 36.
(c) The proof is elementary (Problem 37). ▮

We are mostly interested in the case where T is a compact operator from a separable Hilbert space to itself, so we will not pursue the general case any further (however, see the discussion in the Notes). We denote the Banach space of compact operators on a separable Hilbert space by $\text{Com}(\mathscr{H})$. By the first example and Theorem VI.12 the norm limit of a sequence of finite rank operators is compact. The converse is also true in the Hilbert space case.

Theorem VI.13 Let $\mathscr{H}$ be a separable Hilbert space. Then every compact operator on $\mathscr{H}$ is the norm limit of a sequence of operators of finite rank.

Proof Let $\{\varphi_j\}_{j=1}^{\infty}$ be an orthonormal set in $\mathscr{H}$. Define

$$\lambda_n = \sup_{\substack{\psi \in [\varphi_1, \ldots, \varphi_n]^\perp \\ \|\psi\| = 1}} \|T\psi\|$$

Clearly, $\{\lambda_n\}$ is monotone decreasing so it converges to a limit $\lambda \ge 0$. We first show that $\lambda = 0$. Choose a sequence $\psi_n \in [\varphi_1, \ldots, \varphi_n]^\perp$, $\|\psi_n\| = 1$, with $\|T\psi_n\| \ge \lambda/2$. Since $\psi_n \xrightarrow{w} 0$, $T\psi_n \to 0$ by Theorem VI.11. Thus, $\lambda = 0$. As a result

$$\sum_{j=1}^{n} (\varphi_j, \cdot) T\varphi_j \to T$$

in norm since λ_n is just the norm of the difference. ▮

We have discussed a wide variety of properties of compact operators but we have not yet described any property which explains our special interest in them. The basic principle which makes compact operators important is the Fredholm alternative: If A is compact, then either $A\psi = \psi$ has a solution or $(I - A)^{-1}$ exists. This is not a property shared by all bounded linear transformations. For example, if A is the operator $(A\varphi)(x) = x\varphi(x)$ on $L^2[0, 2]$, then $A\varphi = \varphi$ has no solutions but $(I - A)^{-1}$ does not exist (as a bounded operator). In terms of "solving equations" the Fredholm alternative is especially nice: It tells us that if for any φ there is at most one ψ with $\psi = \varphi + A\psi$, then there is always exactly one. That is, compactness and uniqueness together imply existence; for an example, see the discussion of the Dirichlet problem at the end of the section.

As one might expect, since the Fredholm alternative holds for finite-dimensional matrices, it is possible to prove the Fredholm alternative for compact operators (in the Hilbert space case) by using the fact that any compact operator A can be written as $A = F + R$ where F has finite rank and R has small norm. Compactness combines very nicely with analyticity so we first prove an elegant result which is of great use in itself (see Sections XI.6, XI.7, XIII.4, and XIII.5).

Theorem VI.14 (analytic Fredholm theorem) Let D be an open connected subset of $\mathbb{C}$. Let $f: D \to \mathscr{L}(\mathscr{H})$ be an analytic operator-valued function such that $f(z)$ is compact for each $z \in D$. Then, either

(a) $(I - f(z))^{-1}$ exists for no $z \in D$.

or

(b) $(I - f(z))^{-1}$ exists for all $z \in D \backslash S$ where S is a discrete subset of D (i.e. a set which has no limit points in D). In this case, $(I - f(z))^{-1}$ is meromorphic in D, analytic in $D \backslash S$, the residues at the poles are finite rank operators, and if $z \in S$ then $f(z)\psi = \psi$ has a nonzero solution in $\mathscr{H}$.

Proof We will prove that near any z_0 either (a) or (b) holds. A simple connectedness argument allows one to convert this into a statement about all of D

(Problem 21). Given $z_0 \in D$, choose an r so that $|z - z_0| < r$ implies $\|f(z) - f(z_0)\| < \frac{1}{2}$ and pick F, an operator with finite rank so that

$$\|f(z_0) - F\| < \tfrac{1}{2}$$

Then, for $z \in D_r$, the disc of radius r about z_0, $\|f(z) - F\| < 1$. By expanding in a geometric series we see that $(I - f(z) + F)^{-1}$ exists and is analytic.

Since F has finite rank, there are independent vectors $\psi_1, \ldots, \psi_N$ so that $F(\varphi) = \sum_{i=1}^N \alpha_i(\varphi)\psi_i$. The $\alpha_i(\cdot)$ are bounded linear functionals on $\mathscr{H}$ so by the Riesz lemma there are vectors $\phi_1, \ldots, \phi_N$ so that $F(\varphi) = \sum_{i=1}^N (\phi_i, \varphi)\psi_i$ for all $\varphi \in \mathscr{H}$. Let $\phi_n(z) = ((I - f(z) + F)^{-1})^*\phi_n$ and

$$g(z) = F(I - f(z) + F)^{-1} = \sum_{n=1}^N (\phi_n(z), \cdot)\psi_n$$

By writing

$$(I - f(z)) = (I - g(z))(I - f(z) + F)$$

we see that $I - f(z)$ is invertible for $z \in D_r$ if and only if $I - g(z)$ is invertible and that $\psi = f(z)\psi$ has a nonzero solution if and only if $\varphi = g(z)\varphi$ has a solution.

If $g(z)\varphi = \varphi$, then $\varphi = \sum_{n=1}^N \beta_n \psi_n$ and the β_n satisfy

$$\beta_n = \sum_{m=1}^N (\phi_n(z), \psi_m)\beta_m \qquad \text{(VI.5a)}$$

Conversely, if (VI.5a) has a solution $\langle \beta_1, \ldots, \beta_N \rangle$, then $\varphi = \sum_{n=1}^N \beta_n \psi_n$ is a solution of $g(z)\varphi = \varphi$. Thus $g(z)\varphi = \varphi$ has a solution if and only if the determinant

$$d(z) = \det\{\delta_{nm} - (\phi_n(z), \psi_m)\} = 0$$

Since $(\phi_n(z), \psi_m)$ is analytic in D_r so is $d(z)$ which means that either $S_r = \{z \mid z \in D_r, d(z) = 0\}$ is a discrete set in D_r or $S_r = D_r$. Now, suppose $d(z) \neq 0$. Then, given ψ, we can solve $(I - g(z))\varphi = \psi$ by setting $\varphi = \psi + \sum_{n=1}^N \beta_n \psi_n$ if we can find β_n satisfying

$$\beta_n = (\phi_n(z), \psi) + \sum_{m=1}^N (\phi_n(z), \psi_m)\beta_m \qquad \text{(VI.5b)}$$

But, since $d(z) \neq 0$, this equation has a solution. Thus $(I - g(z))^{-1}$ exists if and only if $z \notin S_r$.

The meromorphic nature of $(I - f(z))^{-1}$ and the finite rank residues follow from the fact that there is an explicit formula for the β_n in (VI.5b) in terms of cofactor matrices. ∎

This theorem has four important consequences:

Corollary (the Fredholm alternative) If A is a compact operator on $\mathscr{H}$, then either $(I - A)^{-1}$ exists or $A\psi = \psi$ has a solution.

Proof Take $f(z) = zA$ and apply the last theorem at $z = 1$. ∎

Theorem VI.15 (Riesz–Schauder theorem) Let A be a compact operator on $\mathscr{H}$, then $\sigma(A)$ is a discrete set having no limit points except perhaps $\lambda = 0$. Further, any nonzero $\lambda \in \sigma(A)$ is an eigenvalue of finite multiplicity (i.e. the corresponding space of eigenvectors is finite dimensional).

Proof Let $f(z) = zA$. Then $f(z)$ is an analytic compact operator-valued function on the entire plane. Thus $\{z \mid zA\psi = \psi$ has a solution $\psi \neq 0\}$ is a discrete set (it is not the entire plane since it does not contain $z = 0$) and if $1/\lambda$ is not in this discrete set then

$$(\lambda - A)^{-1} = \frac{1}{\lambda}\left(I - \frac{1}{\lambda}A\right)^{-1}$$

exists. The fact that the nonzero eigenvalues have finite multiplicity follows immediately from the compactness of A. ∎

Theorem VI.16 (the Hilbert–Schmidt theorem) Let A be a self-adjoint compact operator on $\mathscr{H}$. Then, there is a complete orthonormal basis, $\{\phi_n\}$, for $\mathscr{H}$ so that $A\phi_n = \lambda_n \phi_n$ and $\lambda_n \to 0$ as $n \to \infty$.

Proof For each eigenvalue of A choose an orthonormal basis for the set of eigenvectors corresponding to the eigenvalue. The collection of all these vectors, $\{\phi_n\}$, is an orthonormal set since eigenvectors corresponding to distinct eigenvalues are orthogonal. Let $\mathscr{M}$ be the closure of the span of $\{\phi_n\}$. Since A is self-adjoint and $A: \mathscr{M} \to \mathscr{M}$, $A: \mathscr{M}^\perp \to \mathscr{M}^\perp$. Let $\tilde{A}$ be the restriction of A to $\mathscr{M}^\perp$. Then $\tilde{A}$ is self-adjoint and compact since A is. By the Riesz–Schauder theorem, if any $\lambda \neq 0$ is in $\sigma(\tilde{A})$, it is an eigenvalue of $\tilde{A}$ and thus of A. Therefore the spectral radius of $\tilde{A}$ is zero since the eigenvectors of A are in $\mathscr{M}$. Because $\tilde{A}$ is self-adjoint, it is the zero operator on $\mathscr{M}^\perp$ by Theorem VI.6. Thus, $\mathscr{M}^\perp = \{0\}$ since if $\varphi \in \mathscr{M}^\perp$, then $A\varphi = 0$ which implies that $\varphi \in \mathscr{M}$. Therefore, $\mathscr{M} = \mathscr{H}$.

The fact that $\lambda_n \to 0$ is a consequence of the first part of the Riesz–Schauder theorem which says that each nonzero eigenvalue has finite multiplicity and the only possible limit point of the λ_n is zero. ∎

Theorem VI.17 (canonical form for compact operators) Let A be a compact operator on $\mathscr{H}$. Then there exist (not necessarily complete) orthonormal sets $\{\psi_n\}_{n=1}^N$ and $\{\phi_n\}_{n=1}^N$ and positive real numbers $\{\lambda_n\}_{n=1}^N$ with $\lambda_n \to 0$ so that

$$A = \sum_{n=1}^{N} \lambda_n(\psi_n, \cdot)\phi_n \tag{VI.6}$$

The sum in (VI.6), which may be finite or infinite, converges in norm. The numbers, $\{\lambda_n\}$, are called the **singular values** of A.

Proof Since A is compact, so is A^*A (Theorem VI.12). Thus A^*A is compact and self-adjoint. By the Hilbert–Schmidt theorem, there is an orthonormal set $\{\psi_n\}_{n=1}^N$ so that $A^*A\psi_n = \mu_n\psi_n$ with $\mu_n \neq 0$ and so that A^*A is the zero operator on the subspace orthogonal to $\{\psi_n\}_{n=1}^N$. Since A^*A is positive, each $\mu_n > 0$. Let λ_n be the positive square root of μ_n and set $\phi_n = A\psi_n/\lambda_n$. A short calculation shows that the ϕ_n are orthonormal and that

$$A\psi = \sum_{n=1}^{N} \lambda_n(\psi_n, \psi)\phi_n \quad \blacksquare$$

The proof shows that the singular values of A are precisely the eigenvalues of $|A|$.

We conclude with a classical example.

Example (Dirichlet problem) The main impetus for the study of compact operators arose from the use of integral equations in attempting to solve the classical boundary value problems of mathematical physics. We briefly describe this method. Let D be an open bounded region in $\mathbb{R}^3$ with a smooth boundary surface ∂D. The Dirichlet problem for Laplace's equation is: given a continuous function f on ∂D, find a function u, twice differentiable in D and continuous on $\bar{D}$, which satisfies

$$\begin{aligned} \Delta u(x) &= 0 \qquad x \in D \\ u(x) &= f(x) \qquad x \in \partial D \end{aligned}$$

Let $K(x, y) = (x - y, n_y)/2\pi|x - y|^3$ where n_y is the outer normal to ∂D at the point $y \in \partial D$. Then, as a function of x, $K(x, y)$ satisfies $\Delta_x K(x, y) = 0$ in the interior which suggests that we try to write u as a superposition

$$u(x) = \int_{\partial D} K(x, y)\varphi(y)\, dS(y) \tag{VI.6a}$$

where $\varphi(y)$ is some continuous function on ∂D and dS is the usual surface measure. Indeed, for $x \in D$, the integral makes perfectly good sense and

$\Delta u(x) = 0$ in D. Furthermore, if x_0 is any point in ∂D and $x \to x_0$ from *inside* D, it can be proven that

$$u(x) \to -\varphi(x_0) + \int_{\partial D} K(x_0, y)\varphi(y)\, dS(y) \tag{VI.6b}$$

If $x \to x_0$ from outside D, the minus is replaced by a plus. Also,

$$\int_{\partial D} K(x_0, y)\varphi(y)\, dS(y)$$

exists and is a continuous function on ∂D if φ is a continuous function on ∂D. The proof depends on the fact that the boundary of D is smooth which implies that for $x, y \in \partial D$, $(x - y, n_y) \approx c|x - y|^2$ as $x \to y$.

Since we wish $u(x) = f(x)$ on ∂D, the whole question reduces to whether we can find φ so that

$$f(x) = -\varphi(x) + \int_{\partial D} K(x, y)\varphi(y)\, dS(y), \qquad x \in \partial D$$

Let T: $C(\partial D) \to C(\partial D)$ be defined by

$$T\varphi = \int_{\partial D} K(x, y)\varphi(y)\, dS(y)$$

Not only is T bounded but (as we will shortly see) T is also compact. Thus, by the Fredholm alternative, either $\lambda = 1$ is in the point spectrum of T in which case there is a $\psi \in C(\partial D)$ such that $(I - T)\psi = 0$, or $-f = (I - T)\varphi$ has a unique solution for each $f \in C(\partial D)$. If u is defined by (VI.6a) with ψ replacing φ, then $u \equiv 0$ in D by the maximum principle. Further, $\partial u/\partial n$ is continuous across ∂D and therefore equals zero on ∂D. By an integration by parts this implies that $u \equiv 0$ outside ∂D. Therefore, by (VI.6b), $2\psi(x) \equiv 0$ on ∂D, so the first alternative does not hold.

The idea of the compactness proof is the following. Let

$$K_\delta(x, z) = \frac{(x - z, n_z)}{|x - z|^3 + \delta}$$

If $\delta > 0$, the kernel K_δ is continuous, so, by the discussion at the beginning of this section, the corresponding integral operators T_δ, are compact. To prove that T is compact, we need only show that $\|T - T_\delta\| \to 0$ as $\delta \to 0$. By the estimate

$$|(T_\delta f)(x) - (Tf)(x)| \leq \|f\|_\infty \int_{\partial D} |K(x, z) - K_\delta(x, z)|\, dS(z)$$

we must only show that the integral converges to zero uniformly in x as $\delta \to 0$. To prove this, divide the integration region into the set where $|x-z| \geq \varepsilon$ and its complement. For fixed ε, the kernels converge uniformly on the first region. By using the fact that K is integrable, the contribution from the second region can be made arbitrarily small for ε sufficiently small.

VI.6 The trace class and Hilbert–Schmidt ideals

In the last section we saw that compact operators have many nice properties and are useful for applications. It is therefore important to have effective criteria for determining when a given operator is compact or, better yet, general statements about whole classes of operators. In this section we will prove that the integral operator

$$(Tf)(x) = \int_M K(x, y) f(y)\, d\mu(y)$$

on $L^2(M, d\mu)$ is compact if $K(\cdot, \cdot) \in L^2(M \times M, d\mu \otimes d\mu)$. First we will develop the trace, a tool which is of great interest in itself. Theorem VI.12 shows that $\text{Com}(\mathscr{H})$, the compact operators on a separable Hilbert space $\mathscr{H}$, form a Banach space. At the conclusion of the section, we will compute the dual and double dual of $\text{Com}(\mathscr{H})$. These calculations illustrate the difference between the weak Banach space topology on $\mathscr{L}(\mathscr{H})$ and the weak operator topology and give a foretaste of the structure of abstract von Neumann algebras which we will study later.

The trace is a generalization of the usual notion of the sum of the diagonal elements of a matrix, but because infinite sums are involved, not all operators will have a trace. The construction of the trace is analogous to the construction of the Lebesgue integral where one first defines $\int f\, d\mu$ for $f \geq 0$; it has values in $[0, \infty]$, including ∞. Then $\mathscr{L}^1$ is defined as those f so that $\int |f|\, d\mu < \infty$. $\mathscr{L}^1$ is a vector space and $f \mapsto \int f\, d\mu$ a linear functional. Similarly we first define the trace, $\text{tr}(\cdot)$, on the positive operators; $A \to \text{tr}\, A$ has values in $[0, \infty]$. We then define the **trace class**, $\mathscr{I}_1$, to be all $A \in \mathscr{L}(\mathscr{H})$ such that $\text{tr}\, |A| < \infty$. We will then show that $\text{tr}(\cdot)$ is a linear functional on $\mathscr{I}_1$ with the right properties.

Theorem VI.18 Let $\mathscr{H}$ be a separable Hilbert space, $\{\varphi_n\}_{n=1}^\infty$ an orthonormal basis. Then for any positive operator $A \in \mathscr{L}(\mathscr{H})$ we define $\text{tr}\, A = \sum_{n=1}^\infty (\varphi_n, A\varphi_n)$. The number $\text{tr}\, A$ is called the **trace of** A and is independent of the orthonormal basis chosen. The trace has the following properties:

(a) $\text{tr}(A + B) = \text{tr}\, A + \text{tr}\, B$.
(b) $\text{tr}(\lambda A) = \lambda\, \text{tr}\, A$ for all $\lambda \geq 0$.
(c) $\text{tr}(UAU^{-1}) = \text{tr}\, A$ for any unitary operator U.
(d) If $0 \leq A \leq B$, then $\text{tr}\, A \leq \text{tr}\, B$.

Proof Given an orthonormal basis $\{\varphi_n\}_{n=1}^\infty$, define $\text{tr}_\varphi(A) = \sum_{n=1}^\infty (\varphi_n, A\varphi_n)$. If $\{\psi_m\}_{m=1}^\infty$ is another orthonormal basis then

$$\begin{aligned}
\text{tr}_\varphi(A) &= \sum_{n=1}^\infty (\varphi_n, A\varphi_n) = \sum_{n=1}^\infty \|A^{1/2}\varphi_n\|^2 \\
&= \sum_{n=1}^\infty \left(\sum_{m=1}^\infty |(\psi_m, A^{1/2}\varphi_n)|^2 \right) \\
&= \sum_{m=1}^\infty \left(\sum_{n=1}^\infty |(A^{1/2}\psi_m, \varphi_n)|^2 \right) \\
&= \sum_{m=1}^\infty \|A^{1/2}\psi_m\|^2 \\
&= \sum_{m=1}^\infty (\psi_m, A\psi_m) \\
&= \text{tr}_\psi(A)
\end{aligned}$$

Since all the terms are positive, interchanging the sums is allowed.

Properties (a), (b), and (d) are obvious. To prove (c) we note that if $\{\varphi_n\}$ is an orthonormal basis, then so is $\{U\varphi_n\}$. Thus,

$$\text{tr}(UAU^{-1}) = \text{tr}_{(U\varphi)}(UAU^{-1}) = \text{tr}_\varphi(A) = \text{tr}(A)\,. \quad \blacksquare$$

Definition An operator $A \in \mathscr{L}(\mathscr{H})$ is called **trace class** if and only if $\text{tr}\, |A| < \infty$. The family of all trace class operators is denoted by $\mathscr{I}_1$.

The basic properties of $\mathscr{I}_1$ are given in the following:

Theorem VI.19 $\mathscr{I}_1$ is a $*$-ideal in $\mathscr{L}(\mathscr{H})$, that is,

(a) $\mathscr{I}_1$ is a vector space.
(b) If $A \in \mathscr{I}_1$ and $B \in \mathscr{L}(\mathscr{H})$, then $AB \in \mathscr{I}_1$ and $BA \in \mathscr{I}_1$.
(c) If $A \in \mathscr{I}_1$, then $A^* \in \mathscr{I}_1$.

Proof (a) Since $|\lambda A| = |\lambda|\, |A|$ for $\lambda \in \mathbb{C}$, $\mathscr{I}_1$ is closed under scalar multiplication. Now, suppose that A and B are in $\mathscr{I}_1$, we wish to prove that