INTERPOLATORY WAVELETS FOR MANIFOLD-VALUED DATA

PHILIPP GROHS AND JOHANNES WALLNER

Abstract. Geometric wavelet-like transforms for univariate and multivariate manifold-valued data can be constructed by means of nonlinear stationary subdivision rules which are intrinsic to the geometry under consideration. We show that in an appropriate vector bundle setting for a general class of interpolatory wavelet transforms, which applies to Riemannian geometry, Lie groups and other geometries, Hölder smoothness of functions is characterized by decay rates of their wavelet coefficients.

1. Introduction

A great part of work in the analysis of signals, images, and generally real-valued functions concerns the extraction of local information at different levels of resolution, and the conversion of continuous data to a countable collection of coefficients. *Wavelet transforms* are undoubtedly the most prominent concept in this area [4].

Topics relevant to wavelet-type transforms include: the computation of wavelet coefficients; the approximative computation of coefficients from discretely sampled data; re-synthesis of the original continuous data from the coefficient sequences; the effect which quantizing, thresholding, or otherwise perturbing coefficients has on the synthesis; and how properties like smoothness can be read off the coefficients.

The overwhelming majority of wavelet-type constructions are *linear* and their theory is formulated in terms of topological vector spaces and linear operators. It is a trivial point, which however is important for us, that for linear constructions there is often no difference between applying them to real-valued data and to vector-valued data (at least if one works with so-called *scalar subdivision schemes* as opposed to *vector subdivision schemes*, we refer the interested reader to [18, 21]). Things become different in the analysis of *geometric* data, where the structure of a vector space, even if employed for purposes of coordinate representation, is not natural. Functions which take values in surfaces, or Riemannian manifolds, or Lie groups, should be analyzed by *intrinsic* processes. This basically means invariance under appropriate transformation groups: e.g. it is natural to require that constructions applied to data living in a matrix group $G \leq \mathrm{GL}_n$ be invariant with respect to left translations $x \mapsto ax$ where $a \in G$. Likewise, constructions in geometries based on a metric should be invariant under isometries. Linear constructions for the purpose of analyzing vector-valued data occur only as a special case.

Tools common in multiresolution (wavelet) analysis such as spaces spanned by the translates of a refinable function can usually not easily be modified so as to apply to data which take values in more general geometries. Without function spaces, concepts like orthogonality and best approximation are difficult to formulate. The

2000 *Mathematics Subject Classification.* Primary 41AXX, Secondary 41A25, 53B, 22E.

present paper therefore restricts itself to the *interpolating wavelet transforms* introduced in [8, 15] which are computable from samples of a function. We recall their construction and their relation to stationary subdivision rules in Section 1.2. The idea to generalize them to manifold-valued data is not new, but has been proposed some years ago by D. Donoho [9] (see also [23]).

The present paper shows how an interpolating wavelet transform may be constructed for both univariate and multivariate manifold-valued functions in a way which unifies different kinds of geometries, and that this nonlinear construction retains essential properties of the analogous linear construction. In particular we show that smoothness of functions directly corresponds to the decay rate of coefficients.

We mention a few examples of geometries we are thinking of: the Euclidean motion group SE_n (pose data of rigid bodies), the Grassmann manifolds $\mathrm{G}_{n,k}$ (subspace-valued data), and the symmetric space of positive definite matrices Pos_n (multivariate data representing diffusion tensor images).

1.1. **Linear stationary subdivision rules.** We here recall properties of linear stationary subdivision rules [2]. Such a linear rule $\mathcal{S}$ maps real-valued or vector-valued data $p : \mathbb{Z}^s \to \mathbb{R}^n$ $(n \geq 1)$ to data $\mathcal{S}p : \mathbb{Z}^s \to \mathbb{R}^n$ according to

$$(\mathcal{S}p)_\alpha = \sum_{\beta \in \mathbb{Z}^s} a_{\alpha - N\beta} \cdot p_\beta. \tag{1}$$

This definition involves the *mask* $(a_\alpha)_{\alpha \in \mathbb{Z}^s}$ and the dilation factor N (typically, $N = 2$). We require a finite mask $(\#\{\alpha \mid a_\alpha \neq 0\} < \infty)$ and affine invariance of S:

$$\text{For all } \alpha, \quad \textstyle\sum_{\beta \in \mathbb{Z}^s} a_{\alpha - N\beta} = 1. \tag{2}$$

Data p formally defined as a function on the unit grid can be interpreted as samples of a function $\mathcal{F}_j p$ on the grid $N^{-j}\mathbb{Z}$. Vice versa, a function f may be sampled on a finer grid and converted into data $\mathcal{P}_j f$ formally defined on the unit grid. We let

$$(\mathcal{F}_j p)(N^{-j}\alpha) = p_\alpha, \quad (\mathcal{P}_j f)_\alpha = f(N^{-j}\alpha),$$

so $\mathcal{P}_j \mathcal{F}_j p = p$. A subdivision rule $\mathcal{S}$ is *interpolatory*, if the function $\mathcal{F}_1 \mathcal{S}p$ interpolates the original data (i.e., $p = \mathcal{F}_1 \mathcal{S}p|_{\mathbb{Z}^s}$). This is equivalent to $a|_{N\mathbb{Z}^s}$ being zero except for $a_0 = 1$, and it implies that $\mathcal{F}_i \mathcal{S}^i p = \mathcal{F}_j \mathcal{S}^j p|_{N^{-i}\mathbb{Z}^s}$ whenever $i \leq j$.

The sequence $\{\mathcal{F}_j \mathcal{S}^j p\}_{j \geq 0}$ of functions constructed by subdivision has the limit $\mathcal{S}^\infty p := \lim_{j \to \infty} \mathcal{F}_j \mathcal{S}^j p$, which is defined in a dense subset of $\mathbb{R}^s$. *Convergence* and C^k *smoothness* of a subdivision rule $\mathcal{S}$ means that for all input data p, $\mathcal{S}^\infty p$ is continuous and its unique continuous extension to $\mathbb{R}^s$ enjoys C^k smoothness.

We also consider Hölder smoothness of subdivision rules: With the notation $(\Delta_h f)(x) = f(x + h) - f$ we define the Hölder smoothness classes by

$$f \in \operatorname{Lip}\gamma \iff \|f\|_{\operatorname{Lip}\gamma} = \|f\|_\infty + \sup_{h \in \mathbb{R}^s \setminus 0} \big(h^{-\gamma} \|\Delta^{\lfloor \gamma + 1 \rfloor} f\|_\infty\big) < \infty \tag{3}$$

and we say that $\mathcal{S}$ has *critical Hölder regularity* r, if all $S^\infty p \in \operatorname{Lip}\gamma$ whenever $\gamma < r$.

A rule $\mathcal{S}$ has *polynomial reproduction* of degree d if for any polynomial $f \in \mathbb{R}[x_1, \ldots, x_s]$ of total degree $\leq d$ we have $\mathcal{S}f|_{\mathbb{Z}^s} = \mathcal{P}_1 f$, i.e., applying S to regular samples of f produces a denser sampling of the same f. C^k rules have $d \geq k$.

Example 1. Denote by $L^{t_\alpha,\dots,t_\beta}_{\alpha,\dots,\beta}$ the Lagrange interpolation polynomial which maps each subscript integer to the corresponding superscript. Fix $d>0$ and let

$$\mathcal{S}p_{N\alpha}=p_\alpha,\quad \mathcal{S}p_{N\alpha+\beta}=L^{p_{\alpha-d},\dots,p_{\alpha+d+1}}_{\alpha-d,\dots,\alpha+d+1}(\alpha+\tfrac{\beta}{N})\quad (\beta=1,\dots,N-1).$$

Then $\mathcal{S}$ is a subdivision rule with dilation factor N and polynomial reproduction degree $2d+1$ [6]. One can show that $\mathcal{S}$ has C^k limit functions, with $k\approx d\cdot\text{const}$.

1.2. **Linear interpolating wavelet transforms.** Introduced by [15, 8] for the univariate case and $N=2$, they are based on a "father wavelet" $\varphi:\mathbb{R}\to\mathbb{R}$ with $\varphi|_{\mathbb{Z}}=\delta$, where $(\delta_\alpha)_{\alpha\in\mathbb{Z}}$ is the Kronecker delta sequence (i.e., $\delta_0=1$ and $\delta_\alpha=0$ for $\alpha\neq 0$). The major example of [8] is that $\varphi=\mathcal{S}^\infty\delta$ for some interpolatory subdivision rule $\mathcal{S}$. In the following we consider the general multivariate case. Interpolatory wavelet-like constructions have been used in various places, e.g. [20, 16, 17, 5].

The interpolatory wavelet transform associated with an interpolatory subdivision rule $\mathcal{S}$ maps a function $f:\mathbb{R}^s\to\mathbb{R}^n$ to the coefficient collection $(u_\alpha)_{\alpha\in\mathbb{Z}^s}$, $(w^0_\alpha)_{\alpha\in\mathbb{Z}^s}$, $(w^1_\alpha)_{\alpha\in\mathbb{Z}^s},\dots$, which is defined by

$$u=\mathcal{P}_0f=f|_{\mathbb{Z}^s},\quad w^0=\mathcal{P}_1f-\mathcal{S}\mathcal{P}_0f,\quad w^1=\mathcal{P}_2f-\mathcal{S}\mathcal{P}_1f,\quad \dots$$

Smallness of w^j_β expresses agreement between $(j+1)$-st level samples $f|_{N^{-(j+1)}\mathbb{Z}^s}$ and values $(\mathcal{S}\mathcal{P}_jf)_\beta$ predicted from the j-th level samples $f|_{N^{-j}\mathbb{Z}^s}$. We recover f (actually, dense samples of f) by $\mathcal{P}_jf=w^{j-1}+\mathcal{S}w^{j-2}+\dots+\mathcal{S}^{j-1}w^0+\mathcal{S}^ju$.

The following result expresses the fact that smoothness of functions is characterized by decay rates of their wavelet coefficients. Both smoothness and coefficient decay is encoded by finiteness of certain norms. In the present paper we aim at similar results for the geometric (multivariate and nonlinear) case.

Theorem 2 ([8], Th. 2.7). *Assume that the interpolatory univariate subdivision rule $\mathcal{S}$ has polynomial reproduction degree $\geq d$, and that $\varphi=\mathcal{S}^\infty\delta$ has Hölder continuity $\geq r$. If $r,d>\sigma>\frac{1}{p}$ and $0<p,q\leq\infty$, the norm $\|(u,w^0,w^1,\dots)\|:=\|u\|_{\ell^p(\mathbb{Z})}+\|\omega\|_{\ell^q(\mathbb{Z}_0^+)}$, where $\omega_j=2^{j(\sigma-1/p)}\|w^j\|_{\ell^p(\mathbb{Z})}$, on the interpolating wavelet coefficients of a function f is equivalent to the norm of f in the Besov space $B^\sigma_{p,q}(\mathbb{R})$.*

1.3. **Subdivision rules and wavelet transforms in manifolds.** Geometric subdivision rules have been mostly analyzed with regard to smoothness (cf. [13, 24, 25, 26, 27, 28, 29] for the univariate case and [12, 14] for the multivariate case), but also with regard to approximation order [10]. Various definitions have been given.

A very general way to define subdivision in a manifold M relies on analogues of the operation 'point y minus point x' and its inverse 'point x plus vector' (the vector in question is supposed to lie in an appropriate vector space associated with x). We use the notation $v=y\ominus x$ and $y=x\oplus v$ for these mappings.

Example 3. In a Lie group M with Lie algebra $\mathfrak{g}$ we let $y\ominus x=\log(x^{-1}y)$, $x\oplus v=x\exp(v)$, where $v\in\mathfrak{g}$ and log is the inverse of $\exp:\mathfrak{g}\to M$ around $e\in G$. In a Riemannian manifold M we let $y\ominus x=\exp_x^{-1}(y)$, and $x\oplus v=\exp_x(v)$ where $v\in T_xM$, and $\exp_x$ is the Riemannian exponential mapping.

Equation (2) shows that we can express the subdivision rule $\mathcal{S}$ of (1) in terms of the operations $v=y-x$, $y=x+v$ for points x,y and vectors v:

$$(\mathcal{S}p)_{N\gamma+\alpha}=\sum_{\beta\in\mathbb{Z}^s}a_{N\gamma+\alpha-N\beta}\cdot p_\beta=p_\gamma+\sum_{\beta\in\mathbb{Z}^s}a_{N\gamma+\alpha-N\beta}(p_\beta-p_\gamma)$$

for all α, but especially $\alpha \in \{0, \dots, N-1\}^s$. This motivates the following definition:

Definition 4. Assume that $\pi : E \to M$ is a smooth vector bundle over the base manifold M ($\dim E < \infty$), and that $\oplus : E \to M$ and $\ominus : M \times M \to E$ are smooth and defined locally around M and the diagonal $\{(x, x)\} \subset M \times M$, respectively. With the notation $v \in E_x \stackrel{\oplus}{\longmapsto} x \oplus v$ and $(x, y) \stackrel{\ominus}{\longmapsto} x \ominus y$ we require that $y \ominus x \in E_x$, and $x \oplus (y \ominus x) = y$ whenever defined. Then the subdivision rule $\mathcal{T}$ given by

$$(\mathcal{T}p)_{N\gamma+\alpha} = p_\gamma \oplus \sum_{\beta \in \mathbb{Z}^s} a_{N\gamma+\alpha-N\beta}(p_\beta \ominus p_\gamma) \text{ where } \alpha \in \{0, \dots, N-1\}^s,$$

is called the *geometric analogue* of $\mathcal{S}$. It applies to data p where all instances of $\oplus$ and $\ominus$ which contribute to $\mathcal{T}p$ — terms with $a_{N\gamma+\alpha-N\beta} = 0$ do not — are defined.

In the Lie group case of Ex. 3, $E = M \times \mathfrak{g}$ and $\oplus$ is defined globally, while in the Riemannian case, $E = TM$ and $\oplus$ is defined globally for *complete* M. In both cases the domain of $\ominus$ depends on the respective exponential mappings. E.g. in Cartan-Hadamard manifolds, $\ominus$ is globally defined [7].

Example 5. Consider a surface $M \subset \mathbb{R}^n$ where $P : U \to M$ is a smooth retraction onto M (e.g. a matrix group M is considered as a surface in $\mathbb{R}^{m\times m}$ and P is the closest point projection w.r.t. the Frobenius norm). The subdivision rule $\mathcal{T}p := P \circ \mathcal{S}p$ operates on data $p : \mathbb{Z}^s \to M$ and is easily seen to be an instance of Definition 4: We let $E = M \times \mathbb{R}^n$, $y \ominus x = y - x \in \{x\} \times \mathbb{R}^n$, and $x \oplus v := P(x + v)$.

Having transferred subdivision to the manifold setting, we now define:

Definition 6. The wavelet transform with respect to the interpolatory subdivision rule $\mathcal{T}$ for M-valued data maps a function $f : \mathbb{R}^s \to M$ to the coefficients

$$u = \mathcal{P}_0 f = f|_{\mathbb{Z}^s}, \quad w^0 = \mathcal{P}_1 f \ominus \mathcal{T}\mathcal{P}_0 f, \quad w^1 = \mathcal{P}_2 f \ominus \mathcal{T}\mathcal{P}_1 f, \quad \dots$$

Here we let $(p_\alpha)_{\alpha\in\mathbb{Z}^s} \ominus (q_\beta)_{\beta\in\mathbb{Z}^s} = (p_\alpha \ominus q_\alpha)_{\alpha\in\mathbb{Z}^s}$. Each of $w^0, w^1, \dots$ represents bundle-valued data. Note that for topological reasons (worm holes in M) there might be no function f for given u, $\{w^j\}_{j\geq 0}$, not even if all $w^j_\beta = 0$.

Remark 7. Example 3 is due to [9, 23], and Example 5 is considered also in [29]. In the Lie group case, $\mathcal{T}^\infty p \in \operatorname{Lip}\gamma$ if the Hölder regularity of $\mathcal{S}$ exceeds γ and this limit exists (which it does for dense enough input data), as shown in [28, 14]. Analogous results for the univariate retraction case are given by [13].

2. Results

2.1. **Wavelet coefficient decay and smoothness.**
The 'usefulness' of Definition 6 is indicated by the fact that like in the linear case, the smoothness of a function can be read off its wavelet coefficients. The precise statements are as follows:

Theorem 8. *Let $\mathcal{S}$ be a linear interpolatory subdivision rule of Hölder smoothness r and polynomial reproduction degree d, and let $\mathcal{T}$ be its geometric analogue in the bundle $\pi : E \to M$. Assume that $f : \mathbb{R}^s \to M$ is continuous, and that $w^j : \mathbb{Z}^s \to E$ are the wavelet coefficients of the function $x \mapsto f(\sigma x)$ for some $\sigma > 0$ (whose local existence is guaranteed for small σ).*

If $f \in \operatorname{Lip}\alpha$ and $\alpha < d$, then $\|w^i_\beta\| \leq CN^{-\alpha i}$. Conversely, if $\|w^i_\beta\| \leq CN^{-\alpha i}$ and $\alpha < r$, then $f \in \operatorname{Lip}\alpha$. The constant C is understood to be uniform for data values in a compact set.

Here the symbols $\|w^i_\beta\|$ refer to a smooth bundle norm for E (e.g. the Riemannian metric in $E = TM$) the precise choice of which turns out to be irrelevant.

We break the proof of Theorem 8 into two steps: (i) Localization of the result and transfer to a trivial bundle over an open subset of $\mathbb{R}^m$ (see below); and (ii) Proof for the simplified setting (see Section 3.3).

We start our discussion with the *local nature* of the result. There is $\rho > 0$ such that the mask coefficient $a_\alpha = 0$ whenever α is outside the ball ρB of radius ρ. Consequently the wavelet coefficient w^j_β of the function $x \mapsto f(\sigma x)$ is determined by f's restriction to the ball $\sigma N^{-(j+1)}(\beta + \rho B)$. Smoothness of f (equivalently, smoothness of any $f(\sigma\,\cdot))$, is a local property. We may therefore, without loss of generality, restrict the analysis of smoothness of f, and of the wavelet coefficients of f, to arbitrarily small neighbourhoods.

In particular we assume that we work in the domain of a single bundle chart χ from E to the trivial bundle $\widetilde{\pi} : \widetilde{E} = U \times \mathbb{R}^n \to U$, where U is open in some $\mathbb{R}^m$. Each $\widetilde{E}$-fiber $\{\widetilde{x}\} \times \mathbb{R}^n$ is equipped with the χ-image $\|\cdot\|_{\widetilde{x}}$ of the original bundle metric, which smoothly depends on $\widetilde{x}$. Thus by making U smaller if necessary we can achieve that $\|\cdot\|_{\widetilde{x}}$ is uniformly equivalent to the standard metric in $\mathbb{R}^n$.

It is therefore sufficient to show Theorem 8 for the case that the bundle is $U \times \mathbb{R}^n$, each fiber being equipped with the canonical norm. For convenience, we still use the notation $\oplus$, $\ominus$, $\mathcal{T}$ for the χ-transforms of the respective entities.

2.2. **Proximity results.** The proof of Theorem 8 relies on the *proximity inequality* of Theorem 9 which assumes the viewpoint that $\mathcal{T}$ is a perturbation of $\mathcal{S}$ and which quantifies the distance of $\mathcal{S}$ from $\mathcal{T}$. Such proximity results are widely employed in the analysis of nonlinear subdivision schemes.

A result similar to Theorem 9 is contained in [14], which allows us to keep the proof short by referring to lemmas also found there. The part which is new in contrast to [14] is that Theorem 9 applies not only to nonlinear rules $\mathcal{T}$ defined in matrix groups via the matrix exponential function, but the much more general class of geometric analogues considered here. Nevertheless the algebraic part of the proof is very similar. Theorem 9 considers only subdivision rules in trivial bundles $U \times \mathbb{R}^m$, but in view of the previous section this is sufficient for our purposes. We make use of the following notation: Consider data $p : \mathbb{Z}^s \to \mathbb{R}^n$ and the canonical basis vectors $e_1, \ldots, e_s$ of $\mathbb{R}^s$. Let

$$(\Delta_i p)_\beta = p_{\beta+e_i} - p_\beta, \quad (\boldsymbol{\Delta} p)_\beta = (\Delta_1 p_\beta, \ldots, \Delta_s p_\beta) \in \mathbb{R}^{ns}.$$

Iterating this construction yields data $\boldsymbol{\Delta}^k p : \mathbb{Z}^s \to \mathbb{R}^{ns^k}$. Further, let $\|p\| := \sup_{\alpha\in\mathbb{Z}^s} \|p_\alpha\|_\infty$. With these preparations, we formulate:

Theorem 9. *Assume that $\mathcal{S}$ is a linear interpolatory rule with polynomial reproduction of degree k, and $\mathcal{T}$ is its geometric analogue in the bundle $U \times \mathbb{R}^m$ (U open in $\mathbb{R}^n$). For any compact $K \subset U$ there is $C > 0$ such that for K-valued data p,*

$$\|\mathcal{S}p - \mathcal{T}p\| \le C \sum_{\substack{i_1,\ldots,i_k \in \mathbb{Z}_0^+ \\ i_1 + 2i_2 + \cdots + k i_k = k+1}} \|\boldsymbol{\Delta} p\|^{i_1} \ldots \|\boldsymbol{\Delta}^k p\|^{i_k} \quad (k > 0). \tag{4}$$

For $k = 0$ we have the better estimate $\|\mathcal{S}p - \mathcal{T}p\| \le C\|\boldsymbol{\Delta} p\|^2$.

3. Proofs

3.1. **Proof of the proximity inequality.** Recall that the subdivision rule $\mathcal{T}$ reads

$$(\mathcal{T}p)_{N\gamma+\alpha} = p_\gamma \oplus \sum_{\beta\in\mathbb{Z}^s} a_{N\gamma+\alpha-N\beta}(p_\beta \ominus p_\gamma).$$

The relation $x \oplus (y \ominus x) = y$ implies that the linear rule $\mathcal{S}$ of (1) is expressible as $(\mathcal{S}p)_{N\gamma+\alpha} = \sum_{\beta\in\mathbb{Z}^s} a_{N\gamma+\alpha-N\beta}(p_\gamma \oplus (p_\beta \ominus p_\gamma))$. By introducing some auxiliary notation we can further rewrite $\mathcal{S}$, $\mathcal{T}$:

$$\Psi_p : \mathbb{R}^m \to U,\ v \mapsto p \oplus v, \quad v_\gamma^\beta := p_\beta \ominus p_\gamma \quad \Longrightarrow$$
$$\mathcal{T}p_{N\gamma+\alpha} = \Psi_{p_\gamma}\Big(\sum_{\beta\in\mathbb{Z}^s} a_{N\gamma+\alpha-N\beta} v_\gamma^\beta\Big),\ \mathcal{S}p_{N\gamma+\alpha} = \sum_{\beta\in\mathbb{Z}^s} a_{N\gamma+\alpha-N\beta}\Psi_{p_\gamma}(v_\gamma^\beta).$$

The following lemma concerning the Taylor expansion of $\mathcal{S}p - \mathcal{T}p$ is worded in terms of the r-linear mappings $d^r\Psi_x\big|_0$ which occur in the Taylor expansion $x \oplus v = x + d\Psi_x\big|_0(v) + \cdots + \frac{1}{k!}d^k\Psi_x\big|_0(v,\ldots,v) + \frac{1}{(k+1)!}d^{k+1}\Psi_x\big|_{\theta v}(v,\ldots,v)$, where $0<\theta<1$. We also introduce the right inverse $\Phi_x : y \mapsto y \ominus x$ of the function Ψ_x and consider its expansion $y \ominus x = \sum_{l=1}^k \frac{1}{l!}d^l\Phi_x\big|_x(y-x,\ldots,y-x) + O(\|y-x\|^{k+1})$.

Lemma 10. *The difference $\mathcal{T}p - \mathcal{S}p$ can be expanded around $\gamma\in\mathbb{Z}^s$ as*

$$(\mathcal{T}p)_{N\gamma+\alpha} - (\mathcal{S}p)_{N\gamma+\alpha} = \sum_{l=0}^k B_l + O(\|\mathbf{\Delta}p\|^{k+1}), \tag{5}$$

where $B_l = \frac{1}{l!}\sum_{\beta_1,\ldots,\beta_l\in\mathbb{Z}^s} A_{\beta_1,\ldots,\beta_l} d^l\Psi_{p_\gamma}\big|_0(v_\gamma^{\beta_1},\ldots,v_\gamma^{\beta_l})$, and the coefficients $A_{\ldots}$ are defined as $A_{\beta,\ldots,\beta} := (a_{N\gamma+\alpha-N\beta})^l - a_{N\gamma+\alpha-N\beta}$, if all indices are equal, and as $A_{\beta_1,\ldots,\beta_l} := a_{N\gamma+\alpha-N\beta_1}\ldots a_{N\gamma+\alpha-N\beta_l}$ otherwise.

Proof. The proof is the same as for the special case $x\oplus v = x\exp(v)$ in [14]: We expand $\mathcal{S}p - \mathcal{T}p$ and estimate the remainder term via $v_\gamma^\beta \approx d\Phi_{p_\gamma}(p_\beta - p_\gamma)$. ☐

By substituting $v_\gamma^\beta = \Phi_{p_\gamma}(p_\beta)$ in B_l, we express B_l in terms of the input data p:

$$B_l = \frac{1}{l!}\sum_{I\in\{1,\ldots,k\}^l}\frac{1}{I!}\sum_{\beta_1,\ldots,\beta_l\in\mathbb{Z}^s} A_{\beta_1,\ldots,\beta_l}\cdot F^I_{\beta_1,\ldots,\beta_l} + O(\|\mathbf{\Delta}p\|^{k+1}),$$

where $I = (i_1,\ldots,i_l)$, $I! = i_1!\cdots i_l!$, and the symbol $F^I_{\beta_1,\ldots,\beta_l}$ stands for

$$F^I_{\beta_1,\ldots,\beta_l} = C_I\big([p_{\beta_1}-p_\gamma]^{i_1\text{ times}},\ldots,[p_{\beta_l}-p_\gamma]^{i_l\text{ times}}\big),\ \text{where} \tag{6}$$
$$C_I(x_1,\ldots,x_{|I|}) = d^l\Psi_{p_\gamma}\big|_0(d^{i_1}\Phi_{p_\gamma}(x_1,\ldots,x_{i_1}),\ \ldots,\ d^{i_l}\Phi_{p_\gamma}(\ldots,x_{|I|})). \tag{7}$$

$C_I : (\mathbb{R}^n)^{|I|} \to \mathbb{R}^n$ is multilinear. Lemma 12 below, which gives bounds for B_l not in terms of $\mathbf{\Delta}p$ (which would be easy), but in terms of higher differences, needs

Lemma 11. *Assume that $(v_\tau)_{\tau\in\mathbb{Z}^s}$ are V-valued data, $B : V^r \to W$ is a multilinear mapping, and $\mathcal{A}(v) = \sum_{\tau_1,\ldots,\tau_r\in\mathbb{Z}^s} s_{\tau_1,\ldots,\tau_r}B(v_{\tau_1},\ldots,v_{\tau_r})$. With the notation $L(n_1,\ldots,n_r) = \{(l_i^j) \mid 1\le j\le r, 1\le i\le n_j, 1\le l_i^j\le s\}$, $\mathcal{A}(v)$ is expressible as*

$$\sum_{\substack{\tau_1,\ldots,\tau_r\in\mathbb{Z}^s\\ n_1+\cdots+n_r=k+1\\ n_j<k+1}}\ \sum_{l\in L(n_1,\ldots,n_r)} b^{(n_1,\ldots,n_r),l}_{\tau_1\ldots\tau_r} B\big(\Delta_{l_1^1}\cdots\Delta_{l^1_{n_1}}v_{\tau_1},\ \ldots,\ \Delta_{l_1^r}\cdots\Delta_{l^r_{n_r}}v_{\tau_r}\big),$$

if and only if all derivatives of order $\leq k$ *and all partial derivatives* $\frac{\partial^{k+1}}{\partial \mathbf{x}_{j_0}^\tau}$ *with* $j_0 \in \{1, \dots, r\}$, $\tau \in \mathbb{N}^s$ *and* $|\tau|_1 := \sum_{r=1}^s \tau_r = k+1$ *of the Laurent polynomial*

$$f_{\mathcal{A}}(\mathbf{x}_1, \dots, \mathbf{x}_r) = \sum_{\tau_1,\dots,\tau_r \in \mathbb{Z}^s} s_{\tau_1,\dots,\tau_r} \mathbf{x}_1^{\tau_1} \dots \mathbf{x}_r^{\tau_r}$$

vanish for $(\mathbf{x}_1, \dots, \mathbf{x}_r) = (1, \dots, 1) \in \mathbb{R}^{rs}$.

The special case that B is matrix multiplication and $V = \mathbb{R}^{n\times n}$ is [14, Lemma 1], whose proof carries over unchanged. We use the notation $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_s^{\alpha_s}$, $\alpha \in \mathbb{Z}^s$.

Lemma 12. *If* $\mathcal{S}$ *reproduces polynomials of degree* $\leq k$ *then, in the notation of Equation* (7)*, there exists a constant* $C = C(p_\gamma, I, \alpha) > 0$*, such that*

$$\Big\| \sum_{\beta_1,\dots,\beta_l \in \mathbb{Z}^s} A_{\beta_1,\dots,\beta_l} F^I_{\beta_1,\dots,\beta_l} \Big\| \leq C \sum_{n_1+\dots+n_k = k+1} \|\mathbf{\Delta}^{n_1} p\| \dots \|\mathbf{\Delta}^{n_k} p\|.$$

Proof. The left hand sum has the form of the expression "$\mathcal{A}(v)$" in Lemma 11, if we let $B = C_I$ and $s_{\tau_1,\dots,\tau_{|I|}} = 0$ zero except for $s_{\tau_1,\dots,\tau_{|I|}} = A_{\beta_1,\dots,\beta_l}$, if $(\tau_1, \dots, \tau_{|I|}) = ([\beta_1]^{i_1 \text{ times}}, \dots, [\beta_l]^{i_1 \text{ times}})$. Clearly the associated Laurent polynomial reads

$$f_{\mathcal{A}} = \Big(\sum_{\beta_1 \in \mathbb{Z}^s} a_{N\gamma+\alpha-N\beta_1} \mathbf{x}_1^{\beta_1} \cdots \mathbf{x}_{i_1}^{\beta_1} \Big) \cdots \Big(\sum_{\beta_l \in \mathbb{Z}^s} a_{N\gamma+\alpha-N\beta_l} \mathbf{x}_{|I|-i_l+1}^{\beta_l} \cdots \mathbf{x}_{|I|}^{\beta_l} \Big) - \sum_{\beta \in \mathbb{Z}^s} a_{N\gamma+\alpha-N\beta} \mathbf{x}_1^{\beta} \cdots \mathbf{x}_{|I|}^{\beta}.$$

If D is any differential operator, then $Df_{\mathcal{A}}(\mathbf{x}_1, \dots, \mathbf{x}_{|I|})|_{(1,\dots,1)}$ equals

$$\prod_{j=1}^l \Big(\sum_{\beta \in \mathbb{Z}^s} p_j(\beta) a_{N\gamma+\alpha-N\beta} \Big) - \sum_{\beta \in \mathbb{Z}^s} \prod_{j=1}^l p_j(\beta) a_{N\gamma+\alpha-N\beta},$$

where p_j are polynomials with $\deg \prod_{j=1}^l p_j = \deg(D)$. This expression has an interpretation in terms of samples and the subdivision rule $\mathcal{S}$ such that the polynomial reproduction property applies: If $\deg(D) \leq k$, $Df_{\mathcal{A}}(1, \dots, 1)$ can be expressed as

$$\prod_j (\mathcal{S} p_j|_{\mathbb{Z}^s})_{N\gamma+\alpha} - \big(\mathcal{S}\big(\prod_j p_j|_{\mathbb{Z}^s}\big)\big)_{N\gamma+\alpha} = \prod_j p_j(\gamma + \frac{\alpha}{N}) - \big(\prod_j p_j\big)\big|_{x=\gamma+\frac{\alpha}{N}} = 0.$$

If $D = \frac{\partial^{k+1}}{\partial \mathbf{x}_{j_0}^\tau}$, we have $\prod_j p_j = p_{j_0}$ and we can express $Df_{\mathcal{A}}(1, \dots, 1)$ as

$$\prod_j (\mathcal{S} p_j|_{\mathbb{Z}^s})_{N\gamma+\alpha} - \big(\mathcal{S}\big(\prod_j p_j|_{\mathbb{Z}^s}\big)\big)_{N\gamma+\alpha} = \big(\mathcal{S} p_{j_0}|_{\mathbb{Z}^s}\big)_{N\gamma+\alpha} - \big(\mathcal{S} p_{j_0}|_{\mathbb{Z}^s}\big)_{N\gamma+\alpha} = 0.$$

By Lemma 11 we can rewrite $\mathcal{A}(v)$ in terms of higher order differences $\mathbf{\Delta}^j p$. Taking norms yields the desired upper bound. □

We now complete the proof of Theorem 9. First, since we work in a compact set, we can make the constant $C(p_\gamma, I, \alpha)$ in Lemma 12 independent of p_γ. As there are only finitely many indices I and α, it is likewise independent of them. By substituting the upper bounds of Lemma 12 back into Lemma 10, we obtain

$$\|\mathcal{S}p - \mathcal{T}p\| \leq C \sum_{n_1+\dots+n_k=k+1,\ n_j<k+1} \|\mathbf{\Delta}^{n_1} p\| \dots \|\mathbf{\Delta}^{n_k} p\|$$

for some constant $C > 0$. Sorting the right hand terms by the exponents n_i yields the estimate required by Theorem 9 in the case $k \geq 1$. If $k = 0$, we observe that (2) causes the terms of orders $0, 1$ in the expansion (5) to vanish. Thus, $\|\mathcal{T}p - \mathcal{S}p\| = O(\|v_\gamma^\beta\|^2)$. Since $v_\gamma^\beta = d\Phi_{p_\gamma}(p_\beta - p_\gamma)$ of first order, the result follows. □

3.2. **Wavelet coefficient decay in the linear case.** The following theorem, which concerns linear schemes, has the flavor of a known result. We were however unable to locate a literature reference and therefore give a complete proof here. It is interesting that the implication (ii) $\Longrightarrow$ (i) can be shown using subdivision. Similar results and proofs can be found e.g. in [3].

Theorem 13. *Consider the set $\Omega = \bigcup_{j\geq 0} N^{-j}\mathbb{Z}^s$ of N-adic points, a function $f : \Omega \to \mathbb{R}^n$, and the wavelet coefficients u, $\{w^j\}_{j\geq 0}$ of f with respect to a fixed linear interpolatory rule $\mathcal{S}$. If $\mathcal{S}$ has $\operatorname{Lip}\gamma$ limit functions, the following are equivalent:*

(i) There exists $f' \in \operatorname{Lip}\alpha$ with $f'|_\Omega = f$;
(ii) f is bounded and there is an integer $k > \alpha$ with $\|\mathbf{\Delta}^k \mathcal{P}_j f\| = O(N^{-\alpha j})$;
(iii) $\|u\| < \infty$ and $\|w^j\| = O(N^{-j\alpha})$, provided $\alpha < \gamma$.

Proof. Without loss of generality we let $n = 1$. We first show (i) $\Longleftrightarrow$ (iii), using approximation methods and discrete interpolation spaces according to [1]. Let X consist of the uniformly continuous bounded functions $f : \mathbb{R}^s \to \mathbb{R}$, and let $Y = \operatorname{Lip}\gamma \subset X$. Define the approximation process $\mathcal{V}_j$ by letting $\mathcal{V}_j f(x) = \sum_{\beta\in\mathbb{Z}^s} f(\frac{\beta}{N^j})\phi(N^j x - \beta)$, with $\phi = \mathcal{S}^\infty\delta$ — in other notation, $\mathcal{V}_j f(x) = (\mathcal{S}^\infty \mathcal{P}_j f)(N^j x)$. It obeys $\lim_{j\to\infty} \mathcal{V}_j f = f$, if $f \in X$. As $\mathcal{S}$ is a convergent rule, the norms $\|\mathcal{V}_j\|$ w.r.t. $\|\cdot\|_\infty$ are bounded independently of j.

It is easy to show the Bernstein-type inequality $\|\mathcal{V}_j f\|_Y \leq C(N^j)^\gamma \|f\|_\infty$, as ϕ has compact support and for all $\lambda > 1$, $\|f(\lambda\,\cdot)\|_{\operatorname{Lip}\gamma} \leq C\lambda^\gamma \|f\|_{\operatorname{Lip}\gamma}$.

We also show the Jackson inequality $\|\mathcal{V}_j f - f\|_\infty \leq CN^{-j\gamma}\|f\|_Y$: Any $x \in \mathbb{R}^s$ has the form $x = h + y$ with $y \in N^{-j}\mathbb{Z}^s$ and $\|h\| < s^{\frac{1}{2}} N^{-j}$. Let g locally equal the Taylor polynomial of degree $\lceil\gamma\rceil - 1$ of f at y, so that $|f(x) - g(x)| \leq CN^{-j\gamma}\|f\|_Y$, with C independent of x, y. By polynomial reproduction, $\mathcal{V}_j g = g$, and

$$\|\mathcal{V}_j f - f\|_\infty \leq \|\mathcal{V}_j f - \mathcal{V}_j g\|_\infty + \|g - f\|_\infty \leq C(\|\mathcal{V}_j\|_\infty + 1)(N^{-j})^\gamma \|f\|_Y.$$

Now [1, Th. 3.3.1] implies the norm equivalence $[X, Y]^+_{\alpha,\infty,K} \cong X^J_{\alpha,\infty,\mathcal{V}}$ for $\alpha \in (0, \gamma)$, in the terminology of [1]. The former space, by interpolation, equals $\operatorname{Lip}\alpha$ [22], the latter equals $\{f \in X \mid \sup_{j\geq 0} N^{j\alpha}\|\mathcal{V}_j f - \mathcal{V}_{j-1} f\|_\infty < \infty\}$. By observing $w^j = (f - \mathcal{V}_{j-1} f)|_{N^{-j}\mathbb{Z}^s} = (\mathcal{V}_j f - \mathcal{V}_{j-1} f)|_{N^{-j}\mathbb{Z}^s}$ we obtain $\|w^j\| \leq \|\mathcal{V}_j f - \mathcal{V}_{j-1} f\|$, i.e., (iii) $\Longrightarrow$ (i). Conversely, $\mathcal{V}_{j-1} = \mathcal{V}_j \circ \mathcal{V}_{j-1}$ implies that $\|\mathcal{V}_j f - \mathcal{V}_{j-1} f\| = \|\mathcal{V}_j (f - \mathcal{V}_{j-1} f)\| \leq C\|\mathcal{V}_j\|\|w^j\| \leq C'\|w^j\|$, so (iii) implies $f \in X^J_{\alpha,\infty,\mathcal{V}} = \operatorname{Lip}\alpha$.

The implication (i) $\Longrightarrow$ (ii) follows e.g. from [19, Lemma 2]. For (ii) $\Longrightarrow$ (i), we employ an auxiliary interpolatory rule $\widetilde{\mathcal{S}}$ which has a k-th derived scheme $\widetilde{\mathcal{S}}^{[k]}$ obeying $N^k \mathbf{\Delta}^k \widetilde{\mathcal{S}} = \widetilde{\mathcal{S}}^{[k]} \mathbf{\Delta}^k$ (cf. [11]) and with C^k limit functions (take e.g. tensor products of the rules of Examples 1). Assuming $\|\mathbf{\Delta}^k \mathcal{P}_j f\| \leq CN^{-\alpha j}$, we estimate the interpolatory wavelet coefficients $\widetilde{w}^j$ of f with respect to $\widetilde{\mathcal{S}}$:

$$\begin{aligned}\|\mathbf{\Delta}^k \widetilde{w}^j\| &= \|\mathbf{\Delta}^k(\widetilde{\mathcal{S}}\mathcal{P}_j - \mathcal{P}_{j+1})f\| \leq \|\mathbf{\Delta}^k \widetilde{\mathcal{S}}\mathcal{P}_j f\| + \|\mathbf{\Delta}^k \mathcal{P}_{j+1} f\| \\ &\leq N^{-k}\|\widetilde{\mathcal{S}}^{[k]}\|\|\mathbf{\Delta}^k \mathcal{P}_j f\| + \|\mathbf{\Delta}^k \mathcal{P}_{j+1} f\| = O(N^{-j\alpha}).\end{aligned}$$

Now $\widetilde{w}^j|_{N^{-j+1}\mathbb{Z}^s} = 0$ implies that $\widetilde{w}^j$ itself, not only its k-th differences, is bounded by $O(N^{-j\alpha})$. Applying (iii) $\Longrightarrow$ (i) for the rule $\widetilde{\mathcal{S}}$ completes the proof. □

Obviously (i) $\Longleftrightarrow$ (ii) does not have to do anything with subdivision a priori.

3.3. **Proof of Theorem 8.** Recall that we can restrict ourselves to the bundle $U \times \mathbb{R}^n$, with U open in $\mathbb{R}^m$, and the Euclidean metric in each fiber $\{x\} \times \mathbb{R}^n$. We further assume that we work on data which take values in a compact set K. By locality this is justified, as we can simply consider dense enough samples of f. The proof employs Lipschitz constants $C_1, C_2 > 0$ for the function $\ominus$:

$$C_1\|p \ominus q\| \leq \|p - q\| \leq C_2\|p \ominus q\|. \tag{8}$$

For the first implication of Theorem 8, we assume that $f \in \operatorname{Lip}\alpha$, $\alpha < d$ and observe

$$\|\mathcal{T}\mathcal{P}_j f - \mathcal{P}_{j+1} f\| \leq \|\mathcal{S}\mathcal{P}_j f - \mathcal{P}_{j+1} f\| + \|\mathcal{S}\mathcal{P}_j f - \mathcal{T}\mathcal{P}_j f\|. \tag{9}$$

Theorem 13.(iii) bounds the first term with $CN^{-j\alpha}$. We let $k = \lfloor\alpha\rfloor$, so that $k \leq d$. Theorem 13.(ii) shows that $\|\boldsymbol{\Delta}^l \mathcal{P}_j f\| \leq C_l N^{-(l-\varepsilon)j}$ for $l = 1, \ldots, k$ and any $\varepsilon > 0$. The second term in Equation (9) is estimated by Theorem 9, as follows:

$$\begin{aligned}\|\mathcal{S}\mathcal{P}_j f - \mathcal{T}\mathcal{P}_j f\| &\leq C \sum_{i_1+\cdots+ki_k=k+1} \|\boldsymbol{\Delta}\mathcal{P}_j f\|^{i_1} \cdots \|\boldsymbol{\Delta}^k \mathcal{P}_j f\|^{i_k} \\ &\leq C \sum_{i_1+\cdots+ki_k=k+1} (N^{-(1-\varepsilon)j})^{i_1} \cdots (N^{-(k-\varepsilon)j})^{i_k} \\ &= CN^{-(k+1)j-\varepsilon j}.\end{aligned}$$

These estimates for (9) together with (8) show

$$\|w^j\| = \|\mathcal{T}\mathcal{P}_j f \ominus \mathcal{P}_{j+1} f\| \leq C_1^{-1}\|\mathcal{T}\mathcal{P}_j f - \mathcal{P}_{j+1} f\| \leq C(N^{-\alpha j} + N^{-(k+1-\varepsilon)j}),$$

with $\varepsilon > 0$ arbitrary. This proves the desired decay rate as stated by Theorem 8.

For the proof of the converse statement of Theorem 8, we assume that wavelet coefficients u, w^j are given, samples $\mathcal{P}_j f$ for $j \geq 0$ are defined, and that coefficients decay according to $w^j \sim N^{-j\alpha}$. Part (i) below makes an additional contractivity assumption, which is justified in part (iii).

Part (i): $\alpha < 1$. For now we assume that $\mathcal{T}$ is contractive in the sense that

$$\|\boldsymbol{\Delta}\mathcal{T}p\| \leq \mu\|\boldsymbol{\Delta}p\| \quad (\mu < 1). \tag{10}$$

This allows us to recursively estimate

$$\begin{aligned}\|\boldsymbol{\Delta}\mathcal{P}_j f\| &\leq \|\boldsymbol{\Delta}(\mathcal{P}_j - \mathcal{T}\mathcal{P}_{j-1})f\| + \|\boldsymbol{\Delta}\mathcal{T}\mathcal{P}_{j-1} f\| \\ &\leq 2C_2\|\mathcal{P}_j f \ominus \mathcal{T}\mathcal{P}_{j-1} f\| + \|\boldsymbol{\Delta}\mathcal{T}\mathcal{P}_{j-1} f\| \\ &\leq CN^{-\alpha j} + \mu\|\boldsymbol{\Delta}\mathcal{P}_{j-1} f\| \leq \ldots\end{aligned}$$

which yields

$$\|\boldsymbol{\Delta}\mathcal{P}_j f\| \leq C \sum_{l=0}^{j} \mu^l N^{-\alpha(j-l)}. \tag{11}$$

If $\mu N^\alpha < 1$, then (11), as geometric series, is bounded by $C'N^{-\alpha j}$ and samples $\mathcal{P}_j f$ extend to $f \in \operatorname{Lip}\alpha$ by Theorem 13.(ii). If $\mu N^\alpha \geq 1$, we choose $\nu \in (\mu, 1)$ — this implies $N^{-\alpha}/\nu < 1$ — and gain an estimate by (11) $= C\nu^j \sum_{l=0}^{j} (\mu/\nu)^l (N^{-\alpha}/\nu)^{j-l}$ $\leq C\nu^j \sum_{l=0}^{j} (\mu/\nu)^l \leq C\frac{\nu^j}{1-\mu/\nu}$. With $\nu = N^{-\delta}$ we have obtained $\|\boldsymbol{\Delta}\mathcal{P}_j f\| \leq CN^{-j\delta}$, showing that samples $\mathcal{P}_j f$ extend to $f \in \operatorname{Lip}\delta$. We increase δ by the following 'bootstrapping' argument, which invokes Theorem 9 for $k = 0$:

$$\begin{aligned} \|\mathcal{S}\mathcal{P}_j f - \mathcal{P}_{j+1} f\| &\le \|\mathcal{T}\mathcal{P}_j f - \mathcal{P}_{j+1} f\| + \|\mathcal{T}\mathcal{P}_j f - \mathcal{S}\mathcal{P}_j f\| \\ &\le C_2 \|\mathcal{T}\mathcal{P}_j f \ominus \mathcal{P}_{j+1} f\| + C' \|\mathbf{\Delta}\mathcal{P}_j f\|^2 \\ &\le C N^{-\alpha j} + C'' N^{-2\delta j} \le C''' N^{-\min(\alpha, 2\delta) j}. \end{aligned}$$

Thus $f \in \operatorname{Lip}\min(\alpha, 2\delta)$. By iteration, we obtain $f \in \operatorname{Lip}\alpha$.

Part (ii): $\alpha \ge 1$. Here we use induction. If for an integer $k > 0$ we already know $f \in \operatorname{Lip}(k-\varepsilon)$ for all $\varepsilon > 0$, we show $f \in \operatorname{Lip}\gamma$ for all $\gamma \in [k, k+1)$, provided $\gamma \le \alpha$. As part (i) above serves as an induction base ($k = 1$), this proves $f \in \operatorname{Lip}\alpha$.

We employ as an auxiliary device the wavelet coefficients $\widetilde{w}^j = \mathcal{S}\mathcal{P}_j f - \mathcal{P}_{j+1} f$ with respect to the linear rule $\mathcal{S}$. $\mathcal{S}$ reproduces polynomials of degree k (because $k \le \gamma \le \alpha < r$). We invoke Theorem 9 to estimate the coefficients $\widetilde{w}^j$:

$$\begin{aligned} \|\widetilde{w}^j\| &\le \|\mathcal{S}\mathcal{P}_j f - \mathcal{T}\mathcal{P}_j f\| + \|\mathcal{T}\mathcal{P}_j f - \mathcal{P}_{j+1} f\| \le ('') + C_2 \|\mathcal{T}\mathcal{P}_j f \ominus \mathcal{P}_{j+1} f\| \\ &\le C\Big(\sum_{i_1 + 2i_2 + \cdots + k i_r = k+1} \|\mathbf{\Delta}\mathcal{P}_j f\|^{i_1} \cdots \|\mathbf{\Delta}^k \mathcal{P}_j f\|^{i_k} + N^{-\alpha j}\Big). \end{aligned}$$

By Theorem 13, $f \in \operatorname{Lip}(k-\varepsilon)$ implies $\|\mathbf{\Delta}^l \mathcal{P}_j f\| \le C_l N^{-(l-\varepsilon)j}$ for $l = 1, \dots, k$, so

$$\|\widetilde{w}^j\| \le C\big(N^{-(k+1-\varepsilon)j} + N^{-\alpha j}\big) \le C N^{-\gamma j}, \quad \text{with } C > 0.$$

Theorem 13.(iii) shows that $f \in \operatorname{Lip}\gamma$. By induction, $f \in \operatorname{Lip}\alpha$.

Part (iii). To complete the proof of Theorem 8, we have to justify (10): It is known (cf. [2]) that for some iterate $\mathcal{S}^m$ there is $\mu' < 1$ with $\|\mathbf{\Delta}\mathcal{S}^m p\| \le \mu' \|\mathbf{\Delta} p\|$. By [24, Lemma 3], the case $k = 1$ of Theorem 9 applies to $\mathcal{S}^m, \mathcal{T}^m$ (since it applies to $\mathcal{S}, \mathcal{T}$). Now [25, Th. 1] says that existence of μ' implies $\|\mathbf{\Delta}\mathcal{T}^m p\| \le \mu \|\mathbf{\Delta} p\|$ for some $\mu < 1$, for dense enough input data. Obviously samples $\mathcal{P}_j f$ are dense enough for j greater than some j_0.

We now estimate the wavelet coefficients of f with respect to the subdivision rule $\mathcal{T}^m$, which has dilation factor N^m. Locally $\mathcal{T}$ is Lipschitz continuous, so that $\|\mathcal{T}p - \mathcal{T}q\| \le D\|p - q\|$ (this follows from the construction of $\mathcal{T}$ from $\mathcal{S}$). Thus,

$$\begin{aligned} &\|\mathcal{T}^m \mathcal{P}_j \ominus \mathcal{P}_{j+m}\| \le C_1^{-1} \|\mathcal{T}^m \mathcal{P}_j - \mathcal{P}_{j+m}\| \\ &\le \sum_{l=0}^{m} \|\mathcal{T}^{m-l} \mathcal{P}_{j+l} - \mathcal{T}^{m-l-1} \mathcal{P}_{j+l+1}\| \le \sum_{l=0}^{m} D^{m-l-1} \|\mathcal{T}\mathcal{P}_{j+l} - \mathcal{P}_{j+l+1}\| \\ &\le C_2 \sum_{l=0}^{m} D^{m-l-1} \|\mathcal{T}\mathcal{P}_{j+l} \ominus \mathcal{P}_{j+l+1}\| \le C N^{-\alpha j} = C (N^m)^{-j\frac{\alpha}{m}} \end{aligned}$$

for some $C > 0$. Part (i) applied to $\mathcal{T}^m$ yields $f \in \operatorname{Lip}\delta$, with $\delta = \frac{\alpha}{m}$, and so $\|\mathbf{\Delta}\mathcal{P}_j f\| = O(N^{-\delta j})$. From here part (i) goes as above. □

3.4. **Remarks on the reconstruction process.** Theorem 8 assumes that wavelet data u, $\{w^j\}_{j \ge 0}$ come from a continuous function f. If we do not know this a priori, we must observe that the bundle-valued sequences w^j are not arbitrary: The reconstruction procedure $\mathcal{P}_j f := \mathcal{T}(\dots \mathcal{T}(\mathcal{T}u \oplus w^0) \oplus w^1 \dots) \oplus w^j$ is well defined if and only if $\pi \circ w^j = \mathcal{T}\mathcal{P}_j f$ for $j \ge 0$. However, if the fibers E_x are canonically isomorphic to a fixed vector space E_0 (as in the Lie group and retraction cases), we can view w^j as E_0-valued sequences, and the consistency condition is void.

It is clear that the proof of Theorem 8 applies to data u, $\{w^j\}$:

Corollary 14. *In the same setting as Theorem 8 assume that coefficients* $u : \mathbb{Z}^s \to M$ *and* $w^j : \mathbb{Z}^s \to E$ $(j = 0, 1, \dots)$ *are consistently chosen such that the reconstruction procedure is defined. If* $\|w^j_\beta\| \leq CN^{-\alpha j}$ *with* C *small enough, and* u *is dense enough, then the samples* $\mathcal{P}_j f$ *extend to a* $\operatorname{Lip}\alpha$ *function* f*.*

The rather unspecific statements on u being dense enough and C small enough cannot be avoided. This is because reconstruction of a function with vanishing wavelet coefficients leads to the limit function $\mathcal{T}^\infty u$, and there are examples where that limit does not exist. More specific statements are possible only for specific smaller classes of subdivision rules. We leave this problem, which appears to exhibit a big difference between the cases $s = 1$ and $s > 1$, as a topic for future research.

Further interesting problems related to our work include analysis of average-interpolating transformations [23], as well as the *Lipschitz stability* of the reconstruction procedure, which is intimately connected with the stability of the underlying subdivision scheme. Stability is the topic of a forthcoming paper.

Acknowledgments

This work is supported by grant P19780 of the Austrian Science Fund (FWF). We want to thank A. Weinmann for his valuable suggestions.

References

[1] P. Butzer, K. Scherer, Approximationsprozesse und Interpolationsmethoden, vol. 826 of BI-Hochschultaschenbücher, Bibliogr. Inst., 1968.

[2] A. S. Cavaretta, W. Dahmen, C. A. Micchelli, Stationary subdivision, AMS, 1991.

[3] W. Dahmen, Wavelet and multiscale methods for operator equations, Acta Numerica 6 (1997) 55–228.

[4] I. Daubechies, Ten Lectures on Wavelets, SIAM, 1992.

[5] I. Daubechies, O. Runborg, W. Sweldens, Normal multiresolution approximation of curves, Constr. Approx. 20 (2004) 399–463.

[6] G. Deslauriers, S. Dubuc, Symmetric iterative interpolation processes, Constr. Approx. 5 (1986) 49–68.

[7] M. P. do Carmo, Riemannian Geometry, Birkhäuser, 1992.

[8] D. L. Donoho, Interpolating wavelet transforms, Tech. rep., Statistics Dep., Stanford (1992).

[9] D. L. Donoho, Wavelet-type representation of Lie-valued data, talk at the IMI "Approximation and Computation" meeting, May 12–17, 2001, Charleston, South Carolina (2001).

[10] N. Dyn, P. Grohs, J. Wallner, Approximation order of interpolatory nonlinear subdivision schemes, J. Comp. Appl. Math., to appear.

[11] N. Dyn, D. Levin, Subdivision schemes in geometric modelling, Acta Numer. 11 (2002) 73–144.

[12] P. Grohs, Smoothness analysis of subdivision schemes on regular grids by proximity, SIAM J. Numer. Anal. 46 (2008) 2169–2182.

[13] P. Grohs, Smoothness equivalence properties of univariate subdivision schemes and their projection analogues, Num. Math., to appear.

[14] P. Grohs, Smoothness of multivariate interpolatory subdivision in Lie groups, IMA J. Numer. Anal., to appear.

[15] A. Harten, Multiresolution representation of data: A general framework, SIAM J. Numer. Anal. 33 (1996) 1205.

[16] A. Harten, S. Osher, Uniformly high-order accurate nonoscillatory schemes. I, SIAM J. Numer. Anal. (1987) 279–309.

[17] M. Holmstrom, Solving hyperbolic PDEs using interpolating wavelets, SIAM J. Sci. Comput. 21 (2) (2000) 405–420.

[18] R. Jia, S. Riemenschneider, D. Zhou, Vector subdivision schemes and multiple wavelets, Math. Comput. 67 (224) (1998) 1533–1564.
[19] H. Johnen, K. Scherer, On the equivalence of the K-functional and moduli of continuity and some applications, in: Constructive Theory of functions of several variables, vol. 571 of Lecture Notes Math., Springer, 1977, pp. 119–140.
[20] G. Kutyniok, T. Sauer, Adaptive directional subdivision schemes and shearlet multiresolution analysis, arXiv preprint 0710.2678v1 [math.NA] (2007).
[21] C. Micchelli, T. Sauer, On vector subdivision, Math. Z. 229 (4) (1998) 621–674.
[22] H. Triebel, Interpolation theory, function spaces, differential operators, North-Holland, 1978.
[23] I. Ur Rahman, I. Drori, V. C. Stodden, D. L. Donoho, P. Schröder, Multiscale representations for manifold-valued data, Multiscale Mod. Sim. 4 (2005) 1201–1232.
[24] J. Wallner, Smoothness analysis of subdivision schemes by proximity, Constr. Approx. 24 (2006) 289–318.
[25] J. Wallner, N. Dyn, Convergence and C^1 analysis of subdivision schemes on manifolds by proximity, Comput. Aided Geom. Des. 22 (2005) 593–622.
[26] J. Wallner, E. Nava Yazdani, P. Grohs, Smoothness properties of Lie group subdivision schemes, Multiscale Mod. Sim. 6 (2007) 493–505.
[27] G. Xie, T. P.-Y. Yu, Smoothness equivalence properties of manifold-valued data subdivision schemes based on the projection approach, SIAM J. Numer. Anal. 45 (2007) 1200–1225.
[28] G. Xie, T. P.-Y. Yu, Smoothness equivalence properties of interpolatory Lie group subdivision schemes, IMA J. Numer. Anal., to appear.
[29] G. Xie, T. P.-Y. Yu, Smoothness equivalence properties of general manifold-valued data subdivision schemes, Multiscale Mod. and Sim. 7 (2009) 1073–1100.

Institut für Geometrie, TU Graz, Kopernikusgasse 24, A 8010 Graz, Austria.