Journal of Computational and Applied Mathematics 368 (2020) 112503

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Support and approximation properties of Hermite splines

Julien Fageot [a], Shayan Aziznejad [a], Michael Unser [a], Virginie Uhlmann [a,b,*]

[a] Biomedical Imaging Group, École polytechnique fédérale de Lausanne (EPFL), Station 17, 1015 Lausanne, Switzerland
[b] European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge CB10 1SD, UK

ARTICLE INFO

Article history:
Received 8 February 2019
Received in revised form 25 July 2019

Keywords:
Hermite interpolation
Minimum-support property
Approximation error

ABSTRACT

In this paper, we formally investigate two mathematical aspects of Hermite splines that are relevant to practical applications. We first demonstrate that Hermite splines are maximally localized, in the sense that the size of their support is minimal among pairs of functions with identical reproduction properties. Then, we precisely quantify the approximation power of Hermite splines for the reconstruction of functions and their derivatives. It is known that the Hermite and B-spline approximation schemes have the same approximation order. More precisely, their approximation error vanishes as $O(T^4)$ when the step size T goes to zero. In this work, we show that they actually have the same asymptotic approximation error constants, too. Therefore, they have identical asymptotic approximation properties. Hermite splines combine optimal localization and excellent approximation power, while retaining interpolation properties and closed-form expression, in contrast to existing similar functions. These findings shed a new light on the convenience of Hermite splines in the context of computer graphics and geometrical design.

© 2019 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In his seminal 1973 monograph on cardinal interpolation and spline functions [1], I.J. Schoenberg explains and characterizes B-spline interpolation, which still inspires researchers and yields exciting applications nowadays. In the same work, he also sets the basis of Hermite interpolation [2,3]. In the classical B-spline framework, a continuous-domain function is constructed from a discrete sequence of samples [4–6]. By contrast, the Hermite interpolation problem involves two sequences of discrete samples. They impose constraints not only on the resulting interpolated function but also on its derivatives up to a given order.

Curves in the plane or tensor-product surfaces in space can be constructed from one-dimensional interpolation schemes by interpolating along each spatial coordinate. The practical value of Hermite splines in this context is to offer tangential control on the interpolated curve. This can be easily understood through their link with Bézier curves [7]. The latter lie at the heart of vector graphics and are popular tools for computer-aided geometrical design and modeling [8–10]. Because of their small support, Hermite splines are also an interesting option for the design of multiwavelets, which are wavelets with multiple generators [11,12]. In practice, Hermite splines thus provide a suitable solution to a number of problems, whether with respect to simplicity of construction, efficiency, or convenience. This hands-on intuition can be translated to formal properties of Hermite splines and mathematically characterized. Wer give as examples the joint interpolation properties of Hermite splines (see Section 1.2) that ensure that, at integer values, the interpolated function exactly matches

* Corresponding author at: European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge CB10 1SD, UK.
E-mail addresses: julien.fageot@epfl.ch (J. Fageot), shayan.aziznejad@epfl.ch (S. Aziznejad), michael.unser@epfl.ch (M. Unser), uhlmann@ebi.ac.uk (V. Uhlmann).

https://doi.org/10.1016/j.cam.2019.112503
0377-0427/© 2019 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

the sequences of samples and derivative samples that were used to build it; their smoothness properties [13], which guarantee low curvature of the interpolated curve under some mild conditions; and their statistical optimality (in terms of MMSE) for the reconstruction of second-order Brownian motion from direct and first-derivative samples [14]. In that spirit, we investigate in this work the theoretical counterpart of two additional features that are observed to grant Hermite splines their practical usefulness.

1.1. Contributions

Our contributions state the minimal-support property of Hermite splines and investigate their approximation power. In the following, we describe the practical observations that motivate them, the results themselves, and related works.

Minimal-support property. The short support of Hermite splines is an important feature that makes them attractive in practice. The size of the support relates to the local extent of modifications on the continuously defined spline curve. In Section 2, we formally demonstrate that Hermite splines have the minimal support among all basis functions that generate cubic and quadratic splines. When dealing with B-splines, there is a tradeoff between the ability to reproduce smooth functions, which increases with the B-spline polynomial degree, and the possibility to allow for more or less sharp transitions, which decreases with the degree. As the question of function reproduction is a central concept when studying both minimal support and approximation errors, we provide a formal definition of it in Section 2.1. On one hand, cubic splines can efficiently reproduce smooth functions are able to capture C^2 transitions, but lack the power to capture C^1 transitions. On the other hand, quadratic splines have a lesser approximation power, but are preferred when dealing with less smooth (C^1) transitions. Hermite splines combine these two strengths in one scheme and are, in terms of support size, better than two-function schemes, including the one composed of the classical cubic and quadratic B-splines. In addition, we also show that one necessarily requires two generators to achieve this optimality. This result relates to similar ones involving a single generator [15,16].

Rate of decay of the approximation error. Hermite splines can provide faithful approximation reasonably fast as the number of parameters increases. This feature relates to the rate of decay of the error of approximation. Numerous works approach these questions by restricting themselves to a specific interpolation framework, such as [17,18] for the specific case of Hermite approximation. Relying on L_∞ norms, they provide precise estimations of the optimal bound on the approximation error, but do not allow for comparisons with other schemes. In contrast, approaches have been developed for single [19] and multiple generators [20] to provide a unifying comparison setting. They offer generalized measures that can be applied to a wide variety of basis functions to estimate approximation constants. Hermite interpolation, however, violates some of the core assumptions of those analyses. There, we take strong inspirations from those previous works and provide a novel study of Hermite approximation. We deploy an analogous analysis strategy by relieving the boundedness assumptions and considering additional spline approximation schemes. In Section 3, we precisely quantify the rate of decay of the error of approximation of the Hermite scheme and quantitatively estimate the corresponding approximation constants. Hermite-spline interpolation offers excellent approximation properties when it comes to the reconstruction of a function and its first derivative. It is actually close to achieving the minimal error obtained when the approximation procedure is modified to correspond to the orthogonal projector. The investigation of the error of approximation on the derivative relates to [21,22], although it follows a completely different line: while [21,22] focused on the reconstruction of the derivative from signal samples, the Hermite scheme grants direct access on the function and derivative samples, allowing one to reconstruct the derivative in a multifunction setting.

Pioneering works cover the study of spline schemes to an impressive degree of generality [23–26]. They include results on minimum support and approximation errors in a large variety of cases. Regarding minimum support, these previous results are however restricted to single-generator schemes. To the best of our knowledge, there are no previous works that would consider multi-generators and cover the Hermite case. These pioneering works also do not focus on providing a way to quantify the approximation-error constants, thus preventing one to compare schemes of the same order.

A better understanding of the approximation and support properties of Hermite splines has several useful consequences. For instance, wavelet schemes are commonly built from Hermite splines [27,28], so that the precise characterization of the approximation power of wavelet bases is important in practical applications such as image compression [29]. Hermite splines are also used to construct parametric deformable contours in the context of image segmentation [13,30], where their small support allows for the representation of open curves with natural conditions at their extremities [31].

1.2. Hermite splines

Schoenberg defines the cardinal cubic-Hermite-interpolation problem as follows [2,3]: knowing the discrete sequences of numbers $c[k]$ and $d[k]$, $k \in \mathbb{Z}$, we look for a continuously defined function $f_{\text{Her}}(t)$, $t \in \mathbb{R}$, that satisfies $f_{\text{Her}}(k) = c[k]$, $f'_{\text{Her}}(k) = d[k]$ for all $k \in \mathbb{Z}$, such that f_{Her} is piecewise polynomial of degree at most 3 and once differentiable with continuous derivative at the integers. The existence and uniqueness of the solution is guaranteed [2, Theorem 1] for any sequences $c = (c[k])$ and $d = (d[k])$ bounded by a polynomial, but we shall restrict to sequences in $\ell_2(\mathbb{Z})$ thereafter. In [3], it is shown that the Hermite spline f_{Her} associated to the sequences c and d can be expressed as

$$f_{\text{Her}}(t) = \sum_{k\in\mathbb{Z}} c[k]\phi_1(t-k) + d[k]\phi_2(t-k), \tag{1}$$

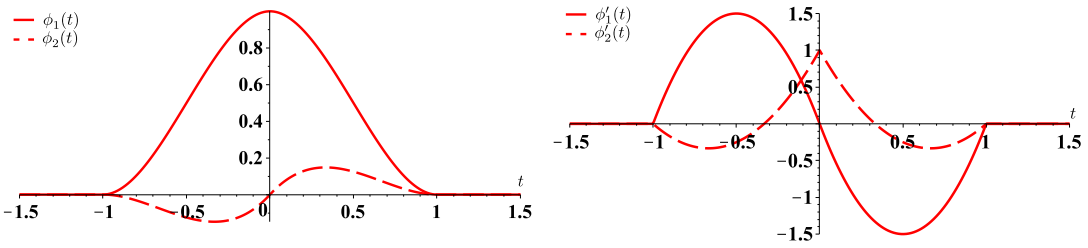

Fig. 1. Cubic Hermite splines ϕ_1 and ϕ_2. The two functions and their derivatives are vanishing at the integers, with the exception of $\phi_1(0) = 1$ and $\phi_2'(0) = 1$ (interpolation properties). They are supported in $[-1, 1]$.

where the functions ϕ_1 and ϕ_2 are given by

$$\phi_1(t) = (2|t| + 1)(|t| - 1)^2 \mathbb{1}_{0 \le |t| \le 1}, \tag{2}$$

$$\phi_2(t) = t(|t| - 1)^2 \mathbb{1}_{0 \le |t| \le 1}. \tag{3}$$

In addition to their fairly simple analytical expression, the cubic Hermite splines have other important properties. First, they are of finite support in $[-1, 1]$. Moreover, the generating functions ϕ_1, ϕ_2 and their derivatives ϕ_1', ϕ_2' satisfy the joint interpolation conditions

$$\phi_1(k) = \delta[k], \quad \phi_2'(k) = \delta[k], \quad \phi_1'(k) = 0, \quad \phi_2(k) = 0, \tag{4}$$

for all $k \in \mathbb{Z}$, where $\delta[k]$ is the discrete unit impulse. The functions and their first derivative are depicted in Fig. 1, where the interpolation properties can easily be observed. The functions ϕ_1 and ϕ_2 are deeply intertwined as $c[k] = f_{\text{Her}}(k)$ and $d[k] = f_{\text{Her}}'(k)$ in (1). The cubic Hermite splines are differentiable with continuous derivatives at the integer knots points $t = k$. As a result, functions generated by cubic Hermite splines are C^1-continuous piecewise-cubic polynomials with knots at integer locations.

2. Minimum-support property of Hermite splines

B-splines are known to be maximally localized, meaning that they are compactly supported with minimal-support property among functions with the same approximation properties [32,33]. Hermite splines possess a similar fundamental minimal-support property: they are of minimal support among the pair of functions that generate both quadratic and cubic splines (Theorem 3). In addition, there exists no single generating function that can take this role (Proposition 5). This demonstrates that Hermite splines are maximally localized for the purpose of representing piecewise-cubic functions that are continuously differentiable, as exploited, for instance, in image processing for the design of deformable parametric contours [13].

2.1. Integer-shift-invariant spaces and support properties

Consider a set of $N \ge 1$ functions $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_N) \in (L_2(\mathbb{R}))^N$ and define the space of functions

$$V(\boldsymbol{\varphi}) = \left\{ \sum_{i=1}^{N} \sum_{k \in \mathbb{Z}} c_i[k] \varphi_i(\cdot - k) : \mathbf{c} = (c_1, \dots, c_N) \in (\ell_2(\mathbb{Z}))^N \right\}. \tag{5}$$

We say that the φ_i are *basis functions* of the set $V(\boldsymbol{\varphi})$. The space $V(\boldsymbol{\varphi})$ is *integer-shift-invariant* in the sense that $f(\cdot - k)$ is in $V(\boldsymbol{\varphi})$ for $f \in V(\boldsymbol{\varphi})$ and $k \in \mathbb{Z}$ [23–25,34].

We consider integer-shift-invariant spaces generated by a single ($N = 1$) or two functions ($N = 2$), which corresponds to B-splines with simple knots and Hermite splines, respectively. We only consider spaces (5) for which the family $(\varphi_i(\cdot - k))_{i=1\dots N, k \in \mathbb{Z}}$ is a Riesz basis and satisfies

$$A \sum_{i=1}^{N} \sum_{k \in \mathbb{Z}} c_i[k]^2 \le \left\| \sum_{i=1}^{N} \sum_{k \in \mathbb{Z}} c_i[k] \varphi_i(\cdot - k) \right\|_{L_2}^2 \le B \sum_{i=1}^{N} \sum_{k \in \mathbb{Z}} c_i[k]^2 \tag{6}$$

for any $c_i \in \ell_2(\mathbb{Z})$ with $i = 1, \dots, N$ and for some positive and finite constants $0 < A \le B < \infty$. This ensures that any $f \in V(\boldsymbol{\varphi})$ has a unique and stable representation in the basis $(\varphi_i(\cdot - k))_{i=1\dots N, k \in \mathbb{Z}}$.

Any function $f \in V(\boldsymbol{\varphi})$ can be *reproduced* by the family $(\varphi_i(\cdot - k))_{i=1\dots N, k \in \mathbb{Z}}$ in the sense that there exist sequences $c_i \in \ell_2(\mathbb{Z})$ such that $f = \sum_{i,k} c_i[k] \varphi_i(\cdot - k)$. However, such functions are necessary in $L_2(\mathbb{R})$ according to (6). It is interesting to investigate the reproduction properties beyond $L_2(\mathbb{R})$; in particular, the ability to reproduce polynomials. For this purpose, one needs to extend the notion. We say that a function f can be *reproduced* by $(\varphi_i(\cdot - k))_{i=1\dots N, k \in \mathbb{Z}}$ if there

exists sequences c_i such that $f(t) = \sum_{i,k} c_i[k]\varphi_i(t-k)$ holds for all $t \in \mathbb{R}$. What is hidden here is that the sum should be well-defined at any time t. When the basis functions are compactly supported, the sum is actually finite for every fixed $t \in \mathbb{R}$ and any sequence c_i. In this section, we restrict ourselves to compactly supported basis functions since our goal is to characterize basis functions of minimal support. For non-compactly supported functions, the same notion is valid up to the condition that the sum is absolutely convergent for any $t \in \mathbb{R}$. The reproduction of polynomials will play a crucial role all along the paper. It is formalized in Definition 1.

Definition 1. A family of $N \geq 1$ basis functions $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_N)$ is *of order L* if they satisfy $\int_{\mathbb{R}} (1+|t|)^L |\varphi_n(t)| \mathrm{d}t < \infty$ for $1 \leq n \leq N$ and if they can reproduce polynomial functions up to degree $(L-1)$, meaning that, for $\ell = 0, \ldots, (L-1)$, there exist sequences $c_{i,\ell}$ such that

$$\forall t \in \mathbb{R}, \quad t^\ell = \sum_{i=1}^{N} \sum_{k \in \mathbb{Z}} c_{i,\ell}[k] \varphi_i(t - Nk). \tag{7}$$

It is worth noting that the sequence of coefficients is not required to take the values of the polynomial at the knots but, instead, that the interpolated function and the polynomial coincide. Condition (7) is known to be equivalent to the so-called Strang and Fix conditions and is very classical in approximation theory [35]. The hypothesis that $\int_{\mathbb{R}} (1+|t|)^L |\varphi_n(t)| \mathrm{d}t < \infty$ ensures that the Fourier transform of the φ_n admits continuous and bounded derivatives up to order L, which is actually crucial in the Stang and Fix formulation. This technical condition appears and is discussed in [20, Section II-D]. Note that it is clearly satisfied as soon as the functions are compactly supported and locally integrable, as will be the case in the rest of this section. In Section 3, we shall consider basis functions that are possibly non-compactly supported.

Support of B-splines. Under the Riesz condition, a natural question is the ability of the basis functions $\boldsymbol{\varphi}$ to exactly reproduce classes of functions. The possibility to perfectly reproduce polynomials is of crucial importance. The constant function 1 can be reproduced if and only if the basis functions satisfy the partition of unity, which is the minimal requirement for a practical approximation scheme [36].

The polynomial B-spline of order $L \geq 1$ is classically known to be able to reproduce polynomials t^ℓ for every $\ell = 0, \ldots, (L-1)$. Following [37], we denote it by $\beta^{(L-1)}$. In particular, the cubic B-spline (of order $L = 4$) can perfectly reproduce any polynomial of degree at most 3. The ability of the basis functions $\boldsymbol{\varphi}$ to perfectly reproduce polynomials is intimately linked to their approximation power, as will be developed in Section 3. B-splines are actually the most localized functions that satisfy this property, as formalized in Proposition 2.

Proposition 2. *Let $\varphi \in L_2(\mathbb{R})$ be a compactly supported function such that $(\varphi(\cdot - k))_{k \in \mathbb{Z}}$ is a Riesz basis and can reproduce polynomials up to degree $(L-1) \geq 0$. Then, the support of φ is at least of size L.*

In particular, the B-spline of order L, whose support is of size L, is optimal in terms of support localization among basis functions that are able to reproduce polynomials of degree at most $(L-1)$.

This result is classical in approximation theory: Schoenberg showed that B-splines effectively have the adequate approximation order [1]. A complete characterization of the piecewise-polynomial functions of minimal support with a given approximation order can be found in [16, Theorem 1]. To the best of our knowledge, very little is known about the localization of basis functions when $N > 1$, which is what we propose to investigate. In the multifunction scheme, we shall characterize reproduction properties by considering the reproduction of B-splines instead of polynomials. This takes advantage of the well-known reproduction properties of B-splines, which are inherited by any family that is able to reproduce them.

2.2. Minimal-support properties for two basis functions

The Hermite splines ϕ_1 and ϕ_2, given by (2) and (3), are able to reproduce both β^2 and β^3, the quadratic and cubic B-splines of order 3 and 4, respectively [13]. In particular, this means that $V(\phi_1, \phi_2)$ contains polynomials of degree at most 3. Many other pairs of basis functions, starting with β^2 and β^3 themselves, can also reproduce quadratic and cubic splines. In line with Proposition 2, the investigation the support of a pair of functions that have the same reproduction properties as the Hermite splines follows naturally. This boils down to the study of basis functions for which $\beta^2, \beta^3 \in V(\varphi_1, \varphi_2)$.

We characterize the support of such pairs of functions in Theorem 3. This result then allows us to deduce the minimal-support property of Hermite splines in Corollary 4.

Theorem 3. *Let $\varphi_1, \varphi_2 \in L_2(\mathbb{R}^d)$ be two compactly supported basis functions. We assume that*

$$\beta^2(t) = \sum_{k \in \mathbb{Z}} a_k \varphi_1(t-k) + b_k \varphi_2(t-k), \tag{8}$$

$$\beta^3(t) = \sum_{k \in \mathbb{Z}} c_k \varphi_1(t-k) + d_k \varphi_2(t-k), \tag{9}$$

with reproduction sequences a, b, c, d that satisfy

$$\sum_{k\in\mathbb{Z}} k^3(|a_k| + |b_k| + |c_k| + |d_k|) < \infty. \tag{10}$$

In particular, the quadratic and cubic B-splines β^2, β^3 are in $V(\varphi_1, \varphi_2)$. Then, we have that

$$|\mathrm{Supp}\varphi_1| + |\mathrm{Supp}\varphi_2| \geq 4. \tag{11}$$

Proof. First of all, one can restrict oneself to compactly supported basis functions φ_1 and φ_2 (otherwise, $|\mathrm{Supp}\varphi_1| + |\mathrm{Supp}\varphi_2| = \infty$). If one of the basis function, for instance φ_2, is identically zero, then the cubic spline $\beta^3 \in V(\varphi_1)$. This means in particular that the basis function φ_1 reproduces polynomials up to degree 3, and its support is therefore at least of size four [16, Theorem 1]. Hence, we again have that $|\mathrm{Supp}\varphi_1| + |\mathrm{Supp}\varphi_2| = |\mathrm{Supp}\varphi_1| \geq 4$. We now assume that φ_1 and φ_2 are not identically 0.

Step 1. We show that the extreme points of the supports of φ_1 and φ_2 are integers. For $x = a, b, c, d$, we set $X(\omega) = \sum_{k\in\mathbb{Z}} x_k \mathrm{e}^{-\mathrm{j}\omega k}$, the 2π-periodic Fourier transform of the sequence x. Condition (10) ensures that X has a periodic continuous third derivative. In the Fourier domain, (8) and (9) become

$$\widehat{\beta^2}(\omega) = \frac{(1-\mathrm{e}^{-\mathrm{j}\omega})^3}{(\mathrm{j}\omega)^3} = A(\omega)\widehat{\varphi_1}(\omega) + B(\omega)\widehat{\varphi_2}(\omega), \tag{12}$$

$$\widehat{\beta^3}(\omega) = \frac{(1-\mathrm{e}^{-\mathrm{j}\omega})^4}{(\mathrm{j}\omega)^4} = C(\omega)\widehat{\varphi_1}(\omega) + D(\omega)\widehat{\varphi_2}(\omega). \tag{13}$$

We set $\det(\omega) = (A(\omega)D(\omega) - B(\omega)C(\omega))$, which is itself a function with continuous third derivative. From (12) and (13), we obtain that

$$\det(\omega)\widehat{\varphi_1}(\omega) = D(\omega)\frac{(1-\mathrm{e}^{-\mathrm{j}\omega})^3}{(\mathrm{j}\omega)^3} - B(\omega)\frac{(1-\mathrm{e}^{-\mathrm{j}\omega})^4}{(\mathrm{j}\omega)^4}, \tag{14}$$

$$\det(\omega)\widehat{\varphi_2}(\omega) = -C(\omega)\frac{(1-\mathrm{e}^{-\mathrm{j}\omega})^3}{(\mathrm{j}\omega)^3} + A(\omega)\frac{(1-\mathrm{e}^{-\mathrm{j}\omega})^4}{(\mathrm{j}\omega)^4}. \tag{15}$$

From (14), we deduce that, at least when $\det(\omega)$ does not vanish, we have the relation

$$(\mathrm{j}\omega)^4\widehat{\varphi_1}(\omega) = (\mathrm{j}\omega)F(\omega)D(\omega) - (1-\mathrm{e}^{-\mathrm{j}\omega})F(\omega)B(\omega), \tag{16}$$

where $F(\omega) = \frac{(1-\mathrm{e}^{-\mathrm{j}\omega})^3}{\det(\omega)}$. The strategy of the proof is to show that the function F is continuous and periodic, and that (16) is therefore valid for any $\omega \in \mathbb{R}$. We study F in two steps: (i) first, we show that $\det(\omega)$ does not vanish for $\omega \notin 2\pi\mathbb{Z}$; and (ii) we then demonstrate that F has a limit at 0 by considering the Taylor expansion of $\det(\omega)$.

(i) Let us start with the first issue. We show that $\det(\omega) \neq 0$ for $\omega \notin 2\pi\mathbb{Z}$ by contradiction. Let us fix $\omega_0 \in (0, 2\pi)$ and assume that $\det(\omega_0) = 0$. We set $\alpha = \frac{1-\mathrm{e}^{-\mathrm{j}\omega_0}}{\mathrm{j}\omega_0}$ and $\beta = \frac{1-\mathrm{e}^{-\mathrm{j}\omega_0}}{\mathrm{j}(\omega_0+2\pi)}$. Then, $\alpha \neq 0$, $\beta \neq 0$, and $\alpha \neq \beta$, while, by periodicity, $\det(\omega_0) = \det(\omega_0 + 2\pi) = 0$. Hence, (14) for $\omega = \omega_0$ and $(\omega_0 + 2\pi)$ implies that

$$\begin{pmatrix} \alpha^3 & -\alpha^4 \\ \beta^3 & -\beta^4 \end{pmatrix} \begin{pmatrix} D(\omega_0) \\ B(\omega_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{17}$$

The matrix being invertible (with determinant $\alpha^3\beta^3(\alpha - \beta) \neq 0$), we deduce that $D(\omega_0) = B(\omega_0) = 0$. Similarly, (15) with $\omega = \omega_0$ and $(\omega_0 + 2\pi)$ implies that $A(\omega_0) = C(\omega_0) = 0$. Injecting this in (12) with $\omega = \omega_0$, we deduce that $\widehat{\beta^2}(\omega_0) = \alpha^3 = 0$, which contradicts our initial assumption.

(ii) We now study $\det(\omega)$ around the origin. The function admits a third-order continuous derivative; hence, it can be McLaurin expanded at 0 as

$$\det(\omega) = \det(0) + \det{}^{(1)}(0)\omega + \frac{1}{2}\det{}^{(2)}(0)\omega^2 + \frac{1}{6}\det{}^{(3)}(0)\omega^3 + o(\omega^3). \tag{18}$$

Assume by contradiction that $\det^{(p)}(0) = 0$ for $p = 0, 1, 2, 3$. Then, (18) gives that $\det(\omega) = o(\omega^3)$ around 0. From (14), we remark that

$$\frac{\det(\omega)(\mathrm{j}\omega)^3}{(1-\mathrm{e}^{-\mathrm{j}\omega})^3}\widehat{\varphi_1}(\omega) = D(\omega) - B(\omega)\frac{1-\mathrm{e}^{-\mathrm{j}\omega}}{\mathrm{j}\omega}. \tag{19}$$

The function $\frac{\det(\omega)(\mathrm{j}\omega)^3}{(1-\mathrm{e}^{-\mathrm{j}\omega})^3}$ vanishes at $\omega = 0$ because det does and $\lim_{\omega\to 0}\frac{(\mathrm{j}\omega)^3}{(1-\mathrm{e}^{-\mathrm{j}\omega})^3} = 1$. Therefore, the left term in (19) vanishes when ω goes to 0. Now, by periodicity, around 2π, we have that $\det(\omega) = o((\omega - 2\pi)^3)$. Hence, $\frac{\det(\omega)}{(1-\mathrm{e}^{-\mathrm{j}\omega})^3} = o(1)$ around 2π and, again, the left term in (19) is also vanishing when ω goes to 2π.

We deduce that the right term in (19) vanishes for both $\omega = 0$ and $\omega = 2\pi$. In other terms, we have that

$$0 = D(0) - B(0) = D(2\pi) - 0 \tag{20}$$

and, since $D(2\pi) = D(0)$ by periodicity, this implies that $D(0) = B(0) = 0$. A similar reasoning shows that $A(0) = C(0) = 0$. From (12) with $\omega = 0$, we obtain that $\widehat{\beta^3}(0) = 0$, which is false.

As a consequence, at least one of the derivative of the McLaurin expansion (18) is nonzero, from which we easily deduce that $F(\omega) = \frac{(1-e^{-j\omega})^3}{\det(\omega)}$ has a limit (possibly 0) at the origin. The function F is well-defined and continuous for $\omega \notin 2\pi\mathbb{Z}$, continuously extendable at 0, and is therefore a continuous periodic function.

At this stage, we obtained that (16) is valid for any $\omega \in \mathbb{R}$. The functions $F(\omega)D(\omega)$ and $(1 - e^{-j\omega})F(\omega)B(\omega)$ are 2π-periodic, hence their inverse Fourier transforms are sums of Dirac impulses located at the integers. It means in particular that we have, in the time domain, that

$$\varphi_1^{(4)}(t) = \sum_{k\in\mathbb{Z}} \left(y_k \delta(t-k) + z_k \delta'(t-k)\right), \tag{21}$$

where y and z are the Fourier sequences of $(1 - e^{-j\omega})F(\omega)B(\omega)$ and $F(\omega)D(\omega)$, respectively. Since $\varphi_1^{(4)}$ is compactly supported, like φ_1, only finitely many y_k and z_k are non-zero. Then, φ_1 is a compactly supported function whose fourth derivative has a support with integer extreme points (due to (21)), and therefore has a support with integer extreme points., too The same reasoning applies for φ_2, which concludes this part of the proof.

Step 2. We know that the supports of φ_1 and φ_2 are of the form $[a, b]$ with $a < b$, $a, b \in \mathbb{Z}$. By contradiction, we assume that $|\text{Supp}\varphi_1| + |\text{Supp}\varphi_2| < 4$. Then, one of the two basis functions has a support of size one, for instance φ_1. We also assume without loss of generality that $\text{Supp}\varphi_1 = [0, 1]$, implying that only y_0, y_1, z_0, z_1 are possibly nonzero in (21). Going back to the Fourier domain, one obtains that

$$(j\omega)^4 \widehat{\varphi}_1(\omega) = y_0 + y_1 e^{-j\omega} + j\omega(z_0 + z_1 e^{-j\omega}). \tag{22}$$

The function φ_1 is compactly supported. Its Fourier transform is hence infinitely smooth, and we can do the McLaurin expansion of both sides in (22) up to order 3. This gives

$$o(\omega^3) = (y_0 + y_1) + j\omega(-y_1 + z_0 + z_1) + \omega^2(-y_1/2 + z_1) + j\omega^3(y_1/6 - z_1/2) + o(\omega^3). \tag{23}$$

In particular, we obtain the relations

$$y_0 + y_1 = z_0 + z_1 - y_1 = z_1 - \frac{y_1}{2} = \frac{y_1}{6} - \frac{z_1}{2} = 0. \tag{24}$$

This imposes that $y_0 = y_1 = z_0 = z_1 = 0$, which is absurd. Finally, it shows that $|\text{Supp}\varphi_1| + |\text{Supp}\varphi_2| \geq 4$, as expected. □

Condition (10) plays an important role in our proof by imposing some regularity in the Fourier domain. In practice, one even expects that compactly supported basis functions can generate the B-splines β^2 and β^3 with finitely many coefficients, in which case (10) is automatically satisfied. However, we believe that (10) can be relaxed up to some extent. From Theorem 3, we easily deduce that Hermite splines have the minimal-support property.

Corollary 4. *The Hermite splines (ϕ_1, ϕ_2) are of minimal support among the pairs of functions that are able to reproduce both quadratic and cubic B-splines with reproduction sequences satisfying (10).*

Proof. From [13, Appendix A], we know that Hermite splines can reproduce both quadratic and cubic B-splines, hence β^2 and β^3 are in $V(\phi_1, \phi_2)$ with compactly supported reproduction sequences obviously satisfying (10). The supports of ϕ_1 and ϕ_2 are of size two, which implies that $|\text{Supp}\phi_1| + |\text{Supp}\phi_2| = 4$. Finally, the pair (ϕ_1, ϕ_2) is maximally localized due to (11). □

It is worth noting that the supports of the pair of Hermite splines jointly has the same size as the B-spline β^3. However, β^3 alone has lesser reproduction properties. Being of class C^2, it is in particular unable to reproduce the quadratic spline β^2, which only has C^1 transitions at the integers. The simplest way of reproducing β^2, β^3 is to consider the basis pair (β^2, β^3) itself, which is not maximally localized since the sum of the supports is 7. An important additional remark is that two functions are needed to reproduce both cubic and quadratic spline, as formalized in Proposition 5.

Proposition 5. *There exists no single function $\varphi \in L_2(\mathbb{R})$ that is able to reproduce β^2 and β^3 with summable reproduction sequences.*

Proof. By contradiction, let us assume that there exists φ such that $\beta^2 = \sum_{k\in\mathbb{Z}} a_k \varphi(\cdot - k)$ and $\beta^3 = \sum_{k\in\mathbb{Z}} b_k \varphi(\cdot - k)$ with $a, b \in \ell_1(\mathbb{Z})$. Then, the Fourier transforms $A(e^{j\omega})$ and $B(e^{j\omega})$ are continuous 2π-periodic functions. In the Fourier domain, we have that

$$\left(\frac{1 - e^{-j\omega}}{j\omega}\right)^3 = A(e^{j\omega})\widehat{\varphi}(\omega), \quad \left(\frac{1 - e^{-j\omega}}{j\omega}\right)^4 = B(e^{j\omega})\widehat{\varphi}(\omega). \tag{25}$$

Set $\omega_0 \in (0, 2\pi)$ and $\omega_1 = \omega_0 + 2\pi$. The relation (25) imposes that $A(e^{j\omega_i})$, $B(e^{j\omega_i})$, and $\widehat{\varphi}(\omega_i)$ are non-zero for $i = 1, 2$, and

$$\frac{1 - e^{-j\omega_i}}{j\omega_i} A(e^{j\omega_i})\widehat{\varphi}(\omega_i) = B(e^{j\omega_i})\widehat{\varphi}(\omega_i). \tag{26}$$

After simplifications, we deduce that

$$\mathrm{j}\omega_i = \frac{(1 - \mathrm{e}^{-\mathrm{j}\omega_i})A(\mathrm{e}^{\mathrm{j}\omega_i})}{B(\mathrm{e}^{\mathrm{j}\omega_i})}. \tag{27}$$

The right term in (27) is equal for ω_0 and ω_1 by periodicity, while the left term is not. This contradicts our initial assumption and implies Proposition 5. □

3. Approximation properties of Hermite splines

Existing approaches have been proposed to characterize the behavior of the approximation error for single [19] and multi-generators [20]. They however assume technical conditions that the Hermite interpolation does not satisfy. We therefore formulate new theoretical tools to quantify the asymptotic constant of the error of approximation of this scheme. Since functions are estimated with derivative samples in the Hermite case, we also propose to investigate the approximation error on the first derivative. In our setting, other existing approximation schemes can be considered as well, allowing us to relate the excellent approximation properties of Hermite splines to other schemes such as cubic B-splines and interlaced derivative sampling.

3.1. Generalized sampling and error of approximation

The approximation of a continuously defined signal from a collection of its samples in a generalized-sampling scheme relies on two ingredients: some basis functions $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_N) \in (L_2(\mathbb{R}))^N$ forming a Riesz basis in the sense of (6), and some *sampling functions* $\tilde{\boldsymbol{\varphi}} = (\tilde{\varphi}_1, \dots, \tilde{\varphi}_N)$ that are rapidly decaying generalized functions. The sampling functions include rapidly decaying L_2-integrable functions together with the Dirac impulse δ, its derivative, and their shifts. We recall that an L_2 function is *rapidly decaying* if it decays faster than the inverse of any polynomial at infinity. A generalized function in $\mathcal{S}'(\mathbb{R})$ is rapidly decaying if its convolution with any infinitely differentiable and rapidly decaying function is a rapidly decaying function [38]. In particular, a rapidly decaying (generalized) function has an infinitely differentiable Fourier transform, a property which we shall rely on thereafter. The term *generalized* is motivated by the fact that sampling functions allow one to access more than just the values of the signal at the sampling points [39].

The set of pairs of basis and sampling functions fully characterizes an approximation scheme. Moreover, because we aim at a fair comparison, the quantity of information per unit of time ought to be equal among the considered schemes. This requires us to slightly adapt the definition of $V(\boldsymbol{\varphi})$ given in (5). From now, we shall consider

$$W(\boldsymbol{\varphi}) = \left\{ \sum_{i=1}^{N} \sum_{k \in \mathbb{Z}} c_i[k] \varphi_i(\cdot - Nk) : \mathbf{c} = (c_1, \dots, c_N) \in (\ell_2(\mathbb{Z}))^N \right\}. \tag{28}$$

The added parameter N ensures that, for any $N \geq 1$, there is on average a single degree of freedom on each interval of size one. The sampling and reconstruction problem is then formally defined as follows: the function f is reconstructed by its approximation $\tilde{f}$ associated to the basis functions $\boldsymbol{\varphi}$ and for the sampling functions $\tilde{\boldsymbol{\varphi}}$, defined as

$$\tilde{f} = \sum_{i=1}^{N} \sum_{k \in \mathbb{Z}} \langle f, \tilde{\varphi}_i(\cdot - Nk) \rangle \varphi_i(\cdot - Nk) \in W(\boldsymbol{\varphi}). \tag{29}$$

We denote by $\mathcal{Q}_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}$ the operator such that $\mathcal{Q}_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}} f = \tilde{f}$. Hermite-spline approximation thus corresponds to $N = 2$ with basis functions $\varphi_1(t) = \phi_1(\frac{t}{2})$ and $\varphi_2(t) = 2\phi_2(\frac{t}{2})$, and sampling functions $\tilde{\varphi}_1(t) = \delta(t)$ and $\tilde{\varphi}_2(t) = -\delta'(t)$.

The best approximation of a given scheme is obtained when the pairs of sampling and basis functions are properly chosen such that $\tilde{f}$ is the orthogonal projection of f onto $W(\boldsymbol{\varphi})$. This implies the imposition of a particular condition [20] on the sampling functions $\tilde{\boldsymbol{\varphi}}$, namely, that

$$\tilde{\boldsymbol{\varphi}} = \begin{pmatrix} \varphi_{1,d} \\ \vdots \\ \varphi_{N,d} \end{pmatrix} = \mathcal{F}^{-1}\{\mathbf{G}_{\boldsymbol{\varphi}}(\cdot)^{-1} \widehat{\boldsymbol{\varphi}}\}, \tag{30}$$

where $\mathbf{G}_{\boldsymbol{\varphi}}$ is the Gram matrix of size $(N \times N)$ associated to $\boldsymbol{\varphi}$, given for $\omega \in \mathbb{R}$ by

$$\mathbf{G}_{\boldsymbol{\varphi}}(\omega) = \sum_{k \in \mathbb{Z}} \widehat{\boldsymbol{\varphi}}(\omega + 2k\pi) \widehat{\boldsymbol{\varphi}}^{*T}(\omega + 2k\pi). \tag{31}$$

This particular collection of sampling functions is called the *dual functions* associated to $\boldsymbol{\varphi}$ and is denoted by $\boldsymbol{\varphi}_d$. Note that $\mathbf{G}_{\boldsymbol{\varphi}}(\omega)$ is invertible for every $\omega \in \mathbb{R}$ and, therefore, (30) is meaningful because (6) is equivalent to $0 < A \leq \lambda_{\min}(\omega) \leq \lambda_{\max}(\omega) \leq B < \infty$, where $\lambda_{\min}(\omega)$ ($\lambda_{\max}(\omega)$, respectively) is the minimum (maximum, respectively) eigenvalue

of $\mathbf{G}_{\boldsymbol{\varphi}}(\omega)$ [40]. When $\tilde{\boldsymbol{\varphi}}$ is defined as (30), the operator $\mathcal{Q}_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}$ is the orthogonal projector over $W(\boldsymbol{\varphi})$, denoted as $\mathcal{P}_{\boldsymbol{\varphi}} = \mathcal{Q}_{\boldsymbol{\varphi}}^{\boldsymbol{\varphi}_d}$. In this situation, (29) is reformulated as

$$\tilde{f}(t) = \mathcal{P}_{\boldsymbol{\varphi}} f = \sum_{i=1}^{N} \sum_{k\in\mathbb{Z}} \langle f, \varphi_{i,d}(\cdot - Nk)\rangle \varphi_i(\cdot - Nk). \tag{32}$$

To simplify the notation, we shall write $\mathcal{P} = \mathcal{P}_{\boldsymbol{\varphi}}$ and $\mathcal{Q} = \mathcal{Q}_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}$ thereafter. The quality of the approximation is then evaluated in terms of the error of approximation which is expressed as

$$E_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}(f) = \|f - \tilde{f}\|_{L_2} = \|f - \mathcal{Q}f\|_{L_2}. \tag{33}$$

When $\tilde{f} = \mathcal{P}f$, the error is denoted as $E_{\boldsymbol{\varphi}}(f)$. A direct implication is that $E_{\boldsymbol{\varphi}}(f) = E_{\boldsymbol{\varphi}}^{\boldsymbol{\varphi}_d}(f) \leq E_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}(f)$, reaching the equality if and only if $\tilde{\boldsymbol{\varphi}} = \boldsymbol{\varphi}_d$.

Up to now, we considered approximation schemes with (generalized) samples taken on the integer grid, which corresponds to the sampling step $T = 1$. This parameter affects the coarseness of the approximation: when $T \to 0$, the error is expected to vanish. For $T > 0$, the approximation space (28) becomes

$$W_T(\boldsymbol{\varphi}) = \left\{ \sum_{i=1}^{N} \sum_{k\in\mathbb{Z}} c_i[k] \varphi_i\left(\frac{\cdot}{T} - Nk\right) \ : \ \mathbf{c} \in (\ell_2(\mathbb{Z}))^N \right\} \tag{34}$$

and the approximation of f is given by

$$\tilde{f}_T = \mathcal{Q}_T f = \sum_{i=1}^{N} \sum_{k\in\mathbb{Z}} \left\langle f, \frac{1}{T}\tilde{\varphi}_i\left(\frac{\cdot}{T} - Nk\right)\right\rangle \varphi_i\left(\frac{\cdot}{T} - Nk\right), \tag{35}$$

with resulting error

$$E_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}(f, T) = \|f - \mathcal{Q}_T f\|_{L_2}. \tag{36}$$

The orthogonal projector (32) and its associated error are easily reformulated accordingly.

A number of hypotheses on the basis functions $\boldsymbol{\varphi}$ and the sampling functions $\tilde{\boldsymbol{\varphi}}$ have to be met in order to study the errors $E_{\boldsymbol{\varphi}}(f, T)$ and $E_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}(f, T)$ in terms of rate of decay and asymptotic constant. The first one is the Riesz-basis condition (6), which ensures a unique and stable representation. The second one is the order of the basis functions. We shall consider basis functions $\boldsymbol{\varphi}$ of a given order in the sense of Definition 1.

When it is met, the decrease of the optimal error $E_{\boldsymbol{\varphi}}(f, T)$ is bounded from above by T^L[20]. The last important condition is that the sampling and basis functions are quasi-biorthonormal of order L.

Definition 6. Two families of bases $\boldsymbol{\varphi}$ and sampling functions $\tilde{\boldsymbol{\varphi}}$ are *quasi-biorthonormal of order L* if the basis functions are of order L and, for the dual function $\boldsymbol{\varphi}_d$ given by (30) and all $\ell = 0, \ldots, (L-1)$, we have that

$$\int_{\mathbb{R}} t^{\ell} \tilde{\boldsymbol{\varphi}}(t) \mathrm{d}t = \int_{\mathbb{R}} t^{\ell} \boldsymbol{\varphi}_d(t) \mathrm{d}t. \tag{37}$$

It is worth noting that (37) is a slight abuse of notation, since the $\tilde{\varphi}_i$ are not necessarily defined pointwise. However, they are assumed to be rapidly decaying generalized functions and can therefore be taken against a slowly growing smooth function such as $t \mapsto t^{\ell}$. The right term in (37) is well-defined due to the condition $\int_{\mathbb{R}} (1 + |t|)^L |\varphi_n(t)| \mathrm{d}t < \infty$ in Definition 1, as discussed in [20].

The quasi-biorthonormality ensures that the decrease of the error $E_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}(f, T)$ is also bounded by T^L[20]. Finally, the rate of decay of the error of approximation being under control in all generality, the asymptotic constant can be obtained as

$$C_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}(f) = \lim_{T\to 0} T^{-L} E_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}}(f, T) \tag{38}$$

which, in practice, can be computed relying on a Fourier-domain approximation-error kernel. For more details, we refer the interested reader to [41–43] and references therein.

3.2. Approximation constants of irregular sampling schemes

Our goal in this section is to extend the main results of [20] so as to include Hermite-spline approximations. The original framework is indeed restricted to sampling functions $\tilde{\varphi}$ with bounded Fourier transforms. While this allows one to consider the case of interpolation, it excludes Hermite-spline approximations since the Fourier transform $\mathrm{j}\omega$ of δ' is unbounded.

In what follows, it will be useful to consider Sobolev spaces of integer order N. For $N \geq 0$, we define $W_2^N(\mathbb{R})$ as the space of functions f such that

$$\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 (1 + \omega^2)^N \mathrm{d}\omega < \infty. \tag{39}$$

For $N = 0$, this corresponds to the space $L_2(\mathbb{R})$. In general, $f \in W_2^N(\mathbb{R})$ if and only if $f, \dots, f^{(N)} \in L_2(\mathbb{R})$. For technical reasons, we also consider Sobolev spaces of fractional order $\gamma \geq 0$, for which we simply replace the integer N by a nonnegative real number γ in (39). Fractional Sobolev spaces are intimately connected to fractional derivatives in the following sense: the fractional derivative of order $\gamma \geq 0$ is defined in the Fourier domain by

$$\mathcal{F}\{f^{(\gamma)}\}(\omega) = (\mathrm{j}\omega)^{\gamma}\widehat{f}(\omega) \tag{40}$$

for any $f \in W_2^{\gamma}(\mathbb{R})$. By definition of the Sobolev space, one has that $f^{(\gamma)} \in L_2(\mathbb{R})$ as soon as $f \in W_2^{\gamma}(\mathbb{R})$. In general, we have the Parseval-type relation $\|f^{(\gamma)}\|_{L_2}^2 = \int_{\mathbb{R}} |\widehat{f}(\omega)|^2 |\omega|^{2\gamma} \mathrm{d}\omega$.

Our development follows the key contributions of [20]. We therefore only detail the adaptations that are required in our case. We start with a brief summary of the general approach, which is common to many works of approximation theory in shift-invariant spaces. We first introduce the kernels associated to the functions $\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}}$ as

$$\mathcal{E}_{\min}(\omega) = 1 + \widehat{\boldsymbol{\varphi}}^{*T}(\omega)\mathbf{G}_{\boldsymbol{\varphi}}^{-1}(\omega)\widehat{\boldsymbol{\varphi}}(\omega), \tag{41}$$

$$\mathcal{E}_{\mathrm{res}}(\omega) = (\widehat{\tilde{\boldsymbol{\varphi}}} - \widehat{\boldsymbol{\varphi}}_d)^{*T}(\omega)\mathbf{G}_{\boldsymbol{\varphi}}(\omega)(\widehat{\tilde{\boldsymbol{\varphi}}} - \widehat{\boldsymbol{\varphi}}_d)(\omega), \tag{42}$$

$$\mathcal{E}(\omega) = \mathcal{E}_{\min}(\omega) + \mathcal{E}_{\mathrm{res}}(\omega), \tag{43}$$

where $\mathbf{G}_{\boldsymbol{\varphi}}$ is the Gram matrix (31). The kernel $\mathcal{E}_{\min}$ relates to the *minimum-error* case achieved using the orthogonal projector (*i.e.*, $\tilde{\boldsymbol{\varphi}} = \boldsymbol{\varphi}_d$), and $\mathcal{E}_{\mathrm{res}}$ to the *residual error* that arise when using sampling functions that differ from the dual functions. We furthermore note that $\mathcal{E}_{\mathrm{res}}(\omega) = 0$ when $\tilde{\boldsymbol{\varphi}} = \boldsymbol{\varphi}_d$, as expected. The key ideas of relying on $\mathcal{E}$ are as follows.

- The kernel $\mathcal{E}$ measures the approximation power of a reasonable approximation scheme $(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}})$, in the sense that $\|f - \mathcal{Q}_T f\|_{L_2} \approx \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega)\mathrm{d}\omega\right)^{1/2}$ for small $T > 0$.
- The precise behavior of $\|f - \mathcal{Q}_T f\|_{L_2}$ is then deduced from that of $\mathcal{E}$ around the origin. It depends on the approximation order L of the scheme and, typically, behaves like CT^L. The constant C depends on the function f to approximate and on the scheme $(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}})$ via the Taylor expansion of $\mathcal{E}$.

We give a precise meaning to these two points in Propositions 8 and 9. Before that, we recall an important lemma taken from [20] which will play a fundamental role in our proofs.

Lemma 7. *Let $f \in W_2^{L+1}(\mathbb{R})$. For $k \geq 0$, we set $\widehat{f}_k(\omega) = \widehat{f}(\omega)1_{k/T \leq |\omega| < (k+1)/T}$. Then, the following relations hold:*

$$\widehat{f} = \sum_{k \geq 0} \widehat{f}_k; \tag{44}$$

$$\|f_k - \mathcal{Q}_T f_k\|_{L_2}^2 = \int_{\mathbb{R}} |\widehat{f}_k(\omega)|^2 \mathcal{E}(T\omega)\mathrm{d}\omega \text{ for } k \geq 0; \text{ and} \tag{45}$$

$$\left| \|f - \mathcal{Q}_T f\|_{L_2} - \|f_0 - \mathcal{Q}_T f_0\|_{L_2} \right| \leq \sum_{k>0} \|f_k - \mathcal{Q}_T f_k\|_{L_2}. \tag{46}$$

Equality (44) is obvious. The two next relations come from [20, Theorem 1]; we have simply reformulated the results with our notation. First, it is proven (see [20, Theorem 1]) that $\|f - \mathcal{Q}_T f\|_{L_2}^2 = \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} |\widehat{f}_k(\omega)|^2 \mathcal{E}(T\omega)\mathrm{d}\omega$ as soon as $\widehat{f}(\omega)\overline{\widehat{f}(\omega - n/T)} = 0$ for any $\omega \in \mathbb{R}$ and $n \in \mathbb{Z}$, a condition that is satisfied by the f_k by construction, giving (45). Moreover, the inequality (46) appears in the proof of [20, Theorem 1] (see (63)).

Proposition 8. *Let $(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}})$ be a set of N basis and sampling functions that are quasi-biorthonormal and provide an approximation scheme of order L. We assume moreover that the kernel $\mathcal{E}$ given by (43) satisfies*

$$|\mathcal{E}(\boldsymbol{\omega})| \leq C^2 \max(1, |\omega|^{2p}) \tag{47}$$

for some $C > 0$, some integer $0 \leq p \leq L$, and every $\omega \in \mathbb{R}$. Then, for every $f \in W_2^{L+1}(\mathbb{R})$, we have that

$$\|f - \mathcal{Q}_T f\|_{L_2} = \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega)\mathrm{d}\omega\right)^{1/2} + O(T^{L+1}). \tag{48}$$

The case $p = 0$ corresponds to a bounded $\mathcal{E}$ and can be found in [20, Theorem 1].

Proof. We recall that $\widehat{f}_k(\omega)$ is defined as $\widehat{f}_k(\omega) = \widehat{f}(\omega)1_{k/T \leq |\omega| < (k+1)/T}$, and that $f_k = \mathcal{F}^{-1}\{\widehat{f}_k\}$. Then, we have that

$$\begin{aligned}
&\left| \|f - \mathcal{Q}_T f\|_{L_2} - \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega)\mathrm{d}\omega\right)^{1/2} \right| \\
&\quad \leq \left| \|f - \mathcal{Q}_T f\|_{L_2} - \|f_0 - \mathcal{Q}_T f_0\|_{L_2} \right| + \left| \|f_0 - \mathcal{Q}_T f_0\|_{L_2} - \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega)\mathrm{d}\omega\right)^{1/2} \right|
\end{aligned}$$

$$\overset{(i)}{\leq} \sum_{k>0} \|f_k - \mathcal{Q}_T f_k\|_{L_2} + \left| \|f_0 - \mathcal{Q}_T f_0\|_{L_2} - \left(\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) \mathrm{d}\omega \right)^{1/2} \right|$$
$$= (I) + (II), \tag{49}$$

where we used (46) in (i). As a consequence, (48) follows if one shows that the two terms (I) and (II) in (49) are $O(T^{L+1})$. Using (45), we have that

$$\begin{aligned} \|f_k - \mathcal{Q}_T f_k\|_{L_2}^2 &= \int_{k/T<|\omega|\leq(k+1)/T} |\widehat{f_k}(\omega)|^2 \mathcal{E}(T\omega) \mathrm{d}\omega \\ &= \int_{k/T<|\omega|\leq(k+1)/T} |\widehat{f_k}(\omega)|^2 |\omega|^{2(L+1)} \frac{T^{2p}}{|\omega|^{2(L-p+1)}} \frac{\mathcal{E}(T\omega)}{T^{2p}|\omega|^{2p}} \mathrm{d}\omega \\ &\leq C^2 \frac{T^{2p}}{(k/T)^{2(L-p+1)}} \int_{\mathbb{R}} |\widehat{f_k}(\omega)|^2 |\omega|^{2(L+1)} \mathrm{d}\omega \\ &= C^2 \frac{T^{2(L+1)}}{k^{2(L-p+1)}} \|f_k^{(L+1)}\|_{L_2}^2, \end{aligned} \tag{50}$$

where the inequality is due to $|\omega| \geq k/T$ over the domain and to (47), which implies that $C^2 \geq \sup_{\omega\in\mathbb{R}} \mathcal{E}(\omega)/\max(1, |\omega|^p) \geq \sup_{|\omega|\geq 1} \mathcal{E}(\omega)/|\omega|^p$. Summing over $k > 0$, we deduce that

$$\begin{aligned} (I) &\leq CT^{L+1} \sum_{k>0} \frac{1}{k^{L-p+1}} \|f_k^{(L+1)}\|_{L_2} \\ &\leq CT^{L+1} \left(\sum_{k>0} \frac{1}{k^{2(L-p+1)}} \right)^{1/2} \left(\sum_{k>0} \|f_k^{(L+1)}\|_{L_2}^2 \right)^{1/2} \\ &= CT^{L+1} \sqrt{\zeta(2(L-p+1))} \|(f - f_0)^{(L+1)}\|_{L_2} = O(T^{L+1}), \end{aligned} \tag{51}$$

where the second inequality is derived from Cauchy–Schwarz, and ζ is the Riemann zeta function. For the second term, we remark that, again due to [20, (27)], $\|f_0 - \mathcal{Q}_T f_0\|_{L_2}^2 = \int_{|\omega|\leq 1/T} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) \mathrm{d}\omega$. Therefore, using the relation $|\|g\|_{L_2} - \|h\|_{L_2}| \leq \|g - h\|_{L_2}$ (a consequence of the Minkowski inequality) we deduce that

$$\begin{aligned} (II) &\leq \left(\int_{|\omega|>1/T} |\widehat{f}(\omega)|^2 \mathcal{E}(T\omega) \mathrm{d}\omega \right)^{1/2} \\ &= \left(\int_{|\omega|>1/T} |\widehat{f}(\omega)|^2 |\omega|^{2(L+1)} \frac{T^{2p}}{|\omega|^{2(L-p+1)}} \frac{\mathcal{E}(T\omega)}{T^{2p}|\omega|^{2p}} \mathrm{d}\omega \right)^{1/2} \\ &\leq CT^p T^{L-p+1} \left(\int_{|\omega|>1/T} |\widehat{f}(\omega)|^2 |\omega|^{2(L+1)} \mathrm{d}\omega \right)^{1/2} \\ &= CT^{L+1} \|(f - f_0)^{(L+1)}\|_{L_2} = O(T^{L+1}), \end{aligned} \tag{52}$$

where the inequality once more follows from $|\omega| \geq 1/T$ over the domain and from our assumption (47). Combining (51) and (52) in (49) completes the proof. □

Our next result connects the expression $\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(Tw) \mathrm{d}\omega$ to the McLaurin expansion of $\mathcal{E}$.

Proposition 9. *Let $(\varphi, \tilde{\varphi})$ be a set of N basis and sampling functions that are quasi-biorthonormal and provide an approximation scheme of order L. We assume moreover that the kernel $\mathcal{E}$ given by (43) is $(2L + 1)$-times continuously differentiable and satisfies*

$$|\mathcal{E}^{(2L+1)}(\omega)| \leq C^2 \max(1, |\omega|^{2p}) \tag{53}$$

for some $C > 0$, some integer $0 \leq p \leq L$, and every $\omega \in \mathbb{R}$. Then, for every $f \in W_2^{L+p+1/2}(\mathbb{R})$, we have that

$$\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(Tw) \mathrm{d}\omega = \frac{\mathcal{E}^{(2L)}(0)}{(2L)!} \|f^{(L)}\|_{L_2}^2 T^{2L} + O(T^{2L+1}). \tag{54}$$

Proof. The kernel $\mathcal{E}$ being symmetric, we deduce that $\mathcal{E}^{(2k+1)}(0) = 0$ for $k = 0, \ldots, (L - 1)$. Moreover, the condition of $(\varphi, \tilde{\varphi})$ ensures that $\mathcal{E}^{(2k)}(0) = 0$ for $k = 0, \ldots, (L - 1)$, together with $\mathcal{E}^{(2L)}(0) \neq 0$. Therefore, $\mathcal{E}$ being $(2L + 1)$-times continuously differentiable, its McLaurin expansion is given by

$$\mathcal{E}(\omega) = \frac{\mathcal{E}^{(2L)}(0)}{(2L)!} \omega^{2L} + \frac{\mathcal{E}^{(2L+1)}(\omega\theta)}{(2L+1)!} w^{2L+1}, \tag{55}$$

with $\theta = \theta(\omega) \in (0, 1)$. Using (55), we deduce that

$$\begin{aligned}\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(Tw)\mathrm{d}\omega &= \int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \frac{\mathcal{E}^{(2L)}(0)}{(2L)!} T^{2L}\omega^{2L}\mathrm{d}\omega \\ &\quad + \frac{T^{2L+1}}{(2L+1)!}\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \omega^{2L+1}\mathcal{E}^{(2L+1)}(\omega T\theta(\omega T))\mathrm{d}\omega \\ &= \frac{\mathcal{E}^{(2L)}(0)}{(2L)!}\|f^{(L)}\|_{L_2}^2 T^{2L} \\ &\quad + \frac{T^{2L+1}}{(2L+1)!}\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \omega^{2L+1}\mathcal{E}^{(2L+1)}(\omega T\theta(\omega T))\mathrm{d}\omega. \end{aligned} \tag{56}$$

Due to (53) and $0 < \theta(\omega T) < 1$,

$$\mathcal{E}^{(2L+1)}(\omega T\theta(\omega T)) \leq C^2 \max(1, |\omega|^{2p}T^{2p}) \leq C^2(1 + |\omega|^{2p}T^{2p}), \tag{57}$$

from which we deduce that $\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \omega^{2L+1}\mathcal{E}^{(2L+1)}(\omega T\theta(\omega T))\mathrm{d}\omega \leq C^2(\|f^{(L+1/2)}\|_{L_2}^2 + T^{2p}\|f^{(L+p+1/2)}\|_{L2}^2)$. Note that we refer here to fractional derivatives, as defined in (40). Injecting this to (56) implies that

$$\begin{aligned}\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \mathcal{E}(Tw)\mathrm{d}\omega - \frac{\mathcal{E}^{(2L)}(0)}{(2L)!}\|f^{(L)}\|_{L_2}^2 T^{2L} &= C^2 \frac{T^{2L+1}}{(2L+1)!}(\|f^{(L+1/2)}\|_{L_2}^2 \\ &\quad + T^{2p}\|f^{(L+p+1/2)}\|_{L2}^2) \\ &= O(T^{2L+1}), \end{aligned} \tag{58}$$

which concludes the proof. □

Finally, we conclude with an extension of [20, Theorem 4] to sampling functions that are not necessarily bounded in the Fourier domain.

Theorem 10. *We consider an approximation scheme $(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}})$ with N basis and sampling functions such that*

- *the basis functions are rapidly decaying L_2 functions such that the family $\{\varphi_i(\cdot - Nk)\}_{i=1\ldots N, k\in\mathbb{Z}}$ is a Riesz basis in the sense of (6), with approximation power of order $L \geq 1$ (see Definition 1);*
- *the sampling functions are rapidly decaying generalized functions such that $(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}})$ is quasi-biorthonormal of order L (see Definition 6);*
- *there exists an integer $0 \leq p \leq L$ such that, for any $\omega \in \mathbb{R}$, any $0 \leq k \leq L$, and any $1 \leq i \leq N$,*

$$|\widehat{\tilde{\varphi}}_i^{(k)}(\omega)| \leq C \max(1, |\omega|^p). \tag{59}$$

Then, for any $f \in W^{L+\max(p+1/2,1)}(\mathbb{R})$, we have that

$$\|f - \mathcal{Q}_T f\|_{L_2} \underset{T\to 0}{\sim} \sqrt{\frac{\mathcal{E}^{(2L)}(0)}{(2L)!}}\|f^{(L)}\|_{L_2} T^L. \tag{60}$$

The two first conditions in Theorem 10 are necessary to ensure that the approximation scheme has an approximation power of order L. The last condition allows us to show that such an approximation power is attained with restricted condition on the sampling functions having a possibly unbounded Fourier transform.

Proof. We prove that the conditions of Theorem 10 imply that we fulfill the hypotheses of both Propositions 8 and 9.

Since $\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}}$ are rapidly decaying (generalized) functions, their Fourier transforms $\widehat{\boldsymbol{\varphi}}$ and $\widehat{\tilde{\boldsymbol{\varphi}}}$ are infinitely differentiable. The same holds true for $\widehat{\varphi}_d = \mathbf{G}_{\boldsymbol{\varphi}}^{-1}\widehat{\varphi}$ since $\mathbf{G}_{\boldsymbol{\varphi}}$ is a matrix-valued infinitely differentiable function with infinitely differentiable inverse $\mathbf{G}_{\boldsymbol{\varphi}}^{-1}$ due to the Riesz basis condition. This implies that $\mathcal{E}$ is infinitely differentiable and therefore $(2L+1)$-times continuously differentiable.

The basis functions are rapidly decaying, hence in $L_1(\mathbb{R})$, which implies that $\widehat{\boldsymbol{\varphi}}$ is bounded. The Riesz-basis condition then implies that both $\mathbf{G}_{\boldsymbol{\varphi}}$ and its inverse are bounded as functions of ω. It then follows that $\mathcal{E}_{\min}(\omega)$ is bounded, while $\mathcal{E}_{\text{res}}(\omega)$ is dominated by $\|(\widehat{\boldsymbol{\varphi}} - \widehat{\boldsymbol{\varphi}}_d)(\omega)\|$, itself being dominated by $\max(1, |\omega|^{2p})$ due to (59). It finally implies (47) for some constant $C > 0$. Similarly, for $k \leq L$, the function $\widehat{\varphi}^{(k)}$ is bounded, being the Fourier transform of $t \mapsto t^k\varphi(t)$, which is in $L_1(\mathbb{R})$ due to the rapid decay of φ. Again, this property is transferred to the derivative of $\mathbf{G}_{\boldsymbol{\varphi}}$ and its inverse. Then, exploiting the Leibnitz rule, one shows that $\mathcal{E}_{\min}^{(2L+1)}$ is bounded, while $\mathcal{E}^{(2L+1)}$ is a sum of products made of bounded terms and of two terms of the form $\widehat{\tilde{\varphi}}_i(\omega)$, which are controlled by (59). Putting things together, we deduce (53) for some constant $C > 0$.

Finally, the hypotheses of Propositions 8 and 9 are satisfied, implying (48) and (54), together with

$$\|f - \mathcal{Q}_T f\|_{L_2} = \sqrt{\frac{\mathcal{E}^{(2L)}(0)}{(2L)!}}\|f^{(L)}\|_{L_2} T^L + O(T^{L+1/2}). \tag{61}$$

This proves (60). □

Note that [20, Theorem 4] corresponds to Theorem 10 with $p = 1$, and with less restrictive assumptions on $\boldsymbol{\varphi}$ and $\tilde{\boldsymbol{\varphi}}$ (essentially, the rapid decay is replaced by a polynomial decay adapted to L). The condition $f \in W^{L+\max(p+1/2,1)}(\mathbb{R})$ reflects the cost of covering the scenario of possibly irregular sampling functions with unbounded Fourier transforms.

3.3. Approximation properties of Hermite splines

We can now evaluate the approximation error on f in different frameworks, including the Hermite scheme. In our analysis, for $(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}})$ being fixed, we also quantify the error on the derivative f' when we approximate it by $(\mathcal{Q}_T f)'$. We therefore study the quantities $\|f - \mathcal{Q}_T f\|_{L_2}$ and $\|f' - (\mathcal{Q}_T f)'\|_{L_2}$. Knowing the order of approximation $L \geq 1$, the quality of the approximation is quantified by the two asymptotic constants

$$\mathbf{C}_{\boldsymbol{\varphi}}^{\tilde{\boldsymbol{\varphi}}} = \begin{pmatrix} C_{\boldsymbol{\varphi},1}^{\tilde{\boldsymbol{\varphi}}} \\ C_{\boldsymbol{\varphi},2}^{\tilde{\boldsymbol{\varphi}}} \end{pmatrix} = \begin{pmatrix} \lim_{T\to 0} T^{-L} \|f^{(L)}\|_{L_2}^{-1} \|f - \mathcal{Q}_T f\|_{L_2} \\ \lim_{T\to 0} T^{-(L-1)} \|f^{(L-1)}\|_{L_2}^{-1} \|f' - (\mathcal{Q}_T f)'\|_{L_2} \end{pmatrix}. \tag{62}$$

The asymptotic constant $C_{\boldsymbol{\varphi},2}^{\tilde{\boldsymbol{\varphi}}}$ can be computed with the same tools as $C_{\boldsymbol{\varphi},1}^{\tilde{\boldsymbol{\varphi}}}$. Using integration by parts, we indeed have that

$$(\mathcal{Q}_T f)' = \frac{1}{T} \sum_{i=1}^{N} \sum_{k\in\mathbb{Z}} \left\langle f, \frac{1}{T} \tilde{\varphi}_i \left(\frac{\cdot}{T} - Nk\right) \right\rangle \varphi_i' \left(\frac{\cdot}{T} - Nk\right) \tag{63}$$

$$= \frac{1}{T} \sum_{i=1}^{N} \sum_{k\in\mathbb{Z}} \left\langle f', \frac{1}{T} \tilde{\varphi}_{i,\mathrm{int}} \left(\frac{\cdot}{T} - Nk\right) \right\rangle \varphi_i \left(\frac{\cdot}{T} - Nk\right), \tag{64}$$

where the new sampling functions are best defined in the Fourier domain as

$$\widehat{\tilde{\varphi}_{i,\mathrm{int}}}(\omega) = -\frac{1}{\mathrm{j}\omega} \widehat{\tilde{\varphi}_i}(\omega). \tag{65}$$

As one power of T gets lost in the differentiation process, the rate of decay of the error on f' is $(L-1)$. Note that the apparent singularity in (65) around 0 is counterbalanced in the analysis since one only approximates functions f' that are derivatives functions, hence, for which $\widehat{f'}(\omega) = \mathrm{j}\omega \widehat{f}(\omega)$.

In the Hermite framework, for which $N = 2$, sampling functions are taken as $\tilde{\varphi}_1 = \delta$ and $\tilde{\varphi}_2 = -\delta'$ and basis functions as (2) and (3). For comparison purpose, we also consider two relevant schemes that fit our analysis framework: classical cubic B-splines and interlaced derivative sampling. Cubic B-spline approximation corresponds to $N = 1$, with $\tilde{\varphi} = \sum_{k\in\mathbb{Z}} (b^3)^{-1}[k]\delta(\cdot - k)$, where $(b^3)^{-1}$ is the direct B-spline filter and $\varphi = \beta^3$ the cubic B-spline [37]. Interlaced derivative sampling can be defined in the more general framework of generalized sampling without band-limited constraints [41]. In this setting, $N = 2$ and the sampling functions correspond to $\tilde{\varphi}_1 = \delta$ and $\tilde{\varphi}_2 = -\delta'\left(\cdot - \frac{1}{2}\right)$. The basis functions are constructed from the cubic B-spline to allow for a fair comparison and are given by

$$\widehat{\varphi}_1(\omega) = \frac{3\mathrm{e}^{-2\mathrm{j}\omega} \left(-1 + \mathrm{e}^{\mathrm{j}\omega}\right)^4}{2\omega^4}, \tag{66}$$

$$\widehat{\varphi}_2(\omega) = \frac{\mathrm{e}^{-2\mathrm{j}\omega} \left(-1 + \mathrm{e}^{\mathrm{j}\omega}\right)^4 \left(1 + \mathrm{e}^{\mathrm{j}\omega} \left(-4 + \mathrm{e}^{\mathrm{j}\omega}\right)\right)}{\left(2 - 2\mathrm{e}^{2\mathrm{j}\omega}\right) \omega^4} \tag{67}$$

in the Fourier domain. A comprehensive description of their derivation is provided in [43]. In particular, the family (φ_1, φ_2) is known to be of order 4. It is worth noting that the sampling functions $(\tilde{\varphi}_1, \tilde{\varphi}_2) = (\delta, -\delta'(\cdot - 1/2))$ do not correspond to the dual functions in these frameworks. This is easily motivated by practical considerations: the sampling process is, in practice, implemented with digital filtering, which excludes dual functions due to their continuous nature. The dual functions can, however, still be constructed from the Gram matrix following (30) so as to estimate the optimal approximation error.

We now reveal the approximation power of these three different schemes. The results are known for the cubic B-splines [19] and are included for comparison purposes. For Hermite approximation and interlaced sampling, they are deduced from Theorem 10 and are not included in the multi-generator framework [20], whose hypotheses exclude the use of derivative samples.

Proposition 11. *Let $(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}})$ be one of the three approximation schemes considered above. Then, we have that*

$$\begin{aligned} \|f - \mathcal{Q}_T f\|_{L_2} &\underset{T\to 0}{\sim} \frac{1}{72\sqrt{70}} \|f^{(4)}\|_{L_2} T^4, \\ \|f' - (\mathcal{Q}_T f)'\|_{L_2} &\underset{T\to 0}{\sim} \frac{1}{12\sqrt{210}} \|f^{(3)}\|_{L_2} T^3, \end{aligned} \tag{68}$$

for every $f \in W_2^5(\mathbb{R})$ (cubic B-splines) or $f \in W_2^{11/2}(\mathbb{R})$ (Hermite splines or interlaced derivative sampling). Moreover, in the three cases, the optimal approximation scheme associated to $\boldsymbol{\varphi}$ leads to:

$$\begin{aligned} \|f - \mathcal{Q}_T f\|_{L_2} &\underset{T\to 0}{\sim} \sqrt{\frac{10}{3}} \|f - \mathcal{P}_T f\|_{L_2}, \\ \|f' - (\mathcal{Q}_T f)'\|_{L_2} &\underset{T\to 0}{\sim} \|f' - (\mathcal{P}_T f)'\|_{L_2}, \end{aligned} \tag{69}$$

for every $f \in W_2^5(\mathbb{R})$ (cubic B-splines) or $f \in W_2^{11/2}(\mathbb{R})$ (Hermite splines and interlaced derivative sampling).

Proof. The three frameworks are readily known to define approximation schemes of order $L = 4$ [13,37,41]. All the considered basis functions specify a Riesz basis, are rapidly decaying (Hermite and cubic B-splines are compactly supported while the basis functions for interlaced derivative-sampling, despite being non-compactly supported, are exponentially decaying [41]), and reproduce polynomials up to degree 3. The sampling functions are rapidly decaying generalized functions. Indeed, they are compactly supported for the Hermite and interlaced derivative-sampling schemes. For cubic B-splines, we have seen that $\tilde{\varphi} = \sum_{k\in\mathbb{Z}} (b^3)^{-1}[k]\delta(\cdot - k)$, where the sequence $(b^3)^{-1}$ is known to be exponentially decaying [37], implying the result. In addition, for the Hermite and interlaced derivative sampling schemes (for cubic splines, respectively), (59) is clearly satisfied with $p = 1$ ($p = 0$, respectively). The conditions of Theorem 10 are therefore satisfied for $L = 4$, implying (60).

The value of $\mathcal{E}^{(2L)}(0)$ is computed by computing the McLaurin expansion of the kernel $\mathcal{E}$.[1] The analysis of the approximation error on the derivative follows the same principle, the kernel being given for $(\boldsymbol{\varphi}', \tilde{\boldsymbol{\varphi}}_{\text{int}})$ according to (65), giving (68).

We obtain the asymptotic behavior of the optimal approximation errors $\|f - \mathcal{P}_T f\|_{L_2}$ and $\|f' - (\mathcal{P}_T f)'\|_{L_2}$ associated to $\tilde{\boldsymbol{\varphi}} = \boldsymbol{\varphi}_d$ in the same way, leading to (69). □

Our findings call for three comments.

- The three schemes being compared have the same approximation order, the same approximation constant, and the same optimal approximation constant (associated to $(\boldsymbol{\varphi}, \boldsymbol{\varphi}_d)$) to reconstruct both f and its derivative.
- In all cases, the sampling functions result in a near-optimal asymptotic constant $C_{\boldsymbol{\varphi},1}^{\tilde{\boldsymbol{\varphi}}}$ for the reconstruction of f. The minor discrepancy with respect to optimality is a factor of $\sqrt{10/3} \approx 1.83$, which is expected from the fact that the sampling functions are not dual functions. It is also remarkable that the reconstruction of the derivative, while not being associated to the dual functions, is associated to an optimal approximation constant $C_{\boldsymbol{\varphi},2}^{\tilde{\boldsymbol{\varphi}}}$.
- Dual functions, although offering the smallest error of approximation, have strong practical disadvantages. First, the $\boldsymbol{\varphi}_d$, are non-compactly supported splines. More importantly, they cannot be easily implemented as they do not have a digital filter equivalent (*i.e.*, computation of the $\langle f, \boldsymbol{\varphi}_d \rangle$ depend on more than just knowing f and its derivative on a fixed grid, in contrast to usual cubic B-splines and Hermite splines). Finally, dual functions do not possess a closed-form expression in general. For these reasons, the sampling functions $\tilde{\boldsymbol{\varphi}}$ classically used in the three considered approximation schemes, although non-optimal, are preferable in practice.

In Table 1, we sum up these findings, including the asymptotic constants for cubic B-spline, interlaced derivative sampling, and cubic Hermite splines. We also provide their comparison with optimal constants. Hermite interpolation is thus not a unique way of approximating a function and its first derivative, even if one wishes the error to remain close to optimal. The notable difference lies in the fact that the Hermite scheme provides functions that are simultaneously of finite support, which is not the case for interlaced derivative sampling (see (66) and (67)), and interpolating, which is not the case for cubic B-splines.

4. Concluding remarks

Our work focused on the formal investigation of two practical aspects of Hermite splines, namely, their short support and their good approximation properties. We show that Hermite splines are of minimal support among pairs of functions with similar reproduction properties and provide a framework to quantify their power of approximation. These results not only allow us to prove that Hermite splines are asymptomatically identical to cubic B-splines, but also offer a general framework for the quantitative approximation of functions and their derivatives.

In summary, Hermite splines are found to offer an approximation scheme that (1) has the same approximation power than the notorious cubic B-splines, (2) is interpolating (possibly with the derivative), (3) is based on maximally localized compactly supported basis functions. The resulting cost is the need to use two basis functions instead of a single one, and the need to have access to derivative samples.

[1] We relied on the technical computing software Mathematica 11 for this task.

Table 1
Comparison of approximation methods.

Approximation method	Cubic B-splines	Interlaced derivative sampling	Hermite splines
Digital-filter implementation	✓	✓	✓
Interpolating	✗	✓	✓
Finite support	✓	✗	✓
Closed-form expression	✓	✗	✓
Rate of decay (L)	4	4	4
Asymptotic constant ($C_{\varphi,1}^{\tilde{\varphi}}$)	$\frac{1}{72\sqrt{70}}$	$\frac{1}{72\sqrt{70}}$	$\frac{1}{72\sqrt{70}}$
Ratio to optimal ($C_{\varphi,1}^{\tilde{\varphi}}$)	$\frac{\sqrt{3}}{\sqrt{10}}$	$\frac{\sqrt{3}}{\sqrt{10}}$	$\frac{\sqrt{3}}{\sqrt{10}}$
Rate of decay ($L-1$)	3	3	3
Asymptotic constant ($C_{\varphi,2}^{\tilde{\varphi}}$)	$\frac{1}{12\sqrt{210}}$	$\frac{1}{12\sqrt{210}}$	$\frac{1}{12\sqrt{210}}$
Ratio to optimal ($C_{\varphi,2}^{\tilde{\varphi}}$)	1	1	1

Acknowledgments

This work was supported by core funding from the European Molecular Biology Laboratory (EMBL), Germany, the Swiss National Science Foundation, Switzerland under Grants 200020_184646/1, and P2ELP2_181759, and the European Research Council under Grant H2020-ERC (ERC grant agreement No 692726 - GlobalBioIm). The authors thank Yoann Pradat for useful remarks on the manuscript.

References

[1] I. Schoenberg, Cardinal Spline Interpolation, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1973.
[2] P. Lipow, I. Schoenberg, Cardinal interpolation and spline functions. III. Cardinal Hermite interpolation, Linear Algebra Appl. 6 (1973) 273–304.
[3] I. Schoenberg, A. Sharma, Cardinal interpolation and spline functions. V. The B-splines for cardinal Hermite interpolation, Linear Algebra Appl. 7 (1) (1973) 1–42.
[4] M. Unser, A. Aldroubi, M. Eden, B-Spline signal processing: Part I—Theory, IEEE Trans. Signal Process. 41 (2) (1993) 821–833.
[5] M. Unser, A. Aldroubi, M. Eden, B-Spline signal processing: Part II—Efficient design and applications, IEEE Trans. Signal Process. 41 (2) (1993) 834–848.
[6] C. De Boor, A Practical Guide to Splines, Springer-Verlag, New York, NY, USA, 1978.
[7] R. Farouki, The Bernstein polynomial basis: A centennial retrospective, Comput. Aided Geom. Design 29 (6) (2012) 379–419.
[8] H. Prautzsch, W. Boehm, M. Paluszny, Bézier and B-Spline Techniques, Springer-Verlag, Berlin, Germany, 2013.
[9] G. Farin, Curves and Surfaces for CAGD: A Practical Guide, Morgan Kaufmann Publishers, Burlington, MA, USA, 2002.
[10] W. Böhm, G. Farin, J. Kahmann, A survey of curve and surface methods in CAGD, Comput. Aided Geom. Design 1 (1) (1984) 1–60.
[11] W. Dahmen, B. Han, R.-Q. Jia, A. Kunoth, Biorthogonal multiwavelets on the interval: Cubic Hermite splines, Constr. Approx. 16 (2) (2000) 221–259.
[12] R. Warming, R. Beam, Discrete multiresolution analysis using Hermite interpolation: Biorthogonal multiwavelets, SIAM J. Sci. Comput. 22 (4) (2000) 1269–1317.
[13] V. Uhlmann, J. Fageot, M. Unser, Hermite snakes with control of tangents, IEEE Trans. Image Process. 25 (6) (2016) 2803–2816.
[14] V. Uhlmann, J. Fageot, H. Gupta, M. Unser, Statistical optimality of Hermite splines, in: Proceedings of the Eleventh International Workshop on Sampling Theory and Applications (SampTA'15), Washington DC, USA, 2015, pp. 226–230.
[15] R. Delgado-Gonzalo, P. Thévenaz, M. Unser, Exponential splines and minimal-support bases for curve representation, Comput. Aided Geom. Design 29 (2) (2012) 109–128.
[16] T. Blu, P. Thévenaz, M. Unser, MOMS: Maximal-order interpolation of minimal support, IEEE Trans. Image Process. 10 (7) (2001) 1069–1080.
[17] C. Hall, On error bounds for spline interpolation, J. Approx. Theory 1 (2) (1968) 209–218.
[18] G. Birkhoff, A. Priver, Hermite interpolation errors for derivatives, Stud. Appl. Math. 46 (1–4) (1967) 440–447.
[19] T. Blu, M. Unser, Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors, IEEE Trans. Signal Process. 47 (10) (1999) 2783–2795.
[20] T. Blu, M. Unser, Approximation error for quasi-interpolators and (multi-) wavelet expansions, Appl. Comput. Harmon. Anal. 6 (2) (1999) 219–251.
[21] L. Condat, Reconstruction of derivatives: Error analysis and design criteria, in: Signal Processing Conference, 2011 19th European, IEEE, 2011, pp. 839–843.
[22] L. Condat, T. Möller, Quantitative error analysis for the reconstruction of derivatives, IEEE Trans. Signal Process. 59 (6) (2011) 2965–2969.
[23] C. de Boor, R. DeVore, A. Ron, Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$, Trans. Amer. Math. Soc. 341 (2) (1994) 787–806.
[24] C. de Boor, R. DeVore, A. Ron, The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^d)$, J. Funct. Anal. 119 (1) (1994) 37–78.
[25] C. de Boor, R. DeVore, A. Ron, Approximation orders of FSI spaces in $L_2(\mathbb{R}^d)$, Constr. Approx. 14 (4) (1998) 631–652.
[26] O. Holtz, A. Ron, Approximation orders of shift-invariant subspaces of $W_2^s(\mathbb{R}^d)$, arXiv preprint math/0512609, 2005.
[27] R.-Q. Jia, S.-T. Liu, Wavelet bases of Hermite cubic splines on the interval, Adv. Comput. Math. 25 (1–3) (2006) 23–39.
[28] T. Goodman, Interpolatory Hermite spline wavelets, J. Approx. Theory 78 (2) (1994) 174–189.
[29] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, Image coding using wavelet transform, IEEE Trans. Image Process. 1 (2) (1992) 205–220.
[30] V. Uhlmann, M. Unser, Tip-seeking active contours for bioimage segmentation, in: Proceedings of the Twelfth IEEE International Symposium on Biomedical Imaging: From Nano To Macro (ISBI'15), Brooklyn NY, USA, 2015, pp. 544–547.

[31] D. Migliozzi, M. Cornaglia, L. Mouchiroud, V. Uhlmann, M. Unser, J. Auwerx, M. Gijs, Multimodal imaging and high-throughput image-processing for drug screening on living organisms on-chip, J. Biomed. Opt. 24 (2) (2019) 021205–1–021205–9.
[32] A. Ron, Factorization theorems for univariate splines on regular grids, Israel J. Math. 70 (1) (1990) 48–68.
[33] T. Blu, P. Thévenaz, M. Unser, MOMS: Maximal-order interpolation of minimal support, IEEE Trans. Image Process. 10 (7) (2001) 1069–1080.
[34] K. Jetter, G. Plonka, A survey on L_2-approximation order from shift-invariant spaces, in: Multivariate Approximation and Applications, Cambridge University Press, 2001, pp. 73–111.
[35] G. Strang, G. Fix, A Fourier analysis of the finite element variational method, in: Constructive Aspect of Functional Analysis, Cremonese, Rome, Italy, 1971, pp. 796–830.
[36] C. De Boor, R. DeVore, Partitions of unity and approximation, Proc. Amer. Math. Soc. 93 (4) (1985) 705–709.
[37] M. Unser, Splines: A perfect fit for signal and image processing, IEEE Signal Process. Mag. 16 (6) (1999) 22–38.
[38] L. Schwartz, Théorie des Distributions, Hermann, 1966.
[39] A. Papoulis, Generalized sampling expansion, IEEE Trans. Circuits Syst. 24 (11) (1977) 652–654.
[40] A. Aldroubi, Oblique projections in atomic spaces, Proc. Amer. Math. Soc. 124 (7) (1996) 2051–2060.
[41] M. Unser, J. Zerubia, A generalized sampling theory without band-limiting constraints, IEEE Trans. Circuits Syst. II 45 (8) (1998) 959–969.
[42] M. Unser, J. Zerubia, Generalized sampling: Stability and performance analysis, IEEE Trans. Signal Process. 45 (12) (1997) 2941–2950.
[43] V. Uhlmann, Landmark active contours for bioimage analysis: A tale of points and curves, EPFL Thesis no. 7951, Swiss Federal Institute of Technology Lausanne (EPFL), 2017, p. 263.