

KERNELS OF TRACE CLASS OPERATORS

CHRIS BRISLAWN

(Communicated by John B. Conway)

ABSTRACT. Let $X \subset \mathbf{R}^n$ and let K be a trace class operator on $L^2(X)$ with corresponding kernel $K(x, y) \in L^2(X \times X)$. An integral formula for $\text{tr } K$, proven by Duflo for continuous kernels, is generalized for arbitrary trace class kernels. This formula is shown to be equivalent to one involving the factorization of K into a product of Hilbert-Schmidt operators. The formula and its derivation yield two new necessary conditions for traceability of a Hilbert-Schmidt kernel, and these conditions are also shown to be sufficient for positive operators. The proofs make use of the boundedness of the Hardy-Littlewood maximal function on $L^2(\mathbf{R}^n)$.

1. Introduction. Given a kernel function, $K(x, y)$, what are necessary and sufficient function-theoretic conditions on $K(x, y)$ so that the corresponding integral operator on $L^2(X)$ is traceable, and how does one compute $\text{tr } K$ in terms of the given kernel? These questions arise quite naturally in, for example, the theory of unitary group representations. For the case when $X \subset \mathbf{R}^n$, the main result of this paper, Theorem 3.1, answers the second question completely by proving a generalization of the formula

$$\text{tr } K = \int K(x, x) dx$$

that holds for arbitrary trace class kernels, and we show how this result can be related to the factorization of a trace class operator into a product of Hilbert-Schmidt operators. Theorem 3.1 also provides a partial answer to the first question by proving that a Hilbert-Schmidt operator, K , is traceable only if

$$\int MK(x, x) dx < \infty,$$

where MK is the Hardy-Littlewood maximal function of the kernel $K(x, y)$. A second necessary condition for traceability is the existence and integrability of a kernel determined by the Lebesgue Differentiation Theorem. In §4 we show that these conditions are also sufficient for traceability of positive operators.

It is difficult to produce conditions on the function-theoretic properties of a kernel that are both strong enough to imply traceability yet weak enough to include a fairly large class of kernels. One technique is to show that $K(x, y)$ is in the domain of a sufficiently high power of some selfadjoint operator on $L^2(\mathbf{R}^{2n})$ with known

Received by the editors October 16, 1987 and, in revised form, December 18, 1987 and February 25, 1988. The results were presented on July 6, 1988 to the AMS 1988 Summer Research Institute on Operator Theory at the University of New Hampshire.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 47B10; Secondary 22E27, 47A70, 47D45.

This research was partially supported by NSF grant DMS-8500941.

©1988 American Mathematical Society
0002-9939/88 \$1.00 + \$.25 per page

spectrum. Sugiura [7, Proposition IV.3.5] uses the two-dimensional Laplacian to prove that if $K(\theta, \phi)$ is a C^2 -function on the torus, \mathbf{T}^2 , then the integral operator on $L^2(\mathbf{T}^1)$ defined by

$$Kf(\theta) = \int_{\mathbf{T}^1} K(\theta, \phi)f(\phi) d\phi$$

is traceable. Following Sugiura, one can prove an analogous theorem for kernel operators on $L^2(\mathbf{R}^n)$ using the $2n$ -dimensional Hermite operator, H_{2n} . For $M \geq n + 1$ the M th power of the inverse Hermite operator, H_{2n}^{-M} , is Hilbert-Schmidt (cf. [4, p. 21]). This fact implies the following:

PROPOSITION 1.1. *Let $K(x, y) \in L^2(\mathbf{R}^{2n})$. If $K \in \text{Dom}(H_{2n}^M)$, $M \geq n + 1$, then the operator*

$$Kf(x) = \int_{\mathbf{R}^n} K(x, y)f(y) dy$$

is a trace class operator on $L^2(\mathbf{R}^n)$.

$\text{Dom}(H_{2n}^M)$ means here a domain in $L^2(\mathbf{R}^{2n})$ on which H_{2n}^M has a bounded inverse. This includes Schwartz functions on \mathbf{R}^{2n} but is otherwise quite restrictive.

The idea of averaging $K(x, y)$ over cubes appears in a paper by Weidmann [9], where he provides a sufficient condition for traceability of a positive Hilbert-Schmidt operator on $L^2(\mathbf{R}^n)$; we shall discuss this result and prove its converse in §4 as an immediate corollary of Theorem 3.1. We will then provide a generalization of the following theorem, which can be found in Tricomi [8] or Riesz and Nagy [3].

PROPOSITION 1.2 (MERCER'S THEOREM). *Let $I = [0, 1]$. Suppose that $P(x, y) \in L^2(I \times I)$ is a continuous, positive-definite Hilbert-Schmidt kernel; then the eigenfunction expansion*

$$P = \sum_{i=1}^{\infty} \lambda_i \phi_i \otimes \overline{\phi_i}$$

converges pointwise in $I \times I$, absolutely and uniformly:

$$P(x, y) = \sum \lambda_i \phi_i(x) \overline{\phi_i(y)}.$$

Consequently,

$$\text{tr } P = \int_0^1 P(x, x) dx.$$

2. Averaging on cubes. Let C_r be the n -dimensional cube of radius r centered at the origin in \mathbf{R}^n , and let $C_r(x)$ be the translated cube centered at $x \in \mathbf{R}^n$:

$$C_r = [-r, r]^n, \quad C_r(x) = x + C_r.$$

Let A_r be the linear operator that averages a function $f \in L^1_{\text{loc}}(\mathbf{R}^n)$ over cubes of radius r :

$$(2.1) \quad A_r f(x) = \frac{1}{|C_r(x)|} \int_{C_r(x)} f(t) dt = \frac{1}{|C_r|} \int_{C_r} f(x+t) dt.$$

We use $|C_r|$ to denote the Lebesgue measure of C_r . For each $x \in \mathbf{R}^n$, $A_r f(x)$ is a continuous function of $r \in (0, \infty)$, and for each $r > 0$, $A_r f(x)$ is a continuous

function of x . The Hardy-Littlewood maximal function of f is

$$(2.2) \quad Mf(x) = \sup_{r>0} \frac{1}{|C_r|} \int_{C_r} |f(x+t)| dt.$$

$Mf(x)$ has a well-defined value in $[0, \infty]$ at every point $x \in \mathbf{R}^n$. The Maximal Theorem states that if $f \in L^p(\mathbf{R}^n)$, $1 \leq p \leq \infty$, then Mf is finite almost everywhere, and if $1 < p \leq \infty$ then

$$(2.3) \quad \|Mf\|_p \leq C_p \|f\|_p,$$

where C_p is a constant depending only on p and the dimension, n . By the Lebesgue Differentiation Theorem, if $f \in L^1_{loc}(\mathbf{R}^n)$ then

$$(2.4) \quad \lim_{r \rightarrow 0} A_r f(x) = f(x) \quad \text{a.e. } [dx].$$

A standard reference on these results is Stein [6].

For $r > 0$, definitions (2.1) and (2.2) show that

$$(2.5) \quad |A_r f(x)| \leq Mf(x)$$

at every point $x \in \mathbf{R}^n$, so if $1 < p \leq \infty$ the Maximal Theorem (2.3) implies that A_r is a bounded linear operator on $L^p(\mathbf{R}^n)$. At every point x at which the limit (2.4) holds, we can extend $A_r f(x)$ to a continuous function of $r \in [0, \infty)$ by defining $A_0 f(x) = f(x)$, and the bound (2.5) then holds when $r = 0$. Define

$$(2.6) \quad \tilde{f}(x) = \lim_{r \rightarrow 0} A_r f(x);$$

then \tilde{f} exists almost everywhere and

$$\tilde{f}(x) = f(x) \quad \text{a.e. } [dx].$$

\tilde{f} agrees with f at each point of continuity.

The operator A_r not only smooths L^p functions, it also smooths L^p convergence:

LEMMA 2.1. *Let $r > 0$, $1 \leq p \leq \infty$; if $f_n \rightarrow f$ in $L^p(\mathbf{R}^n)$ then $A_r f_n \rightarrow A_r f$ uniformly.*

COROLLARY 2.2. *If $\sum f_n$ converges to f in L^p -norm then $\sum A_r f_n$ converges uniformly to $A_r f$.*

We will occasionally use superscripts on A_r or M to indicate the dimension in which averages are being taken. Thus, definition (2.1) for averages of functions $f \in L^1_{loc}(\mathbf{R}^{2n})$ becomes

$$(2.7) \quad A_r^{(2n)} f(x, y) = \frac{1}{|C_r|^2} \int_{C_r} \int_{C_r} f(x+s, y+t) ds dt.$$

The reason for averaging over cubes rather than balls is that the averaging operator on cubes is multiplicative on tensor products of Hilbert spaces:

LEMMA 2.3. *Let $\phi, \psi \in L^2(\mathbf{R}^n)$; then*

$$A_r^{(2n)}(\phi \otimes \psi)(x, y) = A_r^{(n)} \phi(x) A_r^{(n)} \psi(y).$$

As a consequence,

LEMMA 2.4. *The maximal function is submultiplicative and subadditive:*

$$M^{(2n)}(\phi \otimes \psi)(x, y) \leq M^{(n)}\phi(x)M^{(n)}\psi(y)$$

and

$$M(\phi + \psi)(x) \leq M\phi(x) + M\psi(x).$$

REMARKS. Although the theorems in §§3 and 4 about trace class operators on $L^2(X)$ will be presented for the case $X = \mathbf{R}^n$, they remain valid if X is any measurable subset of \mathbf{R}^n . The crucial feature we need to retain in X is the existence almost everywhere of the limits of averages as defined by (2.6). To compute averages of functions defined only on $X \subset \mathbf{R}^n$ we consider them to be equal to zero on $\mathbf{R}^n \setminus X$.

3. Trace class operators. There is an isomorphism between the ideal of Hilbert-Schmidt operators on $L^2(X)$ and the space $L^2(X \times X)$ that identifies each operator, K , with a kernel $K(x, y) \in L^2(X \times X)$. Every positive-definite Hilbert-Schmidt operator, P , has an eigenfunction expansion

$$(3.1) \quad P = \sum_{i=1}^{\infty} \lambda_i \phi_i \otimes \overline{\phi_i}$$

where $(\phi_i)_{i \in \mathbf{N}}$ is an orthonormal sequence in $L^2(X)$, and the eigenvalues, λ_i , are positive and square-summable. The series (3.1) converges in Hilbert-Schmidt norm,

$$(3.2) \quad \|P\|_{HS}^2 = \text{tr}(P^2) = \iint |P(x, y)|^2 d\mu(x) d\mu(y),$$

and thus also in operator norm. The corresponding expansion for the kernel,

$$(3.3) \quad P(x, y) = \sum \lambda_i \phi_i(x) \overline{\phi_i(y)},$$

converges in $L^2(X \times X)$.

If $|K|$ is a trace class operator (i.e., if $\sum \lambda_i < \infty$), to what extent does the formula

$$(3.4) \quad \text{tr } K = \int_X K(x, x) d\mu(x)$$

hold? Clearly, this integral is not even well defined in general since we are only given the kernel as an element of $L^2(X \times X)$. Duflo [1, Theorem V.3.1.1] has proven (3.4) when $K(x, y)$ is a continuous trace class kernel on a σ -compact, locally compact space, X , with a Radon measure, μ . We shall now prove a version of (3.4) that holds for arbitrary trace class kernels when $X \subset \mathbf{R}^n$.

If $K(x, y) \in L^2(\mathbf{R}^{2n})$ then $\tilde{K}(x, y)$ is defined as in (2.6):

$$\tilde{K}(x, y) = \lim_{r \rightarrow 0} A_r^{(2n)} K(x, y).$$

$\tilde{K}(x, y)$ is defined pointwise a.e. $[dx dy]$. Since $\tilde{K}(x, y)$ is uniquely determined by the operator K , the averaging process can be regarded as a way of selecting a "smooth" pointwise representative for the corresponding kernel.

Consider an expansion of the form (3.3) for $P \geq 0$,

$$P(x, y) = \sum \lambda_i \phi_i(x) \overline{\phi_i(y)},$$

where ϕ_i denotes pointwise representatives for the eigenfunctions in $L^2(\mathbf{R}^n)$. We will say that $x \in \mathbf{R}^n$ is a *regular point* of the expansion if, for all $i \in \mathbf{N}$,

$$(3.5) \quad \lim_{r \rightarrow 0} A_r \phi_i(x) = \phi_i(x).$$

Thus, for any given choice of representatives for the functions ϕ_i , almost every point in \mathbf{R}^n is a regular point. By Corollary 2.2 and Lemma 2.3 we have, for $r > 0$,

$$(3.6) \quad A_r^{(2n)} P(x, y) = \sum \lambda_i A_r^{(n)} \phi_i(x) A_r^{(n)} \overline{\phi_i}(y)$$

at every point (x, y) .

THEOREM 3.1. *Let K be a trace class operator on $L^2(\mathbf{R}^n)$; then $M^{(2n)} K(x, x) \in L^1(\mathbf{R}^n)$, $\tilde{K}(x, x)$ exists a.e. $[dx]$, and*

$$(3.7) \quad \text{tr } K = \int \tilde{K}(x, x) dx.$$

PROOF. First, note that every trace class operator, K , can be written as a linear combination of four positive-definite trace class operators:

$$K = P_1 - P_2 + i(P_3 - P_4).$$

Since M is subadditive and since A_r and the trace formula (3.7) are linear, it suffices to prove Theorem 3.1 for positive-definite trace class operators.

We have an expansion for $P \geq 0$ of the form (3.1),

$$(3.8) \quad P = \sum_{i=1}^{\infty} \lambda_i \phi_i \otimes \overline{\phi_i},$$

with $\lambda_i > 0$ and $\sum \lambda_i < \infty$. Let ϕ_i denote representative functions defined pointwise everywhere. Since P is traceable,

$$\int \sum \lambda_i |\phi_i(x)|^2 = \sum \lambda_i < \infty$$

so the series $\sum \lambda_i |\phi_i(x)|^2$ is finite almost everywhere.

Using Lemma 2.4 and the bound (2.3) for the maximal function on $L^2(\mathbf{R}^n)$, we have the following calculation:

$$\begin{aligned} \int M^{(2n)} P(x, x) dx &\leq \int \sum \lambda_i (M^{(n)} \phi_i(x))^2 dx \\ &\leq \sum \lambda_i \|M \phi_i\|_2^2 \\ &\leq C_2^2 \sum \lambda_i \|\phi_i\|_2^2 \\ &< \infty. \end{aligned}$$

This proves the first assertion in the theorem and shows the convergence of the nonnegative series

$$\sum \lambda_i (M \phi_i(x))^2 < \infty \quad \text{a.e. } [dx].$$

Choose a conull set of regular points $Y \subset \mathbf{R}^n$ so that for all $x \in Y$ both of the following series are finite:

$$\sum \lambda_i |\phi_i(x)|^2, \quad \sum \lambda_i (M \phi_i(x))^2.$$

Since Y consists of regular points, the $A_r\phi_i(x)$ are continuous functions of $r \in [0, \infty)$ for all $x \in Y$ and all i . Now,

$$|A_r\phi_i(x)|^2 \leq (M\phi_i(x))^2$$

for $x \in Y$ and $r \in [0, \infty)$ so for each $x \in Y$ the series

$$\sum \lambda_i |A_r\phi_i(x)|^2$$

converges absolutely and uniformly with respect to $r \in [0, \infty)$. By (3.6) we know that for $r > 0$

$$A_r^{(2n)}P(x, x) = \sum \lambda_i |A_r\phi_i(x)|^2$$

so, letting $r \rightarrow 0$, we have

$$\begin{aligned} \tilde{P}(x, x) &= \lim_{r \rightarrow \infty} A_r^{(2n)}P(x, x) \\ &= \sum \lambda_i \lim_{r \rightarrow 0} |A_r\phi_i(x)|^2 \\ &= \sum \lambda_i |\phi_i(x)|^2, \end{aligned}$$

for each $x \in Y$. Since this series converges in $L^1(\mathbf{R}^n)$, we have

$$\begin{aligned} \text{tr } P &= \sum (P\phi_i, \phi_i) \\ &= \sum \int (\lambda_i\phi_i(x))\overline{\phi_i(x)} dx \\ &= \int \sum \lambda_i |\phi_i(x)|^2 dx \\ &= \int \tilde{P}(x, x) dx. \end{aligned}$$

For an illustration of the limitations of Theorem 3.1 we have the following example; applications of the Theorem will occur in the form of Corollaries 4.2 and 4.4.

EXAMPLE 3.2. The Volterra integral operator is the Hilbert-Schmidt operator on $L^2(I)$ given by the kernel

$$K(x, y) = \begin{cases} 1 & \text{if } y < x, \\ 0 & \text{if } y \geq x. \end{cases}$$

Although K is discontinuous, one easily sees that $\tilde{K}(x, x) = \frac{1}{2} = MK(x, x)$ for all $x \in (0, 1)$ so

$$\int_0^1 \tilde{K}(x, x) dx = \frac{1}{2} = \int_0^1 MK(x, x) dx.$$

But it is elementary to show that the Volterra operator's singular values are $\lambda_n = 2(\pi(2n + 1))^{-1}$, so K is *not* traceable. This example shows that integrability of $MK(x, x)$ or $\tilde{K}(x, x)$ are not sufficient conditions for traceability.

Now let us consider the relation between the trace formula (3.7) of Theorem 3.1 and the factorization of a trace class operator into a product of Hilbert-Schmidt operators. K is traceable if and only if K factors as $K = LJ$, where L and J are Hilbert-Schmidt. The corresponding factorization of the kernel in $L^2(\mathbf{R}^{2n})$ is

$$(3.9) \quad K(x, y) = L * J(x, y) \quad \text{a.e. } [dx dy],$$

where the “convolution” of kernels is

$$L * J(x, y) = \int L(x, z)J(z, y) dz.$$

Gaal [2] proves the following result (Theorem VI.7.17):

PROPOSITION 3.3. *If K is traceable, $K = LJ$, then*

$$(3.10) \quad \text{tr } K = \int L * J(x, x) dx.$$

The utility of (3.10) is limited, however, by the fact that it is not possible to explicitly compute the kernels $L(x, y)$ and $J(x, y)$ in terms of a given kernel $K(x, y)$, and Gaal fails to indicate when pointwise equality holds for (3.9) on the diagonal in \mathbf{R}^{2n} . This issue is resolved by our next theorem.

Let $L(x, y)$ and $J(x, y)$ be pointwise representatives for the kernels of the Hilbert-Schmidt operators L and J . $A_{r(n)}(M_{(n)})$ and $A_r^{(n)}(M^{(n)})$ denote the n -dimensional averaging (maximal) operators acting on the first or second variables of a kernel, respectively; e.g.

$$A_{r(n)}L(x, y) = \frac{1}{|C_r|} \int_{C_r} L(x + s, y) ds.$$

LEMMA 3.4. *If $L(x, y) \in L^2(\mathbf{R}^{2n})$, then*

$$M_{(n)}L(x, y), M^{(n)}L(x, y) \in L^2(\mathbf{R}^{2n}).$$

PROOF. For almost every y , $L(\cdot, y) \in L^2(\mathbf{R}^n)$ so the Maximal Theorem implies

$$\iint |M_{(n)}L(x, y)|^2 dx dy \leq \int \left[C_2^2 \int |L(x, y)|^2 dx \right] dy < \infty.$$

THEOREM 3.5. *If $K = LJ$ is a factorization of the trace class operator K into a product of Hilbert-Schmidt operators, then*

$$\tilde{K}(x, x) = L * J(x, x) \quad \text{a.e. } [dx].$$

PROOF. For $r > 0$, note that

$$(3.11) \quad |A_{r(n)}L(x, z)A_r^{(n)}J(z, y)| \leq M_{(n)}L(x, z)M^{(n)}J(z, y)$$

and

$$(3.12) \quad \left[\int M_{(n)}L(x, z)M^{(n)}J(z, y) dz \right]^2 \leq \int |M_{(n)}L(x, z)|^2 dz \int |M^{(n)}J(z, y)|^2 dz$$

where

$$\int |M_{(n)}L(x, z)|^2 dz < \infty \quad \text{a.e. } [dx]$$

and

$$\int |M^{(n)}J(z, y)|^2 dz < \infty \quad \text{a.e. } [dy]$$

by Lemma 3.4. This justifies Fubini’s Theorem:

$$(3.13) \quad \begin{aligned} A_r^{(2n)}K(x, y) &= \frac{1}{|C_r|^2} \int_{C_r} \int_{C_r} \int_{\mathbf{R}^n} L(x + s, z)J(z, y + t) dz ds dt \\ &= \int A_{r(n)}L(x, z)A_r^{(n)}J(z, y) dz, \end{aligned}$$

for almost every x and almost every y .

Now, for almost every z we have

$$\lim_{r \rightarrow 0} A_r^{(n)} L(x, z) = L(x, z) \quad \text{a.e. } [dx]$$

and

$$\lim_{r \rightarrow 0} A_r^{(n)} J(z, y) = J(z, y) \quad \text{a.e. } [dy].$$

Estimates (3.11) and (3.12) allow us to pass to the limit in (3.13):

$$\begin{aligned} \tilde{K}(x, x) &= \lim_{r \rightarrow 0} \int A_r^{(n)} L(x, z) A_r^{(n)} J(z, x) dz \\ &= L * J(x, x) \quad \text{a.e. } [dx]. \end{aligned}$$

4. Positive operators. We now consider the problem of giving necessary and sufficient conditions for a positive Hilbert-Schmidt kernel, $P(x, y)$, to be traceable. Weidmann gives a sufficient condition in Satz 1 of [9]:

PROPOSITION 4.1. *Let $P(x, y) \in L^2(\mathbf{R}^{2n})$ be the kernel of a positive operator, P , on $L^2(\mathbf{R}^n)$. If there exists a function $f \in L^1(\mathbf{R}^n)$ such that for all x and all r , $0 < r < r_0$, we have*

$$A_r^{(2n)} P(x, x) \leq f(x),$$

then P is traceable.

Theorem 3.1 now provides us with the converse:

COROLLARY 4.2. *The positive Hilbert-Schmidt kernel $P(x, y)$ is traceable if and only if*

$$\int M^{(2n)} P(x, x) dx < \infty.$$

If $K(x, y)$ is not traceable, then there is no assurance that $\tilde{K}(x, x)$ exists unless K is continuous on the diagonal. Even then, $\tilde{K}(x, x)$ does not contain the information about the global behavior of K that $MK(x, x)$ does. It is interesting that for positive kernels the integrability of $\tilde{P}(x, x)$ implies traceability, modulo the existence of $\tilde{P}(x, x)$. This is a generalization of Mercer's Theorem (Proposition 1.2):

THEOREM 4.3. *Let P be a positive Hilbert-Schmidt operator on $L^2(\mathbf{R}^n)$ with eigenfunction expansion*

$$(4.1) \quad P = \sum_{i=1}^{\infty} \lambda_i \phi_i \otimes \bar{\phi}_i, \quad \lambda_i \geq 0.$$

If $\tilde{P}(x, x)$ exists for almost every x then

$$\tilde{P}(x, x) \geq 0 \quad \text{a.e. } [dx]$$

and

$$(4.2) \quad \text{tr } P \leq \int \tilde{P}(x, x) dx.$$

Note that we allow the extended real value $\tilde{P}(x, x) = +\infty$. If $\tilde{P}(x, x) < \infty$ a.e. $[dx]$ then expansion (4.1) converges pointwise:

$$(4.3) \quad \tilde{P}(x, y) = \sum \lambda_i \phi_i(x) \overline{\phi_i(y)} \quad \text{a.e. } [dx dy].$$

PROOF. As before, we let the functions ϕ_i be pointwise representatives for the eigenfunctions. Then

$$S_m(x, y) = \sum_{i=1}^m \lambda_i \phi_i(x) \overline{\phi_i(y)}$$

is defined everywhere, and

$$H_m(x, y) = \tilde{P}(x, y) - S_m(x, y)$$

is defined wherever $\tilde{P}(x, y)$ exists. Since

$$H_m = \sum_{i=m+1}^{\infty} \lambda_i \phi_i \otimes \overline{\phi_i},$$

H_m is a positive operator.

Let $Y \subset \mathbb{R}^n$ be the set of all regular points, x , for the expansion (4.1) at which $\tilde{P}(x, x)$ exists, and note that Y contains almost all of \mathbb{R}^n . Then for all $x \in Y$, Lemma 2.3 and (3.5) justify the calculation

$$\begin{aligned} H_m(x, x) &= \tilde{P}(x, x) - \sum_{i=1}^m \lambda_i |\phi_i(x)|^2 \\ &= \lim_{r \rightarrow 0} A_r^{(2n)} P(x, x) - \sum_{i=1}^m \lambda_i \lim_{r \rightarrow 0} A_r^{(2n)} (\phi_i \otimes \overline{\phi_i})(x, x) \\ &= \lim_{r \rightarrow 0} A_r^{(2n)} (P - S_m)(x, x) \\ &= \lim_{r \rightarrow 0} \frac{1}{|C_r|^2} \int_{C_r} \int_{C_r} H_m(x + s, x + t) ds dt \\ &= \lim_{r \rightarrow 0} \frac{1}{|C_r|^2} (H_m \mathbf{1}_{C_r(x)}, \mathbf{1}_{C_r(x)}) \\ &\geq 0, \end{aligned}$$

where $\mathbf{1}_{C_r(x)}$ is the characteristic function of $C_r(x)$. If $x \in Y$ then, for all m ,

$$\tilde{P}(x, x) - S_m(x, x) = H_m(x, x) \geq 0,$$

implying that

$$(4.4) \quad \tilde{P}(x, x) \geq \sum \lambda_i |\phi_i(x)|^2 \geq 0.$$

Now integrate (4.4):

$$(4.5) \quad \text{tr } P = \sum \lambda_i \int |\phi_i(x)|^2 dx \leq \int \tilde{P}(x, x) dx.$$

If $\tilde{P}(x, x) < \infty$ for all $x \in Y$ then

$$\begin{aligned} \left[\sum |\lambda_i \phi_i(x) \overline{\phi_i(y)}| \right]^2 &\leq \sum \lambda_i |\phi_i(x)|^2 \sum \lambda_i |\phi_i(y)|^2 \\ &< \infty \quad \text{if } (x, y) \in Y \times Y, \end{aligned}$$

which says that $S_m(x, y)$ converges absolutely on $Y \times Y$. But $S_m \rightarrow \tilde{P}$ in $L^2(\mathbb{R}^{2n})$ so the pointwise limit agrees with \tilde{P} almost everywhere.

We can combine estimate (4.5) and Theorem 3.1 to obtain the following useful result:

COROLLARY 4.4. If $P \geq 0$ and $\tilde{P}(x, x)$ exists a.e. $[dx]$ then

$$(4.6) \quad \operatorname{tr} P = \int \tilde{P}(x, x) dx.$$

Thus, P is traceable if and only if the integral (4.6) is finite.

REMARKS. Unlike the classical version of Mercer's Theorem on the compact domain $[0, 1]$, even if $P(x, y)$ is continuous on \mathbf{R}^n the best we can hope for is almost everywhere convergence of (4.3) since P need not have continuous eigenfunctions on \mathbf{R}^n . If P is continuous, though, one can prove that (4.3) converges to P almost everywhere on the diagonal:

$$P(x, x) = \sum \lambda_i |\phi_i(x)|^2 \quad \text{a.e. } [dx].$$

For continuous kernels, Corollary 4.4 provides a new proof of a result in [1, Theorem V.3.3.1 and 5, Theorem 2.12]. It would be interesting to know whether $\tilde{P}(x, x)$ exists in $[0, \infty]$ a.e. $[dx]$ for every positive Hilbert-Schmidt operator, P .

REFERENCES

1. M. Duflo, *Généralités sur les représentations induites*, Représentations des Groupes de Lie Résolubles, Monographies de la Soc. Math. de France, vol. 4, Dunod, Paris, 1972, pp. 93–119.
2. S. Gaal, *Linear analysis and representation theory*, Grundlehren Math. Wiss., 198, Springer-Verlag, Berlin, Heidelberg, New York, 1973.
3. F. Riesz and B. Sz.-Nagy, *Functional analysis*, Ungar, New York, 1955.
4. C. Rockland, *Hypoellipticity on the Heisenberg group—representation-theoretic criteria*, Trans. Amer. Math. Soc. **240** (1978), 1–52.
5. B. Simon, *Trace ideals and their applications*, London Math. Soc. Lecture Notes, vol. 35, Cambridge Univ. Press, Cambridge, 1979.
6. E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, N.J., 1970.
7. M. Sugiura, *Unitary representations and harmonic analysis*, Kodansha, Tokyo, 1975.
8. F. G. Tricomi, *Integral equations*, Dover, New York, 1985.
9. J. Weidmann, *Integraloperatoren der spurklasse*, Math. Ann. **163** (1966), 340–345.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO, BOULDER, COLORADO 80309-0426