Reproducing kernels, spline interpolation and Gaussian processes *

Michael Unser †

March 29, 2017

Contents

1 Preliminaries **4**

1.1 Brief overview of continuous-domain operators 4

1.2 Inner products versus duality product 5

1.2.1 Hilbert spaces and inner products 6

1.2.2 Duality product 6

1.2.3 Hilbert-conjugate vs. adjoint operators 7

1.3 Self-adjoint operators and positive-definite kernels 8

1.4 Schwartz's space of test function: density properties 10

2 Reproducing kernel Hilbert spaces (RKHS) **13**

2.1 Definition of RKHS 13

2.2 Decay and continuity properties 17

2.3 RKHS: the simplified finite-dimensional story 23

2.4 RKHS associated with an invertible operator 27

2.5 Factorization of a reproducing kernel 28

2.6 RKHS associated with the derivative operator 33

2.7 Operators with non-trivial null spaces 39

2.7.1 Hilbert-space structure of the null space 40

2.7.2 Conditional positive-definiteness 44

2.7.3 Admissible regularization operators 48

2.7.4 RKHS associated with an admissible operator 50

*The research leading to these results has received funding from the Swiss National Science Foundation under Grant 200020-162343.

†Biomedical Imaging Group, École polytechnique fédérale de Lausanne (EPFL), Station 17, CH-1015, Lausanne, Switzerland (michael.unser@epfl.ch).

2.7.5 Determination of the reproducing kernel 53
2.8 Factorization and Banach space extensions 64

3 Variational splines and representer theorems 68
3.1 Regularization functional induced by an inner product 69
3.1.1 Smoothing splines and ridge regression 73
3.1.2 Representer theorem for statistics and machine learning 75
3.2 Non-coercive regularization functionals 77
3.2.1 Generalized boundary value problem 78
3.2.2 Proper regularization of an inverse problem 80
3.2.3 Representer theorem for linear inverse problems 82
3.3 Discretization and numerical solutions 84
3.3.1 Least-squares approximation problems 86
3.3.2 Specific examples 87
3.3.3 Generalized interpolation revisited 87
3.4 Epilogue: back to the finite-dimensional world 89

4 Gaussian processes 92
4.1 Generalized stochastic processes (GSP) 92
4.1.1 GSPs as random generalized functions 94
4.1.2 GSPs as random linear functionals 95
4.1.3 GSPs as consistent generators of random variables . . 98
4.1.4 Examples 98
4.2 Mean and covariance forms 99
4.2.1 Reproducing kernels and mean-square continuity . . . 102
4.2.2 Effect of a linear transformation 104
4.2.3 Stationary processes 105
4.3 The characteristic functional 105
4.4 Characterization of Gaussian processes 111
4.5 Gaussian solutions of stochastic differential equations 119
4.6 MMSE solution of linear inverse problems 122

A Positivity and positive-definiteness 127

B Quadratic optimization in Hilbert spaces 134
B.1 Linear and bilinear forms 134
B.2 Connection between bilinear forms and operators 136
B.3 Quadratic minimization problems 137

C Foundations of convex optimization 140

D Nuclear-Fréchet structure of $\mathcal{S}(\mathbb{R}^d)$ and the kernel theorem 143

E Finite-dimensional probability theory 145

E.1 Background: The probabilistic formalism 145

E.2 Probability density functions and expectations 147

E.3 Characteristic function 149

E.4 Multivariate Gaussian distributions 153

E.5 Gaussian conditional probabilities 154

F Basic definitions from measure theory 155

1 Preliminaries

1.1 Brief overview of continuous-domain operators

The proper mathematical context is given by Schwartz' theory of generalized functions where $\mathcal{S}'(\mathbb{R}^d)$ is the space of tempered distributions. It is arguably the most comprehensive theory for continuous-domain signals and linear operators acting on such signals because:

- it provides a complete characterization of linear operators in terms of some generalized "integral" equation (see Schwartz' kernel theorem, which is explained next);
- it supports the use of the Fourier transform in its full generality (i.e., the generalized Fourier transform $\mathcal{F}$ is a continuous reversible map $\mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ that coincides with the usual definition for functions that are absolutely integrable).

Formally, a tempered distribution $f : \varphi \mapsto \langle f, \varphi \rangle$ is a linear (and continuous) functional that associates a real number denoted by $\langle f, \varphi \rangle$ to each test function $\varphi \in \mathcal{S}(\mathbb{R}^d)$ (Schwartz's space of smooth and rapidly decaying functions). For instance, the Dirac impulse at location $\boldsymbol{x}_0 \in \mathbb{R}^d$ is the generalized function $\delta(\cdot - \boldsymbol{x}_0) \in \mathcal{S}'(\mathbb{R}^d)$ defined as

$$\varphi \mapsto \langle \delta(\cdot - \boldsymbol{x}_0), \varphi \rangle = \varphi(\boldsymbol{x}_0)$$

Here, the dot "$\cdot$" is used as placeholder for the domain variable (i.e., φ, $\varphi(\boldsymbol{x})$, or $\varphi(\cdot)$ are equivalent notations for the same object which is a function $\varphi : \mathbb{R}^d \to \mathbb{R}$), while $\boldsymbol{x}_0$ is a fixed offset that indicates the location of the impulse. In the case where f is an ordinary function of the variable $\boldsymbol{x} \in \mathbb{R}^d$, the so-called "duality product" is given by

$$\langle f, \varphi \rangle = \int_{\mathbb{R}^d} f(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{1}$$

which is a conventional integral.

The fundamental result for our purpose is Schwartz' kernel theorem, which states that any continuous linear operator $\mathrm{G} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ admits an "integral" representation as

$$\mathrm{G}\{\varphi\}(\boldsymbol{x}) = \langle g(\boldsymbol{x}, \cdot), \varphi \rangle = \int_{\mathbb{R}^d} g(\boldsymbol{x}, \boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y} \tag{2}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ with $g(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$. In essence, $g(\cdot,\cdot)$ is the continuous-domain analog of the matrix that specifies a finite-dimensional linear operator. The kernel of G is identified by formally applying the operator to a shifted Dirac impulse

$$g(\boldsymbol{x},\boldsymbol{y}) = \mathrm{G}\{\delta(\cdot - \boldsymbol{y})\}(\boldsymbol{x}), \tag{3}$$

which is the reason why $g(\boldsymbol{x},\boldsymbol{y})$ is also called the *generalized impulse response* of the operator. Note, however, that the kernel $g(\cdot,\cdot)$ is not always an "ordinary" bivariate function $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, but rather a tempered distribution in the cross-product space $\mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$. The truly powerful aspect of the kernel theorem is that the implication also goes the other way around: Any kernel $g(\cdot,\cdot) \in \mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ specifies a continuous operator $\mathrm{G} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ via equation (2). Moreover, two operators are identical if and only if their kernels are equal (in the sense of distributions).

To take us back to a more classical setting where signals and kernels are ordinary functions of the index variables $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d$, we invoke an extended version of the kernel theorem, due to the famous mathematician Grothendieck, that guarantees the existence of two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}_0$

$$\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{H}, \mathcal{H}_0 \subseteq \mathcal{S}'(\mathbb{R}^d)$$

such that the operator has a continuous extension $\mathrm{G} : \mathcal{H} \to \mathcal{H}_0$, which is also defined by (2), but with $\varphi \in \mathcal{H}$. This is equivalent to the existence of a constant $C_0 > 0$ such that

$$\|\mathrm{G}\{f\}\|_{\mathcal{H}_0} \le C_0 \|f\|_{\mathcal{H}}.$$

From a pragmatic point of view, this means that, given any kernel $g(\boldsymbol{x},\boldsymbol{y})$, it is always possible to specify a Hilbert space on which the corresponding operator is well defined. Among those spaces, the reproducing kernel Hilbert spaces (RKHS) are the ones with the strongest practical appeal because their members are conventional functions of the variable $\boldsymbol{x}$, although not necessarily square integrable (see Theorem 7).

1.2 Inner products versus duality product

Having a good grasp of the distinction between the two parallel notions of "inner product" and "duality product" is essential since it is driving the whole formulation. The duality product, on the one hand, is unique (and universal) as it expresses the pairing of a function space to its topological dual. An inner product, on the other hand, specifies a Hilbert space and is typically tied to a given problem, an operator, or a positive-definite kernel. There is no single inner product—the variations on the theme are essentially limitless.

1.2.1 Hilbert spaces and inner products

We briefly recall the defining properties of an inner product. For simplicity, all (generalized) functions are assumed to be real-valued.

Definition 1 (Inner product)**.** *Let $\mathcal{H}$ be a linear (or vector) space. A real-valued inner product on $\mathcal{H}$ is a bilinear form that associates to each pair $f, g \in \mathcal{H}$ a real number denoted by $\langle f, g\rangle_{\mathcal{H}}$ that satisfies the following properties for all $f, g, h \in \mathcal{H}$ and $\alpha \in \mathbb{R}$.*

- *Linearity:* $\langle \alpha f, g\rangle_{\mathcal{H}} = \alpha\langle f, g\rangle_{\mathcal{H}}$ *and* $\langle f+g, h\rangle_{\mathcal{H}} = \langle f, h\rangle_{\mathcal{H}} + \langle g, h\rangle_{\mathcal{H}}$.
- *Symmetry:* $\langle f, g\rangle_{\mathcal{H}} = \langle g, f\rangle_{\mathcal{H}}$.
- *Non-negativity:* $\langle f, f\rangle_{\mathcal{H}} \geq 0$.
- *Unicity:* $\langle f, f\rangle_{\mathcal{H}} = 0 \Leftrightarrow f = 0$.

If all conditions except the last are met, then $\langle f, g\rangle_{\mathcal{H}}$ is called a semi-inner product.

An inner product automatically determines a norm by the formula $\|f\|_{\mathcal{H}} = \sqrt{\langle f, f\rangle_{\mathcal{H}}}$. A Hilbert space is a complete (i.e., closed) normed space whose norm is induced by an inner product; it is separable if it admits a countable basis. The classical example of a separable Hilbert space is Lebesgue's space of square-integrable functions $L_2(\mathbb{R}^d)$.

1.2.2 Duality product

The canonical duality product $\langle \cdot, \cdot\rangle$ is the continuous bilinear form $\mathcal{S}'(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d) \to \mathbb{R}$ that represents the action of a linear functional $f \in \mathcal{S}'(\mathbb{R}^d)$ on a test function $\varphi \in \mathcal{S}(\mathbb{R}^d)$:

$$f : \varphi \mapsto \langle f, \varphi\rangle. \tag{4}$$

One then extends the notion to any dual pair of Banach spaces $(\mathcal{X}', \mathcal{X})$ in $\mathcal{S}'(\mathbb{R}^d)$ (see Definition 5) by writing the action of $f \in \mathcal{X}'$ on a function $g \in \mathcal{X}$ as

$$f : g \mapsto \langle f, g\rangle.$$

The implicit understanding here is that the bilinear form remains continuous in both arguments because of the bound

$$|\langle f, g\rangle| \leq \|f\|_{\mathcal{X}'}\|g\|_{\mathcal{X}}$$

for all $f \in \mathcal{X}'$ and $g \in \mathcal{X}$, which actually defines duality. This extension, which is compatible with the canonical form (4) when $\varphi \in \mathcal{S}(\mathbb{R}^d)$, is supported by the Hahn-Banach theorem.

In particular when $\mathcal{X}' = \mathcal{H}'$ and $\mathcal{X} = \mathcal{H}$ are a dual pair of Hilbert spaces, we have that

$$\langle f, g \rangle = \langle \mathrm{R}\{f\}, g \rangle_{\mathcal{H}} = \langle f, \mathrm{R}^{-1}\{g\} \rangle_{\mathcal{H}'} \tag{5}$$

where the unitary pair of operators $\mathrm{R} : \mathcal{H}' \to \mathcal{H}$ and $\mathrm{R}^{-1} = \mathrm{R}^H : \mathcal{H}' \to \mathcal{H}$ are the Riesz maps that are encoding the isometric isomorphism between the two spaces (see Theorem 4).

1.2.3 Hilbert-conjugate vs. adjoint operators

Two distinct notions are also required to describe the interaction of operators with inner products versus the duality product: the Hilbert conjugate (which is dependent upon the inner product) vs. the adjoint operator (whose definition is universal). The fact that the two concepts are often represented using the same symbol "*" can be a source of confusion.

Let $(\mathcal{H}', \mathcal{H})$ be a dual pair of Hilbert spaces and G a continuous operator $\mathrm{G} : \mathcal{H} \to \mathcal{H}'$. Since we are dealing with normed spaces, the continuity assumption is equivalent to the existence of a constant (the induced norm of the operator) denoted by $\|\mathrm{G}\|$ such that

$$\|\mathrm{G}f\|_{\mathcal{H}'} \leq \|\mathrm{G}\| \, \|f\|_{\mathcal{H}}$$

for all $f \in \mathcal{H}$.

The Hilbert (or Hermitian) conjugate of G is then defined as the (unique) operator $\mathrm{G}^H : \mathcal{H}' \to \mathcal{H}$ that satisfies

$$\langle \mathrm{G}f, g \rangle_{\mathcal{H}'} = \langle f, \mathrm{G}^H g \rangle_{\mathcal{H}}$$

for all $f \in \mathcal{H}$ and $g \in \mathcal{H}'$.

Since $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{H}$ and $\mathcal{H}' \subseteq \mathcal{S}'(\mathbb{R}^d)$, the operator $\mathrm{G} : \mathcal{H} \to \mathcal{H}'$ is represented by its kernel via (2) (Schwartz' kernel theorem), while its restriction to $\mathcal{S}(\mathbb{R}^d)$ is guaranteed to be continuous. One then defines the adjoint of $\mathrm{G} : \mathcal{H} \to \mathcal{H}'$ as the unique operator $\mathrm{G}^* : \mathcal{H}' \to \mathcal{H}$ such that

$$\langle \mathrm{G}\{\varphi_1\}, \varphi_2 \rangle = \langle \mathrm{G}^*\{\varphi_2\}, \varphi_1 \rangle = \langle \varphi_1, \mathrm{G}^*\{\varphi_2\} \rangle,$$

for all $\varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R}^d)$. Note that the right-hand side of the above identity (which is easier to remember) uses an extended interpretation of the duality product for $\mathcal{H}' \times \mathcal{H} \to \mathbb{R}$, which is supported by the Hahn-Banach theorem with the bound

$$|\langle \mathrm{G}\{\varphi_1\}, \varphi_2\rangle| = |\langle \varphi_1, \mathrm{G}^*\{\varphi_2\}\rangle| \leq \|\mathrm{G}\| \, \|\varphi_1\|_{\mathcal{H}} \|\varphi_2\|_{\mathcal{H}}.$$

The (unique) kernel representation of G^* is

$$\mathrm{G}^*\{\varphi\}(\boldsymbol{x}) = \langle g(\cdot, \boldsymbol{x}), \varphi\rangle = \int_{\mathbb{R}^d} g(\boldsymbol{y}, \boldsymbol{x})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y} \tag{6}$$

which is the "transposed" version of (2) where the index variables $\boldsymbol{x}$ and $\boldsymbol{y}$ have been simply interchanged.

By applying the above definitions, we have that

$$\begin{aligned}\langle \mathrm{G}\{\varphi_1\}, \varphi_2\rangle = \langle \mathrm{G}\{\varphi_1\}, \mathrm{R}^{-1}\{\varphi_2\}\rangle_{\mathcal{H}'} &= \langle \varphi_1, \mathrm{G}^H\mathrm{R}^{-1}\{\varphi_2\}\rangle_{\mathcal{H}} \\ &= \langle \mathrm{R}\mathrm{G}^H\mathrm{R}^{-1}\{\phi_2\}, \varphi_1\rangle = \langle \mathrm{G}^*\{\varphi_2\}, \varphi_1\rangle\end{aligned}$$

which shows that $\mathrm{G}^* = \mathrm{R}\mathrm{G}^H\mathrm{R}^{-1}$ with $\mathrm{R}^{-1} : \mathcal{H} \to \mathcal{H}'$, $\mathrm{G}^H : \mathcal{H}' \to \mathcal{H}$ and $\mathrm{R} : \mathcal{H} \to \mathcal{H}'$. In other words, the Hilbert conjugate and adjoint operators of $\mathrm{G} : \mathcal{H} \to \mathcal{H}'$ are equivalent iff. the Riesz map R is the identity; that is, when $\mathcal{H} = \mathcal{H}' = L_2(\mathbb{R}^d)$.

1.3 Self-adjoint operators and positive-definite kernels

Equation (6) implies that the kernel of a *self-adjoint* operator is symmetric; i.e.,

$$\mathrm{H} = \mathrm{H}^* \quad \Leftrightarrow \quad h(\boldsymbol{x}, \boldsymbol{y}) = h(\boldsymbol{y}, \boldsymbol{x}) \text{ for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d.$$

Among the class of self-adjoint operators, the most favorable ones are those whose kernel is positive-definite.

Definition 2. *A kernel function $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that*

$$h(\boldsymbol{x}, \boldsymbol{y}) = h(\boldsymbol{y}, \boldsymbol{x})$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ is said to be symmetric. Moreover, it is positive semi-definite (or positive definite, for short) if

$$\sum_{m=1}^{N}\sum_{n=1}^{N} z_m h(\boldsymbol{x}_m, \boldsymbol{x}_n) z_n \geq 0$$

for any $N \in \mathbb{N}$, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N \in \mathbb{R}^d$, and $z_1, \dots, z_N \in \mathbb{R}$.

Remarkably, there is a formal equivalence between positive-definite kernels and inner products.

Theorem 1 (Moore-Aronszajn [**?**]). *The kernel function* $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *is symmetric positive (semi-)definite if and only if there exists some Hilbert space* $\mathcal{H}$ *and a families of elements* $\{f_{\boldsymbol{x}}\}_{\boldsymbol{x}\in\mathbb{R}^d}$ *in* $\mathcal{H}$ *such that*

$$h(\boldsymbol{x}, \boldsymbol{y}) = \langle f_{\boldsymbol{x}}, f_{\boldsymbol{y}} \rangle_{\mathcal{H}}. \tag{7}$$

In particular, there is a unique Hilbert space $\mathcal{H}_{\mathrm{rep}}$*—the reproducing kernel Hilbert space of* h*—such that* (7) *holds with* $f_{\boldsymbol{x}} = h(\cdot, \boldsymbol{x})$*.*

Proof. We shall only prove the (easy) direct part of the statement. To that end, we use the announced form of $h(\cdot, \cdot)$ to evaluate

$$\begin{aligned}\sum_{m=1}^{N}\sum_{n=1}^{N} z_m h(\boldsymbol{x}_m, \boldsymbol{x}_n) z_n &= \sum_{m=1}^{N}\sum_{n=1}^{N} z_m \langle f_{\boldsymbol{x}_m}, f_{\boldsymbol{x}_n} \rangle_{\mathcal{H}} z_n \\ &= \langle \sum_{i=1}^{n} z_m f_{\boldsymbol{x}_m}, \sum_{j=1}^{n} z_n f_{\boldsymbol{x}_n} \rangle_{\mathcal{H}} \quad \text{(bilinearity of } \langle \cdot, \cdot \rangle) \\ &= \| \sum_{i=1}^{n} z_m f_{\boldsymbol{x}_m} \|_{\mathcal{H}}^2 \geq 0\end{aligned}$$

The remainder of the proof is more technical—we refer to [, Section 1.3, pp. 13-23] for a comprehensive exposition. □

The fundamental outcome of Theorem 1 is that there is a perfect equivalence between positive-definite kernels and RKHS, which are defined in Section 2.

1.4 Schwartz's space of test function: density properties

As complement, we now present some important topological properties of Schwartz' space of test functions $\mathcal{S}(\mathbb{R}^d)$, which will be invoked in some of our derivations. Since the content of this section is of more abstract nature, it may be skipped on first reading.

The notion of a locally-convex topological space is a generalization of the idea of normed space that retains the key topological properties associated with a norm (or, rather, a countable sequence of (semi-)norms). For our purpose, it is sufficient to know that this family includes all Banach spaces (such as $L_p(\mathbb{R}^d)$ for $p \geq 1$), Fréchet spaces such as $\mathcal{S}(\mathbb{R}^d)$, as well as their duals—e.g., $\mathcal{S}'(\mathbb{R}^d)$: Schwartz' space of tempered distributions.

Definition 3 (Continuous embedding)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two locally-convex topological spaces such that $\mathcal{X} \subset \mathcal{Y}$ (set inclusion). $\mathcal{X}$ is said to be continuously embedded in $\mathcal{Y}$, which is denoted by $\mathcal{X} \subseteq \mathcal{Y}$, if the inclusion map $i : \mathcal{X} \to \mathcal{Y} : x \mapsto x$ is continuous. In particular, if $\mathcal{X}$ and $\mathcal{Y}$ are two Banach spaces with respective norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Y}}$, it is equivalent to the existence of a constant C_0 such that*

$$\|x\|_{\mathcal{Y}} \leq C_0 \|x\|_{\mathcal{X}},$$

for all $x \in \mathcal{X}$.

Definition 4. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two locally-convex topological vector spaces so that $\mathcal{X}$ is continuous embedded in $\mathcal{Y}$; i.e., $\mathcal{X} \subseteq \mathcal{Y}$. We say that $\mathcal{X}$ is dense in $\mathcal{Y}$ if for any $y \in \mathcal{Y}$, there exists some sequence (x_k) in $\mathcal{X}$ such that $\lim_{k\to\infty} x_k = y$ in the topology of $\mathcal{Y}$.*

Theorem 2 (Denseness of $\mathcal{S}(\mathbb{R}^d)$)**.** *Let $\mathcal{X}$ be any locally-convex topological vector space such that $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{X} \subseteq \mathcal{S}'(\mathbb{R}^d)$ where the embedding is continuous. Then, $\mathcal{S}(\mathbb{R}^d)$ is dense in $\mathcal{X}$.*

Proof. The method is constructive: any $f \in \mathcal{X}$ can be approached as closely as desired by the sequence of functions $\varphi_k = (f * \hat{u}_k)u_k \in \mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{X}$ where (u_k) is a series of window functions in $\mathcal{S}(\mathbb{R}^d)$ such that $\lim_{k\to\infty} u_k = 1$ (e.g., $u_k(\boldsymbol{x}) = \mathrm{e}^{-(\|\boldsymbol{x}\|/k)^2}$). The enabling property for this construction is that the convolution of any tempered distribution $\big($e.g., $f \in \mathcal{X} \subseteq \mathcal{S}'(\mathbb{R}^d)\big)$ with a test function $\big($here $\hat{u}_k \in \mathcal{S}(\mathbb{R}^d)\big)$ necessarily yields a function that is infinity differentiable but still possibly of slow growth. The subsequent multiplication with u_k imposes the rapid descent property, which ensures that $\varphi_k \in \mathcal{S}(\mathbb{R}^d)$ for any $k \in \mathbb{N}^+$. As $k \to \infty$, the test function $\hat{u}_k$ converges to the Dirac impulse, which then acts as the convolution identity. The powerful aspect of the argument is that the reasoning holds for any $\mathcal{X} \subseteq \mathcal{S}'(\mathbb{R}^d)$, including the limit case $\mathcal{X} = \mathcal{S}'(\mathbb{R}^d)$. □

In the sequel, we shall often exploit this property to establish algebraic properties of functionals and operators (such boundary conditions and positivity). The practical advantage of considering test functions first is that it allows us to split sums or take limits without having to worry about technicalities. Once the desired property is established over $\mathcal{S}(\mathbb{R}^d)$, it is then readily transferred to some larger topological vector space $\mathcal{X}$ that is of interest to us.

Corollary 1. *We now consider some Banach space $\mathcal{X}$ equipped with the norm $\|\cdot\|_{\mathcal{X}}$ with the property that $\mathcal{S}(\mathbb{R}^d)$ is dense in $\mathcal{X} \subseteq \mathcal{S}'(\mathbb{R}^d)$. Then, the following holds true.*

- *The map $\varphi \mapsto \|\varphi\|_{\mathcal{X}}$ specifies a continuous functional on $\mathcal{S}(\mathbb{R}^d)$ that fullfills the defining properties of a norm; i.e., $\|\alpha\varphi\|_{\mathcal{X}} = \alpha\|\varphi\|_{\mathcal{X}}$, $\|\varphi + \phi\|_{\mathcal{X}} \le \|\varphi\|_{\mathcal{X}} + \|\phi\|_{\mathcal{X}}$ and $\|\varphi\|_{\mathcal{X}} = 0 \;\Leftrightarrow\; \varphi = 0$, for any $\alpha \in \mathbb{R}$ and $\varphi, \phi \in \mathcal{S}(\mathbb{R}^d)$. Then, the closure of $\mathcal{S}(\mathbb{R}^d)$ with respect to the norm $\|\cdot\|_{\mathcal{X}}$ is precisely the Banach space $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$.*
- *Transfer of boundary condition: Let $\phi \in \mathcal{X}'$ with $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{X} \subseteq \mathcal{S}'(\mathbb{R}^d)$. Then,*
$$\forall \varphi \in \mathcal{S}(\mathbb{R}^d), \langle \phi, \varphi\rangle = 0 \Rightarrow \forall f \in \mathcal{X}, \langle \phi, f\rangle = 0$$
- *Extension of the domain of an operator $\mathrm{T} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$. Let $\mathcal{X}$ and $\mathcal{Y}$ be two Banach space such that $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{X}, \mathcal{Y} \subseteq \mathcal{S}'(\mathbb{R}^d)$. If*
$$\|\mathrm{T}\varphi\|_{\mathcal{Y}} \le C\|\varphi\|_{\mathcal{X}}$$
for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, then the operator has a continuous extension $\mathrm{T} : \mathcal{X} \to \mathcal{Y}$.

The first and second statements are immediate consequences of the denseness of $\mathcal{S}(\mathbb{R}^d)$ in $\mathcal{X}$. The third condition ensures that the operator is bounded in the $\|\cdot\|_{\mathcal{Y}}$ norm; we then invoke the Hahn-Banach theorem to justify the extension of its domain.

Proposition 1. *Let L be a continuous linear operator $\mathcal{S}(\mathbb{R}^d) \to \mathcal{X}$ where $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ is a Banach subspace of $\mathcal{S}'(\mathbb{R}^d)$. Then:*

- *The closure of the $\mathcal{S}(\mathbb{R}^d)$ with respect to the semi-norm $\varphi \mapsto \|\mathrm{L}\varphi\|_{\mathcal{X}} \stackrel{\triangle}{=} \|\varphi\|_{\mathrm{L},\mathcal{X}}$ is a semi-normed subspace of $\mathcal{S}'(\mathbb{R}^d)$ denoted by $\mathcal{X}_{\mathrm{L,ext}}$.*

- *The (extended) null space of* L, $\mathcal{N}_{\mathrm{L}} = \{x_0 \in X_{\mathrm{L,ext}} : \mathrm{L}x_0 = 0\}$*, is a closed subspace of* $\mathcal{S}'(\mathbb{R}^d)$ *(???).*

- *The quotient space* $\mathcal{X}_{\mathrm{L,ext}}/\mathcal{N}_{\mathrm{L}}$ *equipped with the norm* $\|\cdot\|_{\mathrm{L},\mathcal{X}}$ *is a Banach space.*

- *If* $\mathcal{N}_{\mathrm{L}}$ *equipped with the norm* $\|\cdot\|_{\mathcal{N}_{\mathrm{L}}}$ *is a Banach space, then the direct sum* $(\mathcal{X}_{\mathrm{L,ext}}/\mathcal{N}_{\mathrm{L}}) \oplus \mathcal{N}_{\mathrm{L}}$ *is a Banach space that is isometrically isomorphic to* $\mathcal{X}_{\mathrm{L,ext}}$*.*

Theorem 3. *Let* $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ *be a Banach space such that* $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{X} \subseteq \mathcal{S}'(\mathbb{R}^d)$*. Then, its continuous dual* $(\mathcal{X}', \|\cdot\|_{\mathcal{X}'})$ *has the same property—i.e.,* $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{X}' \subseteq \mathcal{S}'(\mathbb{R}^d)$*—so that both spaces are dense in* $\mathcal{S}'(\mathbb{R}^d)$*.*

Proof. Due to the nuclear-Fréchet structure of $\mathcal{S}(\mathbb{R}^d)$ (see Appendix D), the continuous embedding $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{X} \subseteq \mathcal{S}'(\mathbb{R}^d)$ implies the existence of a Hilbert space $\mathcal{H} = \mathcal{S}_m(\mathbb{R}^d)$ for some $m \in \mathbb{N}$ such that $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{H} \subseteq \mathcal{X} \subseteq \mathcal{H}' \subseteq \mathcal{S}'(\mathbb{R}^d)$. This is equivalent to the existence of two constants $C'_m, C_m > 0$ such that

$$\frac{1}{C'_m}\|\varphi\|_{\mathcal{H}'} \le \|\varphi\|_{\mathcal{X}} \le C_m\|\varphi\|_{\mathcal{H}} \tag{8}$$

for all $\varphi \in \mathcal{H}$.

Since $\mathcal{S}(\mathbb{R}^d)$ is dense in $\mathcal{X}$ (by Theorem 2), we can expressed the (dual) norm of $\mathcal{X}'$ (see Definition 5) as

$$\|f\|_{\mathcal{X}'} = \sup_{\varphi \in \mathcal{S}(\mathbb{R}^d)\backslash\{0\}} \left(\frac{\langle f, \varphi\rangle}{\|\varphi\|_{\mathcal{X}}}\right)$$

for any $f \in \mathcal{X}'$.

Let us now take $f \in \mathcal{S}(\mathbb{R}^d)$. Since the space $\mathcal{H}$ and $\mathcal{H}'$ are duals of each other, we have that $|\langle f, \varphi\rangle| \le \|f\|_{\mathcal{H}}\|\varphi\|_{\mathcal{H}'}$, which, for $\varphi \neq 0$, implies that

$$\|f\|_{\mathcal{X}'} \le \frac{\|f\|_{\mathcal{H}}\|\varphi\|_{\mathcal{H}'}}{\|\varphi\|_{\mathcal{X}}} \le C'_m\|f\|_{\mathcal{H}} < +\infty.$$

Moreover, we clearly have that $\|\varphi\|_{\mathcal{X}'} = 0 \Leftrightarrow \varphi = 0$, which shows that $\|\cdot\|_{\mathcal{X}'}$ is a valid norm over $\mathcal{S}(\mathbb{R}^d)$. In other words, we can view $\mathcal{X}'$ as the completion of $\mathcal{S}(\mathbb{R}^d)$ with respect to the $\|\cdot\|_{\mathcal{X}'}$ norm.

Likewise, by interchanging the role of $\mathcal{X}'$ and $\mathcal{H}'$, we use the same argument to show that

$$\|f\|_{\mathcal{H}'} = \sup_{\varphi \in \mathcal{S}(\mathbb{R}^d)\backslash\{0\}} \left(\frac{\langle f, \varphi\rangle}{\|\varphi\|_{\mathcal{H}}}\right) \le \frac{\|f\|_{\mathcal{X}'}\|\varphi\|_{\mathcal{X}}}{\|\varphi\|_{\mathcal{H}}} \le C_m\|f\|_{\mathcal{X}'},$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$ and, by extension, for all $f \in \mathcal{X}'$. The final outcome is

$$\frac{1}{C_m}\|f\|_{\mathcal{H}'} \leq \|f\|_{\mathcal{X}'} \leq C'_m\|f\|_{\mathcal{H}},$$

which is the dual counterpart of (8). The statement on density then follows from Theorem 2. □

In view of Theorem 3, the norm for the Banach space $\mathcal{X} \supseteq \mathcal{S}(\mathbb{R}^d)$ admits the following equivalent form

$$\|f\|_{\mathcal{X}} = \sup_{\varphi \in \mathcal{S}(\mathbb{R}^d)\setminus\{0\}} \left(\frac{\langle f, \varphi\rangle}{\|\varphi\|_{\mathcal{X}'}}\right)$$

where the supremum is taken over $\mathcal{S}(\mathbb{R}^d)$, rather than $\mathcal{X}' \supseteq \mathcal{S}(\mathbb{R}^d)$. This yields a meaning to the alternative definition of $\mathcal{X}$ as

$$\mathcal{X} = \left\{v \in \mathcal{S}'(\mathbb{R}^d) : \|v\|_{\mathcal{X}} < \infty\right\}.$$

2 Reproducing kernel Hilbert spaces (RKHS)

In essence, any Hilbert space $\mathcal{H} = \{f : \mathbb{R}^d \to \mathbb{R} : \|f\|_{\mathcal{H}} = \sqrt{\langle f, f\rangle_{\mathcal{H}}} < \infty\} \subseteq \mathcal{S}'(\mathbb{R}^d)$ whose members f are "ordinary"—but, not necessarily bounded—functions on $\mathbb{R}^d$ is a RKHS and vice versa. In other words, $f(\boldsymbol{x})$ is well-defined for any $\boldsymbol{x} \in \mathbb{R}^d$ in contrast with the elements of $\mathcal{S}'(\mathbb{R}^d)$ (generalized functions) that do not necessarily have a pointwise interpretation. This will be made explicit by relating the abstract definition (Definition 6) and the properties of the reproducing kernel (Proposition 2) to the functional characteristics (continuity, rate of decay or growth) of the RKHS (Theorem 7).

2.1 Definition of RKHS

Let us start by recalling a few standard definitions from functional analysis.

Definition 5 (Dual of a Banach space)**.** *The dual of the Banach space $\mathcal{X} \supseteq \mathcal{S}(\mathbb{R}^d)$ is the vector space $\mathcal{X}' \subseteq \mathcal{S}'(\mathbb{R}^d)$ that consists of all continuous linear functional on $\mathcal{X}$. It is can be specified as the completion of $\mathcal{S}(\mathbb{R}^d)$ in the dual norm*

$$\|v\|_{\mathcal{X}'} \triangleq \sup_{u \in \mathcal{X}\setminus\{0\}} \left(\frac{\langle v, u\rangle}{\|u\|_{\mathcal{X}}}\right) \tag{9}$$

where $\langle\cdot,\cdot\rangle : \mathcal{X}' \times \mathcal{X} \to \mathbb{R}$ is the duality pairing that represents the action of the linear functional $v : u \mapsto v(u) = \langle v, u\rangle$.

To keep the notation simple, we shall write

$$\mathcal{X} = \{u \in \mathcal{S}'(\mathbb{R}^d) : \|u\|_{\mathcal{X}} < \infty\}$$
$$\mathcal{X}' = \{v \in \mathcal{S}'(\mathbb{R}^d) : \|v\|_{\mathcal{X}'} < \infty\}$$

with the implicit understanding that the rigorous specification of these spaces involves a completion/density argument (see Theorem 3 and accompanying explanations).

In particular, the dual of a Hilbert space $\mathcal{H} = \mathcal{X}$ is another Hilbert space $\mathcal{H}'$ with the following remarkable property.

Theorem 4 (Riesz representation theorem)**.** *Let $(\mathcal{H}, \mathcal{H}')$ be a dual pair of Hilbert spaces. Then, for any continuous linear functional $v \in \mathcal{H}'$, there is a unique element $v^* = \mathrm{R}\{v\} \in \mathcal{H}$ (the so-called conjugate of v) such that*

$$v(u) = \langle v^*, u\rangle_{\mathcal{H}} \tag{10}$$

for all $u \in \mathcal{H}$. Conversely, for any $v^* \in \mathcal{H}$, the linear functional $v : \mathcal{H} \to \mathbb{R}$ defined by (10) is continuous with $\|v\| = \|v\|_{\mathcal{H}'} = \|v^*\|_{\mathcal{H}} = \|\mathrm{R}v\|_{\mathcal{H}}$, and hence included in $\mathcal{H}'$. The linear isometric map $\mathrm{R} : \mathcal{H}' \to \mathcal{H}$ that associates any element $v \in \mathcal{H}'$ to its conjugate $v^* \in \mathcal{H}$ is called the Riesz map.

The existence of the Riesz map implies that $\langle v_1, v_2\rangle_{\mathcal{H}'} = \langle v_1^*, v_2^*\rangle_{\mathcal{H}} = \langle \mathrm{R}v_1, \mathrm{R}v_2\rangle_{\mathcal{H}}$ for any $v_1, v_2 \in \mathcal{H}'$. This specifies the inner product for $\mathcal{H}'$, while it also shows that the two spaces are isometrically isomorphic.

Definition 6 (RKHS)**.** *The Hilbert space $\mathcal{H} \subseteq \mathcal{S}'(\mathbb{R}^d)$ is said to be a* reproducing kernel Hilbert space *(RKHS) if the shifted Dirac impulse $\delta(\cdot - \boldsymbol{x}_0) \in \mathcal{H}'$ for any $\boldsymbol{x}_0 \in \mathbb{R}^d$.*

Let us momentarily denote the Dirac "sampling" functional $\delta(\cdot - \boldsymbol{x}_0)$ by $\delta_{\boldsymbol{x}_0}$. By the Riesz representation theorem, the RKHS condition $\delta_{\boldsymbol{x}_0} \in \mathcal{H}'$ implies the existence and uniqueness of the conjugate element $\delta^*_{\boldsymbol{x}_0} \in \mathcal{H}$ such that

$$f(\boldsymbol{x}_0) = \delta_{\boldsymbol{x}_0}(f) = \langle \delta(\cdot - \boldsymbol{x}_0), f\rangle = \langle \delta^*_{\boldsymbol{x}_0}, f\rangle_{\mathcal{H}}, \tag{11}$$

for all $f \in \mathcal{H}$ and for any $\boldsymbol{x}_0 \in \mathbb{R}^d$. This brings us to the concept of *reproducing kernel*, which is a reformulation of (11) with the change of notation $\delta^*_{\boldsymbol{x}_0}(\boldsymbol{x}) = h(\boldsymbol{x}_0, \boldsymbol{x})$.

Definition 7 (Reproducing kernel)**.** *The reproducing kernel of a RKHS on $\mathbb{R}^d$ is the function $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that*

$$\begin{aligned} &(i) \quad h(\boldsymbol{x}_0, \cdot) \in \mathcal{H} \text{ for all } \boldsymbol{x}_0 \in \mathbb{R}^d \\ &(ii) \quad f(\boldsymbol{x}_0) = \langle h(\boldsymbol{x}_0, \cdot), f\rangle_{\mathcal{H}} \text{ for all } f \in \mathcal{H} \text{ and } \boldsymbol{x}_0 \in \mathbb{R}^d. \end{aligned} \tag{12}$$

Proposition 2. *Let $\mathcal{H}$ be a RKHS on $\mathbb{R}^d$. Then, its reproducing kernel $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ has the following properties.*

1. *It is unique.*
2. $h(\boldsymbol{x}, \boldsymbol{y}) = \langle h(\boldsymbol{x}, \cdot), h(\cdot, \boldsymbol{y})\rangle_{\mathcal{H}}$
3. *Symmetry:* $h(\boldsymbol{x}, \boldsymbol{y}) = h(\boldsymbol{y}, \boldsymbol{x})$ *for all* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$
4. *Positive (semi-)definiteness.*
5. *The linear span of* $\{h(\boldsymbol{x}, \cdot), \boldsymbol{x} \in \mathbb{R}^d\}$ *is dense in* $\mathcal{H}$.
6. *Link with the Riesz map: The operator*
$$\mathrm{R} : \varphi \mapsto \mathrm{R}\{\varphi\} = \int_{\mathbb{R}^d} h(\cdot, \boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y} \tag{13}$$
is a unitary mapping $\mathcal{H}' \to \mathcal{H}$ *with the property that* $\langle u, \mathrm{R}u\rangle = \|u\|^2_{\mathcal{H}'}$ *for all* $u \in \mathcal{H}'$ *(the dual space of* $\mathcal{H}$*).*
7. *Invertibility: The operator specified by* (13) *admits a unique inverse* $\mathrm{R}^{-1} : \mathcal{H} \to \mathcal{H}'$ *with the property that* $\langle \mathrm{R}^{-1}f, f\rangle = \|f\|^2_{\mathcal{H}}$ *for all* $f \in \mathcal{H}$.

Proof. The reproducing kernel is obtained from (11) by setting $h(\boldsymbol{x}_0, \boldsymbol{y}) = \delta^*_{\boldsymbol{x}_0}(\boldsymbol{y})$; it is unique and included in $\mathcal{H}$ as a consequence of the Riesz representation theorem.

Property 2 is derived by applying the reproduction formula (12) to the kernel itself: $\langle h(\boldsymbol{x}, \cdot), h(\cdot, \boldsymbol{y})\rangle_{\mathcal{H}} = h(\boldsymbol{x}, \boldsymbol{y})$. The symmetry follows from the symmetry of the inner product: $h(\boldsymbol{x}, \boldsymbol{y}) = \langle h(\boldsymbol{x}, \cdot), h(\boldsymbol{y}, \cdot)\rangle_{\mathcal{H}} = \langle h(\boldsymbol{y}, \cdot), h(\boldsymbol{x}, \cdot)\rangle_{\mathcal{H}} = h(\boldsymbol{y}, \boldsymbol{x})$.

The positive definiteness is a consequence of Property 3 where $h(\boldsymbol{x}, \boldsymbol{y})$ is expressed as an inner product (see also Theorem 1).

To establish Property 5, we consider a function $g \in \mathcal{H}$ that is orthogonal to the linear span of $\{h(\boldsymbol{x}, \cdot)\}_{\boldsymbol{x}\in\mathbb{R}^d}$. Then, $\langle g, h(\boldsymbol{x}, \cdot)\rangle = 0$ for every $\boldsymbol{x} \in \mathbb{R}^d$, which, by the reproducing kernel property, is equivalent to $g = 0$.

As for Properties 6 and 7, we refer once more to Theorem 4, which guarantees the existence and unicity of an invertible pair of operators $\mathrm{R} : \mathcal{H}' \to \mathcal{H}$ and $\mathrm{R}^{-1} : \mathcal{H} \to \mathcal{H}'$ (the Riesz maps of $\mathcal{H}'$ and $\mathcal{H}$, respectively) such that

$$\langle v, u\rangle = \langle v^*, u\rangle_{\mathcal{H}} = \langle \mathrm{R}v, u\rangle_{\mathcal{H}} = \langle v, \mathrm{R}^{-1}u\rangle_{\mathcal{H}'} = \langle v, u^*\rangle_{\mathcal{H}} \tag{14}$$

for all $v \in \mathcal{H}'$ and $u \in \mathcal{H}$. Next, we rephrase the definition of the reproducing kernel as

$$\delta^*_{\boldsymbol{y}}(\boldsymbol{x}) = \mathrm{R}\{\delta(\cdot - \boldsymbol{y})\}(\boldsymbol{x}) = h(\boldsymbol{y}, \boldsymbol{x}) = h(\boldsymbol{x}, \boldsymbol{y}),$$

which shows that $h(\cdot, \cdot)$ is the generalized impulse reponse of the Riesz map R. This is precisely what is indicated by Eq. (6) (see (3) and accompanying explanations). Finally, we set $v = \mathrm{R}^{-1}u$ (resp., $u = \mathrm{R}v$) in (14), which yields

$$\langle \mathrm{R}^{-1}u, u\rangle = \langle u, u\rangle_{\mathcal{H}} = \|u\|^2_{\mathcal{H}}$$
$$\langle v, \mathrm{R}v\rangle = \langle v, v\rangle_{\mathcal{H}'} = \|v\|^2_{\mathcal{H}'}.$$

□

There is a striking parallel in (11) between the central equation (the defining property of the shifted Dirac impulse) and the right-hand side, which specifies $\delta^*_{\boldsymbol{x}_0}$. The crucial difference, of course, is that $\delta^*_{\boldsymbol{x}_0} = h(\boldsymbol{x}_0, \cdot)$ is an ordinary function whose action is tied to the inner product of $\mathcal{H}$, whereas $\delta(\cdot - \boldsymbol{x}_0)$ is a tempered distribution whose definition involves the duality product for $\big(\mathcal{S}'(\mathbb{R}^d), \mathcal{S}(\mathbb{R}^d)\big)$. While it would be tempting to interpret $\delta(\boldsymbol{x} - \boldsymbol{y})$ as the reproducing kernel for the canonical Hilbert space $L_2(\mathbb{R}^d)$ for which $f^* = f$ (because $L_2(\mathbb{R}^d)$ is its own dual), there are two reasons why the RKHS property cannot apply there: i) most functions $f \in L_2(\mathbb{R}^d)$ are not continuous which makes the sampling operation ill-defined, and, (ii) $\delta(\boldsymbol{x}_0 - \cdot) \notin L_2(\mathbb{R}^d)$, which would contradict the first requirement in Definition 7.

We shall therefore focus our attention on the Hilbert spaces whose members are continuous functions.

Supplementary material: For completeness, we also list some higher-level topological properties of reproducing kernel Hilbert spaces that are given here without proof; see for the details.

Theorem 5. *Let $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the reproducing kernel of some RKHS $\mathcal{H}$. Then, the following properties hold.*

1. *Any converging (or Cauchy) sequence of functions (f_n) in $\mathcal{H}$ also converges pointwise to the same limit; i.e.,*

$$\lim_{n\to\infty} \|f_n - f\|_{\mathcal{H}} = 0 \quad \Rightarrow \quad \lim_{n\to\infty} |f_n(\boldsymbol{x}) - f(\boldsymbol{x})| = 0 \text{ for every } \boldsymbol{x} \in \mathbb{R}^d.$$

2. *The set $\mathcal{H}_{\mathrm{pre}} = \big\{ \sum_{n=1}^N a_n h(\cdot, \boldsymbol{y}_n) : N \in \mathbb{N}, a_n \in \mathbb{R}, \boldsymbol{y}_n \in \mathbb{R}^d \big\}$ is dense in $\mathcal{H}$. In other words, we can represent any function $f \in \mathcal{H}$ as closely as desired by a linear combination of the form $\tilde{f} = \sum_{n=1}^N a_n h(\cdot, \boldsymbol{y}_n)$ with a finite number N of terms.*

2.2 Decay and continuity properties

We shall monitor the algebraic rate of decay/growth of functions of the variable $\boldsymbol{x} \in \mathbb{R}^d$ via their inclusion in "weighted" function spaces. To that end, we first define

$$L_{\infty,\alpha}(\mathbb{R}^d) = \Big\{f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{\infty,\alpha} < +\infty\Big\}$$

with $\alpha \in \mathbb{R}$, where

$$\|f\|_{\infty,\alpha} = \operatorname*{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}^d} \left(|f(\boldsymbol{x})|(1 + \|\boldsymbol{x}\|)^{\alpha}\right).$$

Note that

$$f \in L_{\infty,\alpha}(\mathbb{R}^d) \quad \Leftrightarrow \quad |f(\boldsymbol{x})| \leq \frac{\|f\|_{\infty,\alpha}}{(1 + \|\boldsymbol{x}\|)^{\alpha}} \quad \textit{almost everywhere}, \tag{15}$$

meaning that $f(\boldsymbol{x})$ decays at least (or growth at most) as $1/\|\boldsymbol{x}\|^{\alpha}$ at infinity. To constrain the setting to "classical" functions that are well defined pointwise, we introduce the function space

$$C_{\mathrm{b},\alpha}(\mathbb{R}^d) = \Big\{f : \mathbb{R}^d \to \mathbb{R} \text{ continuous and s.t. } \|f\|_{\infty,\alpha} < +\infty\Big\}.$$

The main difference there is that the upper bound on the rhs of (15) becomes valid *everywhere*; this allows for the substitution of the "essential supremum" in the definition of the weighted L_∞ norm by the simpler supremum (i.e., $\sup_{\boldsymbol{x}\in\mathbb{R}^d}$).

It can be checked that $C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ equipped with the $\|\cdot\|_{\infty,\alpha}$ norm is complete and hence a Banach space. In particular, $C_{\mathrm{b},0}(\mathbb{R}^d) = C_{\mathrm{b}}(\mathbb{R}^d)$, which is the classical space of bounded continuous functions. Clearly, $\|f\|_{\infty,\alpha_1} \leq \|f\|_{\infty,\alpha_2}$ for any $f \in C_{\mathrm{b},\alpha_2}(\mathbb{R}^d)$ with $\alpha_2 \geq \alpha_1$, which implies that $C_{\mathrm{b},\alpha_2}(\mathbb{R}^d)$ is continuously embedded in $C_{\mathrm{b},\alpha_1}(\mathbb{R}^d)$ (see Definition 3 below) and, a fortiori, in $L_{\infty,\alpha_1}(\mathbb{R}^d)$. This functional embedding is summarized as

$$\mathcal{S}(\mathbb{R}^d) \subseteq \quad C_{\mathrm{b},\alpha+\beta}(\mathbb{R}^d) \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d) \subseteq L_{\infty,\alpha}(\mathbb{R}^d) \quad \subseteq \mathcal{S}'(\mathbb{R}^d),$$

for any $\beta \geq 0$.

A standard result from the theory of tempered distributions is that, for any continuous function $f \in \mathcal{S}'(\mathbb{R}^d)$, there exists some (critical) exponent $\alpha_0 \in \mathbb{R}$ such that $f \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ for all $\alpha \leq \alpha_0$. This exponent qualifies the rate of algebraic decay of f. The function is said to be of *slow growth* if

$\alpha_0 \le 0$. On the contrary, it has a *rapid decay* if the inclusion holds for all $\alpha \in \mathbb{R}$.

The same considerations apply for continuous kernel functions $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ which are then described in terms of the cross-product extension of these spaces; namely, $\mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$, $L_{\infty,\alpha}(\mathbb{R}^d \times \mathbb{R}^d)$ and $C_{\mathrm{b},\alpha}(\mathbb{R}^d \times \mathbb{R}^d)$ with

$$\|h(\cdot,\cdot)\|_{\infty,\alpha} = \sup_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d} |h(\boldsymbol{x},\boldsymbol{y})|(1+\|\boldsymbol{x}\|)^{\alpha}(1+\|\boldsymbol{y}\|)^{\alpha}.$$

To emphasize the central role of continuity in the characterization of RKHS, let us consider some Hilbert space $\mathcal{H}$ that is composed of continuous functions with an algebraic rate of decay no worse than α; that is, $\mathcal{H} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ for some $\alpha \in \mathbb{R}$. This embedding implies the existence of a constant $C > 0$ such that

$$\|f\|_{\infty,\alpha} \le C\|f\|_{\mathcal{H}}, \text{ for all } f \in \mathcal{H}.$$

The condition $f \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ then gives that

$$|f(\boldsymbol{x}_0)| = |\delta_{\boldsymbol{x}_0}(f)| \le (1+\|\boldsymbol{x}_0\|)^{-\alpha}\|f\|_{\infty,\alpha} \le C_{\boldsymbol{x}_0}\|f\|_{\mathcal{H}},$$

which shows that $\delta_{\boldsymbol{x}_0} : \mathcal{H} \to \mathbb{R}$ is continuous or, equivalently, $\delta(\cdot - \boldsymbol{x}_0) \in \mathcal{H}'$ for any $\boldsymbol{x}_0 \in \mathbb{R}^d$. This proves that continuity together with some form of boundedness is sufficient to ensure the reproducing kernel property (see Definition 6). What is more remarkable is that the implication also goes the other way around in the sense that one can tightly control the continuity and rate of decay of $f \in \mathcal{H}$ based on the properties of the reproducing kernel.

The first element of this equivalence is the transfer of continuity between h and the members of $\mathcal{H}$, which is covered by the following theorem, the proof of which can be found in [**?**, Theorem 17, p. 34].

Theorem 6 (RKHS of continuous functions)**.** *Let $\mathcal{H}$ be a RKHS of functions on $\mathbb{R}^d$ with reproducing kernel $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$. Then, any element of $\mathcal{H}$ is continuous if and only if*

1. *the map $\boldsymbol{x} \mapsto h(\boldsymbol{x},\boldsymbol{y})$ is continuous for all $\boldsymbol{y} \in \mathbb{R}^d$;*
2. *for every $\boldsymbol{x} \in \mathbb{R}^d$, there exist $\epsilon > 0$ such that the function $\boldsymbol{y} \mapsto h(\boldsymbol{y},\boldsymbol{y})$ is bounded on the open ball $B(\boldsymbol{x},\epsilon) = \{\boldsymbol{y} \in \mathbb{R}^d : \|\boldsymbol{y}-\boldsymbol{x}\| < \epsilon\}$.*

We now show that in the present setting where $\mathcal{H} \subseteq \mathcal{S}'(\mathbb{R}^d)$, the local boundedness constraint (Condition 2) in Theorem 6 can be substituted by a simpler and more intuitive slow-growth constraint; that is, the existence of a critical rate of decay/growth $\alpha \in \mathbb{R}$ such that $h \in L_{\infty,\alpha}(\mathbb{R}^d \times \mathbb{R}^d)$.

Theorem 7 (Characterization of RKHS in $C_{\mathrm{b},\alpha}(\mathbb{R}^d)$)**.** *A bivariate function $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is the reproducing kernel of a RKHS $\mathcal{H} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ with $\alpha \in \mathbb{R}$ if and only if it is positive-definite, separately continuous in each variable, and such that $h(\cdot,\cdot) \in L_{\infty,\alpha}(\mathbb{R}^d \times \mathbb{R}^d)$. In particular, this implies that*

1. $h(\boldsymbol{x}_0, \cdot) \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ *for any* $\boldsymbol{x}_0 \in \mathbb{R}^d$
2. $\|h(\cdot,\cdot)\|_{\infty,\alpha} = \sup_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d} |h(\boldsymbol{x},\boldsymbol{y})|\ (1+\|\boldsymbol{x}\|)^\alpha(1+\|\boldsymbol{y}\|)^\alpha < \infty$
3. $A_{\alpha,h} = \sup_{\boldsymbol{x}\in\mathbb{R}^d} h(\boldsymbol{x},\boldsymbol{x})\ (1+\|\boldsymbol{x}\|)^{2\alpha} < \infty$

with Conditions 2 and 3 being equivalent.

An interesting outcome of Theorem 7 is that Condition 1—the minimalistic choice dictated by Definition 7—is not sufficient on its own. The natural extension is $h(\cdot,\cdot) \in C_{\mathrm{b},\alpha}(\mathbb{R}^d \times \mathbb{R}^d)$, which is sufficient for the inclusion $\mathcal{H} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$, but slightly too conservative. Indeed, the joint continuity of h over $\mathbb{R}^d \times \mathbb{R}^d$ is stronger than the separate continuity—a topic that is further developed in Proposition 3.

Proof of Theorem 7. We recall that the positive-definiteness of h is equivalent to the RKHS property (see Theorem 1). Likewise, the necessity of Condition 1 (which implies the continuity of h in each argument) for $\mathcal{H} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ is obvious since $h(\boldsymbol{x}_0,\cdot) \in \mathcal{H}$ from the definition of a reproducing kernel. The remainder of the proof is divided in three parts.

Part I: $\mathcal{H} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d) \Rightarrow$ Condition 3.
By applying the reproducing property twice, we get

$$\begin{aligned} |f(\boldsymbol{x})| = |\langle h(\boldsymbol{x},\cdot), f\rangle_{\mathcal{H}}| &\le \sqrt{\langle h(\boldsymbol{x},\cdot), h(\cdot,\boldsymbol{x})\rangle_{\mathcal{H}}}\ \|f\|_{\mathcal{H}} \qquad \text{(Cauchy-Schwarz)} \\ &\le \sqrt{h(\boldsymbol{x},\boldsymbol{x})}\ \|f\|_{\mathcal{H}}, \end{aligned}$$

from which we deduce that

$$\|f\|_{\infty,\alpha} = \sup_{\boldsymbol{x}\in\mathbb{R}^d} (1+\|\boldsymbol{x}\|)^\alpha |f(\boldsymbol{x})| \le C_0\ \|f\|_{\mathcal{H}} \tag{16}$$

where

$$C_0 = \sup_{\boldsymbol{x}\in\mathbb{R}^d} (1+\|\boldsymbol{x}\|)^\alpha \sqrt{h(\boldsymbol{x},\boldsymbol{x})}.$$

By defining $\boldsymbol{x}_0 = \arg\sup_{\boldsymbol{x}\in\mathbb{R}^d}(1+\|\boldsymbol{x}\|)^\alpha \sqrt{h(\boldsymbol{x},\boldsymbol{x})}$ and taking $f = h(\cdot,\boldsymbol{x}_0)$, we then verify that the inequality (16) is sharp. This allows us to identify

$C_0 = \sqrt{A_{\alpha,h}}$ as the operator norm of the identity map $i : \mathcal{H} \to C_{\mathrm{b},\alpha}(\mathbb{R}^d)$, which is bounded by hypothesis.

Part II: Equivalence of Conditions 2 and 3.
The connection between reproducing kernels and inner products yields

$$|h(\boldsymbol{x}, \boldsymbol{y})|^2 \leq h(\boldsymbol{x}, \boldsymbol{x})h(\boldsymbol{y}, \boldsymbol{y}),$$

which is the kernel equivalent of the Cauchy-Schwarz inequality. Accordingly,

$$\sup_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d} |(1 + \|\boldsymbol{x}\|)^\alpha h(\boldsymbol{x}, \boldsymbol{y})(1 + \|\boldsymbol{y}\|)^\alpha|^2 \leq \sup_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d} (1 + \|\boldsymbol{x}\|)^{2\alpha} h(\boldsymbol{x}, \boldsymbol{x})(1 + \|\boldsymbol{y}\|)^{2\alpha} h(\boldsymbol{y}, \boldsymbol{y}) = A_{h,\alpha}^2,$$

which implies that

$$\|h(\cdot, \cdot)\|_{\infty,\alpha} \leq A_{h,\alpha}.$$

Conversely,

$$(1 + \|\boldsymbol{x}\|)^\alpha h(\boldsymbol{x}, \boldsymbol{y})(1 + \|\boldsymbol{y}\|)^\alpha \leq \|h(\cdot, \cdot)\|_{\infty,\alpha}$$

so that

$$A_{h,\alpha} = \sup_{\boldsymbol{x}\in\mathbb{R}^d} (1 + \|\boldsymbol{x}\|)^\alpha h(\boldsymbol{x}, \boldsymbol{x})(1 + \|\boldsymbol{x}\|)^\alpha \leq \|h(\cdot, \cdot)\|_{\infty,\alpha},$$

from which we deduce that $A_{h,\alpha} = \|h(\cdot, \cdot)\|_{\infty,\alpha}$.

Part III: Condition 3 and continuity of $h(\boldsymbol{x}_0, \cdot) \Rightarrow \mathcal{H} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$.
The boundedness of $\|f\|_{\infty,\alpha}$ for all $f \in \mathcal{H}$ follows directly from (16) with $C_0 = \sqrt{A_{h,\alpha}}$. We then prove that $\mathcal{H} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d) \subseteq L_{\infty,\alpha}(\mathbb{R}^d)$ by invoking Theorem 6, which ensures that all the members of $\mathcal{H}$ are continuous functions. The two required hypotheses are: (i) the separate continuity of $h(\boldsymbol{x}, \boldsymbol{y})$ in $\boldsymbol{x}$ and $\boldsymbol{y}$, and, (ii) the local boundedness of h along its diagonal, which is met, thanks to Condition 3. Specifically, for any $\epsilon > 0$, we have that

$$h(\boldsymbol{y}, \boldsymbol{y}) \leq A_{\alpha,h}(1 + \|\boldsymbol{y}\|)^{-2\alpha} \leq A_{\alpha,h}(1 + \epsilon + \|\boldsymbol{x}\|)^{\min(0,-2\alpha)} = M_{\boldsymbol{x}} < \infty$$

for all $\boldsymbol{y} \in B(\boldsymbol{x}, \epsilon)$. □

All the Hilbert spaces $\mathcal{H}$ that will be considered in the sequel are implicitly assumed to meet this minimal requirement for a RKHS; i.e., the existence of a critical exponent $\alpha_0 \in \mathbb{R}$ such that $\mathcal{H} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ for all $\alpha \leq \alpha_0$, the classical scenario being $\alpha_0 = 0$ (continuity and boundedness).

The intuitive explanation for the identity $A_{h,\alpha} = \|h(\cdot,\cdot)\|_{\infty,\alpha}$ is that a positive-definite matrix is dominated on the diagonal (by the Cauchy-Schwarz inequality) so that the max norm of the matrix is equal to the max norm of its diagonal.

The Cauchy-Schwarz inequality also helps us get a clearer understanding of the transfer of continuity from h to $\mathcal{H}$. The relevant estimate there is

$$\begin{aligned}|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)|^2 &= |\langle h(\boldsymbol{x},\cdot) - h(\boldsymbol{x}_0,\cdot), f\rangle_{\mathcal{H}}|^2 \\ &\le \|h(\boldsymbol{x},\cdot) - h(\boldsymbol{x}_0,\cdot)\|_{\mathcal{H}}^2 \, \|f\|_{\mathcal{H}}^2 \qquad \text{(Cauchy-Schwarz)} \\ &= \big(h(\boldsymbol{x},\boldsymbol{x}) + h(\boldsymbol{x}_0,\boldsymbol{x}_0) - 2h(\boldsymbol{x},\boldsymbol{x}_0)\big) \, \|f\|_{\mathcal{H}}^2. \end{aligned} \tag{17}$$

This implies that one can achieve the continuity of any $f \in \mathcal{H}$ by imposing that

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} h(\boldsymbol{x},\boldsymbol{x}) + h(\boldsymbol{x}_0,\boldsymbol{x}_0) - 2h(\boldsymbol{x},\boldsymbol{x}_0) = 0 \tag{18}$$

for all $\boldsymbol{x}_0 \in \mathbb{R}^d$. This latter condition amounts to the continuity of $h(\boldsymbol{x},\boldsymbol{y})$ along the diagonal—i.e., in the neighbourhood of the points $(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{x}_0,\boldsymbol{x}_0)$ for $\boldsymbol{x}_0 \in \mathbb{R}^d$—which is not the same as the (separate) continuity of $\boldsymbol{x} \mapsto h(\boldsymbol{x},\boldsymbol{y})$ and $\boldsymbol{y} \mapsto h(\boldsymbol{x},\boldsymbol{y})$ required in Theorem 7. There is no contradiction, however, because the former condition implies the latter, as stated in Proposition 3 below. In fact, (18) is equivalent to the continuity of h over its whole domain $\mathbb{R}^d \times \mathbb{R}^d$. This also means that the combination of (18) and the boundedness requirement $h \in L_{\infty,\alpha}(\mathbb{R}^d \times \mathbb{R}^d)$ in Theorem 7 is equivalent to the simpler-looking condition $h \in C_{\mathrm{b},\alpha}(\mathbb{R}^d \times \mathbb{R}^d)$

Proposition 3. *Let $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ be a positive-definite kernel. Then, the following conditions are equivalent:*

1. *continuity of h over $\mathbb{R}^d \times \mathbb{R}^d$:*

$$\lim_{(\boldsymbol{x},\boldsymbol{y})\to(\boldsymbol{x}_0,\boldsymbol{y}_0)} |h(\boldsymbol{x},\boldsymbol{y}) - h(\boldsymbol{x}_0,\boldsymbol{y}_0)| = 0$$

2. *(joint) continuity of h along the diagonal:*

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} |h(\boldsymbol{x},\boldsymbol{x}_0) - h(\boldsymbol{x}_0,\boldsymbol{x}_0)| = 0$$
$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} |h(\boldsymbol{x},\boldsymbol{x}) - h(\boldsymbol{x}_0,\boldsymbol{x}_0)| = 0$$

3. *continuity of h in the norm across rows or columns*

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \|h(\boldsymbol{x},\cdot) - h(\boldsymbol{x}_0,\cdot)\|_{\mathcal{H}} = \lim_{\boldsymbol{y}\to\boldsymbol{y}_0} \|h(\cdot,\boldsymbol{y}) - h(\cdot,\boldsymbol{y}_0)\|_{\mathcal{H}} = 0$$

for all $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathbb{R}^d$.

Proof. First, we observe that the positive-definiteness of h implies the symmetry of the kernel (see Appendix A). It is also known from Theorem 1 that h uniquely specifies a Hilbert space $\mathcal{H}$ with the property that $h(\boldsymbol{x}, \boldsymbol{y}) = \langle h(\boldsymbol{x}, \cdot), h(\cdot, \boldsymbol{y}) \rangle_{\mathcal{H}}$. The latter implies that

$$\|h(\boldsymbol{x}, \cdot) - h(\boldsymbol{x}_0, \cdot)\|_{\mathcal{H}}^2 = h(\boldsymbol{x}, \boldsymbol{x}) + h(\boldsymbol{x}_0, \boldsymbol{x}_0) - 2h(\boldsymbol{x}, \boldsymbol{x}_0),$$

which shows the equivalence between Properties 2 and 3. By applying (17) to $f(\boldsymbol{x}) = h(\boldsymbol{x}, \boldsymbol{y})$ with $\boldsymbol{y}$ fixed and by taking the squareroot, we find that

$$|h(\boldsymbol{x}, \boldsymbol{y}) - h(\boldsymbol{x}_0, \boldsymbol{y})| \le \sqrt{h(\boldsymbol{y}, \boldsymbol{y})} \big[h(\boldsymbol{x}, \boldsymbol{x}) + h(\boldsymbol{x}_0, \boldsymbol{x}_0) - 2h(\boldsymbol{x}, \boldsymbol{x}_0)\big]^{\frac{1}{2}},$$

which proves that Property 2 implies the separate continuity of h in each variable. By using this inequality twice, we get

$$\begin{aligned} |h(\boldsymbol{x}, \boldsymbol{y}) - h(\boldsymbol{x}_0, \boldsymbol{y}_0)| &\le |h(\boldsymbol{x}, \boldsymbol{y}) - h(\boldsymbol{x}, \boldsymbol{y}_0)| + |h(\boldsymbol{x}, \boldsymbol{y}_0) - h(\boldsymbol{x}_0, \boldsymbol{y}_0)| \\ &\le \sqrt{h(\boldsymbol{x}, \boldsymbol{x})} \big[h(\boldsymbol{y}, \boldsymbol{y}) + h(\boldsymbol{y}_0, \boldsymbol{y}_0) - 2h(\boldsymbol{y}, \boldsymbol{y}_0)\big]^{\frac{1}{2}} \\ &\quad + \sqrt{h(\boldsymbol{y}_0, \boldsymbol{y}_0)} \big[h(\boldsymbol{x}, \boldsymbol{x}) + h(\boldsymbol{x}_0, \boldsymbol{x}_0) - 2h(\boldsymbol{x}, \boldsymbol{x}_0)\big]^{\frac{1}{2}}, \end{aligned}$$

which proves that $h(\boldsymbol{x}, \boldsymbol{y})$ tends to $h(\boldsymbol{x}_0, \boldsymbol{y}_0)$ as $(\boldsymbol{x}, \boldsymbol{y}) \to (\boldsymbol{x}_0, \boldsymbol{y}_0)$.

□

While we have already pointed out that $L_2(\mathbb{R}^d)$ is not an RKHS, there is a simple generative mechanism for turning it into one by applying a reversible smoothing operator—typically, some kind of integrator—to it. We now provide a sufficient condition on the generalized impulse response of the operator for controlling the rate of decay/growth of the output.

Proposition 4. *Let* $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *be a kernel such that* $g(\boldsymbol{x}, \cdot) \in L_2(\mathbb{R}^d)$ *for any fixed* $\boldsymbol{x} \in \mathbb{R}^d$. *Then, the output of the corresponding linear operator* $\mathrm{G} : w \mapsto f = \int_{\mathbb{R}^d} g(\cdot, \boldsymbol{y}) w(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}$ *is well defined pointwise for any* $w \in L_2(\mathbb{R}^d)$. *If, in addition, there is some* $\alpha \in \mathbb{R}$ *such that*

$$\sup_{\boldsymbol{x} \in \mathbb{R}^d} (1 + \|\boldsymbol{x}\|)^\alpha \|g(\boldsymbol{x}, \cdot)\|_{L_2(\mathbb{R}^d)} < \infty, \tag{19}$$

then G *is bounded from* $L_2(\mathbb{R}^d) \to L_{\infty,\alpha}(\mathbb{R}^d)$.

Proof. By invoking the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}|f(\boldsymbol{x})| &= \left|\int_{\mathbb{R}^d} g(\boldsymbol{x},\boldsymbol{y})w(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\right| \\ &\le \int_{\mathbb{R}^d} \big|g(\boldsymbol{x},\boldsymbol{y})w(\boldsymbol{y})\big|\mathrm{d}\boldsymbol{y} \le \|g(\boldsymbol{x},\cdot)\|_{L_2(\mathbb{R}^d)}\,\|w\|_{L_2(\mathbb{R}^d)}.\end{aligned}$$

This shows that $f(\boldsymbol{x})$ is well-defined pointwise provided that $\|g(\boldsymbol{x},\cdot)\|_{L_2(\mathbb{R}^d)} < \infty$. Based on this estimate, we also get

$$\begin{aligned}\|\mathrm{G}\{w\}\|_{\infty,\alpha} &= \operatorname*{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}^d} |f(\boldsymbol{x})|(1+\|\boldsymbol{x}\|)^\alpha \\ &\le \left(\sup_{\boldsymbol{x}\in\mathbb{R}^d}(1+\|\boldsymbol{x}\|)^\alpha\|g(\boldsymbol{x},\cdot)\|_{L_2(\mathbb{R}^d)}\right)\|w\|_{L_2(\mathbb{R}^d)},\end{aligned}$$

which proves that (72) implies the boundedness of $\mathrm{G} : L_2(\mathbb{R}^d) \to L_{\infty,\alpha}(\mathbb{R}^d)$. □

2.3 RKHS: the simplified finite-dimensional story

To get a hands-on understanding of RKHS, a helpful exercise is to transpose the concept to $\mathbb{R}^N$, the standard vector space of linear algebra. We recall that $\mathbb{R}^N$ is endowed with the Euclidean inner product:

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{n=1}^N x_n y_n,$$

for any pair of vector $\mathbf{x} = (x_1, \dots, x_N)$, $\mathbf{y} = (y_1, \dots, y_N) \in \mathbb{R}^N$. We can also interpret $\mathbf{x}$ as a linear functional acting on the vector $\mathbf{y}$, meaning that $\mathbb{R}^N$ coincides with its own dual (i.e., $(\mathbb{R}^N)' = \mathbb{R}^N$). This duality pairing $(\mathbb{R}^N)' \times \mathbb{R}^N \to \mathbb{R}^N : (\mathbf{x}, \mathbf{y}) \mapsto \langle \mathbf{x}, \mathbf{y}\rangle$ is continuous and controlled by the Cauchy-Schwarz inequality

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \le \|\mathbf{x}\|_2\|\mathbf{y}\|_2$$

with $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}$.

Linear algebra is founded on the property that every continuous linear operator $\mathrm{G} : \mathbb{R}^N \to \mathbb{R}^N$ can be represented by a square matrix $\mathbf{G} \in \mathbb{R}^{N\times N}$ whose entries are denoted by $[\mathbf{G}]_{m,n} = G[m, n]$. (This is the finite-dimensional equivalent of Schwartz' kernel theorem.) Specifically, we have that $\mathrm{G}: \mathbf{x} \mapsto \mathbf{y} = \mathbf{G}\mathbf{x}$ with

$$y_m = \sum_{n=1}^{N} G[m,n] x_n = \langle G[m,\cdot], \mathbf{x} \rangle$$

where the array $G[\cdot,\cdot]$ constitutes the "discrete" kernel of the operator G.

The finite-dimensional equivalent of a reproducing kernel is a *symmetric positive-definite matrix* $\mathbf{R} \in \mathbb{R}^{N \times N}$, which is such that $\langle \mathbf{R}\mathbf{x}, \mathbf{x} \rangle \geq 0$ for all $\mathbf{x} \in \mathbb{R}^N$ (see Appendix A). The defining property of such matrices is that their eigenvalues are non-negative: $\lambda_1 \geq \cdots \geq \lambda_N \geq 0$. The corresponding eigenvectors are denoted by $\{\mathbf{u}_n\}_{n=1}^{N}$ and are such that

$$\mathbf{R}\mathbf{u}_n = \lambda_n \mathbf{u}_n \quad \text{with} \quad \|\mathbf{u}_n\|_2 = 1. \tag{20}$$

They specify the orthonormal transformation matrix $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \ldots \ \mathbf{u}_N]$. This results in the eigen-decomposition of $\mathbf{R}$ as

$$\mathbf{R} = \sum_{n=1}^{N} \lambda_n \mathbf{u}_n \mathbf{u}_n^T = \mathbf{U} \operatorname{diag}(\lambda_1, \ldots, \lambda_N)\ \mathbf{U}^T \tag{21}$$

which comes as a direct consequence of (20) and the orthonormality property $\mathbf{U}^T\mathbf{U} = \mathbf{I}$.

The rank of the matrix $\mathbf{R}$ is given the number N' of its non-zero eigenvalues. Clearly, the inverse matrix $\mathbf{R}^{-1}$ is well defined only when $\mathbf{R}$ is of full rank; that is, when $\mathbf{R}$ is *strictly* positive-definite. Otherwise, when $N' < N$, we need to consider the generalized (Moore-Penrose) inverse of $\mathbf{R}$,

$$\mathbf{R}^{\dagger} = \mathbf{U} \operatorname{diag}(1/\lambda_1, \ldots, 1/\lambda_{N'}, 0, \ldots, 0)\ \mathbf{U}^T, \tag{22}$$

which satisfies the pseudo-inverse property $\mathbf{R}\mathbf{R}^{\dagger}\mathbf{R} = \mathbf{R}$. We also note that $\mathbf{R}^{\dagger} = \mathbf{R}^{-1}$ when $N = N'$.

Proposition 5. *Let $\mathbf{R} \in \mathbb{R}^{N \times N}$ be a positive-definite matrix of rank $N' \leq N$. Then, the RKHS induced by $\mathbf{R}$ is the space $\mathcal{H} \subseteq \mathbb{R}^N$ spanned by its primary[1] eigenvectors $\{\mathbf{u}_n\}_{n=1}^{N'}$ equipped with the inner product*

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{H}} = \mathbf{x}^T \mathbf{R}^{\dagger} \mathbf{y} = \langle \mathbf{x}, \mathbf{R}^{\dagger} \mathbf{y} \rangle.$$

Proof. By restricting the eigen-decomposition of $\mathbf{R}$ to its primary part

$$\mathbf{R} = \sum_{n=1}^{N'} \lambda_n \mathbf{u}_n \mathbf{u}_n^T = [\mathbf{r}_1 \ \ldots \ \mathbf{r}_N],$$

[1]The term primary refers to the components associated with non-zero eigenvalues.

we find that $\mathbf{r}_m = \sum_{n=1}^{N'} y_n \mathbf{u}_n$ with $y_n = \lambda_n [\mathbf{u}_n]_m$, which shows that the column vectors of $\mathbf{R}$, $\mathbf{r}_m$, are included in $\mathcal{H} = \mathrm{span}\{\mathbf{u}_n\}_{n=1}^{N'}$. Consequently, we have

$$\text{Property } (i): \quad \mathbf{r}_m = R[\cdot, m] \in \mathcal{H}, \quad (m = 1, \dots, N)$$

(or, equivalently, $\mathbf{r}_m^T = R[m, \cdot]$ since the matrix $\mathbf{R}$ is symmetric), which is the first requirement for a finite-dimensional reproducing kernel $R : E \times E \to \mathbb{R}$ with $E = \{1, \dots, N\}$, in direct analogy with Definition 7.

Next, we use the explicit form of $\mathbf{R}^\dagger$ in (23) to calculate

$$\mathbf{u}_n^* = \mathbf{R}^\dagger \mathbf{u}_n = \begin{cases} \mathbf{u}_n / \lambda_n, & \text{if } n \le N' \\ \mathbf{0} & \text{otherwise,} \end{cases} \tag{23}$$

which, when combined with (20), yields

$$\mathbf{R}\mathbf{R}^\dagger \mathbf{u}_n = \begin{cases} \mathbf{u}_n, & \text{if } n \le N' \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Since $\{\mathbf{u}_n\}_{n=1}^{N}$ is an orthonormal basis of $\mathbb{R}^N$, this allows us to identify $\mathbf{R}\mathbf{R}^\dagger$ as the orthogonal projector $\mathbb{R}^N \to \mathcal{H}$. In fact, we have the direct sum decomposition $\mathbb{R}^N = \mathcal{H} \oplus \mathcal{N}$ where $\mathcal{N} = \mathrm{span}\{\mathbf{u}_n\}_{n=N'+1}^{N}$, meaning that every vector $\mathbf{x} \in \mathbb{R}^N$ has a unique decomposition as $\mathbf{x} = \mathrm{Proj}_{\mathcal{H}}\{\mathbf{x}\} + \mathrm{Proj}_{\mathcal{N}}\{\mathbf{x}\}$ with

$$\mathrm{Proj}_{\mathcal{H}}\{\mathbf{x}\} = \mathbf{R}\mathbf{R}^\dagger \mathbf{x} \quad \text{and} \quad \mathrm{Proj}_{\mathcal{N}}\{\mathbf{x}\} = (\mathbf{I} - \mathbf{R}\mathbf{R}^\dagger)\mathbf{x}.$$

In particular, if $\mathbf{f} = (f_1, \dots, f_N) \in \mathcal{H}$, then $\mathbf{f} = \mathrm{Proj}_{\mathcal{H}}\{\mathbf{f}\} = \mathbf{R}\mathbf{R}^\dagger \mathbf{f}$. The latter identity is equivalent to

$$\text{Property } (ii): \quad f_n = \langle \mathbf{r}_n, \mathbf{f} \rangle_{\mathcal{H}} = \langle R[n, \cdot], \mathbf{f} \rangle_{\mathcal{H}}, \text{ for all } \mathbf{f} \in \mathcal{H},$$

which is the finite-dimensional counterpart of the second property in Definition 7. □

With this interpretation, the matrix $\mathbf{R}^\dagger$ is the Riesz map $\mathcal{H} \to \mathcal{H}' = \mathrm{span}\{\mathbf{u}_m^*\}_{n=1}^{N'}$ where the dual space $\mathcal{H}'$ is equipped with the inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{H}'} = \langle \mathbf{x}, \mathbf{R}\mathbf{y} \rangle.$$

The conjugate basis $\mathbf{u}_n^* = \mathbf{R}^\dagger \mathbf{u}_n \in \mathcal{H}'$ is given by (23), which shows that the dual space is actually spanned by the same eigenvectors as $\mathcal{H}$, although the underlying norms are different. Moreover, $\mathcal{H}'$ also happens to be the RKHS associated with the positive-definite matrix $\mathbf{R}^\dagger$.

Likewise, the positive-definite matrix $\mathbf{R}$ is the Riesz map $\mathcal{H}' \to \mathcal{H}$ and is (isometrically) invertible over $\mathcal{H}$. In other words, we have that

$$\mathbf{R}^\dagger \mathbf{R}\mathbf{x} = \mathbf{x} \quad \text{and} \quad \mathbf{R}\mathbf{R}^\dagger \mathbf{y} = \mathbf{y}.$$

for all $\mathbf{x} \in \mathcal{H}'$ and $\mathbf{y} \in \mathcal{H}$. Since both $\mathbf{R}$ and $\mathbf{R}^\dagger$ are symmetric and the (Riesz) conjugate of a vector $\mathbf{x} \in \mathcal{H}'$ is given by $\mathbf{x}^* = \mathbf{R}\mathbf{x} \in \mathcal{H}$, the above projection identities are equivalent to

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}^*, \mathbf{y} \rangle_{\mathcal{H}} = \langle \mathbf{R}\mathbf{x}, \mathbf{y} \rangle_{\mathcal{H}} = \langle \mathbf{x}, \mathbf{R}^\dagger \mathbf{y} \rangle_{\mathcal{H}'} = \langle \mathbf{x}, \mathbf{y}^* \rangle_{\mathcal{H}'}$$

for all $\mathbf{x} \in \mathcal{H}'$ and $\mathbf{y} \in \mathcal{H}$, which summarizes the Riesz isomorphism between $\mathcal{H}$ and $\mathcal{H}'$.

These various identities suggest that we can also proceed the other way around and define $\mathcal{H}$ based on the symmetric positive-definite matrix $\mathbf{R}^\dagger$. Instead of considering $\mathbf{R}^\dagger$ directly, it is usually more convenient to work with the factorized form

$$\mathbf{R}^\dagger = \mathbf{L}^T \mathbf{L}$$

where the matrix $\mathbf{L} \in \mathbb{R}^{N \times N}$ specifies a linear operator $\mathbb{R}^N \to \mathbb{R}^N$ whose null space is $\mathcal{N} = \mathrm{span}\{\mathbf{u}_n\}_{n=N'+1}^{N}$, corresponding to the vanishing eigenvalues of $\mathbf{R}^\dagger$ (or $\mathbf{R}$). This allows us to rewrite the inner product associated with the RKHS $\mathcal{H}$ in the simpler form

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{H}} = \langle \mathbf{R}^\dagger \mathbf{x}, \mathbf{y} \rangle_{\mathcal{H}} = \langle \mathbf{L}\mathbf{x}, \mathbf{L}\mathbf{y} \rangle,$$

where $\mathbf{L}$ is our finite-dimensional analog of the *regularization* operator. While this form can be more attractive computationally, two remarks are in order: First, the factorization is non-unique: there are many equivalent solutions such as the positive square-root of $\mathbf{R}^\dagger$ and the Cholesky decomposition where $\mathbf{L}^T$ and $\mathbf{L}$ are lower and upper triangular, respectively. The second point is that $\langle \mathbf{L}\mathbf{x}, \mathbf{L}\mathbf{y} \rangle$ does not define a valid inner product over the whole space $\mathbb{R}^N$ unless $\mathbf{L}$ is invertible, which only happens when $N' = N$. Concretely, this means that one needs to impose additional constraints to factor out the null space $\mathcal{N}$; for instance, the orthogonally conditions $\langle \mathbf{x}, \mathbf{u}_n \rangle = 0$ for $n = N' + 1, \ldots, N$.

We shall now develop similar schemes in the continuous domain starting with the simplest case where the regularization operator is invertible. The main difficulty with the extended theory is that there is no infinite-dimensional counterpart of the eigen-decomposition (21) unless the underlying operator R is compact, which is usually not the case.

2.4 RKHS associated with an invertible operator

As first constructive example of a RKHS that is a "regularized" version of $L_2(\mathbb{R}^d)$, we consider the space

$$\mathcal{H}_{\mathrm{L}} = \{f \in \mathcal{S}'(\mathbb{R}^d) : \|\mathrm{L}f\|_{L_2(\mathbb{R}^d)} < \infty\} \tag{24}$$

where $\mathrm{L} : \mathcal{H}_{\mathrm{L}} \to L_2(\mathbb{R})$ is a coercive linear operator such that

$$c\|f\|_{L_2(\mathbb{R}^d)} \leq \|\mathrm{L}f\|_{L_2(\mathbb{R}^d)} \tag{25}$$

for all $f \in L_2(\mathbb{R}^d)$ and some constant $c > 0$. The effect of L in (24) is to induce some smoothing on f, which is the reason why it is often called the *regularization operator* of the RKHS.

One then easily shows that the bilinear form

$$\langle f, g\rangle_{\mathcal{H}_{\mathrm{L}}} = \langle \mathrm{L}f, \mathrm{L}g\rangle$$

is a valid inner product for $\mathcal{H}_{\mathrm{L}}$. In addition, the coercivity property (lower bound in (25)) implies that $\mathcal{H}_{\mathrm{L}}$ is continuously embedded in $L_2(\mathbb{R}^d)$, with the two Hilbert spaces being isometric. In other words, L is a unitary map $\mathcal{H}_{\mathrm{L}} \to L_2(\mathbb{R}^d)$ that admits a well-defined (i.e., continuous) inverse $\mathrm{G} = \mathrm{L}^{-1} : L_2(\mathbb{R}^d) \to \mathcal{H}_{\mathrm{L}}$. It is important here to emphasize that it is the coercivity property (25) that makes the mapping bijective. In particular, (25) ensures that the null space of L is empty.

The space $\mathcal{H}'_{\mathrm{L}}$ is the continuous dual of $\mathcal{H}_{\mathrm{L}}$ as well as the range (resp., domain) of the adjoint operator $\mathrm{L}^* : L_2(\mathbb{R}^d) \to \mathcal{H}'_{\mathrm{L}}$ (resp. $\mathrm{G}^* : \mathcal{H}'_{\mathrm{L}} \to L_2(\mathbb{R}^d)$). It is also easy to see that G^* and L^* are inverse of each other.

Under the implicit assumption of continuity (i.e., $\mathcal{H}_{\mathrm{L}} \subseteq C_{\mathrm{b},0}(\mathbb{R}^d)$), $\mathcal{H}_{\mathrm{L}}$ is a RKHS whose reproducing kernel is the impulse response of the Riesz map R from $\mathcal{H}'_{\mathrm{L}} \to \mathcal{H}_{\mathrm{L}}$ (see Property 6 in Proposition 2). Moreover, R is the inverse of $\mathrm{R}^{-1} : \mathcal{H}_{\mathrm{L}} \to \mathcal{H}'_{\mathrm{L}}$ (the Riesz map from $\mathcal{H}_{\mathrm{L}} \to \mathcal{H}'_{\mathrm{L}}$). From the definition of $\mathcal{H}_{\mathrm{L}}$ and Property 7 in Proposition 2, we have the isometry

$$\|f\|^2_{\mathcal{H}_{\mathrm{L}}} = \|\mathrm{L}f\|^2_{L_2(\mathbb{R}^d)} = \langle \mathrm{L}^*\mathrm{L}f, f\rangle = \langle \mathrm{R}^{-1}f, f\rangle$$

for all $f \in \mathcal{H}_{\mathrm{L}}$, which implies that $\mathrm{R}^{-1} = \mathrm{L}^*\mathrm{L}$. It then follows that

$$h(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{R}\{\delta(\cdot - \boldsymbol{y})\}(\boldsymbol{x})$$

with $\mathrm{R} = (\mathrm{L}^*\mathrm{L})^{-1} : \mathcal{H}'_{\mathrm{L}} \to \mathcal{H}_{\mathrm{L}}$ where we are taking advantage of the fact that both L and L^* are invertible.

In the special case where L is LSI, we obtain the simplified form

$$h(\boldsymbol{x},\boldsymbol{y}) = \rho_{\mathrm{L^*L}}(\boldsymbol{x}-\boldsymbol{y})$$

where $\rho_{\mathrm{L^*L}}$ is the (unique) symmetric Green's function of $\mathrm{L^*L}$. This function can be conveniently obtained by inverse Fourier transformation:

$$\rho_{\mathrm{L^*L}}(\boldsymbol{x}) = \mathcal{F}^{-1}\left\{\frac{1}{|\widehat{L}(\boldsymbol{\omega})|^2}\right\}(\boldsymbol{x})$$

where $\widehat{L}(\boldsymbol{\omega})$ is the Fourier symbol of L.

Some useful example: Exponential, Bessel potentials.

2.5 Factorization of a reproducing kernel

In the example from the previous section, the Riesz map from $\mathcal{H}'_{\mathrm{L}}$ to $\mathcal{H}_{\mathrm{L}}$ can be written as $\mathrm{R} = \mathrm{GG}^*$ where the inverse operator $\mathrm{G} = \mathrm{L}^{-1}$ continuously maps $L_2(\mathbb{R}^d) \to \mathcal{H}_{\mathrm{L}}$. This implies that the reproducing kernel of $\mathcal{H}_{\mathrm{L}}$ has a factorized representation as

$$h(\boldsymbol{x},\boldsymbol{y}) = \int_{\mathbb{R}^d} g(\boldsymbol{x},\boldsymbol{z})g(\boldsymbol{y},\boldsymbol{z})\mathrm{d}\boldsymbol{z}$$

where $g(\boldsymbol{x},\boldsymbol{y}) = \mathrm{G}\{\delta(\cdot-\boldsymbol{y})\}(\boldsymbol{x})$ is the generalized impulse response of G. We shall now drop the coercivity requirement and prove that this factorization property extends for RKHS that are linked to a much broader class of regularization operators. The idea is that one can always identify an intermediate Hilbert space $\mathcal{H}_0 \subseteq L_2(\mathbb{R}^d)$ (the actual range of the operator $\mathrm{L}: \mathcal{H}_{\mathrm{L}} \to \mathcal{H}_0$) and an operator $\mathrm{L}^{-1}: \mathcal{H}_0 \to \mathcal{H}_{\mathrm{L}}$ that is a proper inverse of L on $\mathcal{H}_0 = \mathrm{Im}(\mathcal{H}_{\mathrm{L}})$. While the domain of L^{-1} can also be extended to $L_2(\mathbb{R}^d)$, the extended operator is generally only a left inverse of L with the property that

$$\mathrm{L}^{-1}\mathrm{L}f = f$$

for all $f \in \mathcal{H}_{\mathrm{L}}$. On the other hand, there is no guarantee a priori that $\mathrm{LL}^{-1}w = w$ for all $w \in L_2(\mathbb{R}^d)$ (right inverse property) unless $\mathcal{H}_0 = L_2(\mathbb{R}^d)$.

Theorem 8 (Factorization of reproducing kernel). *Let $\mathcal{H} \subseteq C_{\mathrm{b},\alpha_0}(\mathbb{R}^d)$ be a RKHS with regularization operator $\mathrm{L}: \mathcal{H} \to L_2(\mathbb{R})$ such that*

$$\langle f, g\rangle_{\mathcal{H}} = \langle \mathrm{L}f, \mathrm{L}g\rangle \tag{26}$$

for all $f \in \mathcal{H}$. *Then, there exits a unique continuous operator* $\mathrm{L}^{-1} : L_2(\mathbb{R}^d) \to L_{\infty,\alpha_0}(\mathbb{R}^d)$ *and a Hilbert space* $\mathcal{H}_0 \subseteq L_2(\mathbb{R}^d)$ *such that the Riesz map from* $\mathcal{H}' \to \mathcal{H}$ *factors though* $\mathcal{H}_0$ *as* $\mathrm{R} = \mathrm{L}^{-1}\mathrm{L}^{-1*} : \mathcal{H}' \to \mathcal{H}_0 \to \mathcal{H}$. *The Schwartz kernel of the inverse operator* L^{-1} *denoted by* $g(\cdot,\cdot)$ *satisfies the estimate*

$$\sup_{\boldsymbol{x}\in\mathbb{R}^d} (1+\|\boldsymbol{x}\|)^{\alpha_0} \|g(\boldsymbol{x},\cdot)\|_{L_2(\mathbb{R}^d)} < \infty \tag{27}$$

and is linked to the reproducing kernel $h(\cdot,\cdot)$ *of* $\mathcal{H}$ *by*

$$h(\boldsymbol{x},\boldsymbol{y}) = \int_{\mathbb{R}^d} g(\boldsymbol{x},\boldsymbol{z})g(\boldsymbol{y},\boldsymbol{z})\mathrm{d}\boldsymbol{z} \tag{28}$$

$$g(\boldsymbol{y},\boldsymbol{x}) = \mathrm{L}\{h(\cdot,\boldsymbol{y})\}(\boldsymbol{x}). \tag{29}$$

Proof. By using the definition of the inverse Riesz map $\mathrm{R}^{-1} : \mathcal{H} \to \mathcal{H}'$ and pluging $f = g$ in (26), we get

$$\|f\|_{\mathcal{H}}^2 = \langle \mathrm{L}f, \mathrm{L}f\rangle = \langle \mathrm{L}^*\mathrm{L}f, f\rangle = \langle \mathrm{R}^{-1}f, f\rangle,$$

which implies that $\mathrm{R}^{-1} = (\mathrm{L}^*\mathrm{L})$. Moreover, the fact that $\langle \mathrm{L}f, \mathrm{L}g\rangle$ specifies a valid inner product is equivalent to the existence of an "intermediate" Hilbert space $\mathcal{H}_0 \subseteq L_2(\mathbb{R}^d)$ such that L is a unitary mapping $\mathcal{H} \to \mathcal{H}_0$ with

$$\langle f, g\rangle_{\mathcal{H}} = \langle \mathrm{L}f, \mathrm{L}g\rangle_{\mathcal{H}_0} = \langle \mathrm{L}f, \mathrm{L}g\rangle = \langle \mathrm{R}^{-1}f, g\rangle$$

for all $f, g \in \mathcal{H}$. Another way to put it is that R^{-1} has a unitary factorization through $\mathcal{H}_0$ as $\mathrm{R}^{-1} = (\mathrm{L}^*\mathrm{L}) : \mathcal{H} \to \mathcal{H}_0 \to \mathcal{H}'$. Since all these mapping are unitary, there exists some corresponding (unique) inverse operators $\mathrm{L}^{-1^*} : \mathcal{H}' \to \mathcal{H}_0$, $\mathrm{L}^{-1} : \mathcal{H}_0 \to \mathcal{H}$ and $\mathrm{R} = (\mathrm{L}^{-1}\mathrm{L}^{-1^*}) : \mathcal{H}' \to \mathcal{H}_0 \to \mathcal{H}$ such that

$$\langle u, v\rangle_{\mathcal{H}'} = \langle \mathrm{L}^{-1*}u, \mathrm{L}^{-1*}v\rangle_{\mathcal{H}_0} = \langle u, \mathrm{L}^{-1}\mathrm{L}^{-1*}v\rangle = \langle \mathrm{u}, \mathrm{R}v\rangle \tag{30}$$

for all $u, v \in \mathcal{H}'$. In particular, this implies that the operator $\mathrm{L}^{-1} : \mathcal{H}_0 \to \mathcal{H}$ (resp., $\mathrm{L}^{-1*} : \mathcal{H}' \to \mathcal{H}_0$) is unitary and a proper inverse of $\mathrm{L} : \mathcal{H} \to \mathcal{H}_0$ (resp., $\mathrm{L}^* : \mathcal{H}_0 \to \mathcal{H}'$).

Next, we recall that the reproducing kernel is the generalized impulse response of the Riesz map, so that

$$\begin{aligned} h(\boldsymbol{x},\boldsymbol{y}) &= \mathrm{R}\{\delta(\cdot - \boldsymbol{y})\}(\boldsymbol{x}) \\ &= \mathrm{L}^{-1}\mathrm{L}^{-1*}\{\delta(\cdot - \boldsymbol{y})\}(\boldsymbol{x}) \\ &= \mathrm{L}^{-1}\{g(\boldsymbol{y},\cdot)\}(\boldsymbol{x}) \\ &= \int_{\mathbb{R}^d} g(\boldsymbol{x},\boldsymbol{z})g(\boldsymbol{y},\boldsymbol{z})\mathrm{d}\boldsymbol{z} \end{aligned} \tag{31}$$

The fact that $\mathcal{H}$ is a RKHS implies that $\delta(\cdot - \boldsymbol{y}) \in \mathcal{H}'$ for any $\boldsymbol{y} \in \mathbb{R}^d$, which allows us to infer that $g(\boldsymbol{y}, \cdot) = \mathrm{L}^{-1*}\{\delta(\cdot - \boldsymbol{y})\} \in \mathcal{H}_0$. We then invoke (31) and the inverse property of L^{-1} on $\mathcal{H}_0$ to show that

$$\mathrm{L}\{h(\cdot, \boldsymbol{y})\}(\boldsymbol{x}) = \mathrm{L}\mathrm{L}^{-1}\{g(\boldsymbol{y}, \cdot)\}(\boldsymbol{x}) = g(\boldsymbol{y}, \boldsymbol{x}).$$

Finally, to reveal the boundedness of $\mathrm{L}^{-1} : L_2(\mathbb{R}^d) \to L_{\infty,\alpha_0}(\mathbb{R}^d)$, we use the hypothesis $\mathcal{H} \subseteq C_{\mathrm{b},\alpha_0}(\mathbb{R}^d)$, which, by Theorem 7, is equivalent to

$$A_{h,\alpha} = \sup_{\boldsymbol{x} \in \mathbb{R}^d} |h(\boldsymbol{x}, \boldsymbol{x})| \, (1 + \|\boldsymbol{x}\|)^{2\alpha} < \infty.$$

Based on (28), we then rewrite this condition as

$$\sup_{\boldsymbol{x} \in \mathbb{R}^d} (1 + \|\boldsymbol{x}\|)^{\alpha_0} \|g(\boldsymbol{x}, \cdot)\|_{L_2(\mathbb{R}^d)} = \sqrt{A_{h,\alpha}} < \infty,$$

which is precisely the bound in Proposition 4 that ensures the continuity of $\mathrm{L}^{-1} : L_2(\mathbb{R}^d) \to L_{\infty,\alpha_0}(\mathbb{R}^d)$. Hence, we can safely extend the domain of the operator from $\mathcal{H}_0$ to $L_2(\mathbb{R}^d)$ (by the Hahn-Banach theorem). □

Corollary 2. *Let* $\mathcal{H}_{\mathrm{L}} \subseteq C_{\mathrm{b},\alpha_0}(\mathbb{R}^d)$ *be the RKHS associated with the regularization operator* $\mathrm{L} : \mathcal{H} \to \mathcal{H}_0 \subseteq L_2(\mathbb{R}^d)$ *and the inner product* $\langle \mathrm{L}f, \mathrm{L}g \rangle$*. Then, the unique inverse operator* L^{-1} *identified in Theorem 8 continuously maps* $L_2(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$*, while its adjoint* L^{-1*} *continuously maps* $\mathcal{S}(\mathbb{R}^d) \to L_2(\mathbb{R}^d)$*.*

Proof. The bound in Theorem 8 ensures that the inverse operator $\mathrm{L}^{-1} : \mathcal{H}_0 \to \mathcal{H}_{\mathrm{L}}$ has a proper extension $L_2(\mathbb{R}^d) \to L_{\infty,\alpha_0}(\mathbb{R}^d)$. The result then follows from the fact that $L_{\infty,\alpha_0}(\mathbb{R}^d) \subseteq \mathcal{S}'(\mathbb{R}^d)$ and the property that $L_2(\mathbb{R}^d)$ is its own dual. □

Since the factorization result in Theorem 8 is stated in terms of the regularization operator L, it raises the issue of the existence of such an operator for an arbitrary RKHS.

Let us first address the simpler related question of unicity. If $\mathrm{L} : \mathcal{H} \to \mathcal{H}_0 \subseteq L_2(\mathbb{R}^d)$ is a regularization operator for the RKHS as stated in Theorem 8, then the same holds true for $\tilde{\mathrm{L}} = \mathrm{UL}$ where $\mathrm{U} : L_2(\mathbb{R}^d) \to L_2(\mathbb{R}^d)$ is an arbitrary L_2-isometry with $\mathrm{UU}^* = \mathrm{U}^*\mathrm{U} = \mathrm{Id}$. This obviously leaves the reproducing kernel unchanged since $\mathrm{R}^{-1} = \tilde{\mathrm{L}}^*\tilde{\mathrm{L}} = \mathrm{LU}^*\mathrm{UL} = \mathrm{L}^*\mathrm{L}$. Likewise, we have that $\tilde{\mathrm{L}}^{-1} = \mathrm{L}^{-1}\mathrm{U}$ and $\mathrm{R} = \tilde{\mathrm{L}}^{-1}\tilde{\mathrm{L}}^{-1*} = \mathrm{L}^{-1}\mathrm{UU}^*\mathrm{L}^{-1*} = \mathrm{L}^{-1}\mathrm{L}^{-1*}$.

We can also guarantee the existence of L when $\mathcal{H} \subseteq L_2(\mathbb{R}^d)$. The simplest solution is provided by the natural embedding (or identity map) $i : L_2(\mathbb{R}^d) \to$

$\mathcal{H}$ (see Definition 3), which can usually be translated into an orthogonal projection operator $\mathrm{Proj}_{\mathcal{H}} : L_2(\mathbb{R}^d) \to \mathcal{H}$. We also note that the coercivity condition (25) in Section 2.4 implies that $\mathcal{H}_{\mathrm{L}}$ is embedded in $L_2(\mathbb{R}^d)$.

Before considering the more challenging cases where $\mathcal{H}$ is not embedded in $L_2(\mathbb{R}^d)$, we shall review three basic examples.

Example 1 (RKHS generated by an orthonormal system)**.** *Let* $\{\phi_n\}_{n\in\mathbb{N}}$ *with* $\phi_n \in C_{\mathrm{b}}(\mathbb{R}^d)$ *be an orthonormal system with* $\langle \phi_m, \phi_n\rangle = \delta_{m,n}$*. Then, one easily checks that* $\mathcal{V} = \overline{\mathrm{span}\{\phi_n\}}_{n\in\mathbb{Z}}$ *equipped with the inner product*

$$\langle f, g\rangle_{\mathcal{V}} = \sum_{n\in\mathbb{N}} \langle \phi_n, f\rangle\langle \phi_n, g\rangle$$

is a Hilbert space. Its reproducing kernel has the generic form

$$h_{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{n\in\mathbb{N}} \phi_n(\boldsymbol{x})\phi_n(\boldsymbol{y}). \tag{32}$$

This kernel can also be used to specify the orthogonal projector $\mathrm{Proj}_{\mathcal{V}} : L_2(\mathbb{R}^d) \to \mathcal{V}$ *as*

$$\mathrm{Proj}_{\mathcal{V}}\{f\}(\boldsymbol{x}) = \langle f, h_{\mathcal{V}}(\boldsymbol{x}, \cdot)\rangle = \sum_{n\in\mathbb{N}} \phi_n(\boldsymbol{x})\langle f, \phi_n\rangle,$$

which is obviously also self-adjoint. Finally, we have the factorization

$$h_{\mathcal{V}}(\boldsymbol{x}, \boldsymbol{y}) = \langle h_{\mathcal{V}}(\boldsymbol{x}, \cdot), h_{\mathcal{V}}(\cdot, \boldsymbol{y})\rangle, \tag{33}$$

which follows from the property that $\mathrm{Proj}_{\mathcal{V}} = \mathrm{Proj}_{\mathcal{V}}\mathrm{Proj}_{\mathcal{V}}^*$*. While* (33) *reminds us of the reproducing kernel Property 2 in Proposition 2, it is structurally not the same because it involves the "duality product" rather than the "inner product" of* $\mathcal{V}$*.*

We like to think of Example 1 as the simplest possible scenario covered by Theorem 8 with $\mathrm{L}^{-1} = \mathrm{Proj}_{\mathcal{V}} = \mathrm{Proj}_{\mathcal{V}}^* = \mathrm{L}$. However, there are many other interesting configurations where the factors are not self-adjoint (in the spirit of the LU decomposition of a symmetric matrix).

Our next example is of great significance for communication engineering and information sciences. It is interesting structurally because it is a cross between the previous example and the LSI regularization operators of Section 2.4, but without the coercivity property (because the null space is non-trivial).

Example 2 (RKHS of bandlimited functions)**.** *The subspace of bandlimited functions in* $L_2(\mathbb{R})$ *with Nyquist frequency* $\omega_{\max} = \pi$ *can be defined as*

$$\mathcal{H}_{\rm sinc} = \left\{ f \in \mathcal{S}'(\mathbb{R}) : ({\rm sinc} * f) \in L_2(\mathbb{R}) \right\}.$$

The underlying regularization operation $\mathrm{L}_{\rm sinc} : \varphi \mapsto {\rm sinc} * \varphi$ *is linear shift-invariant since it is a convolution with the sinus cardinalis:*

$${\rm sinc}(x) = \frac{\sin(\pi x)}{(\pi x)} = \mathcal{F}^{-1}\{\mathbb{1}_{[-\pi,\pi]}\}(x).$$

The sinc *function is endowed with the remarkable reproduction property* ${\rm sinc}(x) = ({\rm sinc} * {\rm sinc})(x)$*, and, more generally,*

$$({\rm sinc} * f)(x) = \int_{\mathbb{R}} {\rm sinc}(x-y) f(y) {\rm d}y = f(x), \text{ for all } f \in \mathcal{H}_{\rm sinc}.$$

This immediately leads to the conclusion that the reproduction kernel for $\mathcal{H}_{\rm sinc}$ *is* ${\rm sinc}(x-y)$*. On the other hand, it is also well known that the set of functions* $\{{\rm sinc}(\cdot - k)\}_{k \in \mathbb{Z}}$ *is an orthonormal basis of* $\mathcal{H}_{\rm sinc}$*. By applying the result of Example 2, we obtain the equivalent representation*

$$h_{\rm sinc}(x,y) = {\rm sinc}(x-y) = \sum_{k \in \mathbb{Z}} {\rm sinc}(x-k){\rm sinc}(y-k).$$

This translates into the reproduction formula

$$\begin{aligned} f(x) = ({\rm sinc} * f)(x) &= \sum_{k \in \mathbb{Z}} \langle f, {\rm sinc}(\cdot - k) \rangle \, {\rm sinc}(x-k) \\ &= \sum_{k \in \mathbb{Z}} f(k) \, {\rm sinc}(x-k) \end{aligned}$$

for all $f \in \mathcal{H}_{\rm sinc}$*, which is the RKHS restatement of Shannon's celebrated Sampling theorem.*

The last example of this section illustrates the point that, in contrast with matrices, the factorization of a symmetric operator is not always feasible. Moreover, it is an instance of a RKHL that is not embedded in $L_2(\mathbb{R})$ and that does not admit a regularization operator.

Example 3 (Subspace of polynomials)**.** *We now take* $\mathcal{N} = {\rm span}\{p_1\}$ *with* $p_1(x) = 1$ *as the simplest (one-dimensional) polynomial subspace of* $C_{{\rm b},\alpha}(\mathbb{R})$ *with* $\alpha \leq 0$*. It is easy to verify that* $\mathcal{N}$ *equipped with the inner product*

$$\langle f, g \rangle_{\mathcal{N}} = f(0)g(0) = \langle \delta, f \rangle \langle \delta, g \rangle$$

is a RKHS and that its reproducing kernel is given by $h_{\mathcal{N}}(x,y) = p_1(x)p_1(y) = 1$*, which is similar to* (32)*. Since* (δ, p_1) *form a biorthogonal pair, we can define the corresponding projection operator* $C_{\mathrm{b},\alpha}(\mathbb{R}) \to \mathcal{N}$ *as*

$$\mathrm{Proj}_{\mathcal{N}}\{f\} = \langle \delta, f\rangle p_1,$$

which is such that $\mathrm{Proj}_{\mathcal{N}_1}\{p\} = p$ *for all* $p \in \mathcal{N}$*. However, we are not able to find a factor* $g(x,y)$ *such that*

$$h(x,y) = \langle g(x,\cdot), g(y,\cdot)\rangle,$$

the fundamental reason being that $p_1 \notin L_2(\mathbb{R})$ *so that it cannot be orthonormalized.*

We can also use this last example to gain more insight on the abstract notion of dual space. Since $\mathcal{N}$ is one-dimensional, the same hold true for its continuous dual $\mathcal{N}'$, which is spanned by the Dirac functional $p = c_1 p_1 \mapsto \langle \delta, p\rangle = c_1$. However, since $\mathcal{N}'$ is embedded in $\mathcal{S}'(\mathbb{R}^d)$, the dual functional of p_1 is actually the equivalence class of all (generalized) functions $\phi_1 \in \mathcal{S}'(\mathbb{R})$ such that $\langle \phi_1, p_1\rangle = \int_{\mathbb{R}} \phi_1(x)\mathrm{d}x = 1$. The Dirac impulse δ is just one representer among an infinity of possibilities. In other words, we could as well have used any other biorthogonal function ϕ_1, which changes the form of the projector $\mathrm{Proj}_{\mathcal{N}}$, but leaves the reproducing kernel unchanged. This is a possibility that will be exploited in the subsequent sections.

2.6 RKHS associated with the derivative operator

As next step in the progression, it is instructive to investigate the derivative operator $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}x}$, which is the simplest differential operator that has a non-trivial null space. The functional characteristics that are relevant for our purpose are:

- The causal Green's function of D (or Heaviside function): $\mathbb{1}_+(x)$. The defining property is $\mathrm{D}\{\mathbb{1}_+\} = \delta$, which is equivalent to $\mathrm{D}\{\mathbb{1}_+(\cdot - y)\} = \delta(\cdot - y)$, due to the shift-invariance of D.

- The null space of dimension $N_0 = 1$:

$$\mathcal{N}_{\mathrm{D}} = \{q \in \mathcal{S}'(\mathbb{R}) : \mathrm{D}\{q\} = 0\} = \mathrm{span}\{p_1\} \quad (34)$$

  where $p_1(x) = 1$ is the constant function.

Let us also recall that any function of the form $\rho_{\mathrm{D}}(x) = \mathbb{1}_+(x) + c_0$ is a Green's function of D. The so-called canonical solution is

$$\tfrac{1}{2}\mathrm{sign}(x) = \mathcal{F}\left\{\frac{1}{\mathrm{j}\omega}\right\}(x) = \mathbb{1}_+(x) - \tfrac{1}{2}.$$

The adjoint of D is $\mathrm{D}^* = -\mathrm{D}$. We shall also need the canonical Green's function of $(\mathrm{D}^*\mathrm{D}) = -\mathrm{D}^2$, which is given by

$$\rho_{\mathrm{D}^*\mathrm{D}}(x) = \mathcal{F}\left\{\frac{1}{\omega^2}\right\}(x) = -\tfrac{1}{2}|x|.$$

Since the null space of D is non-trivial, the operator $\mathrm{D} : \mathcal{H}_{\mathrm{D}} \to L_2(\mathbb{R})$ does not fall in the coercive category of Section 2.4. Yet, we will now see that the function space

$$\mathcal{H}_{\mathrm{D}} = \{f : \mathbb{R} \to \mathbb{R} \text{ s.t. } \mathrm{D}f \in L_2(\mathbb{R})\}$$

can still be identified as a RKHS. To ensure unicity, we have to impose some additional boundary condition: for example, fixing the value of $f(x)$ at $x = 0$.

Hence, our first claim is that $\mathcal{H}_{\mathrm{D}}$ equipped with the inner product

$$\langle f, g\rangle_{\mathcal{H}_{\mathrm{D}}} = \langle \mathrm{D}f, \mathrm{D}g\rangle + f(0)g(0) \tag{35}$$

is a RKHS with critical rate of growth $-\alpha_0 = 1$; that is, $\mathcal{H}_{\mathrm{D}} \subseteq C_{\mathrm{b},-1}(\mathbb{R})$. It is obvious that $\langle \cdot, \cdot\rangle_{\mathcal{H}_{\mathrm{D}}}$ satisfies the three first properties of an inner product in Definition 1. As for the unicity, we observe that $\|f\|^2_{\mathcal{H}_{\mathrm{D}}} = \|\mathrm{D}f\|^2_{L_2(\mathbb{R})} + |f(0)|^2$ so that $\|f\|^2_{\mathcal{H}_{\mathrm{D}}} = 0$ implies that: (i) $\|\mathrm{D}f\|_{L_2(\mathbb{R})} = 0$, and (ii) $f(0) = 0$. Condition (i) is equivalent to $f \in \mathcal{N}_{\mathrm{D}}$, while (ii) removes the ambiguity by forcing the constant to be zero, which proves our assertion.

Next, we claim that the corresponding reproducing kernel is

$$h_{\mathrm{D}}(x, y) = \frac{1}{2}\big(|x| + |y| - |x - y|\big) + 1. \tag{36}$$

While there is a constructive mechanism for obtaining this formula (see Proposition 7), we shall first convince ourselves of its correctness by checking that it fulfils the required conditions. To that end, we first evaluate

$$\begin{aligned}
\mathrm{D}\{h_{\mathrm{D}}(\cdot, y)\}(x) &= \tfrac{1}{2}\big(\mathrm{sign}(x) - \mathrm{sign}(x - y)\big)\\
\mathrm{D}^*\mathrm{D}\{h_{\mathrm{D}}(\cdot, y)\}(x) &= -\mathrm{D}\{\tfrac{1}{2}\big(\mathrm{sign}(\cdot) - \mathrm{sign}(\cdot - y)\big)\}(x) = -\delta(x) + \delta(x - y)\\
h_{\mathrm{D}}(0, y) &= \tfrac{1}{2}\left(|0| + |y| - |0 - y|\right) + 1 = 1
\end{aligned}$$

Hence, for all $f \in \mathcal{H}_{\mathrm{D}}$,

$$\begin{aligned}\langle f, h_{\mathrm{D}}(\cdot,y)\rangle_{\mathcal{H}_{\mathrm{D}}} &= \langle \mathrm{D}f, \mathrm{D}\{h_{\mathrm{D}}(\cdot,y)\}\rangle + f(0)\times 1\\ &= \langle f, \mathrm{D}^*\mathrm{D}\{h_{\mathrm{D}}(\cdot,y)\}\rangle + f(0)\\ &= \big(-f(0)+f(y)\big) + f(0) = f(y),\end{aligned}$$

which proves that the kernel satisfies the reproduction property. What is less obvious is that $h_{\mathrm{D}}(\cdot,x_0) \in \mathcal{H}_{\mathrm{D}}$ for any $x_0 \in \mathbb{R}$. The explanation is that $\mathrm{D}\{h_{\mathrm{D}}(\cdot,x_0)\}(x) = \mathbb{1}_{[0,x_0)}(x)$ is actually compactly supported of size x_0 (thanks to some convenient cancellation mechanism) so that is included in $L_2(\mathbb{R})$.

Supplementary material: To gain a deeper understanding of this construction, we now pick a generic biorthogonal analysis function $\phi \in \mathcal{H}'_{\mathrm{D}}$ such that $\langle \phi, p_1\rangle = 1$ and define the space

$$\mathcal{H}_{\mathrm{D},\phi} = \{f \in \mathcal{H}_{\mathrm{D}} : \langle \phi, f\rangle = 0\}\,. \tag{37}$$

We shall now show that $\mathcal{H}_{\mathrm{D},\phi}$ and $\mathcal{N}_{\mathrm{D}}$ are two complementary Hilbert spaces associated with the inner products $\langle f, g\rangle_{\mathcal{H}_{\mathrm{D},\phi}} = \langle \mathrm{D}f, \mathrm{D}g\rangle$ and $\langle p, q\rangle_{\mathcal{N}_{\mathrm{D}}} = \langle \phi, p\rangle\langle \phi, q\rangle$, respectively. The proposed characterization of $\mathcal{N}_{\mathrm{D}}$ is a slight generalization of Example 3 in Section 2.5 where δ is substituted by the generic analysis function ϕ. As already mentioned, this does not affect the form of the reproducing kernel, which is still given by

$$p_1(x)p_1(y) = 1. \tag{38}$$

The corresponding projection operation $\mathrm{Proj}_{\mathcal{N}_{\mathrm{D}}} : \mathcal{H}_{\mathrm{D}} \to \mathcal{N}_{\mathrm{D}}$ is

$$\mathrm{Proj}_{\mathcal{N}_{\mathrm{D}}}\{f\} = \langle \phi, f\rangle p_1,$$

with the property that $\mathrm{Proj}_{\mathcal{N}_{\mathrm{D}}}\{q\} = q$ for all $q = c_0 p_1 \in \mathcal{N}_{\mathrm{D}}$ (due to the biorthogonality of ϕ and p_1).

To prove that the semi-inner product $\langle \mathrm{D}f, \mathrm{D}g\rangle$ is actually an inner product for $\mathcal{H}_{\mathrm{D},\phi}$, we recall that $\langle \mathrm{D}f, \mathrm{D}f\rangle = 0$ is equivalent to $f \in \mathcal{N}_{\mathrm{D}}$, so that $f = \mathrm{Proj}_{\mathcal{N}_{\mathrm{D}}}\{f\} = \langle \phi, f\rangle p_1$. On the other hand, we have that $\langle \phi, f\rangle = 0$ from the definition of $\mathcal{H}_{\mathrm{D},\phi}$, which gives $f = 0$ and hence proves unicity.

Next, we define the operator $\mathrm{D}_\phi^{-1} : \mathcal{H}_{\mathrm{D},\phi} \to L_2(\mathbb{R})$ whose kernel is given by

$$g_\phi(x,y) = \mathbb{1}_+(x-y) - p_1(x)q_1(y) \tag{39}$$

with $q_1(y) = \langle \phi, \mathbb{1}_+(\cdot - y)\rangle$ and $p_1(x) = 1$. Observe that the function q_1 is bounded with $|q_1(y)| \le \|\phi\|_{\mathrm{TV}}$ where $\|\phi\|_{\mathrm{TV}} = \sup_{|\varphi|_\infty \le 1}\langle \phi, \varphi\rangle$ is the "total variation" of ϕ with $\|\phi\|_{\mathrm{TV}} = \|\phi\|_{L_1(\mathbb{R})}$ when $\phi \in L_1(\mathbb{R})$. This implies that the kernel defined by (39) is bounded—i.e, $g_\phi(\cdot,\cdot) \in L_\infty(\mathbb{R}\times\mathbb{R})$—irrespective of the choice of ϕ.

Based on the Green's function property $\mathrm{D}\{\mathbb{1}_+\} = \delta$ and the definition of $q_1(y)$, we easily verify that

$$\mathrm{D}\{g_\phi(\cdot, y)\}(x) = \delta(\cdot - y) \tag{40}$$

$$\langle \phi, g_\phi(\cdot, y)\rangle = \langle \phi, \mathbb{1}_+(\cdot - y)\rangle - q_1(y) = 0. \tag{41}$$

Equation (40) implies that D_ϕ^{-1} is a right-inverse of D, while (41) enforces the boundary condition $\langle \phi, \mathrm{D}_\phi^{-1}\{w\}\rangle = 0$ for all $w \in L_2(\mathbb{R}^d)$. Indeed,

$$\begin{aligned}\langle \phi, \mathrm{D}_\phi^{-1}\{w\}\rangle &= \langle \phi, \int_{\mathbb{R}} g_\phi(\cdot, y) w(y)\mathrm{d}y\rangle \\ &= \int_{\mathbb{R}} \underbrace{\langle \phi, g_\phi(\cdot, y)\rangle}_{=0} w(y)\mathrm{d}y = 0\end{aligned}$$

This allows us to redefine the Hilbert space $\mathcal{H}_{\mathrm{D},\phi}$ as

$$\mathcal{H}_{\mathrm{D},\phi} = \left\{ f = \mathrm{D}_\phi^{-1}\{w\} : w \in L_2(\mathbb{R}) \right\},$$

which also comes hand-in-hand with the norm-conservation properties

$$\begin{aligned}\|\mathrm{D}_\phi^{-1}\{w\}\|_{\mathrm{D},\phi} &= \|w\|_{L_2(\mathbb{R})} \\ \|f\|_{\mathrm{D},\phi} &= \|\mathrm{D}f\|_{L_2(\mathbb{R})},\end{aligned}$$

for all $f \in \mathcal{H}_{\mathrm{D},\phi}$ and $w \in L_2(\mathbb{R})$. By rewriting the inner product as $\langle f, g\rangle_{\mathrm{D},\phi} = \langle (\mathrm{D}^*\mathrm{D})f, g\rangle$, we deduce that $\mathrm{R}^{-1} = (\mathrm{D}^*\mathrm{D})$ is the (inverse) Riesz map from $\mathcal{H}_{\mathrm{D},\phi} \to \mathcal{H}'_{\mathrm{D},\phi}$. Similarly, we identify the direct Riesz map as $\mathrm{R} = \mathrm{D}_\phi^{-1}\mathrm{D}_\phi^{-1*} : \mathcal{H}'_{\mathrm{D},\phi} \to \mathcal{H}_{\mathrm{D},\phi}$, which is consistent with the property that $\mathrm{D}_\phi^{-1} : L_2(\mathbb{R}) \to \mathcal{H}_{\mathrm{D},\phi}$ is an isometry.

For completeness, we also provide the explicit form of the reproducing kernel of $\mathcal{H}_{\mathrm{D},\phi}$.

Proposition 6. *Let $g_\phi(x, z) = \rho_{\mathrm{D}}(x - y) - \langle \rho_{\mathrm{D}}(\cdot - y), \phi\rangle$ where ρ_{D} and $\phi \in \mathcal{H}'_{\mathrm{D}}$ are such that $\mathrm{D}\{\rho_{\mathrm{D}}\} = \delta$ (Green's function property) and $\langle \phi, 1\rangle = 1$, respectively. Then, the reproducing kernel of the space*

$$\mathcal{H}_{\mathrm{D},\phi} = \{f \in \mathcal{S}'(\mathbb{R}) : \mathrm{D}f \in L_2(\mathbb{R}) \textit{ and } \langle \phi, f\rangle = 0\}$$

is given by

$$h_\phi(x,y) = \int_{\mathbb{R}} g_\phi(x,z) g_\phi(y,z)\mathrm{d}z$$
$$= \rho_{\mathrm{D}^*\mathrm{D}}(x-y) - q_{\mathrm{D},\phi}(y) - q_{\mathrm{D},\phi}(x) + r_{1,1} \tag{42}$$

where

$$\rho_{\mathrm{D}^*\mathrm{D}}(x) = \mathcal{F}^{-1}\left\{\frac{1}{\omega^2}\right\}(x) = -\tfrac{1}{2}|x|$$
$$q_{\mathrm{D},\phi}(y) = \langle \phi, \rho_{\mathrm{D}^*\mathrm{D}}(\cdot - y)\rangle = (\phi * \rho_{\mathrm{D}^*\mathrm{D}})(y)$$
$$r_{1,1} = \int_{\mathbb{R}^2} \phi(x)\rho_{\mathrm{D}^*\mathrm{D}}(x-y)\phi(y)\mathrm{d}x\mathrm{d}y.$$

Proof. The direct calculation of the composed operator is in principle feasible (either in the signal or Fourier domain), but rather technical for it involves singular integrals. Instead, we shall take a softer route: use (42) to evaluate $\mathrm{D}\{h_\phi(\cdot,y)\}$ and $\langle \phi, h_\phi(\cdot,y)\rangle$ and check that the required conditions are met. We rely on the fact that $\mathrm{D}\{\rho_{\mathrm{D}^*\mathrm{D}}\}(x) = \rho_{\mathrm{D}^*}(x) = -\frac{1}{2}\mathrm{sign}(x)$. The other key observation is that the definition of $g_\phi(x,z)$ is independent of the actual choice of $\rho_{\mathrm{D}}(x) = \mathbb{1}_+(x) + c_0$, as long as it satisfies the Green's function property. In particular, we have that

$$g_\phi(x,y) = \tfrac{1}{2}\mathrm{sign}(x-y) - \langle \tfrac{1}{2}\mathrm{sign}(\cdot - y), \phi\rangle$$
$$= \tfrac{1}{2}\mathrm{sign}(x-y) - \tfrac{1}{2}(\mathrm{sign}^\vee * \phi)(y)$$
$$= \tfrac{1}{2}\mathrm{sign}(x-y) + \tfrac{1}{2}(\mathrm{sign} * \phi)(y)$$

where $f^\vee(x) = f(-x)$. Using the same kind of manipulation, we show that

$$\mathrm{D}\{h_\phi(\cdot,y)\} = \mathrm{D}\{\rho_{\mathrm{D}^*\mathrm{D}}\}(x-y) - \mathrm{D}\{\rho_{\mathrm{D}^*\mathrm{D}} * \phi\}(x)$$
$$= -\tfrac{1}{2}\mathrm{sign}(x-y) + \tfrac{1}{2}(\mathrm{sign} * \phi)(x)$$
$$= \tfrac{1}{2}\mathrm{sign}(y-x) + \tfrac{1}{2}(\mathrm{sign} * \phi)(x) = g_\phi(y,x),$$

which is the transpose of the kernel, as expected. We then invoke (27) in Theorem 8, which ensures that $g_\phi(y,\cdot) \in L_2(\mathbb{R})$ for any $y \in \mathbb{R}$. As for the boundary condition, we have that

$$\langle \phi, h_\phi(\cdot,y)\rangle = \langle \phi, \rho_{\mathrm{D}^*\mathrm{D}}(\cdot - y)\rangle - \langle \phi, \rho_{\mathrm{D}^*\mathrm{D}} * \phi\rangle - \langle \phi, \rho_{\mathrm{D}^*\mathrm{D}} * \phi(y)\rangle + \langle \phi, r_{1,1}\rangle$$

Thanks to the property that $\langle \phi, 1\rangle = 1$, this simplifies to

$$\langle \phi, h_\phi(\cdot,y)\rangle = (\phi * \rho_{\mathrm{D}^*\mathrm{D}})(y) - \langle \phi, \rho_{\mathrm{D}^*\mathrm{D}} * \phi\rangle - (\rho_{\mathrm{D}^*\mathrm{D}} * \phi)(y) + r_{1,1}$$
$$= -\langle \phi, \rho_{\mathrm{D}^*\mathrm{D}} * \phi\rangle + r_{1,1} = 0.$$

Since the reproducing kernel is unique, this proves that the provided formula is the correct one. □

Now, in the particular case where $\phi = \delta$, the expression of the reproducing kernel (42) simplifies to

$$h_\delta(x, y) = -\tfrac{1}{2}|x - y| + \tfrac{1}{2}|x| + \tfrac{1}{2}|y| = \min(|x|, |y|), \tag{43}$$

which can be recognized as the correlation function of Brownian motion. We can also easily verify that $h_\delta(\cdot, \cdot) \in C_{\mathrm{b},\alpha_0}(\mathbb{R} \times \mathbb{R})$ with $\alpha_0 = -1$. In view of Theorem 8, this suggests that the kernel of the inverse operator D_ϕ^{-1} satisfies the stability bound

$$\sup_{x \in \mathbb{R}} (1 + |x|)^{-1} \|g_\phi(x, \cdot)\|_{L_2(\mathbb{R})} < \infty. \tag{44}$$

This is not obvious a priori from (39), especially since the condition fails for the leading term $\mathbb{1}_+(x - \cdot)$ (standard LSI integrator) whose L_2 norm is unbounded for all $x \in \mathbb{R}$.

We conclude the section with a summary of these findings.

Proposition 7. *Let $\phi \in \mathcal{H}'_{\mathrm{D}}$ and $p_1 \in \mathcal{N}_{\mathrm{D}}$ be a biorthogonal pair such that $\langle \phi, p_1 \rangle = 1$. Then, the space*

$$\mathcal{H}_{\mathrm{D}} = \{f \in \mathcal{S}'(\mathbb{R}) : \|\mathrm{D}f\|_{L_2(\mathbb{R})} < \infty\}$$

admits a direct sum decomposition as $\mathcal{H}_{\mathrm{D},\phi} \oplus \mathcal{N}_{\mathrm{D}}$ where the two latter Hilbert spaces are defined by (37) and (34), respectively. Moreover, any $f \in \mathcal{H}_{\mathrm{D}}$ has a unique decomposition as

$$f = \mathrm{D}_\phi^{-1} w + q = f_\phi + q$$

where $w = \mathrm{D}f \in L_2(\mathbb{R})$, $f_\phi = \mathrm{D}_\phi^{-1} w \in \mathcal{H}_{\mathrm{D},\phi}$, $q = \langle f, \phi \rangle p_1 \in \mathcal{N}_{\mathrm{D}}$, and the inverse operator $\mathrm{D}_\phi^{-1} : L_2(\mathbb{R}) \to \mathcal{H}_{\mathrm{D},\phi}$ is defined by (39).

Finally, $\mathcal{H}_{\mathrm{D}}$ equipped with the inner product

$$\langle f, g \rangle_{\mathcal{H}_{\mathrm{D}}} = \langle \mathrm{D}f, \mathrm{D}g \rangle + \langle \phi, f \rangle \langle \phi, g \rangle,$$

is a RKHS and its reproducing kernel is the sum of the reproducing kernels of $\mathcal{H}_{\mathrm{D},\phi}$ and $\mathcal{N}_{\mathrm{D}}$ specified by (42) and (38), respectively.

2.7 Operators with non-trivial null spaces

There is a powerful association between splines and operators, the idea being that the selection of an admissible operator L specifies a corresponding type of splines [5][6, Chapter 6]. As we shall see here, we can rely on the same class of operators to specify a corresponding family of RKHS. The procedure is more involved as in Section 2.4 because of the greater difficulty of inverting operators when their null space is non-empty, similar to the previous example of the derivative. The payoff, however, is that the spaces become more interesting with a greater range of applications in a variety of disciplines (approximation theory, machine learning, stochastic processes, etc.).

Our first inclination is to define the native space associated with the differential operator L as

$$\mathcal{H}_{\mathrm{L,ext}} = \{f \in \mathcal{S}'(\mathbb{R}^d) : \|\mathrm{L}f\|^2_{L_2(\mathbb{R}^d)} < \infty\},$$

while its null space is given by

$$\mathcal{N}_{\mathrm{L,ext}} = \{q \in \mathcal{H}_{\mathrm{L,ext}} : \mathrm{L}\{q\} = 0\}.$$

This simple definition is appropriate when $\mathcal{N}_{\mathrm{L,ext}}$ is finite-dimensional, which is the case for most differential operators defined on a 1D domain ($d = 1$). Unfortunately, the situation tends to be more complicated for $d > 1$ because the *extended* null space ($\mathcal{N}_{\mathrm{L,ext}}$) of many partial differential operators (such as the laplacian) is infinite-dimensional. This forces us to constrain the native space to $\mathcal{H}_{\mathrm{L}} \subseteq \mathcal{H}_{\mathrm{L,ext}}$ such as to meet the finite-dimensional constraint

$$\mathcal{N}_{\mathrm{L}} = \{q \in \mathcal{H}_{\mathrm{L}} : \mathrm{L}\{q\} = 0\} = \mathrm{span}\{p_n\}_{n=1}^{N_0} \subseteq \mathcal{N}_{\mathrm{L,ext}}; \tag{45}$$

that is, a null space $\mathcal{N}_{\mathrm{L}}$ that admits a finite basis $\boldsymbol{p} = (p_1, \dots, p_{N_0})$ with $p_n \in \mathcal{S}'(\mathbb{R}^d)$. While this reduction from $\mathcal{H}_{\mathrm{L,ext}}$ to $\mathcal{H}_{\mathrm{L}}$ is required to specify a valid reproducing kernel, the actual restriction of the "extended" space only happens on the side of the null space so that $\mathcal{H}_{\mathrm{L}}$ remains rich enough to represent any function as closely as desired. In other words, we shall define the native space $\mathcal{H}_{\mathrm{L}}$ in a way that leaves the quotient space $\mathcal{H}_Q$ unchanged:

$$\mathcal{H}_Q = \mathcal{H}_{\mathrm{L,ext}}/\mathcal{N}_{\mathrm{L,ext}} = \mathcal{H}_{\mathrm{L}}/\mathcal{N}_{\mathrm{L}} \tag{46}$$

subject to the finite-dimensionality condition (45). This results in the abstract specification of our native space as the direct sum of two Hilbert spaces.

Proposition 8. *Let $\mathcal{H}_Q$ be the quotient space defined by* (46) *and* $\mathrm{L} : \mathcal{H}_{\mathrm{L}} = \mathcal{H}_Q \oplus \mathcal{N}_{\mathrm{L}} \to L_2(\mathbb{R}^d)$ *a linear operator whose null space $\mathcal{N}_{\mathrm{L}}$ is finite-dimensional and endowed with some inner product $\langle\cdot,\cdot\rangle_{\mathcal{N}_{\mathrm{L}}}$. Then, the native space of* L, $\mathcal{H}_{\mathrm{L}} = \mathcal{H}_Q \oplus \mathcal{N}_{\mathrm{L}}$, *is a Hilbert space for the inner product*

$$\langle f, g\rangle_{\mathcal{H}_{\mathrm{L}}} = \langle \mathrm{L}f, \mathrm{L}g\rangle_{L_2} + \langle \mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{f\}, \mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{g\}\rangle_{\mathcal{N}_{\mathrm{L}}}.$$

where $\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}$ *is a projection operator from $\mathcal{H}_{\mathrm{L}}$ into $\mathcal{N}_{\mathrm{L}}$.*

Our guiding principle will be to take $\mathcal{N}_{\mathrm{L}}$ as small as possible, but still large enough for $\mathcal{H}_{\mathrm{L}}$ to be dense in $\mathcal{S}'(\mathbb{R}^d)$. The relevant notion to achieve this controlled reduction of the null space is the property of conditional positive-definiteness, which is reviewed in Section 2.7.2. The other functional ingredient is the projection operator $\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}} : \mathcal{H}_{\mathrm{L}} \to \mathcal{N}_{\mathrm{L}}$, which, as we shall see, can be defined in a rather flexible fashion.

2.7.1 Hilbert-space structure of the null space

While the abstract representation $\mathcal{H}_{\mathrm{L}} = \mathcal{H}_Q \oplus \mathcal{N}_{\mathrm{L}}$ is a first hint that our native space for L is a RKHS, we are aiming at an explicit characterization. This requires that we invert the operator L from the left, which is feasible, but requires special care. The idea is to resolve the non-uniqueness problem by imposing N_0 linear boundary conditions to fix the null-space component.

Definition 8 (Admissible boundary functionals)**.** *Let* $\mathrm{L} : \mathcal{H}_{\mathrm{L}} \to L_2(\mathbb{R}^d)$ *be a linear operator with a finite-dimensional null space $\mathcal{N}_{\mathrm{L}}$ of dimension N_0 equipped with some norm $\|\cdot\|_{\mathcal{N}_{\mathrm{L}}}$. The linear map $\boldsymbol{\phi} : \mathcal{H}_{\mathrm{L}} \to \mathbb{R}^{N_0} : f \mapsto \boldsymbol{\phi}(f) = (\langle\phi_1, f\rangle, \dots, \langle\phi_{N_0}, f\rangle)$, which is composed of N_0 "boundary" functionals $\{\phi_n\}_{n=1}^{N_0}$, is said to be admissible if there exist two constants $B \geq A > 0$ such that*

$$A\,\|\mathrm{Proj}_{\mathcal{N}}\{f\}\|_{\mathcal{N}_{\mathrm{L}}} \leq \|\boldsymbol{\phi}(f)\|_2 \leq B\big(\|\mathrm{L}f\|_{L_2} + \|\mathrm{Proj}_{\mathcal{N}}\{f\}\|_{\mathcal{N}_{\mathrm{L}}}\big), \quad \forall f \in \mathcal{H}_{\mathrm{L}}. \tag{47}$$

In effect, this definition imposes two kind of constraints on $\boldsymbol{\phi}$: (i) the continuity of the boundary functionals, and (ii) their completeness with respect to the null space. Indeed, since $\|\mathrm{L}f\|_{L_2} + \|\mathrm{Proj}_{\mathcal{N}}\{f\}\|_{\mathcal{N}_{\mathrm{L}}}$ is a valid norm for $\mathcal{H}_{\mathrm{L}}$, the upper bound in (47) implies the boundedness (and hence the continuity) of the linear functionals $\phi_n : \mathcal{H}_{\mathrm{L}} \to \mathbb{R}$, while the converse is also true (by the triangle inequality). Hence, the abstract equivalent of the upper bound in (47) is $\phi_n \in \mathcal{H}'_{\mathrm{L}}$ for $n = 1, \dots, N_0$ where the space $\mathcal{H}'_{\mathrm{L}}$ is the continuous topological dual of $\mathcal{H}_{\mathrm{L}}$. To reveal the completeness condition (ii), we specialize the inequality for $q \in \mathcal{N}_{\mathrm{L}}$ as

$$A\|q\|_{\mathcal{N}} \leq \|\boldsymbol{\phi}(q)\|_2 \leq B\|q\|_{\mathcal{N}}, \quad \forall q \in \mathcal{N}_{\mathrm{L}}, \tag{48}$$

which is a norm equivalence reminiscent of the definition of a frame. The crucial point here is the existence of the lower bound that ensures the invertibility of the linear map $q \mapsto \boldsymbol{\phi}(q)$ (see Proposition 9 below).

To turn these abstract considerations into a concrete characterization, we now put an inner-product structure on the null space by selecting a basis $\boldsymbol{p} = (p_1, \dots, p_{N_0})$ and using it to expand $q \in \mathcal{N}_{\mathrm{L}}$ as $q = \boldsymbol{p}^T \mathbf{b}$. The corresponding norm is then given by $\|q\|_{\mathcal{N}_{\mathrm{L}}} = \|\mathbf{b}\|_2$. The most convenient design is to choose $\boldsymbol{p} = \boldsymbol{p}_{\boldsymbol{\phi}}$ such that the basis is biorthogonal to the boundary functionals. In such a scenario, $\mathbf{b} = \boldsymbol{\phi}(q)$ and we have a perfect norm equivalence in (48) with $A = B = 1$.

Definition 9 (Biorthogonal system)**.** *The pair* $(\boldsymbol{\phi}, \boldsymbol{p})$ *with* $\boldsymbol{\phi} = (\phi_1, \dots, \phi_{N_0})$ *and* $\boldsymbol{p} = (p_1, \dots, p_{N_0})$ *is called a* biorthogonal system *for the finite-dimensional subspace* $\mathcal{N}_{\mathrm{L}} = \mathrm{span}\{p_n\}_{n=1}^{N_0} \subseteq \mathcal{S}'(\mathbb{R}^d)$ *if any* $p \in \mathcal{N}_{\mathrm{L}}$ *admits a unique expansion of the form*

$$p = \sum_{n=1}^{N_0} \langle \phi_n, p \rangle p_n. \tag{49}$$

The natural norm induced by such a system is

$$\|p\|_{\mathcal{N}_{\mathrm{L}}} = \|\boldsymbol{\phi}(p)\|_2 = \left(\sum_{n=1}^{N_0} |\langle \phi_n, p \rangle|^2 \right)^{\frac{1}{2}}.$$

The unicity of the representation in Definition 9 implies that $\boldsymbol{p}$ should be a basis of $\mathcal{N}_{\mathrm{L}}$, while the validity of (49) for $p = p_n$ implies that the underlying functions should be biorthogonal; i.e.,

$$\langle \phi_m, p_n \rangle = \delta_{m,n} = \begin{cases} 1, & m = n \\ 0, & \text{otherwise.} \end{cases}$$

The existence of a such system for any choice of basis $\boldsymbol{p}$ is backed by the following result in functional analysis, which is closely related to the Hahn-Banach theorem.

Theorem 9 ([**?**, Theorem 3.5, p. 60])**.** *Let* $\mathcal{M}$ *be a subspace of a locally convex space* $\mathcal{X}$*, and* x_0 *be an element of* $\mathcal{X}$*. If* x_0 *is not in the closure of* $\mathcal{M}$*, then there exists a continuous linear functional* ϕ *on* $\mathcal{X}$ *such that* $\langle \phi, x_0 \rangle = 1$ *but* $\langle \phi, x \rangle = 0$ *for every* $x \in \mathcal{M}$*.*

One then easily proves the existence of a full biorthogonal set $\{\phi_n\}_{m=1}^{N_0}$ by successive exclusion of $x_0 = p_n$ with $\phi = \phi_n$ and $\mathcal{M} = \mathrm{span}_{m\neq n}\{p_m\}$ (the finite dimensionality of $\mathcal{M}$ and the linear independence of the p_m's ensures that $p_n \notin \overline{\mathcal{M}} = \mathcal{M}$).

Conversely, we may pick an admissible set of linear functionals $\{\phi_n\}_{n=1}^{N_0}$ and appropriately modify the basis to meet our requirements. This results in the specification of the corresponding projection operator.

Proposition 9 (Projector onto null space of L)**.** *Let $\widetilde{\boldsymbol{p}} = (\tilde{p}_1, \dots, \tilde{p}_{N_0})$ be a basis of $\mathcal{N}_{\mathrm{L}} \subseteq \mathcal{H}_{\mathrm{L}}$ and $\boldsymbol{\phi} = (\phi_1, \dots, \phi_{N_0})$ some admissible set of boundary functionals such that* (48) *(or* (47)*) is satisfied. Let us also define the $N_0 \times N_0$ cross-product matrix*

$$\mathbf{C}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}} = \langle \boldsymbol{\phi}, \widetilde{\boldsymbol{p}}^T \rangle = [\boldsymbol{\phi}(p_1) \ \cdots \ \boldsymbol{\phi}(p_{N_0})] \tag{50}$$

with entry $[\mathbf{C}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}}]_{m,n} = \langle \phi_m, \tilde{p}_n \rangle$. Then, $\mathbf{C}_{\widetilde{\boldsymbol{p}},\boldsymbol{\phi}} = \mathbf{C}^T_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}}$ is invertible and

$$\boldsymbol{p} = (p_1, \dots, p_{N_0}) = \mathbf{C}^{-1}_{\widetilde{\boldsymbol{p}},\boldsymbol{\phi}} \, \widetilde{\boldsymbol{p}} \tag{51}$$

is the unique basis of $\mathcal{N}_{\mathrm{L}}$ that is biorthogonal to $\boldsymbol{\phi}$. Furthermore, the projector of $\mathcal{H}_{\mathrm{L}}$ onto $\mathcal{N}_{\mathrm{L}}$ perpendicular to $\mathcal{N}'_{\mathrm{L}} = \mathrm{span}\{\phi_n\}_{n=1}^{N_0}$ is specified by

$$\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}} : f \mapsto \sum_{n=1}^{N_0} \langle \phi_n, f \rangle p_n,$$

with the property that $\|\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\| = 1$.

Proof. First, we observe that the matrix $\mathbf{C}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}}$ defined by (50) is well-defined, thanks to the upper bound in (48). Since $\widetilde{\boldsymbol{p}}$ is a basis of $\mathcal{N}_{\mathrm{L}}$, every $q \in \mathcal{N}_{\mathrm{L}}$ has a unique expansion $q = \sum_{n=1}^{N_0} b_n \tilde{p}_n = \widetilde{\boldsymbol{p}}^T \boldsymbol{b}$ with $\|q\|_{\mathcal{N}} = \|\boldsymbol{b}\|_2$. Therefore, by linearity, we have that

$$\boldsymbol{\phi}(q) = \begin{bmatrix} \langle \phi_1, q \rangle \\ \vdots \\ \langle \phi_{N_0}, q \rangle \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{N_0} \langle \phi_1, \tilde{p}_n \rangle b_n \\ \vdots \\ \sum_{n=1}^{N_0} \langle \phi_{N_0}, \tilde{p}_n \rangle b_n \end{bmatrix} = \mathbf{C}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}} \, \boldsymbol{b}.$$

This allows us to rewrite the norm inequality (48) as

$$A\|\boldsymbol{b}\|_2 \le \|\boldsymbol{\phi}(q)\|_2 = \|\mathbf{C}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}} \, \boldsymbol{b}\|_2 \le B\|\boldsymbol{b}\|_2,$$

which ensures that the singular values of $\mathbf{C}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}}$ (resp., $\mathbf{C}_{\widetilde{\boldsymbol{p}},\boldsymbol{\phi}} = \mathbf{C}^T_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}}$) are bounded from above and below. Since we are dealing with a square matrix,

Description	Operator	Kernel
Riesz map $\mathcal{N}'_{\mathrm{L}} \to \mathcal{N}_{\mathrm{L}}$	$\mathrm{R}_{\boldsymbol{p}}$	$\sum_{n=1}^{N_0} p_n(\boldsymbol{x})p_n(\boldsymbol{y})$
Riesz map $\mathcal{N}_{\mathrm{L}} \to \mathcal{N}'_{\mathrm{L}}$	$\mathrm{R}_{\boldsymbol{\phi}}$	$\sum_{n=1}^{N_0} \phi_n(\boldsymbol{x})\phi_n(\boldsymbol{y})$
Projector $\mathcal{H}_{\mathrm{L}} \to \mathcal{N}_{\mathrm{L}}$	$\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}$	$\sum_{n=1}^{N_0} p_n(\boldsymbol{x})\phi_n(\boldsymbol{y})$
Projector $\mathcal{H}'_{\mathrm{L}} \to \mathcal{N}'_{\mathrm{L}}$	$\mathrm{Proj}_{\mathcal{N}'_{\mathrm{L}}}$	$\sum_{n=1}^{N_0} \phi_n(\boldsymbol{x})p_n(\boldsymbol{y})$

Table 1: Complete set of operators associated with the biorthogonal system $(\boldsymbol{\phi}, \boldsymbol{p})$ with $\mathcal{N}_{\mathrm{L}} = \mathrm{span}\{p_n\}_{n=1}^{N_0}$ and $\mathcal{N}'_{\mathrm{L}} = \mathrm{span}\{\phi_n\}_{n=1}^{N_0}$.

the existence of the lower bound guarantees that $\mathbf{C}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}}$ (resp., $\mathbf{C}_{\widetilde{\boldsymbol{p}},\boldsymbol{\phi}}$) is invertible. The cross-product matrix for the new basis $\boldsymbol{p}$ defined by (51) is then given by

$$\begin{aligned}\mathbf{C}_{\boldsymbol{\phi},\boldsymbol{p}} = \langle \boldsymbol{\phi}, \boldsymbol{p}^T \rangle &= \langle \boldsymbol{\phi}, (\mathbf{C}^{-1}_{\widetilde{\boldsymbol{p}},\boldsymbol{\phi}}\, \widetilde{\boldsymbol{p}})^T \rangle \\ &= \langle \boldsymbol{\phi}, \widetilde{\boldsymbol{p}}^T \rangle \mathbf{C}^{-1}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}} = \mathbf{C}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}}\, \mathbf{C}^{-1}_{\boldsymbol{\phi},\widetilde{\boldsymbol{p}}} = \mathbf{I}_{N_0},\end{aligned}$$

which confirms that the functions $\{p_n\}$ and $\{\phi_n\}$ are biorthogonal. This property also yields $\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{p_n\} = p_n$ and, more generally, $\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{q\} = q$ for all $q \in \mathcal{N}_{\mathrm{L}}$. Likewise, we have that $\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{f\} = \mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{f\} \in \mathcal{N}_{\mathrm{L}}$ for all $f \in \mathcal{H}_{\mathrm{L}}$, which implies that the operator is a projector from $\mathcal{H}_{\mathrm{L}}$ onto $\mathcal{N}_{\mathrm{L}}$.

Let us now consider the residual $r = f - \mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{f\}$. By projecting once more, we get $\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{r\} = 0$. Due to the structure of the operator, this is equivalent to $\langle r, \phi_n \rangle = 0$ for all n, which translates into the projection error being perpendicular to $\mathrm{span}\{\phi_n\}_{n=1}^{N_0}$ for all $f \in \mathcal{H}_{\mathrm{L}}$. The boundedness of the operator simply follows from $\|\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{f\}\|_{\mathcal{N}_{\mathrm{L}}} = \|f\|_{\mathcal{N}} \le \|\mathrm{L}\{f\}\|_{L_2} + \|f\|_{\mathcal{N}_{\mathrm{L}}}$. □

In the language of RKHS, this construction is characterized as follows.

Corollary 3. *Let $(\boldsymbol{\phi}, \boldsymbol{p})$ be a biorthogonal system for $\mathcal{N}_{\mathrm{L}} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$. Then,*

$\mathcal{N}_{\rm L} = \mathrm{span}\{p_n\}_{n=1}^{N_0}$ *equipped with the inner product*

$$\langle f, g\rangle_{\mathcal{N}_{\rm L}} = \sum_{n=1}^{N_0} \langle \phi_n, f\rangle\langle \phi_n, g\rangle, \quad f, g \in \mathcal{N}_{\rm L} \tag{52}$$

is a RKHS with reproducing kernel

$$h_{\mathcal{N}_{\rm L}}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{n=1}^{N_0} p_n(\boldsymbol{x}) p_n(\boldsymbol{y}). \tag{53}$$

Likewise, we can easily show that the dual of $\mathcal{N}_{\rm L} = \mathrm{span}\{p_n\}_{n=1}^{N_0}$ is the Hilbert space $\mathcal{N}'_{\rm L} = \mathrm{span}\{\phi_n = p_n^*\}_{n=1}^{N_0}$ equipped with the inner product

$$\langle f^*, g^*\rangle_{\mathcal{N}'_{\rm L}} = \sum_{n=1}^{N_0} \langle p_n, f^*\rangle\langle p_n, g^*\rangle = \langle \mathrm{R}_{\boldsymbol{p}} f^*, g^*\rangle, \quad f^*, g^* \in \mathcal{N}'_{\rm L}$$

where $\mathrm{R}_{\boldsymbol{p}}$ is the Riesz map $\mathcal{N}'_{\rm L} \to \mathcal{N}_{\rm L}$ whose kernel is given by (53). The complementary (or reverse) Riesz map $\mathcal{N}_{\rm L} \to \mathcal{N}'_{\rm L}$ is the positive-definite operator

$$\mathrm{R}_{\boldsymbol{\phi}} : f \mapsto f^* = \sum_{n=1}^{N_0} \langle \phi_n, f\rangle \phi_n.$$

In particular, we have that $\phi_n = p_n^* = \mathrm{R}_{\boldsymbol{\phi}}\{p_n\}$, while the biorthogonality of $\boldsymbol{\phi}$ and $\boldsymbol{p}$ also implies that $\mathrm{R}_{\boldsymbol{p}}\mathrm{R}_{\boldsymbol{\phi}}\{f\} = f$ for all $f \in \mathcal{N}_{\rm L}$. Finally, we can identify the projection operator $\mathrm{Proj}_{\mathcal{N}'_{\rm L}} : \mathcal{H}'_{\rm L} \to \mathcal{N}'_{\rm L}$, which happens to be the adjoint of $\mathrm{Proj}_{\mathcal{N}_{\rm L}}$. The complete set of these null-space-related operators is summarized in Table 1.

2.7.2 Conditional positive-definiteness

The price to pay for considering a regularization operator L whose null space is non-trivial is that this prevents the reproducing kernel from being *strictly* positive definite. Instead, one has to settle for a weaker form of the property that factors out the components that compromise unicity.

Definition 10 (Conditional positive-definiteness)**.** *Let* $\mathcal{N} = \mathrm{span}\{p_n\}_{n=1}^{N_0}$ *be a finite-dimensional subspace of* $C_{\mathrm{b},\alpha}(\mathbb{R}^d)$*. Then, the kernel function* $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *is said to* $\mathcal{N}$*-conditionally positive-definite if*

$$\sum_{m=1}^{N} \sum_{n=1}^{N} z_m h(\boldsymbol{x}_m, \boldsymbol{x}_n) z_n \geq 0$$

for any $N \in \mathbb{N}$, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N \in \mathbb{R}^d$, *and* $z_1, \dots, z_N \in \mathbb{R}$, *subject to the condition*

$$\sum_{m=1}^{N} z_m p_n(\boldsymbol{x}_m) = 0$$

for $n = 1, \dots, N_0$. *The conditional positive-definiteness is said to be* strict *if*

$$\sum_{m=1}^{N} \sum_{n=1}^{N} z_m h(\boldsymbol{x}_m, \boldsymbol{x}_n) z_n > 0,$$

under the same conditions with $(z_1, \dots, z_N) \in \mathbb{R}^N \backslash \{\mathbf{0}\}$ *and the* $\boldsymbol{x}_n$ *all being distinct.*

We shall sometimes refer to this property as $\boldsymbol{p}$-conditional positivity where the vector $\boldsymbol{p} = (p_1, \dots, p_{N_0})$ represents a basis of $\mathcal{N}$. Not too surprizingly, there is also an extended version of the property that applies to general linear operators.

Definition 11 (Positive-definite operator)**.** *Let* A *be a continuous operator* $\mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ *and* $\mathcal{N}$ *some finite-dimensional subspace of* $\mathcal{S}'(\mathbb{R}^d)$ *that is spanned by* $\boldsymbol{p} = (p_1, \dots, p_{N_0})$. *The operator* A *is said to be:*

- *Symmetric or self-adjoint if, for all* $\varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R}^d)$,

$$\langle \mathrm{A}\varphi_1, \varphi_2 \rangle = \langle \mathrm{A}\varphi_2, \varphi_1 \rangle.$$

- *Positive-definite if, for any* $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\langle \mathrm{A}\varphi, \varphi \rangle \geq 0.$$

- $\mathcal{N}$*-conditionally positive-definite (or* $\boldsymbol{p}$*-conditionally positive), if*

$$\langle \mathrm{A}\varphi, \varphi \rangle \geq 0$$

  for any $\varphi \in \mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d) = \mathcal{S}(\mathbb{R}^d) \cap \mathcal{N}^{\perp} = \{\varphi \in \mathcal{S}(\mathbb{R}^d) : \boldsymbol{p}(\varphi) = \mathbf{0}\}$.

- *Strictly* $\mathcal{N}$*-conditionally positive-definite if*

$$\langle \mathrm{A}\varphi, \varphi \rangle > 0$$

  for all $\varphi \in \mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d) \backslash \{0\}$.

We conclude our discussion of positive definiteness by showing that these two definitions are equivalent when the Schwartz kernel of the operator is a bivariate function that is (separately) continuous in each argument.

Theorem 10 (Kernel of a positive-definite operator). *Let us consider a symmetric operator* $\mathrm{A} : \varphi \mapsto \int_{\mathbb{R}^d} a(\cdot, \boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$ *whose Schwartz kernel* a *is such that* $a(\boldsymbol{x}_0, \cdot) = a(\cdot, \boldsymbol{x}_0) \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ *for any* $\boldsymbol{x}_0 \in \mathbb{R}^d$ *and some* $\alpha \in \mathbb{R}$. *Then, the (*$\boldsymbol{p}$*-conditional) positive definiteness of* $\mathrm{A} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ *is equivalent to the (*$\boldsymbol{p}$*-conditional) positive-definiteness of its kernel* $a : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *in the sense of Definition 10.*

Proof. Let us start with the unconditional version of the property. For any $z_1, \cdots, z_N \in \mathbb{R}$ and $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N \in \mathbb{R}^d$, we specify the sequence of test functions $\varphi_k = \sum_{n=1}^{N} z_n u_k(\cdot - \boldsymbol{x}_n) \in \mathcal{S}(\mathbb{R}^d)$ where $u_k(\boldsymbol{x}) = k^d \mathrm{e}^{-\frac{1}{2}\|k\boldsymbol{x}\|^2}$ is a rescaled and renormalized Gaussian pulse with $\langle u_k, 1\rangle = 1$. This construction is such that u_k converges to the Dirac distribution as $k \to \infty$; i.e.,

$$\lim_{k\to\infty} \langle f, u_k(\cdot - \boldsymbol{x}_m)\rangle = \langle f, \lim_{k\to\infty} u_k(\cdot - \boldsymbol{x}_m)\rangle = f(\boldsymbol{x}_m), \tag{54}$$

which is valid for any continuous function $f : \mathbb{R}^d \to \mathbb{R}$ of slow growth. By invoking the continuity of A and the sampling relation (54), we then find that

$$\begin{aligned}
\lim_{k\to\infty} \langle \mathrm{A}\{\varphi_k\}, \varphi_k\rangle &= \langle \mathrm{A}\{\lim_{k\to\infty} \varphi_k\}, \lim_{k\to\infty} \varphi_k\rangle \\
&= \langle \mathrm{A}\left\{\sum_{m=1}^{N_0} z_m \delta(\cdot - \boldsymbol{x}_m)\right\}, \sum_{n=1}^{N_0} z_n \delta(\cdot - \boldsymbol{x}_n)\rangle \\
&= \sum_{m=1}^{N_0}\sum_{n=1}^{N_0} z_m z_n \langle \mathrm{A}\{\delta(\cdot - \boldsymbol{x}_m)\}, \delta(\cdot - \boldsymbol{x}_n)\rangle \qquad \text{(bilinearity)} \\
&= \sum_{m=1}^{N_0}\sum_{n=1}^{N_0} z_m z_n \underbrace{\mathrm{A}\{\delta(\cdot - \boldsymbol{x}_m)\}(\boldsymbol{x}_n)}_{a(\boldsymbol{x}_m, \boldsymbol{x}_n)} \geq 0,
\end{aligned}$$

which proves the direct part of the statement.

For the converse implication, we first consider the truncated integral

$$J_R(\varphi) = \int_{[-R,+R]^d} \int_{[-R,+R]^d} \varphi(\boldsymbol{x}) a(\boldsymbol{x}, \boldsymbol{y}) \varphi(\boldsymbol{y}) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y}$$

where the domain of the kernel is limited to the hypercube $[-R, R]^d \subseteq \mathbb{R}^d$ with $R > 0$. Since both a and φ are continuous, $J_R(\varphi)$ is a $2d$-dimensional

Riemann integral, which can be expressed as $J_R(\varphi) = \lim_{k\to\infty} J_{R,k}(\varphi)$ where

$$J_{R,k}(\varphi) = \left(\frac{R}{k}\right)^{2d} \sum_{\boldsymbol{m}=-(k,\dots,k)}^{(k,\dots,k)} \sum_{\boldsymbol{n}=-(k,\dots,k)}^{(k,\dots,k)} \varphi(\boldsymbol{x_m}) a(\boldsymbol{x_m}, \boldsymbol{x_n}) \varphi(\boldsymbol{x_n}) \quad (55)$$

with $\boldsymbol{x_m} = \frac{\boldsymbol{m}R}{k} \in \mathbb{R}^d$. Now, the positive definiteness of a (Definition 10) implies that $J_{R,k}(\varphi) \geq 0$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and any $k \in \mathbb{N}^+$, $R \in \mathbb{R}^+$. To have this sum converge to $\langle \mathrm{A}\varphi, \varphi\rangle$, we need to let $R \to \infty$, while ensuring that the step size (R/k) goes to zero. This is achieved by setting $R = i$ and $k = i^2$, which yields $\langle \mathrm{A}\varphi, \varphi\rangle = \lim_{i\to\infty} J_{i,i^2}(\varphi) \geq 0$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

To extend the argument to the case of $\boldsymbol{p} = (p_1, \dots, p_{N_0})$-conditional positivity, we pick a set of functions in $\mathcal{S}(\mathbb{R}^d)$, $\{\phi_n\}_{n=1}^{N_0}$, that satisfy the biorthogonality condition:

$$\langle p_m, \phi_n\rangle = \delta_{m-n},$$

which is always feasible (see Theorem 9 with $\mathcal{X} = \mathcal{S}'(\mathbb{R}^d)$ and accompanying explanations).

This allows us to express $\mathcal{S}(\mathbb{R}^d)$ as the direct sum $\mathcal{S}(\mathbb{R}^d) = \mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d) \oplus \mathcal{N}'$ with $\mathcal{N}' = \mathrm{span}\{\phi_n\}_{n=1}^{N_0} \subseteq \mathcal{S}(\mathbb{R}^d)$. Concretely, this means that any $\varphi \in \mathcal{S}(\mathbb{R}^d)$ has a unique decomposition as $\varphi = \tilde{\varphi} + \phi$ with

$$\phi = \mathrm{Proj}_{\mathcal{N}'}\{\varphi\} = \sum_{m=1}^{N_0} \langle p_m, \varphi\rangle \phi_m \in \mathcal{N}' \subseteq \mathcal{S}(\mathbb{R}^d)$$

and

$$\tilde{\varphi} = (\mathrm{Id} - \mathrm{Proj}_{\mathcal{N}'})\{\varphi\} = \varphi - \phi \in \mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d).$$

Armed with this latter projector, we then revisit the direct part of the proof by replacing φ_k by $\tilde{\varphi}_k = (\mathrm{Id} - \mathrm{Proj}_{\mathcal{N}'})\{\varphi_k\} \in \mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d)$. In the limit, we get

$$\begin{aligned}\lim_{k\to\infty} \tilde{\varphi}_k &= \lim_{k\to\infty} (\mathrm{Id} - \mathrm{Proj}_{\mathcal{N}'})\{\varphi_k\} \\ &= z_n \delta(\cdot - \boldsymbol{x}_n) - \sum_{m=1}^{N_0} \underbrace{\sum_{n=1}^{N_0} z_n p_m(\boldsymbol{x}_n)}_{=0} \phi_m = z_n \delta(\cdot - \boldsymbol{x}_n)\end{aligned}$$

where the double sum vanishes as a consequence of our hypothesis on z_n, so that the conclusion remains the same.

Conversely, in the conditional scenario of Definition 10, the positivity of the bilinear form (55) with $R = i$ and $k = i^2$ only holds for the subclass of test functions $\tilde{\varphi}$ such that, for $n = 1, \dots, N_0$,

$$\sum_{\boldsymbol{m}=-(i^2,\dots,i^2)}^{(i^2,\dots,i^2)} \underbrace{\tilde{\varphi}(\boldsymbol{x}_{\boldsymbol{m}})}_{z_{\boldsymbol{m}}} p_n(\boldsymbol{x}_{\boldsymbol{m}}) = 0,$$

with $\boldsymbol{x}_{\boldsymbol{m}} = \frac{\boldsymbol{m}}{i} \in \mathbb{R}^d$. As i increases to infinity, the above sum converges to the Riemann integral

$$\int_{\mathbb{R}^d} \tilde{\varphi}(\boldsymbol{x}) p_n(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = 0,$$

which is equivalent to $\tilde{\varphi} \in \mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d)$.

□

2.7.3 Admissible regularization operators

The regularization operators that are admissible for our construction are those that admit a Green's function of slow growth, subject to some conditional positivity constraint.

Definition 12. *The kernel* $G_{\mathrm{L}} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *is a Green's function of* L *if* $\mathrm{L}\{G_{\mathrm{L}}(\cdot, \boldsymbol{y})\} = \delta(\cdot - \boldsymbol{y})$*, or equivalently, if* $\mathrm{L}\mathrm{L}^{-1}\{\varphi\} = \varphi$ *for any* $\varphi \in \mathcal{S}(\mathbb{R}^d)$ *where* $\mathrm{L}^{-1} : \varphi \mapsto \int_{\mathbb{R}^d} G_{\mathrm{L}}(\cdot, \boldsymbol{y}) \varphi(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}$.

In other words, knowing the Green's function of L is equivalent to having a right-inverse of L at our disposal. It is generally not unique as we may construct many equivalent instances of the form $\widetilde{G}_{\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y}) = G_{\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y}) + q_{\boldsymbol{y}}(\boldsymbol{x})$ with $q_{\boldsymbol{y}} \in \mathcal{N}_{\mathrm{L}}$ for any fixed value $\boldsymbol{y} \in \mathbb{R}^d$.

Definition 13 (Admissible operator)**.** *A linear operator* $\mathrm{L} : \mathcal{H}_{\mathrm{L}} \to L_2(\mathbb{R}^d)$ *is called* spline-admissible *if there exists a symmetric kernel* $G_{\mathrm{L}^*\mathrm{L}} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$*, a finite-dimensional subspace* $\mathcal{N}_{\mathrm{L}} = \mathrm{span}\{p_n\}_{n=1}^{N_0}$ *and an order* $\alpha \in \mathbb{R}$ *of algebraic growth such that :*

1. $G_{\mathrm{L}^*\mathrm{L}}$ *is a Green's function of* $(\mathrm{L}^*\mathrm{L})$ *with the property that*

$$\begin{aligned} \mathrm{L}\{G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y})\}(\boldsymbol{x}) &= G_{\mathrm{L}^*}(\boldsymbol{x}, \boldsymbol{y}) = G_{\mathrm{L}}(\boldsymbol{y}, \boldsymbol{x}) \\ \mathrm{L}^*\mathrm{L}\{G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y})\} &= \mathrm{L}^*\{G_{\mathrm{L}^*}(\cdot, \boldsymbol{y})\} = \delta(\cdot - \boldsymbol{y}) \end{aligned} \tag{56}$$

   where δ *is the Dirac distribution and* $G_{\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y})$ *a Green's function of* L.

2. *Null-space property:* $\mathrm{L}\{q\} = 0$ *for all* $q \in \mathcal{N}_{\mathrm{L}} \subseteq \mathcal{H}_{\mathrm{L}}$.

3. *$G_{\mathrm{L}^*\mathrm{L}}$ is strictly $\mathcal{N}_{\mathrm{L}}$-conditionally positive-definite.*

4. *Continuity and polynomial growth: $G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y}_0) \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ for any $\boldsymbol{y}_0 \in \mathbb{R}^d$ and $\mathcal{N}_{\mathrm{L}} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$.*

5. *Boundedness on the diagonal: $\sup_{\boldsymbol{x}\in\mathbb{R}^d} |G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x}, \boldsymbol{x})| \, (1 + \|\boldsymbol{x}\|)^{2\alpha} < \infty$.*

The scenario of greatest practical interest is when the operator L is linear shift-invariant (LSI); that is, when $\mathrm{L}\{f(\cdot - \boldsymbol{x}_0)\} = \mathrm{L}\{f\}(\cdot - \boldsymbol{x}_0)$ for any function $f \in \mathcal{H}_{\mathrm{L}}$. In such a case, we can determine the canonical Green's kernels of L in Definition 13 by (generalized) inverse Fourier transformation. Specifically, we have that

$$G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y}) = \rho_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x} - \boldsymbol{y}) \quad \text{with} \quad \rho_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x}) = \mathcal{F}^{-1}\left\{\frac{1}{|\widehat{L}(\boldsymbol{\omega})|^2}\right\}(\boldsymbol{x})$$

$$G_{\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y}) = \rho_{\mathrm{L}}(\boldsymbol{x} - \boldsymbol{y}) \quad \text{with} \quad \rho_{\mathrm{L}}(\boldsymbol{x}) = \mathcal{F}^{-1}\left\{\frac{1}{\widehat{L}(\boldsymbol{\omega})}\right\}(\boldsymbol{x}),$$

where $\widehat{L}(\boldsymbol{\omega}) = \mathcal{F}\big\{\mathrm{L}\{\delta\}\big\}(\boldsymbol{\omega})$ is the frequency response of L. Due to the one-to-one relation between the shift-invariant kernel G_{L} and ρ_{L}, we shall therefore also refer to ρ_{L} as the Green's function of L, with a slight abuse of language.

Likewise, we may infer that the composition of the null space of a LSI operator is determined by the zeros of $\widehat{L}(\boldsymbol{\omega})$. Specifically, each zero $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ of order $\boldsymbol{n}_0$ contributes a series of exponential polynomials of the form $\boldsymbol{x}^{\boldsymbol{m}} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{\omega}_0, \boldsymbol{x}\rangle}$ with $\boldsymbol{m} < \boldsymbol{n}_0$. This also implies that the frequency response of an admissible operator L can only have a finite number of zeros.

Example 4. *The prototypical example of an admissible operator for $d = 1$ is the mth-order derivative operator D^m, which is LSI. Its Fourier symbol $(\mathrm{j}\omega)^m$ has an mth-order zero at $\omega = 0$. Its causal Green's function is the one-sided power function*

$$\mu_m(x) = \frac{x_+^{m-1}}{(m-1)!}$$

with $x_+ = \max(0, x)$, while the canonical solution is

$$\rho_{\mathrm{D}^m}(x) = \mathcal{F}^{-1}\left\{\frac{1}{(\mathrm{j}\omega)^m}\right\}(x) = \tfrac{1}{2}\mathrm{sign}(x)\frac{x^{m-1}}{(m-1)!}$$

The null space of D^m *is the space of polynomials of degree* $(m-1)$ *with* $N_0 = m$*; that is,*

$$\mathcal{N}_{\mathrm{D}^m} = \mathrm{span}\{p_n\}_{n=1}^{N_0} \text{ with } p_n(x) = \frac{x^{n-1}}{(n-1)!}.$$

Finally, the corresponding Green's function in Definition 13 is

$$G(x,y) = \rho(x-y) \quad \text{with} \quad \rho(x) = \mathcal{F}^{-1}\left\{\frac{1}{|\omega|^{2m}}\right\}(x) = \frac{(-1)^m}{2}\frac{|x|^{2m-1}}{(2m-1)!},$$

which happens to be $\mathcal{N}_{\mathrm{D}^m}$*-conditionally positive definite.*

One can also specify a corresponding dual basis

$$\{\phi_n\}_{n=1}^{N_0} \quad \text{with} \quad \phi_n = \delta^{(n-1)},$$

which is such that $\langle \phi_n, p_{n'}\rangle = (-1)^{(n-1)}\mathrm{D}^{(n-1)}p_{n'}(0) = \delta_{n-n'}$ *(biorthogonality property). While there are many other possible choices of dual bases, the proposed one is special as it is composed of point distributions entirely located at the origin.*

2.7.4 RKHS associated with an admissible operator

Using the same biothogonal pair $(\boldsymbol{\phi}, \boldsymbol{p})$ as in Section 2.7.1, we define the "orthogonal" complement of $\mathcal{N}_{\mathrm{L}}$ in $\mathcal{H}_{\mathrm{L}}$ as

$$\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}} = \big\{f \in \mathcal{H}_{\mathrm{L}} : \boldsymbol{\phi}(f) = \mathbf{0}\big\}, \tag{57}$$

which will now be used to specify a proper right inverse of the operator $\mathrm{L} : \mathcal{H}_{\mathrm{L}} \to L_2(\mathbb{R}^d)$. We may think of this space as a "concrete" transcription[2] of the quotient space $\mathcal{H}_{\mathrm{L,ext}}/\mathcal{N}_{\mathrm{L,ext}}$ alluded to in the proof of Proposition 8. Note, however, that this association is not unique: there is a whole family of spaces $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ with a corresponding inverse operator $\mathrm{L}_{\boldsymbol{\phi}}^{-1}$, each instance being associated with a specific $\boldsymbol{\phi}$.

Theorem 11 (Construction of stable right-inverse operator)**.** *Let* L *be an admissible operator and* $(\boldsymbol{\phi}, \boldsymbol{p})$ *with* $\boldsymbol{p} = (p_1, \dots, p_{N_0})$ *a corresponding biorthogonal system for* $\mathcal{N}_{\mathrm{L}}$*. Then,* $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ *defined by* (57) *is a Hilbert space equipped*

[2] While the nature of the elements of the two spaces is different—that is, functions f vs. equivalence classes of functions $(f + \mathcal{N}_{\mathrm{L}})$—they are isometrically isomorphic, and hence, topologically equivalent.

with the inner product $\langle f, g\rangle_{\mathrm{L}} = \langle \mathrm{L}\{f\}, \mathrm{L}\{g\}\rangle$. *Moreover, there exists an isometric map* $\mathrm{L}_{\boldsymbol{\phi}}^{-1} : L_2(\mathbb{R}^d) \to \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ *such that*

$$\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}} = \big\{ f = \mathrm{L}_{\boldsymbol{\phi}}^{-1} w : w \in L_2(\mathbb{R}^d) \big\}. \tag{58}$$

The operator $\mathrm{L}_{\boldsymbol{\phi}}^{-1}$ *is uniquely specified through the following properties*

1. *Right-inverse property:* $\mathrm{L}\mathrm{L}_{\boldsymbol{\phi}}^{-1} w = w$ *for all* $w \in L_2(\mathbb{R}^d)$
2. *Boundary conditions:* $\langle \boldsymbol{\phi}, \mathrm{L}_{\boldsymbol{\phi}}^{-1} w\rangle = \mathbf{0}$ *for all* $w \in L_2(\mathbb{R}^d)$

and its kernel is given by

$$g_{\boldsymbol{\phi}}(\boldsymbol{x}, \boldsymbol{y}) = G_{\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y}) - \sum_{n=1}^{N_0} p_n(\boldsymbol{x}) q_n(\boldsymbol{y}), \tag{59}$$

with G_{L} *such that* $\mathrm{L}\{G_{\mathrm{L}}(\cdot, \boldsymbol{y})\} = \delta(\cdot - \boldsymbol{y})$ *(Green's function property) and*

$$q_n(\boldsymbol{y}) = \langle \phi_n, G_{\mathrm{L}}(\cdot, \boldsymbol{y})\rangle. \tag{60}$$

Finally, $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}} \subseteq C_{\mathrm{b},\alpha_0}(\mathbb{R}^d)$ *if and only if* $g_{\boldsymbol{\phi}}$ *satisfies the stability condition*

$$\sup_{\boldsymbol{x}\in\mathbb{R}^d} (1 + \|\boldsymbol{x}\|)^{\alpha_0} \|g_{\boldsymbol{\phi}}(\boldsymbol{x}, \cdot)\|_{L_2(\mathbb{R}^d)} < \infty.$$

Proof. We start by proving that $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ equipped with the inner product $\langle f_1, f_2\rangle_{\mathrm{L}} = \langle \mathrm{L}f_1, \mathrm{L}f_2\rangle$ is a Hilbert space. Thanks to the linearity of L, one immediately deduces that $\langle\cdot,\cdot\rangle_{\mathrm{L}}$ satisfies the easy properties of an inner product: linearity, symmetry and non-negativity. We now show that the only feasible solution for $\langle \mathrm{L}f_0, \mathrm{L}f_0\rangle = 0$ with $f_0 \in \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ is trivial. The constraint on the L_2-norm is equivalent to $\mathrm{L}f_0 = 0$ (almost everywhere) which restricts the possible solutions in $\mathcal{H}_{\mathrm{L}}$ to $f_0 \in \mathcal{N}_{\mathrm{L}}$. We then use the condition $\boldsymbol{\phi}(f_0) = \mathbf{0}$ to project the solution set on $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$. This yields $f_0 = \sum_{n=1}^{N_0} \langle \phi_n, f_0\rangle p_n = 0$ where $\{p_n\}_{n=1}^{N_0}$ is the unique basis of $\mathcal{N}_{\mathrm{L}}$ that is biorthogonal to $\boldsymbol{\phi}$, which proves that $\langle f_0, f_0\rangle_{\mathrm{L}} = 0 \Leftrightarrow f_0 = 0$.

The idea then is to first establish Properties 1) and 2) of the operator $\mathrm{L}_{\boldsymbol{\phi}}^{-1}$ on Schwartz's space of smooth and rapidly-decreasing signals $\mathcal{S}(\mathbb{R}^d)$ to avoid any technical problems related to the splitting of the sum and the interchange of integrals. Since the space $\mathcal{S}(\mathbb{R}^d)$ equipped with the standard Schwartz-Fréchet topology is dense in $L_2(\mathbb{R}^d)$, we are then able to extend the properties by continuity.

To that end, we introduce the operator $\mathrm{G} : \varphi \mapsto \int_{\mathbb{R}^d} G_{\mathrm{L}}(\cdot, \boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$, which is well defined over $\mathcal{S}(\mathbb{R}^d)$ as long as $G_{\mathrm{L}}(\cdot,\cdot) \in \mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ (by

Schwartz's kernel theorem). Under the hypothesis that $w \in \mathcal{S}(\mathbb{R}^d)$, we then rewrite $f = \mathrm{L}_{\boldsymbol{\phi}}^{-1} w$ as

$$f = \mathrm{L}_{\boldsymbol{\phi}}^{-1}\{w\} = \mathrm{G}\{w\} - \sum_{n=1}^{N_0} p_n \langle q_n, w \rangle.$$

Next, we apply the operator L, which yields

$$\begin{aligned} \mathrm{L}\mathrm{L}_{\boldsymbol{\phi}}^{-1}\{w\} = \mathrm{L}\{f\} = \mathrm{L}\left\{ \int_{\mathbb{R}^d} w(\boldsymbol{y}) G_{\mathrm{L}}(\cdot, \boldsymbol{y}) \mathrm{d}\boldsymbol{y} \right\} &- \sum_{n=1}^{N_0} \underbrace{\mathrm{L}\{p_n\}}_{=0} \langle q_n, w \rangle \\ = \int_{\mathbb{R}^d} w(\boldsymbol{y}) \mathrm{L}\{G_{\mathrm{L}}(\cdot, \boldsymbol{y})\} \mathrm{d}\boldsymbol{y} & \qquad \text{(by linearity)} \\ = \int_{\mathbb{R}^d} w(\boldsymbol{y}) \delta(\cdot - \boldsymbol{y}) \mathrm{d}\boldsymbol{y} = w & \end{aligned}$$

where we have used the defining property $\mathrm{L}\{G_{\mathrm{L}}(\cdot, \boldsymbol{y})\} = \delta(\cdot - \boldsymbol{y})$ of the Green's function and $\mathrm{L}\{p_n\} = 0$ for $n = 1, \dots, N_0$. In particular, this implies that

$$\|\mathrm{L}_{\boldsymbol{\phi}}^{-1}\{w\}\|_{\mathrm{L}}^2 = \langle \mathrm{L}_{\boldsymbol{\phi}}^{-1}\{w\}, \mathrm{L}_{\boldsymbol{\phi}}^{-1}\{w\} \rangle_{\mathrm{L}} = \|w\|_{L_2(\mathbb{R}^d)}^2 \tag{61}$$

for all $w \in \mathcal{S}(\mathbb{R}^d)$, which shows that $\mathrm{L}_{\boldsymbol{\phi}}^{-1}$ is bounded in the L_2 norm.

As for the boundary conditions, we first observe that

$$\begin{aligned} q_n(\boldsymbol{y}) &= \langle G_{\mathrm{L}}(\cdot, \boldsymbol{y}), \phi_n \rangle \\ &= \int_{\mathbb{R}^d} G_{\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y}) \phi_n(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \mathrm{G}^*\{\phi_n\}(\boldsymbol{y}) \end{aligned}$$

where G^* is the adjoint of G. We then make use of the biorthogonality property $\langle \phi_m, p_n \rangle = \delta_{m-n}$ to evaluate the inner product of ϕ_m with $\mathrm{L}_{\boldsymbol{\phi}}^{-1} w$ as

$$\begin{aligned} \langle \phi_m, \mathrm{L}_{\boldsymbol{\phi}}^{-1}\{w\} \rangle &= \langle \phi_m, \mathrm{G}\{w\} \rangle - \sum_{n=1}^{N_0} \langle \phi_m, p_n \rangle \langle q_n, w \rangle \\ &= \langle \phi_m, \mathrm{G}\{w\} \rangle - \langle q_{\mathrm{L},m}, w \rangle \\ &= \langle \mathrm{G}^*\{\phi_m\}, w \rangle - \langle \mathrm{G}^*\{\phi_m\}, w \rangle = 0, \end{aligned}$$

which shows that the boundary conditions are satisfied. In doing so, we have effectively shown that $\mathrm{L}_{\boldsymbol{\phi}}^{-1}$ continuously maps $\mathcal{S}(\mathbb{R}^d)$ into $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$.

Next, we invoke the Hahn-Banach theorem in conjunction with the L_2 bound (61) to extend the domain of the operator to all of $L_2(\mathbb{R}^d)$. By

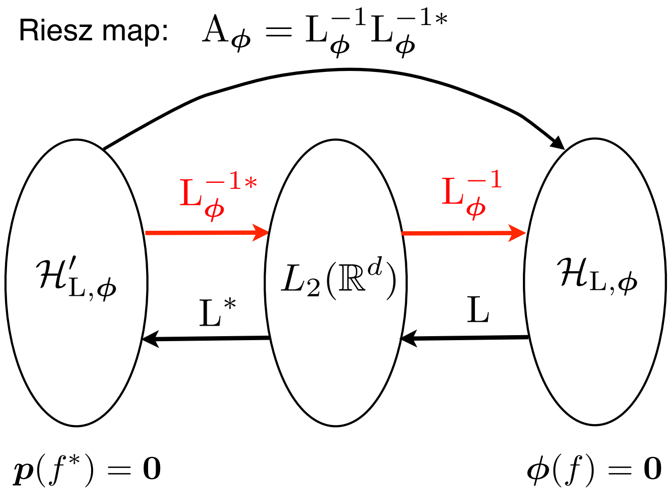

Figure 1: Factorization of the Riesz map and schematic representation of the underlying operators and Hilbert spaces.

recalling that $\mathcal{S}(\mathbb{R}^d)$ is dense in $L_2(\mathbb{R}^d)$, we then extend the boundary conditions for $w \in L_2(\mathbb{R}^d)$ by continuity. This establishes an isometric (and reversible) mapping between $L_2(\mathbb{R}^d)$ and $\mathcal{H}_{L,\boldsymbol{\phi}}$, and allows us to conclude that $L_{\boldsymbol{\phi}}^{-1} : L_2(\mathbb{R}^d) \to \mathcal{H}_{L,\boldsymbol{\phi}}$ is a stable inverse of the operator $L : \mathcal{H}_{L,\boldsymbol{\phi}} \to L_2(\mathbb{R}^d)$. Finally, we make the connection with Theorem 8 by identifying $L^{-1} = L_{\boldsymbol{\phi}}^{-1}$ as the unique inverse of L that factorizes the reproducing kernel of $\mathcal{H}_{L,\boldsymbol{\phi}}$ through $\mathcal{H}_0 = L_2(\mathbb{R}^d)$. This together with the assumption that $\mathcal{H}_{L,\boldsymbol{\phi}} \subseteq C_{b,\alpha_0}(\mathbb{R}^d)$ then yields the stability bound on $g_{\boldsymbol{\phi}}(\cdot,\cdot)$. The converse implication is supported by Theorem 7. □

2.7.5 Determination of the reproducing kernel

Having characterized the relevant right-inverse operator, we can now invoke the second part of Theorem 8 to deduce that $\mathcal{H}_{L,\boldsymbol{\phi}}$ is a RKHS whose reproducing kernel is the generalized impulse response of the composed operator $A_{\boldsymbol{\phi}} = L_{\boldsymbol{\phi}}^{-1} L_{\boldsymbol{\phi}}^{-1*}$. We shall actually take the argument one step further by expressing this kernel in terms of the symmetric Green's function of L^*L in Definition 13.

Theorem 12. *The reproducing kernel of the Hilbert space $\mathcal{H}_{L,\boldsymbol{\phi}}$ specified in Theorem 11 is $a_{\boldsymbol{\phi}}(\boldsymbol{x},\boldsymbol{y}) = \int_{\mathbb{R}^d} g_{\boldsymbol{\phi}}(\boldsymbol{x},\boldsymbol{z}) g_{\boldsymbol{\phi}}(\boldsymbol{y},\boldsymbol{z}) \mathrm{d}\boldsymbol{z}$ where $g_{\boldsymbol{\phi}}$ is given by*

(59). *Moreover, if* $G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x},\boldsymbol{y})$ *is the symmetric Green's function of* $\mathrm{L}^*\mathrm{L}$ *of Definition 13, then the reproducing kernel can be expressed as*

$$a_{\boldsymbol{\phi}}(\boldsymbol{x},\boldsymbol{y}) = G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x},\boldsymbol{y}) - \sum_{n=1}^{N_0} p_n(\boldsymbol{x})v_n(\boldsymbol{y}) - \sum_{n=1}^{N_0} v_n(\boldsymbol{x})p_n(\boldsymbol{y}) + \sum_{m=1}^{N_0}\sum_{n=1}^{N_0} r_{m,n}p_m(\boldsymbol{x})p_n(\boldsymbol{y}) \quad (62)$$

with

$$v_n(\boldsymbol{y}) = \langle \phi_n, G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y})\rangle = \int_{\mathbb{R}^d} \phi_n(\boldsymbol{z})G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{z},\boldsymbol{y})\mathrm{d}\boldsymbol{z} \quad (63)$$

$$r_{m,n} = \langle \phi_m \otimes \phi_n, G_{\mathrm{L}^*\mathrm{L}}\rangle = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \phi_m(\boldsymbol{x})\phi_n(\boldsymbol{y})G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y}, \quad (64)$$

where the integrals on the r.h.s. are symbolic representations of the underlying linear functionals.

Proof. The first statement directly follows from Theorem 8 with $\mathrm{L}^{-1} = \mathrm{L}_{\boldsymbol{\phi}}^{-1}$ and $g(\cdot,\cdot) = g_{\boldsymbol{\phi}}(\cdot,\cdot)$. Hence, we only need to establish the validity of (62). By invoking the symmetry of $G_{\mathrm{L}^*\mathrm{L}}(\cdot,\cdot)$ and the linearity of L, we first calculate the quantity

$$\begin{aligned}
\mathrm{L}\{v_n\}(\boldsymbol{x}) &= \mathrm{L}\Big\{\int_{\mathbb{R}^d} \phi_n(\boldsymbol{z})G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{z})\mathrm{d}\boldsymbol{z}\Big\}(\boldsymbol{x})\\
&= \int_{\mathbb{R}^d} \phi_n(\boldsymbol{z})\mathrm{L}\big\{G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{z})\big\}(\boldsymbol{x})\mathrm{d}\boldsymbol{z}\\
&= \int_{\mathbb{R}^d} \phi_n(\boldsymbol{z})G_{\mathrm{L}^*}(\boldsymbol{x},\boldsymbol{z})\mathrm{d}\boldsymbol{z} = \langle \phi_n, G_{\mathrm{L}^*}(\boldsymbol{x},\cdot)\rangle\\
&= \langle \phi_n, G_{\mathrm{L}}(\cdot,\boldsymbol{x})\rangle = q_n(\boldsymbol{x})
\end{aligned}$$

where q_n is defined by (60). Thanks to this identity, we then evaluate

$$\begin{aligned}
\mathrm{L}\{a_{\boldsymbol{\phi}}(\cdot,\boldsymbol{y})\}(\boldsymbol{x}) &= \mathrm{L}\{G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y})\}(\boldsymbol{x}) - \sum_{n=1}^{N_0} \mathrm{L}\{v_n\}(\boldsymbol{x})\\
&= G_{\mathrm{L}^*}(\boldsymbol{x},\boldsymbol{y}) - \sum_{n=1}^{N_0} q_n(\boldsymbol{x})p_n(\boldsymbol{y}),\\
&= G_{\mathrm{L}}(\boldsymbol{y},\boldsymbol{x}) - \sum_{n=1}^{N_0} q_n(\boldsymbol{x})p_n(\boldsymbol{y}) = g_{\boldsymbol{\phi}}(\boldsymbol{y},\boldsymbol{x}),
\end{aligned}$$

which is the transposed version of (59), as expected. The additional ingredient is $g_{\boldsymbol{\phi}}(\boldsymbol{y},\cdot) = \mathrm{L}_{\boldsymbol{\phi}}^{-1*}\{\delta(\cdot - \boldsymbol{y})\} \in L_2(\mathbb{R})$ for any $\boldsymbol{y} \in \mathbb{R}^d$, which follows from the characterization of the inverse operator in Theorem 11. To verify that $a_{\boldsymbol{\phi}}(\cdot, \boldsymbol{y})$ with $\boldsymbol{y}$ fixed satisfies the boundary conditions, we first observe that

$$\langle \phi_m, v_n \rangle = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi_n(\boldsymbol{x})\phi_n(\boldsymbol{z}) G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x}, \boldsymbol{z}) \mathrm{d}\boldsymbol{z} \mathrm{d}\boldsymbol{x} = r_{m,n}$$

for $m, n = 1, \dots, N_0$. This helps us evaluate

$$\begin{aligned}
\langle \phi_m, a_{\boldsymbol{\phi}}(\cdot, \boldsymbol{y})\}\rangle &= \langle \phi_m, G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y})\rangle - \sum_{n=1}^{N_0} \langle \phi_m, p_n\rangle \langle \phi_n, G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y})\rangle \\
&\quad - \sum_{n=1}^{N_0} \langle \phi_m, v_n\rangle p_n(\boldsymbol{y}) + \sum_{n'=1}^{N_0}\sum_{n=1}^{N_0} r_{n',n} \langle \phi_m, p_{n'}\rangle p_n(\boldsymbol{y}) \\
&= \langle \phi_m, G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y})\rangle - \langle \phi_m, G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y})\rangle \\
&\quad - \sum_{n=1}^{N_0} r_{m,n} p_n(\boldsymbol{y}) + \sum_{n=1}^{N_0} r_{m,n} p_n(\boldsymbol{y}) = 0,
\end{aligned}$$

where we have used the biorthogonality property $\langle \phi_m, p_n\rangle = \delta_{m-n}$ to simplify the sums. Since the reproduction kernel is the unique bivariate function that satisfies these properties, we have proved that (62) is the correct formula.

□

The final ingredient to complete the picture in Figure 1 is the characterization of the Hilbert space $\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$, which is the continuous dual of $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$. The interesting twist is that, contrary to $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ which stands for a whole family of spaces, there is actually a single space $\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$ that is independent of $\boldsymbol{\phi}$.

Proposition 10. *Let* L *be an admissible operator and* $(\boldsymbol{\phi}, \boldsymbol{p})$ *a corresponding biorthogonal system for* $\mathcal{N}_{\mathrm{L}}$*. Then, the continuous dual of the RKHS* $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ *is the Hilbert space*

$$\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}} = \{f^* = \mathrm{L}^* w : w \in L_2(\mathbb{R}^d)\}$$

equipped with the inner product

$$\langle f, g\rangle_{\mathcal{H}'} = \langle \mathrm{L}_{\boldsymbol{\phi}}^{-1*} f, \mathrm{L}_{\boldsymbol{\phi}}^{-1*} g\rangle_{L_2(\mathbb{R}^d)} = \langle \mathrm{A} f^*, g^*\rangle$$

where $\mathrm{L}_{\boldsymbol{\phi}}^{-1*}$ *is the adjoint of the stable inverse operator defined in Theorem 11 and* $\mathrm{A} : \varphi \mapsto \int_{\mathbb{R}^d} G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$ *where is* $G_{\mathrm{L}^*\mathrm{L}}$ *is the symmetric Green's* *function of* $(\mathrm{L}^*\mathrm{L})$ *specified in Definition 13. The operator* $(\mathrm{L}^*\mathrm{L})$ *is the Riesz map* $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}} \to \mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$ *so that any* $f^* = \mathrm{L}^*\mathrm{L}\{f\} \in \mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$ *can be viewed as the Riesz conjugate of some corresponding* $f \in \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$*. Finally, we have the "orthogonality" property*

$$\boldsymbol{p}(f^*) = \mathbf{0} \quad \Leftrightarrow \quad \boldsymbol{p}\big(\mathrm{L}^*\mathrm{L}\{f\}\big) = \mathbf{0} \tag{65}$$

for any $f^* \in \mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$ *and/or* $f \in \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$*.*

Proof. Since $\mathrm{L}_{\boldsymbol{\phi}}^{-1}$ is a right inverse of L, we immediately deduce $\mathrm{L}_{\boldsymbol{\phi}}^{-1*}$ is a left inverse of L^*. This establishes the isometric isomorphism between $\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$ and the pivot space $L_2(\mathbb{R}^d)$, which is itself isomorphic to $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ (see Figure 1). The conjugate relation between f^* and f follows from Riesz' representation theorem (Theorem 4). To establish the orthogonality property, we simply note that

$$\langle p_n, f^*\rangle = \langle p_n, \mathrm{L}^* w\rangle = \langle \mathrm{L}p_n, w\rangle = 0$$

where $w = \mathrm{L}_{\boldsymbol{\phi}}^{-1*} f^* \in L_2(\mathbb{R}^d)$. This latter property is crucial as it implies that the inner product $\langle f, g\rangle_{\mathcal{H}'} = \langle \mathrm{L}_{\boldsymbol{\phi}}^{-1*} f, \mathrm{L}_{\boldsymbol{\phi}}^{-1*} g\rangle_{L_2(\mathbb{R}^d)}$ is independent of $\boldsymbol{\phi}$. Indeed, under the Green's function assumption (56), we have that

$$\begin{aligned}
\mathrm{L}_{\boldsymbol{\phi}}^{-1*}\{f^*\} &= \int_{\mathbb{R}^d} G_{\mathrm{L}}(\boldsymbol{y}, \cdot) f^*(\boldsymbol{y}) \mathrm{d}\boldsymbol{y} - \sum_{n=1}^{N_0} q_n \overbrace{\int_{\mathbb{R}^d} p_n(\boldsymbol{y}) f^*(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}}^{=0} \\
&= \int_{\mathbb{R}^d} G_{\mathrm{L}^*}(\cdot, \boldsymbol{y}) f^*(\boldsymbol{y}) \mathrm{d}\boldsymbol{y} = \mathrm{G}^*\{f^*\}
\end{aligned}$$

for all $f^* \in \mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$. Likewise, using the explicit form of the symmetric kernel $a_{\boldsymbol{\phi}}$ in Theorem 12, we readily verify that

$$\begin{aligned}
\mathrm{L}_{\boldsymbol{\phi}}^{-1}\mathrm{L}_{\boldsymbol{\phi}}^{-1*}\{f^*\} &= \int_{\mathbb{R}^d} G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{y}, \cdot) f^*(\boldsymbol{y}) \mathrm{d}\boldsymbol{y} - \sum_{n=1}^{N_0} q_n \overbrace{\langle p_n, f^*\rangle}^{=0} \\
&\quad - \sum_{n=1}^{N_0} p_n \langle q_n, f^*\rangle + \sum_{n=1}^{N_0}\sum_{m=1}^{N_0} r_{m,n} p_n \overbrace{\langle p_m, f^*\rangle}^{=0} \\
&= \int_{\mathbb{R}^d} G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y}) f^*(\boldsymbol{y}) \mathrm{d}\boldsymbol{y} - \sum_{n=1}^{N_0} p_n \langle q_n, f^*\rangle,
\end{aligned}$$

where the null-space component on the left-hand side of the resulting expression should be interpreted as a finite-part correction of the primary integral.

L	ρ_{L}	$\rho_{\mathrm{L}^*\mathrm{L}}$	N_0	$\{(p_n,\phi_n)\}_{n=1}^{N_0}$
D	$\frac{1}{2}\mathrm{sign}(x)$	$-\frac{1}{2}\lvert x\rvert$	1	$\{(p_1(x)=1,\phi_1=\delta)\}$
D^m	$\frac{1}{2}\mathrm{sign}(x)\frac{x^{m-1}}{(m-1)!}$	$\frac{(-1)^m}{2}\frac{\lvert x\rvert^{2m-1}}{(2m-1)!}$	m	$\left\{\left(\frac{x^{n-1}}{(n-1)!},\delta^{(n-1)}\right)\right\}$

Table 2: Differential operators encountered in spline theory with associated Green's functions and biorthogonal systems.

This allows us to conclude that

$$\langle f^*,g^*\rangle_{\mathcal{H}'}=\langle \mathrm{A}f^*,g^*\rangle-\sum_{n=1}^{N_0}\overbrace{\langle p_n,g^*\rangle}^{=0}\langle q_n,f^*\rangle=\langle \mathrm{A}f^*,g^*\rangle.$$

where $\mathrm{A}:\varphi\mapsto\int_{\mathbb{R}^d}G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$. □

Let us now briefly discuss the selection of the appropriate bivariate function $G_{\mathrm{L}^*\mathrm{L}}:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ in Theorem 12. Since $G_{\mathrm{L}^*\mathrm{L}}$ is a Green's function of $(\mathrm{L}^*\mathrm{L})$, we have that

$$\begin{aligned}\mathrm{L}\{G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y})\}&=G_{\mathrm{L}^*}(\cdot,\boldsymbol{y})\\ \mathrm{L}^*\{G_{\mathrm{L}^*}(\cdot,\boldsymbol{y})\}&=\mathrm{L}^*\mathrm{L}\{G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y})\}=\delta(\cdot-\boldsymbol{y}),\end{aligned}$$

which implies that $G_{\mathrm{L}^*}(\boldsymbol{x},\boldsymbol{y})=\mathrm{L}\{G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y})\}(\boldsymbol{x})$ is a Green's function of the adjoint operator L^*. Besides the symmetry of $G_{\mathrm{L}^*\mathrm{L}}$, the enabling condition for (62) to hold is that $\mathrm{G}_{\mathrm{L}^*}(\boldsymbol{y},\boldsymbol{x})$ (the transposed version of G_{L^*}) should be a valid Green's function of L. While finding such an acceptable $G_{\mathrm{L}^*\mathrm{L}}$ may not always be easy, we can at least guarantee its existence. In particular, we note that the condition is met by all the reproducing kernels within the family, irrespective of the choice of $\boldsymbol{\phi}$. Another way to put it is that the knowledge of a single representative is enough to specify the whole family via equation (62). Finally, we note that there is a systematic method of construction (by generalized inverse Fourier transformation) in the favorable scenario where the operator shift-invariant (see Section 2.7.3).

Thanks to the characterization of the spaces $\mathcal{N}_{\mathrm{L}}$ and $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ provided by Proposition 3 and Theorem 12, respectively, and the fact that they are complementary with $\mathcal{N}_{\mathrm{L}}\cap\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}=\{0\}$, we are now able to extract the direct sum RKHS structure of the native space $\mathcal{H}_{\mathrm{L}}=\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}\oplus\mathcal{N}_{\mathrm{L}}$. Again, let us emphasize that this representation is not unique as there is one associated with each admissible $\boldsymbol{\phi}=(\phi_1,\dots,\phi_{N_0})$ (see Definition 13).

Theorem 13 (Characterization of native space). *Let* L *be an admissible operator and* $(\boldsymbol{\phi}, \boldsymbol{p})$ *a corresponding biorthogonal system for its null space* $\mathcal{N}_{\mathrm{L}}$. *Then, any* $f \in \mathcal{H}_{\mathrm{L}}$ *has a unique representation as*

$$f = \mathrm{L}_{\boldsymbol{\phi}}^{-1} w + q$$

where $w = \mathrm{L}f \in L_2(\mathbb{R}^d)$, $q = \mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{f\} = \sum_{n=1}^{N_0} \langle f, \phi_n \rangle p_n \in \mathcal{N}_{\mathrm{L}}$ *and* $\mathrm{L}_{\boldsymbol{\phi}}^{-1} : L_2(\mathbb{R}^d) \to \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ *is the right-inverse operator specified by Theorem 11. Moreover,* $\mathcal{H}_{\mathrm{L}}$ *equipped with the inner product*

$$\langle f, g \rangle_{\mathrm{L},\boldsymbol{\phi}} = \langle \mathrm{L}f, \mathrm{L}g \rangle + \sum_{n=1}^{N_0} \langle \phi_n, f \rangle \langle \phi_n, g \rangle \tag{66}$$

is a RKHS whose reproducing kernel is

$$h_{\boldsymbol{\phi}}(\boldsymbol{x}, \boldsymbol{y}) = a_{\boldsymbol{\phi}}(\boldsymbol{x}, \boldsymbol{y}) + \sum_{n=1}^{N_0} p_n(\boldsymbol{x}) p_n(\boldsymbol{y}) \tag{67}$$

where $a_{\boldsymbol{\phi}}(\boldsymbol{x}, \boldsymbol{y})$ *is given by* (62).

Finally, if A *(or the Green's function* $G_{\mathrm{L}^*\mathrm{L}}$*) meets the admissibility conditions in Definition 13 (strict conditional positivity and* α*-boundedness) and the* ϕ_n *are such that* $\mathrm{A}\{\phi_n\} \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$*, we have the continuous embedding* $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{H}_{\mathrm{L}} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d) \subseteq \mathcal{S}'(\mathbb{R}^d)$ *with the insurance that* $\mathcal{H}_{\mathrm{L}}$ *is* dense *in* $\mathcal{S}'(\mathbb{R}^d)$.

Proof. **Under construction**: The additive form of the inner product in (66) and the global RKHS property follow directly from the representation of the native space as the direct sum of two (reproducing kernel) Hilbert spaces: $\mathcal{H}_{\mathrm{L}} = \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}} \oplus \mathcal{N}_{\mathrm{L}}$.

The novel element here is the continuous embedding $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{H}_{\mathrm{L}} \subseteq \mathcal{S}'(\mathbb{R}^d)$. It will be established indirectly by showing that $\mathcal{S}(\mathbb{R}^d)$ is dense in $\mathcal{H}'_{\mathrm{L}} = \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}} \oplus \mathcal{N}'_{\mathrm{L}} \subseteq \mathcal{S}'(\mathbb{R}^d)$. To that end, we first assume that the boundary functionals ϕ_n are all included in $\mathcal{S}(\mathbb{R}^d)$. By using the same technique as in the proof of Theorem 10, we then decompose $\mathcal{S}(\mathbb{R}^d)$ as the direct sum $\mathcal{S}(\mathbb{R}^d) = \mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d) \oplus \mathcal{N}'_{\mathrm{L}}$ where

$$\mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d) = \{\tilde{\varphi} \in \mathcal{S}(\mathbb{R}^d) : \boldsymbol{\phi}(\tilde{\varphi}) = \boldsymbol{0}\}$$

and $\mathcal{N}'_{\mathrm{L}} = \mathrm{span}\{\phi_n\}_{n=1}^{N_0} \subseteq \mathcal{S}(\mathbb{R}^d)$. This means that every $\varphi \in \mathcal{S}(\mathbb{R}^d)$ has a unique decomposition as $\varphi = \tilde{\varphi} + \phi$ with $\phi = \mathrm{Proj}_{\mathcal{N}'_{\mathrm{L}}}\{\varphi\} = \sum_{n=1}^{N_0} \langle p_n, \varphi \rangle \phi_n$. Next, we recall that the inner product for $\mathcal{H}'_{\mathrm{L}}$ can be written as

$$\langle f, g \rangle_{\mathcal{H}'_{\mathrm{L}}} = \langle (\mathrm{A}_{\boldsymbol{\phi}} + \mathrm{R}_{\boldsymbol{p}}) f, g \rangle$$

where the two operators $\mathrm{A}_{\boldsymbol{\phi}}$ and $\mathrm{R}_{\boldsymbol{p}}$ are positive-definite by construction. Moreover, due to the direct sum decomposition, we have that

$$\begin{aligned}\|\varphi\|^2_{\mathcal{H}'_{\mathrm{L}}} &= \langle \mathrm{A}_{\boldsymbol{\phi}}\{\tilde{\varphi}\},\tilde{\varphi}\rangle + \langle \mathrm{R}_{\boldsymbol{p}}\{\phi\},\phi\rangle \\ &= \langle \mathrm{A}\{\tilde{\varphi}\} + \sum_{n=1}^{N_0} \langle \mathrm{A}\{\phi_n\},\tilde{\varphi}\rangle \underbrace{\langle p_n,\tilde{\varphi}\rangle}_{=0} + \sum_{n=1}^{N_0}\langle p_n,\phi\rangle\langle p_n,\phi\rangle \\ &= \langle \mathrm{A}\{\tilde{\varphi}\},\tilde{\varphi}\rangle + \|\boldsymbol{p}(\phi)\|_2^2\end{aligned}$$

where we have made use of the result in Proposition 10 to express $\|\varphi\|^2_{\mathcal{H}'_{\mathrm{L}}}$ in terms of A.

Since A is strictly positive-definite over $\mathcal{S}_{\boldsymbol{p}}(\mathbb{R}^d)$ and the same obviously holds true for $\mathrm{R}_{\boldsymbol{p}}$ over the complementary space $\mathcal{N}'_{\mathrm{L}}$, the map

$$\varphi = (\tilde{\varphi},\phi) \mapsto \|\varphi\|_{\mathcal{H}'_{\mathrm{L}}} = \|\tilde{\varphi}+\phi\|_{\mathcal{H}'_{\mathrm{L}}} = \sqrt{\langle \mathrm{A}\{\tilde{\varphi}\},\tilde{\varphi}\rangle + \|\boldsymbol{p}(\phi)\|_2^2}$$

specifies a valid norm over $\mathcal{S}(\mathbb{R}^d)$. Hence, we can view $\mathcal{H}'_{\mathrm{L}}$ as the completion of $\mathcal{S}(\mathbb{R}^d)$ in the $\|\cdot\|_{\mathcal{H}'_{\mathrm{L}}}$-norm, which is equivalent to the density property. Finally, we extend the argument to the general scenario $\phi_n \notin \mathcal{S}(\mathbb{R}^d)$ by considering an appropriate sequence of test functions $(\phi_{n,k})$ in $\mathcal{S}(\mathbb{R}^d)$ such that $\lim_{k\to\infty}\phi_{n,k} = \phi_n \in \mathcal{S}'(\mathbb{R}^d)$. The continuous embedding $\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{H}_{\mathrm{L}} \subseteq \mathcal{S}'(\mathbb{R}^d)$—and hence the denseness of $\mathcal{H}_{\mathrm{L}}$ in $\mathcal{S}'(\mathbb{R}^d)$—then follows from Theorem 3. Lastly, the hypothesis that $\mathrm{A}\{\phi_n\}, p_n \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ and the specific form of the kernel $a_{\boldsymbol{\phi}}(\boldsymbol{x},\boldsymbol{y})$ in (62) allows us to deduce that $h_{\boldsymbol{\phi}}(\cdot,\boldsymbol{y}) \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ for any fixed $\boldsymbol{y} \in \mathbb{R}^d$. Since $h_{\boldsymbol{\phi}}$ is positive-define and bounded on the diagonal (as a result of our assumptions), we readily conclude that $\mathcal{H}_{\mathrm{L}} \subseteq C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ by invoking Theorem 7.

□

Proposition 11. *Let $\mathcal{H}_{\mathrm{L}}$ be the RKHS specified in Theorem 13 and* $\mathrm{A}: \varphi \mapsto \int_{\mathbb{R}^d} G_{\mathrm{L}^*\mathrm{L}}(\cdot)\varphi(\boldsymbol{y})\mathrm{d}y$. *Then,* $G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y}_0) \in \mathcal{H}_{\mathrm{L}}$ *for any* $\boldsymbol{y}_0 \in \mathbb{R}^d$ *and* $\mathrm{A}\{\phi_n\} \in \mathcal{H}_{\mathrm{L}}$ *for all n. In particular, this implies that the set* $\mathcal{H}_{\mathrm{pre,L}} = \big\{\sum_{k=1}^{K} a_k G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y}_k) + \sum_{n=1}^{N_0} b_n p_n : K \in \mathbb{N}, a_k, b_n \in \mathbb{R}, \boldsymbol{y}_k \in \mathbb{R}^d\big\}$ *is dense in $\mathcal{H}_{\mathrm{L}}$. In other words, we can represent any function $f \in \mathcal{H}_{\mathrm{L}}$—and, by extension, $f \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$—as closely as desired by using a linear combination of the form*

$$\tilde{f}(\boldsymbol{x}) = \sum_{k=1}^{K} a_k G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x},\boldsymbol{y}_k) + \sum_{n=1}^{N_0} b_n p_n(\boldsymbol{x})$$

with a finite number $K + N_0$ of terms and (adaptive) centers $\boldsymbol{y}_k \in \mathbb{R}^d$.

Proof. First, we observe that the projection of $\mathrm{A}\{\phi_m\}$ onto $\mathcal{N}_{\mathrm{L}}$ can be written as

$$\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{\mathrm{A}\{\phi_m\}\} = \sum_{n=1}^{N_0} \langle \phi_m, \mathrm{A}\{\phi_m\}\rangle p_n = \sum_{n=1}^{N_0} r_{m,n} p_n$$

where the constants $r_{m,n}$ are defined in (64). It then follows that

$$\begin{aligned}\|\mathrm{A}\{\phi_m\}\|^2_{\mathcal{H}_L} &= \langle (\mathrm{L}^*\mathrm{L})\mathrm{A}\{\phi_m\}, \mathrm{A}\{\phi_m\}\rangle + \|\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}}\{\mathrm{A}\{\phi_n\}\}\|^2_{\mathcal{N}_{\mathrm{L}}}\\ &= \underbrace{\langle \phi_m, \mathrm{A}\{\phi_m\}\rangle}_{r_{m,m}} + \sum_{n=1}^{N_0} r^2_{m,n} < \infty\end{aligned}$$

which proves that $\mathrm{A}\{\phi_m\} \in \mathcal{H}_{\mathrm{L}}$. As for $G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y}_0)$ with $\boldsymbol{y}_0$ fixed, we use the expression of the reproducing kernel (62) to rewrite it as

$$\begin{aligned}G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y}_0) =& a_{\boldsymbol{\phi}}(\cdot, \boldsymbol{y}_0) + \sum_{n=1}^{N_0} p_n(\boldsymbol{y}_0)\,\mathrm{A}\{\phi_n\}(\cdot) + \sum_{n=1}^{N_0} \mathrm{A}\{\phi_n\}(\cdot)\,p_n(\boldsymbol{y}_0)\\ &- \sum_{m=1}^{N_0}\sum_{n=1}^{N_0} r_{m,n} p_m(\cdot)\,p_n(\boldsymbol{y}_0)\end{aligned}$$

Since $p_n(\boldsymbol{y}_0)$ and $\mathrm{A}\{\phi_n\}(\boldsymbol{y}_0)$ are constants and $a_{\boldsymbol{\phi}}(\cdot, \boldsymbol{y}_0) \in \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ (reproducing kernel property), all the functions on the right-hand side are included in $\mathcal{H}_{\mathrm{L}}$ so that the same holds true for $G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y}_0)$ (due to the vector-space property of $\mathcal{H}_{\mathrm{L}}$). □

We conclude this section by revealing the functional properties of the operator $\mathrm{A}_{\boldsymbol{\phi}} = \mathrm{L}^{-1}_{\boldsymbol{\phi}}\mathrm{L}^{-1*}_{\boldsymbol{\phi}}$ associated with the kernel $a_{\boldsymbol{\phi}}(\cdot, \boldsymbol{y})$. Operationally, the latter constitutes a regularized version of the symmetric operator $\mathrm{A} : \varphi \mapsto \int_{\mathbb{R}^d} G_{\mathrm{L}^*\mathrm{L}}(\cdot, \boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$, which cannot generally be ensured to be bounded $\mathcal{H}'_{\mathrm{L}} \to \mathcal{H}_{\mathrm{L}}$. We recall that both operators are right-inverses of $(\mathrm{L}^*\mathrm{L})$ and that they are equivalent only when the null space of L is trivial.

Theorem 14. *Let* L *be an admissible operator and* $(\boldsymbol{\phi}, \boldsymbol{p})$ *a corresponding biorthogonal system for its null space* $\mathcal{N}_{\mathrm{L}}$*. Then, the operator* $\mathrm{A}_{\boldsymbol{\phi}} : \varphi \mapsto \int_{\mathbb{R}^d} a_{\boldsymbol{\phi}}(\cdot, \boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$*, where* $a_{\boldsymbol{\phi}}(\cdot, \boldsymbol{y})$ *is given by* (62) *in Theorem 12, has the following properties:*

1. *It is the Riesz map* $\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}} \to \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}} = \{f \in \mathcal{H}_{\mathrm{L}} : \boldsymbol{\phi}(f) = \boldsymbol{0}\}$.
2. *It is bounded* $\mathcal{H}'_{\mathrm{L}} \to \mathcal{H}_{\mathrm{L}} = \mathcal{H}_{\mathrm{L},\boldsymbol{\phi}} \oplus \mathcal{N}_{\mathrm{L}}$.

Description	Operator	Kernel
Right-inverse of L	G	$G_{\mathrm{L}}(\boldsymbol{x},\boldsymbol{y})$
Right-inverse of $(\mathrm{L}^*\mathrm{L})$	$\mathrm{A}=\mathrm{GG}^*$	$G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x},\boldsymbol{y})$
Stable inverse of L	$\mathrm{L}_{\boldsymbol{\phi}}^{-1}$	$G_{\mathrm{L}}(\boldsymbol{x},\boldsymbol{y})-\sum_{n=1}^{N_0}p_n(\boldsymbol{x})\mathrm{G}^*\{\phi_n\}(\boldsymbol{y})$
Riesz map $\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}\to\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$	$\mathrm{A}_{\boldsymbol{\phi}}$	$a_{\boldsymbol{\phi}}(\boldsymbol{x},\boldsymbol{y})$
Riesz map $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}\to\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$	$(\mathrm{L}^*\mathrm{L})$	
Riesz map $\mathcal{H}'_{\mathrm{L}}\to\mathcal{H}_{\mathrm{L}}$	$\mathrm{A}_{\boldsymbol{\phi}}+\mathrm{R}_{\boldsymbol{p}}$	$a_{\boldsymbol{\phi}}(\boldsymbol{x},\boldsymbol{y})+\sum_{n=1}^{N_0}p_n(\boldsymbol{x})p_n(\boldsymbol{y})$
Riesz map $\mathcal{H}_{\mathrm{L}}\to\mathcal{H}'_{\mathrm{L}}$	$(\mathrm{L}^*\mathrm{L})+\mathrm{R}_{\boldsymbol{\phi}}$	

Table 3: Primary operators that enter the definition of the Hilbert spaces $\mathcal{H}_{\mathrm{L}}$, $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ and their duals $\mathcal{H}'_{\mathrm{L}}$, $\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$.

3. *It has a finite-dimensional null space* $\mathcal{N}_{\mathrm{A}_{\boldsymbol{\phi}}}=\mathrm{span}\{\phi_n\}_{n=1}^{N_0}$ *that is isomorphically equivalent to* $\mathcal{N}_{\mathrm{L}}=\mathrm{span}\{p_n\}_{n=1}^{N_0}$. *In fact,* $\mathcal{N}_{\mathrm{A}_{\boldsymbol{\phi}}}=\mathcal{N}'_{\mathrm{L}}$ *is the topological dual of* $\mathcal{N}_{\mathrm{L}}$ *equipped with the norm* $\|q\|_{\mathcal{N}_{\mathrm{L}}}=\|\boldsymbol{\phi}(q)\|_2$.

4. *It imposes the boundary conditions:* $\boldsymbol{\phi}(\mathrm{A}_{\boldsymbol{\phi}}\{f^*\})=\boldsymbol{0}$ *for all* $f^*\in\mathcal{H}'_{\mathrm{L}}$. *In other words,* $\mathrm{A}_{\boldsymbol{\phi}}$ *continuously maps* $\mathcal{H}'_{\mathrm{L}}\to\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$.

5. *Let* $\mathrm{A}:\varphi\mapsto\int_{\mathbb{R}^d}G_{\mathrm{L}^*\mathrm{L}}(\cdot,\boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$ *where* $G_{\mathrm{L}^*\mathrm{L}}$ *is the symmetric Green's function of* $\mathrm{L}^*\mathrm{L}$ *specified in Theorem 12. Then, for any* φ *for which* $\mathrm{A}\{\varphi\}$ *is well-defined, there exists* $(d_n),(c_n)\in\mathbb{R}^{N_0}$ *such that*

$$\mathrm{A}_{\boldsymbol{\phi}}\{\varphi\}=\mathrm{A}\{\varphi\}+\sum_{n=1}^{N_0}d_np_n+\sum_{n=1}^{N_0}c_n\mathrm{A}\{\phi_n\}. \tag{68}$$

In particular, we have that $\mathrm{A}_{\boldsymbol{\phi}}\{f^*\}=\mathrm{A}\{f^*\}-\sum_{n=1}^{N_0}\langle\mathrm{A}\{\phi_n\},f^*\rangle p_n$ *for all* $f^*\in\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$ *and* $\mathrm{A}_{\boldsymbol{\phi}}\{\phi\}=0$ *for all* $\phi\in\mathcal{N}'_{\mathrm{L}}$.

6. $\mathrm{A}_{\boldsymbol{\phi}}=(\mathrm{Id}-\mathrm{Proj}_{\mathcal{N}'_{\mathrm{L}}})^*\mathrm{A}(\mathrm{Id}-\mathrm{Proj}_{\mathcal{N}'_{\mathrm{L}}})=(\mathrm{Id}-\mathrm{Proj}_{\mathcal{N}_{\mathrm{L}}})\mathrm{A}(\mathrm{Id}-\mathrm{Proj}_{\mathcal{N}'_{\mathrm{L}}})$ *and is the Moore-Penrose pseudoinverse of* $(\mathrm{L}^*\mathrm{L}):\mathcal{H}_{\mathrm{L}}\to\mathcal{H}'_{\mathrm{L}}$ *in the underlying direct-sum topology; that is, it has the property that* $\mathrm{A}_{\boldsymbol{\phi}}(\mathrm{L}^*\mathrm{L})\{f\}=f$ *for all* $f\in\mathcal{N}_{\mathrm{L}}^{\perp}=\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ *and* $\mathrm{A}_{\boldsymbol{\phi}}\{\phi\}=0$ *for all* $\phi\in\mathrm{Im}(\mathrm{L}^*\mathrm{L})^{\perp}=\mathcal{N}'_{\mathrm{L}}$.

Proof. Property 1 is a restatement of the fact that $a_{\boldsymbol{\phi}}(\cdot,\cdot)$ is the reproducing kernel of $\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$ (see Theorem 12). We then reveal $\mathcal{N}_{\mathrm{A}_{\boldsymbol{\phi}}}$ by showing that $\mathrm{A}_{\boldsymbol{\phi}}\{\phi_n\}=0$ for $n=1,\dots,N_0$. To that end, we start by observing that

$$\mathrm{A}\{\phi_m\}(\boldsymbol{x})=\langle G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x},\cdot),\phi_m\rangle=v_m(\boldsymbol{x}), \tag{69}$$

in agreement with Definition (63). Similarly, we find that

$$\langle v_n,\phi_m\rangle=\langle \mathrm{A}\{\phi_n\},\phi_m\rangle=\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}G_{\mathrm{L}^*\mathrm{L}}(\boldsymbol{x},\boldsymbol{y})\phi_n(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\right)\phi_m(\boldsymbol{x})\mathrm{d}\boldsymbol{x}=r_{m,n}$$

where $r_{m,n}=r_{n,m}$ is defined by (64). Based on these identifications, we get

$$\begin{aligned}\mathrm{A}_{\boldsymbol{\phi}}\{\phi_m\}&=v_m(\boldsymbol{x})-\sum_{n=1}^{N_0}p_n(\boldsymbol{x})r_{m,n}-\sum_{n=1}^{N_0}v_n(\boldsymbol{x})\langle p_n,\phi_m\rangle\\&\quad+\sum_{n=1}^{N_0}\sum_{n'=1}^{N_0}r_{n,n'}p_n(\boldsymbol{x})\langle p_{n'},\phi_m\rangle\\&=v_m(\boldsymbol{x})-\sum_{n=1}^{N_0}p_n(\boldsymbol{x})r_{m,n}-v_m(\boldsymbol{x})+\sum_{n=1}^{N_0}r_{n,m}p_n(\boldsymbol{x})=0\end{aligned}$$

where we have made use of the biorthogonality of $\{p_n\}$ and $\{\phi_n\}$ to reduce the sums. Since $\mathcal{H}_{\mathrm{L}}=\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}\oplus\mathcal{N}_{\mathrm{L}}$ (by Corollary 13) and $\mathrm{A}_{\boldsymbol{\phi}}$ isometrically maps $\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}\to\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}$, we have that $\mathcal{H}'_{\mathrm{L}}=\mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}\oplus\mathcal{N}'_{\mathrm{L}}$ with $\mathcal{N}'_{\mathrm{L}}=\mathcal{N}_{\mathrm{A}_{\boldsymbol{\phi}}}=\mathrm{span}\{\phi_n\}_{n=1}^{N_0}$. This, in turn, allows us to deduce that $\mathrm{A}_{\boldsymbol{\phi}}$ continuously maps $\mathcal{H}'_{\mathrm{L}}\to\mathcal{H}_{\mathrm{L},\boldsymbol{\phi}}\subseteq\mathcal{H}_{\mathrm{L}}$, which yields Properties 2 and 4. Note that we could have anticipated these boundary conditions based on the property that $\mathrm{A}_{\boldsymbol{\phi}}=\mathrm{L}_{\boldsymbol{\phi}}^{-1}\mathrm{L}_{\boldsymbol{\phi}}^{-1*}$.

As for the last property, we also rely on (69) and expand $\mathrm{A}_{\boldsymbol{\phi}}\{\varphi\}$ as

$$\begin{aligned}\mathrm{A}_{\boldsymbol{\phi}}\{\varphi\}&=\int_{\mathbb{R}^d}a_{\boldsymbol{\phi}}(\cdot,\boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\\&=\mathrm{A}\{\varphi\}-\sum_{n=1}^{N_0}p_n\langle \mathrm{A}\{\phi_n\},\varphi\rangle-\sum_{n=1}^{N_0}\mathrm{A}\{\phi_n\}\langle\varphi,p_n\rangle\\&\quad+\sum_{m=1}^{N_0}\sum_{n=1}^{N_0}r_{m,n}p_n\langle p_m,\varphi\rangle\\&=\mathrm{A}\{\varphi\}+\sum_{n=1}^{N_0}c_n\mathrm{A}\{\phi_n\}+\sum_{n=1}^{N_0}d_np_n\end{aligned}$$

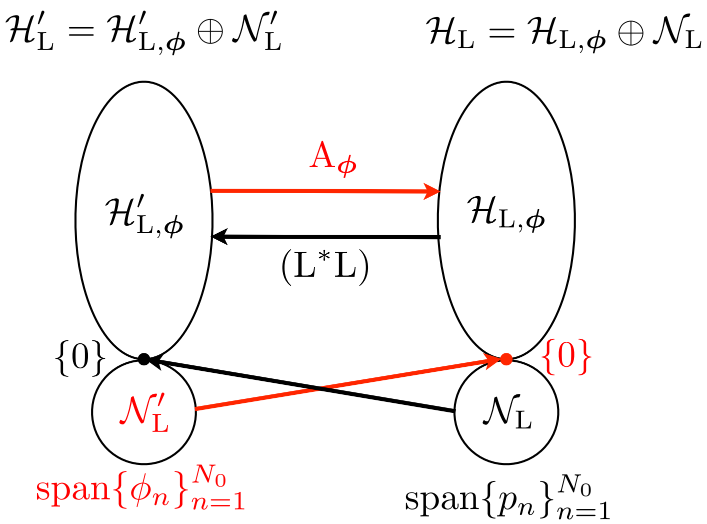

Figure 2: Schematic illustration of the mapping between the various Hilbert spaces.

with

$$c_n = -\langle \varphi, p_n \rangle$$
$$d_n = -\langle \mathrm{A}\{\phi_n\}, \varphi \rangle + \sum_{m=1}^{N_0} r_{m,n} \langle p_m, \varphi \rangle.$$

In particular, if $\varphi \in \mathcal{H}'_{\mathrm{L},\boldsymbol{\phi}}$, then $c_n = \langle \varphi, p_n \rangle = 0$ (see Proposition 10). Since $\mathrm{A}_{\boldsymbol{\phi}}\{\varphi\} \in \mathcal{H}_{\mathrm{L}}$ for any $\varphi \in \mathcal{H}'_{\mathrm{L}}(\mathbb{R}^d)$, it suffices that $\mathrm{A}\{\varphi\}$ be well-defined for the splitting (68) to be legitimate. □

The complete picture of those functional mappings is given in Figure 2, while the relevant operators are summarized in Table 3.