Differential and Integral Equations Volume 19, Number 2 (2006), 189–210

POSITIVE DEFINITE MATRICES AND INTEGRAL EQUATIONS ON UNBOUNDED DOMAINS

Jorge Buescu and A. C. Paixão
Dep. Matemática, Inst. Sup. Técnico, Portugal
Dep. Mecânica, ISEL, Portugal

(Submitted by: Viorel Barbu)

Abstract. The kernel of a continuous positive integral operator on an interval I is a Moore matrix on I. We show that, under minimal differentiability assumptions, this implies that the kernel satisfies a 2-parameter family of differential inequalities. These inequalities ensure that, for unbounded I, the corresponding integral operator is exceptionally well behaved: it is compact and thus the eigenfunctions for its discrete spectrum have the differentiability of the kernel and satisfy sharp Sobolev bounds, the symmetric mixed partial derivatives are again kernels of positive operators and the differentiated eigenfunction series converge uniformly and absolutely. Converse results are derived.

1. Positive definite kernels as Moore matrices

Given an interval $I \subseteq \mathbb{R}$, a linear operator $K : L^2(I) \to L^2(I)$ is integral if there exists a measurable function $k \in L^2(I^2)$ such that for all $\phi \in L^2(I)$

$$K(\phi) = \int_I k(x,y)\,\phi(y)\,dy$$

almost everywhere. The function $k(x,y)$ is called the kernel of K. If in addition K satisfies the condition

$$\int_I \int_I k(x,y)\,\overline{\phi(x)}\,\phi(y)\,dx\,dy \geq 0 \tag{1.1}$$

for all $\phi \in L^2(I)$, then it is a positive operator and the corresponding kernel $k(x,y)$ is a *positive definite kernel*. A positive definite kernel satisfies $k(x,y) = \overline{k(y,x)}$ for almost all $(x,y) \in I^2$, so the associated operator K is self-adjoint. Moreover, all eigenvalues of K are real and nonnegative.

Accepted for publication: October 2005.
AMS Subject Classifications: 45P05; 15A45, 46E22.
The first author acknowledges partial support by CAMGSD through FCT/POCTI/FEDER.

Definition 1.1. *A positive definite kernel $k(x,y)$ in a closed (possibly unbounded) interval $I \subseteq \mathbb{R}$ is called a* Mercer-like kernel on I *if*

(1) *it is continuous in I,*
(2) $k(x,x) \in L^1(I)$,
(3) $k(x,x)$ *is uniformly continuous in I.*

Remark 1.2. If I is compact the first condition trivially implies the other two. This is the classical setting for Mercer's theorem (see e.g. [11]). Definition 1.1 is especially meaningful for noncompact domains. It has been shown [3] that it specifies necessary and sufficient conditions for the noncompact analog of Mercer's theorem to hold; hence the name *Mercer-like* for this class of kernels.

We briefly recall some well-known facts from operator theory. Let $I \subseteq \mathbb{R}$ be an interval. If $k(x,y) \in L^2(I^2)$ then K is Hilbert-Schmidt, thus compact, so eigenvalues accumulate only at 0. If K is self-adjoint, its spectral series may be written as

$$k(x,y) = \sum_{i \geq 1} \lambda_i \, \phi_i(x) \, \overline{\phi_i(y)}, \tag{1.2}$$

where the $\{\phi_i\}_{i \geq 1}$ are an $L^2(I)$-orthonormal set of eigenfunctions spanning the range of K, the $\{\lambda_i\}_{i \geq 1}$ are the nonzero eigenvalues of K and the series (1.2) converges in $L^2(I^2)$. If K is a positive operator, $\lambda_i \geq 0$ and we may assume $\{\lambda_i\}_{i \geq 1}$ is nonincreasing. If in addition k is a Mercer-like kernel on I, then eigenfunctions ϕ_i associated to nonzero eigenvalues are uniformly continuous, convergence of the series (1.2) is absolute and uniform in I^2 and the operator K is trace class and satisfies the trace formula

$$\int_I k(x,x) \, dx = \sum_{i \geq 1} \lambda_i. \tag{1.3}$$

In the case where I is compact, this is the content of the classical theorem of Mercer (see e.g. [11]); the noncompact case is treated in [3].

Positive definite kernels are closely related to a class of functions known as *positive definite matrices in the sense of Moore* (see e.g. Moore [10], Krein [8], Aronszajn [2]). Given an abstract set E, a positive definite matrix in the sense of Moore is a function $k : E \times E \to \mathbb{C}$ such that

$$\sum_{i,j=1}^{n} k(x_i, x_j) \, \overline{\xi_i} \, \xi_j \geq 0 \tag{1.4}$$

for all $n \in \mathbb{N}$, $(x_1, \cdots, x_n) \in E^n$ and $(\xi_1, \cdots, \xi_n) \in \mathbb{C}^n$. Informally, k is a function such that all finite square matrices M of elements $m_{ij} = k(x_i, x_j)$, $i, j = 1, \ldots, n$, are positive semidefinite. From (1.4) it is easily seen that a positive definite matrix in the sense of Moore satisfies $k(x,y) = \overline{k(y,x)}$ for all $x, y \in E$, $k(x,x) \geq 0$ for all $x \in E$ and $|k(x,y)|^2 \leq k(x,x)k(y,y)$ for all $x, y \in E$ (in this context diagonal dominance holds pointwise without continuity assumptions; in fact, the set E is not assumed to be endowed with a topology).

It is straightforward, as first done by Mercer (see e.g. Stewart [14]), to show that if $E = I \subset \mathbb{R}$ is a compact interval and $k : I^2 \to \mathbb{C}$ is continuous then conditions (1.1) and (1.4) are equivalent. The following result shows how this property extends to arbitrary intervals.

Lemma 1.3. *Let $I \subseteq \mathbb{R}$ be an interval and suppose $k \in L^2(I^2) \cap C(I^2)$. Then k is a positive definite matrix in the sense of Moore if and only if k is an $L^2(I^2)$ positive definite kernel.*

Proof. Suppose k is an $L^2(I^2) \cap C(I^2)$ positive definite matrix in the sense of Moore and let $I_c \subset I$ be compact. Continuity of k and Mercer's characterization for compact domains imply that $\iint_{I_c^2} k(x,y)\,\overline{\phi(x)}\,\phi(y)\,dx\,dy \geq 0$ for every $\phi \in L^2(I)$. Arbitrariness of I_c then implies that

$$\iint_{I^2} k(x,y)\,\overline{\phi(x)}\,\phi(y)\,dx\,dy \geq 0$$

for every $\phi \in L^2(I)$, showing that k is an $L^2(I^2)$ positive definite kernel.

For the converse implication, consider any sum $S = \sum_{i,j=1}^{n} k(x_i, x_j)\,\overline{\xi_i}\,\xi_j$ with $n \in \mathbb{N}$, $(x_1, \ldots, x_n) \in I^n$ and $(\xi_1, \ldots, \xi_n) \in \mathbb{C}^n$. If I is compact all such sums S are greater than or equal to 0 as a consequence of Mercer's characterization of continuous positive definite kernels. For noncompact I, the closed convex hull H of $\{x_1, \ldots, x_n\}$ is a compact subset of I, and again by Mercer's result $S \geq 0$. Arbitrariness of n, $(x_1, \ldots, x_n)$ and $(\xi_1, \ldots, \xi_n)$ implies that k is a positive definite matrix in the sense of Moore. $\square$

Although elementary, this result shifts the focus from analytic to algebraic properties of positive definite kernels. This will be crucial for the rest of the paper; for instance, it will be shown in the following sections how integrability conditions on unbounded domains may be derived from the algebraic characterization (1.4). As a simple application let us note that consideration of the cases $n = 1$ and $n = 2$ in (1.4) yields respectively the well-known facts that continuous positive definite kernels on an interval I satisfy the inequalities for all $x \in I$ $k(x,x) \geq 0$ and for all x, $y \in I$ $|k(x,y)|^2 \leq k(x,x)k(y,y)$.

We usually refer to the latter as the *diagonal dominance inequality*; note that it implies the previous one.

The relationship between positive definite kernels and positive definite matrices in the sense of Moore is, however, much deeper. From the fundamental theorem of the theory of reproducing kernels (Moore [10] – Aronszajn [2]), $k : E \times E \to \mathbb{C}$ is a positive definite matrix in the sense of Moore if and only if there exists a (uniquely determined) Hilbert space H_k composed of functions on E such that

$$\forall y \in E,\ k(x,y) \in H_k \text{ as a function of } x, \tag{1.5}$$

and

$$\forall x \in E \text{ and any } f \in H_k,\ f(x) = \langle f(y), k(y,x)\rangle_{H_k}. \tag{1.6}$$

Properties (1.5) and (1.6) are jointly called the *reproducing property of* k *in* H_k. The function k itself is called a *reproducing kernel* and the associated Hilbert space H_k a *reproducing kernel Hilbert space*.

The theorem of Moore-Aronszajn provides an equivalent characterization of positive definite matrices in the sense of Moore as reproducing kernels. These are the object of a deep and elegant general theory with an extensive literature and a broad range of applications; see e.g. [10], [2], [8], or more recent reviews such as the ones by Saitoh [12, 13] or Cucker and Smale [5] and references therein for mainstream applications. In this paper we shall not make use of the Moore-Aronszajn theorem or the general theory of reproducing kernels but work directly in E using the linear algebraic structure provided by (1.4).

2. A family of differential inequalities for Moore matrices

Throughout this section $I \subset \mathbb{R}$ will denote a nontrivial but otherwise arbitrary interval. If $x \in I$ is a boundary point of I, a limit at x will mean the one-sided limit as $y \to x$ with $y \in I$.

Definition 2.1. *Let* $I \subset \mathbb{R}$ *be an interval. A function* $k : I^2 \to \mathbb{C}$ *is said to* be of class $\mathcal{S}_n(I)$ *if, for every* $m_1 = 0, 1, \ldots n$ *and* $m_2 = 0, 1, \ldots n$*, the partial derivatives* $\dfrac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}}k(x,y)$ *are continuous in* I^2.

Remark 2.2. Clearly from the definition $C^{2n}(I^2) \subset \mathcal{S}_n(I) \subset C^n(I^2)$. It is also clear that a function of class $\mathcal{S}_n(I)$ will not in general be in $C^{n+1}(I^2)$. Note however that in the class $\mathcal{S}_n(I)$ equality of all intervening mixed partial derivatives holds. In fact, $\mathcal{S}_n(I)$ is precisely the weakest differentiability class on which Schwarz's theorem for equality of mixed partial derivatives holds.

A function $k : I^2 \to \mathbb{C}$ is conjugate symmetric if $k(x,y) = \overline{k(y,x)}$ for all $x,\, y \in I$. $L^2(I)$-positive definite kernels are almost everywhere conjugate symmetric.

Lemma 2.3. *Let $I \subset \mathbb{R}$ be an interval and $k(x,y)$ be a conjugate symmetric function such that $\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}}k(x,y)$ exists for all $x,\, y \in I$. Then for all $x,\, y \in I$ the partial derivative $\frac{\partial^{m_1+m_2}}{\partial x^{m_2}\partial y^{m_1}}k(y,x)$ exists and satisfies*

$$\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}}k(x,y) = \overline{\frac{\partial^{m_1+m_2}}{\partial x^{m_2}\partial y^{m_1}}k(y,x)}. \tag{2.1}$$

Proof. Fix $y \in I$ and define $f_y : I \to \mathbb{C}$ by $f_y(t) = k(t,y)$. According to the hypothesis, the m_1^{th} derivative of f_y exists and $f_y(t) = \overline{k(y,t)}$. We then have

$$\frac{\partial^{m_1}}{\partial x^{m_1}}k(x,y) = f_y^{(m_1)}(x) = \overline{\frac{\partial^{m_1}}{\partial y^{m_1}}k(y,x)}.$$

For fixed $x \in I$ we now define $g_x : I \to \mathbb{C}$ by $g_x(t) = \frac{\partial^{m_1}}{\partial y^{m_1}}k(x,t)$. According to the hypothesis, the m_2^{th} derivative of g_x exists and $g_x(t) = \overline{\frac{\partial^{m_1}}{\partial x^{m_1}}k(t,x)}$. We then have

$$\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}}k(x,y) = g_x^{(m_2)}(y) = \overline{\frac{\partial^{m_1+m_2}}{\partial x^{m_2}\partial y^{m_1}}k(y,x)},$$

completing the proof. □

Remark 2.4. Since this paper deals with positive definite kernels (which are automatically conjugate symmetric), from now on the kernels $k(x,y)$ under consideration will always be assumed conjugate symmetric. Notice for instance that, in view of Lemma 2.3, assuming $\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}}k(x,y)$ to be continuous on I^2 for $0 \le m_1 \le m_2 \le n$ would be sufficient to ensure that k is of class $\mathcal{S}_n(I)$.

Given an interval $I \subset \mathbb{R}$, let $f : I^2 \to \mathbb{C}$. We define the operators $\Delta_{x,h}$ and $\Delta_{y,h}$ by

$$\Delta_{x,h}f(x,y) = f(x+h,y) \text{ and } \Delta_{y,h}f(x,y) = f(x,y+h) \tag{2.2}$$

for any $h \in \mathbb{R}$ such that $(x+h,y)$ and $(x,y+h)$ lie in I^2. Note that $\Delta_{x,h}$ and $\Delta_{y,h}$ commute.

Proposition 2.5. *Let $I \subset \mathbb{R}$ be an interval and suppose $k : I^2 \to \mathbb{C}$ is of class $\mathcal{S}_n(I)$. Then, for every $m_1,\, m_2 = 0,\, \ldots,\, n$,*

$$\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}}k(x,y) = \lim_{h\to 0}\frac{1}{h^{m_1+m_2}}\left(\Delta_{y,h}-1\right)^{m_2}\left(\Delta_{x,h}-1\right)^{m_1}\, k(x,y)$$

for every $(x, y) \in \text{int}(I^2)$.

Proof. Suppose first that k is real valued. For $R > 0$ and $y \in I$ let $I_{y,R} = (y - R, y + R)$. Choose R such that $I_{y,R} \subset I$, suppose $x \in I$ and let $m_2 > 0$. For $|h| < R$ we write

$$(\Delta_{y,h} - 1)^{m_2} (\Delta_{x,h} - 1)^{m_1}\ k(x, y) = (\Delta_{y,h} - 1)\, g(x, y)$$

where $g(x, y) = (\Delta_{y,h} - 1)^{m_2 - 1} (\Delta_{x,h} - 1)^{m_1}\ k(x, y)$. Since k is in class $\mathcal{S}_n(I)$, we have that

$$(\Delta_{y,h} - 1)\, g(x, y) = h\, \frac{\partial g}{\partial y}(x, y_1)$$

for some y_1 in the interior of the interval of extremes y and $y + h$. Hence we may write

$$(\Delta_{y,h} - 1)^{m_2} (\Delta_{x,h} - 1)^{m_1}\, k(x, y) = h\, (\Delta_{y,h} - 1)^{m_2 - 1} (\Delta_{x,h} - 1)^{m_1}\, \frac{\partial k}{\partial y}(x, y_1).$$

Now choose R_1 and R_2 such that $I_{x,R_1} \times I_{y,R_2} \subset I^2$ and let $|h|$ be sufficiently small that $m_1|h| < R_1$, $m_2|h| < R_2$. Using the same argument as the above m_1 times with respect to x and m_2 times with respect to y we construct sequences $x_0, x_1, \dots, x_{m_1}$, $y_0, y_1, \dots, y_{m_2}$ such that $x_0 = x$, x_i lies in the interval of extremes x_{i-1} and $x_{i-1} + h$ for $i = 1, \dots, m_1$, y_j lies in the interval of extremes y_{j-1} and $y_{j-1} + h$ for $j = 1, \dots, m_2$ and with

$$(\Delta_{y,h} - 1)^{m_2} (\Delta_{x,h} - 1)^{m_1}\ k(x, y) = h^{m_2 + m_1} \frac{\partial^{m_2 + m_1}}{\partial y^{m_2} \partial x^{m_1}} (x_{m_1}, y_{m_2}).$$

Notice that our construction ensures that $(x_{m_1}, y_{m_2}) \to (x, y)$ as $h \to 0$. Therefore, continuity of $\frac{\partial^{m_2+m_1}}{\partial y^{m_2} \partial x^{m_1}}$ implies that

$$\lim_{h \to 0} \frac{1}{h^{m_1 + m_2}} (\Delta_{y,h} - 1)^{m_2} (\Delta_{x,h} - 1)^{m_1}\ k(x, y) = \frac{\partial^{m_1 + m_2}}{\partial y^{m_2} \partial x^{m_1}} k(x, y).$$

This concludes the proof in the case where k is real valued. If k is complex valued, application of the above conclusions to its real and imaginary parts separately yields the proposition. □

Theorem 2.6. *Let* $I \subset \mathbb{R}$ *be an interval and* $k(x, y)$ *be a Moore matrix on* I *of class* $\mathcal{S}_n(I)$. *Then for all* x, $y \in I$ *and all* $0 \le m \le n$ *we have*

$$\left| \frac{\partial^m}{\partial x^m} k(x, y) \right|^2 \le \frac{\partial^{2m}}{\partial y^m \partial x^m} k(x, x)\, k(y, y). \tag{2.3}$$

Proof. For $n = 0$ the statement yields

$$|k(x,y)|^2 \le k(x,x)\,k(y,y), \tag{2.4}$$

which is just the classical diagonal dominance inequality for positive matrices in the sense of Moore cited in Section 1.

For $n > 0$ we begin by fixing $y \in I$ and supposing that $k(y,y) = 0$. Using diagonal dominance, (2.4), we have that $|k(x,y)|^2 = 0$ for every $x \in I$, implying that $\frac{\partial^m}{\partial x^m}k(x,y) = 0$ for every $0 \le m \le n$. This proves the statement in the case $k(y,y) = 0$. From now on we assume without loss of generality that $k(y,y) \ne 0$ (and thus $k(y,y) > 0$ by diagonal dominance).

Since $k(x,y)$ is of class $\mathcal{S}_n(I)$, Proposition 2.5 implies that

$$\frac{\partial^{2m}}{\partial y^m \partial x^m}k(x,x) = \lim_{h\to 0}\frac{1}{h^{2m}}(\Delta_{y,h}-1)^m(\Delta_{x,h}-1)^m\,k(x,x) \tag{2.5}$$

and

$$\frac{\partial^m}{\partial x^m}k(x,y) = \lim_{h\to 0}\frac{1}{h^m}(\Delta_{x,h}-1)^m\,k(x,y) \tag{2.6}$$

for every $0 \le m \le n$.

Using (2.5) and (2.6) we may write

$$\begin{aligned}\Big(\frac{\partial^{2m}}{\partial y^m \partial x^m}k(x,x)\Big)\,k(y,y) - \Big|\frac{\partial^m k}{\partial x^m}(x,y)\Big|^2 \\ = \lim_{h\to 0}\frac{1}{h^{2m}}\Big(\Big[(\Delta_{y,h}-1)^m(\Delta_{x,h}-1)^m\,k(x,x)\Big]\,k(y,y) \\ - \Big[(\Delta_{x,h}-1)^m\,k(x,y)\Big]\,\overline{\Big[(\Delta_{x,h}-1)^m\,k(x,y)\Big]}\Big).\end{aligned} \tag{2.7}$$

We now define the functions

$$\Psi(x,y,h) = \Big[(\Delta_{y,h}-1)^m(\Delta_{x,h}-1)^m\,k(x,x)\Big]\,k(y,y), \tag{2.8}$$

$$\Phi(x,y,h) = \Big[(\Delta_{x,h}-1)^m\,k(x,y)\Big]\,\overline{\Big[(\Delta_{x,h}-1)^m\,k(x,y)\Big]}. \tag{2.9}$$

The rest of the proof will be devoted to showing that $\Psi(x,y,h) - \Phi(x,y,h) \ge 0$ for all x, y, h where both quantities are defined. This will imply that, for all x, $y \in \mathrm{int}I$ and sufficiently small h the finite increment on the right-hand side of (2.7) is greater than or equal to zero, which in turn implies, taking the limit as $h \to 0$, that (2.7) is greater than or equal to zero. Inequality (2.3) will thus be established for x, $y \in \mathrm{int}(I)$ and, by continuity, for all $x, y \in I^2$.

Expanding binomially the finite differences (2.8) and (2.9), we obtain

$$\begin{aligned}
\Psi(x,y,h) \\
&= \Big[\Big(\sum_{i=0}^{m}(-1)^{m-i}\binom{m}{i}\Delta_{y,h}^{i}\Big)\Big(\sum_{j=0}^{m}(-1)^{m-j}\binom{m}{j}\Delta_{x,h}^{j}\Big)k(x,x)\Big]k(y,y) \\
&= \Big(\sum_{i,j=0}^{m}(-1)^{2m-i-j}\binom{m}{i}\binom{m}{j}\Delta_{y,h}^{i}\Delta_{x,h}^{j}\,k(x,x)\Big)\,k(y,y) \\
&= \Big(\sum_{i,j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}k(x+ih,x+jh)\Big)k(y,y) \qquad (2.10)
\end{aligned}$$

and

$$\begin{aligned}
\Phi(x,y,h) \\
&= \Big(\sum_{i=0}^{m}(-1)^{m-i}\binom{m}{i}\Delta_{x,h}^{i}k(x,y)\Big)\overline{\Big(\sum_{j=0}^{m}(-1)^{m-j}\binom{m}{j}\Delta_{x,h}^{j}k(x,y)\Big)} \\
&= \sum_{i,j=0}^{m}(-1)^{2m-i-j}\binom{m}{i}\binom{m}{j}\Delta_{x,h}^{i}k(x,y)\,\overline{\Delta_{x,h}^{j}k(x,y)} \\
&= \sum_{i,j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}k(x+ih,y)\,\overline{k(x+jh,y)}. \qquad (2.11)
\end{aligned}$$

Hence,

$$\Psi(x,y,h)-\Phi(x,y,h) \qquad (2.12)$$

$$=\sum_{i,j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}\Big[k(x+ih,x+jh)k(y,y)-k(x+ih,y)\overline{k(x+jh,y)}\Big].$$

Using the conjugate symmetry of k, we obtain

$$\begin{aligned}
&\Psi(x,y,h)-\Phi(x,y,h) \\
&=\sum_{i,j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}[k(x+ih,x+jh)k(y,y)-k(x+ih,y)k(y,x+jh)] \\
&= \sum_{i,j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}A_{ij}, \qquad (2.13)
\end{aligned}$$

where the A_{ij} are defined by

$$A_{ij} = k(x+ih, x+jh)k(y,y) - k(x+ih, y)k(y, x+jh) \tag{2.14}$$

for i, $j = 0, \dots, m$.

Consider the set of real numbers $X = \{y, x, x+h, \dots, x+mh\} = \{y, x+ih : i = 0, \dots, m\}$. Define the Hermitian $(m+2) \times (m+2)$ matrix T_X associated to X by

$$T_X = \begin{bmatrix} k(y,y) & k(y,x) & k(y,x+h) & \cdots & k(y,x+mh) \\ k(x,y) & k(x,x) & k(x,x+h) & & k(x,x+mh) \\ k(x+h,y) & k(x+h,x) & k(x+h,x+h) & \cdots & k(x+h,x+mh) \\ \vdots & & & & \\ k(x+ih,y) & & \cdots & & \vdots \\ \vdots & & & & \\ k(x+mh,y) & & \cdots & & k(x+mh,x+mh) \end{bmatrix}.$$

The fact that k is by hypothesis a positive definite matrix in the sense of Moore implies that the matrix T_X is positive semidefinite. Recalling that by construction $k(y,y) > 0$, we may use this entry of T_X as a pivot and perform vertical condensation. Multiplying each row of T_X by $k(y,y)$ and subtracting from the resulting row corresponding to the index i the product of the first row by $k(x+ih, y)$, we obtain the matrix T'_X defined by

$$T'_X = \begin{bmatrix} k(y,y) & k(y,x) & k(y,x+h) & \cdots & k(y,x+jh) & \cdots & k(y,x+mh) \\ 0 & A_{00} & A_{01} & & A_{0j} & & A_{0m} \\ 0 & A_{10} & A_{11} & & \cdots & & A_{1m} \\ \vdots & & & & & & \\ 0 & A_{i0} & \cdots & & A_{ij} & & \vdots \\ \vdots & & & & & & \\ 0 & A_{m0} & \cdots & & & & A_{mm} \end{bmatrix}$$

where the A_{ij} are those in (2.14).

We now show that the $(m+1) \times (m+1)$ matrix $A = \left[A_{ij}\right]_{i,j=0}^{n}$ is also positive semidefinite. To see this, the simplest way is to observe that A_{ij} is $k(y,y)$ times the Schur complement of $k(y,y)$ in T_X. Since $k(y,y) > 0$, it follows from [9], Theorem 7.7.6 and [6], Theorem 2.7, that A_{ij} has the same definiteness properties as T_X. Since T_X is positive semidefinite, so is A_{ij}.

Finally, we show that $\Psi(x,y,h) - \Phi(x,y,h) \geq 0$. Define $\xi_i = (-1)^i \binom{m}{i}$ for $0 \leq i \leq m$. Since A is positive definite we have by definition

$$\sum_{i,j}^{m} A_{ij}\xi_i\,\xi_j = \sum_{i,j}^{m} (-1)^{i+j}\binom{m}{i}\binom{m}{j} A_{ij} \geq 0.$$

Hence, by (2.12) we have $\Psi(x,y,h) - \Phi(x,y,h) \geq 0$. This completes the proof. □

Corollary 2.7. *Let $I \subset \mathbb{R}$ be an interval and $k(x,y)$ be a Moore matrix on I of class $\mathcal{S}_n(I)$. Then for all $x \in I$ and all $0 \leq m \leq n$ we have*

$$\frac{\partial^{2m}}{\partial y^m \partial x^m} k(x,x) \geq 0.$$

Proof. Immediate from Theorem 2.6. □

Corollary 2.8. *Let $I \subset \mathbb{R}$ be an interval and $k(x,y)$ be a Moore matrix on I of class $\mathcal{S}_n(I)$. Then for all $x,\, y \in I$ and all $0 \leq m \leq n$ we have*

$$\left|\frac{\partial^m}{\partial y^m}(x,y)\right|^2 \leq \frac{\partial^{2m} k}{\partial y^m \partial x^m} k(y,y)\, k(x,x).$$

Proof. Immediate from Lemma 2.3 and Theorem 2.6. □

3. Positive definite $\mathcal{S}_n(\mathbb{R})$ and $\mathcal{A}_n(\mathbb{R})$ kernels

Definition 3.1. *A function $k : \mathbb{R}^2 \to \mathbb{C}$ is said to belong to class $\mathcal{A}_0(\mathbb{R})$ if:*

(1) *$k(x,y)$ is continuous in $\mathbb{R}^2$;*
(2) *$k(x,x) \in L^1(\mathbb{R})$;*
(3) *$k(x,x) \to 0$ as $|x| \to +\infty$.*

Remark 3.2. If k is an $L^2(\mathbb{R}^2)$ positive definite kernel in class $\mathcal{A}_0(\mathbb{R})$ then it is uniformly continuous in $\mathbb{R}^2$. In fact it is easily seen that for positive definite kernels in $\mathbb{R}$ definitions 3.1 and 1.1 are equivalent: a positive definite kernel is in class $\mathcal{A}_0(\mathbb{R})$ if and only if it is Mercer-like on $\mathbb{R}$. Equivalently, the class of Mercer-like kernels on $\mathbb{R}$ is precisely the intersection of the class $\mathcal{A}_0(\mathbb{R})$ with that of positive definite kernels.

In the sequel we assume that the positive definite kernel $k(x,y)$ is written in terms of its Schmidt series (1.2); that is, $k(x,y) = \sum_{i\geq 1} \lambda_i\, \phi_i(x)\, \overline{\phi_i(y)}$ where each ϕ_i is a normalized eigenfunction of the integral operator with kernel k associated with the eigenvalue $\lambda_i \neq 0$.

According to Remark 3.2 and the observations of Section 1, positive definite kernels in the class $\mathcal{A}_0(\mathbb{R})$ coincide with Mercer-like kernels on $\mathbb{R}$; eigenfunctions ϕ_i associated with nonzero eigenvalues are uniformly continuous and convergence of the series (1.2) is absolute and uniform on $\mathbb{R}^2$.

We will frequently denote $k_m(x,y) = \frac{\partial^{2m}}{\partial y^m \partial x^m} k(x,y)$.

Lemma 3.3. *Let $k(x,y)$ be a Mercer-like kernel in the class $\mathcal{S}_n(\mathbb{R})$, $n \geq 0$. Then the following statements hold.*

(1) *If ϕ_i is an eigenfunction associated to $\lambda_i \neq 0$, then ϕ_i is in $C^n(\mathbb{R})$;*
(2) *each k_m is a Moore matrix in class $\mathcal{S}_{n-m}(\mathbb{R})$ and*

$$k_m(x,y) = \frac{\partial^{2m} k}{\partial y^m \partial x^m}(x,y) = \sum_{i \geq 1} \lambda_i \phi_i^{(m)}(x) \overline{\phi_i^{(m)}(y)} \tag{3.1}$$

absolutely, pointwise in $\mathbb{R}^2$ and uniformly on compact sets of $\mathbb{R}^2$ for each $m = 0, \ldots, n$.

Proof. For $m = 0$, all the statements follow from the hypothesis that k is a Mercer-like kernel.

Let ϕ_i be an eigenfunction associated with an eigenvalue $\lambda_i \neq 0$, that is

$$\phi_i(x) = \frac{1}{\lambda_i} \int_{-\infty}^{+\infty} k(x,y)\, \phi_i(y) dy. \tag{3.2}$$

Suppose k is in $\mathcal{S}_n(\mathbb{R})$. Regarding (3.2) as a parametric integral, we now show that $\phi_i^{(m)}$ exists and is continuous for $m = 1, \ldots, n$.

By hypothesis, for every $1 \leq m \leq n$ the functions $\frac{\partial^m k(x,y)}{\partial x^m}$ and $\frac{\partial^{2m} k(x,y)}{\partial x^m \partial y^m}$ exist and are continuous. We have

$$\begin{aligned} \Big| \frac{\partial^m}{\partial x^m} k(x,y)\, \phi_i(y) \Big| &= \Big| \frac{\partial^m}{\partial x^m} k(x,y) \Big|\, |\phi_i(y)| \\ &\leq \Big(\frac{\partial^{2m}}{\partial y^m \partial x^m} k(x,x) \Big)^{1/2} k(y,y)^{1/2}\, |\phi_i(y)|, \end{aligned} \tag{3.3}$$

where we have used Theorem 2.6.

The fact that $g(y) = k(y,y)^{1/2}\, |\phi_i(y)|$ is in $L^1(\mathbb{R})$ follows from the Cauchy-Schwartz inequality since

$$\int_{-\infty}^{+\infty} k(y,y)^{1/2}\, |\phi_i(y)|\, dy \leq \Big(\int_{-\infty}^{+\infty} k(y,y)\, dy \Big)^{1/2} \|\phi_i(y)\|_{L^2(\mathbb{R})}. \tag{3.4}$$

Thus, differentiation under the integral sign holds (see e.g. [1], Theorem 10.39), the integral (3.2) is n times differentiable, and so are the eigenfunctions ϕ_i. From (3.3), (3.4) and continuity of $k_m(x,x)$ we conclude (see e.g. [1], Theorem 10.38) that the integral corresponding to the m^{th} derivative under the integral sign is continuous in any compact subinterval of $\mathbb{R}$ (since k_m is bounded on such subintervals) and therefore in $\mathbb{R}$ for $m = 1, \ldots, n$. Thus eigenfunctions corresponding to nonzero eigenvalues are $C^n(\mathbb{R})$. This finishes the proof of statement (1).

To prove statement (2) we begin by showing that the series (3.1) with $m = 1$ converges absolutely and pointwise in $\mathbb{R}^2$. Consider

$$k^{[l]}(x,y) = k(x,y) - \sum_{i=1}^{l} \lambda_i \phi_i(x)\overline{\phi_i(y)} = \sum_{i=l+1}^{\infty} \lambda_i \phi_i(x)\overline{\phi_i(y)}. \tag{3.5}$$

The first equality implies that $k^{[l]}(x,y)$ is continuous in $\mathbb{R}$; the second implies that it is a Moore matrix. Since $k_1(x,y)$ is continuous and the ϕ_i are in $C^1(\mathbb{R})$, Corollary 2.7 implies that, for every $l \geq 0$,

$$k_1^{[l]}(x,x) = k_1(x,x) - \sum_{i=1}^{l} \lambda_i |\phi_i'(x)|^2 \geq 0$$

and therefore

$$0 \leq \sum_{i=1}^{l} \lambda_i |\phi_i'(x)|^2 \leq k_1(x,x). \tag{3.6}$$

Thus the series $\sum_{i=1}^{\infty} \lambda_i |\phi_i'(x)|^2$ converges pointwise to a function $d(x)$, since its partial sums are nondecreasing and, by (3.6), bounded by $k_1(x,x)$.

Suppose $L > 0$ and consider the restrictions of k and k_1 to $[-L, L]^2$. Since $k_1(x,x)$ is continuous and positive there exists $M > 0$ such that $k_1(x,x) \leq M$ for $|x| \leq L$. Using (3.6), we have

$$\Big| \sum_{i=m}^{n} \lambda_i \phi_i'(x)\overline{\phi_i'(y)} \Big|^2 \leq \sum_{i=m}^{n} \lambda_i |\phi_i'(x)|^2 \sum_{i=m}^{n} \lambda_i |\phi_i'(y)|^2 \leq M \sum_{i=m}^{n} \lambda_i |\phi_i'(y)|^2. \tag{3.7}$$

Thus, $\sum_{i=1}^{\infty} \lambda_i \phi_i(x)' \, \overline{\phi_i'(y)}$ converges uniformly in $x \in [-L, L]$ for every fixed y. Reversing the roles of x and y in (3.7), we conclude that it also converges uniformly in $y \in [-L, L]$ for every fixed x. Denoting by $k_1^*(x,y)$ the sum of the series, it follows from the continuity of ϕ_i' that $k_1^*(x,y)$ is continuous in x for every fixed y and continuous in y for every fixed x.

Since both k_1 and k_1^* are measurable, it follows that

$$\int_0^y \int_0^x (k_1(u,v) - k_1^*(u,v)) \, du \, dv \tag{3.8}$$

$$= \int_0^y \int_0^x k_1(u,v)\,du\,dv - \sum_i \lambda_i \int_0^x \phi_i'(u)\,du \int_0^y \phi_i'(v)\,dv$$
$$= \int_0^y \int_0^x \frac{\partial^2 k}{\partial y \partial x}(u,v)\,du\,dv - (k(x,y) - k(x,0) - k(0,y) + k(0,0)) = 0$$

for every x and y. Thus $k_1 = k_1^*$ almost everywhere $[dx\,dy]$. Fubini's theorem then implies that, for almost every x, $k_1(x,y) = k_1^*(x,y)$ for almost every y. But, since for every fixed x both k_1 and k_1^* are continuous in y, it follows that for almost every x equality holds for every y. Hence, for every y equality holds for almost every x. But for fixed y both k_1 and k_1^* are continuous in x. Thus $k_1(x,y) = k_1^*(x,y)$ for all $x, y \in [-L, L]^2$. Since L is arbitrary, we have $k_1(x,y) = k_1^*(x,y)$ for all $x, y \in \mathbb{R}$.

The fact that k_1 is a Moore matrix follows directly from the fact that it is the pointwise limit of Moore matrices.

Now observe that the fact that the $\{\phi_i\}_{i\in\mathbb{N}}$ are eigenfunctions of k is necessary only for statement (1), which has already been proved once and for all. Also notice that the proof of pointwise convergence of the series (3.1) in the case $m = 1$ only uses the facts that $\lambda_i \geq 0$, k is a Moore matrix in class $\mathcal{S}_1(\mathbb{R})$, Theorem 2.6 and pointwise convergence of the series (3.1) in the case $m = 0$. To show pointwise convergence of this series for general $m > 1$ we use induction. The remarks in the previous paragraph show that the proof for $m = 1$ adapts immediately, upon replacement of the corresponding objects, to show that pointwise convergence of (3.1) holds for k_{j+1} if it holds for k_j Thus the proof for $m > 1$ proceeds by induction on j, $0 \leq j \leq n-1$. To prove uniform convergence in compacts, we finally consider

$$k_m(x,x) = \sum_{i=1}^{+\infty} \lambda_i |\phi_i^{(m)}(x)|^2.$$

The partial sums of the series form a nondecreasing sequence of continuous functions converging to a continuous function, whence by Dini's theorem convergence is uniform in every compact interval I. The Cauchy-Schwarz inequality

$$\Big| \sum_{i=M}^{n} \lambda_i \phi_i^{(m)}(x) \overline{\phi_i^{(m)}(y)} \Big|^2 \leq \sum_{i=M}^{n} \lambda_i |\phi_i^{(m)}(x)|^2 \sum_{i=M}^{n} \lambda_i |\phi_i^{(m)}(y)|^2$$

then implies that convergence of $\sum_{i\geq 1} \lambda_i \phi_i^{(m)}(x)\overline{\phi_i^{(m)}(y)}$ to $k_m(x,x)$ is uniform in I^2. □

Corollary 3.4. *Let $k(x,y)$ be a Mercer-like kernel in class $\mathcal{S}_n(\mathbb{R})$, $n \geq 0$. Then, in addition to the statements in Lemma 3.3, we have*

$$\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}} k(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i^{(m_1)}(x)\overline{\phi_i^{(m_2)}(y)}$$

pointwise, absolutely in $\mathbb{R}^2$ and uniformly in compact sets for every m_1, $m_2 = 0, 1, \ldots, n$.

Proof. Suppose $L > 0$ and consider the restriction of k to the square $[-L,L]^2$. Fix $m_1, m_2 \leq n$. By Lemma 3.3, the series

$$\sum_{i=1}^{\infty} \lambda_i \phi_i^{(m_1)}(x)\overline{\phi_i^{(m_1)}(y)} \quad \text{and} \quad \sum_{i=1}^{\infty} \lambda_i \phi_i^{(m_2)}(x)\overline{\phi_i^{(m_2)}(y)}$$

converge uniformly and absolutely on $[-L,L]^2$ in both variables x and y. Then $\sum_{i=1}^{\infty} \lambda_i \phi_i^{(m_1)}(x)\overline{\phi_i^{(m_2)}(y)}$ is also absolutely and uniformly convergent since, for any $j_1 < j_2$, x, $y \in [-L,L]$,

$$\Big| \sum_{i=j_1}^{j_2} \lambda_i \phi_i^{(m_1)}(x)\overline{\phi_i^{(m_2)}(y)} \Big| \leq \Big(\sum_{i=j_1}^{j_2} \lambda_i |\phi_i^{(m_1)}(x)|^2 \Big)^{1/2} \Big(\sum_{i=j_1}^{j_2} \lambda_i |\phi_i^{(m_2)}(y)|^2 \Big)^{1/2}.$$

Therefore, (see e.g. [1], Theorem 9.14) successive differentiation of the identity

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

may be performed with respect to x and y, yielding the partial derivatives of k on the left-hand side and the uniformly and absolutely convergent series of continuous functions on the right-hand side. Hence

$$\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}} k(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i^{(m_1)}(x)\overline{\phi_i^{(m_2)}(y)}$$

uniformly and absolutely in $[-L,L]^2$. Since $L > 0$ is arbitrary the result follows. □

Definition 3.5. *Let $n \geq 1$ be an integer. A function $k : \mathbb{R}^2 \to \mathbb{C}$ is said to belong to the class $\mathcal{A}_n(\mathbb{R})$ if $k \in \mathcal{S}_n(\mathbb{R})$ and*

$$k(x,y),\ \frac{\partial^2}{\partial y \partial x} k(x,y),\ \ldots,\ \frac{\partial^{2n}}{\partial y^n \partial x^n} k(x,y)$$

are in the class $\mathcal{A}_0(\mathbb{R})$.

Remark 3.6. Trivially $\mathcal{A}_n(\mathbb{R}) \subset \mathcal{A}_{n-1}(\mathbb{R}) \subset \ldots \subset \mathcal{A}_1(\mathbb{R}) \subset \mathcal{A}_0(\mathbb{R})$. For the significance of the differentiability condition, recall Remark 2.2.

For k in class $\mathcal{A}_n(\mathbb{R})$ and for $m = 0, \ldots, n$, we will write

$$\mathcal{K}_m \equiv \int_{-\infty}^{\infty} k_m(x,x)\,dx.$$

Note that, from Corollary 2.7, $\mathcal{K}_m > 0$ for each $m = 0, \ldots, n$ unless k_m vanishes identically.

Lemma 3.3 and Corollary 3.4 naturally apply to positive definite kernels in $\mathcal{A}_0(\mathbb{R}) \cap \mathcal{S}_n(\mathbb{R})$. Additional properties arise if k is assumed to be in the class $\mathcal{A}_n(\mathbb{R})$, as the following result shows. Here $H^n(I)$ denotes the Sobolev Hilbert space $W^{n,2}(I)$ normed by $\|\phi\|^2_{H^n(I)} = \sum_{m=0}^{n} \|\phi^{(m)}\|^2_{L^2(I)}$.

Theorem 3.7. *Let $k(x,y)$ be a Mercer-like kernel in the class $\mathcal{A}_n(\mathbb{R})$, $n \geq 0$. Then the following statements hold.*

(1) *If ϕ_i is a normalized eigenfunction associated to $\lambda_i \neq 0$, then ϕ_i is in $C^n(\mathbb{R}) \cap H^n(\mathbb{R})$ and*

$$\|\phi_i^{(m)}\|_{L^2(\mathbb{R})} \leq \left(\frac{\mathcal{K}_m}{\lambda_i}\right)^{1/2} \tag{3.9}$$

for every $m = 0, 1, \ldots, n$.

(2) *Each k_m is a positive definite kernel in the class $\mathcal{A}_{n-m}(\mathbb{R})$ and*

$$k_m(x,y) = \sum_{i\geq 1} \lambda_i \phi_i^{(m)}(x)\overline{\phi_i^{(m)}(y)} \tag{3.10}$$

uniformly and absolutely in $\mathbb{R}^2$ for each $m = 0, 1, \ldots, n$.

(3) *The $L^2(\mathbb{R})$ integral operator K_m with kernel k_m is trace class with*

$$\operatorname{tr}(K_m) = \mathcal{K}_m = \int_{-\infty}^{\infty} k_m(x,x)\,dx = \sum_{i=1}^{\infty} \lambda_i \|\phi_i^{(m)}\|^2_{L^2(\mathbb{R})}. \tag{3.11}$$

Proof. Since K is in the class $\mathcal{A}_n(\mathbb{R})$ it verifies the hypothesis of Lemma 3.3; therefore eigenfunctions ϕ_i associated with nonzero eigenvalues are $C^n(\mathbb{R})$.

To prove (2) we recall the corresponding statement in Lemma 3.3. Since by hypothesis k is in the class $\mathcal{A}_n(\mathbb{R})$, it follows that k_m is in the class $\mathcal{A}_{n-m}(\mathbb{R})$ and is a positive definite kernel. It also follows by hypothesis that $k_m(x,x) \to 0$ as $|x| \to \infty$. It is easily seen that, in these circumstances, a simple extension of Dini's theorem to the one-point compactification of $\mathbb{R}$ shows that convergence of the series (3.10) in compact sets implies uniform convergence in $\mathbb{R}^2$.

Thus, $k_m(x,y)$ is the kernel of a trace class operator K_m. Denoting by $\lambda_i^{[m]} \geq 0$ its eigenvalues (with the understanding $\lambda_i^{[0]} \equiv \lambda_i$), uniform convergence of the series yields

$$\mathcal{K}_m = \operatorname{tr}(K_m) = \sum_{i=1} \lambda_i^{[m]} = \int_{\mathbb{R}} k_m(x,x)\,dx = \sum_{i=1} \lambda_i \|\phi_i^{(m)}\|_{L^2}^2. \tag{3.12}$$

Since by hypothesis each $\mathcal{K}_m$ is finite, convergence of the series on the right-hand side of (3.12) implies that $\|\phi_i^{(m)}\|_{L^2}^2$ is finite for each $m = 0, \ldots, n$ and thus $\phi_i \in H^n(\mathbb{R})$. Analogously, from (3.12) it follows that

$$\lambda_i \|\phi_i^{(m)}\|_{L^2}^2 \leq \sum_{i=1} \lambda_i \|\phi_i^{(m)}\|_{L^2}^2 = \mathcal{K}_m$$

from which the norm bound (3.9) is derived, finishing the proof of statement (1). This completes the proof. □

Corollary 3.8. *Let $k(x,y)$ be a positive definite kernel in the class $\mathcal{A}_n(\mathbb{R})$. Then*

$$\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}} k(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i^{(m_1)}(x)\overline{\phi_i^{(m_2)}(y)}$$

uniformly and absolutely in $\mathbb{R}^2$ for every m_1, $m_2 = 0, 1, \ldots, n$.

Proof. The proof is similar to the proof of Corollary 3.4, yielding in this case, according to Theorem 3.7, uniform convergence in $\mathbb{R}^2$.

Theorem 3.7 admits the following converse.

Theorem 3.9. *Let $\{\lambda_i\}_{i\geq 1}$ be a sequence of nonnegative real numbers and $\{\phi_i\}_{i\geq 1}$ be a sequence of $C^n(\mathbb{R})$ functions such that, for each $m = 0, 1, \ldots n$, the series $\sum_{i=1}^{\infty} \lambda_i \phi_i^{(m)}(x)\overline{\phi_i^{(m)}(y)}$ converge uniformly and absolutely on $\mathbb{R}^2$. Set $k(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i(x)\overline{\phi_i(y)}$. Then*

(1) *$k_m(x,y) = \dfrac{\partial^{2m}}{\partial y^m \partial x^m} k(x,y)$ is a continuous Moore matrix on $\mathbb{R}^2$ for each $m = 0, 1, \ldots n$ and*

$$\frac{\partial^{2m}}{\partial y^m \partial x^m} k(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i^{(m)}(x)\overline{\phi_i^{(m)}(y)}. \tag{3.13}$$

(2) *For every m_1, $m_2 = 0, \ldots, n$, we have*

$$\frac{\partial^{m_1+m_2}}{\partial y^{m_2}\partial x^{m_1}} k(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i^{(m_1)}(x)\overline{\phi_i^{(m_2)}(y)} \tag{3.14}$$

uniformly and absolutely in $\mathbb{R}^2$.

(3) *If, in addition, for every* $0 \le m \le n$ *we have* $\phi_i^{(m)} \in L^2(\mathbb{R})$, $\sum_{i=1}^{\infty} \lambda_i \|\phi_i^{(m)}\|_{L^2(\mathbb{R})}^2 < \infty$ *and* $\sum_{i=1}^{\infty} \lambda_i |\phi_i^{(m)}(x)|^2 \to 0$ *as* $|x| \to \infty$, *then each* $k_m(x,y)$ *is a positive definite kernel in the class* $\mathcal{A}_{n-m}(\mathbb{R})$.

Proof. We begin by proving statement (1). Fixing n, choosing $0 \le m \le n$ and setting

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

it is immediate to conclude that k_m^* is a Moore matrix since it is the limit of a sequence of Moore matrices. Continuity of k_m^* follows from the uniform convergence of the series and continuity of the ϕ_i. Since $k(x,y) = k_0^*(x,y)$ the first statement is proved for $n = 0$. We now proceed by induction.

Given $n \in \mathbb{N}$, suppose k_m is defined by (3.13) for every $(x,y) \in \mathbb{R}^2$ and every $m = 0, 1, \ldots n$ and that $\{\phi_i\}_{i \ge 1}$ is a sequence of $C^{n+1}(\mathbb{R})$ functions. To complete the proof we are required to show that if the series $\sum_{i=1}^{\infty} \lambda_i \phi_i^{(n+1)}(x)\overline{\phi_i^{(n+1)}(y)}$ converges uniformly and absolutely on $\mathbb{R}^2$, then

$$k_{n+1}(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i^{(n+1)}(x)\overline{\phi_i^{(n+1)}(y)}.$$

The argument is similar to the one used in Corollary 3.4.

We first observe that, since both the series $\sum_{i=1}^{\infty} \lambda_i \phi_i^{(n+1)}(x)\overline{\phi_i^{(n+1)}(y)}$ and $\sum_{i=1}^{\infty} \lambda_i \phi_i^{(n)}(x)\overline{\phi_i^{(n)}(y)}$ are uniformly convergent, the series

$$\sum_{i=1}^{\infty} \lambda_i \phi_i^{(n)}(x)\overline{\phi_i^{(n+1)}(y)} \tag{3.15}$$

is also uniformly convergent, since

$$\Big| \sum_{i=j_1}^{j_2} \lambda_i \phi_i^{(n)}(x)\overline{\phi_i^{(n+1)}(y)} \Big| \le \Big(\sum_{i=j_1}^{j_2} \lambda_i |\phi_i^{(n)}(x)|^2 \Big)^{1/2} \Big(\sum_{i=j_1}^{j_2} \lambda_i |\phi_i^{(n+1)}(y)|^2 \Big)^{1/2}.$$

Denote the sum (3.15) by $k_{n;y}^*$. For fixed y, the series

$$\sum_{i=1}^{\infty} \lambda_i \phi_i^{(n)}(x)\overline{\phi_i^{(n+1)}(y)} \quad \text{and} \quad \sum_{i=1}^{\infty} \lambda_i \phi_i^{(n+1)}(x)\overline{\phi_i^{(n+1)}(y)}$$

satisfy sufficient conditions for the former to be the derivative of the latter [1]. Therefore $\frac{\partial}{\partial x} k_{n;y}^*$ exists and coincides with $k_{n+1;y}^*$.

In a similar way, for fixed x we apply the same result to the series

$$\sum_{i=1}^{\infty} \lambda_i \phi_i^{(n)}(x)\overline{\phi_i^{(n)}(y)} \quad \text{and} \quad \sum_{i=1}^{\infty} \lambda_i \phi_i^{(n)}(x)\overline{\phi_i^{(n+1)}(y)},$$

concluding that $\frac{\partial}{\partial y} k_n = k^*_{n;y}$. Consequently, $\frac{\partial^2}{\partial y \partial x} k_n = k_{n+1}$ exists and we have $k_{n+1} = k^*_{n+1}$. Similar arguments may now be used to derive statement (2) from statement (1) as in the proof of Corollary 3.4.

We now prove statement (3). Fix n and m such that $0 \le m \le n$. To prove that $k_m(x,y)$ is a positive definite kernel in the class $\mathcal{A}_{n-m}(\mathbb{R})$ it is now sufficient to show that $k_n(x,y)$ is in the class $\mathcal{A}(\mathbb{R})$.

By (3.13),

$$k_n(x,y) = \frac{\partial^{2n} k}{\partial y^n \partial x^n}(x,y) = \sum_{i=1}^{\infty} \lambda_i \phi_i^{(n)}(x)\overline{\phi_i^{(n)}(y)},$$

where the series is absolutely and uniformly convergent. Therefore, according to the hypothesis,

$$\begin{aligned} \int_{-\infty}^{+\infty} k_n(x,x)\,dx &= \int_{-\infty}^{+\infty} \sum_{i=1}^{\infty} \lambda_i \phi_i^{(n)}(x)\overline{\phi_i^{(n)}(x)}\,dx \\ &= \sum_{i=1}^{\infty} \lambda_i \int_{-\infty}^{+\infty} |\phi_i^{(n)}(x)|^2\,dx = \sum_{i=1}^{\infty} \lambda_i \,\|\phi_i^{(m)}\|^2_{L^2(\mathbb{R})} < \infty. \end{aligned} \tag{3.16}$$

The last hypothesis in statement (2) finally yields

$$k_n(x,x) = \sum_{i=1}^{\infty} \lambda_i |\phi_i^{(n)}(x)|^2 \to 0 \text{ as } |x| \to \infty,$$

concluding the proof. □

Remark 3.10. It is straightforward to verify that the hypotheses required in Theorem 3.9 are sharp, in the sense that if any of them fails to hold the corresponding conclusion is no longer generally true.

Remark 3.11. The expansion

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

in Theorem 3.9 is not supposed to be the Schmidt series of the kernel of an integral operator. In particular, the functions ϕ_i are not necessarily eigenfunctions of k, not even orthogonal. Accordingly, the λ_i are not necessarily the eigenvalues of the corresponding integral operator. However, even in these conditions the k_m are positive definite kernels and the associated operators K_m are trace class with trace given by

$$\operatorname{tr}(K_m) = \sum_{i=1}^{\infty} \lambda_i \|\phi^{(m)}\|^2_{L^2(\mathbb{R})}.$$

Remark 3.12. In the conditions of Theorem 3.7, the eigenvalues λ_i and the derivatives $\phi_i^{(m)}$ of the corresponding eigenfunctions of K are not, in general, respectively eigenvalues and eigenfunctions of K_m, $m \geq 1$; explicit examples are easily constructed. Thus the trace formula (3.11) becomes especially interesting: it does not originate in a Hilbert-Schmidt series for k_m but, nevertheless, it allows us to conclude that the operator K_m with kernel k_m is trace class and (3.11) provides us a nontrivial way to compute this trace.

Using the Hilbert-Schmidt series for k_l, the following generalization of Theorem 3.7 may be derived. We write

$$k_l(x,y) = \sum_{i\geq 1} \lambda_i^{[l]} \phi_i^{[l]}(x)\overline{\phi_i^{[l]}(y)},$$

where $\lambda_i^{[l]} \neq 0$ are the nonzero eigenvalues of the operator K_l with kernel k_l and $\{\phi_i^{[l]}\}_{i\geq 1}$ the corresponding set of orthonormalized eigenfunctions of K_l. It is understood that $\lambda_i^{[0]} \equiv \lambda_i$, $\phi_i^{[0]} \equiv \phi_i$.

Corollary 3.13. *Let $k(x,y)$ be a positive definite kernel in the class $\mathcal{A}_n(\mathbb{R})$, $n \geq 0$. Then the following statements hold.*

(1) *If $\phi_i^{[l]}$ is a normalized eigenfunction associated to $\lambda_i^{[l]} \neq 0$, then $\phi_i^{[l]}$ is in $C^{n-l}(\mathbb{R}) \cap H^{n-l}(\mathbb{R})$ and*

$$\|\phi_i^{[l](m-l)}\|_{L^2(\mathbb{R})} \leq \left(\frac{\mathcal{K}_m}{\lambda_i^{[l]}}\right)^{1/2} \tag{3.17}$$

for every $0 \leq l \leq m \leq n$.

(2) *Each k_m is a positive definite kernel in class $\mathcal{A}_{n-m}(\mathbb{R})$ and*

$$k_m(x,y) = \sum_{i\geq 1} \lambda_i^{[l]} \phi_i^{[l](m-l)}(x)\overline{\phi_i^{[l](m-l)}(y)} \tag{3.18}$$

uniformly and absolutely in $\mathbb{R}^2$ for $0 \leq l \leq m \leq n$.

(3) *The $L^2(\mathbb{R})$ integral operator K_m with kernel k_m is trace class with*

$$\mathrm{tr}(K_m) = \mathcal{K}_m = \int_{-\infty}^{\infty} k_m(x,x)\,dx = \sum_{i=1}^{\infty} \lambda_i^{[l]} \|\phi_i^{[l](m-l)}\|^2_{L^2(\mathbb{R})} \tag{3.19}$$

for $0 \le l \le m \le n$.

Proof. We proceed exactly as in the proof of Theorem 3.7, replacing k with k_l, ϕ_i with $\phi_i^{[l]}$, $\phi_i^{(m)}$ with $\phi_i^{[l](m-l)}$ and λ_i with $\lambda_i^{[l]}$. □

Remark 3.14. The norm estimates (3.9) in Theorem 3.7 and (3.17) are sharp, in the sense that for the rank 1 example $k_l(x,y) = \phi(x)\,\overline{\phi(y)}$, where $\phi \in L^2(\mathbb{R})$ is $C^{n-l}(\mathbb{R})$, $0 \le l \le n$ the equalities hold.

Recalling that, for each l, $\{\lambda_i^{[l]}\}_{i\in\mathbb{N}}$ is a nondecreasing null sequence, denote by $E_N^{[l]} = \oplus_{i=1}^N E_{\lambda_i^{[l]}}$ the direct sum of the eigenspaces associated with the first N nonzero eigenvalues of k_l. Statement (1) of Corollary 3.13 then admits the following extension.

Corollary 3.15. *Suppose $k(x,y)$ is a positive definite kernel in class $\mathcal{A}_n(\mathbb{R})$ and let $0 \le l \le m \le n$. Then for any $\phi \in E_N^{[l]}$*

$$\|\phi^{(m-l)}\|_{L^2(\mathbb{R})} \le \left(\frac{\mathcal{K}_m}{\lambda_N^{[l]}}\right)^{1/2} \|\phi\|_{L^2(\mathbb{R})}. \tag{3.20}$$

Proof. Since $\{\phi_i^{[l]}\}_{i=1}^N$ constitute a basis for $E_N^{[l]}$, we have $\phi = \sum_{i=1}^N c_i \phi_i^{[l]}$. For $l \le m \le n$

$$\begin{aligned}\|\phi^{(m-l)}\|^2_{L^2(I)} &= \Big\| \sum_{i=1}^N c_i \phi_i^{[l](m-l)} \Big\|^2_{L^2(I)} \le \Big(\sum_{i=1}^N |c_i| \Big\| \phi_i^{[l](m-l)} \Big\|_{L^2(I)} \Big)^2 \\ &\le \Big(\sum_{i=1}^N \frac{|c_i|^2}{\lambda_i^{[l]}} \Big) \Big(\sum_{i=1}^N \lambda_i^{[l]} \Big\| \phi_i^{[l](m-l)} \Big\|^2_{L^2(I)} \Big) = \Big(\sum_{i=1}^N \frac{|c_i|^2}{\lambda_i^{[l]}} \Big) \mathcal{K}_m, \end{aligned} \tag{3.21}$$

where in the last step use has been made of the expression for the trace formula in Corollary 3.13. Since $\{\lambda_i^{[l]}\}_{i\in\mathbb{N}}$ is nondecreasing and $\|\phi\|^2_{L^2(I)} = \sum_{i=1}^N |c_i|^2$ because the $\{\phi_i^{[l]}\}_{i=1}^N$ are $L^2(I)$-orthonormal, (3.21) implies

$$\|\phi^{(m-l)}\|^2_{L^2(I)} \le \frac{\mathcal{K}_m}{\lambda_N^{[l]}} \|\phi\|^2_{L^2(I)},$$

proving (3.20). □

Corollary 3.16. *Suppose $k(x,y)$ is a positive definite kernel in class $\mathcal{A}_n(\mathbb{R})$ and let $0 \leq l \leq n$. Then for any $\phi \in E_N^{[l]}$*

$$\|\phi\|_{H^{n-l}(\mathbb{R})} \leq \left(\frac{\sum_{m=l}^{n} \mathcal{K}_m}{\lambda_N^{[l]}}\right)^{1/2} \|\phi\|_{L^2(\mathbb{R})}.$$

Proof. Immediate from Corollary 3.15 and the definition of the Sobolev norm. ☐

Remark 3.17. The results in Section 3 are formulated, for the sake of definiteness, for the case $I = \mathbb{R}$. This is because the compact case is very well studied, and the main novelty of our methods lies in the noncompact case. Of course, all our results remain valid for an arbitrary interval $I \subseteq \mathbb{R}$, bounded or unbounded, and Section 3 could have been thus formulated. For compact I many of our constructions collapse trivially: for instance, the class $\mathcal{A}_0(I)$ is merely the class of continuous functions, and the differential class $\mathcal{S}_n(I)$ is automatically contained in the integral class $\mathcal{A}_n(I)$. Thus, for instance, our results for I compact yield as special cases, for $n = 0$, the classical theorem of Mercer, and for $n > 0$ Kadota's theorems [7]. It is only in the unbounded case that the distinction and role of each type of condition becomes meaningful. On the other hand, all our results adapt in the obvious way to the other types of noncompact intervals $I \subset \mathbb{R}$.

REFERENCES

[1] T. Apostol, "Mathematical Analysis," (2nd ed.), Addison-Wesley, 1978.

[2] N. Aronszajn, *Theory of reproducing kernels,* Trans. Amer. Math. Soc., 68 (1950), 337–404.

[3] J. Buescu, *Positive integral operators in unbounded domains,* J. Math. Anal. Appl., 296 (2004), 244–255.

[4] J. Buescu, F. Garcia, I. Lourtie, and A. Paixão, *Positive definiteness, integral equations and Fourier transforms,* Jour. Int. Eq. Appl., 16 (2004), 33–52.

[5] F. Cucker and S. Smale, *On the mathematical foundations of learning,* Bull. Am. Math. Soc., 39 (2002) 1, 1-49.

[6] C. Fitzgerald and R. Horn, *On the structure of Hermitian-symmetric inequalities,* J. Lond. Math. Soc., 15 (1977), 419–430.

[7] T. Kadota, *Term-by-term differentiability of Mercer's expansion,* Proc. Amer. Math. Soc., 18 (1967), 69–72.

[8] M.G. Krein, *Hermitian positive kernels on homogeneous spaces I,* Amer. Math. Soc. Transl., 34 (1963), 69–108.

[9] R. Horn and C. Johnson, "Matrix Analysis," Cambridge University Press, Cambridge, 1985.

[10] E. H. Moore, "General Analysis," Memoirs of Amer. Philos. Soc. Part I (1935), Part II (1939).

[11] F. Riesz and B. Nagy, "Functional Analysis," Ungar, New York, 1952.
[12] S. Saitoh, "Theory of Reproducing Kernels and its Applications," Pitman Research Notes in Mathematics Series, 189, Longman, 1988.
[13] S. Saitoh, "Integral Transforms, Reproducing Kernels and Their Applications," Pitman Research Notes in Mathematics Series, 369, Longman, 1997.
[14] J. Stewart, *Positive definite functions and generalizations, an historical survey,* Rocky Mountain Jour. Math., 6 (1976), 409-434.