6 Splines and wavelets

A fundamental aspect of our formulation is that the whitening operator $L$ is naturally tied to some underlying B-spline function, which will play a crucial role in the sequel. The spline connection also provides a strong link with wavelets [UB03].

In this chapter, we review the foundations of spline theory and show how one can construct B-spline basis functions and wavelets that are tied to some specific operator $L$. The chapter starts with a gentle introduction to wavelets that exploits the analogy with Legos blocks. This naturally leads to the formulation of a multiresolution analysis of $L^2(\mathbb{R})$ using piecewise-constant functions and a de visu identification of Haar wavelets. We then proceed in Section 6.2 with a formal definition of our generalized brand of splines—the cardinal $L$-splines—followed by a detailed discussion of the fundamental notion of Riesz basis. In Section 6.3, we systematically cover the first-order operators with the construction of exponential B-splines and wavelets, which have the convenient property of being orthogonal. We then address the theory in its full generality and present two generic methods for constructing B-spline basis functions (Section 6.4) and semi-orthogonal wavelets (Section 6.5). The pleasing aspect is that these results apply to the whole class of shift-invariant differential operators $L$ whose null space is finite-dimensional (possibly trivial), which are precisely those that can be safely inverted to specify sparse stochastic processes.

6.1 From Legos to wavelets

It is instructive to get back to our introductory example of piecewise-constant splines in Chapter 1 (§1.3) and to show how these are naturally connected to wavelets. The fundamental idea in wavelet theory is to construct a series of fine-to-coarse approximations of a function $s(r)$ and to exploit the structure of the multiresolution approximation errors, which are orthogonal across scale. Here, we shall consider a series of approximating signals $s_i \in \mathbb{Z}$, where $s_i$ is a piecewise-constant spline with knots positioned on $2^i \mathbb{Z}$. These multiresolution splines are represented by their B-spline expansion

$$s_i(r) = \sum_{k \in \mathbb{Z}} c_i[k] \phi_{i,k}(r), \quad (6.1)$$
where the B-spline basis functions (rectangles) are dilated versions of the cardinal ones by a factor of $2^i$

$$\phi_{i,k}(r) = \beta^0_i \left( \frac{r - 2^i k}{2^i} \right) = \begin{cases} 1, & \text{for } r \in [2^i k, 2^i (k + 1)] \\ 0, & \text{otherwise.} \end{cases} \quad (6.2)$$

The variable $i$ is the scale index that specifies the resolution (or knot spacing) $a = 2^i$, while the integer $k$ encodes for the spatial location. The B-spline of degree 0, $\phi = \phi_{0,0} = \beta^0_1$, is the scaling function of the representation. Interestingly, it is the identification of a proper scaling function that constitutes the most fundamental step in the construction of a wavelet basis of $L_2(\mathbb{R})$.

**Definition 6.1 (Scaling function)** $\phi \in L_2(\mathbb{R})$ is a valid scaling function if and only if it satisfies the following three properties:

- Two-scale relation

$$\phi(r/2) = \sum_{k \in \mathbb{Z}} h[k]\phi(r - k), \quad (6.3)$$

where the sequence $h \in \ell_1(\mathbb{Z})$ is the so-called refinement mask.

- Partition of unity

$$\sum_{k \in \mathbb{Z}} \phi(r - k) = 1 \quad (6.4)$$

- The set of functions $\{\phi(-k)\}_{k \in \mathbb{Z}}$ forms a Riesz basis.

In practice, a given brand of (orthogonal) wavelets (e.g., Daubechies or splines) is often summarized by its refinement filter $h$ since the latter uniquely specifies $\phi$, subject to the admissibility constraints (6.4) and $\phi \in L_2(\mathbb{R})$. In the case of the B-spline of degree 0, we have that $h[k] = \delta[k] + \delta[k - 1]$, where

$$\delta[k] = \begin{cases} 1, & \text{for } k = 0 \\ 0, & \text{otherwise} \end{cases}$$

is the discrete Kronecker impulse. This translates into what we jokingly refer to as the Lego-Duplo relation

$$\beta^0_i (r/2) = \beta^0_i (r) + \beta^0_i (r - 1). \quad (6.5)$$

The fact that $\beta^0_i$ satisfies the partition of unity is obvious. Likewise, we already observed in Chapter 1 that $\beta^0_i$ generates an orthogonal system which is a special case of a Riesz basis.

By considering the rescaled version of such a basis, we specify the subspace of splines at scale $i$ as

$$V_i = \left\{ \sum_{k \in \mathbb{Z}} c_j[k] \phi_{i,k}(r) : c_j \in \ell_2(\mathbb{Z}) \right\} \subseteq L_2(\mathbb{R})$$

1. The Duplos are the larger-scale versions of the Lego building blocks and are more suitable for smaller children to play with. The main point of the analogy with wavelets is that Legos and Duplos are compatible; they can be combined to build more complex shapes. The enabling property is that a Duplo is equivalent to two smaller Legos placed next to each other, as expressed by (6.5)
which, in our example, contains all the finite-energy functions that are piecewise-
constant on the intervals \([2^i k, 2^i (k + 1)]\) with \(k \in \mathbb{Z}\). The two-scale relation (6.3) im-
plies that the basis functions at scale \(i = 1\) are contained in \(V_0\) (the original space
of cardinal splines) and, by extension, in \(V_i\) for \(i \leq 0\). This translates into the gen-
eral inclusion property \(V_i \subset \mathcal{V}\) for any \(i > 0\), which is fundamental to the theory.
A subtler point is that the closure of \(S_i\) is equal to \(L^2(\mathbb{R})\), which follows from
the fact that any finite-energy function can be approximated arbitrarily well by a
piecewise-constant spline when the sampling step \(2^i\) tends to zero \((i \to -\infty)\). The
necessary and sufficient condition for this asymptotic convergence is the partition
of unity (6.4), which ensures that the representation is complete.

Having set the notation and specified the underlying hierarchy of function spaces,
we now proceed with the multiresolution approximation procedure starting from the
fine-scale signal \(s_0(x)\), as illustrated in Figure 6.1. Given the sequence \(c_0[\cdot]\) of fine-
scale coefficients, our task is to construct the best spline approximation at scale 1
which is specified by its B-spline coefficients \(c_1[\cdot]\) in (6.1) with \(i = 1\). It is easy to see
that the minimum-error solution (orthogonal projection of \(s_0\) onto \(V_1\)) is obtained by
taking the mean of two consecutive samples. The procedure is then repeated at the
next coarser scale and so forth until one reaches the bottom of the pyramid, as shown
on the left-hand side of Figure 6.1. The description of this coarsening algorithm is

\[
c_i[k] = \frac{1}{2} c_{i-1}[2k] + \frac{1}{2} c_{i-1}[2k + 1] = (c_{i-1} * \tilde{h})[2k]. \tag{6.6}
\]
It is run recursively for \( i = 1, \ldots, i_{\text{max}} \) where \( i_{\text{max}} \) denotes the bottom level of the pyramid. The outcome is a multiresolution analysis of our input signal \( s_0 \).

In order to uncover the wavelets, it is enlightening to look at the residual signals \( r_i(r) = s_{i-1}(r) - s_i(r) \in V_{i-1} \) on the right of Figure 6.1. While these are splines that live at the same resolution as \( s_{i-1} \), they actually have half the apparent degrees of freedom. These error signals exhibit a striking sign-alternation pattern due to the fact that two consecutive samples \( (c_{i-1}[2k], c_{i-1}[2k + 1]) \) are at an equal distance from their mean value \( (c_i[k]) \). This suggests rewriting the residuals more concisely in terms of oscillating basis functions (wavelets) at scale \( i \), like

\[
r_i(r) = s_{i-1}(r) - s_i(r) = \sum_{k \in \mathbb{Z}} d_i[k] \psi_{i,k}(r),
\]

where the (non-normalized) Haar wavelets are given by

\[
\psi_{i,k}(r) = \psi_{\text{Haar}} \left( \frac{r - 2^i k}{2^i} \right)
\]

with the Haar wavelet being defined by (1.19). The wavelet coefficients \( d_{i}[\cdot] \) are given by the consecutive half differences

\[
d_{i}[k] = \frac{1}{2} c_{i-1}[2k] - \frac{1}{2} c_{i-1}[2k + 1] = (c_{i-1} * \tilde{g})[2k].
\]

More generally, since the wavelet template at scale \( i = 1, \psi_{1,0} \in V_0 \), we can write

\[
\psi(r/2) = \sum_{k \in \mathbb{Z}} g[k] \phi(r - k)
\]

which is the wavelet counterpart of the two-scale relation (6.3). In the present example, we have \( g[k] = (-1)^k h[k] \), which is a general relation that is characteristic of an orthogonal design. Likewise, in order to gain in generality, we have chosen to express the decomposition algorithms (6.6) and (6.8) in terms of discrete convolution (filtering) and downsampling operations where the corresponding Haar analysis filters are \( \tilde{h}[k] = \frac{1}{\sqrt{2}} h[-k] \) and \( \tilde{g}[k] = \frac{1}{\sqrt{2}} (-1)^k h[-k] \). The Hilbert-space interpretation of this approximation process is that \( r_i \in W_i \), where \( W_i \) is the orthogonal complement of \( V_i \) in \( V_{i-1} \); that is, \( V_{i-1} = V_i + W_i \) with \( V_i \perp W_i \) (as a consequence of the orthogonal-projection theorem).

Finally, we can close the loop by observing that

\[
s_0(r) = s_{i_{\text{max}}}(r) + \sum_{i=1}^{i_{\text{max}}} \left( \frac{s_{i-1}(r) - s_i(r)}{r(r)} \right) = \sum_{k \in \mathbb{Z}} c_{i_{\text{max}}}[k] \phi_{i_{\text{max}}, k}(r) + \sum_{i=1}^{i_{\text{max}}} \sum_{k \in \mathbb{Z}} d_i[k] \psi_{i, k}(r),
\]

which provides an equivalent, one-to-one representation of the signal in an orthogonal wavelet basis, as illustrated in Figure 6.2.
Figure 6.2  Decomposition of a signal into orthogonal scale components. The error signals $r_i = s_{i-1} - s_i$ between two successive signal approximations are expanded using a series of properly scaled wavelets.

More generally, we can push the argument to the limit and apply the decomposition to any finite-energy function

$$s = \sum_{i \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} d_i[k] \psi_{i,k},$$

(6.11)

where $d_i[k] = \langle s, \psi_{i,k} \rangle_{L^2}$ and $\{\psi_{i,k}\}_{i,k \in \mathbb{Z}}$ is a suitable (bi-)orthogonal wavelet basis with the property that $\langle \psi_{i,k}, \psi_{i',k'} \rangle_{L^2} = \delta_{k-k',i-i'}$.

Remarkably, the whole process described above—except the central expressions in (6.6) and (6.8), and the equations explicitly involving $\tilde{\phi}_0^+$—is completely generic and applicable to any other wavelet basis of $L^2(\mathbb{R})$ that is specified in terms of a wavelet filter $g$ and a scaling function $\phi$ (or, equivalently, an admissible refinement filter $h$). The bottom line is that the wavelet decomposition and reconstruction algorithm is fully described by the four digital filters ($h, g, \tilde{h}, \tilde{g}$) that form a perfect reconstruction filterbank. The Haar transform is associated with the shortest-possible filters. Its less favorable aspects are that the basis functions are discontinuous and that the scale-truncated error decays only like the first power of the sampling step $a = 2^i$ (first order of approximation).

The fundamental point of our formulation is that the Haar wavelet is matched to the pure derivative operator $D = \frac{d}{dr}$, which goes hand in hand with Lévy processes (see Chapter 1). In that respect, the critical observations relating to spline and wavelet theory are as follows:

- all piecewise-constant functions can be interpreted as D-splines;
– the Haar wavelet acts as a smoothed version of the derivative in the sense that 
\( \psi_{\text{Haar}} = D\phi \), where \( \phi \) is an appropriate kernel (triangle function);
– the B-spline of degree 0 can be expressed as \( \beta^0 = \beta D = D_0 D^{-1} \delta \), where the finite-
difference operator \( D_0 \) is the discrete counterpart of \( D \).

We shall now show how these ideas are extendable to a much broader class of differ-
ential operators \( L \).

### 6.2 Basic concepts and definitions

#### 6.2.1 Spline-admissible operators

Let \( L : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d) \) be a generic Fourier-multiplier operator with frequency
response \( \hat{L}(\omega) \). We further assume that \( L \) has a continuous extension \( L : \mathcal{S}(\mathbb{R}^d) \rightarrow 
\mathcal{S}'(\mathbb{R}^d) \) to some larger space of functions \( \mathcal{S}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \).

The null space of \( L \) is denoted by \( \mathcal{N}_L \) and defined as

\[ \mathcal{N}_L = \{ p_0(r) : Lp_0(r) = 0 \}. \]

The immediate consequence of \( L \) being LSI is that \( \mathcal{N}_L \) is shift-invariant as well, in the
sense that \( p_0(r) \in \mathcal{N}_L \Rightarrow p_0(r-r_0) \in \mathcal{N}_L \) for any \( r_0 \in \mathbb{R}^d \). We shall use this property
to argue that \( \mathcal{N}_L \) generally consists of generalized functions whose Fourier trans-
forms are point distributions. In the space domain, they correspond to modulated
polynomials, which are linear combinations of exponential monomials of the form
\( e^{(z_0,r)} r^n \) with \( z_0 \in \mathbb{R}^d \) and multi-index \( n = (n_1, \ldots, n_d) \in \mathbb{N}^d \). It actually turns out that
the existence of a single such element in \( \mathcal{N}_L \) has direct implications on the structure
and dimensionality of the underlying function space.

**Proposition 6.1 (Characterization of null space)** If \( L \) is LSI and \( p_n(r) = e^{(z_0,r)} r^n \in 
\mathcal{N}_L \) with \( z_0 \in \mathbb{C}^d \), then \( \mathcal{N}_L \) does necessarily include all exponential monomials of the form \( p_m(r) = e^{(z_0,r)} r^m \) with \( m \leq n \). In addition, if \( \mathcal{N}_L \) is finite-dimensional, it can
only consist of atoms of that particular form.

**Proof** The LSI property implies that \( p_n(r - r_0) \in \mathcal{N}_L \) for any \( r_0 \in \mathbb{R}^d \). To make our
point about the inclusion of the lower-order exponential polynomials in \( \mathcal{N}_L \), we start
by expanding the scalar term \( (r_i - r_{0,i})^{n_i} \) as

\[
(r_i - r_{0,i})^{n_i} = \sum_{m=0}^{n_i} \binom{n_i}{m} r_i^m (-1)^{n_i-m} r_{0,i}^{n_i-m} = \sum_{m+k=n_i} \frac{n_i!}{m! k!} (-1)^k r_i^m r_{0,i}^k.
\]

By proceeding in a similar manner with the other monomials and combining the
results, we find that

\[
(r - r_0)^n = \sum_{m+k=n} \frac{n!}{m! k!} (-1)^k r^m r_0^k = \sum_{m \leq n} b_m(r_0) r^m
\]

with polynomial coefficients \( b_m(r_0) \) that depend upon the multi-index \( m \) and the
Finally, we note that the exponential factor $e^{i\omega_0 \cdot r}$ can be shifted by $r_0$ by simple multiplication with a constant (see (6.12) below). These facts taken together establish the structure of the underlying vector space. As for the last statement, we rely on the theory of Lie group that tells us that the only finite-dimensional collection of functions that are translation-invariant is made of exponential polynomials. The pure exponentials $e^{i\omega_0 \cdot r}$ (with $n = 0$) are special in that respect: They are the eigenfunctions of the shift operator in the sense that

$$e^{i(\omega_0 \cdot r + r_0)} = \lambda(r_0) e^{i\omega_0 \cdot r}$$

(6.12)

with the (complex) eigenvalue $\lambda(r_0) = e^{i\omega_0 \cdot r_0}$, and hence the only elements that specify shift-invariant subspaces of dimension 1.

Since our formulation relies on the theory of generalized functions, we shall focus on the restriction of $N_L$ to $S_0(\mathbb{R}^d)$. This rules out the exponential factors $\omega_0 = \alpha_0 + j\omega_0$ in Proposition 6.1 with $\omega_0 \not\in \mathbb{R}^d \setminus \{0\}$ for which the Fourier-multiplier operator is not necessarily well-defined. We are then left with null-space atoms of the form $e^{i(\omega_0 \cdot r + r_0)}$ with $\omega_0 \in \mathbb{R}^d$, which are functions of slow growth.

The next important ingredient is the Green’s function $\Omega_L$ of the operator $L$. Its defining property is $L \Omega_L = \delta$, where $\delta$ is the $d$-dimensional Dirac distribution. Since there are many equivalent Green’s functions of the form $\rho_L + p_0$ where $p_0 \in N_L$ is an arbitrary component of the null space, we resolve the ambiguity by defining the (primary) Green’s function of $L$ as

$$\rho_L(r) = \mathcal{F}^{-1} \left\{ \frac{1}{\hat{L}(\omega)} \right\}(r),$$

(6.13)

with the requirement that $\rho_L \in \mathcal{S}'(\mathbb{R}^d)$ is an ordinary function of slow growth. In other words, we want $\rho_L(r)$ to be defined pointwise for any $r \in \mathbb{R}^d$ and to grow no faster than a polynomial. The existence of the generalized inverse Fourier transform (6.13) imposes some minimal continuity and decay conditions on $1/\hat{L}(\omega)$ and also puts some restrictions on the number and nature of its singularities (e.g., the zeros of $\hat{L}(\omega)$).

**Definition 6.2 (Spline admissibility)** The Fourier-multiplier operator $L : \mathcal{S}[\mathbb{R}^d] \to \mathcal{S}'(\mathbb{R}^d)$ with frequency response $\hat{L}(\omega)$ is called spline admissible if (6.13) is well-defined and $\rho_L(r)$ is an ordinary function of slow growth.

An important characteristic of spline-admissible operators is the rate of growth of their frequency response at infinity.

**Definition 6.3 (Order of a Fourier multiplier)** The Fourier-multiplier $\hat{L}(\omega)$ is of (asymptotic) order $\gamma \in \mathbb{R}^+$ if there exists a radius $R \in \mathbb{R}^+$ and a constant $C$ such that

$$C|\omega|^\gamma \leq |\hat{L}(\omega)|$$

(6.14)

for all $|\omega| \geq R$ where $\gamma$ is critical in the sense that the condition fails for any larger value.
The order is in direct relation with the degree of smoothness of the Green's function $\rho_L$. In the case of a scale-invariant operator, it also coincides with the scaling order (or degree of homogeneity) of $\hat{L}(\omega)$. For instance, the fractional derivative operator $D^\gamma$, which is defined via the Fourier multiplier $(j\omega)^\gamma$, is of order $\gamma$. Its Green's function is given by (see Table A.1 in Appendix A)

$$\rho_{D^\gamma}(r) = \mathcal{F}^{-1}\left\{ \frac{1}{(j\omega)^\gamma} \right\}(r) = \frac{r^{\gamma-1}}{\Gamma(\gamma)}, \quad (6.15)$$

where $\Gamma$ is Euler's gamma function (see Appendix C.2) and $r^{\gamma-1} = \max(0, r^{\gamma-1})$. Clearly, the latter is a function of slow growth. It has a single singularity at the origin whose Hölder exponent is $(\gamma - 1)$, and is infinitely differentiable everywhere else. It follows that $\rho_{D^\gamma}$ is uniformly Hölder-continuous of degree $(\gamma - 1)$. This is one less than the order of the operator. On the other hand, the null space of $D^\gamma$ consists of the polynomials of degree $N = \lfloor \gamma - 1 \rfloor$ since $\frac{d^n}{dr^n} (j\omega)^\gamma \propto (j\omega)^{\gamma-n}$ is vanishing at the origin up to order $N$ with $(\gamma - 1) \leq N < \gamma$ (see argumentation in Section 6.4.1).

A fundamental result is that all partial differential operators with constant coefficients are spline-admissible. This follows from the Malgrange-Ehrenpreis theorem which guarantees the existence of their Green's function [Mal56,Wag09]. The generic form of such operators is

$$L_N = \sum_{|n| < N} a_n \partial^n$$

with $a_n \in \mathbb{R}^d$, where $\partial^n$ is the multi-index notation for $\frac{\partial^{n_1} \ldots \partial^{n_d}}{\partial_{x_1} \ldots \partial_{x_d}}$. The corresponding Fourier multiplier is $\hat{L}_N(\omega) = \sum_{|n| < N} a_n |n|! \omega^n$, which is a polynomial of degree $N = |n|$. The operator is elliptic if $\hat{L}_N(\omega)$ vanishes at the origin and nowhere else. More generally, it is called quasi-elliptic of order $\gamma$ if $\hat{L}_N(\omega)$ fulfills the growth condition in Definition 6.3. For $d = 1$, it is fairly easy to determine $\rho_L$ using standard Fourier-inversion techniques (see Chapter 5). Moreover, the condition for quasi-ellipticity of order $N$ is automatically satisfied. When moving to higher dimensions, the study of partial differential operators and the determination of their Green's function becomes more challenging because of the absence of a general multidimensional factorization mechanism. Yet, it is possible to treat special cases in full generality such as the scale-invariant operators (with homogeneous, but not necessarily rotation-invariant, Fourier multipliers), or the class of rotation-invariant operators that are polynomials of the Laplacian ($-\Delta$).

6.2.2 Splines and operators

The foundation of our formulation is the direct correspondence between a spline-admissible operator $L$ and a particular brand of splines.

**Definition 6.4 (Cardinal L-spline)** A function $s(r)$ (possibly of slow growth) is called a cardinal L-spline if and only if

$$Ls(r) = \sum_{k \in \mathbb{Z}^d} a[k] \delta(r - k).$$
The location of the Dirac impulses specifies the spline discontinuities (or knots). The term “cardinal” refers to the particular setting where these are located on the Cartesian grid $\mathbb{Z}^d$.

The remarkable aspect in this definition is that the operator $L$ has the role of a mathematical A-to-D converter since it maps a continuously defined signal $s$ into a discrete sequence $a = (a[k])$. Also note that the weighted sum of Dirac impulses in the r.h.s. of the above equation can be interpreted as the continuous-domain representation of the discrete signal $a$—it is a hybrid-type representation that is commonly used in the theory of linear systems to model ideal sampling (multiplication with a train of Dirac impulses).

The underlying concept of spline is fairly general and it naturally extends to nonuniform grids.

**Definition 6.5 (Nonuniform spline)** Let $\{r_k\}_{k \in S}$ be a set of points (not necessarily finite) that specifies a (nonuniform) grid in $\mathbb{R}^d$. Then, a function $s(r)$ (possibly of slow growth) is a nonuniform L-spline with knots $\{r_k\}_{k \in S}$ if and only if

$$Ls(r) = \sum_{k \in S} a_k \delta(r - r_k).$$

The direct implication of this definition is that a (nonuniform) L-spline with knots $\{r_k\}$ can generally be expressed as

$$s(r) = p_0(r) + \sum_{k \in S} a_k \rho_L(r - r_k),$$

where $\rho_L = L^{-1} \delta$ is the Green’s function of $L$ and $p_0 \in \mathcal{M}_L$ is an appropriate null-space component that is typically selected to fulfill some boundary conditions.

In the case where the grid is uniform, it is usually more convenient to express splines in terms of localized B-spline functions which are shifted replicates of a simple template $\Omega_L$, or some other equivalent generator. An important requirement is that the set of B-spline functions constitutes a Riesz basis.

### 6.2.3 Riesz bases

To quote Ingrid Daubechies [Dau92], a Riesz basis is the next best thing after an orthogonal basis. The reason for not enforcing orthogonality is to leave more room for other desirable features such as simplicity of the construction, maximum localization of the basis function (e.g., compact support), and, last but not least, fast computational solutions.

**Definition 6.6 (Riesz basis)** A sequence of functions $|\phi_k(r)|_{k \in Z}$ in $L_2(\mathbb{R}^d)$ forms a Riesz basis if and only if there exist two constants $A$ and $B$ such that

$$A \|c\|_{\ell_2} \leq \| \sum_{k \in Z} c_k \phi_k(r) \|_{L_2(\mathbb{R}^d)} \leq B \|c\|_{\ell_2}$$

for any sequence $c = (c_k) \in \ell_2$. More generally, the basis is $L_p$-stable if there exist two
6.2 Basic concepts and definitions

constants $A_p$ and $B_p$ such that

$$A_p \|c\|_{\ell_p} \leq \left\| \sum_{k \in \mathbb{Z}} c_k \phi_k(r) \right\|_{L_p(\mathbb{R}^d)} \leq B_p \|c\|_{\ell_p}.$$  

This definition imposes an equivalence between the $L_2$ ($L_p$, resp.) norm of the continuously defined function $s(r) = \sum_{k \in \mathbb{Z}} c_k \phi_k(r)$ and the $\ell_2$ ($\ell_p$, resp.) norm of its expansion coefficients $(c_k)$. This ensures that the representation is stable in the sense that a small perturbation of the expansion coefficients results in a perturbation of comparable magnitude on $s(r)$ and vice versa. Also note that the lower inequality implies that the functions $\{\phi_k\}$ are linearly independent (by setting $s(r) = 0$), which is the defining property of a basis in finite dimensions—but which, on its own, does not ensure stability in infinite dimensions. When $A = B = 1$, we have a perfect norm equivalence which translates into the basis being orthonormal (Parseval’s relation).

Finally, we like to point out that the the existence of the bounds $A$ and $B$ ensures that the (infinite) Gram matrix is positive-definite so that it can be readily diagonalized to yield an equivalent orthogonal basis.

In the (multi-)integer shift-invariant case where the basis functions are given by $\phi_k(r) = \phi(r - k)$, $k \in \mathbb{Z}^d$, there is a simpler equivalent reformulation of the Riesz basis requirement of Definition 6.6.

**Theorem 6.2** Let $\phi(r) \in L_2(\mathbb{R}^d)$ be a B-spline-like generator whose Fourier transform is denoted by $\hat{\phi}(\omega)$. Then, $\{\phi(r - k)\}_{k \in \mathbb{Z}^d}$ forms a Riesz basis with Riesz bounds $A$ and $B$ if and only if

$$0 < A^2 \sum_{n \in \mathbb{Z}^d} |\hat{\phi}(\omega + 2\pi n)|^2 \leq B^2 < \infty$$  \hspace{1cm} (6.17)

for almost every $\omega$. Moreover, the basis is $L_p$-stable for all $1 \leq p \leq +\infty$ if, in addition,

$$\sup_{r \in [0,1]^d} \sum_{k \in \mathbb{Z}^d} |\phi(r - k)| = A_{2,\infty} < +\infty.$$  \hspace{1cm} (6.18)

Under such condition(s), the induced function space

$$V_{\phi} = \left\{ s(r) = \sum_{k \in \mathbb{Z}^d} c[k] \phi(r - k) : c \in \ell_p(\mathbb{Z}^d) \right\}$$

is a closed subspace of $L_p(\mathbb{R}^d)$, including the standard case $p = 2$.

Observe that the central quantity in (6.17) corresponds to the discrete-domain Fourier transform of the Gram sequence $a_{\phi}[k] = \langle \phi(-k), \phi \rangle_{L_2}$. Indeed, we have that

$$A_{\phi}(e^{j\omega}) = \sum_{k \in \mathbb{Z}^d} a_{\phi}[k] e^{-j\omega \cdot k} = \sum_{n \in \mathbb{Z}^d} |\hat{\phi}(\omega + 2\pi n)|^2$$  \hspace{1cm} (6.19)

where the r.h.s. follows from Poisson’s summation formula applied to the sampling at the integers of the autocorrelation function $(\hat{\phi}^\vee * \phi)(r)$. Formula (6.19) is especially advantageous in the case of compactly supported B-splines for which the autocorrelation is often known explicitly (as a B-spline of twice the order) since it reduces the calculation to a finite sum over the support of the Gram sequence (discrete-domain Fourier transform).
Theorem 6.2 is a fundamental result in sampling and approximation theory [Uns00]. It is instructive here to briefly run through the $L_2$ part of the proof which also serves as a refresher on some of the standard properties of the Fourier transform. In particular, we like to emphasize the interaction between the continuous and discrete aspects of the problem.

**Proof** We start by computing the Fourier transform of $s(r) = \sum_{k \in \mathbb{Z}^d} c[k] \phi(r - k)$, which gives

$$\mathcal{F}\{s\}(\omega) = \sum_{k \in \mathbb{Z}^d} c[k] e^{-j\omega \cdot k} \hat{\phi}(\omega)$$  
(by linearity and shift property)

$$= C(e^{j\omega}) \cdot \hat{\phi}(\omega),$$

where $C(e^{j\omega})$ is recognized as the discrete-domain Fourier transform of $c[\cdot]$. Next, we invoke Parseval's identity and manipulate the Fourier-domain integral as follows

$$\|s\|_{L_2}^2 = \int_{\mathbb{R}^d} |C(e^{j\omega})|^2 \left| \hat{\phi}(\omega) \right|^2 \frac{d\omega}{(2\pi)^d}$$

$$= \sum_{n \in \mathbb{Z}^d} \int_{[0,2\pi]^d} |C(e^{j(\omega + 2\pi n)})|^2 \left| \hat{\phi}(\omega + 2\pi n) \right|^2 \frac{d\omega}{(2\pi)^d}$$

$$= \int_{[0,2\pi]^d} \left| C(e^{j\omega}) \right|^2 \sum_{n \in \mathbb{Z}^d} \left| \hat{\phi}(\omega + 2\pi n) \right|^2 \frac{d\omega}{(2\pi)^d}$$

$$= \int_{[0,2\pi]^d} \left| C(e^{j\omega}) \right|^2 A_\phi(e^{j\omega}) \frac{d\omega}{(2\pi)^d}.$$

There, we have used the fact that $C(e^{j\omega})$ is $2\pi$-periodic and the nonnegativity of the integrand to interchange the summation and the integral (Fubini). This naturally leads to the inequality

$$\inf_{\omega \in [0,2\pi]^d} A_\phi(e^{j\omega}) \cdot \|c\|^2_{L_2} \leq \|s\|^2_{L_2} \leq \sup_{\omega \in [0,2\pi]^d} A_\phi(e^{j\omega}) \cdot \|c\|^2_{L_2},$$

where we are now making use of Parseval's identity for sequences, so that

$$\|c\|^2_{L_2} = \int_{[0,2\pi]^d} \left| C(e^{j\omega}) \right|^2 \frac{d\omega}{(2\pi)^d}.$$

The final step is to show that these bounds are sharp. This can be accomplished through the choice of some particular (bandlimited) sequence $c[\cdot]$. \qed

Note that the almost everywhere part of (6.17) can be dropped when $\phi \in L_1(\mathbb{R}^d)$ because the Fourier transform of such a function is continuous (Riemman-Lebesgue Lemma).

While the result of Theorem 6.2 is restricted to the classical $L_p$ spaces, there is no fundamental difficulty in extending it to wider classes of weighted (with negative powers) $L_p$ spaces by imposing some stricter condition than (6.18) on the decay of $\phi$. For instance, if $\phi$ has exponential decay, then the definition of the function space $V_\phi$ can be extended for all sequences $c$ that are growing no faster than a polynomial. This happens to be the appropriate framework for sampling generalized stochastic processes which do not live in the $L_p$ spaces since they are not decaying at infinity.
6.2.4 Admissible wavelets

The other important tool for analyzing stochastic processes is the wavelet transform whose basis functions must be “tuned” to the object under investigation.

**Definition 6.7** A wavelet function $\psi$ is called $L$-admissible if it can be expressed as $\psi = L^H \phi$ with $\phi \in L_1(\mathbb{R}^d)$.

Observe that we are now considering the Hermitian transpose operator $L^H = L^*$ which is distinct from the adjoint operator $L^*$ when the impulse response has some imaginary component. The reason for this is that the wavelet-analysis step involves a Hermitian inner product $\langle \cdot, \cdot \rangle_{L^2}$ whose definition differs by a complex conjugation from that of the distributional scalar product $\langle \cdot, \cdot \rangle_{H^2}$ used in our formulation of stochastic processes when the second argument is complex valued; specifically, $\langle f, g \rangle_{L^2} = \langle f, \overline{g} \rangle_{L^2} = \int_{\mathbb{R}^d} \overline{f(r)} g(r) \, dr$.

The best matched wavelet is the one for which the wavelet kernel $\phi$ is the most localized—ideally, the shortest-possible support assuming that it is at all possible to construct a compactly-supported wavelet basis. The very least is that $\phi$ should be concentrated around the origin and exhibit a sufficient rate of decay; for instance, $|\phi(r)| \lesssim \frac{C}{1 + |r|^\alpha}$ for some $\alpha > d$.

A direct implication of Definition 6.7 is that the wavelet $\psi$ will annihilate all the components (e.g., polynomials) that are in the null space of $L$ because $\langle \psi(\cdot - r_0), p_0 \rangle = \langle \phi(\cdot - r_0), Lp_0 \rangle = 0$, for all $p_0 \in \mathcal{H}_1$ and $r_0 \in \mathbb{R}^d$. In conventional wavelet theory, this behavior is achieved by designing “$N$th-derivative-like” wavelets with vanishing moments up to polynomial degree $N - 1$.

6.3 First-order exponential B-splines and wavelets

Rather than aiming for the highest level of generality right away, we propose to first examine the 1-D first-order scenario in some detail. First-order differential models are important theoretically because they go hand in hand with the Markov property. In that respect, they constitute the next level of generalization just beyond the Lévy processes. Mathematically, the situation is still quite comparable to that of the derivative operator in the sense that it leads to a nice and self-contained construction of (exponential) B-splines and wavelets. The interesting aspect, though, is that the underlying basis functions are no longer conventional wavelets that are dilated versions of a single prototype: they now fall into the lesser-known category of non-stationary wavelets.

The (causal) Green’s function of our canonical first-order operator $P_a = (D - a I)$ is identical to the impulse response $p_a$ of the corresponding differential system, while the (one-dimensional) null space of the operator is given by $\mathcal{N}_a = \{ a e^{at} : a_0 \in \mathbb{R} \}$.

---

2. In the terminology of wavelets, the term “non-stationary” refers to the property that the shape of the wavelet changes with scale, but not with respect to the location, as the more usual statistical meaning of the term would suggest.
Some examples of such Green's functions are shown in Figure 6.3. The case $\alpha = 0$ (red curve) is the classical one already treated in Section 6.1.

6.3.1 B-spline construction

The natural discrete approximation of the differential operator $P_\alpha = D - \alpha I$ is the first-order weighted difference operator

$$\Delta_\alpha s(r) = s(r) - e^{\alpha} s(r-1). \quad (6.20)$$

Observe that $\Delta_\alpha$ annihilates the exponentials $a_0 e^{\alpha r}$ so that its null space includes $\mathcal{N}_\alpha$. The corresponding B-spline is obtained by applying $\Delta_\alpha$ to $\rho_\alpha$, which yields

$$\beta_\alpha(r) = \mathcal{F}^{-1} \left\{ \frac{1 - e^{\alpha} e^{-j\omega}}{j\omega - \alpha} \right\} (r) = \begin{cases} e^{\alpha r}, & \text{for } 0 \leq r < 1 \\ 0, & \text{otherwise.} \end{cases} \quad (6.21)$$

In effect, the localization by $\Delta_\alpha$ results in a “chopped off” version of the causal Green's function that is restricted to the interval $[0,1)$ (see Figure 6.3b). Importantly, the scheme remains applicable in the unstable scenario $\text{Re}(\alpha) \geq 0$. It always results in a well-defined Fourier transform due to the convenient pole-zero cancellation in the central expression of (6.21). The marginally unstable case $\alpha = 0$ results in the rectangular function shown in red in Figure 6.3, which is the standard basis function for representing piecewise-constant signals. Likewise, $\beta_\alpha$ generates an orthogonal basis for the space of cardinal $P_\alpha$-splines in accordance with Definition 6.4. This allows us
to specify our prototypical exponential spline space as \( V_0 = \text{span}(\beta_a(\cdot - k))_{k \in \mathbb{Z}} \) with knot spacing \( 2^0 = 1 \).

### 6.3.2 Interpolator in augmented-order spline space

The second important ingredient is the interpolator for the “augmented-order” spline space generated by the autocorrelation \((\hat{\beta}_a \ast \hat{\beta}_a)(r)\) of the B-spline. Constructing it is especially easy in the first-order case because it involves the simple normalization

\[
\varphi_{\text{int}, a}(r) = \frac{1}{(\hat{\beta}_a \ast \hat{\beta}_a)(0)} (\hat{\beta}_a \ast \hat{\beta}_a)(r)
\]  

(6.22)

Specifically, \( \varphi_{\text{int}, a} \) is the unique cardinal \( P^H \) \( P_a \)-spline function that vanishes at all the integers except at the origin where it takes the value one (see Figure 6.3c). Its classical use is to provide a sinc-like kernel for the representation of the corresponding family of splines, and also for the reconstruction of spline-related signals, including special brands of stochastic processes, from their integer samples [UB05b]. Another remarkable and lesser known property is that this function provides the proper smoothing kernel for defining an operator-like wavelet basis.

### 6.3.3 Differential wavelets

In the generalized spline framework, instead of specifying a hierarchy of multiresolution subspaces of \( L_2(\mathbb{R}) \) (the space of finite-energy functions) via the dilation of a scaling function, one considers the fine-to-coarse sequence of \( L \)-spline spaces

\[
V_i = \{ s(r) \in L_2(\mathbb{R}) : Ls(r) = \sum_{k \in \mathbb{Z}} a_i[k] \delta(r - 2^i k) \},
\]

where the embedding \( V_i \supset V_j \) for \( i \leq j \) is obvious from the (dyadic) hierarchy of spline knots, so that \( s_j \in V_j \) implies that \( s_j \in V_i \) with an appropriate subset of its coefficients \( a_i[k] \) being zero.

We now detail the construction of a wavelet basis at resolution 1 such that \( W_1 = \text{span}(\psi_{1,k})_{k \in \mathbb{Z}} \) with \( W_1 \perp V_1 \) and \( V_1 + W_1 = V_0 = \text{span}(\beta_a(\cdot - k))_{k \in \mathbb{Z}} \). The recipe is to take \( \psi_{1,k}(r) = \psi_a(r - 1 - 2k)/\|\psi_a\|_{L^2} \) where \( \psi_a \) is the mother wavelet given by

\[
\psi_a(r) = P^H_a \varphi_{\text{int}, a}(r) \propto \Delta_a^H \beta_a(r).
\]

There, \( \Delta_a^H \) is the Hermitian adjoint of the finite-difference operator \( \Delta_a \). Examples of such exponential-spline wavelets are shown in Figure 6.3d, including the classical Haar wavelet (up to a sign change) which is obtained for \( a = 0 \) (red curve). The basis functions \( \psi_{1,k} \) are shifted versions of \( \psi_a \) that are centered at the odd integers and normalized to have a unit norm. Since these wavelets are non-overlapping, they form an orthonormal basis. Moreover, the basis is orthogonal to the coarser spline space \( V_1 \) as a direct consequence of the interpolating property of \( \varphi_{\text{int}, a} \) (Proposition 6.6 in
Section 6.5). Finally, based on the fact that \( \psi_{1,k} \in V_0 \) for all \( k \in \mathbb{Z} \), one can show that these wavelets span \( W_1 \), which translates into

\[
W_1 = \left\{ v(r) = \sum_{k \in \mathbb{Z}} v_1[k] \psi_{1,k}(r) : v_1 \in \ell_2(\mathbb{Z}) \right\}.
\]

This method of construction extends to the other wavelet subspaces \( W_i \) provided that the interpolating kernel \( \varphi_{int,a} \) is substituted by its proper counterpart at resolution \( a = 2^{i-1} \) and the sampling grid adjusted accordingly. Ultimately, this results in a wavelet basis of \( L_2(\mathbb{R}) \) whose members are all \( \text{P}_a \)-splines—that is, piecewise-exponential with parameter \( a \)—but not dilates of the same prototype unless \( a = 0 \). Otherwise, the corresponding decomposition is not fundamentally different from a conventional wavelet expansion. The basis functions are equally well localized and the scheme admits the same type of fast reversible filterbank algorithm, albeit with scale-dependent filters [KU06].

### 6.4 Generalized B-spline basis

The procedure of Section 6.3.1 remains applicable for the broad class of spline-admissible operators (see Definition 6.2) in one or multiple dimensions. The two ingredients for constructing a generalized B-spline basis are: 1) the knowledge of the Green's function \( \Omega_L \) of the operator \( L \), and 2) the availability of a discrete approximation (finite-difference-like) of the operator of the form

\[
L_d s(r) = \sum_{k \in \mathbb{Z}^d} d_k[k] s(r-k)
\]

with \( d_k \in \ell_1(\mathbb{Z}^d) \) that fulfills the null-space matching constraint \(^3\)

\[
L_d p_0(r) = L p_0(r) = 0 \quad \text{for all } p_0 \in \mathcal{M}_L.
\]

The generalized B-spline associated with the operator \( L \) is then given by

\[
\beta_L(r) = L_d \rho_L(r) = \mathcal{F}^{-1} \left\{ \sum_{k \in \mathbb{Z}^d} d_k[k] e^{-j(k,\omega)} \frac{L(\omega)}{\hat{L}(\omega)} \right\}(r),
\]

where the numerator and denominator in the r.h.s. expression correspond to the frequency responses of \( L_d \) and \( L \), respectively. The null-space matching constraint is especially helpful for the unstable cases where \( \rho_L \notin L_1(\mathbb{R}^d) \): it ensures that the zeros of \( \hat{L}(\omega) \) (singularities) are cancelled by some corresponding zeros of \( \hat{L}_d(\omega) \) so that the Fourier transform of \( \beta_L \) remains bounded.

**Definition 6.8** The function \( \beta_L \) specified by (6.25) is an admissible B-spline for \( L \) if and only if 1) \( \beta_L \in L_2(\mathbb{R}^d) \cap L_2(\mathbb{R}^d) \), and 2) it generates a Riesz basis of the space of cardinal \( L \)-splines.

\(^3\) We want the null space of \( L_d \) to include \( \mathcal{M}_L \) and to remain the smallest possible. In that respect, it is worth noting that the null space of a discrete operator will always be much larger than that of its continuous-domain counterpart. For instance, the derivative operator \( D \) suppresses constant signals, while its finite-difference counterpart annihilates all 1-periodic functions, including the constants.
In light of Theorem 6.2, the latter property requires the existence of the two Riesz bounds $A$ and $B$ such that

$$0 < A^2 \leq \sum_{n \in \mathbb{Z}^d} |\hat{\beta}_L(\omega + 2\pi n)|^2 = \frac{\sum_{k \in \mathbb{Z}^d} d_k[k]e^{-j(k,\omega)}}{\sum_{n \in \mathbb{Z}^d} |L(\omega + 2\pi n)|^2} \leq B^2; \quad (6.26)$$

A direct consequence of (6.25) is that

$$L\beta_L(r) = \sum_{k \in \mathbb{Z}^d} d_k[k]\delta(r - k) \quad (6.27)$$

so that $\beta_L$ is itself a cardinal L-spline in accordance with Definition 6.4. The bottom line in Definition 6.8 is that any cardinal L-spline admits a unique representation in the B-spline basis ($\beta_L(\cdot - k)$)$_{k \in \mathbb{Z}^d}$ as

$$s(r) = \sum_{k \in \mathbb{Z}^d} c[k]\beta_L(r - k) \quad (6.28)$$

where the $c[k]$ are the B-spline coefficients of $s$.

While (6.25) provides us with a nice recipe for constructing B-splines, it does not guarantee that the Riesz-basis condition (6.26) is satisfied. This needs to be established on a case-by-case basis. The good news for the present theory of stochastic processes is that B-splines are available for virtually all the operators that have been discussed so far.

### 6.4.1 B-spline properties

To motivate the use of B-splines, we shall first restrict our attention to the space $V_L$ of cardinal L-splines with finite energy, which is formally defined as

$$V_L = \left\{ s(r) \in L_2(\mathbb{R}^d) : s(r) = \sum_{k \in \mathbb{Z}^d} a[k]\delta(r - k) \right\}. \quad (6.29)$$

The foundation of spline theory is that there are two complementary ways of representing splines using different types of basis functions: Green's functions versus B-splines. The first representation follows directly from the Definition 6.4 (see also (6.16)) and is given by

$$s(r) = p_0(r) + \sum_{k \in \mathbb{Z}^d} a[k]\rho_L(r - k), \quad (6.30)$$

where $p_0 \in \mathcal{M}_L$ is a suitable element of the null space of $L$ and where $\rho_L = L^{-1}\delta$ is the Green's function of the operator. The functions $\rho_L(\cdot - k)$ are nonlocal and very far from being orthogonal. In many cases, they are not even part of $V_L$, which raises fundamental issues concerning the $L_2$ convergence of the infinite sum $^4$ in (6.30) and the conditions that must be imposed upon the expansion coefficients $a[\cdot]$. The second type of B-spline expansion (6.28) does not have such stability problems. This is the primary reason why it is favored by practitioners.

$^4$ Without further assumptions on $\rho_L$ and $a$, (6.30) is only valid in the weak sense of distributions.
Stable representation of cardinal L-splines

The equivalent B-spline specification of the space \( V_L \) of cardinal splines is

\[
V_L = \left\{ s(r) = \sum_{k \in \mathbb{Z}^d} c[k] \beta_L(r - k) : c[\cdot] \in \ell_2(\mathbb{Z}^d) \right\},
\]

where the generalized B-spline \( \beta_L \) satisfies the conditions in Definition 6.8. The Riesz-basis property ensures that the representation is stable in the sense that, for all \( s \in V_L \), we have that

\[
A \| c \|_{\ell_2} \leq \| s \|_{L^2} \leq B \| c \|_{\ell_2}.
\]

There, \( \| c \|_{\ell_2} = \left( \sum_{k \in \mathbb{Z}^d} |c[k]|^2 \right)^{\frac{1}{2}} \) is the \( \ell_2 \)-norm of the B-spline coefficients \( c \). The fact that the underlying functions are cardinal L-splines is a simple consequence of the atoms being splines themselves. Moreover, we can easily make the link with (6.30) by using (6.27), which yields

\[
L_s(r) = \sum_{k \in \mathbb{Z}^d} c[k] \beta_L(r - k) = \sum_{k \in \mathbb{Z}^d} (c * d_L)[k] \delta(r - k).
\]

The less obvious aspect, which is implicit in the definition of the B-spline, is the completeness of the representation in the sense that the B-spline basis spans the space \( V_L \) defined by (6.29). We shall establish this by showing that the B-splines are capable of reproducing \( \rho_L \) as well as any component \( p_0 \in \mathcal{N}_L \) in the null space of \( L \). The implication is that any function of the form (6.30) admits a unique expansion in a B-spline basis. This is also true when the function is not in \( L^2(\mathbb{R}^d) \), in which case the B-spline coefficients \( c \) are no longer in \( \ell_2(\mathbb{Z}^d) \) due to the discrete-continuous norm equivalence (6.31).

Reproduction of Green's functions

The reproduction of Green's functions follows from the special form of (6.25). To reveal it, we consider the inverse \( L^{-1}_d \) of the discrete localization operator \( L_d \) specified by (6.23), whose continuous-domain impulse response is written as

\[
L^{-1}_d \delta(r) = \sum_{k \in \mathbb{Z}^d} p[k] \delta(r - k) = \mathcal{F}^{-1} \left\{ \frac{1}{\sum_{k \in \mathbb{Z}^d} d_L[k] e^{-j[k,\omega]} \right\}.
\]

The sequence \( p \), which can be determined by generalized inverse Fourier transform, is of slow growth with the property that \( (p * d)[k] = \delta[k] \). The Green's function reproduction formula is then obtained by applying \( L^{-1}_d \) to the B-spline \( \beta_L \) and making use of the left-inverse property of \( L^{-1}_d \). Thus,

\[
L^{-1}_d \beta_L(r) = L^{-1}_d L_d \rho_L(r) = \rho_L(r)
\]

results into

\[
\rho_L(r) = \sum_{k \in \mathbb{Z}^d} p[k] \beta_L(r - k).
\]

To illustrate the concept, let us get back to our introductory example in Section in
6.4 Generalized B-spline basis

6.3.1 with \( L = P_\alpha = (D - \alpha I) \) where \( \text{Re}(\alpha) < 0 \). The frequency response of this first-order operator is

\[
\tilde{P}_\alpha(\omega) = j\omega - \alpha,
\]

while its Green's function is given by

\[
\rho_\alpha(r) = l_+(r) e^{ar} = \mathcal{F}^{-1} \left\{ \frac{1}{j\omega - \alpha} \right\}(r).
\]

On the discrete side of the picture, we have the finite-difference operator \( \Delta_\alpha \) with

\[
\Delta_\alpha(\omega) = 1 - e^{\alpha j\omega},
\]

and its inverse \( \Delta^{-1}_\alpha \) whose expansion coefficients are

\[
p_\alpha[k] = l_+[k] e^{\alpha k} = \mathcal{F}_d^{-1} \left\{ \frac{1}{1 - e^{\alpha e^{-j\omega}}} \right\}[k],
\]

where \( \mathcal{F}_d^{-1} \) denotes the discrete-domain inverse Fourier transform \(^5\). The application of (6.32) then yields the exponential-reproduction formula

\[
I_+(r) e^{ar} = \sum_{k=0}^{\infty} e^{\alpha k} \beta_\alpha(t - k)
\]

where \( \beta_\alpha \) is the exponential B-spline defined by (6.21). Note that the range of applicability of (6.33) extends to \( \text{Re}(\alpha) \leq 0 \).

**Reproduction of null-space components**

A fundamental property of B-splines is their ability to reproduce the components that are in the null space of their defining operator. In the case of our working example, we can simply extrapolate (6.33) for negative indices, which yields

\[
e^{ar} = \sum_{k \in \mathbb{Z}} e^{\alpha k} \beta_\alpha(r - k).
\]

It turns out that this reproduction property is induced by the matching null-space constraint (6.24) that is imposed upon the localization filter. While the reproduction of exponentials is interesting in its own right, we shall focus here on the important case of polynomials and provide a detailed Fourier-based analysis. We start by recalling that the general form of a multidimensional polynomial of total degree \( N \) is

\[
q_N(r) = \sum_{|n| \leq N} a_n r^n
\]

using the multi-index notation with \( n = (n_1, \ldots, n_d) \in \mathbb{N}^d \), \( r^n = r_1^{n_1} \cdots r_d^{n_d} \), and \( |n| = n_1 + \cdots + n_d \). The generalized Fourier transform of \( q_N \in \mathcal{S}'(\mathbb{R}^d) \) (see Table 3.3 and entry \( r^n f(r) \) with \( f(r) = 1 \)) is given by

\[
\hat{q}_N(\omega) = \sum_{|n| \leq N} (2\pi)^d \alpha^n |n|! |\varepsilon^n(\omega)|.
\]

\(^5\) Our definition of the inverse discrete Fourier transform in 1-D is

\[
\mathcal{F}_d^{-1} \left\{ H(\omega) \right\}[k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\omega) e^{i\omega k} d\omega \text{ with } k \in \mathbb{Z}.\]
where \( \partial^n \delta \) denotes the \( n \)th partial derivative of the multidimensional Dirac impulse \( \delta \). Hence, the Fourier multiplier \( \tilde{L} \) will annihilate the polynomials of order \( N \) if and only if \( \tilde{L}(\omega) \partial^n \delta(\omega) = 0 \) for all \( |n| \leq N \). To understand when this condition is met, we expand \( \tilde{L}(\omega) \partial^n \delta(\omega) \) in terms of \( \partial^k \tilde{L}(0), |k| \leq |n| \) by using the general product rule for the manipulation of Dirac impulses and their derivatives given by

\[
f(r) \partial^n \delta(r - r_0) = \sum_{k+l=n} \frac{n!}{k!l!} (-1)^{n+l} \partial^k f(r_0) \partial^l \delta(r - r_0).
\]

The latter follows from Leibnitz’ rule for partially differentiating a product of functions as

\[
\partial^n(f \varphi) = \sum_{k+l=n} \frac{n!}{k!l!} \partial^k f \partial^l \varphi,
\]

and the adjoint relation \( \langle \varphi, f \partial^n \delta(\cdot - r_0) \rangle = \langle \partial^n \varphi, \delta(\cdot - r_0) \rangle \) with \( \partial^n = (-1)^{|n|} \partial^n \).

This allows us to conclude that the necessary and sufficient condition for the inclusion of the polynomials of order \( N \) in the null space of \( L \) is

\[
\partial^n \tilde{L}(0) = 0, \text{ for all } n \in \mathbb{Z}^d \text{ with } |n| \leq N,
\]

which is equivalent to \( \tilde{L}(\omega) = O(\|\omega\|^{N+1}) \) around the origin. Note that this behavior is prototypical of scale-invariant operators such as fractional derivatives and Laplacians. The same condition has obviously to be imposed upon the localization filter \( \tilde{L}_d \) for the Fourier transform of the B-spline in (6.25) to be nonsingular at the origin. Since \( \tilde{L}_d(\omega) \) is 2\( \pi \)-periodic, we have that

\[
\partial^n \tilde{L}_d(2\pi k) = 0, \quad k \in \mathbb{Z}^d, n \in \mathbb{N}^d \text{ with } |n| \leq N.
\]

For practical convenience, we shall assume that the B-spline \( \beta_L \) is normalized to have a unit integral\(^6\) so that \( \hat{\beta}_L(0) = 1 \). Based on (6.35) and \( \hat{\beta}_L(\omega) = \tilde{L}_d(\omega)/\tilde{L}(\omega) \), we find that

\[
\begin{align*}
\hat{\beta}_L(0) &= 1, \\
\partial^n \hat{\beta}_L(2\pi k) &= 0, \quad k \in \mathbb{Z}^d \setminus \{0\}, n \in \mathbb{N}^d \text{ with } |n| \leq N,
\end{align*}
\]

which are the so-called Strang-Fix conditions of order \( N \). Recalling that \( j^{n} \partial^n \hat{\beta}_L(\omega) \) is the Fourier transform of \( r^n \beta_L(r) \) and that periodization in the signal domain corresponds to a sampling in the Fourier domain, we finally deduce that

\[
\sum_{k \in \mathbb{Z}^d} (r - k)^n \beta_L(r - k) = j^n \partial^n \hat{\beta}_L(0) = C_n, \quad n \in \mathbb{N}^d \text{ with } 0 < |n| \leq N,
\]

with the implicit assumption that \( \beta_L \) has a sufficient order of algebraic decay for the above sums to be convergent. The special case of (6.37) with \( n = 0 \) reads

\[
\sum_{k \in \mathbb{Z}^d} \beta_L(r - k) = 1
\]

\(^6\) This is always possible thanks to Condition (6.24), which ensures that \( \hat{\beta}_L(0) \neq 0 \) due to a proper cancellation of poles and zeros in the r. h. s. of (6.25).
and is called the partition of unity. It reflects the fact that $\beta_L$ reproduces the constants. More generally, Condition (6.37) (or (6.36)) is equivalent to the existence of sequences $p_n$ such that

$$r^n = \sum_{k \in \mathbb{Z}^d} p_n[k] \beta_L(r - k) \quad \text{for all } |n| \leq N,$$

(6.39)

which is a more direct statement of the polynomial-reproduction property. For instance, (6.37) with $n = (1, \ldots, 1)$ implies that

$$\sum_{k \in \mathbb{Z}^d} p_n[k] \beta_L(r - k) = C_{(1, \ldots, 1)}$$

from which one deduces that $p_{(1, \ldots, 1)}[k] = k + C_{(1, \ldots, 1)}$. The other sequences $p_n$, which are polynomials in $k$, may be determined in a similar fashion by proceeding recursively. Another equivalent way of stating the Strang-Fix conditions of order $N$ is that the sums

$$\sum_{k \in \mathbb{Z}^d} k^r \beta_L(r - k) = \sum_{k \in \mathbb{Z}^d} j^{|n|} \beta_L(\omega) \beta_L(-\omega)\big|_{\omega = 2\pi t}$$

are polynomials with leading term $r^n$ for all $|n| \leq N$. The left-hand-side expression follows from Poisson’s summation formula\footnote{The standard form of Poisson’s summation formula is $\sum_{k \in \mathbb{Z}^d} f(k) = \sum_{l \in \mathbb{Z}^d} \hat{f}(2\pi l)$: It is valid for any Fourier pair $f, \hat{f} = \mathcal{F}(f) \in L_1(\mathbb{R}^d)$ with sufficient decay for the two sums to be convergent.} applied to the function $f(x) = x^n \beta_L(r - x)$ with $r$ being considered as a constant shift.

**Localization**

The guiding principle for designing B-splines is to produce basis functions that are maximally localized on $\mathbb{R}^d$. Ideally, B-splines should have the smallest possible support which is the property that makes them so useful in applications. When it is not possible to construct compactly supported basis functions, the B-spline should at least be concentrated around the origin and satisfy some decay bound with the tightest possible constants. The primary types of spatial localization, by order of preference, are

1) Compact support: $\beta_L(r) = 0$ for all $r \notin \Omega$ where $\Omega \subset \mathbb{R}^d$ is a convex set with the smallest-possible Lebesgue measure.

2) Exponential decay: $|\beta_L(r - r_0)| \leq C \exp(-a|r|)$ for some $r_0 \in \mathbb{R}^d$ and the largest-possible $a \in \mathbb{R}^+$.

3) Algebraic decay: $|\beta_L(r - r_0)| \leq C \frac{1}{1 + |r|^2a}$ for some $r_0 \in \mathbb{R}^d$ and the largest possible $a \in \mathbb{R}^+$.

By relying on the classical relations that link spatial decay to the smoothness of the Fourier transform, one can get a good estimate of spatial decay based on the knowledge of the Fourier transform $\hat{\beta}_L(\omega) = \tilde{\beta}_L(\omega)/\tilde{\beta}(\omega)$ of the B-spline. Since the localization filter $\tilde{\beta}_L(\omega)$ acts by compensating the (potential) singularities of $\tilde{\beta}(\omega)$, the
guiding principle is that the rate of decay is essentially determined by the degree of differentiability of $\hat{L}(\omega)$.

Specifically, if $\hat{\beta}_L(\omega)$ is differentiable up to order $N$, then the B-spline $\beta_L$ is guaranteed to have an algebraic decay of order $N$. To show this, we consider the Fourier transform pair $r^n \beta_L(r) \leftrightarrow j^n \partial^n \hat{\beta}_L$, subject to the constraint that $\partial^n \hat{\beta}_L \in L_1(\mathbb{R}^d)$ for all $|n| < N$. From the definition of the inverse Fourier integral, it immediately follows that

$$|r^n \beta_L(r)| \leq \frac{1}{(2\pi)^d} \|\partial^n \hat{\beta}_L\|_{L_1},$$

which, when properly combined over all multi-indices $|n| < N$, yields an algebraic decay estimate with $\alpha = N$. By pushing the argument to the limit, we see that exponential decay (which is faster than any order of algebraic decay) requires that $\hat{\beta}_L \in C^\infty(\mathbb{R}^d)$ (infinite order of differentiability), which is only possible if $\hat{L}(\omega) \in C^\infty(\mathbb{R}^d)$ as well.

The ultimate limit in Fourier-domain regularity is when $\hat{\beta}_L$ has an analytic extension that is an entire function. In fact, by the Paley-Wiener theorem (Theorem 6.3 below), one achieves compact support of $\beta_L$ if and only if $\hat{\beta}_L(\zeta)$ is an entire function of exponential type. To explain this concept, we focus on the one-dimensional case where the B-spline $\beta_L$ is supported in the finite interval $[-A, +A]$. We then consider the holomorphic Fourier (or Fourier-Laplace) transform of the B-spline given by

$$\hat{\beta}_L(\zeta) = \int_{-A}^{+A} \beta_L(r)e^{-\zeta r} \, dr \quad (6.40)$$

with $\zeta = \sigma + j\omega \in \mathbb{C}$, which formally amounts to substituting $j\omega$ by $\zeta$ in the expression of the Fourier transform of $\beta_L$. In order to obtain a proper analytic extension, we need to verify that $\hat{\beta}_L(\zeta)$ satisfies the Cauchy-Riemann equation. We shall do so by applying a dominated-convergence argument. To that end, we construct the exponential bound

$$|\hat{\beta}_L(\zeta)| \leq e^{A|\zeta|} \int_{-A}^{+A} |\beta_L(r)| \, dr \leq e^{A|\zeta|} \sqrt{\int_{-A}^{+A} 1 \, dr} \sqrt{\int_{-A}^{+A} |\beta_L(r)|^2 \, dr} = e^{A|\zeta|} \sqrt{2A} \|\beta_L\|_{L_2}$$

where we have applied Cauchy-Schwarz’ inequality to derive the lower inequality. Since $e^{-\zeta r}$ for $r$ fixed is itself an entire function and (6.40) is convergent over the whole complex plane, the conclusion is that $\hat{\beta}_L(\zeta)$ is entire as well, in addition to being a function of exponential type $A$ as indicated by the bound. The whole strength of the Paley-Wiener theorem is that the implication also works the other way around.

**Theorem 6.3 (Paley-Wiener)** Let $f \in L_2(\mathbb{R})$. Then, $f$ is compactly supported in $\mathbb{R}$.
Generalized B-spline basis

\[-A, A\] if and only if its Laplace transform

\[F(\xi) = \int_{\mathbb{R}} f(r) e^{-\xi r} \, dr\]

is an entire function of exponential type \(A\), meaning that there exists a constant \(C\) such that

\[|F(\xi)| \leq Ce^{A|\xi|}\]

for all \(\xi \in \mathbb{C}\).

The result implies that one can deduce the support of \(f\) from its Laplace transform. We can also easily extend the result to the case where the support is not centered around the origin by applying the Paley-Wiener theorem to the autocorrelation function \((f * f^*) (r)\). The latter is supported in the interval \([-2A, 2A]\), which is twice the size of the support of \(f\) irrespective of its center. This suggests the following expression for the determination of the support of a B-spline:

\[
\text{support}(\hat{\beta}_L) = \limsup_{R \to \infty} R \sup_{|\xi| \leq R} \left| \hat{\beta}_L(\xi) \hat{\beta}_L(-\xi) \right|
\]

(6.41)

It returns twice the exponential type of the recentered B-spline which gives \(\text{support}(\hat{\beta}_L) = 2A\). While this formula is only strictly valid when \(\hat{\beta}_L(\xi)\) is an entire function, it can be used otherwise as an operational measure of localization when the underlying B-spline is not compactly supported. Interestingly, (6.41) provides a measure that is additive with respect to convolution and proportional to the order \(\gamma\). For instance, the support of an (exponential) B-spline associated with an ordinary differential operator of order \(N\) is precisely \(N\), as a consequence of the factorization property of such B-splines (see Sections 6.4.2 and 6.4.4).

To get some insight into (6.41), let us consider the case of the polynomial B-spline of order 1 (or degree 0) with \(\beta_D(r) = \mathbb{1}_{[0,1]}(r)\) and Laplace transform

\[
\hat{\beta}_D(\xi) = \left\{ \frac{1 - e^{-\xi}}{\xi} \right\}
\]

The required product in (6.41) is

\[
\hat{\beta}_D(\xi) \hat{\beta}_D(-\xi) = \frac{-e^\xi + 2 - e^{-\xi}}{\xi^2},
\]

which is analytic over the whole complex plane because of the pole-zero cancellation at \(\xi = 0\). For \(R\) sufficiently large, we clearly have that

\[
\max_{|\xi| \leq R} \left| \hat{\beta}_D(\xi) \hat{\beta}_D(-\xi) \right| = \frac{e^R + 2 + e^{-R}}{R^2} = \frac{e^R}{R^2}
\]

By plugging the above expression in (6.41), we finally get

\[
\text{support}(\hat{\beta}_D) = \limsup_{R \to \infty} \frac{R - 2\log R}{R} = 1,
\]
which is the desired result. While the above calculation may look like overkill for the determination of the already-known support of $\hat{\beta}_D$, it becomes quite handy for making predictions for higher-order operators. To illustrate the point, we now consider the B-spline of order $\gamma$ associated with the (possibly fractional) derivative operator $D^\gamma$ whose Fourier-Laplace transform is

$$\hat{\beta}_{D^\gamma}(\zeta) = \left(\frac{1 - e^{-\zeta}}{\zeta}\right)^\gamma.$$ 

We can then essentially replicate the previous manipulation while moving the order out of the logarithm to deduce that

$$\text{support}(\hat{\beta}_{D^\gamma}) = \limsup_{R \to \infty} \gamma R - 2\gamma \log R = \gamma.$$ 

This shows that the “support” of the B-spline is equal to its order, with the caveat that the underlying Fourier-Laplace transform $\hat{\beta}_{D^\gamma}(\zeta)$ is only analytic (and entire) when the order $\gamma$ is a positive integer. This points to the fundamental limitation that a B-spline associated with a fractional operator—that is, when $L(\omega)$ is not an entire function—cannot be compactly supported.

**Smoothness**

The smoothness of a B-spline refers to its degree of continuity and/or differentiability. Since a B-spline is a linear combination of shifted Green’s functions, its smoothness is the same as that of $\rho_L$.

Smoothness descriptors come in two flavors—Hölder continuity versus Sobolev differentiability—depending on whether the analysis is done in the signal or Fourier domain. Due to the duality between Fourier decay and order of differentiation, the smoothness of $\beta_L$ may be predicted from the growth of $L(\omega)$ at infinity without need for the explicit calculation of $\rho_L$. To that end, one considers the Sobolev spaces $W^a_2$ which are defined as

$$W^a_2(\mathbb{R}^d) = \left\{ f : \int_{\mathbb{R}^d} (1 + \|\omega\|^2)^a |\hat{f}(\omega)|^2 \, d\omega < \infty \right\}.$$ 

Since the partial differentiation operator $\partial^\alpha$ corresponds to a Fourier-domain multiplication by $(i\omega)^\alpha$, the inclusion of $f$ in $W^a_2(\mathbb{R}^d)$ requires that its (partial) derivatives be well-defined in the $L_2$ sense up to order $a$. The same is also true for the “Bessel potential” operators $(\mathrm{id} - \Delta)^{\alpha/2}$ of order $\alpha$, or, alternatively, the fractional Laplacians $(-\Delta)^{\alpha/2}$ with Fourier multiplier $\|\omega\|^\alpha$.

**Proposition 6.4** Let $\beta_L$ be an admissible B-spline that is associated with a Fourier multiplier $L(\omega)$ of order $\gamma$. Then, $\beta_L \in W^a_2(\mathbb{R}^d)$ for any $\alpha < \gamma - \frac{d}{2}$.

**Proof** Because of Parseval’s identity, the statement $\beta_L \in W^a_2(\mathbb{R}^d)$ is equivalent to $\beta_L \in L_2(\mathbb{R}^d)$ and $(-\Delta)^{\alpha/2}\beta_L \in L_2(\mathbb{R}^d)$. Since the first inclusion is part of the definition,  

9 The underlying Fourier multipliers are of comparable size in the sense that there exist two constants $c_1$ and $c_2$ such that $c_1(1 + \|\omega\|^2)^\alpha \leq 1 + \|\omega\|^2 \leq c_2(1 + \|\omega\|^2)^\alpha$. 

9
it is sufficient to check for the second. To that end, we recall the stability conditions
\( \beta_l \in L_1(\mathbb{R}^d) \) and \( d_l \in \ell_1(\mathbb{R}^d) \), which are implicit to the B-spline construction (6.25).

These, together with the order condition (6.14), imply that
\[
|\hat{\beta}_l(\omega)| = \frac{\hat{L}_d(\omega)}{\hat{L}(\omega)} \leq \min \left( \|\beta_l\|_{L_1}, C \frac{\|d_l\|_{\ell_1}}{\|\omega\|^r} \right).
\]

This latter bound allows us to control the \( L_2 \) norm of \((-\Delta)^{\alpha/2} \beta_l\) by splitting the spectral range of integration as
\[
\|(-\Delta)^{\alpha/2} \beta_l\|^2_{L_2} = \int_{\mathbb{R}^d} \|\omega\|^{2\alpha} |\hat{\beta}_l(\omega)|^2 \frac{d\omega}{(2\pi)^d}
= \int_{|\omega|<R} \|\omega\|^{2\alpha} |\hat{\beta}_l(\omega)|^2 \frac{d\omega}{(2\pi)^d} + \int_{|\omega|>R} \|\omega\|^{2\alpha} |\hat{\beta}_l(\omega)|^2 \frac{d\omega}{(2\pi)^d}
\leq \|\beta_l\|^2_{L_1} \int_{|\omega|<R} \|\omega\|^{2\alpha} \frac{d\omega}{(2\pi)^d} + C^2 \|d_l\|^2_{\ell_1} \int_{|\omega|>R} \|\omega\|^{2\alpha-2r} \frac{d\omega}{(2\pi)^d}.
\]

The first integral \( I_1 \) is finite due the boundedness of the domain. As for \( I_2 \), it is convergent provided that the rate of decay of the argument is faster than \( d \), which corresponds to the critical Sobolev exponent \( \alpha = \gamma - d/2 \).

As final step of the analysis, we invoke the Sobolev embedding theorems to infer that \( \beta_l \) is Hölder-continuous of order \( r \) with \( r < \alpha - \frac{d}{2} = (\gamma - d) \), which essentially means that \( \hat{\beta}_l \) is differentiable up to order \( r \) with bounded derivatives. One should keep in mind, however, that the latter estimate is a lower bound on Hölder continuity, unlike the Sobolev exponent in Proposition 6.4, which is sharp. For instance, in the case of the 1-D Fourier multiplier \((\omega \gamma)^r\), we find that the corresponding (fractional) B-spline—if it exists—should have a Sobolev smoothness \((\gamma - \frac{1}{2})\), and a Hölder regularity \( r < (\gamma - 1) \). Note that the latter is arbitrarily close (but not equal) to the true estimate \( r_0 = (\gamma - 1) \) that is readily deduced from the Green’s function (6.15).

### 6.4.2 B-spline factorization

A powerful aspect of spline theory is that it is often possible to exploit the factorization properties of differential operators to recursively generate whole families of B-splines. Specifically, if the operator can be decomposed as \( L = L_1 L_2 \), where the B-splines associated to \( L_1 \) and \( L_2 \) are already known, then \( \beta_l = \beta_{l_1} * \beta_{l_2} \) is the natural choice of B-spline for \( L \).

**Proposition 6.5** Let \( \beta_{l_1}, \beta_{l_2} \) be admissible B-splines for the operators \( L_1 \) and \( L_2 \), respectively. Then, \( \beta_l(r) = \beta_{l_1} * \beta_{l_2}(r) \) is an admissible B-spline for \( L = L_1 L_2 \) if and only if there exists a constant \( A > 0 \) such that
\[
\sum_{n \in \mathbb{Z}^d} \left| \hat{\beta}_{l_1}(\omega + 2\pi n) \hat{\beta}_{l_2}(\omega + 2\pi n) \right|^2 \geq A > 0.
\]
for all $\omega \in [0,2\pi]^d$. When $L_1 = L_2^\gamma$ for $\gamma \geq 0$, then the auxiliary condition is automatically satisfied.

**Proof** Since $\beta_{L_1}, \beta_{L_2} \in L_1(\mathbb{R}^d)$, the same holds true for $\beta_L$ (by Young's inequality). From the Fourier-domain definition (6.25) of the B-splines, we have

$$\hat{\beta}_{L_1}(\omega) = \sum_{k \in \mathbb{Z}^d} d_{L_1}(k) e^{-i(k,\omega)} = \frac{\tilde{L}_{L_1}(\omega)}{\tilde{L}_1(\omega)},$$

which implies that

$$\beta_L = \mathcal{F}^{-1} \left\{ \frac{\tilde{L}_{L_1}(\omega)}{\tilde{L}_1(\omega)} \right\} = \mathcal{F}^{-1} \left\{ \frac{\tilde{L}_2(\omega)}{\tilde{L}_1(\omega)} \right\} = L^{-1}\delta$$

with $L^{-1} = L_2^{-1} L_1^{-1}$ and $L_4 = L_{d,1} L_{d,2}$. This translates into the combined localization operator $L_4 s(r) = \sum_{k \in \mathbb{Z}^d} d_k |s(r-k)|$ with $d_k = (d_1 * d_2)(k)$, which is factorizable by construction. To establish the existence of the upper Riesz bound for $\beta_L$, we perform the manipulation

$$A_{\beta_L}(\omega) = \sum_{n \in \mathbb{Z}^d} |\hat{\beta}_L(\omega + 2\pi n)|^2$$

$$= \sum_{n \in \mathbb{Z}^d} |\hat{\beta}_{L_1}(\omega + 2\pi n)|^2 |\hat{\beta}_{L_2}(\omega + 2\pi n)|^2$$

$$\leq \left( \sum_{n \in \mathbb{Z}^d} |\hat{\beta}_{L_1}(\omega + 2\pi n)| |\hat{\beta}_{L_2}(\omega + 2\pi n)| \right)^2$$

$$\leq \sum_{n \in \mathbb{Z}^d} |\hat{\beta}_{L_1}(\omega + 2\pi n)|^2 \sum_{n \in \mathbb{Z}^d} |\hat{\beta}_{L_2}(\omega + 2\pi n)|^2$$

$$\leq B_1^2 B_2^2 < +\infty$$

where the third line follows from the norm inequality $\|a\|_{L_2} \leq \|a\|_{L_1}$ and the fourth from Cauchy-Schwarz; $B_1$ and $B_2$ are the upper Riesz bounds of $\beta_{L_1}$ and $\beta_{L_2}$, respectively. The additional condition in the proposition takes care of the lower Riesz bound. \qed

### 6.4.3 Polynomial B-splines

The factorization property is directly applicable to the construction of the polynomial B-splines (we use the equivalent notation $\beta^n_D = \beta^n_{D^n}$ in Section 1.3.2) via the iterated convolution of a B-spline of degree 0. Specifically,

$$\beta_{D^{n+1}}(r) = (\beta_D * \beta_{D^n})(r) = (\beta_D * \cdots * \beta_D)(r)$$

with $\beta_D = \beta_0^0 = 1_{[0,1]}$ and the convention that $\beta_{D^0} = \beta_D = \delta$. 


6.4 Generalized B-spline basis

6.4.4 Exponential B-splines

More generally, one can consider a generic \( N \)-th-order differential operator of the form \( P_\alpha = P_{\alpha_1} \cdots P_{\alpha_N} \) with parameter vector \( \alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{C}^N \) and \( P_{\alpha_n} = D - \alpha_n \text{Id} \). The corresponding basis function is an exponential B-spline of order \( N \) with parameter vector \( \alpha \), which can be decomposed as

\[
\beta_\alpha(r) = (\beta_{\alpha_1} \ast \beta_{\alpha_2} \ast \cdots \ast \beta_{\alpha_N})(r)
\]

where \( \beta_\alpha = \beta_{\rho_\alpha} \) is the first-order exponential spline defined by (6.21). The Fourier-domain counterpart of (6.42) is

\[
\hat{\beta}_\alpha(\omega) = \prod_{n=1}^{N} \frac{1 - e^{i\alpha_n \omega}}{i\omega - \alpha_n},
\]

which also yields

\[
\beta_\alpha(r) = \Delta_\alpha \rho_\alpha(r),
\]

where \( \Delta_\alpha = \Delta_{\alpha_1} \cdots \Delta_{\alpha_N} \) (with \( \Delta_\alpha \) defined by (6.20)) is the corresponding \( N \)-th-order localization operator (weighted differences) and \( \rho_\alpha \) the causal Green's function of \( P_\alpha \). Note that the complex parameters \( \alpha_n \), which are the roots of the characteristic polynomial of \( P_\alpha \), are the poles of the exponential B-spline, as seen in (6.43). The actual recipe for localization is that each pole is cancelled by a corresponding (2\( \pi \)-periodic) zero in the numerator.

Based on the above equations, one can infer the following properties of the exponential B-splines (see [UB05a] for a complete treatment of the topic):

- They are causal, bounded, and compactly supported in \([0, N]\), simply because all elementary constituents in (6.42) are bounded and supported in \([0, 1]\).
- They are piecewise-exponential with joining points at the integers and a maximal degree of smoothness (spline property). The first part follows from (6.44) using the well-known property that the causal Green's function of an \( N \)-th-order ordinary differential operator is an exponential polynomial restricted to the positive axis. As for the statement about smoothness, the B-splines are Hölder-continuous of order \((N - 1)\). In other words, they are differentiable up to order \((N - 1)\) with bounded derivatives. This follows from the fact that \( D\hat{\beta}_{\alpha_n}(r) = \delta(r) - e^{i\alpha_n \delta(r - 1)} \), which implies that every additional elementary convolution factor in (6.42) improves the differentiability of the resulting B-spline by one.
- They are the shortest elementary constituents of exponential splines (maximally localized kernels) and they each generate a valid Riesz basis (by integer shifting) of the spaces of cardinal \( P_\alpha \)-splines if and only if \( \alpha_n - \alpha_m \neq j2\pi k, k \in \mathbb{Z} \), for all distinct, purely imaginary poles.
- They reproduce the exponential polynomials that are in the null space of the operator \( P_\alpha \), as well as any of its Green's functions \( \rho_\alpha \), which all happen to be special types of \( P_\alpha \)-splines (with a minimum number of singularities).
- For \( \alpha = (0, \ldots, 0) \), one recovers Schoenberg's classical polynomial B-splines of
The system-theoretic interpretation is that the classical polynomial spline of degree \( n \) has a pole of multiplicity \( n + 1 \) at the origin: It corresponds to an (unstable) linear system that is an \( (n + 1) \)-fold integrator.

There is also a corresponding B-spline calculus whose main operations are

- Convolution by concatenation of parameter vectors:
  \[
  (\beta_{\alpha_1} * \beta_{\alpha_2})(r) = \beta_{(\alpha_1, \alpha_2)}(r).
  \]
- Mirroring by sign change
  \[
  \beta_{\alpha}(-r) = \left( \prod_{n=1}^{N} e^{\alpha_n} \right) \beta_{-\alpha}(r + N).
  \]
- Complex-conjugation
  \[
  \overline{\beta_{\alpha}(r)} = \beta_{\overline{\alpha}}(r).
  \]
- Modulation by parameter shifting
  \[
  e^{j\alpha j} \beta_{\alpha}(r) = \beta_{\alpha + j\alpha}(r)
  \]

with the convention that \( j = (j, \ldots, j) \).

Finally, we like to point out that exponential B-splines can be computed explicitly on a case-by-case basis using the mathematical software described in [Uns05, Appendix A].

### 6.4.5 Fractional B-splines

The fractional splines are an extension of the polynomial splines for all non-integer degrees \( \alpha > -1 \). The most notable members of this family are the causal fractional B-splines \( \beta_{\alpha}^n \), whose basic constituents are piecewise-power functions of degree \( \alpha \) [UB00]. These functions are associated with the causal fractional derivative operator \( D^{\alpha + 1} \) whose Fourier-based definition is

\[
D^{\alpha} \phi(r) = \int_{\mathbb{R}} (j \omega)^{\alpha} \hat{\phi}(\omega) e^{j\omega r} \frac{d\omega}{2\pi}
\]

in the sense of generalized functions. The causal Green's function of \( D^{\alpha} \) is the one-sided power function of degree \( \gamma - 1 \) specified by (6.15). One constructs the corresponding B-splines through a localization process similar to the classical one, replacing finite differences by the fractional differences defined as

\[
\Delta_{\gamma}^{\alpha} \phi(r) = \int_{\mathbb{R}} (1 - e^{-j\gamma})^{\alpha} \hat{\phi}(\omega) e^{j\omega r} \frac{d\omega}{2\pi}.
\]
In that respect, it is important to note that \((1 - e^{-j\omega})^\tau = (j\omega)^\tau + O(|\omega|^{2\tau})\) which justifies this particular choice. By applying (6.25), we readily obtain the Fourier-domain representation of the fractional B-splines

\[
\hat{\beta}_\alpha^\tau(\omega) = \left(\frac{1 - e^{-j\omega}}{j\omega}\right)^{\alpha + 1}
\]

which can then be inverted to provide the explicit time-domain formula

\[
\beta_\alpha^\tau(r) = \Delta_{\alpha + 1} \frac{r^\alpha}{\Gamma(\alpha + 1)}
\]

\[
= \sum_{m=0}^{\infty} (-1)^m \binom{\alpha + 1}{m} \frac{(r - m)^\alpha}{\Gamma(\alpha + 1)},
\]

where the generalized fractional binomial coefficients are given by

\[
\binom{\alpha + 1}{m} = \frac{\Gamma(\alpha + 2)}{\Gamma(m + 1)\Gamma(\alpha + 2 - m)} = \frac{(\alpha + 1)!}{m!(\alpha + 1 - m)!}.
\]

What is remarkable with this construction is the way in which the classical B-spline formulas of Section 1.3.2 carry over to the fractional case almost literally by merely replacing \(n\) by \(\alpha\). This is especially striking when we compare (6.47) to (1.11), as well as the expanded versions of these formulas given below, which follow from the (generalized) binomial expansion of \((1 - e^{-j\omega})^{\alpha + 1}\).

Likewise, it is possible to construct the \((\alpha, \tau)\) extension of these B-splines. They are associated with the operators \(L = \partial_t^{\alpha + 1} \longrightarrow (j\omega)^{\alpha + 1/\tau} (-j\omega)^{\alpha + 1/\tau - \tau}\) and \(\tau \in \mathbb{R}\) [BU03]. This family covers the entire class of translation- and scale-invariant operators in 1-D (see Proposition 5.6).

The fractional B-splines share virtually all the properties of the classical B-splines, including the two-scale relation, and can also be used to define fractional wavelet bases with an order \(\gamma = \alpha + 1\) that varies continuously. They only lack positivity and compact support. Their most notable properties are summarized below.

- **Generalization:** For \(\alpha\) integer, they are equivalent to the classical polynomial splines. The fractional B-splines interpolate the polynomial ones in very much the same way as the gamma function interpolates the factorials.
- **Stability:** All brands of fractional B-splines satisfy the Riesz-basis condition in Theorem 6.2.
- **Regularity:** The fractional splines are \(\alpha\)-Hölder continuous; their critical Sobolev exponent (degree of differentiability in the \(L^2\) sense) is \(\alpha + 1/2\) (see Proposition 6.4).
- **Polynomial reproduction:** The fractional B-splines reproduce the polynomials of degree \(N = [\alpha + 1]\) that are in the null space of the operator \(D^{\alpha + 1}\) (see Section 6.2.1).
- **Decay:** The fractional B-splines decay at least like \(|r|^{-\alpha - 2}\); the causal ones are compactly supported for \(\alpha\) integer.
- **Order of approximation:** The fractional splines have the non-integer order of approximation \(\alpha + 1\), a property that is rather unusual in approximation theory.
Fractional derivatives: Simple formulas are available for obtaining the fractional derivatives of B-splines. In addition, the corresponding fractional spline wavelets essentially behave like fractional-derivative operators.

6.4.6 Additional brands of univariate B-splines

To be complete, we briefly mention some additional types of univariate B-splines that have been investigated systematically in the literature.

- The generalized exponential B-splines of order $N$ that cover the whole class of differential operators with rational transfer functions [Uns05]. These are parameterized by their poles and zeros. Their properties are very similar to those of the exponential B-splines of the previous section, which are included as a special case.

- The Matérn splines of (fractional) order $\gamma$ and parameter $\alpha \in \mathbb{R}^+$ with $L = (D + \alpha 1d)^\gamma \rightarrow (j\omega + \alpha)^\gamma$ [RU06]. These constitute the fractionalization of the exponential B-spline with a single pole of multiplicity $N$.

In principle, it is possible to construct even broader families via the convolution of existing components. The difficulty is that it may not always be possible to obtain explicit signal-domain formulas, especially when some of the constituents are fractional.

6.4.7 Multidimensional B-splines

While the construction of B-splines is well understood and covered systematically in 1-D, the task becomes more challenging in multiple dimensions because of the inherent difficulty of imposing compact support. Apart from the easy cases where the operator $L$ can be decomposed in a succession of 1-D operators (tensor-product B-splines and box splines), the available collection of multidimensional B-splines is much more restricted than in the univariate case. The construction of B-splines is still considered an art where the ultimate goal is to produce the most localized basis functions. The primary families of multidimensional B-splines that have been investigated so far are

- The polyharmonic B-splines of (fractional) order $\gamma$ with $L = (-\Delta)^\gamma \rightarrow \|\omega\|^\gamma$ [MN90b,Rab92a,Rab92b,VDVBU05].

- The box splines of multiplicity $N \geq d$ with $L = D_{u_1} \cdots D_{u_N} \rightarrow \prod_{n=1}^{N} (j\omega, u_n)$ with $\|u_n\| = 1$, where $D_{u_n} = (\nabla, u_n)$ is the directional derivative along $u_n$ [dBHR93]. The box splines are compactly supported functions in $L_1(\mathbb{R}^d)$ if and only if the set of orientation vectors $\{u_n\}_{n=1}^{N}$ forms a frame of $\mathbb{R}^d$.

We encourage the reader who finds the present list incomplete to work on expanding it. The good news for the present study is that the polyharmonic B-splines are particularly relevant for image-processing applications because they are associated with the class of operators that are scale- and rotation-invariant. They naturally come into play when considering isotropic fractal-type random fields.
The principal message of this section is that B-splines—no matter the type—are localized functions with an equivalent width that increases in proportion to the order. In general, the fractional brands and the non-separable multidimensional ones are not compactly supported. The important issue of localization and decay is not yet fully resolved in higher dimensions. Also, since $L_d s = \beta L_s$, it is clear that the search for a “good” B-spline $\beta L$ is intrinsically related to the problem of finding an accurate numerical approximation $L_d$ of the differential operator $L$. Looking at the discretization issue from the B-spline perspective leads to new insights and sometimes to nonconventional solutions. For instance, in the case of the Laplacian $L = \Delta$, the continuous-domain localization requirement points to the choice of the 2D discrete operator $\Delta_d$ described by the $3 \times 3$ filter mask

$$
\text{Isotropic discrete Laplacian: } \frac{1}{6} \begin{pmatrix}
-1 & -4 & -1 \\
-4 & 20 & -4 \\
-1 & -4 & -1
\end{pmatrix}
$$

which is not the standard version used in numerical analysis. This particular set of weights produces a much nicer, bell-shaped polyharmonic B-spline than the conventional finite-difference mask which induces significant directional artifacts, especially when one starts iterating the operator [VDVBU05].

### 6.5 Generalized operator-like wavelets

In direct analogy with the first-order scenario in Section 6.3.3, we shall now take advantage of the general B-spline formalism to construct a wavelet basis that is matched to some generic operator $L$.

#### 6.5.1 Multiresolution analysis of $L_2(\mathbb{R}^d)$

The first step is to lay out a fine-to-coarse sequence of (multidimensional) $L$-spline spaces in essentially the same way as in our first-order example. Specifically,

$$
V_i = \{s(r) \in L_2(\mathbb{R}^d) : Ls(r) = \sum_{k \in \mathbb{Z}^d} a_i[k] \delta(r - D^i k)\}
$$

where $D$ is a proper dilation matrix with integer entries (e.g., $D = 2I$ in the standard dyadic configuration). These spline spaces satisfy the general embedding relation $V_i \supseteq V_j$ for $i \leq j$.

The reference space ($i = 0$) is the space of cardinal $L$ splines which admits the standard B-spline representation

$$
V_0 = \{s(r) = \sum_{k \in \mathbb{Z}^d} c[k] \beta L(r - k) : c \in \ell_2(\mathbb{Z}^d)\},
$$

where $\beta L$ is given by (6.25). Our implicit assumption is that each $V_i$ admits a similar B-spline representation

$$
V_i = \{s(r) = \sum_{k \in \mathbb{Z}^d} c_i[k] \beta_{L,i}(r - D^i k) : c_i \in \ell_2(\mathbb{Z}^d)\},
$$
which involves the multiresolution generators $\beta_{L,i}$ described in the next section.

### 6.5.2 Multiresolution B-splines and the two-scale relation

In direct analogy with (6.25), the multiresolution B-splines $\beta_{L,i}$ are localized versions of the Green's function $\rho_{L}$ with respect to the grid $D^i Z^d$. Specifically, we have that

$$
\beta_{L,i}(r) = \sum_{k \in \mathbb{Z}^d} d_i[k] \rho_{L}(r - D^i k) = L_d|L^{-1}\delta(r),
$$

where $L_d$ is the discretized version of $L$ on the grid $D^i Z^d$. The Fourier-domain counterpart of this equation is

$$
\hat{\beta}_{L,i}(\omega) = \sum_{k \in \mathbb{Z}^d} d_i[k] e^{-j\omega \cdot D^i k} \frac{L(\omega)}{L(\omega)}.
$$

(6.48)

The implicit requirement for the multiresolution decomposition scheme to work is that $\beta_{L,i}$ generates a Riesz basis. This needs to be asserted on a case-by-case basis.

A particularly favorable situation occurs when the operator $L$ is scale-invariant with $L(\omega) = |a|^j L(\omega)$. Let $i' > i$ be two multiresolution levels of the pyramid such that $D^i = mD^{i'}$ where $m$ is a proportionality constant. It is then possible to relate the B-spline at resolution $i'$ to the one at the finer level $i$ via the simple dilation relation

$$
\beta_{L,i'}(r) \propto \beta_{L,i}(r / m).
$$

This is shown by considering the Fourier transform of $\beta_{L,i}(r / m)$, which is written as

$$
|m|d \hat{\beta}_{L,i}(m\omega) = |m|d \sum_{k \in \mathbb{Z}^d} d_i[k] e^{-j\omega \cdot mD^i k} \frac{L(\omega)}{L(\omega)}
$$

$$
= |m|d \sum_{k \in \mathbb{Z}^d} d_i[k] e^{-j\omega \cdot D^{i'} k} \frac{L(\omega)}{L(\omega)}
$$

$$
= |m|d \sum_{k \in \mathbb{Z}^d} d_i[k] e^{-j\omega \cdot D' k} \frac{L(\omega)}{L(\omega)},
$$

and found to be compatible with the form of $\hat{\beta}_{L,i'}(\omega)$ given by (6.48) by taking $d_i[k] \propto d_i[k]$. The prototypical scenario is the dyadic configuration $D = 2I$ for which the B-splines at level $i$ are all constructed through the dilation of the single prototype $\beta_L \propto \beta_{L,0}$, subject to the scale-invariance constraint on $L$. This happens, for instance, for the classical polynomial splines which are associated with the Fourier multipliers $(j\omega)^N$.

A crucial ingredient for the fast wavelet-transform algorithm is the two-scale relation that links the B-splines basis functions at two successive levels of resolution. Specifically, we have that

$$
\beta_{L,i+1}(r) = \sum_{k \in \mathbb{Z}^d} h_i[k] \beta_{L,i}(r - D^i k),
$$

where the sequence $h_i$ specifies the scale-dependent refinement filter. The frequency
6.5 Generalized operator-like wavelets

The response of $h_i$ is obtained by taking the ratios of the Fourier transforms of the corresponding B-splines as

$$
\hat{h}_i(\omega) = \frac{\hat{\beta}_{L,i+1}(\omega)}{\hat{\beta}_{L,i}(\omega)}
$$

which is $2\pi(D^T)^{-i}$ periodic and hence defines a valid digital filter with respect to the spatial grid $D^iZ^d$.

To illustrate those relations, we return to our introductory example in Section 6.1: the Haar wavelet transform, which is associated with the Fourier multipliers $j\omega$ (derivative) and $(1 - e^{-j\omega})$ (finite-difference operator). The dilation matrix is $D = 2$ and the localization filter is the same at all levels because the underlying derivative operator is scale-invariant. By plugging those entities into (6.48), we obtain the Fourier transform of the corresponding B-spline at resolution $i$ as

$$
\hat{\beta}_{D,i}(\omega) = 2^{-i/2} \frac{1 - e^{2i\omega}}{j\omega},
$$

where the normalization by $2^{-i/2}$ is included to standardize the norm of the B-splines. The application of (6.49) then yields

$$
\hat{h}_i(\omega) = \frac{1}{\sqrt{2}} \frac{1 - e^{2i\omega}}{1 - e^{2i\omega}} = \frac{1}{\sqrt{2}} (1 + e^{2i\omega}),
$$

which, up to the normalization by $\sqrt{2}$, is the expected refinement filter with coefficients proportional to $(1,1)$ that are independent upon the scale.

6.5.3 Construction of an operator-like wavelet basis

To keep the notation simple, we concentrate on the specification of the wavelet basis at the scale $i = 1$ with $W_1 = \text{span}\{\psi_{1,k}\}_{k \in Z^d}$ such that $W_1 \perp V_1$ and $V_0 = V_1 + W_1$, where $V_0 = \text{span}\{\beta_{L}(-k)\}_{k \in Z^d}$ is the space of cardinal L-splines.

The relevant smoothing kernel is the interpolation function $\varphi_{\text{int}} = \varphi_{\text{int},0}$ for the space of cardinal $L^d$-L-splines, which is generated by $(\hat{\beta}_{L}^T \ast \hat{\beta}_{L})(r)$ (autocorrelation of the generalized B-spline). This interpolator is best described in the Fourier domain using the formula

$$
\varphi_{\text{int}}(r) = \mathcal{F}^{-1} \left\{ \frac{|\hat{\beta}_{L}(\omega)|^2}{\sum_{n \in Z^d} |\hat{\beta}_{L}(\omega + 2\pi n)|^2} \right\}(r),
$$

where $\hat{\beta}_{L}$ (resp., $\overline{\hat{\beta}_{L}}$) is the Fourier transform of the generalized B-spline $\beta_{L}$ (resp.,
\( \tilde{\beta}_L \). It satisfies the fundamental interpolation property
\[
\varphi_{\text{int}}(k) = \delta[k] = \begin{cases} 
1, & \text{for } k = 0 \\
0, & \text{for } k \in \mathbb{Z}^d \setminus \{0\}.
\end{cases}
\] (6.52)

The existence of such a function is guaranteed whenever \( \beta_L = \beta_{L,0} \) is an admissible B-spline. In particular, the Riesz basis condition (6.26) implies that the denominator of \( \tilde{\varphi}_{\text{int}}(\omega) \) in (6.51) is non-vanishing.

The sought-after wavelets are then constructed as \( \psi_{1,k}(r) = \psi_L(r - k)/\|\psi_L\|_{L^2} \), where the operator-like mother wavelet \( \psi_L \) is given by
\[
\psi_L(r) = L^H \varphi_{\text{int}}(r),
\] (6.53)

where \( L^H \) is the adjoint of \( L \) with respect to the Hermitian-symmetric \( L_2 \) inner product. Also, note that we are removing the functions located on the next coarser resolution grid \( DZ^d \) associated with \( V_1 \) (critically sampled configuration).

The proof of the following result is illuminating because it relies heavily on the notion of duality which is central to our whole argumentation.

**Proposition 6.6** The operator-like wavelet \( \psi_L = L^H \varphi_{\text{int}} \) satisfies the property \( \langle s_1, \psi_{1,(-k)} \rangle_{L^2} = \langle s_1, \psi_L(-k) \rangle = 0, \forall k \in \mathbb{Z}^d \setminus DZ^d \) for any spline \( s_1 \in V_1 \). Moreover, it can be written as \( \psi_L(r) = L^H \beta_L(r) = \sum_{k \in \mathbb{Z}^d} \delta_L[-k] \tilde{\beta}_L[r - k] \) where \( \delta_L[-k] \) is the dual basis of \( V_0 \) such that \( \langle \delta_L[-k], \beta_L[-k'] \rangle_{L^2} = \delta[k - k'] \). This implies that \( W_1 = \text{span}\{\psi_{1,k}\}_{k \in \mathbb{Z}^d \setminus DZ^d} \subseteq V_0 \) and \( W_1 \perp V_1 \).

**Proof** We pick an arbitrary spline \( s_1 \in V_1 \) and perform the inner-product manipulation
\[
\langle s_1, \psi_{1,(-k)} \rangle_{L^2} = \langle s_1, L^* \varphi_{\text{int}}(-k) \rangle
\]
(by shift-invariance of \( L \))
\[
= (Ls_1, \varphi_{\text{int}}(-k))
\]
(by duality)
\[
= \sum_{k \in \mathbb{Z}^d} a_1[k] \delta[-Dk] \varphi_{\text{int}}(-k)
\]
(by definition of \( V_1 \))
\[
= \sum_{k \in \mathbb{Z}^d} a_1[k] \varphi_{\text{int}}(Dk - k_0).
\]
(by definition of \( \delta \))

Due to the interpolation property of \( \varphi_{\text{int}} \), the kernel values in the sum are vanishing if \( Dk - k_0 \in \mathbb{Z}^d \setminus \{0\} \) for all \( k \in \mathbb{Z}^d \), which proves the first part of the statement.

As for the second claim, we consider the Fourier-domain expression of \( \psi_L \):
\[
\hat{\psi}_L(\omega) = \hat{L}(\omega) \varphi_{\text{int}}(\omega) = \hat{L}(\omega) \hat{\beta}_L(\omega) \hat{\beta}_L(\omega)
\]
where
\[
\hat{\beta}_L(\omega) = \frac{\hat{\beta}_L(\omega)}{\sum_{n \in \mathbb{Z}^d} |\hat{\beta}_L(\omega + 2\pi n)|^2}
\]
is the Fourier transform of the dual B-spline \( \hat{\beta}_L \). The above factorization implies that \( \varphi_{\text{int}}(r) = (\hat{\beta}_L \ast \hat{\beta}_L)(r) \), which ensures the biorthonormality \( \langle \beta_L[-k], \beta_L[-k'] \rangle_{L^2} = \delta[k] = \varphi_{\text{int}}(k - k') \) of the basis functions.
6.5 Generalized operator-like wavelets

Finally, by replacing \( \hat{\beta}_L(\omega) \) by its explicit expression (6.25), we show that \( \overline{L_d(\omega)} \hat{\beta}_L(\omega) = L_d(\omega) \), where \( \hat{L}_d(\omega) = \sum_{k \in \mathbb{Z}^d} d_t[k] e^{-i(k,\omega)} \) is the frequency response of the (discrete) operator \( L_d \). This implies that \( \hat{\psi}_L = \overline{L_d} \hat{\beta}_L \), which is the Fourier equivalent of \( \psi_L = L_d^H \hat{\beta}_L \).

This interpolation-based method of construction is applicable to all the wavelet subspaces \( W_i \) and leads to the specification of operator-like Riesz bases of \( L_2(\mathbb{R}^d) \) under relatively mild assumptions on \( L \) [KUW13]. Specifically, we have that \( W_i = \text{span}\{\psi_i, k\}_{k \in \mathbb{Z}^d} \) with

\[
\psi_{i,k}(r) \propto L^H_{\text{int},i-1}(r - D^{i-1} k),
\]

(6.54)

where \( \varphi_{\text{int},i-1} \) is the \( L^H L \)-spline interpolator on the grid \( D^{i-1} \mathbb{Z}^d \). The fact that the interpolator is specified with respect to the grid of the next finer spline space \( V_{i-1} = \text{span}(\hat{\beta}_{L,i-1}(-D^{i-1} k))_{k \in \mathbb{Z}^d} \) is essential to ensure that \( W_i \subset V_{i-1} \). This kernel satisfies the fundamental interpolation property

\[
\varphi_{\text{int},i-1}(D^{i-1} k) = \delta[k],
\]

(6.55)

which results in \( W_i \) being orthogonal to \( V_i = \text{span}(\hat{\beta}_{L,i}(-D^i k))_{k \in \mathbb{Z}^d} \) (the reasoning is the same as in the proof of Proposition 6.6 which covers the case \( i = 1 \)). For completeness, we also provide the general expression of the Fourier transform of \( \varphi_{\text{int},i} \),

\[
\hat{\varphi}_{\text{int},i}(\omega) = |\text{det}(D)|^i \sum_{n \in \mathbb{Z}^d} |\hat{\beta}_{L,i}(\omega + 2\pi(D^T)^{-1} n)|^2 / |\hat{L}_{d,i}(\omega)|^2 / \sum_{n \in \mathbb{Z}^d} |\hat{L}(\omega + 2\pi(D^T)^{-1} n)|^2,
\]

(6.56)

which can be used to show that \( L^H_{\text{int},i}(-D^i k) \propto L^H_{d,i} \hat{\beta}_{L,i}(-D^i k) \in V_i \) for any \( k \in \mathbb{Z}^d \).

While we have seen that this scheme produces an orthonormal basis for the first-order operator \( P_d \) in Section 6.3.3, the general procedure does only guarantee semi-orthogonality. More precisely, it ensures the orthogonality between the wavelet subspaces \( W_i \). If necessary, one can always fix the intra-scale orthogonality a posteriori by forming appropriate linear combinations of wavelets at a given resolution. The resulting orthogonal wavelets will still be \( L \)-admissible in the sense of Definition 6.7. However, for \( d > 1 \), intra-scale orthogonalization is likely to spoil the simple, convenient structure of the above construction which uses a single generator per scale, irrespective of the number of dimensions. Indeed, the examples of multidimensional orthogonal wavelet transforms that can be found in the literature—either separable, or non-separable—systematically involve \( M = (\text{det}(D) - 1) \) distinct wavelet generators per scale. Moreover, unlike the present operator-like wavelets, they do generally not admit an explicit analytical description.
In summary, wavelets generally behave like differential operators and it is possible to match them to a given class of stochastic processes. The wavelet transforms that are currently most widely used in applications act as multiscale derivatives or Laplacians. They are therefore best suited for the representation of fractal-type stochastic processes that are defined by scale-invariant SDEs [TVDVU09].

The general theme that emerges is that a signal transform will behave appropriately if it has the ability to suppress the signal components (polynomial or sinusoidal trends) that are in the null space of the whitening operator $L$. This will result in a stationarizing effect that is well-documented in the Gaussian context [Fla89, Fla92]. This is the fundamental reason why vanishing moments are so important.

### Bibliographical notes

#### Section 6.1

Alfréd Haar constructed the orthogonal Haar system as part of his Ph.D. thesis, which he defended in 1909 under the supervision of David Hilbert [Haa10]. From then on, the Haar system remained relatively unnoticed until it was revitalized by the discovery of wavelets nearly one century later. Stéphane Mallat set the foundation of the multiresolution theory of wavelets in [Mal89] with the help of Yves Meyer, while Ingrid Daubechies constructed the first orthogonal family of compactly supported wavelets [Dau88]. Many of the early constructions of wavelets are based on splines [Mal89, CW91, UAE92, UAE93]. The connection with splines is actually quite fundamental in the sense that all multiresolution wavelet bases, including the non-spline brands such as Daubechies’, necessarily include a B-spline as a convolution factor—the latter is responsible for their primary mathematical properties such as vanishing moments, differentiability, and order of approximation [UB03]. Further information on wavelets can be found in several textbooks [Dau92, Mey90, Mal09].

#### Section 6.2

Splines constitute a beautiful topic of investigation in their own right with hundreds of papers specifically devoted to them. The founding father of the field is Schoenberg who, during war time, was asked to develop a computational solution for constructing an analytic function that fits a given set of equidistant noisy data points [Sch88]. He came up with the concept of spline interpolation and proved that polynomial spline functions have a unique expansion in terms of B-splines [Sch46]. While splines can also be specified for nonuniform grids and extended in a variety of ways [dB78, Sch81a], the cardinal setting is especially pleasing because it lends itself to systematic treatment with the aid of the Fourier transform [Sch73a]. The relation between splines and differential operators was recognized early on and led to the generalization known as L-splines [SV67].

The classical reference on partial differential operators and Fourier multipliers is [Hör80]. A central result of the theory is the Malgrange-Ehrenpreiss theorem [Mal56,
Ehr54], and its extension stating that the convolution with a compactly supported
generalized function is invertible [Hör05].

The concept of a Riesz basis is standard in functional analysis and approximation
theory [Chr03]. The special case where the basis functions are integer translates of a
single generator is treated in [AU94]. See also [Uns00] for a review of such representa-
tions in the context of sampling theory.

Section 6.3
The first-order illustrative example is borrowed from [UB05a, Figure 1] for the con-
struction of the exponential B-spline, and from [KU06, Figure 1] for the wavelet part
of the story.

Section 6.4
The 1-D theory of cardinal L-splines for ordinary differential operators with con-
stant coefficients is due to Micchelli [Mic76]. In the present context, we are especially
concerned with ordinary differential equations, which go hand-in-hand with the ex-
tended family of cardinal exponential splines [Uns05]. The properties of the relevant
B-splines are investigated in full detail in [UB05a], which constitutes the ground ma-
terial for Section 6.4. A key property of B-splines is their ability to reproduce poly-
nomials. It is ensured by the Strang-Fix conditions (6.37) which play a central role
in approximation theory [dB87, SF71]. While there is no fundamental difficulty in
specifying cardinal-spline interpolators in multiple dimensions, it is much harder
to construct compactly supported B-splines, except for the special cases of the box
splines [dBH82, dBH93] and exponential box splines [Ron88]. For elliptic operators
such as the Laplacian, it is possible to specify exponentially decaying B-splines, with
the caveat that the construction is not unique [MN90b, Rab92a, Rab92b]. This calls for
some criterion to identify the most-localized solution [VDVBU05]. B-splines, albeit
non-compactly supported ones, can also be specified for fractional operators [UB07].
This line of research was initiated by Unser and Blu with the construction of the frac-
tional B-splines [UB00]. As suggested by the name, the (Gaussian) stochastic coun-
terparts of these B-splines are Mandelbrot’s fractional Brownian motions [MVN68],
as we shall see in Chapters 7 and 8. The association is essentially the same as the
connection between the B-spline of degree 0 (rect) and Brownian motion, or by ex-
tension, the whole family of Lévy processes (see Section 1.3).

Section 6.5
de Boor et al. were among the first to extend the notion of multiresolution ana-
lysis beyond the idea of dilation and to propose a general framework for construct-
ing “non-stationary” wavelets [dBR93]. Khalidov and Unser proposed a systematic
method for constructing wavelet-like basis functions based on exponential splines
and proved that these wavelets behave like differential operators [KU06]. The ma-
terial in Section 6.5 is an extension of those ideas to the case of a generic Fourier-
multiplier operator in multiple dimensions; the full technical details can be found
in [KUW13]. Operator-like wavelets have also been specified within the framework
of conventional multiresolution analysis; in particular, for the Laplacian and its it-
erates [VDVBU05, TVDVU09] and for the various brands of 1-D fractional derivat-
ives [VDVFHUB10], which have the common property of being scale-invariant. Fi-
nally, we mention that each exponential-spline wavelet has a compactly supported
Daubechies’ counterpart that is orthogonal and operator-like in the sense of having
the same vanishing exponential moments [VBU07].