Approximation on partially ordered sets of regular grids

P.W. Hemker and C. Pflaum

Department of Numerical Mathematics

NM-R9611 Juny 31, 1996
Approximation on Partially Ordered Sets of Regular Grids

P.W. Hemker

CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

C. Pflaum
Institut für Angewandte Mathematik und Statistik
Universität Würzburg, Am Hubland, D-97074 Würzburg, BRD

Abstract

In this paper we analyse the approximation of functions on partially ordered sequences of regular grids. We start with the formulation of minimal requirements for useful grid transfer operators in such a partially ordered context, and we continue with the introduction of hierarchical decompositions and the identification of piecewise constant and piecewise linear approximations as special instances of the tensor product case.

In the second part of the paper we derive error estimates for approximation in different norms on more-dimensional dyadic sequences of regular and sparse grids. We give special attention to a convenient notation.


Keywords & Phrases: sparse grids, multigrid, grid transfer operators, error estimates

Note: The major part of this report will be published elsewhere.

1. INTRODUCTION AND NOTATION

1.1. Introduction

Recently, in the research on multigrid methods for problems in three dimensions more and more attention is paid to semi-coarsening \cite{5, 13, 15, 19} and sparse grid approaches \cite{2, 6, 8, 9, 11, 18, 14}. This can be understood if we notice that the classical multigrid approach, where a linear sequence of nested grids is used for the approximation on different grids, requires very strong relaxation techniques. The selection of a suitable relaxation is difficult because of the large number of possible choices, each with their particular advantages and disadvantages.

In semi-coarsening, where different coarser grids are introduced, each coarsened in a single direction, the role of the smoothing procedure is reduced, and simpler relaxation procedures can be applied \cite{12}. This makes it attractive to study partially ordered sets of grids, rather than sequentially ordered ones.

Another difficulty, that particularly arises when regular grids are used for the approximation of functions, is the curse of three dimensions: the number of cells increases cubically with each refinement in all directions. This results in enormous amounts of degrees of freedom in the approximation, and in very large systems of algebraic equations to solve. This difficulty can be removed to a large extent by adaptive refinement, i.e. by adding only those degrees of freedom that contribute significantly to the improvement of the accuracy. Of course, what degree of freedom adds to the higher accuracy depends on the choice of the set of basis-functions that span the approximating function space. If, on a regular rectangular grid, a hierarchical basis is chosen, for a sufficiently smooth function the degrees of freedom associated with a "sparse grid" are the optimal choice. The sparse grid can be seen as a combination of regular grids, each with a different cell aspect ratio.

In this way semi-coarsening multigrid and sparse grid approximations are much related and make an interesting match.
It appears that the relations between the approximations on the different grids in the partially ordered set are not always clear [4], and that the requirements for the prolongations and restrictions between the approximations on the different grids are often chosen in an ad hoc way. In this paper we study the approximation of functions on partially ordered sets of regular grids (on a grids of grids). In particular we are interested in the minimal requirements that are needed to introduce the necessary grid transfer operators. Analysing these requirements results naturally in the introduction of a hierarchical decomposition of the approximation on the grid of grids, and we are able to show how the usual approximation by piecewise constant and piecewise linear basis-functions appear as a special case of tensor product form. At that point we can also introduce an MRA (multi-resolution analysis) in more dimensions, for a partially ordered set of approximations.

In the next sections we concentrate on the piecewise constant and piecewise linear approximation. We define their construction in a systematic way and we derive error estimates for the approximations in different norms.

Studying the approximation on a grids of grids, it appeared that a simple and convenient notation was lacking and that the data structures that are used in practice to realise the related algorithms, are rather complicated. Therefore, in the treatment much attention is given to a convenient notation, that can be used in general for the description and analysis of algorithms on a grid of grids.

Sparse grids yield a way for obtaining an approximation with a high accuracy relative to the number of degrees of freedom (unknowns) used. This was first observed by Smoljak [16] for numerical integration and interpolation with trigonometric functions. A different approach of constructing sparse grids is presented in [18]. This approach uses hierarchical basis functions for interpolation with piecewise multilinear functions. Error estimates in different norms and with different assumptions are found in [2, 14, 18]. For obtaining optimal estimates, it is necessary to assume that suitable derivatives of the functions are bounded. In case of singularities these assumptions may not hold. Then, optimal estimates can be obtained on adaptive sparse grids, which can be constructed in a natural way with hierarchical basis functions [2]. High order finite elements on sparse grids were analysed for example in [17].

1.2. Notation

Let \( k \in \mathbb{Z}^d \) be a multi-integer in \( d \) dimensions, then \( k = (k_1, k_2, \ldots, k_d) \), with \( k_i \in \mathbb{Z} \) for \( i = 1,2, \ldots, d \). We define relational operators between multi-integers by

\[
    k < m \iff (k_1 < m_1 \text{ and } k_2 < m_2 \text{ and } \ldots \text{ and } k_d < m_d),
\]

and analogously we define \( k \leq m \), \( k > m \), \( k \geq m \) and \( k = m \). Further we define \( \max(m,n) = (\max(m_1,n_1), \max(m_2,n_2), \ldots, \max(m_d,n_d)) \), and \( \min(m,n) \) similarly. In a few instances we will use these operators with the same meaning for real vectors \( x = (x_1, \ldots, x_d) \in \mathbb{R}^d \).

With \( n = (n_1, \ldots, n_d) \in \mathbb{Z}^d \) we denote \( \|n\| = n_1 + \cdots + n_d \). We also use the notation \( o = (0, \ldots, 0) \in \mathbb{N}^d \); \( 2^n = (2^{n_1}, \ldots, 2^{n_d}) \); \( 2^n x = (2^{n_1}x_1, \ldots, 2^{n_d}x_d) \); \( n o m = \sum_{i=1}^{d} n_i m_i \), and \( \|n\| = n_1 + \cdots + n_d \). Further we introduce in \( \mathbb{Z}^d \) the unit vectors \( e_k, k = 1, \ldots, d \), as follows: \( e_1 = (1,0, \ldots, 0) \); \( e_2 = (0,1,0, \ldots, 0) \); \( e_3 = (0, \ldots, 0,1) \), and we use \( e = (1, \ldots, 1) \). Finally we define \( E = \{e_1, \ldots, e_d\} \).

Let either \( \Omega = \mathbb{R}^d \) be the \( d \)-dimensional Euclidean space, or let \( \Omega = (0,1)^d \subset \mathbb{R}^d \) be the \( d \)-dimensional open unit cube. With any multi-integer \( n \in \mathbb{Z}^d \) we associate a function space \( V_n \), e.g., the space of piecewise constant or piecewise linear (bi-linear, tri-linear, d-linear) functions on a uniform grid with mesh size \( h = (h_1, \ldots, h_d) = (2^{-n_1}, \ldots, 2^{-n_d}) \). These grids are uniformly spaced in each of the \( d \) coordinate directions, but possibly with a different mesh size in the different directions. The volume of these cells is denoted by \( \|h\| = 2^{-\|n\|} \). The functions in \( V_n \) all are constant or d-linear on each dyadic block or cell

\[
    \Omega_{n,k} = [k_1 2^{-n_1}, (k_1 + 1)2^{-n_1}] \times \cdots \times [k_d 2^{-n_d}, (k_d + 1)2^{-n_d}],
\]
and this family of cells forms the grid \( \Omega_n = \{ \Omega_{n,k} \mid \Omega_{n,k} \subset \Omega, k \in \mathbb{Z}^d \} \). The family of cell centers or cell nodes is denoted by \( \Omega^*_n = \{ z_{n,k} \mid z_{n,k} = (k + e/2)2^{-n} ; k \in \mathbb{Z}^d ; z_{n,k} \in \Omega \} \). Other grids are obtained by considering the cell vertices or vertex nodes of the cells in \( \Omega_n \) as a grid of points. We denote these grids by \( \Omega^*_n \).

Apparently, all grids are identified by a multi-integer \( n \); the number \( |n| \) is called the level of the grid \( n \). Notice that different from classical multigrid theory we make a clear distinction between the grid-identification \( n \) and the level number \( |n| \).

It is clear that all cells in the grids \( \Omega_n \) are nested in some way. Therefore, also piecewise polynomial spaces defined on these cells are nested. In the following sections we will formalise this.

We also use the following notation for partial derivatives, with \( n \in \mathbb{N}_0^d \),

\[
D^n = D^{n_1 \cdots n_d} = \left( \frac{\partial}{\partial x_1} \right)^{n_1} \cdots \left( \frac{\partial}{\partial x_d} \right)^{n_d}.
\]

For the Banach spaces of continuously differentiable functions we use, with \( n \in \mathbb{N}_0^d \), the notation

\[
C^n(\Omega) = \left\{ u \mid \max_{0 \leq m \leq n} \max_{x \in \Omega} |D^m u(x)| < \infty \right\},
\]

with norm \( \| u \|_{C^n} = \max_{0 \leq m \leq n} \max_{x \in \Omega} |D^m u(x)| \).

For \( l \in \mathbb{N}_0 \), we introduce the notation \( C^{n,l}(\Omega) = \bigcap_{|m|=l} C^{n+m}(\Omega) \). This is a generalisation which combines \( C^n(\Omega) = C^{n,0}(\Omega) \) with the usual space of \( l \times \) continuously differentiable functions \( C^l(\Omega) = C^{0,l}(\Omega) \). With a \( C^n(\Omega) \) and \( C^{n,l}(\Omega) \) we denote the corresponding subspaces with homogeneous boundary conditions.

For the Banach spaces of integrable functions, \( 1 \leq p \leq \infty \), we, similarly, use the notation

\[
W^p_n(\Omega) = \left\{ u \mid \sum_{0 \leq m \leq n} \int_{x \in \Omega} |D^m u(x)|^p < \infty \right\},
\]

and for the semi-norm and norm

\[
|u|_{W^p_n} = \left( \int_{x \in \Omega} |D^m u(x)|^p \right)^{1/p} \quad \text{and} \quad \| u \|_{W^p_n} = \sqrt[p]{\sum_{0 \leq m \leq n} |u|^p_{W^p_n}}.
\]

or, with \( 0 \leq k \leq d \),

\[
\| u \|_{W^p_{n,k}} = \left( \sum_{0 \leq m \leq |e| \leq n, |m|=k} \| D^{e+m} u \|_{L^p}^p \right)^{1/p}.
\]

For \( l \in \mathbb{N}_0 \), we write \( W^{n,l}_p(\Omega) = \bigcap_{|m|=l} W^{n+m}_p(\Omega) \) and we obtain the Sobolev space \( W^l_p(\Omega) = W^{0,l}_p(\Omega) \). Again, with a \( W^{l}_p(\Omega) \) and \( W^{n,0}_p(\Omega) \) we denote the corresponding subspaces with homogeneous boundary conditions. For \( p = \infty \) we use the standard modifications, and for \( W_2 \) we also write \( H \).

Thus, for the Hilbert spaces of square integrable functions we use the notation \( H^n(\Omega) = W^n_2(\Omega) \), and for the semi-norm and norm \( |u|_{H^n} = |u|_{W^n_{2,0}} \) and \( \| u \|_{H^n} = \| u \|_{W^n_{2,0}} \). For \( l \in \mathbb{N}_0 \), we write \( H^{n,l}(\Omega) = \bigcap_{|m|=l} H^{n+m}(\Omega) \) and we obtain the usual Sobolev space \( H^l(\Omega) = H^{0,l}(\Omega) \). Again, with a \( H^0_0(\Omega) \) and \( H^{0,0}_0(\Omega) \) we denote the corresponding subspaces with homogeneous boundary conditions.
2. Space decomposition

2.1. Nested restrictions and prolongations

Let $X$ be a Banach space; e.g., $X = C^0(\Omega)$, $X = L_p(\Omega)$ or $X = L_p^{\text{loc}}(\Omega)$, where $\Omega \subset \mathbb{R}^d$. Let $k \in \mathbb{Z}^d$ and let

$$R_k : X \to V_k$$

(2.1)

be a restriction, i.e., a linear surjection. Possibly $V_k \subset X$, but this is not necessary. We notice that for any such $R_k$, because of the surjection, there exists the right-inverse or reconstruction

$$P_k : V_k \to X,$$

(2.2)

such that

$$R_k P_k = I_k$$

(2.3)

is the identity operator on $V_k$. We notice that $P_k$ is an injection (and hence a prolongation) and $\text{Ran}(P_k) \subset X$, but $P_k$ is not uniquely determined by a given $R_k$. In this section we study properties of such sets of transfer operators $\{R_k\}_{k \in \mathbb{Z}^d}$ and $\{P_k\}_{k \in \mathbb{Z}^d}$.

It is a consequence of (2.3) that

$$\Pi_k = P_k R_k$$

is a projection

$$\Pi_k : X \to \text{Ran}(\Pi_k) = \text{Ran}(P_k) \subset X,$$

as is

$$I - \Pi_k : X \to \text{Ker}(\Pi_k) = \text{Ker}(R_k) \subset X,$$

and we observe that $X$ can be written as a direct sum $X = \text{Ran}(P_k) \oplus \text{Ker}(R_k)$.

**Definition 2.1**

A set $\{R_k\}_{k \in \mathbb{Z}^d}$ is called a nested set of restrictions (or NSR) iff

$$k \geq m \Rightarrow \text{Ker}(R_k) \subset \text{Ker}(R_m).$$

(2.4)

A set $\{P_k\}_{k \in \mathbb{Z}^d}$ is called a nested set of prolongations (or NSP) iff

$$k \geq m \Rightarrow \text{Ran}(P_k) \supset \text{Ran}(P_m).$$

(2.5)

**Remark:**

It is obvious that for an NSR $\{R_k\}$ a set of corresponding reconstructions is not necessarily an NSP. On the other hand, given an NSP, the corresponding set of restrictions is not necessarily an NSR. However, in some cases both the restrictions and their reconstructions may form nested sets. Then we say that the transfer operators are nested and $\{V_k\}$ forms a nested set of representations of functions in $X$.

**Theorem 2.2**

Let $\{R_k\}_{k \in \mathbb{Z}^d}$ be a nested set of restrictions, then

$$\forall n \geq m \exists! R_{mn} : V_n \to V_m,$$

with the properties:

1. $R_{mn}$ is a restriction;
2. $R_{mn} R_n = R_m$;
3. $R_{mn} = R_m P_n$ (independent of the choice of $P_n!$).
Proof: (i) Define $R_m^1 = R_m P_n^1$ and $R_m^{2n} = R_m P_n^2$. Then we know that $R_n P_n^1 = I_n = R_n P_n^2$ and hence $\forall v_n \in V_n$ $R_n (P_n^1 - P_n^2) v_n = 0$. Because $\{R_k\}$ is an NSR and $m \leq n$ it follows that $R_m (P_n^1 - P_n^2) v_n = 0$ and hence $R_{mn} = R_m^{2n}$. So that there exists a unique $R_{mn}$. This means that we can write $R_{mn} = R_m P_n$, and $R_{mn}$ is independent of the choice of $P_n$.

(ii) $R_{mn} R_n = R_m P_n R_n = R_m P_n R_n$ on $\text{Ran}(P_n)$. Now, because $X = \text{Ran}(P_n) \oplus \text{Ker}(R_n)$ we may write $\forall v \in X : v = v_p + v_n$ so that $R_{mn} R_n v = R_{mn} R_n v_p + R_{mn} R_n v_n = R_m v_p + 0 = R_m v_p$. Further, because of $\text{Ker}(R_n) \subset \text{Ker}(R_m)$ we see $R_m v = R_m v_p + R_m v_n = R_m v_p + 0 = R_m v_p$, so that $R_{mn} R_n v = R_m v_p = R_m v \forall v \in X$ and hence $R_{mn} R_n = R_m$.

(iii) Because $R_m$ is a surjection, and by (2.6), $R_{mn}$ is necessarily a surjection. Of course, $R_{mn}$ is linear (trivial). Therefore $R_{mn}$ is a restriction. □

Given an NSR $\{R_m\}_{m \in \mathbb{Z}^d}$, and a set of corresponding reconstructions $\{P_m\}_{m \in \mathbb{Z}^d}$, we introduce, for $m \leq n$,

$$ P_{mn} = R_m P_n : V_m \rightarrow V_n. $$  

(2.7)

Notice that there are many possible choices of $P_m$ for a given $R_m$. Of course, some actual properties of $P_{mn}$ may depend on this choice!

**Lemma 2.3** $P_{nm}$ is a right inverse of $R_{mn}$:

$$ R_{mn} P_{nm} = I_m. $$  

(2.8)

**Proof:** $R_{mn} P_{nm} = R_m P_n R_n P_m = R_m P_m - R_m (I - P_n R_n) P_m = I_m + 0$, because $\text{Ker}(R_n) \subset \text{Ker}(R_m)$. □

**Corollary 2.4**

With $n \geq m$:

1. $P_{nm}$ is a prolongation (i.e. a linear injection);
2. $P_{nm} R_{mn}$ is a projection $V_n \rightarrow \text{Ran}(P_{nm}) \subset V_n$;
3. $I_n - P_{nm} R_{mn}$ is a projection $V_n \rightarrow \text{Ker}(R_{mn}) \subset V_n$;
4. $V_n = \text{Ran}(P_{nm}) \oplus \text{Ker}(R_{mn})$;
5. $P_{nm} : V_m \rightarrow \text{Ran}(P_{nm}) \subset V_n$ is a bijection.

**Lemma 2.5** Let $\{R_k\}_{k \in \mathbb{Z}^d}$ be an NSR and let $k \geq n \geq l$, then with a given set of corresponding reconstructions $\{P_k\}_{k \in \mathbb{Z}^d}$ we have

$$ P_{kl} = P_{kn} P_{nl}. $$

**Proof:** $P_{kn} P_{nl} = R_k P_n R_n P_l = R_k P_l - R_k (I - P_n R_n) P_l = R_k P_l - R_k (I - \Pi_l) P_l = R_k P_l - R_k P_l = 0$, because $\text{Ker}(R_k) \subset \text{Ker}(\Pi_l)$. Hence $P_{kl} = R_k P_l = P_{kn} P_{nl}$. □

**Lemma 2.6** Let $\{R_k\}_{k \in \mathbb{Z}^d}$ be an NSR with the corresponding $\{P_k\}_{k \in \mathbb{Z}^d}$ an NSP, then

$$ m \geq l \Rightarrow P_m P_{ml} = P_l. $$

**Proof:** $P_m P_{ml} = P_m R_m P_l = \Pi_m P_l = P_l$. The last equality holding because $\Pi_m u = u$ for all $u \in \text{Ran}(\Pi_m) = \text{Ran}(P_m) \supset \text{Ran}(P_l)$. □

5
Lemma 2.7: If \( \{R_k\}_{k \in \mathbb{Z}^d} \) is an NSR, then
\[
m \geq l \Rightarrow \Pi_l \Pi_m = \Pi_l.
\]
If, in addition, \( \{P_k\}_{k \in \mathbb{Z}^d} \) is an NSP, then also
\[
m \geq l \Rightarrow \Pi_m \Pi_l = \Pi_l.
\]

**Proof:** The first equality follows by \( \Pi_l \Pi_m = P_l P_l P_m R_m = P_l R_l R_m = P_l P_l = \Pi_l \), and the second equality by \( \Pi_m \Pi_l = P_m R_m P_l R_l = P_m R_m P_m P_m R_m = P_m R_m R_m = P_l R_l = \Pi_l \). □

If \( \{R_k\}_{k \in \mathbb{Z}^d} \) is an NSR, and if \( m \leq n \), then a bijection exists between \( \text{Ran}(P_m) \) and a subset of \( \text{Ran}(P_n) \). We denote this relation by \( \text{Ran}(P_m) \preceq \text{Ran}(P_n) \). I.e., a function that can be found in \( \text{Ran}(P_m) \), can uniquely be associated with a function in \( \text{Ran}(P_n) \). This follows because a bijection exists between \( \text{Ran}(P_n) \) and \( V_n \), and between \( \text{Ran}(P_m) \) and \( V_m \); and also a bijection exists between \( V_m \) and \( \text{Ran}(P_{nm}) \subset V_n \).

Hence, given a Banach space \( X \) with a nested set of restrictions \( \{R_k\}_{k \in \mathbb{Z}^d} \), a family of subspaces \( \text{Ran}(P_n) \) exists, with a partial ordering corresponding with the partial ordering of \( \{n\} \). This means that, although not necessarily \( \text{Ran}(P_m) \subset \text{Ran}(P_n) \), a partial ordering exists such that \( m \leq n \Leftrightarrow \text{Ran}(P_m) \preceq \text{Ran}(P_n) \Leftrightarrow P_{nm} R_m \cap \text{Ran}(P_m) = R_n \text{Ran}(P_m) \subset R_n \text{Ran}(P_n) \).

If \( \{P_k\}_{k \in \mathbb{Z}^d} \) is an NSP, this partial ordering simply reduces to \( \text{Ran}(P_m) \subset \text{Ran}(P_n) \). Then, in the case that \( V_n \subset X \) and we take \( P_n \) to be the natural injection, this means that \( \text{Ran}(P_n) \) can be identified with \( V_n \) and \( m \leq n \Leftrightarrow V_m \subset V_n \) and, thus, we find \( \{V_n\} \) to be a partially ordered family of subsets of \( X \).

**Definition 2.8**
Functions \( f_m \in V_m \) and \( f_n \in V_n \) are mutually coherent iff
\[
\exists f_k \in V_k \text{ with } k \geq m, k \geq n, \text{ such that } f_m = R_{mk} f_k \text{ and } f_n = R_{nk} f_k.
\]

**Theorem 2.9**

Let \( \{R_k\}_{k \in \mathbb{Z}^d} \) be an NSR and \( \{P_k\}_{k \in \mathbb{Z}^d} \) an NSP. Then, if \( f_m \) and \( f_n \) are mutually coherent, we have
\[
\forall l \in \mathbb{Z}^d \text{ with } l \leq m, l \leq n, \text{ we have } R_{lm} f_m = R_{ln} f_n. \tag{2.9}
\]
Moreover, under the additional condition that \( \Pi_{\min}(m,n) = \Pi_m \Pi_n \), also the reverse holds: it follows from (2.9) that \( f_m \) and \( f_n \) are mutually coherent.

**Proof:**  \( \Rightarrow \) First we assume that \( f_m \in V_m \) and \( f_n \in V_n \) are mutually coherent. Then \( \exists f_k \in V_k \) with \( k \geq m, k \geq n \) such that \( f_m = R_{mk} f_k \) and \( f_n = R_{nk} f_k \), and hence \( R_{lm} f_m = R_{lm} R_{mk} f_k = R_{lk} f_k = R_{ln} R_{nk} f_k = R_{ln} f_n \), which proves (2.9)

\( \Leftarrow \) Now we assume (2.9). Let \( k = \max(m,n) \) and take \( l = \min(m,n) \). We introduce \( f_k \) by:
\[
f_k = P_{mk} f_m + P_{kn} f_n - P_{kl} R_{lm} f_m = P_{mk} f_m + P_{kn} f_n - P_{kl} R_{ln} f_n.
\]
Then
\[
R_{mk} f_k = R_{mk} P_{km} f_m + R_{mk} P_{kn} f_n - R_{mk} P_{kl} R_{ln} f_n = R_{mk} f_m + R_{mk} P_{kn} (I_n - P_{nl} R_{ln}) f_n,
\]
Now, because \( k \geq m \) and \( k \geq n \geq l \), we know \( R_{mk} P_{km} = f_m \) (Lemma 2.3).

The additional condition and Lemma (2.7) show \( \Pi_m \Pi_n = \Pi_l = \Pi_m \Pi_l \Pi_n \), so that \( \Pi_m (I - \Pi_l) \Pi_n = 0 \). Hence,
\[
\begin{align*}
R_{mk} P_{kn} (I_n - P_n R_l) P_n &= R_m R_k R_l P_n (I_n - R_n P_l) P_n = \noindent R_m \Pi_m \Pi_k (I - \Pi_n \Pi_l) \Pi_n P_n = \noindent R_m \Pi_m (I - \Pi_l) \Pi_n P_n = 0.
\end{align*}
\]

Thus, we find \( R_{mk} f_k = f_m + 0 = f_m \). Analogously we prove \( R_n k f_k = f_n \). □

2.2. Commutative restrictions and projections

**Definition 2.10** A set \( \{ R_k \}_{k \in \mathbb{Z}^d} \) is called a commutative set of restrictions (or a CSR) iff for all \( m, n \in \mathbb{Z}^d \)
\[
\begin{align*}
\ker(R_n) \cap \ker(R_m) &\oplus \\
\ker(R_{\min(m,n)}) &= \ker(R_n) \cap \operatorname{ran}(P_m) \oplus \\
&\oplus \operatorname{ran}(P_n) \cap \ker(R_m).
\end{align*}
\]

A set \( \{ P_k \}_{k \in \mathbb{Z}^d} \) is called a commutative set of prolongations (or a CSP) iff for all \( m, n \in \mathbb{Z}^d \)
\[
\operatorname{ran}(P_{\min(m,n)}) = \operatorname{ran}(P_n) \cap \operatorname{ran}(P_m).
\]

If \( \{ R_k \}_{k \in \mathbb{Z}^d} \) is a CSR and \( \{ P_k \}_{k \in \mathbb{Z}^d} \) is a CSP, then we say that we have commutative transfer operators.

**Remark:**
It is immediate that each CSR is an NSR and each CSP is an NSP; which simply follows from the equivalence \( n \leq m \iff n = \min(n,m) \). On the other hand not necessarily every NSP is a CSP nor every NSR a CSR.

In the next theorem we show how the above definition of a CSR and an NSP lead to commutative projection operators indeed. To prove the theorem, we first derive a few lemmas.

**Lemma 2.11** If (2.12) holds, then \( \{ R_k \}_{k \in \mathbb{Z}^d} \) is a CSR.

**Proof:** Let \( m, n \in \mathbb{Z}^d \) be arbitrary, and let \( l = \min(m,n) \), then we know by assumption \( \Pi_m \Pi_n = \Pi_l = \Pi_l \Pi_m \). We define \( M = \text{Span}((\ker(R_n) \cap \ker(R_m)) \cup (\ker(R_n) \cap \operatorname{ran}(P_m)) \cup (\operatorname{ran}(P_n) \cap \ker(R_m))) \). To prove the lemma we show (i) \( M \subset \ker(\Pi_l) \), and (ii) \( \ker(\Pi_l) \subset M \), and (iii) \( M = (\ker(R_n) \cap \ker(R_m)) \oplus (\ker(R_n) \cap \operatorname{ran}(P_m)) \oplus (\operatorname{ran}(P_n) \cap \ker(R_m)) \).

To prove (i), let \( x \in M \), then \( x = x_m + x_n + x_l \) with \( x_n \in \ker(\Pi_l) \cap \operatorname{ran}(\Pi_m) \), \( x_m \in \ker(\Pi_l) \cap \ker(\Pi_m) \), then \( \Pi_l x = \Pi_n \Pi_m x_m + \Pi_m \Pi_n x_n + \Pi_m \Pi_n x_l = \Pi_n 0 + \Pi_m 0 + \Pi_m 0 = 0 \). So that \( x \in \ker(\Pi_l) \).

To prove (ii), let \( x \in \ker(\Pi_l) \) be arbitrary and define \( x_m := \Pi_m (I - \Pi_l) x = \Pi_m x \) and \( x_n := \Pi_n (I - \Pi_m) x = \Pi_n x \) and \( x_0 := (I - \Pi_m - \Pi_n) \Pi_l x = (I - \Pi_m - \Pi_n) x \). Then a simple calculation shows: \( x_0 \in \ker(R_n) \cap \ker(R_m) \) and \( x_m \in \ker(R_n) \cap \operatorname{ran}(P_m) \). Then \( x_0 \in \ker(R_n) \cap \ker(R_m) \) and \( x_m \in \operatorname{ran}(P_n) \cap \ker(R_m) \) and \( x_m + x_n = x_0 \).

To prove (iii), we show that \( M \) is a direct sum of the three spanning spaces. For this we have to prove: if \( x_0 \in \ker(R_n) \cap \ker(R_m) \) and \( x_m \in \ker(R_n) \cap \operatorname{ran}(P_m) \) and \( x_n \in \operatorname{ran}(P_n) \cap \ker(R_m) \) and \( x_0 + x_m + x_n = 0 \), then \( x_0 = x_m = x_n = 0 \). This is seen by \( 0 = \Pi_m (I - \Pi_n) (x_0 + x_m + x_n) = x_m \)
and $0 = \Pi_n (I - \Pi_m)(x_0 + x_m + x_n) = x_n$. This implies $x_0 = 0$. □

**Lemma 2.12** If (2.12) holds, then $\{P_k\}_{k \in \mathbb{Z}^d}$ is a NSP.

**Proof:** Let $n < m$ then (2.12) implies $\Pi_m \Pi_n = \Pi_n$, and hence, for all $x \in X$ we have $\Pi_m \Pi_n x = \Pi_n x$. It follows that for all $x \in X$ holds $\Pi_n x \in \text{Ran}(\Pi_m)$. Therefore $\text{Ran}(\Pi_n) \subset \text{Ran}(\Pi_m)$ for all $m, n \in \mathbb{Z}^d$ with $n < m$. Hence $\{P_k\}_{k \in \mathbb{Z}^d}$ is a NSP. □

**Lemma 2.13** If $\{R_k\}_{k \in \mathbb{Z}^d}$ is a CSR and $\{P_k\}_{k \in \mathbb{Z}^d}$ is a NSP then (2.12) holds.

**Proof:** Let $l = \min(m, n)$ and let $x \in X$ be arbitrary. We know that $X = \text{Ran}(P_l) \oplus \ker(R_l) = \text{Ran}(P_l) \oplus (\ker(R_n) \cap \ker(R_m)) \oplus (\ker(R_n) \cap \ker(R_m)) \oplus \ker(R_l) \cap \text{Ran}(P_m))$ because $\{R_k\}$ is a CSR. Hence, we may split $x = x_l + x_o + x_m + x_n = \Pi_m \Pi_n x_l + \Pi_m \Pi_n x_o + \Pi_m \Pi_n x_m + \Pi_m \Pi_n x_n = \Pi_m \Pi_n x_l + \Pi_m 0 + \Pi_m x_m + \Pi_m 0 = \Pi_m \Pi_n x_l$. Because $\{P_k\}_{k \in \mathbb{Z}^d}$ is a NSP, we know that $x_l \in \text{Ran}(P_l) \subset \text{Ran}(P_n)$ and $x_l \in \text{Ran}(P_m) \subset \text{Ran}(P_m)$; hence $\Pi_m \Pi_n x_l = x_l$. We conclude that, for arbitrary $x \in X$ holds $\Pi_m \Pi_n x = x_l = \Pi_l x_l$; which proves the lemma. □

**Theorem 2.14** The two following statements are equivalent:

1. $\{R_k\}_{k \in \mathbb{Z}^d}$ is a CSR, and $\{P_k\}_{k \in \mathbb{Z}^d}$ is an NSP; and
2. $\Pi_m \Pi_n = \Pi_{\min(m, n)}$ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ (2.12)

**Proof:** The theorem is a direct combination of the three lemmas above. □

It is an immediate consequence of the theorem that operators $\Pi_n$ associated with commutative transfer operators $\{R_k\}_{k \in \mathbb{Z}^d}$ and $\{P_k\}_{k \in \mathbb{Z}^d}$ do commute:

$$
\Pi_m \Pi_n = \Pi_{\min(m, n)} = \Pi_n \Pi_m .
$$

Further, combination with Theorem 2.9 gives the following important result.

**Corollary 2.15**

Let $\{R_k\}_{k \in \mathbb{Z}^d}$ be a CSR and $\{P_k\}_{k \in \mathbb{Z}^d}$ an NSP, then $f_m$ and $f_n$ are mutually coherent if and only if

$$
\forall l \in \mathbb{Z}^d, \text{with } l \leq m, l \leq n, \text{ we have } R_{lm} f_m = R_{ln} f_n . \ \ \ \ \ \ \ \ \ (2.13)
$$

2.3. **The merging operator**

Now we have seen how information about a function $u \in X$ can be represented on $V_n$, and how the representations $R_n u$ are related for different $n \in \mathbb{Z}^d$. An important question is how these $R_n u$, given for a limited number of $n \in \mathbb{Z}^d$, can be used to restore the picture of the original function $u$ as complete as possible.

We start with the situation where information is available from two representations, viz. in $V_n, V_m$. Therefore we introduce the **merging operator** $\Pi_{mn}$, which selects for an $x \in X$ the information that can be represented by the combined representations in $V_n$ and $V_m$.

**Definition 2.16** The **merging operator** $\Pi_{mn}$ : $X \to X$ is defined by

$$
\Pi_{mn} = \Pi_m + \Pi_n - \Pi_{\min(m, n)} . \ \ \ \ \ \ \ \ \ (2.14)
$$
Lemma 2.17 Let \( \{R_n\} \) and \( \{P_n\} \) be a set of commutative transfer operators, then

1. \[ \Pi_{mn} = \Pi_{nm}. \]  
2. \[ \text{if } m \leq n \text{ then } \Pi_{mn} = \Pi_n. \]  
3. \[ \Pi_{mn} \text{ is a projection}, \]
   \[ \Pi_{mn} = \Pi_n \cap \Pi_m. \] 
4. \[ \text{Ran}(P_n) \cap \text{Ran}(P_m) \] 
   \[ \text{Ran}(\Pi_{mn}) = \text{Ran}(P_n) \cap \text{Ker}(R_m) \] 
   \[ \text{Ker}(R_n) \cap \text{Ran}(P_m) \]
5. \[ \text{Ker}(\Pi_{mn}) = \text{Ker}(R_n) \cap \text{Ker}(R_m) \]

**Proof:** The first two statements are trivial by the definition of \( \Pi_{mn} \). Now set \( l = \min(m, n) \).

- Because \( \Pi_m \Pi_n = \Pi_n \Pi_m = \Pi_{\min(m,n)} \) we have
  \[ \Pi_{mn} \Pi_{mn} = (\Pi_m + \Pi_n - \Pi_{\min(m,n)}) (\Pi_m + \Pi_n - \Pi_{\min(m,n)}) \]
  \[ = \Pi_m + \Pi_n - \Pi_{\min(m,n)} \]
  \[ = \Pi_m + \Pi_n - \Pi_{l} + \Pi_{l} = \Pi_{mn}. \]

Hence, \( \Pi_{mn} \) is a projection.

- Assume that \( x \in \text{Ran}(\Pi_{mn}) \) implies \( \exists z \in \text{Ran}(\Pi_{mn}) \) such that \( x = \Pi_{mn}z = \Pi_mz + \Pi_nz - \Pi_lz = (\Pi_m - \Pi_l)z + (\Pi_n - \Pi_l)z + \Pi_lz = z_m + z_n + z_l. \) It is clear that this is a decomposition according to the direct sum in (2.18):
  \[ z_l \in \text{Ran}(\Pi_l) \cap \text{Ran}(\Pi_m) \cap \text{Ran}(\Pi_n); z_m \in \text{Ran}(\Pi_m) \text{ and } z_n \in \text{Ran}(\Pi_n) \] because \( \Pi_n z_m = \Pi_n (\Pi_m - \Pi_l) = \Pi_{\min(m,n)} = \Pi_{mn} \).

- On the other hand, if \( z = z_l + z_n + z_m \) is a splitting according to the direct sum, then \( \Pi_lz = \Pi_l(z_l + z_n + z_m) = \Pi_lz_l + \Pi_l\Pi_m z_m + \Pi_l\Pi_n z_n = z_l + \Pi_lz_m + \Pi_lz_n = z_l \) and \( (\Pi_m - \Pi_l)z = (\Pi_m - \Pi_l)(z_l + z_n + z_m) = (\Pi_m - \Pi_l)z_l + (\Pi_m - \Pi_l)z_m + (\Pi_m - \Pi_l)z_n = z_l - z_l + \Pi_m (I - \Pi_l)z_n = \Pi_m z_n = z_m. \)

- Analogously \( (\Pi_n - \Pi_l)z = z_n = z_n = z_n = \Pi_n z_n. \) Thus \( z = z_l + z_n + z_m = \Pi_lz + (\Pi_m - \Pi_l)z + (\Pi_n - \Pi_l)z = \Pi_{mn}z \). Hence \( z \in \text{Ran}(\Pi_{mn}) \).

- Assume that \( 0 = z = z_l + z_n + z_m \) is a splitting as above, then it follows that \( 0 = z_l = z_n = z_m \) because \( 0 = \Pi_l(z) = (\Pi_l \Pi_n)(z_l + z_n + z_m) = z_l \) and \( 0 = (\Pi_m - \Pi_l)(z) = (I - \Pi_l)\Pi_m (z_l + z_n + z_m) = z_m \) and \( 0 = (\Pi_n - \Pi_l)(z) = (I - \Pi_m)\Pi_n (z_l + z_n + z_m) = z_n. \)

- \( \Pi_{mn} = P_m P_m R_m + P_n P_n R_n - P_l R_l, \) hence \( \text{Ker}(\Pi_{mn}) \cap \text{Ker}(R_l) \) is trivial.

- \[ \text{Ker}(\Pi_{mn}) \cap \text{Ker}(R_l) \cap \text{Ker}(R_m) \] is shown as follows. Let \( x \in \text{Ker}(\Pi_{mn}) \), then \( 0 = \Pi_{nm} x = (\Pi_n - \Pi_l) x + (\Pi_m - \Pi_l) x + \Pi_l x = x_n + x_m + x_l. \) This implies \( x_l \in \text{Ran}(\Pi_l) \) and \( x_n, x_m \in \text{Ker}(\Pi_l) \), so that \( x_l = 0 \) and \( x_n + x_m = 0. \) Further, from \( x_n = -x_m \in \text{Ran}(\Pi_m) \cap \text{Ran}(\Pi_n) = \text{Ran}(\Pi_{\min(m,n)}) = \text{Ran}(\Pi_l) \) it follows that \( x_m = x_n = 0 \), and hence \( \text{Ker}(\Pi_{mn}) = \text{Ker}(R_l) \cap \text{Ker}(R_m). \) \( \square \)
2.4. The hierarchical decomposition

Further in this section we assume commutative transfer operators.

**Definition 2.18** Let \( \{R_n\} \) and \( \{P_n\} \) be a commutative set of transfer operators, then we define \( H_{mn} : X \to \text{Ran}(P_{\max(m,n)}) \subset X \), the hierarchical surplus, relative to the grids \( m \) and \( n \), by

\[
H_{mn} = \Pi_{\max(m,n)} - \Pi_{mn}.
\]

**(Remarks):**
- The hierarchical surplus is a projection operator.
- The hierarchical surplus, \( H_{mn}u \), of a function \( u \) represents the amount of information in \( u \in V_{\max(m,n)} \) that cannot be represented on the \( \text{Span}(\text{Ran}(P_n), \text{Ran}(P_m)) \).
- We can also write

\[
H_{mn} = (\Pi_{\max(n,m)} - \Pi_{n}) \left( \Pi_{\max(n,m)} - \Pi_{m} \right).
\]

**Definition 2.19** For a fixed \( n_0 \in \{-\infty, \mathbb{Z}^d\} \), which indicates a coarsest grid, we define for arbitrary \( n \geq n_0, n \in \mathbb{Z}^d \), the operator \( Q_n : X \to \text{Ran}(P_n) \subset X \), by

\[
Q_n = \prod_{j=1}^{d} (\Pi_n - \Pi_{n-e_j}).
\]

We use the convention that \( Q_{n_0} = \Pi_{n_0} \). The operator \( Q_n \) is called the (direct) hierarchical surplus at grid \( n \).

If \( n_0 \in \mathbb{Z}^d \) we call \( n_0 \in \mathbb{Z}^d \) the coarsest grid, and, without loss of generality we may assume \( n_0 = 0 \). If \( n_0 = -\infty \) then no coarsest grid exists.

**Lemma 2.20** \( Q_n \) is a projection and \( Q_m Q_n = 0 \) for all \( m \neq n \).

**Proof:** First we show that \( Q_n \) is a projection. For simplicity of notation, we set \( C = n_0 \).

\[
Q_n Q_n = \prod_{j=1, n_j \neq c_j} (\Pi_n - \Pi_{n-e_j}) (\Pi_n - \Pi_{n-e_j})
= \prod_{j=1, n_j \neq c_j} (\Pi_n - \Pi_{n-e_j})(\Pi_n - \Pi_{n-e_j} + \Pi_{n-e_j} + \Pi_{n-e_j})
= \prod_{j=1, n_j \neq c_j} (\Pi_n - \Pi_{n-e_j} + \Pi_{n-e_j})
= \prod_{j=1, n_j \neq c_j} (\Pi_n - \Pi_{n-e_j}) = Q_n
\]

To show that \( Q_m Q_n = 0 \) for all \( m \neq n \), let \( m \neq n \). Without loss of generality we may assume \( n_i < m_i \) for some \( i \in \{1, \ldots, d\} \). We consider the case \( n_i \neq c_i \); the other cases are similar.

\[
Q_m Q_n = \prod_{i} (\Pi_m - \Pi_{m-e_i}) (\Pi_n - \Pi_{n-e_i})
= (\Pi_m - \Pi_{m-e_i}) (\Pi_n - \Pi_{n-e_i}) \prod_{j \neq i} \cdots
= (\Pi_m - \Pi_{m-e_i}, \Pi_{n-e_i}) \prod_{j \neq i} \cdots
= (\Pi(n_{i-1}, \ldots, n_i, \ldots, n_{d}) - \Pi(n_{i-1}, \ldots, n_i, \ldots, n_{d})) \prod_{j \neq i} \cdots
= \prod_{j \neq i} \cdots = 0
\]

The indices indicated by dots correspond with those of \( \min(n, m) \).
Remark:
Notice that, for $n > N_{\theta}$, the two-dimensional case the relation (2.21) reads
\[ Q_n u = \Pi_n u - \Pi_{n-e_1} u - \Pi_{n-e_2} u + \Pi_{n-e_1-e_2} u , \]  
(2.24)
where $e = (1,1)$, and in the one-dimensional case we have
\[ Q_n u = \Pi_n u - \Pi_{n-e} u . \]  
(2.25)

Corollary 2.21
From Definition 2.19 it is immediately obvious that the projection $\Pi_n$ can be decomposed as
\[ \Pi_n = \sum_{n_0 \leq m \leq n} Q_m , \]  
(2.26)
and, hence, $\text{Ran}(\Pi_n) = \text{Span}(\text{Ran}(Q_m))_{n_0 \leq m \leq n}$. Because of Lemma 2.20 we can write
\[ \text{Ran}(\Pi_n) = \bigoplus_{n_0 \leq m \leq n} \text{Ran}(Q_m) . \]  
(2.27)
If $V_n \subset X$ and $P_n$ is the natural injection, then we see that $\text{Ran}(P_n) = \text{Ran}(\Pi_n) = V_n$ and, defining the pre-wavelet space $W_n = \text{Ran}(Q_m)$, we find
\[ V_n = \bigoplus_{n_0 \leq m \leq n} W_n . \]  
(2.28)

2.5. The tensor product case
For any $i = 1, \ldots, d$, let $\Omega_i \subset \mathbb{R}$ and let $\Omega = \otimes^n_{i=1} \Omega_i \subset \mathbb{R}^d$ be their Cartesian product. Let $X_i(\Omega_i)$ be a function space on $\Omega_i$ with functions $u_{i,a}(x_i)$ so that $X_i(\Omega_i) = \text{Span}\{u_{i,a}\}_{a \in A_i}$. Let $A = \otimes_{i=1}^d A_i$ be the Cartesian product of index sets. Then the tensor product space $X(\Omega)$ is defined by
\[ X(\Omega) = \otimes_{i=1}^d X_i(\Omega_i) = \text{Span} \prod_{a \in A} u_{i,a} , \]  
(2.29)
It is well known, e.g., that for $X_i(\Omega_i) = C^0_0(\mathbb{R})$, the tensor product space $X(\Omega)$ is densely embedded both in $C^0(\mathbb{R}^d)$ and in $H^1(\mathbb{R}^d)$.
For each $i = 1, \ldots, d$, let $R_{i,n} = \otimes_{i=1}^d \Omega_i$. If $V_n(\Omega_i)$, with $R_{i,n} : X_i(\Omega_i) \rightarrow V_{i,n}(\Omega_i) \subset X_i(\Omega_i)$, then, similar to $X(\Omega)$, for each $n \in \mathbb{Z}^d$ we may define a tensor product space $V_n = \otimes_{i=1}^d V_{i,n}(\Omega_i)$.
If the elements of $V_n(\Omega_i)$ are all determined by values associated with $\Omega_{n+1}$ (or $\Omega_{n+1}$ or $\Omega_{n+1}$), then we find a bijection $V_n = V_n(\Omega) = V_n(\Omega_n)$ (or $\Omega_{n+1}$ or $\Omega_{n+1}$). This notation indicates that its elements are determined by their values on the Cartesian product space $\Omega_n = \otimes_{i=1}^d \Omega_i$ or $\Omega_{n+1}$ or $\Omega_{n+1}$).

Definition 2.22 We define the tensor product restriction $R_n : X(\Omega) \rightarrow V_n(\Omega)$ by its action on a typical basis function
\[ u_a(x) = \prod_{i=1}^d u_{i,a}(x_i) \mapsto R_n u_a(x) = \prod_{i=1}^d R_{i,n} u_{i,a}(x_i) . \]  
(2.30)
We also write $R_n = \otimes_{i=1}^d R_{i,n}$. Since $V_n(\Omega) \subset X(\Omega)$, we can take the natural injection $P_n$ as the corresponding reconstruction for $R_n$. This $P_n$ we call the tensor product prolongation.
Theorem 2.23 For each $i \in \{1, \ldots, d\}$ let $\Omega_i \subset \mathbb{R}$, and let \( \{V_{i,n}(\Omega_i)\}_{n \in \mathbb{Z}} \) form a nested sequence of subspaces of the function space $X(\Omega_i)$, with
\[
V_{i,p}(\Omega_i) \subset V_{i,q}(\Omega_i) \subset X(\Omega_i) \quad \text{for} \quad p \leq q,
\]
and let each sequence of (one-dimensional) operators \( \{R_{i,n}\}_{n \in \mathbb{Z}} \) form an NSR. If we take for the corresponding reconstructions \( \{P_{i,n}\}_{n \in \mathbb{Z}} \) the natural injection, then the tensor-product restrictions \( \{R_n\} \) and prolongations \( \{P_n\} \) form a CSR and a CSP respectively. (So, together they form a commutative set of transfer operators.)

**Proof:** By the nesting of the subspaces it is immediate that for each $i$ the set \( \{P_{i,n}\}_{n \in \mathbb{Z}} \) forms an NSP (Section 2.1). Further, using the fact that the prolongation is the natural injection, we can identify restrictions $R_i$ with the corresponding projections $\Pi_i$, and by Lemma 2.7 we see that for each $i$ we have $\Pi_{i,p} \Pi_{i,q} = \Pi_{i,q} \Pi_{i,p} = \Pi_{i,p}$ if $p \leq q$. Thus, for each $i$ it follows that the sequences \( \{P_{i,n}\}_{n \in \mathbb{Z}} \) and \( \{R_{i,n}\}_{n \in \mathbb{Z}} \) are CSP and SCR. Now we prove that $\{R_n\}$ is an CSR by showing $R_m R_n = R_{\min(m,n)}$ as follows.

Let $l = \min(m, n)$ and let $v \in X(\Omega)$ be arbitrary, then we can write
\[
v(x) = \sum_{a \in A} c_a u_a(x)
\]
and
\[
R_m v(x) = \sum_{a} c_a R_m u_a(x) = \sum_{a} c_a \prod_i R_{i,m} u_{i,a}(x_i).
\]
Similarly,
\[
R_n R_m v(x) = \sum_{a} c_a R_n \prod_i R_{i,n} u_{i,a}(x_i)
= \sum_{a} c_a \prod_i R_{i,n} u_{i,a}(x_i)
= \sum_{a} c_a \prod_i R_{i,\min(m,n)} u_{i,a}(x_i)
= \sum_{a} c_a R_{\min(m,n)} \prod_i u_{i,a}(x_i)
= R_{\min(m,n)} v(x).
\]
Hence $R_n R_m = R_{\min(m,n)}$, and, thus, $\Pi_n \Pi_m = \Pi_{\min(m,n)}$. Now Theorem (2.14) shows that $\{R_n\}$ is an CSR and $\{P_n\}$ is an NSP.

To prove that $\{P_n\}$ is an CSP, we have to show $V_{\min(m,n)} = V_m \cap V_n$, or
\[
\oplus_i V_{i,\min(m,n)} = \oplus_i V_{i,m_i} \cap \oplus_i V_{i,n_i}.
\]
As for each $i$ we know that \( \{V_{i,n}(\Omega_i)\}_{n \in \mathbb{Z}} \) is a nested sequence of subspaces of $X_i(\Omega_i)$, we can construct a sequentially ordered set of basis functions $B_i = \{u_{i,b}\}_b$ in $X_i(\Omega_i)$, such that $n_i < m_i$ implies $u_{i,b} \in V_{i,n_i} \Rightarrow u_{i,b} \in V_{i,m_i}$. It follows that we have $V_{i,m_i} = \text{Span}\{u_{i,b} \in B_i | u_{i,b} \in V_{i,m_i}\}$ and similarly
\[
V_{i,\min(n_i,m_i)} = \text{Span}\{u_{i,b} \in B_i | u_{i,b} \in V_{i,n_i} \cap V_{i,m_i}\}.
\]
So we see
\[
\text{Span} \prod_i u_{i,i} = \text{Span} \prod_i u_{i,m_i} \cap \text{Span} \prod_i u_{i,n_i},
\]
which is equivalent with (2.32). $\Box$

**Example 2.24 Piecewise constant approximation.**

If we consider $L^2_0(\Omega) = \bar{X}(\Omega)$ and we choose for $R_{i,n}$ the one-dimensional $L_2$-projection $R_{i,n}$:
\( X(\Omega_i) \to V_n(\Omega_i) \subset X(\Omega_i) \), where \( \Omega_i \subset \mathbb{R} \) and \( V_n(\Omega_i) \) is the space of piecewise constant functions on dyadic intervals (i.e., if \( R_{i,n} \) denotes: taking mean values over intervals \([j2^{-n},(j+1)2^{-n}]\) in the \( i \)-th coordinate direction), then, for each \( i \), the restrictions \( \{R_{i,n}\}_n \) form a one-dimensional NSR. The corresponding reconstructions \( P_{i,n} \) represent piecewise constant interpolation over dyadic intervals. This makes the prolongations \( \{P_{i,n}\}_n \) an NSP.

Then, as a consequence of the above theorem, \( \{R_k\} \) and \( \{P_k\} \) are commutative transfer operators, i.e., \( \{R_k\} \) is a CSR and \( \{P_k\} \) a CSP.

**Example 2.25 Piecewise linear approximation.**

If we select the restriction \( R_{i,n} : C^0(\Omega_i) \to V_n(\Omega_i) \subset \mathbb{R}^d \) to be: taking function values at dyadic points \( j2^{-n} \) in the interval \( \Omega_i \subset \mathbb{R} \), then \( \{R_{i,n}\}_n \) is an NSR. Corresponding reconstructions \( P_{i,n} \), defined by piecewise linear interpolation over dyadic intervals, make the prolongations \( \{P_{i,n}\}_n \) an NSP.

As a consequence of the previous theorem, with \( C^0(\Omega) = \overline{X(\Omega)} \), the tensor product operators \( \{R_k\} \) and \( \{P_k\} \), defined on \( X(\Omega) \), are commutative (are a CSR and a CSP respectively). The restriction \( R_k \) takes the function values at grid points \( \Omega^*_n \), and the prolongation \( P_k \) makes a multi-linear interpolation over cells in \( \Omega_n \).

In the above examples, with \( V_n \subset X \) we took for the reconstruction \( P_n \) the natural injection (the identity in \( X \)). In this way we may identify \( R_n \) and \( \Pi_n \). It appears that in both cases, i.e. for the piecewise constant and the piecewise linear approximation, we have a projection \( \Pi_n \) of the form

\[
\Pi_n = \prod_{j=1}^d \Pi_{n_j} e_j , = \prod_{j=1}^d R_{n_j} e_j .
\]

Here \( R_{n_j} e_j : X(\Omega) \to X(\Omega) \) is the operator on the tensor product space \( X(\Omega) \) such that

\[
R_{n_j} e_j u_\varphi(x) = R_{j,n_j} u_{j,a_j}(x) \cdot \prod_{i \neq j} u_{i,a_i}(x_i)
\]

In the following section we consider the case of nested subspaces \( \{V_n\} \) with \( V_n \subset X \) and \( X = \bigcup_n V_n \), and where all spaces \( V_n \) are spanned by dilations of a single function \( \phi(x) \), together with all its dyadic translations. This leads to the more-dimensional **multiresolution analysis** or MRA. In this case the spaces \( W_m = \text{Ran}(Q_m) \) correspond with more-dimensional wavelet spaces.

### 2.6. More-dimensional MRA and wavelets

It will be convenient if

(i) we can make an arbitrarily accurate approximation of any function \( u \in X \) by taking the multi-
integer \( n \) large enough. Moreover, it will be convenient

(ii) all spaces \( \{\text{Ran}(P_n)\} \) or \( \{V_n\} \) have a similar structure, and

(iii) there is a clear relation between the spaces in \( \{\text{Ran}(P_n)\} \) or \( \{V_n\} \).

In order to create such a structure, in this section we introduce the multidimensional multiresolution analysis. For this purpose we will restrict ourselves to Hilbert spaces. First we introduce the important notion of frame.

**Definition 2.26** A sequence \( \{x_n\} \) in a Hilbert space \( H \) is a frame if there exist numbers \( A, B > 0 \) such that for all \( x \in H \) we have

\[
A\|x\|^2 \leq \sum_n \|(x, x_n)\|^2 \leq B\|x\|^2 .
\] (2.33)
The numbers \(A, B\) are called \textit{frame bounds} . The frame is \textit{tight} if \(A = B\). The frame is \textit{exact} if it ceases to be a frame whenever any single element is deleted from the sequence. If the sequence \(\{x_n\}\) satisfies (only) the second part of the inequality (2.33) then the sequence is called a \textit{Bessel sequence}.

Having introduced the exact frame, we can define the partially ordered, more-dimensional multiresolution analysis. Notice that this is different from the more-dimensional multiresolution analysis introduced in [1], which considers a sequentially ordered nested set of approximating spaces.

\textbf{Definition 2.27} Let \(\Omega = \mathbb{R}^d\) and let \(X(\Omega)\) be a Hilbert space of functions defined on \(\Omega\). A \textit{multidimensional multiresolution analysis} of \(X(\Omega)\), is a \textbf{partially ordered} set of \textbf{closed} linear subspaces

\[
\{V_n | V_n \subset X(\Omega)\}_{n \in \mathbb{Z}^d}
\]

with the four properties:

\[
\bigcap_n V_n = \{0\}; \quad \bigcup_n V_n = X(\Omega),
\]

\[
f(x) \in V_n \iff f(2^m x) \in V_{n+m} \quad \forall n, m \in \mathbb{Z}^d,
\]

\[
f(x) \in V_n \iff f(x - 2^{-n} k) \in V_n \quad \forall n, k \in \mathbb{Z}^d,
\]

\[
\exists \phi \in \Phi_0 : \{\phi(x - k)\}_{k \in \mathbb{Z}^d} \text{ is an exact frame for } \Phi_0.
\]

The function \(\phi(x)\) in (2.34d) is called the \textit{father function} or the \textit{scaling function} of the multiresolution analysis.

For \(\Omega = \mathbb{R}^d\) the tensor product Examples 2.24 and 2.25 in Section 2.5 also yield examples of a multidimensional MRA.

For piecewise constant interpolation we take \(X(\Omega) = L^2(\mathbb{R}^d)\) as the starting point. The characteristic function on the unit cube (the more-dimensional Haar function) is the scaling function \(\phi\). The set \(\{V_n\}\) contains the spaces of piecewise constant functions on \(\Omega_n\), and a CSR is obtained by \(R_n : X(\Omega) \rightarrow V_n\), the \(L^2\)-projection. It is obvious that in this case the set \(\{\phi(x - k)\}\) is an orthonormal basis and hence an exact frame with bounds \(A = B = 1\).

For piecewise linear interpolation we take \(X(\Omega) = H^1(\mathbb{R}^d)\) as the Hilbert space. The set \(\{V_n\}\) contains the space of piecewise \(d\)-linear functions, determined by their nodal values at \(\Omega_n^+\). A CSR is obtained by \(R_n : X(\Omega) \rightarrow V_n\), the piecewise \(d\)-linear interpolation at \(\Omega_n^+\). Here, the \(d\)-linear finite-elem basis function is the scaling function \(\phi\). By Theorem ??? from [10] it is easily seen that in this case \(\{\phi(x - k)\}\) is an exact frame, with as frame bounds the extreme eigenvalues of the frame operator \(S : L^2(\Omega) \rightarrow L^2(\Omega)\) defined by \(Su = \sum_n (u, \phi(x - k))\phi(x - k)\). Bounds for these extreme eigenvalues are \(A = 3^{-d}\) and \(B = 1\) respectively.

As in the tensor product case, we take for the reconstruction the natural injection \(P_n : V_n \rightarrow X\) so that \(R_n \equiv P_n \forall n\in \mathbb{Z}^d\).

\textit{More-dimensional wavelets}

A \textit{wavelet space} \(W_n \subset V_n\), a closed subspace of \(V_n\) which contains those functions in \(V_n\) that cannot be represented in any of the function spaces on the next coarser level, i.e. these functions are in \(V_n\) but not in \(\text{Span}(V_{n-e_1}, \cdots, V_{n-e_d})\). Thus \(W_n \subset V_n\) is a closed subspace so that

\[
V_n = W_n \oplus \text{Span}(V_{n-e_1}, \cdots, V_{n-e_d}),
\]

(2.35)

This means that \(W_n\) contains the ‘difference information’ that is available in the fine grid \(V_n\) but not in the span of the coarser grids \(V_{n-e_1}, V_{n-e_2}, \cdots\) and \(V_{n-e_d}\).
The space $W_n$ is the complement of $\text{Span}(V_{n-e_1}, \cdots, V_{n-e_d})$ in $V_n$. Of course, this complement is not uniquely determined. If we want we can make use of the Hilbert space structure and consider the (unique) orthogonal complement

$$W_n \perp \text{Span}(V_{n-e_1}, \cdots, V_{n-e_d}).$$

(2.36)

This choice corresponds with $R_n : X \to V_n$ being the orthogonal projection. However, in many cases we will use spaces $W_n$ that don't satisfy this orthogonality property!

As soon as we have selected a CSR $\{R_n\}$, then corresponding pre-wavelet spaces are defined as in Section 2.4. These pre-wavelet spaces on an MRA are wavelet spaces.

In the case of an MRA no coarsest grid exists, so that (2.28) gives

$$V_n = \bigoplus_{j \leq n} W_j.$$  

(2.37)

Because of property (2.34.a) we can decompose the space $X(\Omega)$ in

$$X(\Omega) = \bigoplus_{j \in \mathbb{Z}^d} W_j,$$

(2.38)

so that we can write any $u \in X(\Omega)$ as $u = \sum_{j \in \mathbb{Z}^d} w_j$ with $w_j \in W_j$. A restriction $R_n : X(\Omega) \to V_n$ is now determined by

$$R_n \left( \sum_j w_j \right) = \sum_{j \leq n} w_j.$$

(2.39)

By Definition 2.19 we recognise the direct hierarchical surplus

$$Q_n : X(\Omega) \to \text{Ran}(Q_n) = W_n.$$  

(2.40)

We see that there is no coarsest grid and we can decompose $R_n$ as

$$R_n = \sum_{k \leq n} Q_k.$$  

(2.41)

The four relations (2.34.a) to (2.34.d) imply that also the spaces $W_n$ are scaled versions of one space $W_0$,

$$f(x) \in W_n \Leftrightarrow f(2^{-n} x) \in W_0, \quad \forall n \in \mathbb{Z}^d,$$

(2.42)

and, moreover, that they are translation invariant for the discrete translations $2^{-n} \mathbb{Z}^d$,

$$f(x) \in W_0 \Leftrightarrow f(x - k) \in W_0, \quad \forall n \in \mathbb{Z}^d.$$  

(2.43)

As soon as we find a function $\psi(x)$ with the property that $\psi(x - k), k \in \mathbb{Z}^d$, is a basis of $W_0$, then by a simple rescaling we see that $\psi(2^n x - k)$, yields a basis of $W_{n+e}$. Such a function is the more-dimensional generalisation of a wavelet [3]. Because of (2.38) the full collection

$$\left\{ \psi_{n,k}(x) \mid \psi_{n,k}(x) = \psi(2^n x - k), \; n, k \in \mathbb{Z}^d \right\}$$

is a basis of $X(\mathbb{R}^d)$. 

15
3. PIECEWISE APPROXIMATION IN $d$ DIMENSIONS

3.1. Piecewise constant approximation

We approximate $u \in X(\Omega) = L^p_{\text{loc}}(\Omega)$ by $u_n \in V_n$, in the space of piecewise constant functions on $\Omega_n$, i.e. in

$$V_n = \text{Span}\{ \phi_{n} \},$$

(3.1)

with, for some $q \geq 1$ or $q = \infty$,

$$\phi_{n}(x) = 2^{n/q} \phi(2^n x - j),$$

$$\phi(x) = \prod_{j=1}^d \chi_{[0,1]}(x_j),$$

(3.2)

with $\chi_{[0,1]}(x)$ the characteristic function on the unit interval.

This clearly describes a basis for a tensor product space, and we may write

$$V_n = V_n(\Omega) = \bigotimes_{j=1}^d V_n(\Omega_j),$$

(3.3)

the tensor product of spaces $V_n(\Omega_j)$. These $V_n$ are the spaces of piecewise constant functions with meshwidth $h_j = 2^{-n_j}$ on $\Omega_j \subset \mathbb{R}$. The corresponding grid of cells on the Cartesian product of $\{\Omega_j\}$, is denoted by $\Omega_n$. The cell centers are denoted by by $\Omega_n^*$. We define the restriction $R_n$ as the projection

$$R_n : X \times \Omega \ni u \mapsto u_n = R_n u, \text{with}$$

$$u_{n,i} = \int_{\Omega_n} u(\xi) \, d\Omega.$$

(3.4)

This restriction is of type (2.30), and $R_n = R_{n1,\ldots,n_d}$ can be decomposed as

$$R_n = \bigotimes_{j=1}^d R_{n_j} e_j,$$

(3.5)

where $R_{n_j} e_j u(x)$ is the function, piecewise constant in the $j$-th coordinate direction on a partitioning $\Omega_n$, so that

$$R_{n_j} e_j u(x) = 2^{n_j} \int_{x-(h/2) e_j}^{x+(h/2) e_j} u(\xi_1, \ldots, \xi_d) \, d\xi_j$$

for all $x$ with $(x_j2^{n_j} \pm 1/2)\in \mathbb{Z}$.

In the special case $X = L^2(\Omega)$, the space $X = X(\Omega)$ is a Hilbert space, and $\{ \phi_{n_j} \}$ is an orthogonal (orthonormal if $q = 2$) basis in $V_n$. In this case $R_n$ is the orthogonal projection $L^2(\Omega) \to V_n$. For $\Omega = \mathbb{R}^d$, the set $\{ V_n \}$ as defined in (3.1)-(3.2) is a typical MRA. This is no longer the case if we consider a bounded domain $\Omega$, but the decomposition as treated in Section 2 still can be used in the case of a bounded domain.

It is easily checked that the more-dimensional wavelet $\psi(x) \in \mathcal{W}_{e}$, corresponding with the piecewise constant scaling function $\phi(x) \in \mathcal{V}_0$, from the previous section, is the more-dimensional elementary checkerboard function given by

$$\psi(x) = \begin{cases} 0 & \text{if } x \notin \Omega_{0,0}, \\ (-1)^{|k|} & \text{if } x \in \Omega_{0,0} \text{ and } x \in \Omega_{e,k}. \end{cases}$$

(3.6)

This function is the tensor product of the Haar-wavelet.
In wavelet theory the spaces \( W_n \) are labeled \textit{channels}, and the distinct channels are linearly independent. The first decomposition of an arbitrary function from \( X(\Omega) \) consists in writing \( u(x) = \sum w_n(x) \), where \( w_n \in W_n \) with \( n \in \mathbb{Z}^d \), according to (2.38).

Each subspace \( W_n \), has its natural basis, the \textit{standard basis} ,

\[
\left\{ \psi_{n,k}(x) \mid \psi_{n,k}(x) = 2^{n/2}\psi(2^{n/2}x - k), \ k \in \mathbb{Z}^d \right\},
\]

(3.7)
of functions with a minimal support. We see that \( \psi = \psi_{e,0} \in V_e \) is a function with the unit cube \( \Omega_{0,0} \) as support. The basis function \( \psi_{n,k} \) is a scaled, elementary checkerboard function, that may be characterised either by its support, which is a single cell in \( \Omega_{n-e} \), or by the centerpoint of this cell, \( z_{n-e,k} = 2^{n/2}(k + e/2) \).

On the open unit cube \( \Omega = (0,1)^d \) we consider the \( 2^{|n|} \)-dimensional spaces \( V_n = \otimes_{j=1}^d V_{n_j} \), the tensor product of \( V_{n_j}((0,1)) \), the spaces of piecewise constant functions with meshwidth \( h_j = 2^{-n_j} \) in the \( j \)-th coordinate direction. For functions defined on \( \Omega = (0,1)^d \) we can write relation (2.38) as

\[
X(\Omega) = \bigoplus_{n \geq 0} W_n ,
\]

(3.8)
and make a decomposition in channels correspondingly. Each subspace \( W_{n+e} \), now with \( n \geq e \), has its standard basis \( \psi_{n+e,k} \):

\[
\left\{ \psi_{n+e,k}(x) \mid \psi_{n+e,k}(x) = 2^{n/2}\psi(2^n x - k), \ e < k < 2^n \right\} .
\]

(3.9)
For \( \Omega = (0,1)^d \), the exceptions related with the boundary are found in the spaces \( W_n \) with a zero index (i.e. \( \|n\| = 0 \)). These \( W_n \) have basis functions with different shapes. They are derived from the corresponding functions for the unbounded case, but their support is restricted to \( \Omega \). Their corresponding nodal points \( z_{n-e,k} \) are found on the boundary \( \partial \Omega = \overline{\Omega} \setminus \Omega \) instead of in the interior. For \( \|n\| = 0 \) we have \( W_n \) spanned by a basis

\[
\left\{ \psi_{n,k}(x) \begin{array}{ll}
0 \leq k_j < 2^{n_j-1} & \text{if } n_j \neq 0 \\
k_j = 0 & \text{if } n_j = 0 \\
j = 1, \ldots, d
\end{array} \right\} .
\]

(3.10)
Taking such modifications into account, both for \( \Omega = (0,1)^d \) and for \( \Omega = \mathbb{R}^d \), for each \( u \in L^p_{\text{loc}}(\Omega) \) we may write a \textit{hierarchical expansion} (a wavelet expansion) according to (2.38) or (3.8), as

\[
u(x) = \sum_{n \in \mathbb{Z}} w_n = \sum_{n,k} c_{nk} \psi_{n,k}(x) = \sum_{n,k} c_{nk} \hat{\psi}(2^n x - k) ,
\]

(3.11)
where \( \hat{\psi} \) is simply

\[
\hat{\psi}(x) = \begin{cases} = \text{sign } |x| & \text{if } \max(x_1, \ldots, x_d) < 1 , \\
= 0 & \text{if } \max(x_1, \ldots, x_d) > 1 ,
\end{cases}
\]

and \( c_{nk} = 0 \) for all \( k \) with \( \|k\| \) even.

3.2. \textit{Piecewise linear approximation}  

We approximate \( u \in X = C^0(\Omega) \) by \( u_n \in V_n \), in the space of piecewise \( d \)-linear functions on \( \Omega_n \), i.e. in

\[
V_n = \text{Span} \{ \phi_{nj} \} ,
\]

(3.12)
with, for some \( q \geq 1 \) or \( q = \infty \),

\[
\phi_{n,j}(x) = 2^{\lfloor m/q \rfloor} \phi(2^m x - j), \\
\phi(x) = \prod_{j=1}^d \Lambda(x_j),
\]

with \( \Lambda(x) = \max(0, 1 - |x|) \) the usual hat function.

Clearly this is a basis for a tensor product space as (3.3), where \( V_{n,j} \) are spaces of piecewise linear functions on a partitioning of \( \Omega_j \) with meshwidth \( h_j = 2^{-n_j} \). The set of nodal points \( \{ j2^n \}_{j \in \mathbb{Z}^d} \) in \( \Omega \) is denoted by \( \Omega_n^* \).

Here we define the restriction \( R_n \) as the projection

\[
R_n : \mathcal{X} \rightarrow V_n \subset \mathcal{X} \\
u \mapsto u_n = R_n u, \\
u_n(x) = u(x) \forall x \in \Omega_n^*.
\]

The restriction is also of type (2.30) and the operator \( R_n = R_{n_1, \ldots, n_d} \) can be decomposed as (3.5) where \( R_{n_j} u(x) \) is a function, piecewise linear in the \( j \)-th coordinate direction, on a partitioning \( \Omega_n \), such that \( R_{n_j} u(x) = u(x) \) for all \( x \) with \( x_j \in \mathbb{Z} \).

It is clear that there exists a basis-function in \( V_n \) for each nodal point \( x_{nk} \) in \( \Omega_n^* \). If and only if \( \| k \| \) is even, there exists a ‘parent grid’ \( \Omega_m^+ \) with \( m \leq n \) and \( m \neq n \), for which \( x_{nk} \in \Omega_m^* \). Hence, in this case each wavelet space \( W_n \) has its natural basis

\[
\{ \phi_{n,k} \mid k \in \mathbb{Z}^d \text{ with } \| k \| \text{ odd} \}
\]

On the closed cube \( \Omega = [0,1]^d \) we consider the \( \prod_{j=1}^d (2^{n_j} + 1) \)-dimensional spaces \( V_n = V_n(\Omega) = \bigotimes_{j=1}^d V_{n_j}([0,1]) \). With homogeneous Dirichlet boundary conditions, the dimension of the corresponding space \( V_n^0 \subset V_n \) is \( \prod_{j=1}^d (2^{n_j} - 1) \). It is immediately clear that typical FE-basis functions for \( V_n \) are the \( d \)-dimensional hat-functions; functions that vanish on all but one point of \( \Omega_n^+ \). Each such FE basis function of \( V_n^0 \) is characterised by an interior point from \( \Omega_n^+ \).

We notice that for \( \Omega = [0,1]^d \) we have

\[
V_n = \{ 0 \} \text{ except for } n > 0,
\]

and

\[
V_n^0 = \{ 0 \} \text{ except for } n > 1.
\]

With \( W_n \) (or \( W_n^0 \)) we denote the subspace of \( V_n \) (respectively \( V_n^0 \)) of functions that vanish at the gridpoints of all \( \Omega_n^+ \) (respectively \( \Omega_n^+ - e_j \), \( j = 1, \ldots, d \)). From (3.12) we see that for \( \Omega = \mathbb{R}^d \)

\[
W_n = \text{Span}\{ \phi_{n,j} \mid \|j\| \text{ odd} \}.
\]

and for \( \Omega = [0,1]^d \) we see that \( 0 \leq j \leq 2^n \) and

\[
W_n = \text{Span}\left\{ \phi_{n,j} \mid j_i \text{ odd, } 0 < j_i < 2^n \text{ if } n_i > 0, \begin{array}{l} j_i = 0, 1 \text{ if } n_i = 0, \\ i = 1, \ldots, d \end{array} \right\}.
\]

Clearly \( W_n = W_n^0 = \{ 0 \} \), except for \( n > 0 \). If \( \| n \| = 0 \) we see that \( V_n^0 = \{ 0 \} \) and \( W_n \) is spanned by FE basis functions that are characterised by boundary points on the unit cube. Thus, the trace of a
function on the boundary is exclusively approximated by elements of \( W_n \) with \( \| n \| = 0 \). Further we see \( W_n^0 = W_n \) if \( n \geq e \). Apparently
\[
V_n = \bigoplus_{0 \leq k \leq n} W_k, \tag{3.20}
\]
and
\[
V_n^0 = \bigoplus_{e \leq k \leq n} W_k. \tag{3.21}
\]

4. Error estimates for regular grids

The decompositions of type (2.38) allow the approximation of a sufficiently smooth function in \( X(\Omega) \) by a series with elements in \( W_j \). To obtain an impression of the quality of these expansions we derive some error estimates.

4.1. Estimates for piecewise constant approximation

As the case where domain boundaries are present, is the more general one, we study the case \( \Omega = (0, 1)^d \). To quantify the error of approximation on \( \Omega \), we introduce for \( u \in C^e(\Omega) \) the seminorm
\[
|u| = \max_{x \in \Omega} |D^e u(x)| + \max_{0 < |p| < d} \max_{x \in \Omega} |D^p u(x)|. \tag{4.1}
\]

Now we derive the following

**Theorem 4.1** If we consider an expansion of a \( C^e(\Omega) \)-function, \( u \), in piecewise constant functions on the grid \( \Omega_n \), for an arbitrary \( n \in \mathbb{Z}^d \), \( n > 0 \), and if we write
\[
R_n u = \sum_{0 \leq m \leq n} w_m, \tag{4.2}
\]
with \( w_m \in W_m, 0 \leq m \leq n \), then, for \( m \neq 0 \) we have
\[
\| w_m \|_{L^2(\Omega)} \leq 2^{-d/2} 2^{-|m|} |u|, \tag{4.3}
\]
and an estimate for the approximation error
\[
\| u - R_n u \|_{L^2(\Omega)} \leq (2/3)^{d/2} \| h_n \| \| u \|. \tag{4.4}
\]

**Proof:** We take \( \{ \psi_{mk} \} \) as a basis in \( W_m \), \( e \leq m \leq n \). All these functions form an \( L^2(\Omega) \)-orthonormal set (orthonormal Haar basis) and they are orthogonal to all functions in \( W_n \), \( n \neq m \).

Thus, we find (4.2) with \( w_m = \sum_k a_{mk} \psi_{mk} \), where
\[
a_{mk} = (u, \psi_{mk}) = \int_{\Omega} u \psi_{m,k} \, d\Omega = \int_{m-e,k} u \, \psi_{m,k} \, d\Omega.
\]

For \( m \geq e \) the point \( z_{m-e,k} \) lies in the interior of \( \Omega \) and the estimate holds with
\[
|u| = \max_{x} |D^e u(x)|.
\]
Viz., by Taylor expansion around \( z_{m-e,k} \), we have

\[
|a_{m,k}| = \left| \int_{\Omega} u \psi_{m,k} \, d\Omega \right|
\]

\[
\leq \int_{\Omega} \left\| x - z_{m-e,k} \right\| \left| u \psi_{m,k} \right| \, d\Omega
\]

\[
= \int_{\Omega} \left\| x - z_{m-e,k} \right\| \left| u \right| \, |2^{m-e}| \, d\Omega
\]

\[
= \left| u \right| 2^{m-e}/2 \int_{\Omega} \left\| x - z_{m-e,k} \right\| \, d\Omega
\]

\[
= \left| u \right| 2^{m-e}/2 \prod_{j=1}^{d} \int_{0}^{2^{-m_j}} \xi_j \, d\xi_j
\]

\[
= \left| u \right| 2^{m-e}/2 \prod_{j=1}^{d} (1/2^{2m_j})
\]

\[
= \left| u \right| 2^{m-e}/2 - 2|m| \geq 0 \quad \text{for} \quad m \neq 0, \quad m \neq 0,
\]

\[
\text{if} \quad m_1 = 0 \quad \text{and} \quad m_j \geq 1 \quad \text{for} \quad j = 2, ..., d,
\]

\[
\left| u \right| = \max_{x} \left| \frac{\partial^{l-1} u(x)}{\partial x_2 \cdots \partial x_d} \right|
\]

Hence, the estimate (4.5) holds for \( m \geq 0, \quad m \neq 0, \) if we use the seminorm (4.1), and we find

\[
\| w_m \|^2 = \sum_{k} \left| a_{m,k} \right|^2 \leq \sum_{k} 2^{-d} 2^{-3|m|} \left| u \right|^2 = 2^{-d} 2^{-2|m|} \left| u \right|^2.
\]

so that \( \| w_m \| \leq 2^{-d/2} 2^{-|m|/2} \| u \| \), which leads to (4.3) and (4.4) because

\[
\| u - R_n u \|^2 = \sum_{m_1 > n_1 \text{ or } \ldots \text{ or } m_d > n_d} \| w_m \|^2 \leq \sum_{j=1}^{d} \sum_{m_j \geq 0} 2^{-d} 2^{-2|m|} \left| u \right|^2
\]

\[
\leq 2^{-d} \left| u \right|^2 \sum_{j=1}^{d} \sum_{m_j \geq 0} (1/4)^m = 2^{-d} \sum_{j=1}^{d} \sum_{m_j \geq 0} (1/4)^m \left| u \right|^2
\]

\[
\leq 2^{-d} \left| u \right|^2 \left( 4/3 \right)^d \sum_{j=1}^{d} (1/4)^n_j \leq \left| u \right|^2 \left( 2/3 \right)^d \sum_{j=1}^{d} h_{n_j}^2
\]

\[
\leq \left| u \right|^2 \left( 2/3 \right)^d \| h_n \|^2.
\]
4.2. Estimates for piecewise linear approximation

For a function \( u \in C^0(\Omega) \) we consider piecewise linear approximation as in Section 3.2. We approximate \( u \) by \( u_n \in V_n \), where \( V_n \) is the space of piecewise \( d \)-linear functions on \( \Omega_n \). We take \( u_n \) such that \( u_n(x) = u(x) \) for all \( x \in \Omega_n^+ \) and we write

\[
    u_n(x) = \sum_j d_{n,j} \phi_{n,j}(x),
\]

where \( \phi_{n,j}(x) \) is defined by (3.13).

With \( u_n \in V_n \) the piecewise linear approximation on \( \Omega_n \) of the function \( u \in C^0(\Omega) \), we make the hierarchical decomposition \( V_n = \bigoplus_{k \leq n} W_k \), and write

\[
    u_n = \sum_{k \leq n} w_k, \quad w_k \in W_k,
\]

where

\[
    w_k(x) = \sum_j c_{k,j} \phi_{k,j}(x),
\]

with \( c_{k,j} = 0 \) for all \( j \) with \( \|j\| \) even.

In practice the coefficients \( c_{k,j} \), \( \|j\| \) odd, are computed as hierarchical surplus coefficients, by taking the difference between the value \( u(jh_k) \) and the interpolant from coarser grids. This is most conveniently formulated by introducing stencil notation. Therefore, we introduce the difference operator

\[
    \nabla_h u(z) = u(z + h) - u(z),
\]

and the usual central difference approximation for the second derivative by stencil notation, as

\[
    \left[ \frac{1}{2}, -1, \frac{1}{2} \right]_{h_j e_j} u(z) = \frac{1}{2} \nabla_h e_j u(z - h_j e_j).
\]

With this notation we write an expression for the hierarchical coefficients in a piecewise linear approximation. We see that \( d \)-linear interpolation leads to the following expression for the hierarchical surplus coefficient

\[
    c_{k,j} \|h_k\|^{-\frac{1}{d}} = \prod_{j=1}^d \left[ -\frac{1}{2}, 1, -\frac{1}{2} \right]_{h_j e_j} u(jh_k).
\]

Notice that the factor \( \|h_k\|^{-\frac{1}{d}} \) cancels the scaling factor \( 2^{|k|/d} \) in the definition of \( \phi_{k,j} \), so that the function \( u(x) \) is expanded as

\[
    u(x) \approx \sum_{k,j} c_{k,j} \|h_k\|^{-\frac{1}{d}} \phi(2^d x - j).
\]

An expression for the coefficient \( c_{k,j} \) is found in the following lemma.

**Lemma 4.2**

Let \( u \in C^{e+m} \), for a given \( m \) with \( 0 \leq m \leq e \), and let

\[
    L_{n,j}(x) = 2^{-|m|} \phi(2^n x - j),
\]

then, for each \( \phi_{n,j} \in W_n \), \( \|n\| \neq 0 \), \( \|j\| \) odd, we have

\[
    \|h_n\|^{-\frac{1}{d}} |c_{n,j}| = \prod_{i=1}^d \left[ -\frac{1}{2}, 1, -\frac{1}{2} \right]_{h_i e_i} u(jh_n) = (-1)^{e+m} 2^{-d} \int_{\Omega} D^{e+m} u(x) D^{e-m} L_{n,j}(x) d\Omega.
\]
Proof: We see that for $|n| \neq 0$, $|j|$ odd, each $L_{n,j}$ has a support in the interior of $\Omega$. Taking this into account, we give the proof after a coordinate translation with $2^m j$ then we see that for all $i$, $0 \leq i \leq d$, 

$$
\sum_{s_i=-1,1} \int_0^{s_i h_i} D e_i u(z) \, dz_i \\
= \sum_{s_i=-1,1} u(z) \bigg|_{z_i=0}^{s_i h_i} \\
= \sum_{s_i=-1,1} \nabla_{s_i h_i} e_i u(z) \bigg|_{z_i=0} \\
= \{1, -2, 1\}_0 e_i u(z) \bigg|_{z_i=0}
$$

and hence

$$
\prod_{i=1, d} \left[-\frac{1}{2}, \frac{1}{2}\right]_h e_i u(o) \\
= \left(\frac{1}{2}\right)^d \prod_{i=1, d} \left[\sum_{s_i=-1,1} \int_0^{s_i h_i} D e_i \bullet dx_i\right] u(x) \\
= \left(\frac{1}{2}\right)^d \sum_{s_1, \ldots, s_d=-1,1} \int_0^{s_1 h_1} \cdots \int_0^{s_d h_d} D e dx_1 \cdots dx_d \\
= \left(\frac{1}{2}\right)^d \int D e_u(x) \prod_{i=1}^d \left[2^{-n_i} \frac{d}{dx_i} \Lambda(2^n x_i)\right] dx_1 \cdots dx_d \\
= \left(\frac{1}{2}\right)^d \int D e_u(x) \prod_{i=1}^d \left(2^{-n_i} \frac{d}{dx_i} \Lambda(2^n x_i)\right) \, d\Omega \\
= \left(\frac{1}{2}\right)^d \int D e_u(x) \prod_{i=1}^d \left(2^{-n_i} \frac{d}{dx_i} \Lambda(2^n x_i)\right) \, d\Omega \\
= \left(\frac{1}{2}\right)^d \int D e_u(x) \prod_{i=1}^d \left(2^{-n_i} \frac{d}{dx_i} \Lambda(2^n x_i)\right) \, d\Omega \\
= \left(-1\right)^{e+m} \left(2^{-d} \int D e_u(x) \prod_{i=1}^d \left(2^{-n_i} \frac{d}{dx_i} \Lambda(2^n x_i)\right) \, d\Omega \right)
$$

\[ \Box \]

Remarks:

- For $|n| = 0$, (i.e. for boundary points), the same formula holds, provided that the formula is restricted to the lower dimensional boundary manifold (e.g. the face or the edge of the unit cube).

- For an $m$ with $0 \leq m \leq e$ we derive an expression for $\|D^m \phi\|_p$, with $\phi$ given by (3.13) as follows

$$
\|D^m \phi\|_p^p = \int_\Omega \prod_i \left|D^m \Lambda(x_i)\right|^p \, d\Omega \\
= \prod_i \left(\int_1 \left|D^m \Lambda(1-x_i)\right|^p \, dx_i\right) \\
= 2^{d(p+1)} |e-m|
$$

So that

$$
\|D^m \phi\|_p = 2^{d(p+1)} |e-m|/p
$$

(4.13)

- In (3.13) we have $\phi_{n,j} = 2^{|n|/q} \phi(2^n x - j)$, and hence

$$
\|D^m \phi_{n,j}\|_p^p = \int 2^{|n|/q} \left|D^m \phi(2^n x - j)\right|^p \, d\Omega \\
= 2^{|n|/q} \int \left|D^m \phi(2^n x)\right|^p \, d\Omega \\
= 2^{|n|/q} \int 2^{|m|} \left|D^m \phi(2^n x)\right|^p \, d\Omega \\
= 2^{|m|/q} \int 2^{|m|} \left|D^m \phi(2^n x)\right|^p \, d\Omega \\
= 2^{|m|/p} \left|D^m \phi\right|_p^p.
$$

Here, $2^m n = \prod_{i=1}^d 2^m n_i = \prod_{i=1}^d h_{n_i} = h_n$, so that for arbitrary $j$,

$$
\|D^m \phi_{n,j}\|_p = \left\|h_n\right\|^{(1/p-1/q)} h_n^{-m} \|D^m \phi\|_p.
$$

(4.14)

This means that the norm $\|\phi_{n,j}\|_p$ is independent of the level $n$ iff we take $q = p$. 

22
To obtain error estimates in the approximation Theorem 4.3 we compute an expression for 
\[ \|D^m L_{n,j}(x)\|_p \] 
with \( L_{n,j} \) as introduced in (4.12), in particular for \( p = 1, 2, \infty \).

\[ \left\| D^m L_{n,j}(x) \right\|_p = \left\| h_n \right\|^{p+1} \left\| h_n^{-m} D^m \phi \right\|_p^p = \left\| h_n \right\|^{p+1} \left\| h_n^{-m} 2^{-d(p+1)} |e-m| \right\| . \]

So that we may conclude, also considering the special case \( p = \infty \),

\[ \left\| D^m L_{n,j}(x) \right\|_1 = \left\| h_n \right\|^2 \left( h_n / 2 \right)^{-m}, \]
\[ \left\| D^m L_{n,j}(x) \right\|_2 = \left( 2/3 \right)^{d/2} \left\| h_n \right\|^{3/2} \left( h_n / \sqrt{3} \right)^{-m}, \]
\[ \left\| D^m L_{n,j}(x) \right\|_\infty = \left\| h_n \right\| \left( h_n \right)^{-m}. \]

Using the above expressions and Lemma 4.2, we can derive the error estimates in the following theorem.

**Theorem 4.3** Let \( u \in C^{e+m}(\Omega) \) be given for some \( m \) with \( 0 \leq m \leq e \), and let \( u_n \in V_n \) be the piecewise linear approximation on \( \Omega n \) of \( u \), such that \( u_n(x) = u(x) \) for all \( x \in \Omega n^+ \). If we make the hierarchical decomposition \( V_n = \oplus_{k \leq n} W_k \), and write

\[ u_n = \sum_{k \leq n} w_k , \quad w_k \in W_k , \]

then we have the estimates

\[ \| w_k \|_2 \leq \left\| D^{e+m} u \right\|_2 \left\| h_k \right\|^{2-d/2} 3^{-|m|/2} h_k^{-2(e-m)}, \]
\[ \| w_k \|_\infty \leq \left\| D^{e+m} u \right\|_\infty \left\| h_k \right\|^{2} \left( h_k / 2 \right)^{-2(e-m)}, \]
\[ \| u - u_n \|_2 \leq \left\| D^{e+m} u \right\|_2 2^{-d} 3^{-|m|/2} \sum_{i=1}^d h_i^{-1+m}, \]
\[ \| u - u_n \|_\infty \leq \left\| D^{e+m} u \right\|_\infty 6^{-|m|} \sum_{i=1}^d h_i^{-1+m}. \]

**Proof:** Using (4.11) and Lemma 4.2 we can obtain estimates for the hierarchical coefficients \( c_{kj} \).
We fix \( k \) and derive, writing \( h = h_k \),

\[ \left\| h \right\|^{-\frac{1}{2}} |c_{kj}| = \frac{2^{-d} \int_\Omega D^{e+m} u(x) \chi_{kj}(x) \left\| D^{e-m} L_{kj}(x) \right\| dx}{\left\| D^{e+m} u \right\|_2 \left\| \chi_{kj} \right\|_2} \leq \frac{2^{-d} \left\| D^{e+m} u \right\|_\infty \left\| \chi_{kj} \right\|_\infty^{3/2} \left\| h \right\|^{3/2} \left( h / \sqrt{3} \right)^{-2(e-m)}}, \]

where \( \chi_{kj} \) is the characteristic function for the support of \( L_{kj}(x) \), or similarly

\[ \left\| h \right\|^{-\frac{1}{2}} |c_{kj}| = \frac{2^{-d} \int_\Omega D^{e+m} u(x) \chi_{kj} \left\| D^{e-m} L_{kj}(x) \right\| dx}{\left\| D^{e+m} u \right\|_2 \left\| \chi_{kj} \right\|_2} \leq \frac{6^{-d/2} \left\| D^{e+m} u \right\|_2 \left\| \chi_{kj} \right\|_2 \left\| h \right\|^{3/2} \left( h / \sqrt{3} \right)^{-2(e-m)}}. \]

We write \( w_k = \sum_j c_{kj} \phi_{kj} \) with \( |j| \) odd, and we know that these functions \( \{ \phi_{kj} \}_j \), for fixed \( k \) have disjoint supports. Hence, for the hierarchical contribution,

\[ \left\| w_k \right\|_2^2 = \sum_{i,j} \left\| c_{ki} \phi_{ki} \right\|_2^2 = \sum_{i,j} c_{ki}^2 \phi_{ki}^{2/2} \int \phi_{ki} \phi_{kj} d\Omega \leq \left\| h \right\|^{2/\delta} \left\| \phi_{ki} \right\|_2^2 \sum_j \left\| D^{e+m} u \chi_{kj} \right\|_2^2 6^{-d} \left\| h \right\| \left( h / \sqrt{3} \right)^{-2(e-m)} \leq \left\| h \right\|^{2/\delta} \left\| D^{e+m} u \right\|_2 \left( h / \sqrt{3} \right)^{-2(e-m)} 3^{-|m|}. \]

23
For the other norm \(\|w_k\|_\infty\) we obtain similarly
\[
\|w_k\|_\infty = \max_j \|c_j \phi_k j\|_\infty \\
\leq \|h\|^2 \max_j 2^{-d} \|D^{e+m} u \phi_k j\|_\infty \|h\|^2 \left(\frac{h}{2}\right)^{-(e-m)} \\
\leq 2^{-d} \|D^{e+m} u\|_\infty \|h\|^2 \left(\frac{h}{2}\right)^{-(e-m)}
\]

For the error, for \(p = 2\) or \(p = \infty\), we get
\[
\|u - u_n\|_p = \|\sum_k w_k - \sum_{k \leq n} w_k\|_p \\
\leq \sum_{k \leq n} \|w_k\|_p \\
\leq C_p \|D^{e+m} u\|_p \sum_{k \leq n} \|h_k\|^2 h^{-e-m} \\
= C_p \|D^{e+m} u\|_p \sum_{k \leq n} h_k^{e+e-m}
\]
with \(C_2 = 2^{-d/2}3^{|m|/2}\) or \(C_\infty = 2^{-d}2^{e-m}\). This yields the above mentioned estimates, by taking into account that
\[
\sum_{k \leq n} h_k^{e+e-m} = \sum_k h_k^{e+e-m} - \sum_{k \leq n} h_k^{e+e-m} = \\
= \sum_{0 \leq k_1, \ldots, 0 \leq k_d} \prod_{i=1}^d \left(1 - (1/2)k_i(1+m_i)\right) - \sum_{0 \leq k_1, \ldots, 0 \leq k_d \leq n_d} \prod_{i=1}^d h_i^{1+m_i} = \\
= \prod_{i=1}^d \left[1 - \prod_{i=1}^d \left(1 - (1/2)k_i(1+m_i)\right)\right] \\
\leq 2^d (2/3)^{|m|} \sum_{i=1}^d 2^{-1+m_i(1+m_i)} = 3^{-m} \sum_{i=1}^d h_i^{1+m_i}
\]

For \(m = e\) this simply reads \(\sum_{k \leq n} \|h_k\|^2 \leq 3^{-d} \|h_n\|^2\). □

**Corollary 4.4**

As a direct corollary we find
\[
\|u - u_n\|_2 \leq 54^{-d/2} \|D^{e} u\|_2 \|h_n\|^2, \\
\|u - u_n\|_\infty \leq 6^{-d} \|D^{e} u\|_\infty \|h_n\|^2,
\]
and, for \(p = 2\) or \(p = \infty\), and \(0 \leq m \leq e\),
\[
\|w_k\|_p \leq C \|D^{e+m} u\|_p h_k^{e+m},
\]
and
\[
\|u - u_n\|_p \leq C \|D^{e+m} u\|_p \sum_{i=1}^d h_i^{1+m_i}, \\
\|u - u_n\|_p \leq C \|D^{e} u\|_p \|h_n\|^2.
\]
From (14.5) also follows a bound by a \(W_p^{e, \ell}\)-norm, We immediately see that, for \(0 \leq \ell \leq d\),
\[
\|w_k\|_p \leq C \|u\|_{W_p^{e, \ell}} \min_{|m| = \ell} h_k^{e+m} \leq C \|u\|_{W_p^{e, \ell}} \|h_k\|^{1+\ell/d}.
\]
Remark:
We gave the proof of Theorem 4.3 for functions that vanish at the boundary of \( \Omega \). Taking into account the remark following Lemma 4.2, it is clear that similar estimates (with different constants and with terms including derivatives of \( u \) that are restricted to the boundary planes) also hold for functions with non-homogeneous boundary conditions.

5. Error estimates for sparse grids

5.1. Estimates for piecewise constant approximation

Theorem 5.1 Let \( \hat{R}_n u \) be the piecewise constant approximation of a function \( u \in L_2(\Omega) \) on a sparse grid on level \( n \):

\[
\hat{R}_n u = \sum_{|k| \leq n} w_k, \quad w_k \in W_k, \tag{5.1}
\]

then, with \( \| h \| = 2^{-n} \), the volume of the finest cells, we have the estimate

\[
\| u - \hat{R}_n u \|_{L_2(\Omega)} \leq C \| u \| \log^{(d-1)/2} \| h \|. \tag{5.2}
\]

Proof: To prove the theorem for the \( L_2(\Omega) \)-norm, we use (4.6), and the orthogonality of the hierarchical basis functions, to obtain

\[
\| u - \hat{R}_n u \|^2_{L_2(\Omega)} \leq \sum_{|k| > n} \| w_k \|^2_{L_2(\Omega)} \\
\leq \sum_{|k| > n} 2^{-d} 4^{-|k|} |u|^2 \\
= 2^{-d} |u|^2 \sum_{|k| > n} 2^{-2|k|} \\
= 2^{-d} |u|^2 \sum_{l > n} 2^{-2l} \left( \frac{l + d - 1}{d - 1} \right) \tag{5.3}
\]

We know

\[
\sum_{l > n} 2^{-2l} \left( \frac{l + d - 1}{d - 1} \right) \\
= 2^{-2(n+1)} \left( \frac{n + d}{d - 1} \right) F(1,1 + n + d; 2 + n; 1/4) \\
= G(n,d) \tag{5.4}
\]

where \( F \) is the hypergeometric function. It follows that

\[
G(n,d) \sim \frac{n^{d-1} 2^{-d n}}{3(d-1)!} \quad \text{for} \quad n \to \infty ,
\]

where the asymptotic value is reached soon for small values of \( d \). Hence

\[
\| u - R_n u \|^2 \leq 2^{-d} |u|^2 \sum_{l > n} 2^{-2l} \left( \frac{l + d - 1}{d - 1} \right) \\
\leq 2^{-d} |u|^2 \left( \frac{n + d}{d - 1} \right) F(1,1 + n + d; 2 + n; 1/4) \tag{5.5}
\]

where \( C_{nd} \) is a constant that tends to one for large \( n \). So, we conclude that

\[
\| u - R_n u \|_{L_2(\Omega)} \leq C \| u \| n^{(d-1)/2} 2^{-n} ,
\]

which is equivalent with (5.2). \( \square \)
To guarantee a small error on a regular grid, in (4.4) all cell edges $h_j$ need to be small, but in (5.2) for the sparse grid only the volume $|h|$ has to be small. Further, in the two-dimensional case, the estimate (5.2) is of a similar order of accuracy as (4.4), except for a logarithmic small factor. However, the number of degrees of freedom for the approximation (5.2) is significantly less. Namely, in the unit cube, for $R_n u$ the number of degrees of freedom is $2^n|n|$, whereas for $\hat R_n u$ it is $\mathcal{O}(n^{d-1}2^n)$, viz. $2n^{2^n} + 1$ in the 2D case, and in the 3D-case e.g. $(n^2 + n + 2)2^n - 1$. Because significantly less degrees of freedom are involved in the approximation $\hat R_n u$ than in the approximation of $R_{(n,n,n)} u$, i.e. less coefficients $a_{j,k}$ and less gridpoints $z_{j,k}$, in analogy to [7], we call the approximation $\hat R_n u$ the \textit{sparse grid approximation} and

$$\Omega_n^+ = \left\{ z_{j,k} \mid z_{j,k} \in \mathbb{Z}_n, |n| \leq n \right\}$$

is the \textit{sparse (box) grid} for this approximation on level $n$.

\subsection*{5.2. Estimates for piecewise linear approximation}

For piecewise linear approximation we use a \textit{sparse (vertex) grid}. $\Omega_n^+ = \bigcup_{|k| \leq n} \Omega_l^+ \cap \Omega$. A sparse grid approximation is obtained by interpolation on this grid by means of the space spanned by all $W_k$ with $|k| \leq n$.

\textbf{Theorem 5.2} Let $\hat R_n u$ be the piecewise $d$-linear approximation of a function $u \in C^{e+m}_0(\Omega)$, with $0 \leq m \leq e$, on a sparse grid on level $n$:

$$\hat R_n u = \sum_{|k| \leq n} w_k, \quad w_k \in W_k,$$

(5.6)

then, with $|h| = 2^{-n}$, the volume of the finest cells, we have for $p = 2, \infty$, with $m = e$ the estimates

$$\| u - \hat R_n u \|_p \leq C \| D^{e+m} u \|_p \| h \|^2 \log^{d-1} \| h \|^{-1},$$

(5.7)

and with $|m| < d$ the estimates

$$\| u - \hat R_n u \|_p \leq C \| D^{e+m} u \|_p \| h \| \log^{d-1-|m|} \| h \|^{-1},$$

(5.8)

and with $0 \leq \ell \leq d$

$$\| u - \hat R_n u \|_p \leq C \| u \|_{W^{e,\ell}_p} \| h \|^{\ell/d} \log^{d-1} \| h \|^{-1}.$$  

(5.9)

\textbf{Proof:} Using the estimates for $\| w_k \|_p$ from Theorem 4.3, we prove, more generally, for some $m$ with $0 \leq m \leq e$, and for $p = 2$ or $p = \infty$,

$$\| u - \hat R_n u \|_p \leq \sum_{|k| > n} \| w_k \|_p \leq \sum_{|k| > n} C \| D^{e+m} u \|_p h_k^{e+m}$$

$$= C \| D^{e+m} u \|_p \sum_{|k| > n} h_k^{e+m}$$

$$= C \| D^{e+m} u \|_p \sum_{|k| > n} h_k^{e+m}$$

$$= C \| D^{e+m} u \|_p \sum_{j=1}^d \Pi_{j=1}^d 2^{-k_j m_j}$$

$$= C \| D^{e+m} u \|_p \sum_{j=1}^d \Pi_{j=1}^d 2^{-k_j m_j}$$

$$= C \| D^{e+m} u \|_p \sum_{j=1}^d \Pi_{j=1}^d 2^{-k_j m_j}$$

$$\leq C \| D^{e+m} u \|_p 2^{-n} (C_1 2^{-n} |m|^{-1} + C_2 n^{d-|m|-1}),$$

(5.10)

with $C_1 = 0$ if $|m| = 0$, and $C_2 = 0$ if $|m| = d$. Hence, for $m \neq e$ we have

$$\| u - \hat R_n u \|_p \leq C \| D^{e+m} u \|_p \| h \| \log^{d-1-|m|} \| h \|^{-1}.$$
Moreover, (5.10) yields, for \( m = e \),
\[
\| u - \widehat{R}_n u \|_p \leq C \| D^{2e} u \|_p \| h \|^2 \log^{d-1} \| h \|^{-1}.
\]
Further, using the estimate (4.17) we obtain, similar to the proof for Theorem 5.1,
\[
\| u - \widehat{R}_n u \|_p \leq \sum_{|k| > n} \| w_k \|_p \sum_{|k| > n} C \| u \|_{w_{p,e}, l} \| h_k \|^{1 + \ell/d}
\]
\[
= C \| u \|_{w_{p,e}, l} \sum_{|k| > n} 2^{-|k| (1 + \ell/d)}
\]
\[
= C \| u \|_{w_{p,e}, l} \sum_{|k| > n} \left( \frac{l + d - 1}{d - 1} \right) 2^{-l(1 + \ell/d)}
\]
\[
= C \| u \|_{w_{p,e}, l} 2^{-n} F(1, 1 + n + d; 2 + n; 2^{-1 + \ell/d}) 2^{-n(1 + \ell/d)}
\]
\[
\leq C \| u \|_{w_{p,e}, l} 2^{-n(1 + \ell/d)} n^{d-1}/(d - 1)!
\]
\[
\leq C \| u \|_{w_{p,e}, l} \| h \|^{1 + \ell/d} \log^{d-1} \| h \|^{-1}.
\] (5.11)

\( \square \)

**Theorem 5.3** Let \( \widehat{R}_n u \) be the piecewise \( d \)-linear approximation of a function \( u \in C_{0}^{e,1}(\Omega) \) on a sparse grid on level \( n \), as in Theorem 5.2, then, with \( \| h \| = 2^{-n} \), the volume of the finest cells, we have, for \( p = 2, p = \infty \), the estimates
\[
\| u - \widehat{R}_n u \|_{W^1_p} \leq C \| h \| \log^{d-1} \| h \|^{-1} \| u \|_{w_{p,e}, l} =
\] (5.12)

If, moreover, we know \( u \in C^{2e} \), then
\[
\| u - \widehat{R}_n u \|_{W^1_p} \leq C \| h \| \| D^{2e} u \|_p .
\] (5.13)

**Proof:** Let \( u \) be sufficiently differentiable and let \( 0 \leq m \leq e \) and \( |m| \geq 1 \), then

**Part 1:**
\[
\| D^m w_k \|_p = \| D^m \prod_{j=1}^d (R_k - R_{k-e_j}) u \|_p
\]
\[
\leq C^d \| D^m \prod_{j=1}^d h_k D^e_j u \|_p
\]
\[
\leq C^d \| h_k \| \| D^m D^e \|_p
\]
\[
\| D^m (u - \widehat{R}_n u) \|_p \leq \| \sum_{|k| > n} D^m w_k \|_p
\]
\[
\leq \sum_{|k| > n} \| D^m w_k \|_p
\]
\[
\leq \sum_{|k| > n} C^d \| h_k \| \| D^m u \|_p
\]
\[
\leq C^d \| D^m u \|_p \sum_{|k| > n} \| h_k \|^{1 + \ell/d}
\]
\[
\leq C^d \| D^m u \|_p \sum_{|k| > n} \left( \frac{l + d - 1}{d - 1} \right) 2^{-l}
\]
\[
\leq C^d \| D^m u \|_p n^{d-1}/(d - 1)!
\]
\[
\leq C^d \| D^m u \|_p \| h \| \log^{d-1} \| h \|.
\] (5.15)

**Part 2:**
\[
\| D^m (u - \widehat{R}_n u) \|_p \leq \| \sum_{|k| > n} D^m w_k \|_p
\]
\[
\leq \sum_{|k| > n} \| D^m \prod_{j=1}^d (R_k - R_{k-e_j}) u \|_p
\]
\[
\leq \sum_{|k| > n} C \| h_k \| \| D^{e-m} u \|_p
\]
\[
\leq C \| D^{e-m} u \|_p \sum_{|k| > n} \| h_k \|^{2 \ell/m}
\]
\[
\leq C \| D^{e-m} u \|_p \sum_{|k| > n} 2^{-2\ell} \sum_{|k| > n} \| h_k \|^{2 \ell/m}
\]
\[
= C \| D^{e-m} u \|_p \sum_{|k| > n} 2^{-2\ell} \sum_{|k| > n} 2^{-k \| e-m \|}
\]
\[
\leq C \| D^{e-m} u \|_p 2^{-n} (C_1 2^{-n} n^{d+|m|-1} + C_2 n^{|m|-1})
\] (5.16)
with $C_1 = 0$ if $|m| = d$, and $C_2 = 0$ if $|m| = 0$. Because

$$
\|v\|_{W^1_p} = \left( \|v\|_p^p + \sum_{|m|=1,0 \leq m \leq e} \|D^m v\|_p^p \right)^{1/p},
$$

we consider the case $|m| = 1$ and we find

$$
\|u - \hat{R}_n u\|_{W^2_p} \leq C \|D^{2e} u\|_p |H|.
$$

Together with the result of Theorem 5.2 this proves the theorem. \(\Box\)

\textbf{References}


SUMMARY OF NOTATIONS

We use the following notation for partial derivatives, \( n \in \mathbb{N}_0^d \),

\[
D^n = D^{n_1, \cdots, n_d} = \left( \frac{\partial}{\partial x_1} \right)^{n_1} \cdots \left( \frac{\partial}{\partial x_d} \right)^{n_d}.
\]

For the Banach spaces of continuously differentiable functions we use the notations

\[
C^n(\Omega) = \left\{ u \mid \max_{0 \leq m \leq n} \max_{x \in \Omega} |D^m u(x)| < \infty \right\},
\]

\[
\| u \|_{C^n} = \max_{0 \leq m \leq n} \max_{x \in \Omega} |D^m u(x)|,
\]

and, with \( l \in \mathbb{N}_0 \),

\[
C^l(\Omega) = \{ u \in C(\Omega) \mid u \in C^m(\Omega); |m| \leq l \},
\]

\[
\| u \|_{C^l} = \max_{|m| \leq l} \| u \|_{C^m} = \max_{|m| \leq l} \max_{x \in \Omega} |D^m u(x)|.
\]

For the Banach spaces of integrable functions, \( 1 \leq p \leq \infty \), we, similarly, use the notation

\[
W^p_n(\Omega) = \left\{ u \left| \sum_{0 \leq m \leq n} \int_{\Omega} |D^m u(x)|^p < \infty \right\},
\]

\[
\| u \|_{W^p_n} = \sqrt[p]{\int_{\Omega} |D^m u(x)|^p} \text{ and } \| u \|_{W^p_n} = \sqrt[p]{\sum_{0 \leq m \leq n} |u|^p_{W^p_n}}.
\]

For \( 0 \leq k \leq d \) we write \( W^p_n, k(\Omega) = \bigcap_{|m| = k} W^p_n + m(\Omega) \), and

\[
\| u \|_{W^p_n, k} = \left( \sum_{0 \leq m \leq e, |m| = k} \| D^n + m u \|_{L^p}^p \right)^{1/p}.
\]

With a \( W^p_{p,0}(\Omega) \) and \( W^p_{p,0} n,k(\Omega) \) we denote the corresponding subspaces with homogeneous boundary conditions. For \( p = \infty \) we use the standard modifications, and for \( W_2 \) we also write \( H \).

### Table 1. Elementary notations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbb{N} )</td>
<td>natural numbers</td>
</tr>
<tr>
<td>( \mathbb{N}_0 )</td>
<td>( \mathbb{N} \cup {0} )</td>
</tr>
<tr>
<td>( \mathbb{Z} )</td>
<td>integer numbers</td>
</tr>
<tr>
<td>( \mathbb{R} )</td>
<td>real numbers</td>
</tr>
<tr>
<td>( \mathbb{R}_+ )</td>
<td>positive real numbers</td>
</tr>
<tr>
<td>( n )</td>
<td>( (n_1, \cdots, n_d) ), multi-integer</td>
</tr>
<tr>
<td>(</td>
<td>n</td>
</tr>
<tr>
<td>( |n| )</td>
<td>( n_1 \cdot n_2 \cdots n_d ), ( n \in \mathbb{N}_0^d )</td>
</tr>
<tr>
<td>( o \in \mathbb{N}_0^d )</td>
<td>( (0, \cdots, 0) )</td>
</tr>
<tr>
<td>( e_i \in \mathbb{N}_0^d )</td>
<td>( (0, \cdots, 0, 1, \cdots, 0) ), the ( i )-th unit vector</td>
</tr>
<tr>
<td>( e \in \mathbb{N}_0^d )</td>
<td>( (1, \cdots, 1) )</td>
</tr>
<tr>
<td>( E )</td>
<td>( {e_1, e_2, \cdots, e_d} )</td>
</tr>
</tbody>
</table>
Table 2. Grids

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega \subset \mathbb{R}^d$</td>
<td>connected open subset</td>
</tr>
<tr>
<td>$h_n \in \mathbb{R}_+$</td>
<td>$h_n = (2^{-n_1}, 2^{-n_2}, \ldots, 2^{-n_d})$</td>
</tr>
<tr>
<td>$|h_n|$</td>
<td>$\sqrt{\sum_{j=1}^d h_{n_j}^2}$, $h_n \in \mathbb{R}_+$</td>
</tr>
<tr>
<td>$|h_n|$</td>
<td>$\prod_{j=1}^d h_{n_j} = 2^{-</td>
</tr>
<tr>
<td>$\Omega_{n,i}$</td>
<td>${x \mid ih_n \leq x &lt; (i+1)h_n}$, a grid cell</td>
</tr>
<tr>
<td>$\Omega_n$</td>
<td>${\Omega_{n,i} \mid i \in \mathbb{Z}^d}$, a grid of cells</td>
</tr>
<tr>
<td>$N_n \subset \mathbb{R}^d$</td>
<td>${x \mid x = ih_n, i \in \mathbb{Z}^d}$</td>
</tr>
<tr>
<td>$\Omega_n^*$</td>
<td>$\Omega \cap N_n$ a grid of (interior) vertex nodes</td>
</tr>
<tr>
<td>$\Omega_n^+$</td>
<td>$\Omega \cap N_n$, a grid of vertex nodes</td>
</tr>
<tr>
<td>$Z_n \subset \mathbb{R}^d$</td>
<td>${z_{n,i} \mid z_{n,i} = (i + e/2)h_n, i \in \mathbb{Z}^d}$</td>
</tr>
<tr>
<td>$\Omega_n^*$</td>
<td>$\Omega \cap Z_n$, a grid of cell centers</td>
</tr>
</tbody>
</table>

Table 3. Families of grids

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G$</td>
<td>${\Omega_n \mid n \in \mathbb{Z}^d}$, an infinite grid of grids</td>
</tr>
<tr>
<td>$G_n$</td>
<td>${\Omega_m \mid 0 \leq m \leq n, m \in \mathbb{Z}^d}$, a finite grid of grids</td>
</tr>
<tr>
<td>$G^+$</td>
<td>${\Omega_n^+ \mid n \in \mathbb{Z}^d}$</td>
</tr>
<tr>
<td>$G_n^+$</td>
<td>${\Omega_m^+ \mid 0 \leq m \leq n, m \in \mathbb{Z}^d}$</td>
</tr>
<tr>
<td>$G^*$</td>
<td>${\Omega_n^* \mid n \in \mathbb{Z}^d}$</td>
</tr>
<tr>
<td>$G_n^*$</td>
<td>${\Omega_m^* \mid 0 \leq m \leq n, m \in \mathbb{Z}^d}$</td>
</tr>
</tbody>
</table>

Table 4. Sparse grids

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_n$</td>
<td>${\Omega_m \mid m \in \mathbb{Z}^d,</td>
</tr>
<tr>
<td>$\overline{D}_n$</td>
<td>${\Omega_m \mid m \in \mathbb{Z}^d, 0 \leq m,</td>
</tr>
<tr>
<td>$D_n^+$</td>
<td>${x \mid x \in \Omega_m^+, 0 \leq</td>
</tr>
<tr>
<td>$D_n^*$</td>
<td>${x \mid x \in \Omega_m^*, 0 \leq</td>
</tr>
</tbody>
</table>

Table 5. Mappings

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X(\Omega)$</td>
<td>a Banach space of functions defined on $\Omega$</td>
</tr>
<tr>
<td>$V_n$</td>
<td>$\text{Ran}(R_n)$</td>
</tr>
<tr>
<td>$R_n$</td>
<td>$X(\Omega) \rightarrow V_n$, a restriction</td>
</tr>
<tr>
<td>$P_n$</td>
<td>$V_n \rightarrow X(\Omega)$, a prolongation</td>
</tr>
<tr>
<td>$\Pi_n$</td>
<td>$X \rightarrow \text{Ran}(P_n) \subset X$, a projection</td>
</tr>
<tr>
<td>$Rmn$</td>
<td>$V_n \rightarrow V_m$, $m \leq n$, a restriction</td>
</tr>
<tr>
<td>$Pmn$</td>
<td>$V_m \rightarrow V_n$, $m \leq n$, a prolongation</td>
</tr>
<tr>
<td>INDEX</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td>approximation</td>
<td>set</td>
</tr>
<tr>
<td>sparse grid, 26</td>
<td>of prolongations, 4</td>
</tr>
<tr>
<td>Bessel sequence, 14</td>
<td>of restrictions, 4</td>
</tr>
<tr>
<td>cell, 2</td>
<td>transfer operators, 4</td>
</tr>
<tr>
<td>channel, 17</td>
<td>NSP, 4</td>
</tr>
<tr>
<td>characteristic function, 16</td>
<td>NSR, 4</td>
</tr>
<tr>
<td>coarsest grid, 10</td>
<td>operator</td>
</tr>
<tr>
<td>coherent</td>
<td>difference, 21</td>
</tr>
<tr>
<td>mutually, 6</td>
<td>relational, 2</td>
</tr>
<tr>
<td>commutative</td>
<td>orthogonal complement, 15</td>
</tr>
<tr>
<td>transfer operators, 7</td>
<td>pre-wavelet space, 11</td>
</tr>
<tr>
<td>commutative set</td>
<td>prolongation</td>
</tr>
<tr>
<td>of prolongations, 7</td>
<td>tensor product, 11</td>
</tr>
<tr>
<td>of restrictions, 7</td>
<td>prolongations</td>
</tr>
<tr>
<td>complement, 15</td>
<td>commutative set</td>
</tr>
<tr>
<td>orthogonal, 15</td>
<td>nested set, 4</td>
</tr>
<tr>
<td>CSP, 7</td>
<td>reconstruction, 4</td>
</tr>
<tr>
<td>CSR, 7</td>
<td>relational operators, 2</td>
</tr>
<tr>
<td>curse of three dimensions, 1</td>
<td>restriction, 4</td>
</tr>
<tr>
<td>difference operator, 21</td>
<td>tensor product, 11</td>
</tr>
<tr>
<td>dyadic block, 2</td>
<td>restrictions</td>
</tr>
<tr>
<td>father function, 14</td>
<td>commutative set</td>
</tr>
<tr>
<td>frame, 13</td>
<td>nested set, 4</td>
</tr>
<tr>
<td>bounds, 14</td>
<td>scaling function, 14</td>
</tr>
<tr>
<td>exact, 14</td>
<td>sparse grid</td>
</tr>
<tr>
<td>tight, 14</td>
<td>approximation, 26</td>
</tr>
<tr>
<td>grid</td>
<td>box, 26</td>
</tr>
<tr>
<td>coarsest, 10</td>
<td>vertex, 26</td>
</tr>
<tr>
<td>of grids, 2</td>
<td>stencil, 21</td>
</tr>
<tr>
<td>hat function, 18</td>
<td>tensor product</td>
</tr>
<tr>
<td>hierarchical</td>
<td>prolongation, 11</td>
</tr>
<tr>
<td>surplus, 10</td>
<td>restriction, 11</td>
</tr>
<tr>
<td>direct, 10</td>
<td>space, 11</td>
</tr>
<tr>
<td>relative, 10</td>
<td>transfer operators, 4</td>
</tr>
<tr>
<td>surplus coefficients, 21</td>
<td>nested, 4</td>
</tr>
<tr>
<td>merging operator, 8</td>
<td>wavelet, 15</td>
</tr>
<tr>
<td>MRA, 13</td>
<td>space, 14</td>
</tr>
<tr>
<td>multi-integer, 2</td>
<td></td>
</tr>
<tr>
<td>multiresolution analysis, 13</td>
<td></td>
</tr>
<tr>
<td>multidimensional, 14</td>
<td></td>
</tr>
<tr>
<td>mutually coherent, 6</td>
<td></td>
</tr>
<tr>
<td>nested</td>
<td></td>
</tr>
</tbody>
</table>