Christopher Heil

Metrics, Norms, Inner Products and Operator Theory

Chapter 8

Online Extra Chapter on Integral Operators

Last Updated: October 3, 2020

Chapter 8
Integral Operators

In our development of metrics, norms, inner products, and operator theory in Chapters 1–7 we only tangentially considered topics that involved the use of Lebesgue measure, such as the Lebesgue spaces $L^p(E)$. In this extra chapter we will study *integral operators*, which are particular types of operators on $L^p(E)$. These results require the use of the use of Lebesgue measure and the Lebesgue integral, and we will assume in this chapter that the reader is familiar with the theory of Lebesgue measure and the Lebesgue integral. A brief review of Lebesgue measure and integral can be found on the author's website for this text:

`http://people.math.gatech.edu/~heil/metricnote`

Throughout this chapter, a "measurable set" will mean a Lebesgue measurable set, a "measurable function" will mean a Lebesgue measurable function, and all integrals will be Lebesgue integrals.

8.1 Integral Operators

Integral operators are an important special class of linear operators that act on function spaces. We will define these operators and explore some of their properties. As the integral in question will not always exist, we make a "formal" definition of an integral operator. Following mathematical tradition, a *formal statement* often means an entirely *informal* statement. In particular, here it means that integral need not make sense for some f, and for those f we simply leave $L_k f$ undefined.

Definition 8.1.1 (Integral Operator). Let k be a fixed measurable function on $\mathbb{R}^2$. The *integral operator* L_k *with kernel* k is formally defined by

$$L_k f(x) \;=\; \int_{-\infty}^{\infty} k(x,y)\, f(y)\, dy. \tag{8.1}$$

That is, if f is a measurable function on $\mathbb{R}$, then $L_k f$ is the function defined by equation (8.1), as long as this integral is well-defined for a.e. $x \in \mathbb{R}$. Otherwise $L_k f$ is not defined. ♢

The use of the word *kernel* in this definition should not be confused with its other meaning as the nullspace of an operator. It should always be clear from context which meaning of "kernel" is intended.

Remark 8.1.2. An integral operator is a natural generalization of the ordinary matrix-vector product. To see this, let A be an $m \times n$ matrix with entries a_{ij} and let u be a vector in $\mathbb{C}^n$. Then $Au \in \mathbb{C}^m$, and its components are

$$(Au)_i \;=\; \sum_{j=1}^{n} a_{ij}\, u_j, \qquad i = 1, \dots, m.$$

Thus, the function values $k(x,y)$ are entirely analogous to the entries a_{ij} of the matrix A, and the values $L_k f(x)$ are analogous to the entries $(Au)_i$. ♢

As an illustration, we will look at integral operators whose kernels take the especially simple form given in the following definition.

Definition 8.1.3. Given two functions g, h whose domain is the real line, the *tensor product* of g and h is the function $g \otimes h$ on $\mathbb{R}^2$ defined by

$$(g \otimes h)(x,y) \;=\; g(x)\,\overline{h(y)}, \qquad x, y \in \mathbb{R}. \qquad ♢ \tag{8.2}$$

Sometimes the complex conjugate is omitted in the definition of a tensor product $g \otimes h$, or the complex conjugate is placed on g instead of h. It will be most convenient for our purposes to place the complex conjugate on the function h.

Example 8.1.4 (Tensor Product Kernels). Given functions g, $h \in L^2(\mathbb{R})$, let $k = g \otimes h$. If $f \in L^2(\mathbb{R})$, then the inner product $\langle f, h \rangle$ is well-defined and $L_k f$ is

$$L_k f(x) \;=\; \int_{-\infty}^{\infty} g(x)\,\overline{h(y)}\, f(y)\, dy \;=\; g(x) \int_{-\infty}^{\infty} f(y)\,\overline{h(y)}\, dy \;=\; \langle f, h \rangle\, g(x).$$

That is, $L_k f = \langle f, h \rangle\, g$, so L_k maps every vector in $L^2(\mathbb{R})$ to a scalar multiple of g. If $g = 0$ or $h = 0$ then L_k is the zero operator; otherwise the range of L_k is the one-dimensional subspace spanned by g. In either case L_k is a bounded map of $L^2(\mathbb{R})$ into itself, because

$$\|L_k f\|_2 \;=\; |\langle f, h \rangle|\, \|g\|_2 \;\le\; \|f\|_2\, \|h\|_2\, \|g\|_2 \;=\; C\, \|f\|_2,$$

where $C = \|h\|_2\, \|g\|_2 < \infty$. ♢

We often identify a tensor product function $k = g \otimes h$ with the integral operator $L_k = L_{g \otimes h}$ whose kernel is $k = g \otimes h$ in the following way.

Notation 8.1.5. Given g and h in $L^2(\mathbb{R})$, we let the symbols $g \otimes h$ denote either the tensor product function given in equation (8.2), or the operator whose rule is

$$(g \otimes h)(f) \;=\; \langle f, h\rangle\, g, \qquad f \in L^2(\mathbb{R}). \tag{8.3}$$

It is usually clear from context whether $g \otimes h$ is meant to denote a function or an operator. $\diamondsuit$

We extend this notation to arbitrary Hilbert spaces by replacing $L^2(\mathbb{R})$ with H in equation (8.3). That is, if g, h are vectors in a Hilbert space H, then we define $g \otimes h$ to be the *operator*

$$(g \otimes h)(f) \;=\; \langle f, h\rangle\, g, \qquad f \in H.$$

We call $g \otimes h$ the *tensor product* of g and h. If g is a unit vector, then $g \otimes g$ is the orthogonal projection of H onto $\operatorname{span}\{g\}$.

Problems

8.1.6. This problem does not involve integral operators, but it does require the use of Lebesgue measure and integration to determine properties of a linear operator. Let $1 \le p \le \infty$ be fixed, and assume that $\phi\colon \mathbb{R} \to \mathbb{C}$ is measurable. We formally define an operator M_ϕ by $M_\phi f = f\phi$ for measurable functions f.

(a) Prove that if $\phi \in L^\infty(\mathbb{R})$, then M_ϕ is a bounded linear map from $L^p(\mathbb{R})$ to $L^p(\mathbb{R})$, and its operator norm is $\|M_\phi\| = \|\phi\|_\infty$.

(b) Show that if $f\phi \in L^p(\mathbb{R})$ for every $f \in L^p(\mathbb{R})$, then $\phi \in L^\infty(\mathbb{R})$.

(c) Determine a necessary and sufficient condition on ϕ that implies that $M_\phi\colon L^p(\mathbb{R}) \to L^p(\mathbb{R})$ is injective.

(d) Determine a necessary and sufficient condition on ϕ that implies that $M_\phi\colon L^p(\mathbb{R}) \to L^p(\mathbb{R})$ is surjective.

(e) Show directly that if M_ϕ is injective but not surjective, then the inverse mapping $M_\phi^{-1}\colon \operatorname{range}(M_\phi) \to L^p(\mathbb{R})$ is unbounded.

(f) Prove that $M_\phi\colon L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is compact if and only if $\phi = 0$ a.e.

8.2 Integral Operators with Square-Integrable Kernels

It is usually not obvious how properties of a kernel k relate to properties of the integral operator L_k determined by k. We will show that if the kernel k is a square-integrable function on $\mathbb{R}^2$, then L_k is a bounded mapping on $L^2(\mathbb{R})$. In the proof of this theorem, note that f belongs to $L^2(\mathbb{R})$ while $k \in L^2(\mathbb{R}^2)$. We use $\|f\|_2$ and $\|k\|_2$ to denote the L^2-norms of these functions, the domains $\mathbb{R}$ or $\mathbb{R}^2$ being clear from context.

Theorem 8.2.1. *If $k \in L^2(\mathbb{R}^2)$, then the integral operator L_k given by equation* (8.1) *defines a bounded mapping of $L^2(\mathbb{R})$ into itself, and its operator norm satisfies $\|L_k\| \le \|k\|_2$.*

Proof. Suppose that k belongs to $L^2(\mathbb{R}^2)$, and fix $f \in L^2(\mathbb{R})$. By Fubini's Theorem, the function $k_x(y) = k(x,y)$ belongs to $L^2(\mathbb{R})$ for a.e. x. For those x,

$$L_k f(x) \;=\; \int_{-\infty}^{\infty} k_x(y)\, f(y)\, dy \;=\; \int_{-\infty}^{\infty} k_x(y)\, \overline{\overline{f(y)}}\, dy \;=\; \langle k_x, \overline{f}\,\rangle.$$

Therefore $L_k f(x)$ is defined for almost every x. We still must show that $L_k f \in L^2(\mathbb{R})$ and $\|L_k f\|_2 \le \|k\|_2\, \|f\|_2$.

Step 1. Suppose that f and k are both nonnegative a.e. Then $k(x,y)\, f(y)$ is nonnegative and measurable on $\mathbb{R}^2$, so $L_k f(x) = \int k(x,y)\, f(y)\, dy$ is a measurable function of x by Tonelli's Theorem. We estimate its L^2-norm by applying the Cauchy–Bunyakovski–Schwarz Inequality:

$$\begin{aligned}
\|L_k f\|_2^2 \;&=\; \int_{-\infty}^{\infty} |L_k f(x)|^2\, dx \\
&=\; \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} k(x,y)\, f(y)\, dy \right|^2 dx \\
&\le\; \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |k(x,y)|^2\, dy \right) \left(\int_{-\infty}^{\infty} |f(y)|^2\, dy \right) dx \\
&=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |k(x,y)|^2\, dy \;\; \|f\|_2^2\, dx \\
&=\; \|f\|_2^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |k(x,y)|^2\, dy\, dx \\
&=\; \|k\|_2^2\, \|f\|_2^2 \;<\; \infty.
\end{aligned} \tag{8.4}$$

Therefore $L_k f \in L^2(\mathbb{R})$, and $\|L_k f\|_2 \le \|k\|_2\, \|f\|_2$.

Step 2. Now let $f \in L^2(\mathbb{R})$ and $k \in L^2(\mathbb{R}^2)$ be arbitrary functions. Write

$$f = (f_1^+ - f_1^-) + i(f_2^+ - f_2^-) \qquad \text{and} \qquad k = (k_1^+ - k_1^-) + i(k_2^+ - k_2^-),$$

where each function $f_\ell^\pm$ and $k_j^\pm$ is nonnegative a.e. By Step 1, each of the functions $L_{k_j^\pm}(f_\ell^\pm)$ is measurable and belongs to $L^2(\mathbb{R})$. Since $L_k f$ is a finite linear combination of the sixteen functions $L_{k_j^\pm}(f_\ell^\pm)$, we conclude that $L_k f$ is measurable and belongs to $L^2(\mathbb{R})$. Now that we know that $L_k f$ is measurable, we can repeat the same set of calculations that appear in equation (8.4) and conclude that $\|L_k f\|_2 \le \|k\|_2 \|f\|_2$. Since this inequality holds for all $f \in L^2(\mathbb{R})$, it follows that L_k maps $L^2(\mathbb{R})$ boundedly into itself, and its operator norm satisfies the estimate $\|L_k\| \le \|k\|_2$. □

We will give several improvements to Theorem 8.2.1 below. In particular, Theorem 8.2.3 will show that L_k is a *compact* operator when $k \in L^2(\mathbb{R}^2)$, and in Theorem 8.4.8 we will prove that an operator on $L^2(\mathbb{R})$ is *Hilbert–Schmidt* if and only if it can be written as an integral operator whose kernel is square-integrable.

In order to prove that L_k is compact when k is square-integrable, we first need the following construction of a convenient orthonormal basis for $L^2(\mathbb{R}^2)$; the proof is Problem 8.2.4. This construction will also be used later in the proof of Theorem 8.4.8.

Lemma 8.2.2. *Let $\{e_n\}_{n\in\mathbb{N}}$ be an orthonormal basis for $L^2(\mathbb{R})$, and set*

$$e_{mn}(x,y) \;=\; (e_m \otimes e_n)(x,y) \;=\; e_m(x)\,\overline{e_n(y)}, \qquad x,y \in \mathbb{R}.$$

Then $\{e_{mn}\}_{m,n\in\mathbb{N}}$ is an orthonormal basis for $L^2(\mathbb{R}^2)$. ◇

Now we will show that the integral operator L_k is compact when k is in $L^2(\mathbb{R}^2)$. Using the terminology that we will introduce in Section 8.4, this says that if k is square-integrable, then L_k is a *Hilbert–Schmidt operator.*

Theorem 8.2.3. *Choose $k \in L^2(\mathbb{R}^2)$, and let L_k be the corresponding integral operator.*

(a) *Let $\{e_{mn}\}_{m,n\in\mathbb{N}}$ be an orthonormal basis for $L^2(\mathbb{R}^2)$ of the type constructed in Lemma 8.2.2. Define*

$$k_N \;=\; \sum_{m=1}^{N}\sum_{n=1}^{N} \langle k, e_{mn}\rangle\, e_{mn}.$$

Then the corresponding integral operator L_{k_N} is bounded and has finite rank.

(b) *$L_k - L_{k_N}$ is the integral operator whose kernel is $k - k_N$, and*

$$\|L_k - L_{k_N}\| \;\le\; \|k - k_N\|_2.$$

(c) *L_k is a compact operator.*

Proof. By Theorem 8.2.1 we know that L_k is bounded and $\|L_k\| \le \|k\|_2$. Let $\{e_n\}_{n\in\mathbb{N}}$ be any orthonormal basis for $L^2(\mathbb{R})$, and define

$$e_{mn}(x,y) \;=\; (e_m \otimes e_n)(x,y) \;=\; e_m(x)\,\overline{e_n(y)}, \qquad x,y \in \mathbb{R}.$$

Then Lemma 8.2.2 establishes that $\{e_{mn}\}_{m,n\in\mathbb{N}}$ is an orthonormal basis for $L^2(\mathbb{R}^2)$. Since $k \in L^2(\mathbb{R}^2)$, we therefore have

$$k \;=\; \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \langle k, e_{mn}\rangle\, e_{mn},$$

where this series converges unconditionally in $L^2(\mathbb{R}^2)$. For each $N \in \mathbb{N}$ define an approximation to k by setting

$$k_N \;=\; \sum_{m=1}^{N}\sum_{n=1}^{N} \langle k, e_{mn}\rangle\, e_{mn}.$$

Note that $k_N \to k$ in L^2-norm. The integral operator

$$L_{k_N} f(x) \;=\; \int k_N(x,y)\, f(y)\, dy, \qquad f \in L^2(\mathbb{R}).$$

is an approximation to L_k. It is bounded since $k_N \in L^2(\mathbb{R}^2)$. Since the sums involved are finite, interchanges are allowed in the following calculation:

$$\begin{aligned}
L_{k_N} f(x) \;&=\; \int k_N(x,y)\, f(y)\, dy \\
&=\; \int \sum_{m=1}^{N}\sum_{n=1}^{N} \langle k, e_{mn}\rangle\, e_{mn}(x,y)\, f(y)\, dy \\
&=\; \sum_{m=1}^{N}\sum_{n=1}^{N} \langle k, e_{mn}\rangle \int e_m(x)\, \overline{e_n(y)}\, f(y)\, dy \\
&=\; \sum_{m=1}^{N}\sum_{n=1}^{N} \langle k, e_{mn}\rangle\, \langle f, e_n\rangle\, e_m(x).
\end{aligned}$$

This is an equality of functions in $L^2(\mathbb{R})$, i.e., it holds for almost every x. Hence $L_{k_N} f \in \operatorname{span}\{e_1, \dots, e_N\}$, and therefore L_{k_N} has finite rank.

Since $k - k_N \in L^2(\mathbb{R}^2)$, we know that $L_k - L_{k_N}$ is bounded. Also

$$\|L_k - L_{k_N}\| \;\le\; \|k - k_N\|_2 \;\to\; 0 \quad \text{as } N \to \infty,$$

so $L_{k_N} \to L$ in operator norm. Thus L is the limit in operator norm of operators that have finite rank. Since each L_{k_N} is compact, Corollary 7.5.2 implies that L_k is compact. $\square$

Problems

8.2.4. Prove Lemma 8.2.2.

8.2.5. Define $k(x,y) = e^{2\pi i(x-y)}$ for $(x,y) \in [0,1]^2$. Prove that the corresponding integral operator $L_k \colon L^2[0,1] \to L^2[0,1]$ is an orthogonal projection.

8.2.6. Define $\mathcal{L} \colon L^2(\mathbb{R}^2) \to \mathcal{B}(L^2(\mathbb{R}))$ by $\mathcal{L}(k) = L_k$ for $k \in L^2(\mathbb{R}^2)$. Prove that $\mathcal{L}$ is linear and continuous.

8.3 Schur's Test

Theorem 8.2.1 shows that if the kernel k of an integral operator is square-integral then L_k is bounded. Our next theorem will give a different sufficient condition on the kernel k that ensures that L_k is a bounded operator on $L^2(\mathbb{R})$. This result is sometimes called *Schur's Test* (not to be confused with *Schur's Lemma*, which is an entirely distinct result).

Theorem 8.3.1 (Schur's Test). *Let k be a measurable function on $\mathbb{R}^2$ that satisfies the mixed-norm conditions*

$$\begin{aligned} C_1 &= \operatorname*{ess\,sup}_{x\in\mathbb{R}} \int_{-\infty}^{\infty} |k(x,y)|\,dy < \infty, \\ C_2 &= \operatorname*{ess\,sup}_{y\in\mathbb{R}} \int_{-\infty}^{\infty} |k(x,y)|\,dx < \infty. \end{aligned} \tag{8.5}$$

Then the integral operator L_k defined by equation (8.1) *is a bounded mapping of $L^2(\mathbb{R})$ into itself, and its operator norm satisfies $\|L_k\| \le (C_1C_2)^{1/2}$.*

Proof. As in the proof of Theorem 8.2.1, measurability of L_kf is most easily shown by first considering nonnegative f and k, and then extending to the general case. We omit the details and assume that L_kf is measurable for all $f \in L^2(\mathbb{R})$. Applying the Cauchy–Bunyakovski–Schwarz Inequality, we see that

$$\begin{aligned} \|L_kf\|_2^2 &= \int_{-\infty}^{\infty} |L_kf(x)|^2\,dx \\ &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} k(x,y)\,f(y)\,dy \right|^2 dx \\ &\le \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |k(x,y)|^{1/2} \cdot |k(x,y)|^{1/2}\,|f(y)|\,dy \right)^2 dx \end{aligned}$$

$$
\begin{aligned}
&\leq \int_{-\infty}^{\infty} \biggl(\int_{-\infty}^{\infty} |k(x,y)|\,dy\biggr) \biggl(\int_{-\infty}^{\infty} |k(x,y)|\,|f(y)|^2\,dy\biggr)\,dx \\
&\leq \int_{-\infty}^{\infty} C_1 \int_{-\infty}^{\infty} |k(x,y)|\,|f(y)|^2\,dy\,dx \\
&= C_1 \int_{-\infty}^{\infty} |f(y)|^2 \int_{-\infty}^{\infty} |k(x,y)|\,dx\,dy \\
&\leq C_1 \int_{-\infty}^{\infty} |f(y)|^2\, C_2\,dy \\
&= C_1 C_2\, \|f\|_2^2.
\end{aligned}
$$

We were allowed to interchange the order of integration in the preceding calculation because the integrand $|f(y)|^2\,|k(x,y)|$ is nonnegative, and therefore Tonelli's Theorem is applicable. It follows that L_k is bounded, and its operator norm satisfies $\|L_k\| \leq (C_1C_2)^{1/2}$. □

By applying Hölder's Inequality instead of CBS, we can obtain a similar result showing that L_k is bounded from $L^p(\mathbb{R})$ to $L^p(\mathbb{R})$ for all $1 \leq p \leq \infty$ (the proof is assigned as Problem 8.3.5).

Theorem 8.3.2 (Schur's Test). *If k satisfies the conditions given in equation* (8.5)*, then L_k is a bounded mapping of $L^p(\mathbb{R})$ into itself for each index $1 \leq p \leq \infty$, and its operator norm satisfies $\|L_k\| \leq C_1^{1/p'} C_2^{1/p}$.* ♢

8.3.1 Convolution

Convolution is an important operation that plays central roles in harmonic analysis in mathematics and signal processing in engineering. We studied the convolution of sequences in Section 6.7, and now we will consider convolution of functions. For more details on convolution than what is presented here, we refer to [Heil11] and [Heil19].

Let f and g be measurable functions whose domain is the real line $\mathbb{R}$. Formally, the convolution of f and g is the function $f * g$ defined by

$$(f * g)(x) \;=\; \int_{-\infty}^{\infty} f(y)\, g(x-y)\,dy,$$

as long as this integral makes sense. If we set $k(x,y) = g(x-y)$ then, at least formally, $L_k f = f*g$. However, even if g belongs to $L^2(\mathbb{R})$, the kernel $k(x,y) = g(x-y)$ will not belong to $L^2(\mathbb{R}^2)$, so Theorem 8.2.1 is not applicable. On the other hand, if g is *integrable* then Schur's Test can be applied, and it yields the results stated next (the proof is assigned as Problem 8.3.5).

Theorem 8.3.3 (Young's Inequality). *If $1 \le p \le \infty$, $f \in L^p(\mathbb{R})$, and $g \in L^1(\mathbb{R})$, then $f * g \in L^p(\mathbb{R})$ and*

$$\|f * g\|_p \le \|f\|_p \, \|g\|_1. \qquad \diamondsuit$$

Taking $p = 1$, we see that if f and g both belong to $L^1(\mathbb{R})$, then so does $f * g$. Hence $L^1(\mathbb{R})$ is closed under convolution. Moreover, the convolution of integrable functions is commutative and associative and is *submultiplicative* in the sense that $\|f * g\|_1 \le \|f\|_1 \, \|g\|_1$. Using the terminology of functional analysis, this says that $L^1(\mathbb{R})$ is a *commutative Banach algebra* with respect to convolution.

8.3.2 Integral Operators on Other Domains

We have focused so far on functions whose domain is the real line, kernels that are defined on $\mathbb{R}^2$, and the corresponding integral operators. However, all of the results we derived can be extended to functions that are defined on other domains. For example, let $E \subseteq \mathbb{R}^m$ and $F \subseteq \mathbb{R}^n$ be measurable sets, and let k be any measurable function that is defined on $E \times F$. Given a function f defined on F, we formally define $L_k f$ on the domain E by the rule

$$L_k f(x) \;=\; \int_F k(x,y)\, f(y)\, dy, \qquad x \in E.$$

The analogues of Theorem 8.2.1 and 8.3.1 for this setting are stated in the following result (whose proof is assigned as Problem 8.3.5).

Theorem 8.3.4. *Let $E \subseteq \mathbb{R}^m$ and $F \subseteq \mathbb{R}^n$ be measurable sets, and let k be a measurable function on $E \times F$. Then the following statements hold.*

(a) *If $k \in L^2(E \times F)$, then L_k maps $L^2(F)$ boundedly into $L^2(E)$.*

(b) *If*

$$\operatorname*{ess\,sup}_{x \in E} \int_F |k(x,y)|\, dy \;<\; \infty, \qquad \operatorname*{ess\,sup}_{y \in F} \int_E |k(x,y)|\, dx \;<\; \infty,$$

then L_k maps $L^p(F)$ boundedly into $L^p(E)$ for $1 \le p \le \infty$. $\diamondsuit$

Problems

8.3.5. Prove Theorems 8.3.2, 8.3.3, and 8.3.4.

8.3.6. The *Volterra operator* V is the integral operator $Vf(x) = \int_0^x f(y)\,dy$. Prove that V maps $L^p[0,1]$ continuously into itself for each $1 \le p \le \infty$, and likewise $V\colon C[0,1] \to C[0,1]$ is bounded.

8.3.7. Prove the following weighted version of Schur's Test. Assume that k is a measurable function on $\mathbb{R}^2$ and there are strictly positive measurable functions u, v on $\mathbb{R}$ such that

$$\int_{-\infty}^{\infty} |k(x,y)|\,v(y)\,dy \;\le\; C_1\,u(x), \qquad \text{a.e. } x,$$
$$\int_{-\infty}^{\infty} |k(x,y)|\,u(x)\,dx \;\le\; C_2\,v(y), \qquad \text{a.e. } y.$$

Prove that the integral operator L_k whose kernel is k defines a bounded mapping of $L^2(\mathbb{R})$ into itself.

8.3.8. Prove the following matrix version of Schur's Test. Let $A = [a_{ij}]_{i,j\in\mathbb{N}}$ be an infinite matrix such that

$$C_1 \;=\; \sup_{i\in\mathbb{N}} \sum_{j=1}^{\infty} |a_{ij}| \;<\; \infty, \qquad C_2 \;=\; \sup_{j\in\mathbb{N}} \sum_{i=1}^{\infty} |a_{ij}| \;<\; \infty.$$

Given $x = (x_k)_{k\in\mathbb{N}}$, define $Ax = \big((Ax)_i\big)_{i\in\mathbb{N}}$ where $(Ax)_i = \sum_{j=1}^{\infty} a_{ij}x_j$. Prove that $A\colon \ell^p \to \ell^p$ is bounded and linear for each $1 \le p \le \infty$, and $\|A\|_{\ell^p\to\ell^p} \le C_1^{1/p'} C_2^{1/p}$.

8.3.9. The convolution of bi-infinite sequences $x = (x_k)_{k\in\mathbb{Z}}$ and $y = (y_k)_{k\in\mathbb{Z}}$ is formally defined to be the sequence $x*y$ whose kth component is $(x*y)_k = \sum_{j=-\infty}^{\infty} x_j\,y_{k-j}$. Prove Young's Inequality for sequences: If $x \in \ell^p(\mathbb{Z})$ and $y \in \ell^1(\mathbb{Z})$, then $x * y \in \ell^1(\mathbb{Z})$ and $\|x*y\|_p \le \|x\|_p\,\|y\|_1$.

8.4 Hilbert–Schmidt Operators

Throughout Section 8.4, H ***will denote a separable Hilbert space.***

A Hilbert–Schmidt operator is a special type of compact operator on a Hilbert space. Although the definition of a Hilbert–Schmidt operator makes sense when H is nonseparable (by using a complete orthonormal system instead of a countable orthonormal basis), we will restrict our discussion to Hilbert–Schmidt operators on separable spaces. We do allow H to be either finite or infinite dimensional in this section, so when we say "let $\{e_n\}$ be an orthonormal basis for H," this basis has one of the forms $\{e_n\}_{n=1}^d$ or $\{e_n\}_{n=1}^\infty$.

8.4.1 Definition and Basic Properties

Suppose that H is a separable infinite-dimensional Hilbert space and $\{e_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for H. In this case $\sum \|e_n\|^2 = \infty$. A Hilbert–Schmidt operator will be required to perform some "compression" on H, in the sense that the image $\{Te_n\}_{n\in\mathbb{N}}$ of the orthonormal basis will need to satisfy the following condition.

Definition 8.4.1. An operator $T \in \mathcal{B}(H)$ is a *Hilbert–Schmidt operator* if there exists an orthonormal basis $\{e_n\}$ for H such that

$$\sum_n \|Te_n\|^2 < \infty. \qquad \diamondsuit$$

For example, the identity operator on an infinite-dimensional Hilbert space is not a Hilbert–Schmidt operator. At the other extreme, if H is finite-dimensional then every bounded linear operator on H is Hilbert–Schmidt.

The next result shows that the choice of orthonormal basis in Definition 8.4.1 is irrelevant (the proof is assigned as Problem 8.4.9).

Theorem 8.4.2. *Fix $T \in \mathcal{B}(H)$. For each orthonormal basis $\mathcal{E} = \{e_n\}$ for H, define*

$$S(\mathcal{E}) \;=\; \biggl(\sum_n \|Te_n\|^2\biggr)^{1/2}.$$

Then $S(\mathcal{E})$ is independent of the choice of orthonormal basis $\mathcal{E}$. That is, if $S(\mathcal{E})$ is finite for one orthonormal basis then it is finite for all and takes the same value for every orthonormal basis $\mathcal{E}$, while if $S(\mathcal{E})$ is infinite for one orthonormal basis then it is infinite for all. $\diamondsuit$

We use the following notation to denote the space of Hilbert–Schmidt operators on H.

Definition 8.4.3. The space of *Hilbert–Schmidt operators* on H is

$$\mathcal{B}_2(H) \;=\; \bigl\{T \in \mathcal{B}(H) \;:\; T \text{ is Hilbert–Schmidt}\bigr\}.$$

The *Hilbert–Schmidt norm* of $T \in \mathcal{B}_2(H)$ is

$$\|T\|_{\mathcal{B}_2} \;=\; \biggl(\sum_n \|Te_n\|^2\biggr)^{1/2},$$

where $\{e_n\}$ is any orthonormal basis for H. $\diamondsuit$

Remark 8.4.4. Writing

$$\|T\|_{\mathcal{B}_2}^2 \;=\; \sum_n \|Te_n\|^2 \;=\; \sum_n \sum_m |\langle Te_n, e_m\rangle|^2, \tag{8.6}$$

we see that the Hilbert–Schmidt norm can be viewed as the ℓ^2-norm of the matrix representation of T with respect to the orthonormal basis $\{e_n\}$.

In particular, if $A = [a_{ij}]$ is an $m \times m$ matrix and we let $\{e_n\}$ be the standard basis for $\mathbb{C}^m$, then we see that $\|A\|_{\mathcal{B}_2}^2 = \sum_{i,j} |a_{ij}|^2$. In this context $\|A\|_{\mathcal{B}_2}$ is sometimes called the *Frobenius norm* of A. $\diamondsuit$

The following theorem justifies the use of the word "norm" in connection with $\|\cdot\|_{\mathcal{B}_2}$, and presents some of the other basic properties of Hilbert–Schmidt operators (the proof is assigned as Problem 8.4.9).

Theorem 8.4.5. *The following statements hold.*

(a) *The Hilbert–Schmidt norm dominates the operator norm, i.e.,*

$$\|T\| \le \|T\|_{\mathcal{B}_2} \quad \text{for all } T \in \mathcal{B}_2(H).$$

(b) $\|\cdot\|_{\mathcal{B}_2}$ *is a norm, and* $\mathcal{B}_2(H)$ *is complete with respect to* $\|\cdot\|_{\mathcal{B}_2}$.

(c) $\mathcal{B}_2(H)$ *is closed under adjoints, and* $\|T^*\|_{\mathcal{B}_2} = \|T\|_{\mathcal{B}_2}$ *for all* $T \in \mathcal{B}_2(H)$.

(d) *If* $T \in \mathcal{B}_2(H)$ *and* $A \in \mathcal{B}(H)$, *then* $AT,\, TA \in \mathcal{B}_2(H)$, *and*

$$\|AT\|_{\mathcal{B}_2} \le \|A\|\,\|T\|_{\mathcal{B}_2}, \qquad \|TA\|_{\mathcal{B}_2} \le \|A\|\,\|T\|_{\mathcal{B}_2}.$$

Consequently, $\mathcal{B}_2(H)$ *is a* two-sided ideal *in* $\mathcal{B}(H)$.

(e) *Every bounded, finite-rank linear operator on* H *is Hilbert–Schmidt.*

(f) *Every Hilbert–Schmidt operator on* H *is compact.*

(g) *The space* $\mathcal{B}_{00}(H)$ *of finite-rank operators is dense in* $\mathcal{B}_2(H)$ *with respect to both the operator norm and the Hilbert–Schmidt norm.* $\diamondsuit$

Combining statements (e) and (f) from the preceding exercise, we see that

$$\mathcal{B}_{00}(H) \subseteq \mathcal{B}_2(H) \subseteq \mathcal{B}_0(H). \tag{8.7}$$

If H is finite-dimensional, then these three spaces coincide. If H is infinite-dimensional then each of the two inclusions in equation (8.7) is proper (see Problem 8.4.11).

8.4.2 Singular Numbers and the Hilbert–Schmidt Norm

We will give an equivalent formulation of Hilbert–Schmidt operators in terms of their singular numbers (see Section 7.9 for more details on singular numbers and the Singular Value Decomposition).

A compact operator $T \in \mathcal{B}_0(H)$ need not have any eigenvalues. However, T^*T is both compact and self-adjoint, so by the Spectral Theorem there exists a countable orthonormal sequence $\{e_n\}_{n\in J}$ and corresponding nonzero real numbers $(\mu_n)_{n\in J}$ such that

$$T^*Tf \;=\; \sum_{n\in J} \mu_n \, \langle f, e_n\rangle \, e_n, \qquad f \in H.$$

The scalars μ_n are the nonzero eigenvalues of T^*T. Problem 7.3.14 implies that T^*T is a positive operator, so $\mu_n > 0$ for each n.

If T has finite rank then T^*T also has finite rank. Therefore T^*T has only finitely many nonzero eigenvalues, so in this case the index set J is

$$J \;=\; \{1, \dots, N\} \quad \text{where } N = \dim\bigl(\mathrm{range}(T^*T)\bigr).$$

If T does not have finite rank, then T^*T does not have finite rank either (see Problem 7.2.17), so in this case the the index set is $J = \mathbb{N}$. Furthermore, $\mu_n \to 0$ in this case.

Definition 8.4.6 (Singular Numbers). Let $T\colon H \to H$ be compact, and let $(\mu_n)_{n\in J}$ and $\{e_n\}_{n\in J}$ be as constructed above. The *singular numbers* or *singular values* of T are

$$s_n \;=\; \mu_n^{1/2}, \qquad n \in J,$$

taken in decreasing order:

$$s_1 \;\ge\; s_2 \;\ge\; \cdots \;>\; 0.$$

The vectors e_n are corresponding *singular vectors* of T. ♢

Now we reformulate the definition of Hilbert–Schmidt operators in terms of singular numbers (the proof is assigned as Problem 8.4.9).

Theorem 8.4.7. *Let $T\colon H \to H$ be compact, and let $s = (s_n)_{n\in J}$ be the sequence of singular numbers of T. Then the following statements hold.*

(a) *T is Hilbert–Schmidt if and only if $s \in \ell^2(J)$.*

(b) *If T is Hilbert–Schmidt, then*

$$\|T\|_{\mathcal{B}_2} \;=\; \|s\|_2 \;=\; \biggl(\sum_{n\in J} s_n^2\biggr)^{1/2}.$$

(c) *If T is a self-adjoint Hilbert–Schmidt operator, then*

$$\|T\|_{\mathcal{B}_2} \;=\; \biggl(\sum_{n\in J} \lambda_n^2\biggr)^{1/2}$$

where $(\lambda_n)_{n\in J}$ is the sequence of nonzero eigenvalues of T. ♢

8.4.3 Hilbert–Schmidt Integral Operators

Now we focus on integral operators on $L^2(\mathbb{R})$. By Theorem 8.2.1, if a kernel k belongs to $L^2(\mathbb{R}^2)$, then the corresponding integral operator L_k is a bounded operator on $L^2(\mathbb{R})$. Theorem 8.2.3 showed further that L_k is compact in this case. According to the next result, L_k is Hilbert–Schmidt when $k \in L^2(\mathbb{R}^2)$, and conversely every Hilbert–Schmidt operator on $L^2(\mathbb{R})$ can be written as an integral operator whose kernel is square integrable.

Theorem 8.4.8 (Hilbert–Schmidt Kernel Theorem).

(a) *If $k \in L^2(\mathbb{R}^2)$, then the integral operator L_k with kernel k is Hilbert–Schmidt, and its operator norm is $\|L_k\|_{\mathcal{B}_2} = \|k\|_2$.*

(b) *If T is a Hilbert–Schmidt operator on $L^2(\mathbb{R})$, then there exists a function $k \in L^2(\mathbb{R}^2)$ such that $T = L_k$.*

(c) *$k \mapsto L_k$ is an isometric isomorphism of $L^2(\mathbb{R}^2)$ onto $\mathcal{B}_2(L^2(\mathbb{R}))$.*

Proof. (a) Assume that $k \in L^2(\mathbb{R}^2)$. We already know that L_k is compact. To show that L_k is Hilbert–Schmidt, let $\{e_n\}_{n\in\mathbb{N}}$ be any orthonormal basis for $L^2(\mathbb{R})$. If we set

$$e_{mn}(x,y) \;=\; (e_m \otimes e_n)(x,y) \;=\; e_m(x)\,\overline{e_n(y)}, \tag{8.8}$$

then Lemma 8.2.2 shows that $\{e_{mn}\}_{m,n\in\mathbb{N}}$ is an orthonormal basis for $L^2(\mathbb{R}^2)$.

Fix integers $m,\, n \in \mathbb{N}$. Since $k \cdot \overline{e_{mn}} \in L^1(\mathbb{R}^2)$, Fubini's Theorem allows us to interchange integrals in the following calculation:

$$\begin{aligned}
\langle k, e_{mn}\rangle \;&=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(x,y)\,\overline{e_m(x)}\,e_n(y)\,dx\,dy \\
&=\; \int_{-\infty}^{\infty}\biggl(\int_{-\infty}^{\infty} k(x,y)\,e_n(y)\,dy\biggr)\overline{e_m(x)}\,dx \\
&=\; \int_{-\infty}^{\infty} L_k e_n(x)\,\overline{e_m(x)}\,dx \\
&=\; \langle L_k e_n, e_m\rangle.
\end{aligned}$$

Consequently, the Hilbert–Schmidt norm of L_k is

$$\begin{aligned}
\|L_k\|_{\mathcal{B}_2}^2 \;=\; \sum_{n=1}^{\infty}\|L_k e_n\|^2 \;&=\; \sum_{n=1}^{\infty}\biggl(\sum_{m=1}^{\infty}|\langle L_k e_n, e_m\rangle|^2\biggr) \\
&=\; \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|\langle k, e_{mn}\rangle|^2 \;=\; \|k\|_2^2 \;<\; \infty.
\end{aligned}$$

Therefore L_k is Hilbert–Schmidt and $\|L_k\|_{\mathcal{B}_2} = \|k\|_2$. This also shows that $k \mapsto L_k$ is an isometric map of $L^2(\mathbb{R})$ into $\mathcal{B}_2(L^2(\mathbb{R}))$.

(b) Let T be any Hilbert–Schmidt operator on $L^2(\mathbb{R})$. Choose an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ for $L^2(\mathbb{R})$, and let e_{mn} be defined as in equation (8.8). Since T is Hilbert–Schmidt, equation (8.6) shows that the sequence $\{\langle Te_m, e_n\rangle\}_{m,n\in\mathbb{N}}$ is square-summable. As $\{e_{mn}\}_{m,n\in\mathbb{N}}$ is an orthonormal basis for $L^2(\mathbb{R}^2)$, Theorem 5.7.1 therefore implies that the series

$$k \;=\; \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \langle Te_n, e_m\rangle\, e_{mn} \tag{8.9}$$

converges unconditionally in $L^2(\mathbb{R}^2)$, i.e., it converges regardless of what ordering we impose on the index set $\mathbb{N}\times\mathbb{N}$. Let L_k be the integral operator whose kernel is k. By part (a), L_k is Hilbert–Schmidt and $\|L_k\|_{\mathcal{B}_2} = \|k\|_2$. We will show that $L_k = T$.

By part (a), the mapping $\mathcal{L}\colon L^2(\mathbb{R}) \to \mathcal{B}_2(L^2(\mathbb{R}))$ given by $\mathcal{L}(k) = L_k$ is an isometry. Since unconditional convergence is preserved by continuous linear maps (see Problem 6.5.8), we have

$$\begin{aligned} L_k \;=\; \mathcal{L}(k) \;&=\; \mathcal{L}\biggl(\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \langle Te_n, e_m\rangle\, e_{mn}\biggr) \\ &=\; \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \langle Te_n, e_m\rangle\, L_{e_{mn}}, \end{aligned} \tag{8.10}$$

where the series in equation (8.10) converges unconditionally in $\mathcal{B}_2(L^2(\mathbb{R}))$. Since the Hilbert–Schmidt norm dominates the operator norm, this series also converges unconditionally with respect to the operator norm. Consequently, for each function $f \in L^2(\mathbb{R})$ we have

$$L_k f \;=\; \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \langle Te_n, e_m\rangle\, L_{e_{mn}} f, \tag{8.11}$$

where this series converges unconditionally in L^2-norm.

Now, $L_{e_{mn}}$ is the rank-one operator $L_{e_{mn}} f = \langle f, e_n\rangle\, e_m$ (see Example 8.1.4). Substituting this into equation (8.11) and using the unconditionality of the convergence to reorder the summations (see Problem 8.4.15), we compute that

$$\begin{aligned} L_k f \;&=\; \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \langle Te_n, e_m\rangle\, L_{e_{mn}} f, \\ &=\; \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \langle Te_n, e_m\rangle\, \langle f, e_n\rangle\, e_m \\ &=\; \sum_{n=1}^{\infty} \langle f, e_n\rangle \biggl(\sum_{m=1}^{\infty} \langle Te_n, e_m\rangle\, e_m\biggr) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{n=1}^{\infty} \langle f, e_n \rangle\, Te_n \\
&= T\biggl(\sum_{n=1}^{\infty} \langle f, e_n \rangle\, e_n\biggr) \qquad \text{(since } T \text{ is continuous)} \\
&= Tf.
\end{aligned}
$$

Thus $L_k f = Tf$ for every f, so $T = L_k$.

(c) Part (a) shows that $\mathcal{L}(k) = L_k$ is an isometry, and part (b) shows that $\mathcal{L}$ is surjective. □

In summary, $\mathcal{B}_2(L^2(\mathbb{R}))$ is isometrically isomorphic to the Hilbert space $L^2(\mathbb{R}^2)$. Therefore the space of Hilbert–Schmidt operators inherits a Hilbert space structure from $L^2(\mathbb{R}^2)$. As every Hilbert–Schmidt operator is defined by a kernel, we can write the inner product on $\mathcal{B}_2(L^2(\mathbb{R}))$ as

$$\langle L_k, L_h \rangle \;=\; \langle k, h \rangle, \qquad k, h \in L^2(\mathbb{R}^2).$$

With respect to this inner product, $\mathcal{L}(k) = L_k$ defines a unitary map of $L^2(\mathbb{R}^2)$ onto $\mathcal{B}_2(L^2(\mathbb{R}))$.

Problems

8.4.9. Prove Theorems 8.4.2, 8.4.5, and 8.4.7.

8.4.10. Given g, $h \in H$, show that the Hilbert–Schmidt norm of the tensor-product operator $g \otimes h$ is $\|g \otimes h\|_{\mathcal{B}_2} = \|g\|\,\|h\|$.

8.4.11. Show that if H is infinite-dimensional, then $\mathcal{B}_{00}(H) \neq \mathcal{B}_2(H)$ and $\mathcal{B}_2(H) \neq \mathcal{B}_0(H)$.

8.4.12. Define a kernel k on $[0,1]^2$ by $k(x,y) = |x-y|$. Show that the integral operator L_k on $L^2[0,1]$ is Hilbert–Schmidt, and find $\sum \lambda_n^2$, where $\{\lambda_n\}_{n\in\mathbb{N}}$ is the set of nonzero eigenvalues of L_k.

8.4.13. Exhibit an orthonormal basis for $\mathcal{B}(L^2(\mathbb{R}))$.

8.4.14. Let X be a Banach space, and assume that vectors $x_{mn} \in X$, m, $n \in \mathbb{N}$, satisfy $\sum_m \sum_n \|x_{mn}\| < \infty$. Show that if $\sigma\colon \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ is a bijection, then the following series all converge and are equal as indicated:

$$\sum_{m=1}^{\infty}\biggl(\sum_{n=1}^{\infty} x_{mn}\biggr) \;=\; \sum_{n=1}^{\infty}\biggl(\sum_{m=1}^{\infty} x_{mn}\biggr) \;=\; \sum_{k=1}^{\infty} x_{\sigma(k)}.$$

8.4.15. Prove that the conclusion of Problem 8.4.14 remains valid if we replace the hypothesis of absolute convergence with unconditional convergence, i.e., we assume that the series $\sum_{(m,n)\in\mathbb{N}^2} c_{mn}$ converges unconditionally in X.