Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

L^∞ estimation of tensor truncations

(revised version: April 2012)

by

Wolfgang Hackbusch

Preprint no.: 17 2012

L^∞ estimation of tensor truncations

Wolfgang Hackbusch
Max-Planck-Institut *Mathematik in den Naturwissenschaften*
Inselstr. 22, D-04103 Leipzig

Abstract

Tensor truncation techniques are based on singular value decompositions. Therefore, the direct error control is restricted to ℓ^2 or L^2 norms. On the other hand, one wants to approximate multivariate (grid) functions in appropriate tensor formats in order to perform cheap pointwise evaluations, which require ℓ^∞ or L^∞ error estimates. Due to the huge dimensions of the tensor spaces, a direct estimate of $\|\cdot\|_\infty$ by $\|\cdot\|_2$ is hopeless. In the paper we prove that, nevertheless, in cases where the function to be approximated is smooth, reasonable error estimates with respect to $\|\cdot\|_\infty$ can be derived from the Gagliardo-Nirenberg inequality because of the special nature of the SVD truncation.

AMS Subject Classifications: 15A69, 15A18, 35J08, 46E35, 26D10
Key words: tensor calculus, tensor truncation, high-order singular value decomposition (HOSVD), approximation, Gagliardo-Nirenberg inequality

1 Introduction

In the numerical tensor calculus, one is operating with large-scale tensors. It is essential to represent a tensor in a certain format requiring extremely less storage than the number of entries of a tensor. The possible operations between such formatted tensors are addition, scalar products, multiplication of (Kronecker) tensors (representing matrices) by tensors (representing vectors), Hadamard products, convolutions, etc. All these operations have the tendency to increase the so-called representation ranks of the format, i.e., the storage cost increases. To overcome this difficulty, one applies tensor truncations (from higher representation ranks to lower representation ranks). For details of the formats and operations, we refer to Hackbusch [6].

The tensor subspace format (also called Tucker format; cf. [6, §8]) as well as the hierarchical format (this includes the TT format; cf. [6, §§11-12]) allow a black-box like truncation based on singular value decompositions (SVD). In this context, they are called HOSVD (higher order SVD; cf. De Lathauwer et al. [3]). Formally, the tensor is rewritten as a certain matrix—the so-called matricisation of a tensor (cf. [6, §5.2])—whose SVD is computed. Given a decomposition $\sum_\nu \sigma_\nu u_\nu \otimes v_\nu$ with orthonormal systems $\{u_\nu\}$ and $\{v_\nu\}$, one can drop the terms with σ_ν sufficiently small. The very pleasant feature of this HOSVD-based truncation is the black-box character and the fact that standard linear algebra program tools can be applied. Furthermore, one has full control about the truncation error, which is $\sqrt{\sum' \sigma_\nu^2}$, where the sum $\sum'$ involves all dropped terms.

The drawback of the described truncation procedure is the choice of norms. The SVD requires finite or infinite dimensional Hilbert spaces. In the standard case, the Hilbert space is equipped with the ℓ^2 or L^2 scalar product and norm. Consequently, the error control mentioned above holds only with respect to this norm. There are many applications, where the ℓ^2/L^2 norm is the adequate choice. However, below we shall describe situations, where the supremum norm $\|\cdot\|_\infty$ is the desired norm.

Particular examples of tensor spaces are spaces of functions defined on the d-fold Cartesian product $\Omega := \Omega_1 \times \ldots \times \Omega_d$, where, e.g., $\Omega_j \subset \mathbb{R}$. If we aim at a Hilbert space, the simplest choice are the spaces $V_j = L^2(\Omega_j)$, which generate the Hilbert tensor space

$$\mathbf{V} = {}_{\|\cdot\|_{L^2}}\bigotimes_{j=1}^{d} V_j = L^2(\Omega).$$

A possible application is the approximation of a particular d-variate function $f(x_1, \ldots, x_d)$ within one of the tensor formats. Having computed an approximation $\tilde{f}$ of f, one can cheaply evaluate the function f, whose direct evaluation may be very costly, otherwise. In practice, it is not the function $f \in L^2(\Omega)$ which is considered, but its restriction to a regular grid $\mathcal{G} \subset \Omega$, i.e., $f \in \ell^2(\mathcal{G})$. The Euclidean norm of ℓ^2 should be scaled corresponding to the $L^2(\Omega)$ norm. If h is the grid size of G, $\|f\|_{\ell^2(\mathcal{G})} = \sqrt{\sum_{\mathbf{x}\in\mathcal{G}} |h^d f(\mathbf{x})|^2}$ is a suitable choice. Concrete examples of multivariate functions coded in tensor formats can be found in Ballani [1], where f represents involved integrals depending on parameters $x_1, \ldots, x_d$. In this example, the function is analytic in all variables.

In the latter example, the aim is to evaluate $\tilde{f}$ at some argument $\mathbf{x} \in \mathcal{G}$. Therefore, one is interested in the *pointwise* error $|\tilde{f}(\mathbf{x}) - f(\mathbf{x})| \le \|\tilde{f} - f\|_\infty$. We recall that the estimate $\|\cdot\|_\infty \le h^{-d} \|\cdot\|_{\ell^2(\mathcal{G})}$ is sharp for the norm defined above. Obviously, error estimates with respect to $\|\cdot\|_{\ell^2(\mathcal{G})}$ do not really help.

In contrary to these pessimistic remarks, practical experiments of tensor approximations of smooth functions do not show such deficits. The aim of this paper is the justification that, indeed, the supremum norm $\|\cdot\|_\infty$ can be estimated by the $\|\cdot\|_{\ell^2/L^2}$ norm, provided that we approximate smooth functions. The smoothness is described by $|f|_m \le M_m$, where $|\cdot|_m$ is a semi-norm involving m-th derivatives.

In a first step (see §2.1), we recall the interpolation inequality by Gagliardo and Nirenberg, which allows to obtain a supremum norm estimate from the L^2 norm bound, provided that the function is sufficiently smooth. For C^∞ functions, one can—in the best case—derive an inequality of the form $\|\varphi\|_\infty \le C \|\varphi\|_{L^2}$.

The application, we have in mind, is the error $\delta f := \tilde{f} - f$ between the true function and the truncated version. Here, the control of the truncation ensures that $\|\delta f\|_{L^2} \le \varepsilon$ for an ε which is either fixed a priori or determined a posteriori. Concerning f, we may know that $|f|_m \le M_m$. However, this does not imply that some approximation $\tilde{f}$ has similar smoothness properties (in the extreme case, $\tilde{f}$ is chosen from a subspace not belonging to H^m, i.e., $|\delta f|_m = \infty$). In §4.3 we prove that, for the particular case of the HOSVD-based tensor truncation, at least the inequality

$$|\delta f|_m \le 2 |f|_m$$

holds, i.e., the error δf inherits the smoothness from f. Now, the results from §2.1 are applicable and relate the pointwise error $\|\delta f\|_\infty$ to the Euclidean truncation error $\|\delta f\|_{\ell^2/L^2}$. The results are gathered in §2.2 (see Theorem 2.1). We conclude that chapter by examples (§2.3) and comments to local estimates (§2.4).

To get inside into the constants c_m^Ω involved in the Gagliardo-Nirenberg inequality, we give an analysis in §5, where the constant is explicitly characterised. In particular, we determine the limit of c_m^Ω for $m \to \infty$ in the case of $\Omega = \mathbb{R}^d$.

Note that the aim of this paper is not an explicit determination of the pointwise error $\|\tilde{f} - f\|_\infty$. Such an attempt is usually impossible, since, in practice, the quantitative value of $|f|_m$ is not known. Instead, it serves as a justification that, even when $\|\cdot\|_\infty$ error estimates are desired, the HOSVD truncation based on $\|\cdot\|_{\ell^2/L^2}$ estimates makes perfect sense.

2 L^∞ estimates

2.1 Gagliardo-Nirenberg Inequality

We consider functions defined on the Cartesian product $\Omega = \Omega_1 \times \ldots \times \Omega_d$, where $\Omega_j \subset \mathbb{R}$. The model cases of the intervals Ω_j are $\mathbb{R}$, $[0, \infty)$, and $[0, 1]$. In principle, for each direction j another interval can be chosen, but we restrict ourselves to the model cases $\Omega = \mathbb{R}^d$, $\Omega = [0, \infty)^d$, and $\Omega = [0, 1]^d$. The basic Hilbert space is $\mathbf{V} = L^2(\Omega)$, which can be seen as the tensor space $\bigotimes_{j=1}^d V_j$ for $V_j = L^2(\Omega_j)$. In many practical applications, functions are replaced by grid functions, i.e., the underlying Hilbert spaces are of ℓ^2 type. This situation is discussed in §5.7. It turns out that the same results can be obtained as for $L^2(\mathbb{R}^d)$.

For sufficiently smooth functions we define the semi-norm

$$|\varphi|_m := \|\mathbf{D}^m \varphi\|_{L^2(\Omega)} := \sqrt{\int_\Omega \sum\nolimits_{j=1}^d \left| \frac{\partial^m \varphi}{\partial x_j^m} \right|^2 \mathrm{d}x}.$$

The Gagliardo-Nirenberg inequality states that

$$\|\varphi\|_\infty \leq \left\{ \begin{array}{ll} c_m^\Omega \, |\varphi|_m^{\frac{d}{2m}} \, \|\varphi\|_{L^2}^{1-\frac{d}{2m}} & \text{for } \Omega \in \left\{\mathbb{R}^d, [0,\infty)^d\right\} \\ c_m^\Omega \, \left[|\varphi|_m^2 + \|\varphi\|^2\right]^{\frac{d}{4m}} \, \|\varphi\|_{L^2}^{1-\frac{d}{2m}} & \text{for } \Omega = [0,1]^d \end{array} \right\} \qquad (2m > d)$$

(for $\Omega = \mathbb{R}^d$ this inequality is given[1] by Nirenberg [9, Theorem on page 125] and in the monograph of Maz'ja [8, Eq. (2.3.50)], when setting $q = \infty$, $j = 0$, $p = r = 2$, $\ell = m$, $n = d$). For completeness and in order to describe the constant c_m^Ω, we shall give the proof in §5. The condition $2m > d$ corresponds to the Sobolev embedding theorem (cf. [5, Theorem 6.2.30]).

For infinitely differentiable functions with $M_m := |\varphi|_m < \infty$ the asymptotic behaviour of M_m is of interest. If $\log(M_m) = O(m)$, there is some bound μ of $M_m^{1/m} \leq \mu$ and $\|\varphi\|_\infty \leq c_m^\Omega \, \mu^{\frac{d}{2}} \, \|\varphi\|_{L^2}^{1-\frac{d}{2m}}$ holds for all $m > d/2$. At least in the case of $\Omega = \mathbb{R}^d$, the limit $c_m^\Omega \to \pi^{-d/2}$ for $m \to \infty$ is known (cf. Lemma 5.10) and yields

$$\|\varphi\|_\infty \leq (\mu/\pi)^{d/2} \, \|\varphi\|_{L^2} . \tag{2.1}$$

A stronger increase of M_m holds, if $\log(M_m) = O(m(1+q\log m))$. In this case, $M_m^{1/m} \leq \mu^{1+q\log m} = \mu m^p$ with $p = q\log\mu$ follows and yields $\|\varphi\|_\infty \leq c_m^\Omega \mu^{d/2} m^{pd/2} \, \|\varphi\|_{L^2}^{1-\frac{d}{2m}}$ for all $m > d/2$. Assume that $\|\varphi\|_{L^2} < 1/\mathrm{e}$ and set $m^* := \log(1/\|\varphi\|_{L^2})$. Then we obtain

$$\|\varphi\|_\infty \leq c_{m^*}^\Omega \, [\mu \mathrm{e} \log^p(1/\|\varphi\|_{L^2})]^{d/2} \, \|\varphi\|_{L^2} ,$$

since $\|\varphi\|_{L^2}^{-\frac{d}{2m^*}} = \mathrm{e}^{d/2}$.

2.2 Application to Tensor Truncation

In §3.4 (for the tensor subspace format) and Lemma 4.1 (for the hierarchical format) we shall prove that the SDV-based tensor approximation $\tilde{f}$ to f leads to an error

$$\delta f := \tilde{f} - f,$$

which inherits its smoothness from f:

$$|\delta f|_m \leq c_F \, |f|_m \qquad \text{with } c_F = \left\{ \begin{array}{ll} \sqrt{2} & \text{for the tensor subspace format,} \\ 2 & \text{for the hierarchical format.} \end{array} \right. \tag{2.2}$$

The L^2 norm $\|\delta f\|_{L^2} \leq \varepsilon$ will be controlled by means of the singular values and can be assumed to be small. Applying the previous estimates to δf, we get the following result.

Theorem 2.1 *(a) If $|f|_m < \infty$ for some $m > d/2$, the error δf of the tensor truncation allows an estimate with respect to the supremum norm:*

$$\|\delta f\|_\infty \leq c_m^\Omega c_F^{\frac{d}{2m}} \, |f|_m^{\frac{d}{2m}} \, \|\delta f\|_{L^2}^{1-\frac{d}{2m}} , \tag{2.3a}$$

provided that $\Omega \in \left\{\mathbb{R}^d, [0,\infty)^d\right\}$. The analogous statements hold for $\Omega = [0,1]^d$ with $|\cdot|_m^2$ replaced by $|\cdot|_m^2 + \|\cdot\|^2$.
(b) If $|f|_m \lesssim \mu^m$ as $m \to \infty$, the estimate

$$\|\delta f\|_\infty \leq c^\Omega \mu^{d/2} \, \|\delta f\|_{L^2} \qquad \textit{with } c^\Omega := \liminf_{m\to\infty} c_m^\Omega \tag{2.3b}$$

is valid, where $c^\Omega = \pi^{-d/2}$ holds for $\Omega = \mathbb{R}^d$.
(c) If $|f|_m^{1/m} \leq \mu m^p$, the asymptotic behaviour for $\|\delta f\|_{L^2} \to 0$ is described by

$$\|\delta f\|_\infty \lesssim c^\Omega \left[\mu \mathrm{e} \log^p(\tfrac{1}{\|\delta f\|_{L^2}})\right]^{d/2} \|\delta f\|_{L^2} . \tag{2.3c}$$

It remains to prove (2.2), which will be done in §§3-4.

[1] The norm $|\cdot|_m$ is defined in [8] by all derivatives of order m, whereas here we use only the non-mixed derivatives.

2.3 Examples

The first example shows that estimate (2.3b) with $c^{\Omega} = \pi^{-d/2}$ for $\Omega = \mathbb{R}^d$ is sharp. We consider the product of sinc functions:

$$f(x) = \prod_{j=1}^{d} \frac{\sin(Ax_j)}{x_j} \qquad \text{with } A > 0. \tag{2.4}$$

The Fourier transform $\hat{f}$ is the constant $(\pi/2)^{d/2}$ for $|\xi_j| \le A$ $(1 \le j \le d)$ and zero, otherwise (cf. [11, §0.10]). This allows a simple computation of the sum of m-th derivatives:

$$M_m := |f|_m = \|\widehat{\mathbf{D}^m f}\|_{L^2} = (2m+1)^{-1/2}\,\pi^{d/2} A^{m+d/2}.$$

Since $M^{\frac{d}{2m}} \lesssim A^{d/2}$, (2.1) yields

$$\|f\|_\infty \le (A/\pi)^{d/2}\,\|f\|_{L^2}\,.$$

In fact, f from (2.4) satisfies the identity $\|f\|_\infty = (A/\pi)^{d/2}\,\|f\|_{L^2}$. Hence, the estimate from above is sharp.

Another important example is the Gaussian function $f(x) = \exp(-Ax^2)$ with $A > 0$. To estimate the derivative, we use the Fourier transform $\hat{f}(\xi) = (2A)^{-d/2}\exp(-\xi^2/(4A))$:

$$\begin{aligned}
M_m^2 = |f|_m^2 = \|\widehat{\mathbf{D}^m f}\|_{L^2}^2 &= (2A)^{-d}\sum_{j=1}^{d}\int_{\mathbb{R}^d}\xi_j^{2m}\exp(-\frac{1}{2A}\xi^2)\mathrm{d}\xi \\
&= (2A)^{-d}\sum_{j=1}^{d}\left(\Gamma(m+\tfrac{1}{2})\,(2A)^{m+\frac{1}{2}}\right)\left(\Gamma(\tfrac{1}{2})\sqrt{2A}\right)^{d-1} \\
&= d\,(2A)^{m-d/2}\,\pi^{(d-1)/2}\Gamma(m+\tfrac{1}{2}) \approx d\sqrt{2}\,(2A)^{m-d/2}\,\pi^{d/2}\,(m-1/2)^{m-1/2}\,\mathrm{e}^{-m+1/2},
\end{aligned}$$

where the last line follows by Stirling's formula. The asymptotic behaviour of $M_m^{\frac{d}{2m}}$ is $[(2m-1)\,A/\mathrm{e}]^{d/4}$, so that Theorem 2.1 applies with $\mu = \sqrt{2A/\mathrm{e}}$ and $p = 1/2$.

The last example is $f(x) = \prod_{j=1}^{d} 1/\cosh(x_j)$. Thanks to the Fourier transform $\hat{f}(\xi) = (\pi/2)^{\frac{d}{2}} f(\frac{\pi}{2}\xi)$ (cf. Oberhettinger [10, I.§7]) one obtains $|f|_m^{1/m} \le \mu m^p$ with $\mu = 1/\mathrm{e}$ and $p = 1$ in Theorem 2.1.

2.4 Generalisation to Local Estimates

The assumption of smoothness in the sense of $|\cdot|_m$ does not hold for functions with, e.g., pointwise singularities. In fact, one observes that large $\|\cdot\|_\infty$ errors occur close to the point singularity. However, outside a neighbourhood of the singularity, the $\|\cdot\|_\infty$ errors are under control.

Let $\Omega' := \Omega_1' \times \ldots \times \Omega_d'$ be a subset of Ω, which does not contain the singularity. The previous results can be restricted to Ω'. The practical problem is that the $L^2(\Omega)$ norm of the error is only known for Ω, but no better estimate for its $L^2(\Omega')$ norm. Using $\|\delta f\|_{L^2(\Omega')} \le \|\delta f\|_{L^2(\Omega)}$, we obtain the result

$$\|\delta f\|_{\infty,\Omega'} \le c_F c_m^{\Omega'}\,|f|_{m,\Omega'}^{\frac{d}{2m}}\,\|\delta f\|_{L^2(\Omega)}^{1-\frac{d}{2m}}\,,$$

where $\|\cdot\|_{\infty,\Omega'}$ and $|\cdot|_{m,\Omega'}$ are the respective (semi-)norms on Ω'.

It is not obvious whether variants of the inequality by weighted norms apply (cf. Caffarelli-Kohn-Nirenberg [2]).

3 Truncation for the Tensor Subspace Format

In the following, we recall the tensor subspace format (also called Tucker format) and the special bases used for the truncation procedure.

3.1 Format and HOSVD Bases

Let V_j be Hilbert spaces and $\mathbf{V} = \bigotimes_{j=1}^d V_j$ the Hilbert tensor space with induced scalar product[2] (cf. [6, §4.5.1]) and the corresponding norm $\|\cdot\|$. The tensor subspace format of a tensor from $\mathbf{V} = \bigotimes_{j=1}^d V_j$ is characterised by d 'ranks' $r_j \in \mathbb{N} \cup \{\infty\}$, which are gathered in the d-tuple $\mathbf{r} = (r_1, \ldots, r_d)$. The set of rank-$\mathbf{r}$ tensors is defined by

$$\mathcal{T}_{\mathbf{r}} := \left\{ \mathbf{v} \in \mathbf{V} \colon \text{ there are subspaces } U_j \subset V_j \text{ with } \dim(U_j) = r_j \text{ and } \mathbf{v} \in \bigotimes_{j=1}^d U_j \right\}. \tag{3.1}$$

Using basis vectors $b_\nu^{(j)}$, $1 \le \nu \le r_j$, of U_j, we can represent $\mathbf{v}$ by

$$\mathbf{v} = \sum_{i_1=1}^{r_1} \cdots \sum_{i_d=1}^{r_d} a[i_1, \ldots, i_d] \bigotimes_{j=1}^d b_{i_j}^{(j)} \tag{3.2}$$

with suitable coefficients $a[i_1, \ldots, i_d]$. Note that in the supposed Hilbert setting, even for non-separable Hilbert spaces, any tensor from $\mathbf{V}$ has a representation (3.2) with some $r_j \le \infty$.

The practical performance of the truncation as well as the theoretical analysis requires the so-called matricisations and their singular value decompositions (SVD), which in this context are called 'higher order SVD' (HOSVD). Set $D := \{1, \ldots, d\}$ and define

$$\mathbf{V}_\alpha := \bigotimes_{j \in \alpha} V_j \qquad \text{for subsets } \alpha \subset D. \tag{3.3}$$

Then $\mathbf{V}$ is isomorphic to $\mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$, $\emptyset \subsetneqq \alpha \subsetneqq D$, where $\alpha^c := D \backslash \alpha$ is the complement. We denote this isomorphism by $\mathcal{M}_\alpha$ and call it α-matricisation, since, for finite dimensions, $\mathcal{M}_\alpha(\mathbf{v}) \in \mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$ can be considered as a matrix. In general, for each α there is a singular value decomposition (HOSVD)

$$\begin{aligned} &\mathcal{M}_\alpha(\mathbf{v}) = \sum_{i=1}^{r_\alpha} \sigma_i^{(\alpha)} \mathbf{v}_i^{(\alpha)} \otimes \mathbf{v}_i^{(\alpha^c)}, \qquad (3.4)\\ &\sigma_1^{(\alpha)} \ge \sigma_2^{(\alpha)} \ge \ldots > 0, \quad \{\mathbf{v}_i^{(\alpha)}\} \subset \mathbf{V}_\alpha \text{ and } \{\mathbf{v}_i^{(\alpha^c)}\} \subset \mathbf{V}_{\alpha^c} \text{ orthonormal systems.} \end{aligned}$$

In the infinite dimensional case, $r_\alpha = \infty$ may occur. By definition (note that $\sigma_i^{(\alpha)} > 0$), the ranks r_α are the minimal integers in the equation from above. The α-rank $r_\alpha = \operatorname{rank}_\alpha(\mathbf{v})$ has already been introduced by Hitchcock [7].

The connection to the tensor subspace format is given, when we choose $\alpha = \{j\}$, $j \in D$. In this case, $\mathbf{v}_i^{(\alpha)} = v_i^{(j)}$ is a vector from V_j. For simplicity, we change the notation of the complementary part $\mathbf{v}_i^{(\alpha^c)}$ into $\mathbf{v}_i^{[j]}$:

$$\mathcal{M}_j(\mathbf{v}) = \sum_{i=1}^{r_j} \sigma_i^{(j)} v_i^{(j)} \otimes \mathbf{v}_i^{[j]}. \tag{3.5}$$

Then, $\mathbf{v}$ belongs to $\mathcal{T}_{\mathbf{r}}$ with $\mathbf{r} = (r_1, \ldots, r_d)$, $r_j = \operatorname{rank}_j(\mathbf{v})$, and there is no smaller $\mathbf{r}$ with $\mathbf{v} \in \mathcal{T}_{\mathbf{r}}$. In the sequel, we assume for simplicity, that $\mathbf{v} \in \mathcal{T}_{\mathbf{r}}$ is given with this minimal $\mathbf{r}$. The subspaces U_j appearing in (3.1) can be characterised by

$$U_j = \operatorname{span}\{v_i^{(j)} : 1 \le i \le r_j\}. \tag{3.6}$$

Therefore, the basis of U_j can be chosen as HOSVD basis: $b_i^{(j)} := v_i^{(j)}$.

[2] The induced scalar product in $\mathbf{V} \times \mathbf{V}$ is completely defined by $\langle \mathbf{v}, \mathbf{w} \rangle = \prod_{j=1}^d \langle v^{(j)}, w^{(j)} \rangle_{V_j}$ for elementary tensors $\mathbf{v} = \bigotimes_{j=1}^d v^{(j)}$ and $\mathbf{w} = \bigotimes_{j=1}^d w^{(j)}$.

3.2 Truncation

Let $\mathbf{s} \in \mathbb{N}^d$ be a d-tuple with entrywise inequality $\mathbf{s} \leq \mathbf{r}$ (the interesting case is $\mathbf{s} < \mathbf{r}$). Truncation of $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ to rank $\mathbf{s}$ means that we are looking for an approximation $\mathbf{w} \in \mathcal{T}_\mathbf{s}$. The HOSVD truncation follows the same lines as in the matrix case ($d = 2$). Assume that $\mathbf{v}$ is given by (3.2) with HOSVD bases $b_i^{(j)} := v_i^{(j)}$. Then define the truncated tensor by[3]

$$\mathbf{w} := \sum_{i_1=1}^{s_1} \cdots \sum_{i_d=1}^{s_d} a[i_1, \ldots, i_d] \bigotimes_{j=1}^{d} v_{i_j}^{(j)} \in \mathcal{T}_\mathbf{s} \,, \tag{3.7}$$

i.e., the upper summation bounds r_j are replaced by s_j. Differently from the matrix case, $\mathbf{w}$ is not necessarily the best approximation in $\mathcal{T}_\mathbf{s}$, but one can prove quasi-optimality:

$$\|\mathbf{v} - \mathbf{w}\| \leq \sqrt{\sum_{j=1}^{d} \sum_{i_j=s_j+1}^{r_j} \left(\sigma_{i_j}^{(j)}\right)^2} \leq \sqrt{d}\, \|\mathbf{v} - \mathbf{w}_{\mathrm{best}}\| \,, \tag{3.8}$$

where $\mathbf{w}_{\mathrm{best}} \in \mathcal{T}_\mathbf{s}$ is the best approximation (cf. [6, Theorem 10.3], [4, Lemma 2.6]). In the case of a function $f = \mathbf{v}$ and its approximation $\tilde{f} = \mathbf{w}$, the L^2 norm is under control: $\|\tilde{f} - f\|_{L^2} \leq \sqrt{\sum_{j=1}^{d} \sum_{i_j=s_j+1}^{r_j} (\sigma_{i_j}^{(j)})^2}$. Often, the new ranks s_j are chosen adaptively to ensure $\|\tilde{f} - f\|_{L^2} \leq \varepsilon$ for a given $\varepsilon > 0$.

The formal definition of the HOSVD truncation uses projections. For this purpose we consider the subspaces U_j from (3.6) and

$$U_j' := \mathrm{span}\{v_i^{(j)} : 1 \leq i \leq s_j\}, \quad U_j'' := \mathrm{span}\{v_i^{(j)} : s_j + 1 \leq i \leq r_j\}, \tag{3.9}$$

i.e., $U_j = U_j' \oplus U_j''$ and $U_j' \perp U_j''$. Let[4] $\Pi^{(j)} \in \mathcal{L}(V_j, V_j)$ be the orthogonal projection onto U_j'. Then

$$\mathbf{w} = \Pi \mathbf{v} \qquad \text{with } \Pi := \bigotimes_{j=1}^{d} \Pi^{(j)} \in \mathcal{L}(\mathbf{V}, \mathbf{V}) \tag{3.10}$$

yields the tensor from (3.7). The tensor product Π is the orthogonal projection on $\bigotimes_{j=1}^{d} U_j'$. The truncation error is described by the complementary projection $(I - \Pi)\,\mathbf{v}$.

3.3 General Estimate

First we consider (possibly unbounded) operators $A^{(j)}$ on[5] V_j. The extension of A_j to $\mathbf{V}$ is given by

$$\mathbf{A}_j := I \otimes I \otimes \ldots \otimes A^{(j)} \otimes \ldots \otimes I,$$

where $A^{(j)}$ appears at the j-th position. The following result depends essentially on the SVD nature of the truncation.

Lemma 3.1 *Let $\mathbf{w} = \Pi \mathbf{v}$ with Π as in (3.10), where $\mathbf{v}$ belongs to the domain of $\mathbf{A}_j$ defined above. Then*

$$\|\mathbf{A}_j\,(\mathbf{v} - \mathbf{w})\,\| \leq \sqrt{2}\, \|\mathbf{A}_j \mathbf{v}\|$$

holds with respect to the norm of $\mathbf{V}$.

---

[3] Truncation reduces the indices i_j to $\{1, \ldots, s_j\}$ corresponding to the largest singular eigenvalues. The following statements are true for the reduction to *any* index subset of $\{1, \ldots, r_j\}$.

[4] $\mathcal{L}(X, Y)$ denotes the space of bounded linear mappings from X into Y.

[5] $A^{(j)}$ may be a mapping from *domain*$(A^{(j)}) \subset V_j$ into V_j or into another Hilbert space W_j. In the latter case, the operator norm has to be changed accordingly.

Proof. Without loss of generality, we may consider $j = 1$. If $\mathbf{A}_1$ would commute with Π, the proof followed from $\|\mathbf{A}_1(\mathbf{v} - \mathbf{w})\| = \|\mathbf{A}_1(\mathbf{v} - \Pi\mathbf{v})\| = \|(I - \Pi)\mathbf{A}_1\mathbf{v}\| \le \|\mathbf{A}_1\mathbf{v}\|$. However, in general, the operators do not commute.

Set $\Pi_j := I \otimes I \otimes \ldots \otimes \Pi^{(j)} \otimes \ldots \otimes I$. These projections are mutually commutative and their product (in any order) yields Π. Since $\mathbf{A}_1 = A^{(1)} \otimes I \otimes \ldots \otimes I$, $\mathbf{A}_1\Pi_j = \Pi_j\mathbf{A}_1$ holds for all $j \ge 2$, so that

$$\mathbf{A}_1\mathbf{w} = \mathbf{A}_1\Pi_2\Pi_3\cdots\Pi_d\Pi_1\mathbf{v} = \Pi_2\Pi_3\cdots\Pi_d\mathbf{A}_1\Pi_1\mathbf{v}.$$

Now we use a special property of the SDV (3.5). The projection Π_1 annihilates all $v_i^{(1)}$ for $i > s_1$. The same result can be obtained by $\Pi_{[1]} := I \otimes \Pi^{[1]}$, where $\Pi^{[1]} \in \mathcal{L}(\mathbf{V}^{[1]}, \mathbf{V}^{[1]})$ is the orthogonal projection on $\mathrm{span}\{\mathbf{v}_i^{[1]} : 1 \le i \le s_1\}$ ($\mathbf{v}_i^{[j]}$ from (3.5)). Note that $\Pi_{[1]}$ and Π_1 are different projections, but their application to the special tensor $\mathbf{v}$ yields the same result: $\Pi_1\mathbf{v} = \Pi_{[1]}\mathbf{v}$. The operators $\mathbf{A}_1 = A^{(1)} \otimes I \otimes \ldots \otimes I$ and $I \otimes \Pi^{[1]}$ commute. Therefore, we can continue the previous equation,

$$\mathbf{A}_1\Pi\mathbf{v} = \mathbf{A}_1\mathbf{w} = \Pi_2\Pi_3\cdots\Pi_d\mathbf{A}_1\Pi_1\mathbf{v} = \Pi_2\Pi_3\cdots\Pi_d\mathbf{A}_1\Pi_{[1]}\mathbf{v} = \Pi_2\Pi_3\cdots\Pi_d\Pi_{[1]}\mathbf{A}_1\mathbf{v},$$

and obtain

$$\|\mathbf{A}_1(\mathbf{v} - \Pi\mathbf{v})\| = \|(I - \Pi_2\Pi_3\cdots\Pi_d\Pi_{[1]})\mathbf{A}_1\mathbf{v}\| \le \|I - \Pi_2\Pi_3\cdots\Pi_d\Pi_{[1]}\|\|\mathbf{A}_1\mathbf{v}\|.$$

Note that $P_1 := \Pi_2\Pi_3\cdots\Pi_d$ as well as $P_2 := \Pi_{[1]}$ are orthogonal projections. By

$$\|I - P_1P_2\|^2 = \|(I - P_1) + P_1(I - P_2)\|^2 = \|I - P_1\|^2 + \|P_1(I - P_2)\|^2 = 1 + 1 = 2$$

(cf. [6, Lemma 4.123]), the assertion follows. ■

If one is only interested in the properties of $\mathbf{w}$, the same proof shows the following result.

Corollary 3.2 $\|\mathbf{A}_j\mathbf{w}\| \le \|\mathbf{A}_j\mathbf{v}\|$.

There are interesting conclusions from this statement, different from those in the next section; for instance, (a) if $\mathbf{v}$ is a function belonging to the weighted space L^2_Φ with $\Phi = \sum_{j=1}^d \phi_j^2$, then also $\mathbf{w}$ belongs to this space with the same bound; (b) if $\mathbf{v}$ is a multivariate function such that $\mathbf{v}$ as a function of the real variable x_j can be extended holomorphically into the complex domain $\omega_j \subset \mathbb{C}$, this is also true for $\mathbf{w}$ with the same bounds.

Remark 3.3 *There is a sequential modification of the truncation (see [6, §10.1.2]). The statements from above are also valid for this version. The proofs are similar.*

3.4 Smoothness of the Error

Now we consider the function spaces $V_j = L^2(\Omega_j)$, $\Omega_j \subset \mathbb{R}$, and $\mathbf{V} = \bigotimes_{j=1}^d V_j = L^2(\Omega)$. Instead of the notations $\mathbf{v}$ and $\mathbf{w} = \Pi\mathbf{v}$ we use f and $\tilde{f} := \Pi f$. The choice $A^{(j)} = \partial^m/\partial x_j^m$ in Lemma 3.1 yields

$$|f - \tilde{f}|_m^2 = \sum_{j=1}^d \|\mathbf{A}_j(I - \Pi)f\|_{L^2}^2 \le 2\sum_{j=1}^d \|\mathbf{A}_jf\|_{L^2}^2 = 2|f|_m^2.$$

This proves statement (2.2) with $c_F = \sqrt{2}$.

4 Truncation for the Hierarchical Format

The modern applications are based on the hierarchical format. We briefly repeat its definition in §4.1 (cf. [6, §11]) and describe the truncation in §4.2 (cf. [6, §11.4.2]).

4.1 Hierarchical Format

While in the tensor subspace format the set $D = \{1, \ldots, d\}$ is immediately separated into the single directions $\{j\}$, this splitting is now performed via a binary dimension splitting tree T_D. The vertices of T_D are non-empty subsets of D. The root of the tree is D. If $\alpha \in T_D$ is a vertex with $\#\alpha > 1$, it possesses sons $\alpha_1, \alpha_2 \in T_D$ such that $\alpha = \alpha_1 \cup \alpha_2$ is a disjoint union. Leaves of T_D are characterised by $\#\alpha = 1$, i.e., the set of leaves is given by

$$\mathcal{L}(T_D) := \{\alpha = \{j\} : j \in D\}.$$

As in the tensor subspace format, the tensor representation uses subspaces $U_j \subset V_j$, but the use of subspaces is iterated in the tree. If, e.g., $\alpha = \{1, 2\}$ is a vertex of T_D corresponding to the tensor space $\mathbf{V}_\alpha$ (cf. (3.3)), the subspaces U_j at the leaves yield a subspace $\bigotimes_{j\in\alpha} U_j \subset \mathbf{V}_\alpha$, whose dimension is $\prod_{j\in\alpha} \dim(U_j)$, i.e., the dimension is much higher than the single dimensions $\dim(U_j)$. Again, we look for a suitable subspace $\mathbf{U}_\alpha \subset \bigotimes_{j\in\alpha} U_j \subset \mathbf{V}_\alpha$ of smaller dimension. In general, each vertex $\alpha \in T_D \backslash \mathcal{L}(T_D)$ is associated to a subspace $\mathbf{U}_\alpha$ satisfying

$$\mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2} \qquad (\alpha_1, \alpha_2 \text{ sons of } \alpha, \text{i.e., } \alpha = \alpha_1 \dot\cup \alpha_2). \tag{4.1}$$

The subspaces must be chosen such that $\mathbf{v} \in \mathbf{U}_D$ holds for the tensor to be represented. It suffices to set $\mathbf{U}_D = \operatorname{span}\{\mathbf{v}\}$.

Again, each subspace is represented by a basis $\{\mathbf{b}_i^{(\alpha)} : 1 \le i \le r_\alpha\}$, where $r_\alpha := \dim(\mathbf{U}_\alpha)$ is called the α-rank. The truncation procedure is based on the singular vectors $\mathbf{b}_i^{(\alpha)} = \mathbf{v}_i^{(\alpha)}$ from the HOSVD (3.4). In fact, the subspace spanned by the singular vectors $\mathbf{v}_i^{(\alpha)}$ yields the smallest r_α.

For theoretical considerations we use the bases $\{\mathbf{b}_i^{(\alpha)}\}$, but these are not suited for the practical representation. Here, we exploit the nestedness property (4.1). If $\{\mathbf{b}_i^{(\alpha_1)}\}$ and $\{\mathbf{b}_j^{(\alpha_2)}\}$ are bases of $\mathbf{U}_{\alpha_1}$ and $\mathbf{U}_{\alpha_2}$, respectively, the basis $\{\mathbf{b}_\ell^{(\alpha)}\}$ has a representation of the form

$$\mathbf{b}_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\ell,\alpha)} \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}. \tag{4.2}$$

Therefore, it suffices to store the coefficient matrices $C^{(\ell,\alpha)} = (c_{ij}^{(\ell,\alpha)})$, $1 \le \ell \le r_\alpha$. We remark that the computation of the HOSVD data can be achieved such that only the matrices $C^{(\ell,\alpha)}$ are involved (cf. [6, §11.3.3]).

4.2 Truncation

The HOSVD truncation is similar as in (3.7). Given smaller ranks $s_\alpha \le r_\alpha$, we restrict all sums in (4.2) to the smaller upper bound s_α. More precisely, $C^{(\ell,\alpha)}$ is omitted for $\ell > s_\alpha$, while the other $C^{(\ell,\alpha)} \in \mathbb{R}^{r_{\alpha_1} \times r_{\alpha_2}}$ are restricted to the size $s_{\alpha_1} \times s_{\alpha_2}$.

The mathematical description uses the orthogonal projections $\Pi^{(\alpha)} \in \mathcal{L}(\mathbf{V}_\alpha, \mathbf{V}_\alpha)$ onto the subspaces $\mathbf{U}'_\alpha$ defined by

$$\mathbf{U}'_\alpha := \operatorname{span}\{\mathbf{v}_i^{(\alpha)} : 1 \le i \le s_\alpha\}, \quad \mathbf{U}''_\alpha := \operatorname{span}\{\mathbf{v}_i^{(\alpha)} : s_\alpha + 1 \le i \le r_\alpha\} \tag{4.3}$$

(cf. (3.9)), where $\mathbf{v}_i^{(\alpha)}$ are the HOSVD basis vectors. $\Pi^{(\alpha)}$ is extended to $\mathcal{L}(\mathbf{V}, \mathbf{V})$ by $\Pi_\alpha := \Pi^{(\alpha)} \otimes I$ (I identity on $\mathbf{V}_{\alpha^c}$). Differently from the tensor subspace case in §3.3, these projections do not commute in general. The factors in the product

$$\Pi := \prod_{\alpha \in T_D} \Pi_\alpha \tag{4.4}$$

must be ordered such that Π_α is applied before Π_{α_1} and Π_{α_2} (α_1, α_2 sons of α) follow.

In analogy to (3.8), the estimate[6]

$$\|\mathbf{v} - \Pi\mathbf{v}\| \le \sqrt{\sum_\alpha \sum_{i_\alpha > s_\alpha} (\sigma_{i_\alpha}^{(\alpha)})^2} \le \sqrt{2d-3}\, \|\mathbf{v} - \mathbf{w}_{\text{best}}\|$$

[6]The sum $\sum_\alpha$ is taken over all $\alpha \in T_D$ except $\alpha = D$ and one of the sons of D. Hence, the sum contains $2d - 3$ terms.

holds (cf. [6, Theorem 11.58], [4, Remark 3.12]). $\mathbf{w}_{\text{best}}$ is the best approximation in the hierarchical format with ranks s_α.

4.3 General Estimate

As in §3.3, we want to estimate $\mathbf{A}_j(\mathbf{v} - \Pi\mathbf{v})$ by means of $\mathbf{A}_j\mathbf{v}$. Again, it suffices to consider the case $j = 1$. The projections Π_α with $1 \notin \alpha$ contain the identity operator at position 1; hence, $\mathbf{A}_1$ commutes with Π_α. The difficulties arise from Π_α with $1 \in \alpha$. In §3.3, we have replaced Π_α by another projection Π_{α^c} acting on the singular vectors $\mathbf{v}_i^{(\alpha^c)}$ from (3.4). This approach does not work in the present case (only for $\alpha = \{1\}$). Therefore, we apply another construction.

For the following artificial construction let $\hat{V}_j$ be an isomorphic (but disjoint) copy of V_j and embed the Hilbert space V_j into the direct sum $X_j := V_j \oplus \hat{V}_j$, which is again a Hilbert space with $V_j \perp \hat{V}_j$. Correspondingly, the Hilbert tensor space $\mathbf{V}$ is embedded into $\mathbf{X} = \bigotimes_{j\in D} X_j$. The isomorphism $\phi_j : V_j \to \hat{V}_j$ gives rise to isomorphisms $\Phi_\alpha : \mathbf{V}_\alpha = \bigotimes_{j\in\alpha} V_j \to \hat{\mathbf{V}}_\alpha = \bigotimes_{j\in\alpha} \hat{V}_j$.

In the following, several projections will appear:

$$\begin{aligned} &\Pi^{(\alpha)} \in \mathcal{L}(\mathbf{X}_\alpha, \mathbf{X}_\alpha) \text{ orthogonal projection onto } \mathbf{U}'_\alpha, \\ &\Pi_\alpha = \Pi^{(\alpha)} \otimes I^{(\alpha^c)} \qquad (I^{(\alpha^c)}\text{: identity on } \mathbf{X}_{\alpha^c}), \end{aligned} \tag{4.5a}$$

where $\mathbf{U}'_\alpha$ from (4.3) corresponds to the first tensors $\mathbf{v}_i^{(\alpha)}$ in (3.4). In the case of $\Pi^{(\alpha^c)}$, the involved subspace $\mathbf{U}'_{\alpha^c}$ corresponds to the second tensors $\mathbf{v}_i^{(\alpha^c)}$ in (3.4). In the special case of $\alpha = \{j\}$, $1 \leq j \leq d$, we write

$$\begin{aligned} &\Pi^{(j)} \in \mathcal{L}(X_j, X_j) \text{ orthogonal projection onto } U'_j, \\ &\Pi_j = I \otimes I \otimes \ldots \otimes \Pi^{(j)} \otimes \ldots \otimes I. \end{aligned} \tag{4.5b}$$

Furthermore, we introduce

$$\begin{aligned} &\Pi_{V_j} \in \mathcal{L}(X_j, X_j) \text{ orthogonal projection onto } V_j, \\ &\Pi_{\mathbf{V}} = I \otimes \bigotimes_{j=2}^{d} \Pi_{V_j} \text{ orthogonal projection onto } X_1 \otimes V_2 \otimes \ldots \otimes V_d. \end{aligned} \tag{4.5c}$$

which will eliminate all artificial contributions in $\hat{V}_j$.

The vertices α containing 1 can be ordered linearly:[7] $\alpha_0 := D \subset \alpha_1 \subset \ldots \subset \alpha_{L-1} \subset \alpha_L = \{1\}$, where α_{i+1} is a son of α_i. A possible arrangement of the factors in (4.4) is

$$\Pi = (\Pi_2\Pi_3 \cdots \Pi_d)\, \Pi' \left(\Pi_{\alpha_L}\Pi_{\alpha_{L-1}} \cdots \Pi_{\alpha_1}\right) \quad \text{with } \Pi' := \prod_{1\notin\alpha\in T_D\backslash\mathcal{L}(T_D)} \Pi_\alpha \tag{4.6}$$

with suitable ordering of the factors in Π'. The operator $\mathbf{A}_1 = A^{(1)} \otimes I^{[1]}$ commutes with $(\Pi_2\Pi_3 \cdots \Pi_d)$ and Π', since these projections are of the form $I \otimes \ldots$

In §3.3 we have replaced $\Pi_{\alpha_1} = \Pi^{(\alpha_1)} \otimes I$ by $I \otimes \Pi^{(\alpha_1^c)}$. Now we choose a unitary mapping Q_1 with the property[8]

$$Q_1\mathbf{x} = \mathbf{x} \text{ for } \mathbf{x} \perp \mathbf{U}''_{\alpha_1^c} \quad \text{and} \quad Q_1\mathbf{x} = \Phi_{\alpha_1^c}\mathbf{x} \in \hat{\mathbf{V}}_{\alpha_1^c} \text{ for } \mathbf{x} \in \mathbf{U}''_{\alpha_1^c},$$

($\mathbf{U}''_{\alpha_1^c}$ from (4.3), $\Phi_{\alpha_1^c} \in \mathcal{L}(\mathbf{V}_{\alpha_1^c}, \hat{\mathbf{V}}_{\alpha_1^c})$: isomorphism introduced above) and set

$$Q_{\alpha_1} := I \otimes Q_1.$$

The result of $\mathbf{v}_1 := Q_{\alpha_1}\mathbf{v}$ ($\mathbf{v}$ from (3.4)) is $\mathbf{v}' + \mathbf{v}''$, where $\mathbf{v}' = \Pi_{\alpha_1}\mathbf{v}$ and $\mathbf{v}'' = \Phi_{\alpha_1^c}(I - \Pi_{\alpha_1})\mathbf{v}$, i.e., the part $(I - \Pi_{\alpha_1})\mathbf{v}$ is not omitted but moved into the orthogonal complement $\hat{\mathbf{V}} \subset \mathbf{X}$ of $\mathbf{V}$. Note that $\Pi_{\alpha_1}\mathbf{v} = \Pi_{\mathbf{V}}\mathbf{v}_1$ and that $\Pi_{\mathbf{V}}$ commutes with all projections Π_j and Π' (cf. (4.6)).

[7] $\Pi_{\alpha_0} = \Pi_D$ can be omitted, since $\Pi_D = I$ because of $s_D = r_D = 1$.

[8] Note that Q_1 is not only the identity on $\mathbf{U}'_{\alpha_1^c}$, but also on $\hat{\mathbf{V}}_{\alpha_1^c}$.

We recall that α_2 is a son of α_1, i.e., $\alpha_2 \subset \alpha_1$. Thanks to the construction $\mathbf{v}_1 = Q_{\alpha_1}\mathbf{v}$, $\mathbf{v}_1$ possesses an HOSVD

$$\mathcal{M}_{\alpha_2}(\mathbf{v}_1) = \sum_{i=1}^{r_{\alpha_2}} \sigma_i^{(\alpha_2)} \mathbf{v}_i^{(\alpha_2)} \otimes \check{\mathbf{v}}_i^{(\alpha_2^c)}$$

with the same singular values $\sigma_i^{(\alpha_2)}$ and same singular vectors $\mathbf{v}_i^{(\alpha_2)}$ as in (3.4), only $\check{\mathbf{v}}_i^{(\alpha_2^c)}$ is changed:

$$\check{\mathbf{v}}_i^{(\alpha_2^c)} = (I \otimes Q_1)\mathbf{v}_i^{(\alpha_2^c)},$$

where I is the identity on $\mathbf{X}_{\alpha_2^c \backslash \alpha_1^c}$. Since $I \otimes Q_1$ is unitary, $\{\check{\mathbf{v}}_i^{(\alpha_2^c)}\}$ is still orthonormal! Since all $\mathbf{v}_i^{(\alpha_2)}$ are unchanged, $\Pi_{\alpha_2}\mathbf{v}_1 = (\Pi^{(\alpha_2)} \otimes I)\mathbf{v}_1 = (I \otimes \hat{\Pi}^{(\alpha_2^c)})\mathbf{v}_1$ holds, where now $\hat{\Pi}^{(\alpha_2^c)}$ is the orthogonal projection onto $\mathbf{U}''_{\alpha_2^c} := \mathrm{span}\{\hat{\mathbf{v}}_i^{(\alpha_2^c)} : s_{\alpha_2} + 1 \leq i \leq r_{\alpha_2}\}$. Again, we replace the projection $\hat{\Pi}^{(\alpha_2^c)}$ by a unitary map with the properties $Q_2 = \Phi_{\alpha_2^c} : \mathbf{U}''_{\alpha_2^c} \to \hat{\mathbf{V}}_{\alpha_2^c}$ and $Q_2 = I$ on $(\mathbf{U}''_{\alpha_2^c})^\perp$ and define

$$Q_{\alpha_2} := I \otimes Q_2 \qquad (I\text{: identity on } \mathbf{X}_{\alpha_2}).$$

Then, $\Pi_{\alpha_2}\mathbf{v}_1 = \Pi_{\mathbf{V}} Q_{\alpha_2}\mathbf{v}_1$ holds. Together, we obtain $\Pi_{\alpha_2}\Pi_{\alpha_1}\mathbf{v} = \Pi_{\mathbf{V}} Q_{\alpha_2} Q_{\alpha_1}\mathbf{v}$

In the same way, we can construct unitary mappings $Q_{\alpha_\nu} = (I \otimes Q_\nu)$, $Q_\nu \in \mathcal{L}(\mathbf{X}_{\alpha_\nu^c}, \mathbf{X}_{\alpha_\nu^c})$, such that

$$\Pi_{\alpha_L}\Pi_{\alpha_{L-1}} \cdots \Pi_{\alpha_1}\mathbf{v} = \Pi_{\mathbf{V}} Q_{\alpha_L} Q_{\alpha_{L-1}} \cdots Q_{\alpha_1}\mathbf{v}.$$

and $\Pi\mathbf{v} = (\Pi_2\Pi_3 \cdots \Pi_d)\,\Pi' \left(\Pi_{\alpha_L}\Pi_{\alpha_{L-1}} \cdots \Pi_{\alpha_1}\right)\mathbf{v} = \Pi_{\mathbf{V}}\,(\Pi_2\Pi_3 \cdots \Pi_d)\,\Pi' Q_{\alpha_L} Q_{\alpha_{L-1}} \cdots Q_{\alpha_1}\mathbf{v}$. By $U_j \subset V_j$, $\Pi_{\mathbf{V}}\,(\Pi_2\Pi_3 \cdots \Pi_d) = \Pi_2\Pi_3 \cdots \Pi_d$ holds.

Note that all operators in the product $(\Pi_2\Pi_3 \cdots \Pi_d)\,\Pi' Q_{\alpha_L} Q_{\alpha_{L-1}} \cdots Q_{\alpha_1}$ are of the form $I \otimes \cdots$, where I is the identity on X_1. This proves

$$\mathbf{A}_1\Pi\mathbf{v} = (\Pi_2\Pi_3 \cdots \Pi_d)\,\Pi' Q_{\alpha_L} Q_{\alpha_{L-1}} \cdots Q_{\alpha_1}\mathbf{A}_1\mathbf{v}$$

and $\mathbf{A}_1(\mathbf{v} - \Pi\mathbf{v}) = [I - (\Pi_2\Pi_3 \cdots \Pi_d)\,\Pi' Q_{\alpha_L} Q_{\alpha_{L-1}} \cdots Q_{\alpha_1}]\mathbf{A}_1\mathbf{v}$. As the operator norm of the bracket is ≤ 2, we have shown the following counterpart of Lemma 3.1.

Lemma 4.1 *Let $\mathbf{w} = \Pi\mathbf{v}$ with Π as in (4.4), and $\mathbf{A}_j$ as in §3.3. Then*

$$\|\mathbf{A}_j(\mathbf{v} - \mathbf{w})\| \leq 2\|\mathbf{A}_j\mathbf{v}\|$$

holds.

As in §3.4, we obtain from Lemma 4.1 the statement (2.2) with $c_F = 2$. Also in the case of the hierarchical format Corollary 3.2 is valid: $\|\mathbf{A}_j\mathbf{w}\| \leq \|\mathbf{A}_j\mathbf{v}\|$.

Remark 3.3 has the following counterpart.

Remark 4.2 *There are two different sequential modifications of the truncation in the hierarchical format (see [6, §11.4.2.2 and §11.4.2.3]). The statements from above are also valid for these versions.*

5 Analysis of the Gagliardo-Nirenberg Inequality

This chapter is considered as an appendix. The emphasis lies on the concrete characterisation of the constants c_m^Ω appearing in the Gagliardo-Nirenberg inequality. Furthermore, we determine $\lim_{m\to\infty} c_m^\Omega$ for the case $\Omega = \mathbb{R}^d$.

5.1 Notations

Let $\Omega = \Omega_1 \times \ldots \times \Omega_d$ with $\Omega_j \subset \mathbb{R}$. The scalar product of $L^2(\Omega)$ is denoted by[9] $(u, v) = \int_\Omega uv\mathrm{d}x$, the corresponding norm is written as $\|\cdot\|$. For $m \in \mathbb{N}$, we define the bilinear form and semi-norm

$$\langle u, v\rangle_m := \sum_{j=1}^{d} \left(\frac{\partial^m u}{\partial x_j^m}, \frac{\partial^m v}{\partial x_j^m}\right), \qquad |u|_m := \sqrt{\langle u, u\rangle_m}\,.$$

[9] For simplicity, the field $\mathbb{R}$ is assumed.

For positive numbers α, β, we define the bilinear form

$$a(u,v) := a^{\Omega}_{m,\alpha,\beta}(u,v) := \alpha^2 \cdot \langle u, v\rangle_m + \beta^2 \cdot (u,v)\,.$$

The corresponding norm is denoted by

$$|||u||| = |||u|||^{\Omega}_{m,\alpha,\beta} := \sqrt{a(u,u)}\,.$$

$|||u|||^{\Omega}_{m,\alpha,\beta}$ for different $\alpha, \beta > 0$ are equivalent norms of the Sobolev space $H^m(\Omega)$. The Sobolev embedding theorem ensures $H^m(\Omega) \subset C(\Omega)$ for $m > d/2$ (cf. [5, Theorem 6.2.30]), i.e., $\|\cdot\|_\infty \le \gamma \cdot ||| \cdot |||$ holds, where $\|\cdot\|_\infty$ is the supremum norm of $C(\Omega)$. We set

$$\gamma = \gamma^{\Omega}_{m,\alpha,\beta} := \sup\{\|u\|_\infty / |||u|||^{\Omega}_{m,\alpha,\beta} : 0 \neq u \in H^m(\Omega)\}. \tag{5.1}$$

5.2 Green's Function

As a consequence of $H^m(\Omega) \subset C(\Omega)$ for $m > d/2$, the Dirac functional δ_ξ ($\xi \in \Omega$) with $\delta_\xi(u) = u(\xi)$ belongs to $H^m(\Omega)'$. The Green function $G_\xi = G(\cdot,\xi) = G^{\Omega}_{m,\alpha,\beta}(\cdot,\xi)$ is the solution of the variational formulation

$$a^{\Omega}_{m,\alpha,\beta}(G_\xi, v) = v(\xi) \qquad \text{for all } v \in H^m(\Omega) \text{ and a fixed } \xi \in \Omega. \tag{5.2}$$

We gather some trivial facts in the following lemma.

Lemma 5.1 *(a) $G_\xi \in H^m(\Omega)$,*

(b) $G(\xi,\xi) = |||G_\xi|||^2 > 0$,

(c) $|u(\xi)| \le \sqrt{G(\xi,\xi)}|||u|||$ for all $u \in V$ and the maximum of the ratio $|u(\xi)| / |||u|||$ is taken for $u = G_\xi$.

(d) γ from (5.1) satisfies $\gamma^2 = \sup\{|G(x,y)| : x,y \in \Omega\} = \sup\{G(\xi,\xi) : \xi \in \Omega\}$.

(e) $G(x,y) = G(y,x)$,

Proof. 1) $\delta_\xi \in H^m(\Omega)'$ implies $G_\xi \in H^m(\Omega)$.
2) $v = G_\xi$ in (5.2) yields $G(\xi,\xi) = a(G_\xi, G_\xi) > 0$.
3) $|u(\xi)| =_{(5.2)} |a(G_\xi, u)| \le |||G_\xi||| |||u||| =_{(\mathrm{b})} \sqrt{G(\xi,\xi)}|||u|||$. Equality holds for $u = G_\xi$.
4) $u = G_y$ and $\xi = x$ in (c) yield $|G(x,y)| \le \sqrt{G(x,x)}|||G_y||| =_{(\mathrm{b})} \sqrt{G(x,x)G(y,y)}$, i.e., $|G(x,y)| \le \max\{G(x,x), G(y,y)\}$ and the supremum is taken along the diagonal $\{(\xi,\xi) : \xi \in \Omega\}$.
5) $G(x,y) = G_y(x) =_{(5.2)} a(G_x, G_y) = a(G_y, G_x) =_{(5.2)} G_x(y) = G(y,x)$. ∎

5.3 $\Omega = \mathbb{R}^d$ and $[0,\infty)^d$

First, we consider the case $\Omega = \mathbb{R}^d$. The translation operator T_δ ($\delta \in \mathbb{R}^d$) is defined by $(T_\delta u)(x) = u(x+\delta)$. T_δ is unitary in $L^2(\Omega)$: $T_\delta^* = T_\delta^{-1} = T_{-\delta}$, and the bilinear form $a^{\mathbb{R}^d}_{m,\alpha,\beta}$ satisfies

$$a(T_\delta u, v) = a(u, T_{-\delta} v). \tag{5.3}$$

Conclusion 5.2 *Under assumption (5.3), G depends only on the difference of its arguments: $G(x,y) = G(x-y)$. In particular, $G(x,x) = G(y,y)$ for all $x,y \in \mathbb{R}^d$, and γ from Lemma 5.1d can be defined by*

$$\gamma = \sqrt{G(0,0)}.$$

Proof. We have to show that $G(x,y) = G(x+\delta, y+\delta)$. Use $G(x+\delta, y+\delta) = (T_\delta G_{y+\delta})(x) = a(G_x, T_\delta G_{y+\delta}) = a(T_\delta G_{y+\delta}, G_x) = a(G_{y+\delta}, T_{-\delta} G_x) = (T_{-\delta} G_x)(y+\delta) = G_x(y) = G(y,x) = G(x,y)$. ∎

Lemma 5.3 *Also in the case of $\Omega = [0,\infty)^d$, the maximum is taken at $\xi = 0$: $\gamma^2 = \sup_{\xi\in\Omega} G(\xi,\xi) = G(0,0)$.*

Proof. Assume that for some $0 \neq \xi \geq 0$ (pointwise inequality), $G(\xi,\xi) > G(0,0)$ holds. Define $g(x) := G(x+\xi,\xi)$ and note that $g(0) = G(\xi,\xi)$. The squared norm $|||g|||^2$ is an integral over $[0,\infty)^d$ and equal to the integral over $[\xi_1,\infty) \times \ldots \times [\xi_d,\infty)$ with g replaced by G_ξ. Obviously, the latter integral is not larger than $|||G_\xi|||^2$, i.e., $|||g||| \leq |||G_\xi|||$.

By Lemma 5.1c,

$$\frac{G(\xi,\xi)}{|||g|||} = \frac{|g(0)|}{|||g|||} \leq \frac{G(0,0)}{|||G(0,0)|||} = \sqrt{G(0,0)}$$

holds, while the previous inequality yields the contradiction $\frac{G(\xi,\xi)}{|||g|||} \geq \frac{G(\xi,\xi)}{|||G_\xi|||} = \sqrt{G(\xi,\xi)} > \sqrt{G(0,0)}$. ∎

5.4 Dilatations

For $\Omega \in \{[0,\infty)^d, \mathbb{R}^d\}$ we can define[10] the dilatation M_λ $(\lambda > 0)$ centred at $0 \in \mathbb{R}^d$ by

$$(M_\lambda u)(x) := u(\lambda x) \qquad (x \in \Omega).$$

Because of $(M_\lambda u, M_\lambda v) = \lambda^{-d/2}(u,v)$ (substitution rule) and $\frac{\partial^m}{\partial x_j^m}(M_\lambda u) = \lambda^m M_\lambda \frac{\partial^m}{\partial x_j^m} u$ (chain rule), we obtain

$$a^\Omega_{m,\alpha,\beta}(M_\lambda u, M_\lambda v) = a^\Omega_{m,\alpha\lambda^{m-d/2},\beta\lambda^{-d/2}}(u,v) = \lambda^{-d} a^\Omega_{m,\alpha\lambda^m,\beta}(u,v). \tag{5.4}$$

Choosing $\lambda^m = \beta/\alpha$, we can relate $a^\Omega_{m,\alpha,\beta}$ to $a^\Omega_{m,1,1}$:

$$a^\Omega_{m,\alpha,\beta}(M_{(\beta/\alpha)^{1/m}} u, M_{(\beta/\alpha)^{1/m}} v) = (\beta/\alpha)^{-d/m} a^\Omega_{m,\beta,\beta}(u,v) = \beta^2 (\beta/\alpha)^{-d/m} a^\Omega_{m,1,1}(u,v)$$

or equivalently

$$a^\Omega_{m,\alpha,\beta}(u,v) = \beta^2(\beta/\alpha)^{-d/m} a^\Omega_{m,1,1}(M_{(\beta/\alpha)^{-1/m}} u, M_{(\beta/\alpha)^{-1/m}} v).$$

Lemma 5.4 *The Green functions $G^\Omega_{m,\alpha,\beta}$ belonging to $a^\Omega_{m,\alpha,\beta}$ allow a similar relation:*

$$G^\Omega_{m,\alpha,\beta}(x,y) = \beta^{-2}(\beta/\alpha)^{d/m} G^\Omega_{m,1,1}((\beta/\alpha)^{1/m}x, (\beta/\alpha)^{1/m}y). \tag{5.5}$$

Proof. Set $u(x) = \beta^{-2}(\beta/\alpha)^{d/m} G^\Omega_{m,1,1}((\beta/\alpha)^{1/m}x, (\beta/\alpha)^{1/m}y)$ and test with some function v:

$$\begin{aligned} a^\Omega_{m,\alpha,\beta}(u,v) &= a^\Omega_{m,1,1}(M_{(\beta/\alpha)^{-1/m}} G^\Omega_{m,1,1}((\beta/\alpha)^{1/m}\bullet, (\beta/\alpha)^{1/m}y), M_{(\beta/\alpha)^{-1/m}} v) \\ &= a^\Omega_{m,1,1}(G^\Omega_{m,1,1}(\bullet, (\beta/\alpha)^{1/m}y), M_{(\beta/\alpha)^{-1/m}} v) = \delta_{(\beta/\alpha)^{1/m}y}(M_{(\beta/\alpha)^{-1/m}} v) \\ &= (M_{(\beta/\alpha)^{-1/m}} v)((\beta/\alpha)^{1/m}y) = v(y) = \delta_y(v), \end{aligned}$$

i.e., u satisfies the variational problem defining $G^\Omega_{m,\alpha,\beta}(\cdot,y)$. ∎

Since the supremum norm is invariant under dilatation, we obtain the following result from (5.5).

Lemma 5.5 *The quantity $\gamma = \gamma^\Omega_{m,\alpha,\beta}$ defined in Lemma 5.1d is the following function of α,β:*

$$\gamma^\Omega_{m,\alpha,\beta} = \sqrt{\|G^\Omega_{m,\alpha,\beta}(\cdot,0)\|_\infty} = \beta^{-1}(\beta/\alpha)^{\frac{d}{2m}} \sqrt{\|G^\Omega_{m,1,1}(\cdot,0)\|_\infty} = \beta^{-1}(\beta/\alpha)^{\frac{d}{2m}} \gamma^\Omega_{m,1,1}.$$

Let $Q_a := [0,a]^d$ be the cube with side length a. Finally, we mention the case of $\Omega = Q_1 = [0,1]^d$. Application of the dilatation operator M_λ to $u \in H^m(\Omega)$ yields a function $M_\lambda u \in H^m(Q_\lambda)$, where, differently from the situation above, the domain Q_λ depends on λ. Now, (5.5) takes the form

$$G^\Omega_{m,\alpha,\beta}(x,y) = \beta^{-2}(\beta/\alpha)^{d/m} G^{Q_{(\beta/\alpha)^{1/m}}}_{m,1,1}((\beta/\alpha)^{1/m}x, (\beta/\alpha)^{1/m}y). \tag{5.6}$$

Since the Green functions can, in principle, be approximated numerically, at least the size of the value of $\gamma^\Omega_{m,1,1}$ can be determined.

Conjecture 5.6 $G(0,0) \geq G(\xi,\xi)$.

Remark 5.7 *If $\Omega' \subset \Omega''$ and if $\gamma^{\Omega''}_{m,1,1} = \sqrt{G^{\Omega''}_{m,1,1}(\xi,\xi)}$ for some $\xi \in \Omega'$, then $\gamma^{\Omega'}_{m,1,1} \geq \gamma^{\Omega''}_{m,1,1}$.*

[10] In general, $\Omega = \Omega_1 \times \ldots \times \Omega_d$ must be assumed to be a cone (with origin at 0), i.e., $x \in \Omega$ implies $\lambda x \in \Omega$ for all $\lambda \geq 0$.

5.5 Inequality

Theorem 5.8 *Let $\Omega \in \left\{\mathbb{R}^d, [0,\infty)^d\right\}$ and suppose that $m > d/2$. Then*

$$\|u\|_\infty \le c_m^\Omega \cdot |u|_m^{\frac{d}{2m}} \cdot \|u\|^{1-\frac{d}{2m}} \qquad (u \in H^m(\Omega)) \tag{5.7}$$

holds, where the constant is

$$c_m^\Omega = a_m \gamma_{m,1,1}^\Omega \tag{5.8}$$

where

$$a_m := \min_{0<\delta<\pi/2} \sin^{\frac{d}{2m}-1}(\delta)\cos^{-\frac{d}{2m}}(\delta) = \sin^{\frac{d}{2m}-1}\big(\tfrac{1}{2}\arccos(\tfrac{d}{m}-1)\big)\cos^{-\frac{d}{2m}}\big(\tfrac{1}{2}\arccos(\tfrac{d}{m}-1)\big) \le \sqrt{2}$$

has the asymptotic behaviour $a_m = 1 + \frac{d}{4m}\ln\frac{2m}{d} + O(\frac{d}{m}) \to 1$ for $m \to \infty$.

Proof. Fix some $u \in H^m(\Omega)$. Let $M := |u|_m$ and $L := \|u\|$, and set $\alpha := \cos(\delta)/M$ and $\beta := \sin(\delta)/L$ for any $\delta \in (0, \pi/2)$. By construction,

$$|||u|||_{m,\alpha,\beta}^\Omega = 1$$

holds. We infer from (5.1) and Lemma 5.5 that

$$\begin{aligned}\|u\|_\infty &\le \gamma_{m,\alpha,\beta}^\Omega |||u|||_{m,\alpha,\beta}^\Omega = \gamma_{m,\alpha,\beta}^\Omega = \beta^{-1}(\beta/\alpha)^{\frac{d}{2m}}\gamma_{m,1,1}^\Omega \\ &= \gamma_{m,1,1}^\Omega \cdot (\beta^{-1})^{1-\frac{d}{2m}} \cdot \left(\alpha^{-1}\right)^{\frac{d}{2m}} = a(\delta)\cdot\gamma_{m,1,1}^\Omega \cdot \|u\|^{1-\frac{d}{2m}} \cdot |u|_m^{\frac{d}{2m}}.\end{aligned}$$

The factor $a(\delta)$ equals $\sin^{\frac{d}{2m}-1}(\delta)\cos^{-\frac{d}{2m}}(\delta)$. The choice $\delta = \pi/4$ yields $a(\pi/4) = \sqrt{2}$, while the minimum is taken at $\delta = \frac{1}{2}\arccos(\frac{d}{m}-1)$ with the asymptotic behaviour described above. ∎

In the case of the bounded domain $[0,1]^d$, the estimate (5.7) must be modified, since $|u|_m = 0$ holds for polynomials of degree $< m$.

Theorem 5.9 *Let $\Omega = [0,1]^d$ and $m > d/2$. Then*

$$\|u\|_\infty \le c_m^\Omega \cdot \left[|u|_m^2 + \|u\|^2\right]^{\frac{d}{4m}} \cdot \|u\|^{1-\frac{d}{2m}} \qquad (u \in H^m(\Omega))$$

holds, where the constant is $c_m^\Omega = \sqrt{2}\gamma_{m,1,1}^{Q_{(1+|u|_m^2/\|u\|^2)^{1/(2m)}}}$.

Proof. Set $M := \sqrt{|u|_m^2 + \|u\|^2}$, $L := \|u\|$, and[11] $\alpha := 1/(\sqrt{2}M)$, $\beta := 1/(\sqrt{2}L)$. Now we can proceed as in the previous proof. Because of the different definition of M, $|u|_m^2$ has to be replaced by $|u|_m^2 + \|u\|^2$. ∎

Concerning $\gamma_{m,1,1}^{Q_{(1+|u|_m^2/\|u\|^2)^{1/(2m)}}}$, we note that by Remark 5.7 and Conjecture 5.6, larger cubes lead to better bounds: $\gamma_{m,1,1}^{Q_{(1+|u|_m^2/\|u\|^2)^{1/(2m)}}} \le \gamma_{m,1,1}^{Q_1}$.

5.6 Fourier Transform

$\boldsymbol{\Omega} = \mathbb{R}^d$ allows to use Fourier transforms. Functions $u \in L^2(\mathbb{R}^d)$ are characterised by the Fourier transform $\hat{u} \in L^2(\mathbb{R}^d)$: $u(x) := (2\pi)^{-d/2}\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle\xi,x\rangle}\hat{u}(\xi)\mathrm{d}\xi$. The scaling by $(2\pi)^{-d/2}$ ensures isometry: $\langle v, w\rangle = \langle \hat{v}, \hat{w}\rangle$. The bilinear form $a_{m,\alpha,\beta}(v,w)$ corresponds to $\hat{a}_{m,\alpha,\beta}(\hat{v},\hat{w})$ with

$$a_{m,\alpha,\beta}(v,w) = \hat{a}_{m,\alpha,\beta}(\hat{v},\hat{w}) := \int_{\mathbb{R}^d}\left(\beta^2 + \alpha^2\sum\nolimits_{j=1}^d \xi_j^{2m}\right)\hat{v}(\xi)\overline{\hat{w}(\xi)}\mathrm{d}\xi.$$

The Green function $G_y = G(\cdot, y)$ has the transform

$$\hat{G}_y(\xi) = \hat{G}(\xi, y) = (2\pi)^{-d/2}\frac{\mathrm{e}^{-\mathrm{i}\langle\xi,y\rangle}}{\beta^2 + \alpha^2\sum_{j=1}^d \xi_j^{2m}}.$$

[11] The estimate is not optimised concerning the choice of α and β. The present choice corresponds to $\cos(\delta) = \sin(\delta) = 1/\sqrt{2}$ in the previous proof. Hence, $a_m = \sqrt{2}$ may be improved.

The norm $|||G_y||| = \sqrt{a(G_y, G_y)}$ can be rewritten as $\sqrt{\hat{a}(\hat{G}_y, \hat{G}_y)}$ and becomes

$$\sqrt{\int_{\mathbb{R}^d} \left(\alpha^2 \sum\nolimits_{j=1}^{d} \xi_j^{2m} + \beta^2\right) |\hat{G}_y|^2 \mathrm{d}\xi} = (2\pi)^{-d/2} \sqrt{\int_{\mathbb{R}^d} \left(\alpha^2 \sum\nolimits_{j=1}^{d} \xi_j^{2m} + \beta^2\right)^{-1} \mathrm{d}\xi}.$$

Therefore, the quantity $\gamma_{m,\alpha,\beta}^{\Omega}$ has the representation

$$\gamma_{m,\alpha,\beta}^{\Omega} = (2\pi)^{-d/2} \sqrt{\int_{\mathbb{R}^d} \left(\beta^2 + \alpha^2 \sum\nolimits_{j=1}^{d} \xi_j^{2m}\right)^{-1} \mathrm{d}\xi}.$$

In particular, the constant $c_m^{\Omega} = a_m \gamma_{m,1,1}^{\Omega}$ from (5.8) equals

$$c_m^{\Omega} = a_m \cdot (2\pi)^{-d/2} \sqrt{\int_{\mathbb{R}^d} \left(1 + \sum\nolimits_{j=1}^{d} \xi_j^{2m}\right)^{-1} \mathrm{d}\xi}.$$

For $m \to \infty$, the function $\left(1 + \sum_{j=1}^{d} \xi_j^{2m}\right)^{-1}$ tends to 0, when $\|\xi\|_\infty > 1$, and to 1, when $\|\xi\|_\infty < 1$. Together with $a_m \to 1$, the next lemma follows.

Lemma 5.10 $\lim_{m\to\infty} c_m^{\mathbb{R}^d} = \pi^{-d/2}$ *holds.*

5.7 Application to Discrete Grid Functions

In the practical applications, the Hilbert tensor space does not appear in full generality. Either $V = \bigotimes_{j=1}^{d} L^2(\Omega_j)$ is replaced by $\bigotimes_{j=1}^{d} U_j$ with finite dimensional subspaces of $L^2(\Omega_j)$, or the function spaces are replaced by spaces of grid functions, where the interval Ω_j is replaced by an, e.g., equidistant grid ω_j of a certain step size h_j.

We consider the infinite grid $h\mathbb{Z} = \{\nu h : \nu \in \mathbb{Z}\} \subset \mathbb{R}$. Hence, $L^2(\Omega_j)$ is to be replaced by the space $V_j = \ell^2(h\mathbb{Z})$ of grid functions $f : h\mathbb{Z} \to \mathbb{R}$. The entry f_ν is understood as the function value[12] at $x = \nu h$. The scalar product of this Hilbert space is $\langle f, g\rangle_j = h \sum_{\nu\in\mathbb{Z}} f_\nu g_\nu$. The factor h is added to ensure that the limit $h \to 0$ leads to the standard $L^2(\mathbb{R})$ scalar product. The tensor product yields the Hilbert space $\mathbf{V} = \bigotimes_{j=1}^{d} V_j = \ell^2((h\mathbb{Z})^d)$ with the corresponding induced scalar product

$$\langle f, g\rangle = h^d \sum\nolimits_{\boldsymbol{\nu}\in\mathbb{Z}^d} f_{\boldsymbol{\nu}} g_{\boldsymbol{\nu}} \qquad (f, g \in \ell^2(h\mathbb{Z}^d)).$$

The Fourier image of $f = (f_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathbb{Z}^d} \in \ell^2(h\mathbb{Z}^d)$ is the $2\pi/h$-periodic function

$$\hat{f}(\xi) = h^d (2\pi)^{-d/2} \sum\nolimits_{\boldsymbol{\nu}\in\mathbb{Z}^d} f_{\boldsymbol{\nu}} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\nu} h, \xi\rangle} \qquad \text{for } \xi \in [-\pi/h, \pi/h]^d$$

with the back transform $f_{\boldsymbol{\nu}} = (2\pi)^{-d/2} \int_{[-\pi/h,\pi/h]^d} \hat{f}(\xi) \mathrm{e}^{\mathrm{i}\langle \boldsymbol{\nu} h, \xi\rangle} \mathrm{d}\xi$. The scalar product in the Fourier space is $\langle \hat{f}, \hat{g}\rangle = \int_{[-\pi/h,\pi/h]^d} \hat{f}(\xi)\overline{\hat{g}(\xi)} \mathrm{d}\xi$.

The most elegant way of describing the smoothness of f uses the Fourier transform $\hat{f}$. Define the discrete Sobolev semi-norm $|f|_m$ by

$$|f|_m := \left\| \sqrt{\sum\nolimits_{j=1}^{d} \xi_j^{2m}} \hat{f}(\xi) \right\|_{L^2([-\pi/h,\pi/h]^d)}.$$

Instead, one may use the ℓ^2 norm of m-th divided difference quotients. For instance, the Fourier transform of the m-th forward difference is $[\frac{1}{h}(\exp(\xi_j h) - 1)]^m = [\frac{2\mathrm{i}}{h} \sin(\xi_j h/2)]^m$ and yields an equivalent norm.

The discrete 'delta function' at $\boldsymbol{\mu} \in \mathbb{Z}^d$ is $h^{-d}\delta_{\boldsymbol{\mu}}$, where $\delta_{\boldsymbol{\mu}}$ is the vector with components $\delta_{\boldsymbol{\mu},\boldsymbol{\nu}}$ (Kronecker symbol). Its Fourier transform is $h^{-d}\widehat{\delta_{\boldsymbol{\mu}}}(\xi) = (2\pi)^{-d/2} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\mu} h, \xi\rangle}$.

[12] The sequence $f = (f_\nu)$ may be identified with the piecewise constant function $f = f_\nu$ on $[(\nu - 1/2)h, (\nu + 1/2)h)$.

The Green (grid) function is $G_{\boldsymbol{\mu}} = G_{m,\alpha,\beta}(\cdot, \boldsymbol{\mu}h)$, $\boldsymbol{\mu} \in \mathbb{Z}^d$, defined on $(h\mathbb{Z})^d$. The Fourier image $\hat{G}_{\boldsymbol{\mu}} = \hat{G}_{m,\alpha,\beta}(\xi, \boldsymbol{\mu}h)$ of $G_{\boldsymbol{\mu}}$ is the $2\pi/h$-periodic function satisfying

$$\begin{aligned}&\int_{[-\pi/h,\pi/h]^d} \left(\alpha^2 \sum\nolimits_{j=1}^{d} \xi_j^{2m} + \beta^2\right) \hat{G}_{\boldsymbol{\mu}}(\xi)\hat{v}(\xi)\mathrm{d}\xi \\ &= h^{-d} \int_{[-\pi/h,\pi/h]^d} \widehat{\delta_{\boldsymbol{\mu}}}(\xi)\hat{v}(\xi)\mathrm{d}\xi = (2\pi)^{-d/2} \int_{[-\pi/h,\pi/h]^d} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\mu}h,\xi\rangle}\hat{v}(\xi)\mathrm{d}\xi\end{aligned}$$

for all $\hat{v}$. This yields

$$\hat{G}_{m,\alpha,\beta}(\xi, \boldsymbol{\mu}h) = (2\pi)^{-d/2} \left[\beta^2 + \alpha^2 \sum\nolimits_{j=1}^{d} \xi_j^{2m}\right]^{-1} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\mu}h,\xi\rangle} \qquad \text{for } \xi \in [-\pi/h, \pi/h]^d,$$

so that

$$G_{m,\alpha,\beta}(\boldsymbol{\nu}, \boldsymbol{\mu}) = (2\pi)^{-d} \int_{[-\pi/h,\pi/h]^d} \left[\beta^2 + \alpha^2 \sum\nolimits_{j=1}^{d} \xi_j^{2m}\right]^{-1} \mathrm{e}^{\mathrm{i}\langle(\boldsymbol{\nu}-\boldsymbol{\mu})h,\xi\rangle}\mathrm{d}\xi$$

An integral substitution $\xi = \lambda\zeta$ with $\lambda = (\beta/\alpha)^{1/m}$ yields

$$G_{m,\alpha,\beta}(0,0) = (2\pi)^{-d}(\beta/\alpha)^{d/m}\beta^{-2} \int_{[-\pi(\beta/\alpha)^{-1/m},\pi(\beta/\alpha)^{-1/m}]^d} \left[1 + \sum\nolimits_{j=1}^{d} \zeta_j^{2m}\right]^{-1} \mathrm{d}\zeta,$$

which corresponds to (5.5) and allows the same conclusions.

The limit of the integral for $m \to \infty$ is 2^d. Therefore, we obtain the same results as in Lemma 5.10 for $L^2(\mathbb{R}^d)$.

Acknowledgment. We thank Prof. Dr. H. Triebel for bibliographical hints.

References

[1] J. Ballani, *Fast evaluation of BEM integrals based on tensor approximations*, Numer. Math., to appear

[2] L. Caffarelli, R. Kohn, and L. Nirenberg, *First order interpolation inequalities with weights*, Composition Mathematica **53** (1984), 259–275.

[3] L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl. **21** (2000), 1253–1278.

[4] L. Grasedyck, *Hierarchical singular value decomposition of tensors*, SIAM J. Matrix Anal. Appl. **31** (2010), 2029–2054.

[5] W. Hackbusch, *Elliptic differential equations. Theory and numerical treatment*, 2nd ed., SCM, no. 18, Springer, Berlin, 2003.

[6] ______, *Tensor spaces and numerical tensor calculus*, SCM, no. 42, Springer, Berlin, 2012.

[7] F. L. Hitchcock, *Multiple invariants and generalized rank of a p-way matrix or tensor*, Journal of Mathematics and Physics **7** (1927), 40–79.

[8] V. G. Maz'ja, *Sobolev spaces*, 2nd ed., Springer, Berlin, 2011.

[9] L. Nirenberg, *On elliptic partial differential equations*, Ann. Sc. Norm. Sup. Pisa, Cl. Sci. 3 Ser. **13** (1959), 115–162.

[10] F. Oberhettinger, *Tabellen zur Fourier Transformation*, Springer, Berlin, 1957.

[11] E. Zeidler (ed.), *Oxford users' guide to mathematics*, Oxford University Press, Oxford, 2004.