Springer Series in Computational Mathematics 42

Wolfgang Hackbusch

Tensor Spaces and Numerical Tensor Calculus

Springer

Springer Series in Computational Mathematics

42

For further volumes:
http://www.springer.com/series/797

Wolfgang Hackbusch

Tensor Spaces and Numerical Tensor Calculus

Springer

Wolfgang Hackbusch
Max-Planck-Institute
for Mathematics in the Sciences
Leipzig
Germany

ISSN 0179-3632
ISBN 978-3-642-28026-9 e-ISBN 978-3-642-28027-6
DOI 10.1007/978-3-642-28027-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012933366

Mathematics Subject Classification (2010): 15A69, 35C99, 44A35, 46A32, 46B28, 47A80,
65F99 65N99 65L99 65N22 65N99 65Z05

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Dedicated to my grandchildren
Alina and Daniel

Preface

Large-scale problems have always been a challenge for numerical computations. An example is the treatment of fully populated $n \times n$ matrices, when n^2 is close to or beyond the computer's memory capacity. Here, the technique of hierarchical matrices can reduce the storage and the cost of numerical operations from $O(n^2)$ to almost $O(n)$.

Tensors of order (or spatial dimension) d can be understood as d-dimensional generalisations of matrices, i.e., arrays with d discrete or continuous arguments. For large $d \geq 3$, the data size n^d is far beyond any computer capacity. This book concerns the development of compression techniques for such high-dimensional data via suitable data sparse representations. Just as the hierarchical matrix technique was based on a successful application of the low-rank strategy, in recent years, related approaches have been used to solve high-dimensional tensor-based problems numerically. The results are quite encouraging, at least for data arising from suitably smooth problems, and even some problems of size $n^d = 1000^{1000}$ have become computable.

The methods, which can be applied to these multilinear problems, are black box-like. In this aspect they are similar to methods used in linear algebra. On the other hand, most of the methods are approximate (computing suitably accurate approximations to quantities of interest) and in this respect they are similar to some approaches in analysis. The crucial key step is the construction of an efficient new tensor representation, thus overcoming the drawbacks of the traditional tensor formats. In 2009 a rapid progress could be achieved by introducing the hierarchical format as well as the TT format for the tensor representation. Under suitable conditions these formats allow a stable representation and a reduction of the data size from n^d to $O(dn)$. Another recent advancement is the so-called tensorisation technique, which may replace the size n by $O(\log n)$. Altogether, there is the hope that problems of size h^d can be reduced to size $O(d \log(n)) = O(\log(n^d))$, i.e., we reduce the problems to *logarithmical size*.

It turned out that some of the raw material for the methods described in this book was already known in the literature belonging to other (applied) fields outside of mathematics, such as chemistry. However, the particular language used to describe this material, combined with the fact that the algorithms (although potentially of general interest) were given names relating them only to a particular application, prevented the dissemination of the methods to a wider audience.

One of the aims of this monograph is to introduce a more mathematically-based treatment of this topic. Through this more abstract approach, the methods can be better understood, independently of the physical or technical details of the application.

The material in this monograph has been used for as the basis for a course of lectures at the University Leipzig in the summer semester of 2010.

The author's research at the Max-Planck Institute of Mathematics in the Sciences has been supported by a growing group of researchers. In particular we would like to mention: B. Khoromskij, M. Espig, L. Grasedyck, and H.J. Flad. The help of H.J. Flad was indispensable for bridging the terminological gap between quantum chemistry and mathematics. The research programme has also benefited from the collaboration between the group in Leipzig and the group of E. Tyrtyshnikov in Moscow. E. Tyrtyshnikov and I. Oseledets have delivered important contributions to the subject. A further inspiring cooperation[1] involves R. Schneider (TU Berlin, formerly University of Kiel). The author thanks many more colleagues for stimulating discussions.

The author also wishes to express his gratitude to the publisher Springer for their friendly cooperation.

Leipzig, October 2011 *Wolfgang Hackbusch*

[1] The groups in Leipzig and Berlin thank the Germany Research Society (DFG) for financial support within the priority program 1324 (Schwerpunktprogramm "Extraktion quantifizierbarer Information aus komplexen Systemen"). The Russian partner has been supported by the Russian Foundation for Basic Research RFBR.

Contents

Part I Algebraic Tensors

List of Symbols and Abbreviations

Symbols

$[a\ b\ \ldots]$	aggregation of vectors $a, b \in \mathbb{K}^I, \ldots$ into a matrix of size $I \times J$
$[A\ B\ \ldots]$	aggregation of matrices $A \in \mathbb{K}^{I\times J_1}$, $B \in \mathbb{K}^{I\times J_2}, \ldots$ into a matrix of size $I \times (J_1 \cup J_2 \cup \ldots)$
$\lceil\cdot\rceil$	smallest integer $\geq \cdot$
$\lfloor\cdot\rfloor$	largest integer $\leq \cdot$
$\langle\cdot,\cdot\rangle$	scalar product; in $\mathbb{K}^I$ usually the Euclidean scalar product; cf. §2.1, §4.4.1
$\langle\cdot,\cdot\rangle_\alpha$	partial scalar product; cf. (4.66)
$\langle\cdot,\cdot\rangle_H$	scalar product of a (pre-)Hilbert space H
$\langle\cdot,\cdot\rangle_{\mathrm{HS}}$	Hilbert-Schmidt scalar product; cf. Definition 4.117
$\langle\cdot,\cdot\rangle_j$	scalar product of the (pre-)Hilbert space V_j from $\mathbf{V} = \bigotimes_{j=1}^d V_j$
$\langle\cdot,\cdot\rangle_\mathsf{F}$	Frobenius scalar product of matrices; cf. (2.10)
$\#$	cardinality of a set
$\rightharpoonup$	weak convergence; cf. §4.1.7
$\bullet\vert_{\tau\times\sigma}$	restriction of a matrix to the matrix block $\tau \times \sigma$; cf. §1.7
$\bullet^\perp$	orthogonal complement, cf. §4.4.1
$\bullet^\mathsf{H}$	Hermitean transpose of a matrix or vector
$\bullet^\mathsf{T}$	transpose of a matrix or vector
$\bullet^{-\mathsf{T}}, \bullet^{-\mathsf{H}}$	inverse matrix of $\bullet^\mathsf{T}$ or $\bullet^\mathsf{H}$, respectively
$\overline{\bullet}$	either complex-conjugate value of a scalar or closure of a set
$\times$	Cartesian product of sets: $A \times B := \{(a,b) : a \in A, b \in B\}$
$\bigtimes_{j=1}^d$	d-fold Cartesian product of sets
$\times_j$	j-mode product, cf. Footnote 6 on page 5; not used here
$\star$	convolution; cf. §4.6.5
$\wedge$	exterior product; cf. §3.5.1
$\odot$	Hadamard product; cf. (4.72a)
$\oplus$	direct sum; cf. footnote on page 21
$\otimes^d$	d-fold tensor product; cf. Notation 3.23

$v \otimes w$, $\bigotimes_{j=1}^{d} v^{(j)}$	tensor product of two or more vectors; cf. §3.2.1
$V \otimes W$, $\bigotimes_{j=1}^{d} V_j$	tensor space generated by two or more vector spaces; cf. §3.2.1
$V \otimes_a W$, ${}_a\bigotimes_{j=1}^{d} V_j$	algebraic tensor space; cf. (3.11) and §3.2.4
$V \otimes_{\|\cdot\|} W$, ${}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$	topological tensor space; cf. (3.12); §4.2
$\bigotimes_{j \neq k}$	cf. (3.21b)
$\subset$	the subset relation $A \subset B$ includes the case $A = B$
$\dot\cup$	disjoint union
$\sim$	equivalence relation; cf. §3.1.3, §4.1.1
$\bullet \cong \bullet$	isomorphic spaces; cf. §3.2.5
$\bullet \leq \bullet$	semi-ordering of matrices; cf. (2.14)
$\|\cdot\|$	norm; cf. §4.1.1
$\|\cdot\|^*$	dual norm; cf. Lemma 4.18
$\|\cdot\|_2$	Euclidean norm of vector or tensor (cf. (2.12) and Example 4.126) or spectral norm of a matrix (cf. (2.13))
$\|\cdot\|_\mathsf{F}$	Frobenius norm of matrices; cf. (2.9)
$\|\cdot\|_\mathrm{HS}$	Hilbert-Schmidt norm; cf. Definition 4.117
$\|\cdot\|_{\mathrm{SVD},p}$	Schatten norm; cf. (4.17)
$\|\cdot\|_X$	norm of a space X
$\|\cdot\|_{X \leftarrow Y}$	associated matrix norm (cf. (2.11) or operator norm (cf. (4.6a))
$\|\cdot\|_1 \lesssim \|\cdot\|_2$	semi-ordering of norms; cf. §4.1.1
$\|\cdot\|_{\wedge(V,W)}$, $\|\cdot\|_\wedge$	projective norm; cf. §4.2.4
$\|\cdot\|_{\vee(V,W)}$, $\|\cdot\|_\vee$	injective norm; cf. §4.2.7

Greek Letters

α	often a subset of the set D of directions (cf. (5.3a,b)) or vertex of the tree T_D (cf. Definition 11.2)
α^c	complement $D \backslash \alpha$; cf. (5.3c)
α_1, α_2	often sons of a vertex $\alpha \in T_D$; cf. §11.2.1
δ_{ij}	Kronecker delta; cf. (2.1)
$\boldsymbol{\rho}$	tuple of TT ranks; cf. Definition 12.1
$\rho(\cdot)$	spectral radius of a matrix; cf. §4.6.6
$\rho_\mathrm{xyz}(\cdot)$	tensor representation by format ‘xyz’; cf. §7.1
ρ_frame	general tensor subspace format; cf. (8.13c)
ρ_HOSVD	HOSVD tensor subspace format; cf. (8.26)
ρ_HTR	hierarchical format; cf. (11.28)
$\rho_\mathrm{HTR}^\mathrm{HOSVD}$	hierarchical HOSVD format; cf. Definition 11.36
$\rho_\mathrm{HTR}^\mathrm{orth}$	orthonormal hierarchical format; cf. (11.38)
$\rho_\mathrm{HTR}^\mathrm{tens}$	TT format for tensorised vectors; cf. (14.5a)
ρ_hybr, $\rho_\mathrm{orth}^\mathrm{hybr}$, $\rho_\text{r-term}^\mathrm{hybr}$	hybrid formats; cf. §8.2.4
ρ_j	TT rank; cf. (12.1a) and Definition 12.1
ρ_orth	orthonormal tensor subspace format; cf. (8.8b)

$\rho_{\text{r-term}}$	r-term format; cf. (7.7a)
ρ_{sparse}	sparse format; cf. (7.5)
ρ_{TS}	general tensor subspace format; cf. (8.6c)
ρ_{TT}	TT format; cf. (12.7)
$\sigma(\cdot)$	spectrum of a matrix; cf. §4.6.6
σ_i	singular value of the singular value decomposition; cf. (2.19a), (4.59)
$\varSigma$	diagonal matrix of the singular value decomposition; cf. (2.19a)
φ, ψ	often linear mapping or functional (cf. §3.1.4)
$\varPhi, \varPsi$	often linear mapping or operator (cf. §4.1.4)
$\varPhi'$	dual of $\varPhi$; cf. Definition 4.20
$\varPhi^*$	adjoint of $\varPhi$; cf. Definition 4.113

Latin Letters

$\mathbf{a}$	coefficient tensor, cf. Remark 3.29
A, $B, \ldots, A_1, A_2, \ldots$	often used for linear mapping (from one vector space into another one). This includes matrices.
$\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}, \ldots$	tensor products of operators or matrices
$A^{(j)}$	mapping from $L(V_j, W_j)$, j-th component in a Kronecker product
$\mathcal{A}(V)$	tensor algebra generated by V; cf. (3.43)
$\mathfrak{A}(V)$	antisymmetric tensor space generated by V; cf. Definition 3.62
Arcosh	area [inverse] hyperbolic cosine: $\cosh(\mathrm{Arcosh}(x)) = x$
$b_i^{(j)}, \mathbf{b}_i^{(\alpha)}$	basis vectors; cf. (8.5a), (11.20a)
$B, B_j, \mathbf{B}_\alpha$	basis (or frame), $B_j = \left[b_1^{(j)}, \ldots, b_r^{(j)}\right]$, cf. (8.5a-d); in the case of tensor spaces: $\mathbf{B}_\alpha = \left[\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}\right]$, cf. (11.20a)
$c_0(I)$	subset of $\ell^\infty(I)$; cf. (4.4)
$c_{ij}^{(\alpha,\ell)}$	coefficients of the matrix $C^{(\alpha,\ell)}$; cf. (11.24)
$\mathbb{C}$	field of complex numbers
$C(D), C^0(D)$	bounded, continuous functions on D; cf. Example 4.8
$\mathbf{C}_\alpha$	tuple $(C^{(\alpha,\ell)})_{1\leq\ell\leq r_\alpha}$ of $C^{(\alpha,\ell)}$ from below; cf. (11.27)
$C^{(\alpha,\ell)}$	coefficient matrix at vertex α characterising the basis vector $\mathbf{b}_\ell^{(\alpha)}$; cf. (11.24)
$\mathfrak{C}_j, \mathfrak{C}_\alpha$	contractions; cf. Definition 4.130
$C_N(f,h), C(f,h)$	sinc interpolation; cf. Definition 10.31
d	order of a tensor; cf. §1.1.1
D	set $\{1, \ldots, d\}$ of directions; cf. (5.3b)
$\mathfrak{D}_\delta$	analyticity stripe; cf. (10.38)
$depth(\cdot)$	depth of a tree; cf. (11.7)
$\det(\cdot)$	determinant of a matrix
$\mathrm{diag}\{\ldots\}$	diagonal matrix with entries ...
$\dim(\cdot)$	dimension of a vector space
$e^{(i)}$	i-th unit vector of $\mathbb{K}^I$ ($i \in I$); cf. (2.2)

$E_N(f,h), E(f,h)$	sinc interpolation error; cf. Definition 10.31
$E_r(\cdot)$	exponential sum; cf. (9.27a)
$\mathcal{E}_\rho$	regularity ellipse; cf. §10.4.2.2
$\mathcal{F}(W,V)$	space of finite rank operators; cf. §4.2.13
$G(\cdot)$	Gram matrix of a set of vectors; cf. (2.16), (11.35)
H, H_1, $H_2, \ldots$	(pre-)Hilbert spaces
$\mathbf{H}(\mathfrak{D}_\delta)$	Banach space from Definition 10.33
$H^{1,p}(D)$	Sobolev space; cf. Example 4.41
$HS(V,W)$	Hilbert-Schmidt space; cf. Definition 4.117
id	identity mapping
i, j, $k, \ldots$	index variables
$\mathbf{i}, \mathbf{j}, \mathbf{k}$	multi-indices from a product index set $\mathbf{I}$ etc.
I	identity matrix or index set
$\mathcal{I}, \mathcal{I}_{[a,b]}$	interpolation operator; cf. §10.4.3
I, J, K, I_1, $I_2, \ldots,$ J_1, $J_2, \ldots$	often used for index sets
$\mathbf{I}$, $\mathbf{J}$	index sets defined by products $I_1 \times I_2 \times \ldots$ of index sets
j	often index variable for the directions from $\{1, \ldots, d\}$
$\mathbb{K}$	underlying field of a vector space; usually $\mathbb{R}$ or $\mathbb{C}$
$\mathcal{K}(W,V)$	space of compact operators; cf. §4.2.13
$\ell(I)$	vector space $\mathbb{K}^I$; cf. Example 3.1
$\ell_0(I)$	subset of $\ell(I)$; cf. (3.2)
$\ell^p(I)$	Banach space from Example 4.5; $1 \leq p \leq \infty$
$level$	level of a vertex of a tree, cf. (11.6)
L	often depth of a tree, cf. (11.7)
L	lower triangular matrix in Cholesky decomposition; cf. §2.5.1
$\mathcal{L}(T)$	set of leaves of the tree T; cf. (11.9)
$L(V,W)$	vector space of linear mappings from V into W; cf. §3.1.4
$\mathcal{L}(X,Y)$	space of continuous linear mappings from X into Y; cf. §4.1.4
$L^p(D)$	Banach space; cf. Example 4.7; $1 \leq p \leq \infty$
$\mathcal{M}_\alpha, \mathcal{M}_j$	matricisation isomorphisms; cf. Definition 5.3
n, n_j	often dimension of a vector space V, V_j
$\mathbb{N}$	set $\{1, 2, \ldots\}$ of natural numbers
$\mathbb{N}_0$	set $\mathbb{N} \cup \{0\} = \{0, 1, 2, \ldots\}$
$\mathcal{N}(W,V)$	space of nuclear operators; cf. §4.2.13
N_{xyz}	arithmetical cost of 'xyz'
$N_{\mathrm{mem}}^{\mathrm{xyz}}$	storage cost of 'xyz'; cf. (7.8a)
N_{LSVD}	cost of a left-sided singular value decomposition; cf. p. 2.21
N_{QR}	cost of a QR decomposition; cf. Lemma 2.19
N_{SVD}	cost of a singular value decomposition; cf. Corollary 38
$o(\cdot), O(\cdot)$	Landau symbols; cf. (4.12)
P	permutation matrix (cf. (2.18)) or projection
$P_\mathfrak{A}$	alternator, projection onto $\mathfrak{A}(V)$; cf. (3.45)
$P_\mathfrak{S}$	symmetriser, projection onto $\mathfrak{S}(V)$; cf. (3.45)
$\mathbf{P}$, $\mathbf{P}_j$, etc.	often used for projections in tensor spaces
$P_j^{\mathrm{HOSVD}}, P_{j,\mathrm{HOSVD}}^{(r_j)}, \mathbf{P}_\mathbf{r}^{\mathrm{HOSVD}}$	HOSVD projections; cf. Lemma 10.1

$\mathcal{P}, \mathcal{P}_p, \mathcal{P}_{\mathbf{p}}$	spaces of polynomials; cf. §10.4.2.1
Q	unitary matrix of QR decomposition; cf. (2.17a)
r	matrix rank or tensor rank (cf. §2.2), representation rank (cf. Definition 7.3), or bound of ranks
$\mathfrak{r}$	rank $(r_\alpha)_{\alpha \in T_D}$ connected with hierarchical format $\mathcal{H}_{\mathfrak{r}}$; cf. §11.2.2
r_α	components of $\mathfrak{r}$ from above
$\mathbf{r}$	rank $(r_1, \ldots, r_d)$ connected with tensor subspace representation in $\mathcal{T}_{\mathbf{r}}$
r_j	components of $\mathbf{r}$ from above
$\mathbf{r}_{\min}(\mathbf{v})$	tensor subspace rank; cf. Remark 8.4
$range(\cdot)$	range of a matrix or operator; cf. §2.1
$\operatorname{rank}(\cdot)$	rank of a matrix or tensor; cf. §2.2 and (3.24)
$\underline{\operatorname{rank}}(\cdot)$	border rank; cf. (9.11)
$\operatorname{rank}_\alpha(\cdot), \operatorname{rank}_j(\cdot)$	α-rank and j-rank; cf. Definition 5.7
$r_{\max}$	maximal rank; cf. (2.5) and §3.2.6.4
R	upper triangular matrix of QR decomposition; cf. (2.17a)
$\mathbb{R}$	field of real numbers
$\mathbb{R}^J$	set of J-tuples; cf. page 4
$\mathcal{R}_r$	set of matrices or tensors of rank $\leq r$; cf. (2.6) and (3.22)
$S(\alpha)$	set of sons of a tree vertex α; cf. Definition 11.2
$\mathfrak{S}(V)$	symmetric tensor space generated by V; cf. Definition 3.62
$S(k,h)(\cdot)$	see (10.36)
$\operatorname{sinc}(\cdot)$	sinc function: $\sin(\pi x)/(\pi x)$
$\operatorname{span}\{\cdot\}$	subspace spanned by $\cdot$
$\operatorname{supp}(\cdot)$	support of a mapping; cf. §3.1.2
T_α	subtree of T_D; cf. Definition 11.6
T_D	dimension partition tree; cf. Definition 11.2
$T_D^{(\ell)}$	set of tree vertices at level ℓ; cf. (11.8)
T_D^{TT}	linear tree used for the TT format; cf. §12
$\mathcal{T}_{\mathbf{r}}$	set of tensors of representation rank $\mathbf{r}$; cf. Definition 8.1
$\mathbb{T}_\rho$	set of tensors of TT representation rank $\boldsymbol{\rho}$; cf. (12.4)
$\operatorname{trace}(\cdot)$	trace of a matrix or operator; cf. (2.8) and (4.60)
$tridiag\{a,b,c\}$	tridiagonal matrix (a: lower diagonal entries, b : diagonal, c: upper diagonal entries)
U	vector space, often a subspace
U, V	unitary matrices of the singular value decomposition; cf. (2.19b)
u_i, v_i	left and right singular vectors of SVD; cf. (2.21)
u,v,w	vectors
$\mathbf{u}$,$\mathbf{v}$,$\mathbf{w}$	tensors
$\mathbf{U}$	tensor space, often a subspace of a tensor space
$\mathbf{U}_\alpha$	subspace of the tensor space $\mathbf{V}_\alpha$; cf. (11.10)
$\{\mathbf{U}_\alpha\}_{\alpha \in T_D}$	hierarchical subspace family; cf. Definition 11.8
$U', V', W', \ldots$	algebraic duals of $U, V, W, \ldots$; cf. (3.7)
$U_j^I(\mathbf{v}), U_j^{II}(\mathbf{v}), U_j^{III}(\mathbf{v}), U_j^{IV}(\mathbf{v})$	see Lemma 6.12
$U_j^{\min}(\mathbf{v}), \mathbf{U}_\alpha^{\min}(\mathbf{v})$	minimal subspaces of a tensor $\mathbf{v}$; Def. 6.3, (6.10a), and §6.4
v_i	either the i-th component of v or the i-th vector of a set of vectors

$v^{(j)}$	vector of V_j corresponding to the j-th direction of the tensor; cf. §3.2.4
$\mathbf{v}^{[k]}$	tensor belonging to $\mathbf{V}_{[k]}$; cf. (3.21d)
$\mathcal{V}_{\mathrm{free}}(S)$	free vector space of a set S; cf. §3.1.2
$\mathbf{V}_\alpha$	tensor space $\bigotimes_{j\in\alpha} V_j$; cf. (5.3d)
$\mathbf{V}_{[j]}$	tensor space $\bigotimes_{k\neq j} V_j$; cf. (3.21a) and §5.2
$V, W, \ldots, X, Y, \ldots$	vector spaces
$V', W', \ldots, X', Y', \ldots$	algebraically dual vector spaces; cf. (3.7)
$\mathbf{V}, \mathbf{W}, \mathbf{X}, \mathbf{Y}$	tensor spaces
X, Y	often used for Banach spaces; cf. §4.1
$X^*, Y^*, \ldots$	dual spaces containing the continuous functionals; cf. §4.1.5
V^{**}	bidual space; cf. §4.1.5

Abbreviations and Algorithms

ALS	alternating least-squares method, cf. §9.5.2
ANOVA	analysis of variance, cf. §17.4
DCQR	cf. (2.40)
DFT	density functional theory, cf. §13.11
DFT	discrete Fourier transform, cf. §14.4.1
DMRG	density matrix renormalisation group, cf. §17.2.2
FFT	fast Fourier transform, cf. §14.4.1
HOOI	higher-order orthogonal iteration; cf. §10.3.1
HOSVD	higher-order singular value decomposition; cf. §8.3
HOSVD$(\cdot)$, **HOSVD**$^*(\cdot)$, **HOSVD**$^{**}(\cdot)$	procedures constructing the hierarchical HOSVD format; cf. (11.46a-c)
HOSVD-lw, **HOSVD*****-lw**	levelwise procedures; cf. (11.46a-c), (11.47a,b)
HOSVD-TrSeq	sequential truncation procedure; cf. (11.63)
$\mathrm{HOSVD}_\alpha(\mathbf{v}), \mathrm{HOSVD}_j(\mathbf{v})$	computation of HOSVD data; cf. (8.30a)
JoinBases	joining two bases; cf. (2.35)
JoinONB	joining two orthonormal bases; cf. (2.36)
LOBPCG	locally optimal block preconditioned conjugate gradient, cf. (16.13)
LSVD	left-sided reduced SVD; cf. (2.32)
MALS	modified alternating least-squares method, cf. §17.2.2
MPS	matrix product state, matrix product system; cf. §12
PEPS	projected entangled pairs states, cf. footnote 5 on page 384
PGD	proper generalised decomposition, cf. (17.1.1)
PQR	pivotised QR decomposition; cf. (2.30)
QR	QR decomposition; §2.5.2
REDUCE, **REDUCE***	truncation procedure; cf. §11.4.2
RQR	reduced QR decomposition; cf. (2.29)
RSVD	reduced SVD; cf. (2.31)
SVD	singular value decomposition; cf. §2.5.3

Part I
Algebraic Tensors

In *Chap. 1*, we start with an elementary introduction into the world of tensors (the precise definitions are in Chap. 3) and explain where large-sized tensors appear. This is followed by a description of the *Numerical Tensor Calculus*. Section 1.4 contains a preview of the material of the three parts of the book. We conclude with some historical remarks and an explanation of the notation.

The numerical tools which will be developed for tensors, make use of linear algebra methods (e.g., QR and singular value decomposition). Therefore, these matrix techniques are recalled in *Chap. 2*.

The definition of the algebraic tensor space structure is given in *Chap. 3*. This includes linear mappings and their tensor product.

Chapter 1
Introduction

In view of all that ..., the many obstacles we appear to have surmounted, what casts the pall over our victory celebration? It is the curse of dimensionality, a malediction that has plagued the scientist from earliest days.
(Bellman [11, p. 94]).

1.1 What are Tensors?

For a first rough introduction into tensors, we give a preliminary definition of tensors and the tensor product. The formal definition in the sense of multilinear algebra will be given in Chap. 3. In fact, below we consider three types of tensors which are of particular interest in later applications.

1.1.1 Tensor Product of Vectors

While vectors have entries v_i with one index and matrices have entries M_{ij} with two indices, tensors will carry d indices. The natural number[1] d defines the *order* of the tensor. The indices

$$j \in \{1, \ldots, d\}$$

correspond to the 'j-th direction', 'j-th position', 'j-th dimension', 'j-th axis', 'j-th site', or[2] 'j-th mode'. The names 'direction' and 'dimension' originate from functions $f(x_1, \ldots, x_d)$ (cf. §1.1.3), where the variable x_j corresponds to the j-th spatial direction.

For each $j \in \{1, \ldots, d\}$ we fix a (finite) index set I_j, e.g., $I_j = \{1, \ldots, n_j\}$. The Cartesian product of these index sets yields

$$\mathbf{I} := I_1 \times \ldots \times I_d\,.$$

The elements of $\mathbf{I}$ are multi-indices or d-tuples $\mathbf{i} = (i_1, \ldots, i_d)$ with $i_j \in I_j$. A tensor $\mathbf{v}$ is defined by its entries

$$\mathbf{v}_\mathbf{i} = \mathbf{v}[\mathbf{i}] = \mathbf{v}[i_1, \ldots, i_d] \in \mathbb{R}.$$

[1] The letter d is chosen because of its interpretation as spatial $\underline{\text{d}}$imension.

[2] The usual meaning of the term 'mode' is 'eigenfunction'.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series
in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_1,

We may write $\mathbf{v} := (\mathbf{v}[\mathbf{i}])_{\mathbf{i}\in\mathbf{I}}$. Mathematically, we can express the set of these tensors by $\mathbb{R}^{\mathbf{I}}$. Note that for any index set J, $\mathbb{R}^J$ is the vector space

$$\mathbb{R}^J = \{v = (v_i)_{i\in J} : v_i \in \mathbb{R}\}$$

of dimension $\#J$ (the sign $\#$ denotes the cardinality of a set).

Notation 1.1. Both notations, $\mathbf{v}_\mathbf{i}$ with subscript $\mathbf{i}$ and $\mathbf{v}[\mathbf{i}]$ with square brackets are used in parallel. The notation with square brackets is preferred for multiple indices and in the case of secondary subscripts: $\mathbf{v}[i_1,\ldots,i_d]$ instead of $\mathbf{v}_{i_1,\ldots,i_d}$.

There is an obvious entrywise definition of the multiplication $\lambda\mathbf{v}$ of a tensor by a real number and of the (commutative) addition $\mathbf{v}+\mathbf{w}$ of two tensors. Therefore the set of tensors has the algebraic structure of a vector space (here over the field $\mathbb{R}$). In particular in scientific fields more remote from mathematics and algebra, a tensor $\mathbf{v}[i_1,\ldots,i_d]$ is regarded as data structure and called 'd-way array'.

The relation between the vector spaces $\mathbb{R}^{I_j}$ and $\mathbb{R}^{\mathbf{I}}$ is given by the tensor product. For vectors $v^{(j)} \in \mathbb{R}^{I_j}$ $(1 \le j \le d)$ we define the *tensor product*[3],[4]

$$\mathbf{v} := v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes v^{(d)} = \bigotimes_{j=1}^{d} v^{(j)} \in \mathbb{R}^{\mathbf{I}}$$

via its entries

$$\mathbf{v}_\mathbf{i} = \mathbf{v}[i_1,\ldots,i_d] = v^{(1)}_{i_1} \cdot v^{(2)}_{i_2} \cdot \ldots \cdot v^{(d)}_{i_d} \qquad \text{for all } \mathbf{i} \in \mathbf{I}. \tag{1.1}$$

The tensor space is written as tensor product $\bigotimes_{j=1}^{d} \mathbb{R}^{I_j} = \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \ldots \otimes \mathbb{R}^{I_d}$ of the vector spaces $\mathbb{R}^{I_j}$ defined by the span

$$\bigotimes_{j=1}^{d} \mathbb{R}^{I_j} = \operatorname{span}\left\{v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes v^{(d)} : v^{(j)} \in \mathbb{R}^{I_j},\ 1 \le j \le d\right\}. \tag{1.2}$$

The generating products $v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes v^{(d)}$ are called *elementary tensors*.[5] Any element $\mathbf{v} \in \bigotimes_{j=1}^{d} \mathbb{R}^{I_j}$ of the tensor space is called a (general) *tensor*. It is important to notice that, in general, a tensor $\mathbf{v} \in \bigotimes_{j=1}^{d} \mathbb{R}^{I_j}$ is not representable as elementary tensor, but only as a linear combination of such products.

The definition (1.2) implies $\bigotimes_{j=1}^{d} \mathbb{R}^{I_j} \subset \mathbb{R}^{\mathbf{I}}$. Taking all linear combinations of elementary tensors defined by the unit vectors, one easily proves $\bigotimes_{j=1}^{d} \mathbb{R}^{I_j} = \mathbb{R}^{\mathbf{I}}$. In particular, because of $\#\mathbf{I} = \prod_{j=1}^{d} \#I_j$, the dimension of the tensor space is

$$\dim\left(\bigotimes_{j=1}^{d} \mathbb{R}^{I_j}\right) = \prod_{j=1}^{d} \dim(\mathbb{R}^{I_j}).$$

[3] In some publications the term '*outer product*' is used instead of 'tensor product'. This contradicts another definition of the outer product or exterior product satisfying the antisymmetric property $u \wedge v = -(v \wedge u)$ (see page 82).

[4] The index j indicating the 'direction' is written as upper index in brackets, in order to let space for further indices placed below.

[5] Also the term 'decomposable tensors' is used. Further names are 'dyads' for $d = 2$, 'triads' for $d = 3$, etc. (cf. [139, p. 3]).

Remark 1.2. Let $\#I_j = n$, i.e., $\dim(\mathbb{R}^{I_j}) = n$ for $1 \leq j \leq d$. Then the dimension of the tensor space is n^d. Unless both n and d are rather small numbers, n^d is a huge number. In such cases, n^d may exceed the computer memory by far. This fact indicates a practical problem, which must be overcome.

The set of matrices with indices in $I_1 \times I_2$ is denoted by $\mathbb{R}^{I_1 \times I_2}$.

Remark 1.3. (a) The particular case $d = 2$ leads to matrices $\mathbb{R}^{\mathbf{I}} = \mathbb{R}^{I_1 \times I_2}$, i.e., matrices may be identified with tensors of order 2. To be precise, the tensor entry $\mathbf{v_i}$ with $\mathbf{i} = (i_1, i_2) \in \mathbf{I} = I_1 \times I_2$ is identified with the matrix entry M_{i_1,i_2}. Using the matrix notation, the tensor product of $v \in \mathbb{R}^{I_1}$ and $w \in \mathbb{R}^{I_2}$ equals

$$v \otimes w = v\, w^{\mathsf{T}}. \tag{1.3}$$

(b) For $d = 1$ the trivial identity $\mathbb{R}^{\mathbf{I}} = \mathbb{R}^{I_1}$ holds, i.e., vectors are tensors of order 1.

(c) For the degenerate case $d = 0$, the empty product is defined by the underlying field: $\bigotimes_{j=1}^{0} \mathbb{R}^{I_j} = \mathbb{R}$.

1.1.2 Tensor Product of Matrices, Kronecker Product

Let d pairs of vector spaces V_j and W_j $(1 \leq j \leq d)$ and the corresponding tensor spaces

$$\mathbf{V} = \bigotimes_{j=1}^{d} V_j \quad \text{and} \quad \mathbf{W} = \bigotimes_{j=1}^{d} W_j$$

be given together with linear mappings

$$A^{(j)} : V_j \to W_j.$$

The tensor product of the $A^{(j)}$, the so-called *Kronecker product*, is the linear mapping

$$\mathbf{A} := \bigotimes_{j=1}^{d} A^{(j)} : \mathbf{V} \to \mathbf{W} \tag{1.4a}$$

defined by

$$\bigotimes_{j=1}^{d} v^{(j)} \in \mathbf{V} \quad \mapsto \quad \mathbf{A}\left(\bigotimes_{j=1}^{d} v^{(j)}\right) = \bigotimes_{j=1}^{d} \left(A^{(j)} v^{(j)}\right) \in \mathbf{W} \tag{1.4b}$$

for[6] all $v_j \in V_j$. Since $\mathbf{V}$ is spanned by elementary tensors (cf. (1.2)), equation (1.4b) defines $\mathbf{A}$ uniquely on $\mathbf{V}$ (more details in §3.3).

[6] In De Lathauwer et al. [41, Def. 8], the matrix-vector multiplication $\mathbf{A}\mathbf{v}$ by $\mathbf{A} := \bigotimes_{j=1}^{d} A^{(j)}$ is denoted by $\mathbf{v} \times_1 A^{(1)} \times_2 A^{(2)} \cdots \times_d A^{(d)}$, where $\times_j$ is called the *j-mode product*.

In the case of $V_j = \mathbb{R}^{I_j}$ and $W_j = \mathbb{R}^{J_j}$, the mappings $A^{(j)}$ are matrices from $\mathbb{R}^{I_j \times J_j}$. The Kronecker product $\bigotimes_{j=1}^{d} A^{(j)}$ belongs to the matrix space $\mathbb{R}^{\mathbf{I}\times\mathbf{J}}$ with

$$\mathbf{I} = I_1 \times \ldots \times I_d \quad \text{and} \quad \mathbf{J} = J_1 \times \ldots \times J_d.$$

For $d = 2$ let $I_1 = \{1, \ldots, n_1\}$, $I_2 = \{1, \ldots, n_2\}$, $J_1 = \{1, \ldots, m_1\}$, and $J_2 = \{1, \ldots, m_2\}$ be ordered index sets and use the lexicographical ordering[7] of the pairs (i, j) in $\mathbf{I} = I_1 \times I_2$ and $\mathbf{J} = J_1 \times J_2$. Then the matrix $A \otimes B \in \mathbb{R}^{\mathbf{I}\times\mathbf{J}}$ has the block form

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n_2}B \\ \vdots & & \vdots \\ a_{n_1 1}B & \cdots & a_{n_1 n_2}B \end{bmatrix}. \tag{1.5}$$

1.1.3 Tensor Product of Functions

Now, we redefine $I_j \subset \mathbb{R}$ as an interval and consider infinite dimensional vector spaces of functions like $V_j = C(I_j)$ or $V_j = L^2(I_j)$. $C(I_j)$ contains the continuous functions on I_j, while $L^2(I_j)$ are the measurable and square-integrable functions on I_j. The tensor product of univariate functions $f_j(x_j)$ is the d-variate function[8]

$$f := \bigotimes_{j=1}^{d} f_j \quad \text{with } f(x_1, \ldots, x_d) = \prod_{j=1}^{d} f_j(x_j) \quad (x_j \in I_j,\ 1 \le j \le d). \tag{1.6}$$

The product belongs to

$$\mathbf{V} = \bigotimes_{j=1}^{d} V_j, \quad \text{where } \mathbf{V} \subset C(\mathbf{I}) \text{ or } \mathbf{V} \subset L^2(\mathbf{I}), \text{ respectively.}$$

for $V_j = C(I_j)$ or $V_j = L^2(I_j)$ (details in §4 and §4.4).

In the infinite dimensional case, the definition (1.2) must be modified, if one wants to obtain a complete (Banach or Hilbert) space. The span of the elementary tensors must be closed with respect to a suitable norm (here norm of $C(\mathbf{I})$ or $L^2(\mathbf{I})$):

$$\bigotimes_{j=1}^{d} V_j = \overline{\operatorname{span}\{v^1 \otimes v^2 \otimes \ldots \otimes v^d :\ v^j \in V_j,\ 1 \le j \le d\}}. \tag{1.7}$$

[7] This is the ordering $(1,1), (1,2), \ldots, (1,n_2), (2,1), \ldots, (2,n_2), \ldots$ If another ordering or no ordering is defined, definition (1.5) is incorrect.

[8] According to Notation 1.1 we might write $f[x_1, x_2, \ldots, x_d]$ instead of $f(x_1, x_2, \ldots, x_d)$. In the sequel we use the usual notation of the argument list with round brackets.

The tensor structure of functions is often termed *separation of the variables*. This means that a multivariate function f can be written either as an elementary tensor product $\bigotimes_{j=1}^{d} f_j$ as in (1.6) or as a sum (series) of such products.

A particular example of a multivariate function is the polynomial

$$P(x_1,\ldots,x_d) = \sum_{\mathbf{i}} a_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}, \tag{1.8}$$

where each monomial $\mathbf{x}^{\mathbf{i}} := \prod_{j=1}^{d}(x_j)^{i_j}$ is an elementary product.

The definitions in §§1.1.1-3 may lead to the impression that there are different tensor products. This is only partially true. The cases of §§1.1.1-2 follow the same concept. In Chap. 3, the *algebraic* tensor product $\mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$ of general vector spaces V_j $(1 \leq j \leq d)$ will be defined. Choosing $V_j = \mathbb{R}^{I_j}$, we obtain tensors as in §1.1.1, while for matrix spaces $V_j = \mathbb{R}^{I_j \times J_j}$ the tensor product coincides with the Kronecker product.

The infinite dimensional case of §1.1.3 is different, since *topological* tensor spaces require a closure with respect to some norm (see Chap. 4).

1.2 Where do Tensors Appear?

At the first sight, tensors of order $d \geq 3$ do not seem to be used so often. Vectors (the particular case $d = 1$) appear almost everywhere. Since matrices (case $d = 2$) correspond to linear mappings, they are also omnipresent. The theory of vectors and matrices has led to the field of linear algebra. However, there are no standard constructions in *linear* algebra which lead to tensors of order $d \geq 3$. Instead, tensors are studied in the field of *multilinear* algebra.

1.2.1 Tensors as Coefficients

The first purpose of indexed quantities is a simplification of notation. For instance, the description of the polynomial (1.8) in, say, $d = 3$ variables is easily readable if coefficients a_{ijk} with three indices are introduced. In §1.6 we shall mention such an approach used already by Cayley in 1845.

Certain quantities in the partial differential equations of elasticity or in Maxwell's equations are called tensor (e.g., stress tensor). These tensors, however, are of order two, therefore the term 'matrix' would be more appropriate. Moreover, in physics the term 'tensor' is often used with the meaning 'tensor-valued function'.

In differential geometry, tensors are widely used for coordinate transformations. Typically, one distinguishes covariant and contravariant tensors and those of mixed type. The indices of the coefficients are placed either in lower position (covariant case) or in upper position (contravariant). For instance, ${a_{ij}}^{k}$ is a mixed tensor with two covariant and one contravariant component. For coordinate systems in $\mathbb{R}^n$, all indices vary in $\{1, \ldots, n\}$. The notational advantage of the lower and upper indices is the implicit Einstein summation rule: expressions containing a certain index in both positions are to be summed over this index. We give an example (cf. [132]). Let a smooth two-dimensional manifold be described by the function $\mathbf{x}(u^1, u^2)$. First and second derivatives with respect to these coordinates are denoted by $\mathbf{x}_{u^k}$ and $\mathbf{x}_{u^i,u^j}$. Together with the normal vector $\mathbf{n}$, the Gaussian formula for the second derivatives is

$$\mathbf{x}_{u^i,u^j} = {\mathit{\Gamma}_{ij}}^k \, \mathbf{x}_{u^k} + a_{ij}\mathbf{n} \qquad \text{(apply summation over } k\text{)},$$

where ${\mathit{\Gamma}_{ij}}^k$ are the Christoffel symbols[9] of second kind (cf. Christoffel [37], 1869).

The algebraic explanation of co- and contravariant tensor is as follows. The dual space to $V := \mathbb{R}^n$ is denoted by V'. Although V' is isomorphic to V, it is considered as a different vector space. Mixed tensors are elements of $\mathbf{V} = \bigotimes_{j=1}^d V_j$, where V_j is either V (contravariant component) or V' (covariant component). The summation rule performs the dual form $v'(v)$ of $v' \in V'$ and $v \in V$.

1.2.2 Tensor Decomposition for Inverse Problems

In many fields (psychometrics, linguistics, chemometrics,[10] telecommunication, biomedical applications, information extraction,[11] computer vision,[12] etc.) matrix-valued data appear. $M \in \mathbb{R}^{n \times m}$ may correspond to m measurements of different properties j, while i is associated to n different input data. For instance, in problems from chemometrics the input may be an excitation spectrum, while the output is the emission spectrum. Assuming a linear behaviour, we obtain for one substance a matrix ab^T of rank one. In this case, the inverse problem is trivial: the data ab^T allow to recover the vectors a and b up to a constant factor. Having a mixture of r substances, we obtain a matrix

$$M = \sum_{\nu=1}^{r} c_\nu \, a_\nu \, b_\nu^\mathsf{T} \qquad (a_\nu \in \mathbb{R}^n, b_\nu \in \mathbb{R}^m),$$

where $c_\nu \in \mathbb{R}$ is the concentration of substance ν. The componentwise version of the latter equation is

$$M_{ij} = \sum_{\nu=1}^{r} c_\nu \, a_{\nu i} \, b_{\nu j}.$$

[9] This notation is not used in Christoffel's original paper [37].

[10] See, for instance, Smile-Bro-Geladi [173] and De Lathauwer-De Moor-Vandevalle [42].

[11] See, for instance, Lu-Plataniotis-Venetsanopoulos [142].

[12] See, for instance, Wang-Ahuja [193].

With $A = [c_1 a_1 \; c_2 a_2 \; \ldots \; c_r a_r] \in \mathbb{R}^{n\times r}$ and $B = [b_1 \; b_2 \; \ldots \; b_r] \in \mathbb{R}^{m\times r}$, we may write

$$M = AB^{\mathsf{T}}.$$

Now, the inverse problem is the task to recover the factors A and B. This, however, is impossible since $A' = AT$ and $B' = T^{-\mathsf{T}}B$ satisfy $M = A'B'^{\mathsf{T}}$ for any regular matrix $T \in \mathbb{R}^{r\times r}$.

Tensors of order three come into play, when we repeat the experiments with varying concentrations $c_{\nu k}$ (concentration of substance ν in the k-th experiment). The resulting data are

$$M_{ijk} = \sum_{\nu=1}^{r} c_{\nu k}\, a_{\nu i}\, b_{\nu j}.$$

By definition (1.1), we can rewrite the latter equation as[13]

$$M = \sum_{\nu=1}^{r} a_\nu \otimes b_\nu \otimes c_\nu. \tag{1.9}$$

Under certain conditions, it is possible to recover the vectors $a_\nu \in \mathbb{R}^n$, $b_\nu \in \mathbb{R}^m$, $c_\nu \in \mathbb{R}^r$ from the data $M \in \mathbb{R}^{n\times m\times r}$ (up to scaling factors; cf. Remark 7.4b). In these application fields, the above 'inverse problem' is called 'factor analysis' or 'component analysis' (cf. [96], [42]).

These techniques have developed in the second part of the last century: Cattell [31] (1944), Tucker [184] (1966), Harshman [96] (1970), Appellof-Davidson [3] (1981) and many more (see review by Kolda-Bader [128]).

In this monograph, we shall *not* study these inverse problems. In §7.1.3, the difference between tensor *representations* and tensor *decompositions* will be discussed. Our emphasis lies on the tensor representation.

We remark that the tensors considered above cannot really be large-sized as long as all entries M_{ijk} can be stored.

1.2.3 Tensor Spaces in Functional Analysis

The analysis of topological tensor spaces has been started by Schatten [167] (1950) and Grothendieck [79]. Chapter 4 introduces parts of their concepts. However, most of the applications in functional analysis concern tensor products $X = V \otimes W$ of *two* Banach spaces. The reason is that these tensor spaces of order two can be related to certain linear operator spaces. The interpretation of X as tensor product may allow to transport certain properties from the factors V and W, which are easier to be analysed, to the product X which may be of a more complicated nature.

[13] Representations like (1.9) are used by Hitchcock [100] in 1927 (see §1.6).

1.2.4 Large-Sized Tensors in Analysis Applications

In analysis, the approximation of functions is well-studied. Usually, the quality of approximation is related to smoothness properties. If a function is the solution of a partial differential equation, a lot is known about its regularity (cf. [82, §9]). Below, we give an example how the concept of tensors may appear in the context of partial differential equations and their discretisations.

1.2.4.1 Partial Differential Equations

Let $\Omega = I_1 \times I_2 \times I_3 \subset \mathbb{R}^3$ be the product of three intervals and consider an elliptic differential equation $Lu = f$ on Ω, e.g., with Dirichlet boundary conditions $u = 0$ on the boundary $\Gamma = \partial\Omega$. A second order differential operator L is called *separable*, if

$$L = L_1 + L_2 + L_3 \text{ with } L_j = \frac{\partial}{\partial x_j} a_j(x_j) \frac{\partial}{\partial x_j} + b_j(x_j) \frac{\partial}{\partial x_j} + c_j(x_j). \tag{1.10a}$$

Note that any differential operator with constant coefficients and without mixed derivatives is of this kind. According to §1.1.3, we may consider the three-variate function as a tensor of order three. Moreover, the operator L can be regarded as a Kronecker product:

$$L = L_1 \otimes id \otimes id + id \otimes L_2 \otimes id + id \otimes id \otimes L_3. \tag{1.10b}$$

This tensor structure becomes more obvious, when we consider a finite difference discretisation of $Lu = f$. Assume, e.g., that $I_1 = I_2 = I_3 = [0, 1]$ and introduce the uniform grid $G_n = \left\{ \left(\frac{i}{n}, \frac{j}{n}, \frac{k}{n}\right) : 0 \leq i, j, k \leq n \right\}$ of grid size $h = 1/n$. The discrete values of u and f at the nodes of the grid are denoted by[14]

$$\mathbf{u}_{ijk} := u(\tfrac{i}{n}, \tfrac{j}{n}, \tfrac{k}{n}), \quad \mathbf{f}_{ijk} := f(\tfrac{i}{n}, \tfrac{j}{n}, \tfrac{k}{n}), \qquad (1 \leq i, j, k \leq n-1)\,. \tag{1.11a}$$

Hence, $\mathbf{u}$ and $\mathbf{f}$ are tensors of size $(n-1)\times(n-1)\times(n-1)$. The discretisation of the one-dimensional differential operator L_j from (1.10a) yields a tridiagonal matrix $L^{(j)} \in \mathbb{R}^{(n-1)\times(n-1)}$. As in (1.10b), the matrix of the *discrete* system $\mathbf{L}\mathbf{u} = \mathbf{f}$ is the Kronecker product

$$\mathbf{L} = L^{(1)} \otimes I \otimes I + I \otimes L^{(2)} \otimes I + I \otimes I \otimes L^{(3)}. \tag{1.11b}$$

$I \in \mathbb{R}^{(n-1)\times(n-1)}$ is the identity matrix. Note that $\mathbf{L}$ has size $(n-1)^3 \times (n-1)^3$.

The standard treatment of the system $\mathbf{L}\mathbf{u} = \mathbf{f}$ views $\mathbf{u}$ and $\mathbf{f}$ as vectors from $\mathbb{R}^N$ with $N := (n-1)^3$ and tries to solve N equations with N unknowns. If $n \approx 100$, a system with $N \approx 10^6$ equations can still be handled. However, for $n \approx 10000$ or

[14] Because of the boundary condition, $\mathbf{u}_{ijk} = 0$ holds if one the indices equals 0 or n.

even $n \approx 10^6$, a system of size $N \approx 10^{12}$ or $N \approx 10^{18}$ exceeds the capacity of standard computers.

If we regard $\mathbf{u}$ and $\mathbf{f}$ as tensors of $\mathbb{R}^{n-1} \otimes \mathbb{R}^{n-1} \otimes \mathbb{R}^{n-1}$, it might be possible to find tensor representations with much less storage. Consider, for instance, a uniform load $f = 1$. Then $\mathbf{f} = \mathbf{1} \otimes \mathbf{1} \otimes \mathbf{1}$ is an elementary tensor, where $\mathbf{1} \in \mathbb{R}^{n-1}$ is the vector with entries $\mathbf{1}_i = 1$. The matrix $\mathbf{L}$ is already written as Kronecker product (1.11b). In §9.7.2.6 we shall show that at least for positive definite $\mathbf{L}$ a very accurate inverse matrix $\mathbf{B} \approx \mathbf{L}^{-1}$ can be constructed and that the matrix-vector multiplication $\tilde{\mathbf{u}} = \mathbf{B}\mathbf{f}$ can be performed. The required storage for the representation of $\mathbf{B}$ and $\tilde{\mathbf{u}}$ is bounded by $O(n \log^2(1/\varepsilon))$, where ε is related to the error $\left\|\mathbf{L}^{-1} - \mathbf{B}\right\|_2 \leq \varepsilon$. The same bound holds for the computational cost.

The following observations are important:

1) Under suitable conditions, the exponential cost n^d can be reduced to $O(dn)$ (here: $d = 3$). This allows computations in cases, where the standard approach fails and not even the storage of the data $\mathbf{u}, \mathbf{f}$ can be achieved.

2) Usually, tensor computations will not be exact, but yield approximations. In applications from analysis, there are many cases where fast convergence holds. In the example from above the accuracy ε improves exponentially with a certain rank parameter, so that we obtain the logarithmic factor $\log^2(1/\varepsilon)$. Although such a behaviour is typical for many problems from analysis, it does not hold in general, in particular not for random data.

3) The essential key are tensor representations with two requirements. First, low storage cost is an obvious option. Since the represented tensors are involved into operations (here: the matrix-vector multiplication $\mathbf{B}\mathbf{f}$), the second option is that such tensor operations should have a comparably low cost.

Finally, we give an example, where the tensor structure can be successfully applied without any approximation error. Instead of the linear system $\mathbf{L}\mathbf{u} = \mathbf{f}$ from above, we consider the eigenvalue problem $\mathbf{L}\mathbf{u} = \lambda\mathbf{u}$. First, we discuss the undiscretised problem

$$Lu = \lambda u. \tag{1.12}$$

Here, it is well-known that the separation ansatz $u(x, y, z) = u_1(x)u_2(y)u_3(z)$ yields three one-dimensional boundary eigenvalue problems

$$L_1 u_1(x) = \lambda^{(1)} u_1, \quad L_2 u_2(y) = \lambda^{(2)} u_2, \quad L_3 u_3(z) = \lambda^{(3)} u_3$$

with zero conditions at $x, y, z \in \{0, 1\}$. The product $u(x, y, z) := u_1(x)u_2(y)u_3(z)$ satisfies $Lu = \lambda u$ with $\lambda = \lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}$. The latter product can be understood as tensor product: $u = u_1 \otimes u_2 \otimes u_3$ (cf. §1.1.3).

Similarly, we derive from the Kronecker product structure (1.11b) that the solutions of the *discrete* eigenvalue problems

$$L^{(1)} u_1 = \lambda^{(1)} u_1, \quad L^{(2)} u_2 = \lambda^{(2)} u_2, \quad L^{(3)} u_3 = \lambda^{(3)} u_3$$

in $\mathbb{R}^{n-1}$ yield the solution $\mathbf{u} = u_1 \otimes u_2 \otimes u_3$ of $\mathbf{L}\mathbf{u} = \lambda \mathbf{u}$ with $\lambda = \lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}$.

In this example we have exploited that the eigensolution is exactly[15] equal to an elementary tensor. In the discrete case, this implies that an object of size $(n-1)^3$ can be represented by three vectors of size $n-1$.

1.2.4.2 Multivariate Function Representation

The computational realisation of a special function $f(x)$ in one variable may be based on a rational approximation, a recursion etc. or a combination of these tools. The computation of a multivariate function $f(x_1, \ldots, x_p)$ is even more difficult. Such functions may be defined by complicated integrals involving parameters $x_1, \ldots, x_p$ in the integrand or integration domain. Consider the evaluation of f on $\mathbf{I} = I_1 \times \ldots \times I_p \subset \mathbb{R}^p$ with $I_j = [a_j, b_j]$. We may precompute f at grid points $(x_{1,i_1}, \ldots, x_{p,i_p})$, $x_{j,i_j} = a_j + i_j(b_j - a_j)/n$ for $0 \le i_j \le n$, followed by a suitable interpolation at the desired $\mathbf{x} = (x_1, \ldots, x_p) \in \mathbf{I}$. However, we fail as the required storage of the grid values is of size n^p. Again, the hope is to find a suitable tensor approximation with storage $O(pn)$ and an evaluation procedure of similar cost.

To give an example, a very easy task is the approximation of the function

$$\frac{1}{\|\mathbf{x}\|} = \left(\sum_{i=1}^{p} x_i^2 \right)^{-1/2} \qquad \text{for } \|\mathbf{x}\| \geq a > 0.$$

We obtain a uniform accuracy of size $O\big(\exp(-\pi\sqrt{r/2})/\sqrt{a}\big)$ with a storage of size $2r$ and an evaluation cost $O(rp)$. Details will follow in §9.7.2.5.2.

1.2.5 Tensors in Quantum Chemistry

The Schrödinger equation determines ‘wave functions’ $f(x_1, \ldots, x_d)$, where each variable $x_j \in \mathbb{R}^3$ corresponds to one electron. Hence, the spatial dimension $3d$ increases with the size of the molecule. A first ansatz[16] is $f(x_1, \ldots, x_d) \approx \Phi(x_1, \ldots, x_d) := \varphi_1(x_1)\varphi_2(x_2) \cdot \ldots \cdot \varphi_d(x_d)$, which leads to the Hartree-Fock equation. According to (1.6), we can write $\Phi := \bigotimes_{j=1}^{d} \varphi_j$ as a tensor. More accurate approximations require tensors being linear combinations of such products.

The standard ansatz for the three-dimensional functions $\varphi_j(\mathbf{x})$ are sums of Gaussian functions[17] $\Phi_\nu(\mathbf{x}) := \exp(\alpha_\nu \|\mathbf{x} - \mathbf{R}_\nu\|^2)$ as introduced by Boys [23] in 1950. Again, Φ_ν is the elementary tensor $\bigotimes_{k=1}^{3} e_k$ with $e_k(x_k) := \exp(\alpha_\nu (x_k - \mathbf{R}_{\nu,k})^2)$.

[15] This holds only for separable differential operators (cf. (1.10a)), but also in more general cases tensor approaches apply as shown in [91] (see §16.3).

[16] In fact, the product must be antisymmetrised yielding the Slater determinant from Lemma 3.72.

[17] Possibly multiplied by polynomials.

1.3 Tensor Calculus

The representation of tensors (in particular, with not too large storage requirements) is one goal of the efficient numerical treatment of tensors. Another goal is the efficient performance of tensor operations.

In the case of matrices, we apply matrix-vector and matrix-matrix multiplications and matrix inversions. The same operations occur for tensors, when the matrix is given by a Kronecker matrix and the vector by a tensor. Besides of these operations there are entry-wise multiplications, convolutions etc.

In linear algebra, basis transformations are well-known which lead to vector and matrix transforms. Such operations occur for tensors as well. There are matrix decompositions like the singular value decomposition. Generalisations to tensors will play an important rôle.

These and further operations are summarised under the term of 'tensor calculus'.[18] In the same way, as a library of matrix procedures is the basis for all algorithms in linear algebra, the tensor calculus enables computations in the world of tensors.

Note that already in the case of large-sized matrices, special efficient matrix representations are needed (cf. Hackbusch [86]), although the computational time grows only polynomially (typically cubically) with the matrix size. All the more important are efficient algorithms for tensors to avoid exponential run time.

1.4 Preview

1.4.1 Part I: Algebraic Properties

Matrices can be considered as tensors of second order. In Chap. 2 we summarise various properties of matrices as well as techniques applicable to matrices. QR and singular value decompositions will play an important rôle for later tensor operations.

In Chap. 3, tensors and tensor spaces are introduced. The definition of the tensor space in §3.2 requires a discussion of free vectors spaces (in §3.1.2) and of quotient spaces (in §3.1.3). Furthermore, linear and multilinear mappings and algebraic dual spaces are discussed in §3.1.4.

In §3.2 not only the tensor product and the (algebraic) tensor space are introduced, but also the (tensor) rank of a tensor is defined, which generalises the rank of a matrix. Later, we shall introduce further vector-valued ranks of tensors.

In §3.3 we have a closer look to linear and multilinear maps. In particular, tensor products of linear maps are discussed.

[18] The Latin word '*calculus*' is the diminutive of '*calx*' (lime, limestone) and has the original meaning 'pebble'. In particular, it denotes the pieces used in the Roman abacus. Therefore the Latin word '*calculus*' has also the meaning of 'calculation' or, in modern terms, 'computation'.

Tensor spaces with additional algebra structure are different from tensor algebras. Both are briefly described in §3.4.

In particular applications, symmetric or antisymmetric tensors are needed. These are defined in §3.5. Symmetric tensors are connected to quantics (cf. §3.5.2), while antisymmetric tensors are related to determinants (cf. §3.5.3).

1.4.2 Part II: Functional Analysis of Tensors

Normed tensor spaces are needed as soon as we want to approximate certain tensors. Even in the finite dimensional case one observes properties of tensors which are completely unknown from the matrix case. In particular in the infinite dimensional case, one has Banach (or Hilbert) spaces V_j endowed with a norm $\|\cdot\|_j$ as well as the algebraic tensor space $\mathbf{V}_{\text{alg}} = {}_a\bigotimes_{j=1}^d V_j$, which together with a norm $\|\cdot\|$ becomes a normed space. Completion yields the topological Banach space $\mathbf{V}_{\text{top}} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$. The tensor space norm $\|\cdot\|$ is by no means determined by the single norms $\|\cdot\|_j$. In §4.2 we study the properties of tensor space norms. It turns out that continuity conditions on the tensor product limit the choice of $\|\cdot\|$ (cf. §§4.2.2-7). There are two norms induced by $\{\|\cdot\|_j : 1 \leq j \leq d\}$, the *projective norm* (cf. §4.2.4) and the *injective norm* (cf. §4.2.7), which are the strongest and weakest possible norms. Further terms of interest are *crossnorms* (cf. §4.2.2), *reasonable crossnorms* (cf. §4.2.9), and *uniform crossnorms* (cf. §4.2.12). The case $d = 2$ considered in §4.2 allows to discuss nuclear and compact operators (cf. §4.2.13). The extension to $d \geq 3$ discussed in §4.3 is almost straightforward except that we also need suitable norms, e.g., for the tensor spaces ${}_a\bigotimes_{j\in\{1,\ldots,d\}\setminus\{k\}} V_j$ of order $d-1$.

While L^p or C^0 norms of tensor spaces belong to the class of crossnorms, the usual spaces C^m or H^m $(m \geq 1)$ cannot be described by crossnorms, but by intersections of Banach (or Hilbert) tensor spaces (cf. §4.3.6). The corresponding construction by crossnorms leads to so-called mixed norms.

Hilbert spaces are discussed in §4.4. In this case, the scalar products $\langle\cdot,\cdot\rangle_j$ of V_j define the *induced scalar product* of the Hilbert tensor space (cf. §4.4.1). In the Hilbert case, the infinite singular value decomposition can be used to define the Hilbert-Schmidt and the Schatten norms (cf. §4.4.3). Besides the usual scalar product the *partial* scalar product is of interest (cf. §4.5.4).

In §4.6 the tensor operations are enumerated which later are to be performed numerically.

Particular subspaces of the tensor space $\otimes^d V$ are the symmetric and antisymmetric tensor spaces discussed in §4.7.

Chapter 5 concerns algebraic as well as topological tensor spaces. We consider different isomorphisms which allow to regard tensors either as vectors (vectorisation in §5.1) or as matrices (matricisation in §5.2). In particular, the matricisation will become an important tool. The opposite direction is the *tensorisation* considered in

§5.3 and later, in more detail, in Chap. 14. Here, vectors from $\mathbb{R}^n$ are artificially reformulated as tensors.

Another important tool for the analysis and for concrete constructions are the *minimal subspaces* studied in Chap. 6. Given some tensor $\mathbf{v}$, we ask for the smallest subspaces U_j such that $\mathbf{v} \in \bigotimes_{j=1}^{d} U_j$. Of particular interest is their behaviour for sequences $\mathbf{v}_n \rightharpoonup \mathbf{v}$.

1.4.3 Part III: Numerical Treatment

The numerical treatment of tensors is based on a suitable *tensor representation*. Chapters 7 to 10 are devoted to two well-known representations, the r*-term format* (also called canonical or CP format) and the *tensor subspace format* (also called Tucker format). We distinguish the *exact* representation from the approximation task. Exact representations are discussed in Chap. 7 (r-term format) and Chap. 8 (tensor subspace format). If the tensor rank is moderate, the r-term format is a very good choice, whereas the tensor subspace format is disadvantageous for larger tensor order d because of its exponentially increasing storage requirement.

Tensor *approximations* are discussed separately in Chaps. 9 (r-term format) and 10 (tensor subspace format). In the first case, many properties known from the matrix case (see §9.3) do not generalise to tensor orders $d \geq 3$. A particular drawback is mentioned in §9.4: the set of r-term tensors is not closed, which may cause a numerical instability. An approximation of a tensor $\mathbf{v}$ by some $\tilde{\mathbf{v}}$ in the r-term format may be performed numerically using a regularisation (cf. §9.5). In some cases, analytical methods allow to determine very accurate r-term approximations to functions and operators (cf. §9.7).

In the case of tensor subspace approximations (§10), there are two different options. The simpler approach is based on the *higher order singular value decomposition* (HOSVD; cf. §10.1). This allows a projection to smaller rank similar to the standard singular value decomposition in the matrix case. The result is not necessarily the best one, but quasi-optimal. The second option is the best-approximation considered in §10.2. In contrast to the r-term format, the existence of a best-approximation is guaranteed. A standard numerical method for its computation is the alternating least-squares method (ALS, cf. §10.3). For particular cases, analytical methods are available to approximate multivariate functions (cf. §10.4).

While the r-term format suffers from a possible numerical instability, the storage size of the tensor subspace format increases exponentially with the tensor order d. A format avoiding both problems is the *hierarchical format* described in Chap. 11. Here, the storage is strictly bounded by the product of the tensor order d, the maximal involved rank, and the maximal dimension of the vector spaces V_j. Again, HOSVD techniques can be used for a quasi-optimal truncation. Since the format is closed, numerical instability does not occur.

The hierarchical format is based on a dimension partition tree. A particular choice of the tree leads to the *matrix product representation* or *TT format* described in Chap. 12.

The essential part of the numerical tensor calculus is the performance of *tensor operations*. In Chap. 13 we describe all operations, their realisation in the different formats, and the corresponding computational cost.

The *tensorisation* briefly mentioned in §5.3 is revisited in Chap. 14. When applied to grid functions, tensorisation corresponds to a multiscale approach. The tensor truncation methods allow an efficient compression of the data size. As shown in §14.2, the approximation can be proved to be at least as good as analytical methods like hp-methods, exponential sum approximations, or wavelet compression techniques. In §14.3 the performance of the convolution is described. The fast Fourier transform is explained in §14.4. The method of tensorisation can also be applied to functions instead of grid functions as detailed in §14.5.

Chapter 15 is devoted to the *generalised cross approximation*. The underlying problem is the approximation of general tensors, which has several important applications.

In Chap. 16, the application of the tensor calculus to elliptic boundary value problems (§16.2) and eigenvalue problems (§16.3) is discussed.

The final Chap. 17 collects a number of further topics. §17.1 considers general minimisation problems. Another minimisation approach described in §17.2 applies directly to the parameters of the tensor representation. Dynamic problems are studied in §17.3, while the ANOVA method is mentioned in §17.4.

1.4.4 Topics Outside the Scope of the Monograph

As already mentioned in §1.2.2, we do not aim at *inverse problems*, where the parameters of the representation (decomposition) have a certain external interpretation (see references in Footnotes 10-12).

Also in *data mining* high-dimensional tensors arise (cf. Kolda - Sun [129]). However, in contrast to mathematical applications (e.g., in partial differential equations) weaker properties hold concerning data smoothness, desired order of accuracy, and often availability of data.

We do not consider *data completion* (approximation of incomplete data; cf. [140]), which is a typical problem for data from non-mathematical sources. Entries $\mathbf{v}[\mathbf{i}]$ of a tensor $\mathbf{v} \in \bigotimes_{j=1}^{d} \mathbb{R}^{n_j}$ may be available only for $\mathbf{i} \in \mathring{\mathbf{I}}$ of a subset $\mathring{\mathbf{I}} \subset \mathbf{I} := \times_{j=1}^{d} \{1, \ldots, n_j\}$. Another example are cases where data are lost or deleted. Approximation of the remaining data by a tensor $\tilde{\mathbf{v}}$ of a certain format yields the desired completion (cf. Footnote 9 on page 262). Instead, in Chap. 15 we are discussing quite another kind of data completion, where an approximation $\tilde{\mathbf{v}}$ is constructed by a small part of the data, but in contrast to the usual data completion problem, we assume that *all* data are available on demand, although possibly with high arithmetical cost.

Another subject, which is not discussed here, is the detection and determination of *principal manifolds* of smaller dimension ('manifold learning'; see, e.g., Feuersänger-Griebel [59]).

The order d of the tensor considered here, is always finite. In fact, the numerical cost of storage or arithmetical operations is at least increasing linearly in d. Infinite dimensions ($d = \infty$) may appear theoretically (as in §15.1.2.2), but only truncations to finite d are discussed.

There are purely algebraic approaches to tensors which try to generalise terms from linear algebra to multilinear algebra including certain decompositions (cf. [160, §5]). Unfortunately, constructive algorithms in this field are usually NP hard and do not help for large-sized tensors.

1.5 Software

Free software for tensor applications is offered by the following groups:

- MATLAB Tensor Toolbox by Bader-Kolda [4]:
  `http://csmr.ca.sandia.gov/˜tgkolda/TensorToolbox`
- Hierarchical Tucker Toolbox by Tobler-Kressner [131]:
  `http://www.sam.math.ethz.ch/NLAgroup/software.html`
- TT TOOLBOX by I. Oseledets: `http://spring.inm.ras.ru/osel`
- TensorCalculus by H. Auer, M. Espig, S. Handschuh, and P. Wähnert:
  `http://gitorious.org/tensorcalculus/pages/Home`

1.6 Comments about the Early History of Tensors

The word 'tensor' seems to be used for the first time in an article by William Rowan Hamilton [95] from 1846. The meaning, however, was quite different. Hamilton is well-known for his quaternions. Like complex numbers, a modulus of a quaternion can be defined. For this non-negative real number he introduced the name 'tensor'. The word 'tensor' is used again in a book by Woldemar Voigt [192] in 1898 for quantities which come closer to our understanding.[19]

In May 1845, Arthur Cayley [32] submitted a paper, in which he described *hyperdeterminants*.[20] There he considers tensors of general order. For instance, he gives an illustration of a tensor from $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$ (p. 11 in [192]):

[19] From [192, p. 20]: Tensors are "… Zustände, die durch eine Zahlgrösse und eine zweiseitige Richtung charakterisiert sind. … Wir wollen uns deshalb nur darauf stützen, dass Zustände der geschilderten Art bei Spannungen und Dehnungen nicht starrer Körper auftreten, und sie deshalb *tensorielle*, die für sie charakteristischen physikalischen Grössen aber *Tensoren* nennen."

[20] See also [44, §5.3]. The hyperdeterminant vanishes for a tensor $\mathbf{v} \in \mathbb{R}^p \otimes \mathbb{R}^q \otimes \mathbb{R}^r$ if and only if the associated multilinear form $\varphi(x, y, z) := \sum_{i,j,k} \mathbf{v}[i, j, k] x_i y_j z_k$ allows nonzero vectors x, y, z such that $\nabla_x \varphi$, $\nabla_y \varphi$, or $\nabla_z \varphi$ vanish at (x, y, z).

> Soit $n = 3$, posons pour plus simpicité $m = 2$, et prenons
>
> $$\begin{aligned} 111 &= a, & 112 &= e,\\ 211 &= b, & 212 &= b,\\ 121 &= c, & 122 &= g,\\ 221 &= d, & 222 &= h, \end{aligned}$$
>
> de manière que la fonction à considérer est
>
> $$\begin{aligned} U = {} & ax_1y_1z_1 + bx_2y_1z_1 + cx_1y_2z_1 + dx_2y_2z_1 + \\ & + ex_1y_1z_2 + fx_2y_1z_2 + gx_1y_2z_2 + hx_2y_2z_2. \end{aligned}$$

Next, he considers linear transformations $\Lambda^{(1)}$, $\Lambda^{(2)}$, $\Lambda^{(3)}$ in all three directions, e.g., $\Lambda^{(1)}$ is described as follows.

> Les équations pour la transformation sont
>
> $$\begin{aligned} x_1 &= \lambda_1^1 \dot{x}_1 + \lambda_1^2 \dot{x}_2,\\ x_2 &= \lambda_2^1 \dot{x}_1 + \lambda_2^2 \dot{x}_2,\\ &\vdots \end{aligned}$$

The action of the transformations $\Lambda^{(1)}$, $\Lambda^{(2)}$, $\Lambda^{(3)}$ represents already the Kronecker product $\Lambda^{(1)} \otimes \Lambda^{(2)} \otimes \Lambda^{(3)}$.

The paper of Hitchcock [100] from 1927 has a similar algebraic background. The author states that 'any covariant tensor $A_{i_1..i_p}$ can be expressed as the sum of a finite number of which is the product of p covariant vectors'. In [100] the ranks are defined which we introduce in Definition 5.7. Although in this paper he uses the name 'tensor', in the following paper [99] of the same year he prefers the term 'matrix' or 'p-way matrix'. The tensor product of vectors a, b is denoted by ab without any special tensor symbol.

In §1.1.2 we have named the tensor product of matrices 'Kronecker product'. In fact, this term is well-introduced, but historically it seems to be unfounded. The 'Kronecker product' (and its determinant) was first studied by Johann Georg Zehfuss [200] in 1858, while it is questionable whether there exists any notice of Kronecker about this product (see [107] for historical remarks). Zehfuss' result about determinants can be found in Exercise 4.134.

1.7 Notations

A list of symbols, letters etc. can be found on page xix. Here, we collect the notational conventions which we use in connection with vectors, matrices, and tensors.

Index Sets. I, J, K are typical letters used for index sets. In general, we do not require that an index set is ordered. This allows, e.g., to define a new index set

$K := I \times J$ as the product of index sets I, J without prescribing an ordering of the pairs (i,j) of $i \in I$ and $j \in J$.

Fields. A vector space is associated with some field, which will be denoted by $\mathbb{K}$. The standard choices[21] are $\mathbb{R}$ and $\mathbb{C}$. When we use the symbol $\mathbb{K}$ instead of the special choice $\mathbb{R}$, we use the complex-conjugate value $\overline{\lambda}$ of a scalar whenever this is required in the case of $\mathbb{K} = \mathbb{C}$.

Vector Spaces $\mathbb{K}^n$ and $\mathbb{K}^I$. Let $n \in \mathbb{N}$. $\mathbb{K}^n$ is the standard notation for the vector space of the n-tuples $v = (v_i)_{i=1}^n$ with $v_i \in \mathbb{K}$**.** The more general notation $\mathbb{K}^I$ abbreviates the vector space $\{v = (v_i)_{i \in I} : v_i \in \mathbb{K}\}$. Equivalently, one may define $\mathbb{K}^I$ as the space of mappings from I into $\mathbb{K}$. Note that this definition makes sense for non-ordered index sets. If, e.g., $K = I \times J$ is the index set, a vector $v \in \mathbb{K}^K$ has entries $v_k = v_{(i,j)}$ for $k = (i,j) \in K$. The notation $v_{(i,j)}$ must be distinguished from $v_{i,j}$ which indicates a matrix entry. The simple notation $\mathbb{K}^n$ is identical to $\mathbb{K}^I$ for $I = \{1, \ldots, n\}$.

Vectors will be symbolised by small letters. Vector entries are usually denoted by v_i. The alternative notation $v[i]$ is used if the index carries a secondary index (example: $v[i_1]$) or if the symbol for the vector is already indexed (example: $v_\nu[i]$ for $v_\nu \in \mathbb{K}^I$).

Typical symbols for vector spaces are V, W, U, etc. Often, U is used for subspaces.

Matrices and Matrix Spaces $\mathbb{K}^{I \times J}$. Any linear mapping $\Phi : \mathbb{K}^I \to \mathbb{K}^J$ (I, J index sets) can be represented by means of a matrix[22]

$$M \in \mathbb{K}^{I \times J}$$

with entries $M_{ij} \in \mathbb{K}$ and one may write $M = (M_{ij})_{i \in I, j \in J}$ or $M = (M_{ij})_{(i,j) \in I \times J}$. The alternative notation $\mathbb{K}^{n \times m}$ is used for the special index sets $I = \{1, \ldots, n\}$ and $J = \{1, \ldots, m\}$. Even the mixed notation $\mathbb{K}^{I \times m}$ appears if $J = \{1, \ldots, m\}$, while I is a general index set.

Matrices will be symbolised by capital letters. Matrix entries are denoted by $M_{i,j} = M_{ij}$ or by $M[i,j]$. Given a matrix $M \in \mathbb{K}^{I \times J}$, its i-th row or its j-th column will be denoted by

$$M_{i,\bullet} = M[i,\bullet] \in \mathbb{K}^J \quad \text{or} \quad M_{\bullet,j} = M[\bullet,j] \in \mathbb{K}^I, \text{ respectively.}$$

If $\tau \subset I$ and $\sigma \subset J$ are index subsets, the restriction of a matrix is written as

$$M|_{\tau \times \sigma} = (M_{ij})_{i \in \tau, j \in \sigma} \in \mathbb{K}^{\tau \times \sigma}.$$

More about matrix notations will follow in §2.1.

[21] Fields of finite characteristic are of less interest, since approximations do not make sense. Nevertheless, there are applications of tensor tools for Boolean data (cf. Lichtenberg-Eichler [138]).

[22] $M \in \mathbb{K}^{I \times J}$ is considered as matrix, whereas $v \in \mathbb{K}^K$ for $K = I \times J$ is viewed as vector.

Tensors. Tensors are denoted by small bold type letters: $\mathbf{v}, \mathbf{w}, \ldots, \mathbf{a}, \mathbf{b}, \ldots$ Their entries are usually indexed in square brackets: $\mathbf{v}[i_1, \ldots, i_d]$. Only in simple cases, subscripts are used: $\mathbf{v}_{ijk}$. The bold type notation $\mathbf{v}[i_1, \ldots, i_d]$ is also used in the case of a variable d which possibly takes the values $d = 1$ [vector case] or 2 [matrix case].

The standard notation for a tensor space of order d is

$$\mathbf{V} = \bigotimes_{j=1}^{d} V_j .$$

Here, V_j $(1 \leq j \leq d)$ are vector spaces generating the tensor space $\mathbf{V}$. As in this example, tensor spaces are denoted by capital letters in bold type. $\mathbf{U}$ is the typical letter for a subspace of a tensor space.

Elementary tensors from $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ have the form

$$\mathbf{v} = \bigotimes_{j=1}^{d} v^{(j)} = v^{(1)} \otimes \ldots \otimes v^{(d)}.$$

The superscript in round brackets indicates the vector corresponding to the j-th direction. The preferred letter for the direction index is j (or k, if a second index is needed). The entries of $v^{(j)}$ may be written as $v_i^{(j)}$ or $v^{(j)}[i]$. A lower subscript may also denote the ν-th vector $v_\nu^{(j)} \in V_j$ as required in $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$. In this case, the entries of $v_\nu^{(j)}$ are written as $v_\nu^{(j)}[i]$.

To be precise, we have to distinguish between algebraic and topological tensor spaces denoted by ${}_a\bigotimes_{j=1}^{d} V_j$ and ${}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$, respectively. Details can be found in Notation 3.8.

Chapter 2
Matrix Tools

Abstract In connection with tensors, matrices are of interest for two reasons. Firstly, they are tensors of order two and therefore a nontrivial example of a tensor. Differently from tensors of higher order, matrices allow to apply practically realisable decompositions. Secondly, operations with general tensors will often be reduced to a sequence of matrix operations (realised by well-developed software). *Sections* 2.1–2.3 introduce the notation and recall well-known facts about matrices. *Section* 2.5 discusses the important QR decomposition and the singular value decomposition (SVD) and their computational cost. The (optimal) approximation by matrices of lower rank explained in *Sect.* 2.6 will be used later in truncation procedures for tensors. In Part III we shall apply some linear algebra procedures introduced in *Sect.* 2.7 based on QR and SVD.

2.1 Matrix Notations

In this subsection, the index sets I, J are assumed to be finite. As soon as complex conjugate values appear,[1] the scalar field is restricted to $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

We recall the notation $\mathbb{K}^{I\times J}$ explained in §1.7. The *entries* of a matrix $M \in \mathbb{K}^{I\times J}$ are denoted by M_{ij} $(i \in I, j \in J)$. Vice versa, numbers $\alpha_{ij} \in \mathbb{K}$ $(i \in I, j \in J)$ may be used to define $M := (\alpha_{ij})_{i\in I,\, j\in J} \in \mathbb{K}^{I\times J}$.

Let $j \in J$. The j-th *column* of $M \in \mathbb{K}^{I\times J}$ is the vector $M[\bullet, j] = (M_{ij})_{i\in I} \in \mathbb{K}^I$, while vectors $c^{(j)} \in \mathbb{K}^I$ generate a matrix $M := [\, c^{(j)} : j \in J \,] \in \mathbb{K}^{I\times J}$. If J is ordered, we may write $M := [\, c^{(j_1)}, c^{(j_2)}, \ldots]$.

δ_{ij} $(i, j \in I)$ is the *Kronecker symbol* defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \in I, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

[1] In the case of $\mathbb{K} = \mathbb{R}$, $\alpha = \bar{\alpha}$ holds for all $\alpha \in \mathbb{K}$.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_2,

The *unit vector* $e^{(i)} \in \mathbb{K}^I$ ($i \in I$) is defined by

$$e^{(i)} := (\delta_{ij})_{j\in I} \ . \tag{2.2}$$

The symbol $I = (\delta_{ij})_{i,j\in I}$ is used for the *identity matrix*. Since matrices and index sets do not appear at the same place, the simultaneous use of I for a matrix and for an index set should not lead to any confusion (example: $I \in \mathbb{K}^{I\times I}$).

If $M \in \mathbb{K}^{I\times J}$, the *transposed matrix* $M^{\mathsf{T}} \in \mathbb{K}^{J\times I}$ is defined by $M_{ij} = (M^{\mathsf{T}})_{ji}$ ($i \in I$, $j \in J$). A matrix from $\mathbb{K}^{I\times I}$ is *symmetric* if $M = M^{\mathsf{T}}$.

The *Hermitean transposed matrix* $M^{\mathsf{H}} \in \mathbb{K}^{J\times I}$ is $\overline{M^{\mathsf{T}}}$, i.e., $\overline{M_{ij}} = (M^{\mathsf{T}})_{ji}$, where $\overline{\bullet}$ is the complex conjugate value. If $\mathbb{K} = \mathbb{R}$, $M^{\mathsf{H}} = M^{\mathsf{T}}$ holds. This allows us to use $^{\mathsf{H}}$ for the general case of $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. A *Hermitean* matrix satisfies $M = M^{\mathsf{H}}$.

The *range* of a matrix $M \in \mathbb{K}^{I\times J}$ is[2]

$$\mathrm{range}(M) := \{Mx : x \in \mathbb{K}^J\}.$$

The *Euclidean scalar product* in $\mathbb{K}^I$ is given by

$$\langle x, y\rangle = y^{\mathsf{H}}x = \sum_{i\in I} x_i\, \overline{y_i}\,,$$

where in the real case $\mathbb{K} = \mathbb{R}$ the conjugate sign can be ignored. In the case of $\mathbb{K} = \mathbb{C}$, the scalar product is a *sesquilinear* form, i.e., it is antilinear in the second argument.[3]

Two vectors $x, y \in \mathbb{K}^I$ are *orthogonal* (symbolic notation: $x \perp y$), if $\langle x, y\rangle = 0$. A family of vectors $\{x_\nu\}_{\nu\in F} \subset \mathbb{K}^I$ is *orthogonal*, if the vectors are pairwise orthogonal, i.e., $\langle x_\nu, x_\mu\rangle = 0$ for all $\nu, \mu \in F$ with $\nu \neq \mu$.

Similarly, two vectors $x, y \in \mathbb{K}^I$ or a family $\{x_\nu\}_{\nu\in F} \subset \mathbb{K}^I$ is *orthonormal*, if, in addition, all vectors are normalised: $\langle x, x\rangle = \langle y, y\rangle = 1$ or $\langle x_\nu, x_\nu\rangle = 1$ ($\nu \in F$).

A matrix $M \in \mathbb{K}^{I\times J}$ is called *orthogonal*, if the columns of M are orthonormal. An equivalent characterisation is

$$M^{\mathsf{H}}M = I \in \mathbb{K}^{J\times J}. \tag{2.3}$$

Note that the (Hermitean) transpose of an orthogonal matrix is, in general, not orthogonal. $M \in \mathbb{K}^{I\times J}$ can be orthogonal only if $\#J \leq \#I$.

An orthogonal square[4] matrix $M \in \mathbb{K}^{I\times I}$ is called *unitary* (if $\mathbb{K} = \mathbb{R}$, often the term ‘orthogonal’ is preferred). Differently from the remark above, unitary matrices satisfy

$$M^{\mathsf{H}}M = MM^{\mathsf{H}} = I \in \mathbb{K}^{I\times I},$$

i.e., $M^{\mathsf{H}} = M^{-1}$ holds.

Assume that the index sets satisfy either $I \subset J$ or $J \subset I$. Then a (rectangular) matrix $M \in \mathbb{K}^{I\times J}$ is *diagonal*, if $M_{ij} = 0$ for all $i \neq j$, $(i, j) \in I \times J$. Given numbers

[2] Also the notation $\mathrm{colspan}(M)$ exists, since $\mathrm{range}(M)$ is spanned by the columns of M.

[3] A mapping φ is called *antilinear*, if $\varphi(x + \alpha y) = \varphi(x) + \overline{\alpha}\varphi(y)$ for $\alpha \in \mathbb{C}$.

[4] We may assume $M \in \mathbb{K}^{I\times J}$ with $\#I = \#J$ and different I, J. Then $M^{\mathsf{H}}M = I \in \mathbb{K}^{I\times I}$ and $MM^{\mathsf{H}} = I \in \mathbb{K}^{J\times J}$ are the precise conditions.

$\delta_i \ (i \in I \cap J)$, the associated diagonal matrix M with $M_{ii} = \delta_i$ is written as

$$\operatorname{diag}\{\delta_i : i \in I \cap J\}.$$

If the index set $I \cap J$ is ordered, an enumeration of the diagonal entries can be used: $\operatorname{diag}\{\delta_{i_1}, \delta_{i_2}, \ldots\}$.

Assume again $I \subset J$ or $J \subset I$ and a common ordering of $I \cup J$. A (rectangular) matrix $M \in \mathbb{K}^{I \times J}$ is *lower triangular*, if $M_{ij} = 0$ for all $(i, j) \in I \times J$ with $i > j$. Similarly, $M_{ij} = 0$ for all $i < j$ defines the *upper triangular* matrix.

2.2 Matrix Rank

Remark 2.1. Let $M \in \mathbb{K}^{I \times J}$. The following statements are equivalent and may be used as definition of the 'matrix rank' $r = \operatorname{rank}(M)$:
(a) $r = \dim \operatorname{range}(M)$,
(b) $r = \dim \operatorname{range}(M^{\mathsf{T}})$,
(c) r is the maximal number of linearly independent rows in M,
(d) r is the maximal number of linearly independent columns in M,
(e) $r \in \mathbb{N}_0$ is minimal with the property

$$M = \sum_{i=1}^{r} a_i b_i^{\mathsf{T}}, \qquad \text{where } a_i \in \mathbb{K}^I \text{ and } b_i \in \mathbb{K}^J, \tag{2.4}$$

(f) r is maximal with the property that there exists a regular $r \times r$ submatrix[5] of M.
(g) r is the number of positive singular values (see (2.19a)).

In (b) and (e) we may replace $\bullet^{\mathsf{T}}$ by $\bullet^{\mathsf{H}}$. Part (e) states in particular that products $a_i b_i^{\mathsf{T}}$ of non-vanishing vectors represent all rank-1 matrices.

The rank of $M \in \mathbb{K}^{I \times J}$ is bounded by the *maximal rank*

$$r_{\max} := \min\{\#I, \#J\}, \tag{2.5}$$

and this bound is attained for the so-called *full-rank* matrices.

The definition of linear independency depends on the field $\mathbb{K}$. This leads to the following question. A real matrix $M \in \mathbb{R}^{I \times J}$ may also be considered as an element of $\mathbb{C}^{I \times J}$. Hence, in principle, such an M may possess a 'real' rank and a 'complex' rank. However, the equivalent characterisations (f) and (g) are independent of the choice $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and prove the next remark.

Remark 2.2. For $M \in \mathbb{R}^{I \times J} \subset \mathbb{C}^{I \times J}$ the value of $\operatorname{rank}(M)$ is independent of the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Corollary 2.3. Let $r = \operatorname{rank}(M)$ and define $A := [a_1, \ldots, a_r]$ and $B := [b_1, \ldots, b_r]$ by a_i and b_i from (2.4). Then (2.4) is equivalent to $M = AB^{\mathsf{T}}$.

[5] That means that there are $I' \subset I$ and $J' \subset J$ with $\#I' = \#J' = r$ and $M|_{I' \times J'}$ regular.

An interesting matrix family is the set of matrices of rank not exceeding r:

$$\mathcal{R}_r := \left\{ M \in \mathbb{K}^{I \times J} : \operatorname{rank}(M) \le r \right\}. \tag{2.6}$$

Any $M \in \mathcal{R}_r$ may be written in the form (2.4).

Lemma 2.4. *The sets $\mathcal{R}_r \subset \mathbb{K}^{I \times J}$ for $r \in \mathbb{N}_0$ are closed. Any convergent sequence $R^{(k)} \in \mathcal{R}_r$ satisfies*

$$\liminf_{k\to\infty} \operatorname{rank}(R^{(k)}) \ge \operatorname{rank}\left(\lim_{k\to\infty} R^{(k)} \right). \tag{2.7}$$

Proof. For $s \in \mathbb{N}_0$ set $N_s := \left\{ k \in \mathbb{N} : \operatorname{rank}(R^{(k)}) = s \right\} \subset \mathbb{N}$ and

$$r_\infty = \min\{s \in \mathbb{N}_0 : \#N_s = \infty\} = \liminf_{k\to\infty} \operatorname{rank}(R^{(k)}).$$

We restrict $R^{(k)}$ to the subsequence with $k \in N_{r_\infty}$, i.e., $\operatorname{rank}(R^{(k)}) = r_\infty$. For full rank, i.e., $r_\infty = \min\{\#I_1, \#I_2\}$, nothing is to be proved. Otherwise, we use the criterion from Remark 2.1f: all $(r_\infty + 1) \times (r_\infty + 1)$ submatrices $R^{(k)}|_{I_1' \times I_2'}$ ($\#I_1' = \#I_2' = r_\infty + 1$) are singular, in particular, $\det\left(R^{(k)}|_{I_1' \times I_2'}\right) = 0$. Since the determinant is continuous, $0 = \lim \det(R^{(k)}|_{I_1' \times I_2'}) = \det(\lim(R^{(k)})|_{I_1' \times I_2'})$ proves that $\operatorname{rank}(\lim R^{(k)}) \le r_\infty$. □

Remark 2.5. A matrix $M \in \mathbb{K}^{I \times J}$ with random entries has maximal rank $r_{\max}$ with probability one.

Proof. Matrices of smaller rank form a subset of measure zero. □

2.3 Matrix Norms

Before the Euclidean, spectral and Frobenius norms will be discussed, the trace of a square matrix is introduced. For a generalisation of the trace mapping to operators see (4.60).

Definition 2.6. The mapping trace : $\mathbb{K}^{I \times I} \to \mathbb{K}$ is defined by

$$\operatorname{trace}(M) := \sum_{i \in I} M_{ii}. \tag{2.8}$$

Exercise 2.7. (a) $\operatorname{trace}(AB) = \operatorname{trace}(BA)$ for any $A \in \mathbb{K}^{I \times J}$ and $B \in \mathbb{K}^{J \times I}$.
(b) $\operatorname{trace}(M) = \operatorname{trace}(UMU^{\mathsf{H}})$ for $M \in \mathbb{K}^{I \times I}$ and any orthogonal matrix $U \in \mathbb{K}^{J \times I}$ (in particular, for a unitary matrix $U \in \mathbb{K}^{I \times I}$).
(c) Let λ_i ($i \in I$) be all eigenvalues of $M \in \mathbb{K}^{I \times I}$ according to their multiplicity. Then $\operatorname{trace}(M) = \sum_{i \in I} \lambda_i$.

The general definition of norms and scalar products can be found in §4.1.1 and §4.4.1. The *Frobenius norm*

$$\|M\|_{\mathsf{F}} = \sqrt{\sum_{i\in I, j\in J} |M_{i,j}|^2} \qquad \text{for } M \in \mathbb{K}^{I\times J} \tag{2.9}$$

is also called *Schur norm* or *Hilbert-Schmidt norm*. This norm is generated by the scalar product

$$\langle A, B\rangle_{\mathsf{F}} := \sum_{i\in I, j\in J} A_{i,j}\overline{B_{i,j}} = \operatorname{trace}(AB^{\mathsf{H}}) = \operatorname{trace}(B^{\mathsf{H}}A), \tag{2.10}$$

since $\langle M, M\rangle_{\mathsf{F}} = \|M\|_{\mathsf{F}}^2$. In particular, $\|M\|_{\mathsf{F}}^2 = \operatorname{trace}(MM^{\mathsf{H}}) = \operatorname{trace}(M^{\mathsf{H}}M)$ holds.

Remark 2.8. Let $I \times J$ and $I' \times J'$ define two matrix formats with the same number of entries: $\#I \cdot \#J = \#I' \cdot \#J'$. Any bijective mapping $\pi : I \times J \to I' \times J'$ generates a mapping $P : M \in \mathbb{K}^{I\times J} \mapsto P(M) = M' \in \mathbb{K}^{I'\times J'}$ via $M'[i', j'] = M[i, j]$ for $(i', j') = \pi(i, j)$. Then, the Frobenius norm and scalar product are invariant with respect to P, i.e.,

$$\|P(M)\|_{\mathsf{F}} = \|M\|_{\mathsf{F}} \quad \text{and} \quad \langle P(A), P(B)\rangle_{\mathsf{F}} = \langle A, B\rangle_{\mathsf{F}}.$$

Let $\|\cdot\|_X$ and $\|\cdot\|_Y$ be vector norms on $X = \mathbb{K}^I$ and $Y = \mathbb{K}^J$, respectively. Then the *associated matrix norm* is

$$\|M\| := \|M\|_{X\leftarrow Y} := \sup\left\{\frac{\|My\|_X}{\|y\|_Y} : 0 \neq y \in \mathbb{K}^J\right\} \quad \text{for } M \in \mathbb{K}^{I\times J}. \tag{2.11}$$

If $\|\cdot\|_X$ and $\|\cdot\|_Y$ coincide with the *Euclidean vector norm*

$$\|u\|_2 := \sqrt{\sum\nolimits_{i\in K} |u_i|^2} \qquad \text{for } u \in \mathbb{K}^K, \tag{2.12}$$

the associated matrix norm $\|M\|_{X\leftarrow Y}$ is the *spectral norm* denoted by $\|M\|_2$.

Exercise 2.9. Let $M \in \mathbb{K}^{I\times J}$. (a) Another equivalent definition of $\|\cdot\|_2$ is

$$\|M\|_2 = \sup\left\{\frac{y^{\mathsf{H}}Mx}{\sqrt{y^{\mathsf{H}}y \cdot x^{\mathsf{H}}x}} : 0 \neq x \in \mathbb{K}^J,\, 0 \neq y \in \mathbb{K}^I\right\}. \tag{2.13}$$

(b) $\|M\|_2 = \|UM\|_2 = \|MV^{\mathsf{H}}\|_2 = \|UMV^{\mathsf{H}}\|_2$ holds for orthogonal matrices $U \in \mathbb{K}^{I'\times I}$ and $V \in \mathbb{K}^{J'\times J}$.

From Lemma 2.20b we shall learn that the squared spectral norm $\|M\|_2^2$ is the largest eigenvalue of both $M^{\mathsf{H}}M$ and MM^{H}.

Both matrix norms $\|\cdot\|_2$ and $\|\cdot\|_{\mathsf{F}}$ are submultiplicative, i.e., $\|AB\| \leq \|A\|\,\|B\|$. The example of $A = B = I \in \mathbb{R}^{n\times n}$ shows the equality in $1 = \|I \cdot I\|_2 \leq \|I\|_2\,\|I\|_2 = 1$, while $\sqrt{n} = \|I \cdot I\|_{\mathsf{F}} \leq \|I\|_{\mathsf{F}}\,\|I\|_{\mathsf{F}} = n$ is a rather pessimistic estimate. In fact, spectral and Frobenius norms can be mixed to get better estimates.

Lemma 2.10. *The product of $A \in \mathbb{K}^{I\times J}$ and $B \in \mathbb{K}^{J\times K}$ is estimated by*

$$\|AB\|_{\mathsf{F}} \le \|A\|_2 \, \|B\|_{\mathsf{F}} \quad \textit{as well as} \quad \|AB\|_{\mathsf{F}} \le \|A\|_{\mathsf{F}} \, \|B\|_2 \, .$$

Proof. $C[\bullet, j]$ denotes the j-th column of $C \in \mathbb{K}^{I\times K}$. $\|C\|_{\mathsf{F}}^2 = \sum_{j\in J} \|C[\bullet, j]\|_2^2$ involves the Euclidean norm of the columns. For $C := AB$ the columns satisfy $C[\bullet, j] = A \cdot B[\bullet, j]$ and the estimate $\|C[\bullet, j]\|_2 \le \|A\|_2 \, \|B[\bullet, j]\|_2$. Together with the foregoing identity, $\|A\|_{\mathsf{F}}^2 \le \|A\|_2^2 \, \|B\|_{\mathsf{F}}^2$ follows. The second inequality can be concluded from the first one because of $\|X\|_{\mathsf{F}} = \|X^{\mathsf{T}}\|_{\mathsf{F}}$ and $\|X\|_2 = \|X^{\mathsf{T}}\|_2$. □

A particular consequence is $\|A\|_2 \le \|A\|_{\mathsf{F}}$ (use $B = I$ in the second inequality).

Exercise 2.11. Let $U \in \mathbb{K}^{I'\times I}$ and $V \in \mathbb{K}^{J'\times J}$ be orthogonal matrices and prove:
(a) $\|M\|_{\mathsf{F}} = \|UM\|_{\mathsf{F}} = \|MV^{\mathsf{H}}\|_{\mathsf{F}} = \|UMV^{\mathsf{H}}\|_{\mathsf{F}}$ for $M \in \mathbb{K}^{I\times J}$.
(b) $\langle A, B\rangle_{\mathsf{F}} = \left\langle UAV^{\mathsf{H}}, UBV^{\mathsf{H}} \right\rangle_{\mathsf{F}}$ for $A, B \in \mathbb{K}^{I\times J}$.

Exercise 2.12. For index sets I, J, and K let $A \in \mathbb{K}^{I\times K}$ and $B \in \mathbb{K}^{J\times K}$ with $\operatorname{rank}(B) = \#J \le \#K$. Show that the matrix $C \in \mathbb{K}^{I\times J}$ minimising $\|A - CB\|_{\mathsf{F}}$ is given by $C := AB^{\mathsf{H}}(BB^{\mathsf{H}})^{-1}$.

2.4 Semidefinite Matrices

A matrix $M \in \mathbb{K}^{I\times I}$ is called *positive semidefinite*, if

$$M = M^{\mathsf{H}} \quad \text{and} \quad \langle Mx, x\rangle \ge 0 \qquad \text{for all } x \in \mathbb{K}^I.$$

In addition, a *positive definite* matrix has to satisfy $\langle Mx, x\rangle > 0$ for $0 \ne x \in \mathbb{K}^I$.

Remark 2.13. Let $M \in \mathbb{K}^{I\times I}$ be positive [semi]definite. (a) The equation $X^2 = M$ has a unique positive [semi]definite solution in $\mathbb{K}^{I\times I}$, which is denoted by $M^{1/2}$.
(b) M has positive [non-negative] diagonal entries M_{ii} $(i \in I)$.

In the set of Hermitean matrices from $\mathbb{K}^{I\times I}$ a semi-ordering can be defined via

$$A \le B \quad :\Longleftrightarrow \quad B - A \text{ positive semidefinite.} \tag{2.14}$$

When we write $A \le B$, we always tacitly assume that A and B are Hermitean.

Remark 2.14. Let $A, B \in \mathbb{K}^{I\times I}$ be Hermitean. (a) $A \le B$ is equivalent to

$$\langle Ax, x\rangle \le \langle Bx, x\rangle \qquad \text{for all } x \in \mathbb{K}^I. \tag{2.15}$$

(b) For any matrix $T \in \mathbb{K}^{I\times J}$ the inequality $A \le B$ implies $T^{\mathsf{H}}AT \le T^{\mathsf{H}}BT$.
(c) $A \le B$ implies $\operatorname{trace}(A) \le \operatorname{trace}(B)$.
(d) $A \le B$ implies $\operatorname{trace}(T^{\mathsf{H}}AT) \le \operatorname{trace}(T^{\mathsf{H}}BT)$ for all T.

Lemma 2.15. *For $0 \le \hat{E} \le E \in \mathbb{K}^{I\times I}$ and arbitrary $C_i \in \mathbb{K}^{J\times K}$ $(i \in I)$, we have*

$$0 \le \hat{X} := \sum\nolimits_{i,j\in I} \hat{E}_{ij} C_i C_j^{\mathsf{H}} \le X := \sum\nolimits_{i,j\in I} E_{ij} C_i C_j^{\mathsf{H}} \in \mathbb{K}^{J\times J}.$$

Proof. Diagonalisation $E - \hat{E} = U \operatorname{diag}\{\lambda_k : k \in I\} U^{\mathsf{H}}$ holds with $\lambda_k \geq 0$. Set $B_k := \sum_{i \in I} U_{ik} C_i$. Then $X - \hat{X} = \sum_{k \in I} \lambda_k B_k B_k^{\mathsf{H}}$ proves $\hat{X} \leq X$ because of $\lambda_k \geq 0$ and $B_k B_k^{\mathsf{H}} \geq 0$. □

A tuple $\mathfrak{x} := (x_i : i \in I)$ of vectors[6] $x_i \in \mathbb{K}^J$ leads to the scalar products $\langle x_j, x_i \rangle$ for all $i, j \in I$. Then the *Gram matrix* of $\mathfrak{x}$ is defined by

$$G := G(\mathfrak{x}) = \big(\langle x_j, x_i \rangle \big)_{i,j \in I} . \tag{2.16}$$

Exercise 2.16. (a) Gram matrices are always positive semidefinite.
(b) The Gram matrix $G(\mathfrak{x})$ is positive definite if and only if $\mathfrak{x}$ is a tuple of linearly independent vectors.
(c) Any positive definite matrix $G \in \mathbb{K}^{I \times I}$ can be interpreted as a Gram matrix of a basis $\mathfrak{x} := (x_i : i \in I)$ of $\mathbb{K}^I$ by defining a scalar product via $\langle v, w \rangle := b^{\mathsf{H}} G a$ for $v = \sum_{i \in I} a_i x_i$ and $w = \sum_{i \in I} b_i x_i$.

Lemma 2.17. *The spectral norm of $G(\mathfrak{x})$ can be characterised by*

$$\|G(\mathfrak{x})\|_2 = \max \Big\{ \big\| \sum\nolimits_{i \in I} \xi_i x_i \big\|_2^2 : \ \xi_i \in \mathbb{K} \text{ with } \sum\nolimits_{i \in I} |\xi_i|^2 = 1 \Big\} .$$

Proof. Let $\boldsymbol{\xi} := (\xi_i)_{i \in I} \in \mathbb{K}^I$. $\|G(\mathfrak{x})\|_2 = \max \{ |\langle G\boldsymbol{\xi}, \boldsymbol{\xi} \rangle| : \|\boldsymbol{\xi}\| = 1 \}$ holds, since $G(\mathfrak{x})$ is symmetric. $\langle G\boldsymbol{\xi}, \boldsymbol{\xi} \rangle = \sum_{i,j} \langle x_j, x_i \rangle \xi_j \overline{\xi_i} = \big\langle \sum_{j \in I} \xi_j x_j, \sum_{i \in I} \xi_i x_i \big\rangle = \big\| \sum_{i \in I} \xi_i x_i \big\|_2^2$ proves the assertion. □

2.5 Matrix Decompositions

Three well-known decompositions will be recalled. The numbers of arithmetical operations[7] given below are reduced to the leading term, i.e., terms of lower order are omitted.

2.5.1 Cholesky Decomposition

Remark 2.18. Given a positive definite matrix $M \in \mathbb{K}^{n \times n}$, there is a unique lower triangular matrix $L \in \mathbb{K}^{n \times n}$ with positive diagonal entries such that

$$M = LL^{\mathsf{H}}.$$

The computation of L costs $\frac{1}{3} n^3$ operations. Matrix-vector multiplications La or $L^{\mathsf{H}} a$ or the solution of linear systems $Lx = a$ or $L^{\mathsf{H}} x = b$ require n^2 operations.

For semidefinite matrices there are pivotised versions such that M is equal to $PLL^{\mathsf{H}} P^{\mathsf{H}}$ with a permutation matrix P and the condition $L_{ii} \geq 0$ instead of $L_{ii} > 0$.

[6] Here, $\mathbb{K}^J$ can also be replaced by an infinite dimensional Hilbert space.

[7] Here, we count all arithmetical operations ($+, -, *, /, \sqrt{\ }$, etc.) equally. Sometimes, the combination of one multiplication and one addition is counted as one unit ('flop', cf. [20, p. 43]).

2.5.2 QR Decomposition

The letter ‘R’ in ‘QR decomposition’ stands for a right (or upper) triangular matrix. Since an upper triangular matrix R is defined by $R_{ij} = 0$ for all $i > j$, this requires suitably ordered index sets. The QR decomposition (or ‘QR factorisation’) is a helpful tool for orthogonalisation (cf. [161, §3.4.3], [69, §5.2]) and can be viewed as algebraic formulation of the Gram-Schmidt[8] orthogonalisation. Concerning details about different variants and their numerical stability we recommend the book of Björck [20].

Lemma 2.19 (QR factorisation). *Let* $M \in \mathbb{K}^{n\times m}$*. (a) Then there are a unitary matrix* $Q \in \mathbb{K}^{n\times n}$ *and an upper triangular matrix* $R \in \mathbb{K}^{n\times m}$ *with*

$$M = QR \qquad (Q \textit{ unitary, } R \textit{ upper triangular matrix}). \tag{2.17a}$$

Q can be constructed as product of Householder transforms (cf. [178, §4.7]). The computational work is $2mn\min(n,m) - \frac{2}{3}\min(n,m)^3$ for the computation of R (while Q is defined implicitly as a product of Householder matrices), and $\frac{4}{3}n^3$ for forming Q explicitly as a matrix (cf. [69, §5.2.1]).
(b) If $n > m$, the matrix R has the block structure $\left[\begin{smallmatrix} R' \\ 0 \end{smallmatrix}\right]$, where the submatrix R' is an upper triangular matrix of size $m \times m$. The corresponding block decomposition $Q = [Q' \; Q'']$ yields the reduced QR factorisation

$$M = Q'R' \qquad (Q' \in \mathbb{K}^{n\times m}, R' \in \mathbb{K}^{m\times m}). \tag{2.17b}$$

The computational work is[9] $N_{\mathrm{QR}}(n,m) := 2nm^2$ *(cf. [69, Alg. 5.2.5], [161, §3.4]).*
(c) If $r := \operatorname{rank}(M) < \min\{n,m\}$, the sizes of Q' and R' can be further reduced:

$$M = Q'R' \qquad (Q' \in \mathbb{K}^{n\times r}, R' \in \mathbb{K}^{r\times m}). \tag{2.17c}$$

In particular, if M does not possess full rank as in Part (c) of the lemma above, one wants R' from (2.17c) to be of the form

$$R' = [R_1' \; R_2'], \quad R_1' \in \mathbb{K}^{r\times r} \text{ upper triangular}, \quad \operatorname{rank}(R_1') = r, \tag{2.17d}$$

i.e., the diagonal entries of R_1' do not vanish. This form of R' can be achieved if and only if the part $(M_{ij})_{1\le i,j\le r}$ of M has also rank r. Otherwise, one needs a suitable permutation $M \mapsto MP$ of the columns of M. Then the factorisation takes the form

$$MP = Q'\,[R_1' \; R_2'] \quad (P \text{ permutation matrix, } Q', R_1', R_2' \text{ from (2.17c,d)}). \tag{2.18}$$

An obvious pivot strategy for a matrix $M \in \mathbb{K}^{n\times m}$ with $r = \operatorname{rank}(M)$ is the Gram-Schmidt orthogonalisation in the following form (cf. [69, §5.4.1]).

[8] A modified Gram-Schmidt algorithm was already derived by Laplace in 1816 (see reference in [20, p. 61] together with further remarks concerning history).

[9] Half of the cost of $N_{\mathrm{QR}}(n,m)$ is needed for $\frac{1}{2}(m^2+m)$ scalar products. The rest is used for scaling and adding column vectors.

1) Let $m_i \in \mathbb{K}^n$ $(1 \leq i \leq m)$ be the i-th columns of M.

2) for $i := 1$ to r do

2a) Choose $k \in \{i, \ldots, m\}$ such that $\|m_k\| = \max\{\|m_\nu\| : i \leq \nu \leq m\}$. If $k \neq i$, interchange the columns m_i and m_k.

2b) Now m_i has maximal norm. Normalise: $m_i := m_i / \|m_i\|$. Store m_i as i-th column of the matrix Q.

2c) Perform $m_k := m_k - \langle m_k, m_i \rangle m_i$ for $i + 1 \leq k \leq m$.

Here, $\|\cdot\|$ is the Euclidean norm and $\langle \cdot, \cdot \rangle$ the corresponding scalar product. The column exchanges in Step 2a lead to the permutation matrix[10] P in (2.18). The operations in Step 2b and Step 2c define $[R_1' \; R_2']$.

The presupposition $r = \operatorname{rank}(M)$ guarantees that all m_i appearing in Step 2b do not vanish, while $m_i = 0$ $(r + 1 \leq i \leq m)$ holds after the r-th iteration for the remaining columns. In usual applications, the rank is unknown. In that case, one may introduce a tolerance $\tau > 0$ and redefine Step 2b as follows:

2b') If $\|m_i\| \leq \tau$ set $r := i - 1$ and terminate. Otherwise, proceed as in Step 2b.

The principle of the QR decomposition can be generalised to tuples V^m, where the column vectors from $\mathbb{K}^n$ are replaced by functions from the space V (cf. Trefethen [183]).

2.5.3 Singular Value Decomposition

2.5.3.1 Definition and Computational Cost

The singular value decomposition (abbreviation: SVD) is the generalisation of the diagonalisation of square matrices (cf. [161, §1.9]).

Lemma 2.20 (SVD). *(a) Let $M \in \mathbb{K}^{n \times m}$ be any matrix. Then there are unitary matrices $U \in \mathbb{K}^{n \times n}$, $V \in \mathbb{K}^{m \times m}$, and a diagonal rectangular matrix $\Sigma \in \mathbb{R}^{n \times m}$,*

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots\ 0 & 0 \dots 0 \\ 0 & \sigma_2 & \ddots\ 0 & 0 \qquad 0 \\ \vdots & \ddots & \ddots\ \vdots & \vdots \qquad \vdots \\ 0 & \dots\ 0 & \sigma_n & 0 \dots 0 \end{bmatrix} \qquad \left(\begin{array}{l} \textit{illustration} \\ \textit{for the case} \\ \textit{of } n \leq m \end{array} \right), \tag{2.19a}$$

with so-called singular values[11]

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_i = \Sigma_{ii} \geq \ldots \geq 0 \qquad (1 \leq i \leq \min\{n, m\})$$

[10] A permutation matrix $P \in \mathbb{K}^{r \times r}$ (corresponding to a permutation $\pi : \{1, \ldots, r\} \to \{1, \ldots, r\}$) is defined by $(Pv)_i = v_{\pi(i)}$. Any permutation matrix P is unitary.

[11] For indices $\ell > \min\{\#I, \#J\}$ we formally define $\sigma_\ell := 0$.

such that[12]

$$M = U\Sigma V^{\mathsf{T}}. \tag{2.19b}$$

The columns of U are the left singular vectors*, the columns of V are the* right singular vectors.
(b) The spectral norm of M has the value $\|M\|_2 = \sigma_1$.
(c) The Frobenius norm of M equals

$$\|M\|_{\mathsf{F}} = \sqrt{\sum\nolimits_{i=1}^{\min\{n,m\}} \sigma_i^2}\,. \tag{2.19c}$$

Proof. i) Assume without loss of generality that $n \le m$ and set $A := MM^{\mathsf{H}} \in \mathbb{K}^{n\times n}$. Diagonalise the positive semidefinite matrix: $A = UDU^{\mathsf{H}}$ with $U \in \mathbb{K}^{n\times n}$ unitary, $D = \operatorname{diag}\{d_1, \ldots, d_n\} \in \mathbb{R}^{n\times n}$, where the (non-negative) eigenvalues are ordered by size: $d_1 \ge d_2 \ge \ldots \ge 0$. Defining $\sigma_i := \sqrt{d_i}$ in (2.19a), we rewrite

$$D = \Sigma\Sigma^{\mathsf{T}} = \Sigma\Sigma^{\mathsf{H}}.$$

With $W := M^{\mathsf{H}}U = [w_1, \ldots, w_n] \in \mathbb{K}^{m\times n}$ we have

$$D = U^{\mathsf{H}}AU = U^{\mathsf{H}}MM^{\mathsf{H}}U = W^{\mathsf{H}}W.$$

Hence, the columns w_i of W are pairwise orthogonal and $w_i^{\mathsf{H}}w_i = d_i = \sigma_i^2$.

Next, we are looking for a unitary matrix $V = [v_1, \ldots, v_m] \in \mathbb{K}^{m\times m}$ with

$$W = \overline{V}\Sigma^{\mathsf{T}}, \qquad \text{i.e., } w_i = \sigma_i \overline{v_i} \quad (1 \le i \le m)$$

(note that the complex conjugate values $\overline{v_i}$ are used[12]).

Let $r := \max\{i : \sigma_i > 0\}$. For $1 \le i \le r$, the condition above leads to $\overline{v_i} := \frac{1}{\sigma_i} w_i$, i.e., v_i is normalised: $v_i^{\mathsf{H}}v_i = 1$. Since the vectors w_i of W are already pairwise orthogonal, the vectors $\{\overline{v_i} : 1 \le i \le r\}$ are orthonormal.

For $r + 1 \le i \le n$, $\sigma_i = 0$ implies $w_i = 0$. Hence $w_i = \sigma_i\overline{v_i}$ holds for any choice of v_i. To obtain a unitary matrix $\overline{V}$, we may choose any orthonormal extension $\{\overline{v_i} : r + 1 \le i \le m\}$ of $\{\overline{v_i} : 1 \le i \le r\}$. The relation $W = \overline{V}\Sigma^{\mathsf{T}}$ (with $\Sigma^{\mathsf{T}} = \Sigma^{\mathsf{H}}$) implies $W^{\mathsf{H}} = \Sigma V^{\mathsf{T}}$. By Definition of W we have $M = UW^{\mathsf{H}} = U\Sigma V^{\mathsf{T}}$, so that (2.19b) is proved.

ii) Exercises 2.9b and 2.11a imply that $\|M\|_2 = \|\Sigma\|_2$ and $\|M\|_{\mathsf{F}} = \|\Sigma\|_{\mathsf{F}}$ proving the parts (b) and (c). □

If $n < m$, the last $m - n$ columns of V are multiplied by the zero part of Σ. Similarly, for $n > m$, certain columns of U are not involved in the representation of M. Reduction to the first $\min\{n, m\}$ columns yields the following result.

Corollary 2.21. (a) Let $u_i \in \mathbb{K}^I$ and $v_i \in \mathbb{K}^J$ be the (orthonormal) columns of U and V, respectively. Then the statement $M = U\Sigma V^{\mathsf{T}}$ from (2.19b) is equivalent to

[12] The usual formulation uses $M = U\Sigma V^{\mathsf{H}}$ (or $U^{\mathsf{H}}\Sigma V$) with the Hermitean transposed V^{H}. Here we use V^{T} also for $\mathbb{K} = \mathbb{C}$ because of Remark 1.3a.

$$M = \sum_{i=1}^{\min\{n,m\}} \sigma_i \, u_i \, v_i^{\mathsf{T}}. \tag{2.20}$$

The computational cost is about

$$N_{\mathrm{SVD}}(n,m) := \min\left\{14nmN + 8N^3, 6nmN + 20N^3\right\},$$

where $N := \min\{n,m\}$ (cf. [69, §5.4.5]).

(b) The decomposition (2.20) is not unique. Let $\sigma_i = \sigma_{i+1} = \ldots = \sigma_{i+k-1}$ be a k-fold singular value. The part $\sum_{j=i}^{i+k-1} \sigma_j \, u_j \, v_j^{\mathsf{T}}$ in (2.20) equals

$$\sigma_i \left[u_i, \ldots, u_{i+k-1}\right] \left[v_i, \ldots, v_{i+k-1}\right]^{\mathsf{T}}.$$

For any unitary $k \times k$ matrix Q, the transformed vectors

$$\left[\hat{u}_i, \ldots, \hat{u}_{i+k-1}\right] := \left[u_i, \ldots, u_{i+k-1}\right] Q \text{ and } \left[\hat{v}_i, \ldots, \hat{v}_{i+k-1}\right] := \left[v_i, \ldots, v_{i+k-1}\right] \overline{Q}$$

yield the same sum $\sum_{j=i}^{i+k-1} \sigma_j \, \hat{u}_j \, \hat{v}_j^{\mathsf{T}}$. Even in the case $k = 1$ of a simple singular value, each pair u_i, v_i of columns may be changed into $\hat{u}_i := z u_i$, $\hat{v}_i := \frac{1}{z} v_i$ with $z \in \mathbb{K}$ and $|z| = 1$.

(c) In many applications $\operatorname{span}\{u_i : 1 \leq i \leq r\}$ for some $r \leq \min\{n,m\}$ is of interest. This space is uniquely determined if and only if $\sigma_r < \sigma_{r+1}$. The same statement holds for $\operatorname{span}\{v_i : 1 \leq i \leq r\}$.

Next, we consider a convergent sequence $M^{(\nu)} \to M$ of matrices together with their singular value decompositions $M^{(\nu)} = U^{(\nu)} \Sigma^{(\nu)} V^{(\nu)\mathsf{T}}$ and $M = U \Sigma V^{\mathsf{T}}$.

Remark 2.22. (a) Let $M^{(\nu)} = U^{(\nu)} \Sigma^{(\nu)} V^{(\nu)\mathsf{T}} \in \mathbb{K}^{n \times m}$ be the singular value decompositions of $M^{(\nu)} \to M$. Then there is a subsequence $\{\nu_i : i \in \mathbb{N}\} \subset \mathbb{N}$ such that

$$U^{(\nu_i)} \to U, \quad \Sigma^{(\nu_i)} \to \Sigma, \quad V^{(\nu_i)} \to V, \quad M = U \Sigma V^{\mathsf{T}}.$$

(b) Subsequences of the spaces $S_r^{(\nu)} := \operatorname{span}\{u_i^{(\nu)} : 1 \leq i \leq r\}$ converge to $S_r := \operatorname{span}\{u_i : 1 \leq i \leq r\}$, where $u_i^{(\nu)}$ and u_i are the columns of $U^{(\nu)}$ and U from Part (a).

Proof. Eigenvalues depend continuously on the matrix; hence, $\Sigma^{(\nu)} \to \Sigma$. The subset $\{u_1^{(\nu)} : \nu \in \mathbb{N}\} \subset \mathbb{K}^I$ is bounded by 1, thus it is pre-compact and a subsequence yields $u_1 := \lim_i u_1^{(\nu_i)}$ with $\|u_1\| = \lim \|u_1^{(\nu_i)}\| = 1$. Restrict the sequence to the latter subsequence and proceed with $u_2^{(\nu)}$ in the same way. The convergence of the subsequences to U, Σ, V implies $M = U \Sigma V^{\mathsf{T}}$. □

Since the factors U and V are possibly not unique (cf. Corollary 2.21b), it may happen that $M = U \Sigma V^{\mathsf{T}} = \hat{U} \Sigma \hat{V}^{\mathsf{T}}$ are two different decompositions and the limit of $M^{(\nu)} = U^{(\nu)} \Sigma^{(\nu)} V^{(\nu)\mathsf{T}}$ yields $U \Sigma V^{\mathsf{T}}$, while no subsequence converges to $\hat{U} \Sigma \hat{V}^{\mathsf{T}}$. Hence, the reverse statement that, in general, any singular value decomposition $M = U \Sigma V^{\mathsf{T}}$ is a limit of $M^{(\nu)} = U^{(\nu)} \Sigma^{(\nu)} V^{(\nu)\mathsf{T}}$, is wrong.

2.5.3.2 Reduced and One-Sided Singular Value Decompositions

If M is not of full rank, there are singular values $\sigma_i = 0$, so that further terms can be omitted from the sum in (2.20). Let[13] $r := \max\{i : \sigma_i > 0\} = \operatorname{rank}(M)$ as in the proof above. Then (2.20) can be rewritten as

$$M = \sum_{i=1}^{r} \sigma_i \, u_i \, v_i^{\mathsf{T}} \qquad \text{with } \begin{cases} \{u_i\}_{i=1}^{r}, \{v_i\}_{i=1}^{r} \text{ orthonormal,} \\ \sigma_1 \geq \ldots \geq \sigma_r > 0, \end{cases} \tag{2.21}$$

where only nonzero terms appear. The corresponding matrix formulation is

$$M = U' \Sigma' V'^{\mathsf{T}} \text{ with } \begin{cases} U' = [u_1, \ldots, u_r] \in \mathbb{K}^{n \times r} \text{ orthogonal,} \\ V' = [v_1, \ldots, v_r] \in \mathbb{K}^{m \times r} \text{ orthogonal,} \\ \Sigma' = \operatorname{diag}\{\sigma_1, \ldots, \sigma_r\} \in \mathbb{R}^{r \times r}, \sigma_1 \geq \ldots \geq \sigma_r > 0. \end{cases} \tag{2.22}$$

Definition 2.23 (reduced SVD). The identities (2.21) or (2.22) are called the *reduced singular value decomposition* (since the matrices U, Σ, V from (2.19b) are reduced to the essential nonzero part).

There are cases—in particular, when $m \gg n$—where one is interested only in the left singular vectors u_i and the singular values σ_i from (2.21) or equivalently only in U' and Σ' from (2.22). Then we say that we need the *left-sided singular value decomposition*. The proof of Lemma 2.20 has already shown how to solve for U' and Σ' alone:

1) Perform $A := MM^{\mathsf{H}} \in \mathbb{K}^{n \times n}$.

2) Diagonalise $A = UDU^{\mathsf{H}}$ with the non-negative diagonal matrix

$$D = \operatorname{diag}\{d_1, \ldots, d_n\} \in \mathbb{R}^{n \times n}, \quad d_1 \geq d_2 \geq \ldots \geq 0.$$

3) Set $r := \max\{i : d_i > 0\}$, $\sigma_i := \sqrt{d_i}$, and $\Sigma' := \operatorname{diag}\{\sigma_1, \ldots, \sigma_r\}$.

4) Restrict U to the first r columns: $U' = [u_1, \ldots, u_r]$.

Remark 2.24. (a) Steps 1-4 from above define the matrices U' and Σ' from (2.22). The third matrix V' is theoretically available via $V' = (\Sigma')^{-1} M^{\mathsf{H}} \overline{U'}$. The product MM^{H} in Step 1 requires the computation of $\frac{n(n+1)}{2}$ scalar products $\langle m_i, m_j \rangle$ $(i, j \in I)$ involving the rows $m_i := M[i, \bullet] \in \mathbb{K}^J$ of M. The computational cost for these scalar products will crucially depend on the underlying data structure (cf. Remark 7.12). Steps 2-4 are independent of the size of J. Their cost is asymptotically $\frac{8}{3}n^3$ (cf. [69, §8.3.1]).

(b) The knowledge of U' suffices to define $\hat{M} := U'^{\mathsf{H}} M$. $\hat{M}$ has orthogonal rows $\hat{m}_i$ $(1 \leq i \leq n)$ which are ordered by size: $\|\hat{m}_1\| = \sigma_1 > \|\hat{m}_2\| = \sigma_2 > \ldots > 0$.

Proof. Let $M = U' \Sigma' V'^{\mathsf{T}}$ be the reduced singular value decomposition. Since $U'^{\mathsf{H}} U' = I \in \mathbb{K}^{r \times r}$, Part (b) defines $\hat{M} := U'^{\mathsf{H}} M = \Sigma' V'^{\mathsf{T}}$. It follows that $\hat{M} \hat{M}^{\mathsf{H}} = \left(\Sigma' V'^{\mathsf{T}}\right) \left(\overline{V'} \Sigma'\right) = \Sigma'^2$, i.e., $\langle \hat{m}_i, \hat{m}_j \rangle = 0$ for $i \neq j$ and $\|\hat{m}_i\| = \sigma_i$. □

[13] If $\sigma_i = 0$ for all i, set $r := 0$ (empty sum). This happens for the uninteresting case of $M = 0$.

The analogously defined *right-sided singular value decomposition* of M is identical to the left-sided singular value decomposition of the transposed matrix M^{T}, since $M = U'\varSigma'V'^{\mathsf{T}} \Longleftrightarrow M^{\mathsf{T}} = V'\varSigma'U'^{\mathsf{T}}$.

2.5.3.3 Inequalities of Singular Values

Finally, we discuss estimates about eigenvalues and singular values of perturbed matrices. The following lemma states the Fischer-Courant characterisation of eigenvalues. For a general matrix $A \in \mathbb{K}^{n\times n}$ we denote the eigenvalues corresponding to their multiplicity by $\lambda_k(A)$. If $\lambda_k(A) \in \mathbb{R}$, we order the eigenvalues such that $\lambda_k(A) \geq \lambda_{k+1}(A)$. Formally, we set $\lambda_k(A) := 0$ for $k > n$.

Remark 2.25. For matrices $A \in \mathbb{K}^{n\times m}$ and $B \in \mathbb{K}^{m\times n}$ the identity $\lambda_k(AB) = \lambda_k(BA)$ is valid. If A and B are positive semidefinite, the eigenvalues $\lambda_k(AB)$ are non-negative.

Proof. 1) If $e \neq 0$ is an eigenvector of AB with nonzero eigenvalue λ, the vector Be does not vanish. Then $(BA)(Be) = B(AB)e = B(\lambda e) = \lambda(Be)$ proves that Be is an eigenvector of BA for the same λ. Hence, AB and BA share the same nonzero eigenvalues. The further ones are zero (maybe by the setting from above).

2) If $B \geq 0$, the square root $B^{1/2}$ is defined (cf. Remark 2.13). Part 1) shows $\lambda_k(AB) = \lambda_k(AB^{1/2}B^{1/2}) = \lambda_k(B^{1/2}AB^{1/2})$. The latter matrix is positive semidefinite proving $\lambda_k(B^{1/2}AB^{1/2}) \geq 0$. □

Lemma 2.26. *Let the matrix $A \in \mathbb{K}^{n\times n}$ be positive semidefinite. Then the eigenvalues $\lambda_1(A) \geq \ldots \geq \lambda_n(A) \geq 0$ can be characterised by*

$$\lambda_k(A) = \min_{\substack{\mathcal{V} \subset \mathbb{K}^n \text{ subspace} \\ \text{with } \dim(\mathcal{V}) \leq k-1}} \; \max_{\substack{x \in \mathbb{K}^n \text{ with} \\ x^{\mathsf{H}}x = 1 \text{ and } x \perp \mathcal{V}}} x^{\mathsf{H}}Ax. \tag{2.23}$$

Proof. A can be diagonalised: $A = U\varLambda U^{\mathsf{H}}$ with $\varLambda_{ii} = \lambda_i$. Since $x^{\mathsf{H}}Ax = y^{\mathsf{H}}\varLambda y$ for $y = U^{\mathsf{H}}x$, the assertion can also be stated in the form

$$\lambda_k = \min_{\mathcal{W} \text{ with } \dim(\mathcal{W}) \leq k-1} \max\left\{y^{\mathsf{H}}\varLambda y : y \in \mathbb{K}^n \text{ with } y^{\mathsf{H}}y = 1,\ y \perp \mathcal{W}\right\}$$

$(\mathcal{W} = U^{\mathsf{H}}\mathcal{V})$. Fix $\mathcal{W}$ with $\dim(\mathcal{W}) \leq k-1$. All $y \in \mathbb{K}^n$ with $y_i = 0$ for $i > k$ form a k-dimensional subspace $\mathcal{Y}$. Since $\dim(\mathcal{W}) \leq k-1$, there is at least one $0 \neq y \in \mathcal{Y}$ with $y^{\mathsf{H}}y = 1,\ y \perp \mathcal{W}$. Obviously, $y^{\mathsf{H}}\varLambda y = \sum_{i=1}^{k} \lambda_i y_i^2 \geq \sum_{i=1}^{k} \lambda_k y_i^2 \geq \lambda_k$. The choice $\mathcal{W} = \{w \in \mathbb{K}^n : w_i = 0 : k \leq i \leq n\}$ yields equality: $y^{\mathsf{H}}\varLambda y = \lambda_k$. □

In the following we use the notation $\lambda_k(A)$ for the k-th eigenvalue of a positive semidefinite matrix A, where the ordering of the eigenvalues is by size (see Lemma 2.26). Similarly, $\sigma_k(A)$ denotes the k-th singular value of a general matrix A. Note that $\|\cdot\|_2$ is the spectral norm from (2.13).

Lemma 2.27. *(a) Let $A, B \in \mathbb{K}^{n\times n}$ be two positive semidefinite matrices. Then*

$$\lambda_k(A) \le \lambda_k(A+B) \le \lambda_k(A) + \|B\|_2 \qquad \text{for } 1 \le k \le n. \tag{2.24a}$$

In particular, $0 \le A \le B$ implies $\lambda_k(A) \le \lambda_k(B)$ for $1 \le k \le n$.

(b) Let the matrices $A \in \mathbb{K}^{n\times m}$ and $B \in \mathbb{K}^{n\times m'}$ satisfy $AA^{\mathsf{H}} \le BB^{\mathsf{H}}$. Then the singular values[14] *$\sigma_k(A)$ and $\sigma_k(B)$ of both matrices are related by*

$$\sigma_k(A) \le \sigma_k(B) \qquad \text{for } 1 \le k \le n. \tag{2.24b}$$

The same statement holds for $A \in \mathbb{K}^{m\times n}$ and $B \in \mathbb{K}^{m'\times n}$ with $A^{\mathsf{H}}A \le B^{\mathsf{H}}B$.

(c) Let $M \in \mathbb{K}^{n\times m}$ be any matrix, while $A \in \mathbb{K}^{n'\times n}$ and $B \in \mathbb{K}^{m\times m'}$ have to satisfy $A^{\mathsf{H}}A \le I$ and $B^{\mathsf{H}}B \le I$. Then[14]

$$\sigma_k(AMB) \le \sigma_k(M) \qquad \text{for } k \in \mathbb{N}.$$

Proof. 1) $\lambda_k(A) \le \lambda_k(A+B)$ is a consequence of Remark 2.14 and Lemma 2.26.

2) Let $\mathcal{V}_A$ and $\mathcal{V}_{A+B}$ be the subspaces from (2.23), which yield the minimum for A and $A+B$, respectively. Abbreviate the maximum in (2.23) over $x \in \mathbb{K}^n$ with $x^{\mathsf{H}}x = 1$ and $x \perp \mathcal{V}$ by $\max_{\mathcal{V}}$. Then

$$\begin{aligned}\lambda_k(A+B) &= \max_{\mathcal{V}_{A+B}} x^{\mathsf{H}}(A+B)x \le \max_{\mathcal{V}_A} x^{\mathsf{H}}(A+B)x = \max_{\mathcal{V}_A}\left[x^{\mathsf{H}}Ax + x^{\mathsf{H}}Bx\right] \\ &\le \max_{\mathcal{V}_A} x^{\mathsf{H}}Ax + \max_{x^{\mathsf{H}}x=1} x^{\mathsf{H}}Bx = \lambda_k(A) + \|B\|_2 .\end{aligned}$$

3) For Part (b) use $\lambda_k(A) \le \lambda_k(A+B)$ with A and B replaced by AA^{H} and $BB^{\mathsf{H}} - AA^{\mathsf{H}}$ in the case of $AA^{\mathsf{H}} \le BB^{\mathsf{H}}$. Otherwise, use that the eigenvalues of $X^{\mathsf{H}}X$ and XX^{H} coincide (cf. Remark 2.25).

4) Let $M' := AMB$ and use $\sigma_k(M')^2 = \lambda_k(M'^{\mathsf{H}}M')$. Remark 2.14b implies that $M'(M')^{\mathsf{H}} = AMBB^{\mathsf{H}}M^{\mathsf{H}}A^{\mathsf{H}} \le AMM^{\mathsf{H}}A^{\mathsf{H}}$, so that $\lambda_k(M'^{\mathsf{H}}M') \le \lambda_k(AMM^{\mathsf{H}}A^{\mathsf{H}})$. Remark 2.25 states that $\lambda_k(AMM^{\mathsf{H}}A^{\mathsf{H}}) = \lambda_k(M^{\mathsf{H}}A^{\mathsf{H}}AM)$, and from $A^{\mathsf{H}}A \le I$ we infer that $\lambda_k(M^{\mathsf{H}}A^{\mathsf{H}}AM) \le \lambda_k(M^{\mathsf{H}}M) = \sigma_k(M)^2$. □

Let $n = n_1 + n_2$, $A \in \mathbb{K}^{n_1\times m}$ and $B \in \mathbb{K}^{n_2\times m}$. Then the agglomerated matrix $\left[\begin{smallmatrix}A\\B\end{smallmatrix}\right]$ belongs to $\mathbb{K}^{n\times m}$. In the next lemma we compare singular values of A and $\left[\begin{smallmatrix}A\\B\end{smallmatrix}\right]$.

Lemma 2.28. *For general $A \in \mathbb{K}^{n_1\times m}$ and $B \in \mathbb{K}^{n_2\times m}$, the singular values satisfy*

$$\sigma_k(A) \le \sigma_k(\left[\begin{smallmatrix}A\\B\end{smallmatrix}\right]) \le \sqrt{\sigma_k^2(A) + \|B\|_2^2}.$$

The same estimate holds for $\sigma_k([A\ B])$, where $A \in \mathbb{K}^{n\times m_1}$ and $B \in \mathbb{K}^{n\times m_2}$.

Proof. Use $\sigma_k^2(A) = \lambda_k(A^{\mathsf{H}}A)$ and $\sigma_k^2(\left[\begin{smallmatrix}A\\B\end{smallmatrix}\right]) = \lambda_k([A\ B]^{\mathsf{H}}\left[\begin{smallmatrix}A\\B\end{smallmatrix}\right]) = \lambda_k(A^{\mathsf{H}}A + B^{\mathsf{H}}B)$ and apply (2.24a): $\sigma_k^2(\left[\begin{smallmatrix}A\\B\end{smallmatrix}\right]) \le \lambda_k(A^{\mathsf{H}}A) + \|B^{\mathsf{H}}B\|_2 = \sigma_k^2(A) + \|B\|_2^2$. □

[14] See Footnote 11 on page 29.

Exercise 2.29. Prove $\sigma_k(A) \leq \sigma_k\left(\left[\begin{array}{c|c} A & C \\ \hline B & \end{array}\right]\right) \leq \sqrt{\sigma_k^2(A) + \|B\|_2^2 + \|C\|_2^2}$.

2.6 Low-Rank Approximation

Given a matrix M, we ask for a matrix $R \in \mathcal{R}_s$ of lower rank (i.e., $s < \operatorname{rank}(M)$) such that $\|M - R\|$ is minimised. The answer is given by[15] Erhard Schmidt (1907) [168, §18]. In his paper, he studies the infinite singular value decomposition for operators (cf. Theorem 4.114). The following finite case is a particular application.

Lemma 2.30. *(a) Let $M, R \in \mathbb{K}^{n\times m}$ with $r := \operatorname{rank}(R)$. The singular values of M and $M - R$ are denoted by $\sigma_i(M)$ and $\sigma_i(M - R)$, respectively. Then*[16]

$$\sigma_i(M - R) \geq \sigma_{r+i}(M) \qquad \text{for all } 1 \leq i \leq \min\{n, m\}. \tag{2.25}$$

(b) Let $s \in \{0, 1, \ldots, \min\{n, m\}\}$. Use the singular value decomposition $M = U\Sigma V^\mathsf{T}$ to define

$$R := U\Sigma_s V^\mathsf{T} \quad \text{with } (\Sigma_s)_{ij} = \begin{cases} \sigma_i & \text{for } i = j \leq s, \\ 0 & \text{otherwise,} \end{cases} \tag{2.26a}$$

i.e., Σ_s results from Σ by replacing all singular values $\sigma_i = \Sigma_{ii}$ for $i > s$ by zero. Then the approximation error is

$$\|M - R\|_2 = \sigma_{s+1} \qquad \text{and} \qquad \|M - R\|_\mathsf{F} = \sqrt{\sum\nolimits_{i=s+1}^{\min\{n,m\}} \sigma_i^2}. \tag{2.26b}$$

Inequalities (2.25) becomes $\sigma_i(M - R) = \sigma_{s+i}(M)$.

Proof. 1) If $r + i > \min\{n, m\}$, (2.25) holds because of $\sigma_{r+i}(M) = 0$. Therefore suppose $r + i \leq \min\{n, m\}$.

2) First, $\sigma_i(M - R)$ is investigated for $i = 1$. $\lambda_{r+1}(MM^\mathsf{H}) := \sigma_{r+1}^2(M)$ is the $(r + 1)$-th eigenvalue of $A := MM^\mathsf{H}$ (see proof of Lemma 2.20). The minimisation in (2.23) yields

$$\sigma_{r+1}^2(M) \leq \max\left\{x^\mathsf{H} Ax : x \in \mathbb{K}^n \text{ with } x^\mathsf{H} x = 1,\ x \perp \mathcal{V}\right\}$$

for any fixed subspace $\mathcal{V}$ of dimension $\leq r$. Choose $\mathcal{V} := \ker(R^\mathsf{H})^\perp$. As $x \perp \mathcal{V}$ is equivalent to $x \in \ker(R^\mathsf{H})$, we conclude that

$$\begin{aligned} x^\mathsf{H} Ax &= x^\mathsf{H} MM^\mathsf{H} x = \left(M^\mathsf{H} x\right)^\mathsf{H} \left(M^\mathsf{H} x\right) = \left((M - R)^\mathsf{H} x\right)^\mathsf{H} \left((M - R)^\mathsf{H} x\right) \\ &= x^\mathsf{H} (M - R) (M - R)^\mathsf{H} x. \end{aligned}$$

[15] Occasionally, this result is attributed to Eckart-Young [50], who reinvented the statement later in 1936.

[16] See Footnote 11 on page 29.

Application of (2.23) to the first eigenvalue $\lambda_1 = \lambda_1((M-R)(M-R)^{\mathsf{H}})$ of the matrix $(M-R)(M-R)^{\mathsf{H}}$ shows

$$\begin{aligned}
&\max\left\{x^{\mathsf{H}}Ax : x \in \mathbb{K}^n \text{ with } x^{\mathsf{H}}x = 1,\ x \perp \mathcal{V}\right\} \\
&\quad = \max\{x^{\mathsf{H}}(M-R)(M-R)^{\mathsf{H}}x : x^{\mathsf{H}}x = 1,\ x \perp \mathcal{V}\} \\
&\quad \le \max\{x^{\mathsf{H}}(M-R)(M-R)^{\mathsf{H}}x : x \in \mathbb{K}^I \text{ with } x^{\mathsf{H}}x = 1\} \\
&\quad = \lambda_1\big((M-R)(M-R)^{\mathsf{H}}\big)
\end{aligned}$$

(in the case of the first eigenvalue, the requirement $x \perp \mathcal{V}$ with $\dim(\mathcal{V}) = 0$ is an empty condition). Since, again, $\lambda_1\big((M-R)(M-R)^{\mathsf{H}}\big) = \sigma_1^2(M-R)$, we have proved $\sigma_{r+1}^2(M) \le \sigma_1^2(M-R)$, which is Part (a) for $i = 1$.

3) For $i > 1$ choose $\mathcal{V} := \ker(R^{\mathsf{H}})^{\perp} + \mathcal{W}$, where $\mathcal{W}$ with $\dim(\mathcal{W}) \le i-1$ is arbitrary. Analogously to Part 2), one obtains the bound

$$\max\left\{x^{\mathsf{H}}(M-R)(M-R)^{\mathsf{H}}x :\ x \in \mathbb{K}^n \text{ with } x^{\mathsf{H}}x = 1,\ x \perp \mathcal{W}\right\}.$$

Minimisation over all $\mathcal{W}$ yields $\lambda_i\big((M-R)(M-R)^{\mathsf{H}}\big) = \sigma_i^2(M-R)$.

4) The choice from (2.26a) eliminates the singular values $\sigma_1, \ldots, \sigma_s$ so that $\sigma_i(M-R) = \sigma_{s+i}(M)$ for all $i \ge 1$. □

Using the notation $M = \sum_{i=1}^{r} \sigma_i u_i v_i^{\mathsf{T}}$ from (2.21), we write R as $\sum_{i=1}^{s} \sigma_i u_i v_i^{\mathsf{T}}$. A connection with projections is given next.

Remark 2.31. $P_1^{(s)} := \sum_{i=1}^{s} u_i u_i^{\mathsf{H}}$ and $P_2^{(s)} := \sum_{i=1}^{s} v_i v_i^{\mathsf{H}}$ are the orthogonal projections onto $\mathrm{span}\{u_i : 1 \le i \le s\}$ and $\mathrm{span}\{v_i : 1 \le i \le s\}$, respectively. Then R from (2.26a) can be written as $R = P_1^{(s)} M (P_2^{(s)})^{\mathsf{T}} = P_1^{(s)} M = M (P_2^{(s)})^{\mathsf{T}}$.

Conclusion 2.32 (best rank-k approximation). *For $M \in \mathbb{K}^{n \times m}$ construct R as in (2.26a). Then R is the solution of the following two minimisation problems:*

$$\min_{\mathrm{rank}(R) \le r} \|M-R\|_2 \qquad \textit{and} \qquad \min_{\mathrm{rank}(R) \le r} \|M-R\|_{\mathsf{F}}. \tag{2.27}$$

The values of the minima are given in (2.26b). The minimising element R is unique if and only if $\sigma_r > \sigma_{r+1}$.

Proof. 1) Since $\|M-R'\|_2 = \sigma_1(M-R')$ and $\|M-R'\|_{\mathsf{F}}^2 = \sum_{i>0} \sigma_i^2(M-R')$ follows from Lemma 2.20b,c, we obtain from Lemma 2.30a that

$$\|M-R'\|_2 \ge \sigma_{r+1}(M), \ \|M-R'\|_{\mathsf{F}}^2 \ge \sum_{i>r} \sigma_i^2(M) \quad \text{for } R' \text{ with } \mathrm{rank}(R') \le r.$$

Since equality holds for $R' = R$, this is the solution of the minimisation problems.

2) If $\sigma_k = \sigma_{k+1}$, one may interchange the r-th and $(r+1)$-th columns in U and V obtaining another singular value decomposition. Thus, another R results. □

Next, we consider a convergent sequence $M^{(\nu)}$ and use Remark 2.22.

Lemma 2.33. *Consider $M^{(\nu)} \in \mathbb{K}^{n\times m}$ with $M^{(\nu)} \to M$. Then there are best approximations $R^{(\nu)}$ according to (2.27) so that a subsequence of $R^{(\nu)}$ converges to R, which is the best approximation to M.*

Remark 2.34. The optimisation problems (2.27) can also be interpreted as the best approximation of the range of M:

$$\max\left\{\|PM\|_\mathsf{F} : P \text{ orthogonal projection with } \operatorname{rank}(P) = r\right\}. \tag{2.28a}$$

Proof. The best approximation $R \in \mathcal{R}_r$ to M has the representation $R = PM$ for $P = P_1^{(r)}$ (cf. Remark 2.31). By orthogonality,

$$\|PM\|_\mathsf{F}^2 + \|(I-P)M\|_\mathsf{F}^2 = \|M\|_\mathsf{F}^2$$

holds. Hence, minimisation of $\|(I-P)M\|_\mathsf{F}^2 = \|M-R\|_\mathsf{F}^2$ is equivalent to maximisation of $\|PM\|_\mathsf{F}^2$. □

In the following, not only the range of one matrix but of a family of matrices $M_i \in \mathbb{K}^{n\times m_i}$ $(1 \le i \le p)$ is to be optimised:

$$\max\left\{\sum_{i=1}^{p} \|PM_i\|_\mathsf{F}^2 : P \text{ orthogonal projection with } \operatorname{rank}(P) = r\right\}. \tag{2.28b}$$

Problem (2.28b) can be reduced to (2.28a) by agglomerating M_i into

$$M := [M_1\, M_2\, \cdots\, M_p]\,. \tag{2.28c}$$

The optimal projection $P = P_1^{(r)}$ from Remark 2.31 is obtained from the left singular vectors of M. If the left singular value decompositions of M_i are known, the corresponding decomposition of M can be simplified.

Lemma 2.35. *The data of the left singular value decompositions of $M_i = U_i\Sigma_i V_i^\mathsf{T}$ consist of U_i and Σ_i. The corresponding data U, Σ of $M = U\Sigma V^\mathsf{T}$ from (2.28c) can also be obtained from the left singular value decomposition of*

$$M' := [U_1\Sigma_1 \quad U_2\Sigma_2 \quad \cdots \quad U_p\Sigma_p]\,. \tag{2.28d}$$

Proof. Since $MM^\mathsf{H} = \sum_{i=1}^p M_iM_i^\mathsf{H} = \sum_{i=1}^p U_i\Sigma_i^2U_i^\mathsf{H}$ coincides with the product $M'M'^\mathsf{H} = \sum_{i=1}^p U_i\Sigma_i^2U_i^\mathsf{H}$, the diagonalisation $U\Sigma^2U^\mathsf{H}$ is identical. □

2.7 Linear Algebra Procedures

For later use, we formulate procedures based on the previous techniques.

The reduced QR decomposition is characterised by the dimensions n and m, the input matrix $M \in \mathbb{K}^{n\times m}$, the rank r, and resulting factors Q and R. The corresponding procedure is denoted by

$$\begin{array}{ll} \text{procedure } \mathbf{RQR}(n,m,r,M,Q,R); & \{\text{reduced QR decomposition}\} \\ \text{input: } M \in \mathbb{K}^{n\times m}; & \\ \text{output: } r = \operatorname{rank}(M),\ Q \in \mathbb{K}^{n\times r} \text{ orthogonal,} & \\ \qquad\quad R \in \mathbb{K}^{r\times m} \text{ upper triangular.} & \end{array} \tag{2.29}$$

and requires $N_{\mathrm{QR}}(n,m)$ operations (cf. Lemma 2.19).

The modified QR decomposition from (2.18) produces a further permutation matrix P and the decomposition of R into $[R_1\ R_2]$:

$$\begin{array}{ll} \text{procedure } \mathbf{PQR}(n,m,r,M,P,Q,R_1,R_2); & \{\text{pivotised QR decomposition}\} \\ \text{input: } M \in \mathbb{K}^{n\times m}; & \\ \text{output: } Q \in \mathbb{K}^{n\times r} \text{ orthogonal, } P \in \mathbb{K}^{m\times m} \text{ permutation matrix,} & \\ \qquad\quad R_1 \in \mathbb{K}^{r\times r} \text{ upper triangular with } r = \operatorname{rank}(M),\ R_2 \in \mathbb{K}^{r\times(m-r)}. & \end{array} \tag{2.30}$$

A modified version of $\mathbf{PQR}$ will be presented in (2.40).

The (two-sided) reduced singular value decomposition from Definition 2.23 leads to

$$\begin{array}{ll} \text{procedure } \mathbf{RSVD}(n,m,r,M,U,\varSigma,V); & \{\text{reduced SVD}\} \\ \text{input: } M \in \mathbb{K}^{n\times m}; & \\ \text{output: } U \in \mathbb{K}^{n\times r},\ V \in \mathbb{K}^{m\times r} \text{ orthogonal with } r = \operatorname{rank}(M), & \\ \qquad\quad \varSigma = \operatorname{diag}\{\sigma_1,\ldots,\sigma_r\} \in \mathbb{R}^{r\times r} \text{ with } \sigma_1 \geq \ldots \geq \sigma_r > 0. & \end{array} \tag{2.31}$$

Here the integers n, m may also be replaced by index sets I and J. Concerning the cost $N_{\mathrm{SVD}}(n,m)$ see Corollary 2.21a.

The left-sided reduced singular value decomposition (cf. Remark 2.24) is denoted by

$$\begin{array}{ll} \text{procedure } \mathbf{LSVD}(n,m,r,M,U,\varSigma); & \{\text{left-sided reduced SVD}\} \\ \text{input: } M \in \mathbb{K}^{n\times m}; & \\ \text{output: } U,\ r,\ \varSigma \text{ as in (2.31).} & \end{array} \tag{2.32}$$

Its cost is

$$N_{\mathrm{LSVD}}(n,m) := \frac{1}{2}n\,(n+1)\,N_m + \frac{8}{3}n^3,$$

where N_m is the cost of the scalar product of rows of M. In general, $N_m = 2m-1$ holds, but it may be less for structured matrices (cf. Remark 7.12).

In the procedures above, M is a general matrix from $\mathbb{K}^{n\times m}$. Matrices $M \in \mathcal{R}_r$ (cf. (2.6)) may be given in the form

$$M=\sum_{\nu=1}^{r}\sum_{\mu=1}^{r} c_{\nu\mu}a_\nu b_\mu^{\mathsf{H}}=ACB^{\mathsf{H}} \quad \begin{pmatrix} a_\nu\in\mathbb{K}^n,\ A=[a_1a_2\cdots]\in\mathbb{K}^{n\times r},\\ b_\nu\in\mathbb{K}^m,\ B=[b_1\,b_2\cdots]\in\mathbb{K}^{m\times r}\end{pmatrix}. \tag{2.33}$$

Then the following approach has a cost proportional to $n+m$ if $r\ll n,m$ (cf. [86, Alg. 2.5.3]), but also for $r\gg n,m$ it is cheaper than the direct computation[17] of the product $M=ACB^{\mathsf{H}}$ followed by a singular value decomposition.

Remark 2.36. For $M=ACB^{\mathsf{H}}$ from (2.33) compute the reduced QR decompositions[18]

$$A=Q_AR_A \text{ and } B=Q_BR_B \quad \text{with } \begin{cases} Q_A\in\mathbb{K}^{n\times r_A},\ r_A:=\operatorname{rank}(A),\\ Q_B\in\mathbb{K}^{n\times r_B},\ r_B:=\operatorname{rank}(B),\end{cases}$$

followed by the singular value decomposition $R_ACR_B^{\mathsf{H}}=\hat{U}\Sigma\hat{V}^{\mathsf{H}}$. Then the singular value decomposition of M is given by $U\Sigma V^{\mathsf{H}}$ with $U=Q_A\hat{U}$ and $V=Q_B\hat{V}$. The cost of this calculation is

$$N_{\mathrm{QR}}(n,r)+N_{\mathrm{QR}}(m,r)+N_{\mathrm{LSVD}}(r_A,r_B)+2r_Ar_Br+2\bar{r}r^2+2\left(r_A^2n+r_B^2m\right)$$

with $\bar{r}:=\min\{r_A,r_B\}\le\min\{n,m,r\}$. In the symmetric case of $A=B$ and $n=m$ with $\bar{r}:=\operatorname{rank}(A)$, the cost reduces to

$$N_{\mathrm{QR}}(n,r)+2\bar{r}r^2+\bar{r}^2\left(r+2n+\tfrac{8}{3}\bar{r}\right).$$

Let $B'=[b_1',\ldots,b_{r'}']\in\mathbb{K}^{n\times r'}$ and $B''=[b_1'',\ldots,b_{r''}'']\in\mathbb{K}^{n\times r''}$ contain two systems of vectors. Often, B' and B'' correspond to bases of subspaces $U'\subset V$ and $U''\subset V$. A basic task is the construction of a basis[19] $B=[b_1,\ldots,b_r]$ of $U:=U'+U''$. Furthermore, the matrices $T'\in\mathbb{K}^{r\times r'}$ and $T''\in\mathbb{K}^{r\times r''}$ with

$$B'=BT' \text{ and } B''=BT'', \text{ i.e., } b_j'=\sum_{i=1}^{r}T_{ij}'b_i,\ b_j''=\sum_{i=1}^{r}T_{ij}''b_i, \tag{2.34}$$

are of interest. The corresponding procedure is

```
procedure JoinBases($B', B'', r, B, T', T''$);        {joined bases}
input:  $B'\in\mathbb{K}^{n\times r'}$, $B''\in\mathbb{K}^{n\times r''}$,
output: $r=\operatorname{rank}[B'\ B'']$; $B$ basis of range($[B'\ B'']$),        (2.35)
        $T'\in\mathbb{K}^{r\times r'}$ and $T''\in\mathbb{K}^{r\times r''}$ with (2.34).
```

A possible realisation starts from $B=[b_1',\ldots,b_{r'}',b_1'',\ldots,b_{r''}'']\in\mathbb{K}^{n\times(r'+r'')}$ and performs the reduced QR factorisation $B=QR$ by $\mathbf{RQR}(n,r'+r'',r,B,P,Q,R1,R2)$. Then the columns of $P^{\mathsf{T}}Q$ form the basis B, while $[T',T'']=R:=[R1,R2]$.

[17] For instance, the direct computation is cheaper if $n=m=r=r_A=r_B$.

[18] Possibly, permutations according to (2.30) are necessary.

[19] We call $B=[b_1,\ldots,b_r]$ a basis, meaning that the set $\{b_1,\ldots,b_r\}$ is the basis.

If B' is a basis, if may be advantageous to let the basis vectors $b_i = b'_i$ from B' unchanged, whereas for $i > r'$, b_i is the i-th column of Q. Then $T' = \left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right]$ holds, while T'' is as before.

If all bases B', B'', B are orthonormal, the second variant from above completes the system B' to an orthonormal basis B:

$$\begin{array}{ll} \text{procedure } \mathbf{JoinONB}(\mathfrak{b}', \mathfrak{b}'', r, \mathfrak{b}, T', T''); & \{\text{joined orthonormal basis}\} \\ \text{input: } \; B' \in \mathbb{K}^{n\times r'}, \; B'' \in \mathbb{K}^{n\times r''} \text{ orthonormal bases,} & \\ \text{output: } B \text{ orthonormal basis of } \operatorname{range}([B' \; B'']);\ r, T', T'' \text{ as in (2.35).} & \end{array} \tag{2.36}$$

The cost of both procedures is $N_{\mathrm{QR}}(n, r' + r'')$.

2.8 Deflation Techniques

2.8.1 Dominant Columns

We consider again the minimisation problem (2.27): $\min_{\operatorname{rank}(M')\le k} \|M - M'\|$ for $M \in \mathbb{K}^{n\times m}$ and $\|\cdot\| = \|\cdot\|_2$ or $\|\cdot\| = \|\cdot\|_{\mathsf{F}}$. Without loss of generality, we assume that the minimising matrix M_k satisfies $\operatorname{rank}(M_k) = k$; otherwise, replace k by $k' := \operatorname{rank}(M_k)$ and note that $\min\limits_{\operatorname{rank}(M')\le k} \|M - M'\| = \min\limits_{\operatorname{rank}(M')\le k'} \|M - M'\|$.

The minimising matrix $M_k \in \mathcal{R}_k$ is of the form

$$M_k = AB^{\mathsf{T}}, \quad \text{where } A \in \mathbb{K}^{n\times k} \text{ and } B \in \mathbb{K}^{m\times k}$$

and $\operatorname{range}(M_k) = \operatorname{range}(A)$. The singular value decomposition $M = U\Sigma V^{\mathrm{T}}$ yields the matrices $A = U'\Sigma'$ and $B = V'$, where the matrices U', Σ', V' consist of the first k columns of U, Σ, V. Since $A = U'\Sigma' = MV'^{\mathrm{T}}$, the columns of A are linear combinations of *all* columns of M. The latter fact is a disadvantage is some cases. For a concrete numerical approach, we have to represent A. If the columns m_j of M are represented as full vectors from $\mathbb{K}^n$, a linear combination is of the same kind and leads to no difficulty. This can be different, if other representations are involved. To give an example for an extreme case, replace $\mathbb{K}^{n\times m} = (\mathbb{K}^n)^m$ by X^m, where X is a subspace of, say, $L^2([0,1])$. Let the columns be functions like x^ν or $\exp(\alpha x)$. Such functions can be simply coded together with procedures for pointwise evaluation and mutual scalar products. However, linear combinations cannot be simplified. For instance, scalar products of linear combinations must be written as double sums of elementary scalar products. As a result, the singular value decomposition reduces the rank of M to k, but the related computational cost may be larger than before.

This leads to a new question. Can we find $AB^{\mathsf{T}} \in \mathcal{R}_k$ approximating M such that $A = [c_{j_1} \cdots c_{j_k}]$ consists of k (different) columns of M? In this case, AB^{T} involves only k columns of M instead of all. For this purpose we define

$$\mathcal{R}_k(M) := \{AB^{\mathsf{T}} \in \mathcal{R}_k : A = [M[\cdot, j_1], \cdots, M[\cdot, j_k]] \text{ with } 1 \le j_\kappa \le m\}, \tag{2.37}$$

using the notations from above. The minimisation (2.27) is now replaced by

$$\text{find } M_k \in \mathcal{R}_k(M) \text{ with } \quad \|M - M_k\| = \min_{M' \in \mathcal{R}_k(M)} \|M - M'\| . \tag{2.38}$$

Since there are $\binom{m}{k}$ different combinations of columns, we do not try to solve this combinatorial problem exactly. Instead, we are looking for an approximate solution.

By procedure **PQR** from (2.30), we obtain the QR decomposition $MP = QR$ with $R = [R_1 \; R_2]$. First we discuss the case $r := \operatorname{rank}(M) = m$, which is equivalent to M possessing full rank. Then R_2 does not exist ($m - r = 0$ columns) and $R := R_1$ is a square upper triangular matrix with non-vanishing diagonal entries. Thanks to the pivoting strategy, the columns of R are of decreasing Euclidean norm. Let $k \in \{1, \ldots, m-1\}$ be the desired rank from problem (2.38). We split the matrices into the following blocks:

$$R = \begin{array}{|c|c|}\hline R' & S \\ \hline 0 & R'' \\ \hline \end{array} \qquad \text{with} \begin{cases} R' \in \mathbb{K}^{k \times k}, \; R'' \in \mathbb{K}^{(m-k)\times(m-k)} \text{ upper triangular,} \\ S \in \mathbb{K}^{k \times (m-k)}, \end{cases}$$

$$Q = \begin{array}{|c|c|}\hline Q' & Q'' \\ \hline \end{array} \qquad \text{with} \quad Q' \in \mathbb{K}^{n \times k}, \; Q'' \in \mathbb{K}^{n \times (m-k)}.$$

Then $Q'R'$ corresponds to the first k columns of MP. As P is a permutation matrix, these columns form the matrix A as required in (2.37). The approximating matrix is defined by $M_k^{\mathsf{PQR}} := Q' \, [R' \; S]$, where $[R' \; S] \in \mathbb{K}^{k \times m}$, i.e., the matrix B from (2.37) is $B = [R' \; S]^{\mathsf{T}}$.

Proposition 2.37. *The matrix* $M_k^{\mathsf{PQR}} := Q' \, [R' \; S]$ *constructed above belongs to* $\mathcal{R}_k(M)$ *and satisfies the following estimates:*

$$\|M - M_k^{\mathsf{PQR}}\|_2 \le \|R''\|_2 , \quad \|M - M_k^{\mathsf{PQR}}\|_{\mathsf{F}} \le \|R''\|_{\mathsf{F}} , \tag{2.39a}$$

$$\sigma_k(M_k^{\mathsf{PQR}}) \le \sigma_k(M) \le \sqrt{\sigma_k^2(M_k^{\mathsf{PQR}}) + \|R''\|_2^2}. \tag{2.39b}$$

Proof. 1) By construction, $M - M_k^{\mathsf{PQR}} = Q \left[\begin{smallmatrix} 0 & 0 \\ 0 & R'' \end{smallmatrix}\right]$ holds and leads to (2.39a).
2) (2.39b) follows from Lemma 2.28. □

Now, we investigate the case $r < m$. Then the full QR decomposition would lead to QR with $R = \left[\begin{smallmatrix} R' \\ 0 \end{smallmatrix}\right]$ with zeros in the rows $r+1$ to m. These zero rows are omitted by the reduced QR decomposition. The remaining part R' (again denoted by R) is of the shape $R = [R_1 \; R_2]$, where R_1 has upper triangular form. As $\operatorname{rank}(M) = r$, the approximation rank k from (2.38) should vary in $1 \le k < r$. Again the columns of R_1 are decreasing, but the choice of the first k columns may not be the optimal one. This is illustrated by the following example.

Let $M = \left[\begin{smallmatrix} 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{smallmatrix}\right]$ with $r = 2 < m = 4$. Since the first column has largest norm, procedure **PQR** produces $P = I$ (no permutations) and

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad R_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \quad R_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Let $k = 1$. Choosing the first column of Q and first row of R, we obtain $M_1^{[1]} := \left[\begin{smallmatrix}2&1&1&1\\0&0&0&0\end{smallmatrix}\right]$. The approximation error is $\varepsilon_1 := \|M - M_1^{[1]}\| = \left\|\left[\begin{smallmatrix}0&0&0&0\\0&1&1&1\end{smallmatrix}\right]\right\| = \sqrt{3}$. Note that we cannot choose the second column $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ of Q and the second row of R instead, since $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ is not a column of M, i.e., the resulting approximation does not belong to $\mathcal{R}_1(M)$. A remedy is to change the pivot strategy in Step 2a from page 29. We choose the second column of M as first column of Q. For this purpose let P be the permutation matrix corresponding to $1 \leftrightarrow 2$. The QR decomposition applied to $MP = \left[\begin{smallmatrix}1&2&1&1\\1&0&1&1\end{smallmatrix}\right]$ yields

$$Q = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}, \quad R_1 = \sqrt{2}\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}, \quad R_2 = \sqrt{2}\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}.$$

The first column of Q and the first row of R result in $M_1^{[2]}P$ with the smaller approximation error

$$\varepsilon_2 := \|M - M_1^{[2]}\| = \left\|\begin{bmatrix}1\,0\,0\,0\\ -1\,0\,0\,0\end{bmatrix}\right\| = \sqrt{2} < \sqrt{3}.$$

The reason for $\varepsilon_2 < \varepsilon_1$ is obvious: although $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ is of smaller norm than $\left[\begin{smallmatrix}2\\0\end{smallmatrix}\right]$, it has a higher weight because it appears in three columns of M. To take this weight into consideration, we need another pivot strategy.

Let $M = [c_1 \cdots c_m] \in \mathbb{K}^{n\times m}$. Each column $c_j \neq 0$ is associated with the orthogonal projection $P_j := \|c_j\|^{-2} c_j c_j^{\mathsf{H}}$ onto $\operatorname{span}\{c_j\}$. We call c_i a *dominant column*, if $\|P_i M\| = \max_{1\le j\le m} \|P_j M\|$. Equivalently, $\|M - P_j M\|$ is minimal for $j = i$. Let P be the permutation matrix corresponding to the exchange $1 \leftrightarrow i$. Then $MP = QR$ leads to Q with $c_i / \|c_i\|$ as first column. The first row of R is $r_1^{\mathsf{H}} := \|c_i\|^{-1} c_i^{\mathsf{H}} MP$. Hence, the choice of the dominant column ensures that the approximation (2.38) with $k = 1$ is given by $\|c_i\|^{-1} c_i r_1^{\mathsf{H}} P^{\mathsf{H}} = P_i M$.

The calculation of a dominant column is discussed in the next lemma.

Lemma 2.38. *For $M = [c_1 \cdots c_m]$ set*

$$Z = (\zeta_{jk})_{1\le j,k\le m} \qquad \text{with } \zeta_{jk} := \langle c_k, c_j\rangle / \|c_j\|.$$

Then the index $i_{\max} \in \{1, \ldots, m\}$ with $\|\zeta_{i_{\max},\bullet}\| = \max_{1\le j\le m} \|\zeta_{j,\bullet}\|$ characterises the dominant column.

Proof. Because of

$$P_j M = \|c_j\|^{-2}\left(c_j\, c_j^{\mathsf{H}}\, c_k\right)_{1\le k\le m} = \left(\frac{\zeta_{jk}}{\|c_j\|}\, c_j\right)_{1\le k\le m},$$

its norm is $\|P_j M\| = \sqrt{\sum_k |\zeta_{jk}|^2} = \|\zeta_{j,\bullet}\|$. □

The concept of the dominant column leads to the following variant of $\mathbf{PQR}$:

```
procedure $\mathbf{DCQR}(n,m,r,M,P,Q,R_1,R_2)$;     {pivot by dominant column}
input:  $M \in \mathbb{K}^{n\times m}$;
output: $Q \in \mathbb{K}^{n\times r}$ orthogonal, $P \in \mathbb{K}^{m\times m}$ permutation matrix,
        $R_1 \in \mathbb{K}^{r\times r}$ upper triangular with $r = \operatorname{rank}(M)$, $R_2 \in \mathbb{K}^{r\times(m-r)}$.
for $j := 1$ to $r$ do
begin   determine $i \in \{j,\ldots,m\}$ such that $c_i := M[\bullet,i]$          (2.40)
           is the dominant column of $M[\bullet, j:m]$;
        permute $j \leftrightarrow i$ (change of $P$)
        $Q[j,\bullet] := c_i$ ; $M[\bullet,i] := M[\bullet,j]$;
        $M[\bullet,j+1:m] := (I - P_i)\, M[\bullet,j+1:m]$      $(P_i := \|c_i\|^{-2} c_i c_i^{\mathsf{H}})$
end;    {determination of $R_1, R_2$ as usual; here omitted}
```

Corollary 2.39. In practical applications, it suffices to determine $\hat{\imath}_{\max}$ instead of $i_{\max}$, where $\|\zeta_{\hat{\imath}_{\max},\bullet}\|$ is sufficiently close to $\|\zeta_{i_{\max},\bullet}\|$. For this purpose, order the columns by decreasing norm: $MP =: [c_1 \cdots c_m]$ with $\|c_j\| \geq \|c_{j+1}\|$. Choose $m_0 \in \{1,\ldots,m-1\}$ with $\varepsilon^2 := \sum_{k=m_0+1}^{m} \|c_k\|^2$ sufficiently small and reduce the maximisation to the first m_0 columns: $\hat{\imath}_{\max}$ is the maximiser of $\max_{1\leq j\leq m_0} \sqrt{\sum_{k=1}^{m_0} |\zeta_{jk}|^2}$. Then the two maxima are related by

$$\|\zeta_{\hat{\imath}_{\max},\bullet}\| \leq \|\zeta_{i_{\max},\bullet}\| \leq \sqrt{\|\zeta_{\hat{\imath}_{\max},\bullet}\|^2 + \varepsilon^2}.$$

To estimate the cost, we start with the determination of the matrix Z from Lemma 2.38. The $\frac{1}{2}m(m+1)$ scalar products are to be computed only once. After one elimination step ($c_k \mapsto P_i c_k$) they can be updated[20] without new scalar products (cost: $2(m-1)m$). Also the application of P_i to M does not require scalar products, since they are precomputed. The cost of the original procedure $\mathbf{DCQR}$ is

$$N_{\mathbf{DCQR}} = 4\Big(mr - \tfrac{1}{2}r^2\Big)n + m^2 n + 2rm(m-r)$$

plus lower order terms. If m is immediately reduced to m_0, the cost is reduced correspondingly.

Remark 2.40. Assume that $n = \hat{n}p$ with $\hat{n}, p \in \mathbb{N}$. A further reduction of cost is possible, if the scalar product $\langle a,b\rangle = \sum_{i=1}^{n} a_i \overline{b_i}$ is approximated by

$$\langle a,b\rangle_p := p \sum_{i=1}^{\hat{n}} a_i \overline{b_i}.$$

$\langle\cdot,\cdot\rangle_p$ is not a scalar product in $\mathbb{K}^n$, but if $a,b \in \mathbb{K}^n$ is a smooth grid function, $\langle a,b\rangle_p$ approximates $\langle a,b\rangle$. With this modification in $\mathbf{DCQR}$, n in $N_{\mathbf{DCQR}}$ can be reduced to $\hat{n}$. Note that the computed Q is not strictly orthogonal.

Following the ideas from [120], one may start with $\hat{n} = O(1)$ to select $\hat{r}$ columns, where $r < \hat{r} = O(r)$. Then the final call of $\mathbf{DCQR}$ (with exact scalar product) is applied to the reduced matrix $M \in \mathbb{K}^{n\times\hat{r}}$. Note that this procedure yields the same result as $\mathbf{DCQR}$ applied to the original matrix M, if the finally chosen r columns are among the $\hat{r}$ columns selected by the first heuristic step. The total work is $O(mr + m_0^2 r + r^2 n + r^3)$ with m_0 from Corollary 2.39.

[20] To be precise, only $\langle c_k, c_j\rangle$ and $|\zeta_{jk}|^2$ need to be updated.

2.8.2 Reduction of a Basis

A related, but more general problem is discussed next. Let $B \in \mathbb{K}^{I\times J}$ be a matrix containing $\#J$ basis vectors

$$b_\nu = B[\cdot,\nu] \in \mathbb{K}^I \qquad (\nu \in J)\,. \tag{2.41a}$$

Furthermore, p vectors $x_\mu \in \mathrm{range}(B) \subset \mathbb{K}^I$ are given, gathered in the matrix

$$X \in \mathbb{K}^{I\times p}, \quad x_\mu = X[\cdot,\mu] \qquad (1 \le \mu \le p)\,. \tag{2.41b}$$

Because of $x_\mu \in \mathrm{range}(B) = \mathrm{span}\{b_\nu : \nu \in J\}$, there are coefficients $a_{\nu\mu}$ such that

$$x_\mu = \sum\nolimits_{\nu\in J} a_{\nu\mu} b_\nu \qquad (1 \le \mu \le p)\,, \tag{2.41c}$$

which is equivalent to the matrix formulation

$$X = BA \qquad \text{with } A = (a_{\nu\mu})_{\nu\in J, 1\le\mu\le p} \in \mathbb{K}^{J\times p}. \tag{2.41d}$$

Now, we want to approximate X by a reduced basis. Similarly as in §2.8.1, the smaller basis must consist of a subset of the vectors b_ν from (2.41a). Split J into $J = J'\dot\cup J''$, where J' is the remaining index set, while basis vectors b_μ with $\mu \in J''$ are to be omitted. The simultaneous approximation of all x_ν is expressed by minimising the Frobenius norm:

$$\text{given } J', \text{ find } A^{J'} \in \mathbb{K}^{J'\times p} \text{ such that} \tag{2.42a}$$

$$\varepsilon(J') := \|X - B'A^{J'}\|_\mathsf{F} = \min_{A'\in\mathbb{K}^{J'\times p}} \|X - B'A'\|_\mathsf{F}, \tag{2.42b}$$

where B' is the restriction $B|_{I\times J'} \in \mathbb{K}^{I\times J'}$. The best choice of $J' \subset J$ under the constraint $\#J' = q \in \{1,\ldots,\#J\}$ is given by

$$J' \text{ minimiser of } \quad \varepsilon_q := \min\{\varepsilon(J') : J' \subset J \text{ with } \#J' = q\}. \tag{2.42c}$$

The problem from §2.8.1 corresponds to the particular case of $B{=}M$ and $X{=}M$.

The minimisation in (2.42a,b) is a least-squares problem. For its solution we define the Gram matrix

$$G = (g_{\nu\mu})_{\nu,\mu\in J} \qquad \text{with } g_{\nu\mu} = \langle b_\mu, b_\nu\rangle\,.$$

The disjoint partition $J = J' \dot\cup J''$ leads to the block decomposition

$$G = \begin{bmatrix} G' & G^\sharp \\ (G^\sharp)^\mathsf{H} & G'' \end{bmatrix} \text{ with } \begin{cases} G' = (g_{\nu\mu})_{\nu,\mu\in J'}\,, \; G'' = (g_{\nu\mu})_{\nu,\mu\in J''}\,, \\ G^\sharp = (g_{\nu\mu})_{\nu\in J',\mu\in J''}\,. \end{cases} \tag{2.43}$$

The decomposition defines the Schur complement

$$S := G'' - (G^\sharp)^\mathsf{H} G'^{-1} G^\sharp. \tag{2.44}$$

Lemma 2.41. *Let $X = BA$. The solution $A^{J'}$ of (2.42a,b) is given by*

$$A^{J'} := \begin{bmatrix} I & G'^{-1}G^{\sharp} \\ 0 & 0 \end{bmatrix} A. \tag{2.45a}$$

Splitting $A \in \mathbb{K}^{J\times p}$ into $\left[\begin{smallmatrix} A' \\ A'' \end{smallmatrix}\right]$ with $A' \in \mathbb{K}^{J'\times p}$ and $A'' \in \mathbb{K}^{J''\times p}$, we rewrite the latter equation as

$$A^{J'} = A' + G'^{-1}G^{\sharp}A''. \tag{2.45b}$$

The minimum $\varepsilon(J')$ from (2.42b) equals

$$\varepsilon(J') = \sqrt{\sum\nolimits_{\mu=1}^{p} \left\langle S^{-1}a''_{\mu}, a''_{\mu}\right\rangle}, \qquad \text{where } a''_{\mu} := (a_{\nu\mu})_{\nu\in J''} \in \mathbb{K}^{J''}. \tag{2.45c}$$

Proof. First, we may assume $p = 1$, i.e., X and A are vectors. Define the subspaces $U' := \mathrm{span}\{b_\mu : \mu \in J'\}$ and $U'' := \mathrm{span}\{b_\mu : \mu \in J''\}$. Then, $\mathrm{range}(B)$ is the direct sum[21] $U' \oplus U''$. The best approximation in U' is given by the orthogonal projection P' onto U', i.e., $P'X \in U'$ is the desired solution $B'A^{J'}$. It remains to determine the coefficients of $A^{J'}$. Split X into $X = BA = X' + X''$ with $X' = B'A' \in U'$ and $X'' = B''A'' \in U''$ (for A', A'', B', and B'' see the lines after (2.42b) and (2.45a)). Let P be the mapping $X = BA \mapsto B'A^{J'}$ with $A^{J'}$ from (2.45b). We have to show that $P = P'$. Obviously, P is a mapping into U'. Furthermore, $X \in U'$ implies $X = B'A'$ and $A'' = 0$ and, therefore, $PX = X$. This proves that P is a projection. Next, one verifies that $\langle (I-P)\,X, b_\nu\rangle = 0$ for $\nu \in J'$, i.e., $\mathrm{range}(I-P) \subset U'^{\perp}$. Hence, P is the orthogonal projection P'.

$X = X' + X''$ leads to the error $(I-P)\,X = X'' - PX''$ with the squared Frobenius norm

$$\begin{aligned} \|X - B'A^{J'}\|^2 &= \left\| B \begin{bmatrix} -G'^{-1}G^{\sharp} \\ I \end{bmatrix} A'' \right\|^2 = \left\langle G \begin{bmatrix} -G'^{-1}G^{\sharp} \\ I \end{bmatrix} A'', \begin{bmatrix} -G'^{-1}G^{\sharp} \\ I \end{bmatrix} A'' \right\rangle \\ &= \left\langle S^{-1}A'', A''\right\rangle. \end{aligned}$$

For $p > 1$, we have to sum over all columns a''_{μ} of A''. □

The calculation of ε_q from (2.42c) can be quite costly, since there may be very many subsets $J' \subset J$ with $\#J' = q$. For each J' one has to evaluate $\varepsilon(J')$ involving the Schur complement $S = S^{J'}$. Here, it is helpful that all inverse Schur complements can be computed simultaneously.

Lemma 2.42. *Compute the inverse G^{-1}. Then, for any $\emptyset \neq J' \subset J$, the inverse Schur complement $(S^{J'})^{-1} \in \mathbb{K}^{J''\times J''}$ corresponding to J' is the restriction of G^{-1} to the part $J'' \times J''$:*

$$(S^{J'})^{-1} = G^{-1}|_{J''\times J''}.$$

[21] $U := U' \oplus U''$ is called a *direct sum*, if every $u \in U$ has a unique decomposition $u = u' + u''$ with $u' \in U'$ and $u'' \in U''$.

Proof. Note the identity

$$G^{-1} = \begin{bmatrix} G'^{-1} + G'^{-1} G^\sharp S^{-1} \left(G^\sharp\right)^{\mathsf{H}} G'^{-1} & -G'^{-1} G^\sharp S^{-1} \\ -S^{-1} \left(G^\sharp\right)^{\mathsf{H}} G'^{-1} & S^{-1} \end{bmatrix}$$

(only for depicting the matrix, we order J such that $\nu \in J'$ are taken first and $\nu \in J''$ second). □

We make use of Lemma 2.42 for $q := \#J - 1$, i.e., for $\#J'' = 1$. Let $J'' = \{\iota\}$ for $\iota \in J$. In this case, $(S^{J'})^{-1}$ is the 1×1 matrix $(G^{-1})_{\iota\iota}$. The approximation error $\varepsilon(J') = \varepsilon(J \backslash \{\iota\})$ from (2.45c) becomes

$$\varepsilon(\iota) := \left(G^{-1}\right)_{\iota\iota} \sum_{\mu=1}^{p} \left|a''_{\iota\mu}\right|^2 \qquad (\iota \in J)\,, \tag{2.46a}$$

involving the diagonal entries $(G^{-1})_{\iota\iota}$ of G^{-1}. Minimisation over all $\iota \in J$ yields

$$\varepsilon_{\#J-1} := \min\{\varepsilon(\iota) : \iota \in J\}. \tag{2.46b}$$

This leads to the following algorithm.

Algorithm 2.43. (a) Let $X = BA$ with $B \in \mathbb{K}^{I \times J}$. Compute the Gram matrix and its inverse $G^{-1} \in \mathbb{K}^{J \times J}$. Determine $\iota \in J$ with minimal value $\varepsilon(\iota)$ (cf. (2.46a)). Then the reduction of J to $J' := J \backslash \{\iota\}$ yields the best approximation of X by $X' \in \operatorname{span}\{b_\nu : J'\}$ among all J' with $\#J' = \#J - 1$.

(b) If one wants to omit more than one basis vector, the procedure from Part (a) can be iterated. The computation of the reduced Gram matrix $G' := G|_{J' \times J'}$ and its inverse need not be repeated. Decompose G and its inverse into[22]

$$G = \begin{bmatrix} G' & g \\ g^{\mathsf{H}} & g'' \end{bmatrix} \text{ and } G^{-1} = \begin{bmatrix} H' & h \\ h^{\mathsf{H}} & h'' \end{bmatrix} \quad \text{with } G, H \in \mathbb{K}^{J' \times J'},\ g, h \in \mathbb{K}^{J'}.$$

Then the inverse of G' is given by

$$(G')^{-1} = H' \left(I - \frac{1}{1 - h^{\mathsf{H}} g} g h^{\mathsf{H}} \right).$$

Corollary 2.44. The general assumption of this section is that the columns of B form a basis. If this is not the case, i.e., $r := \operatorname{rank}(B) < \#J$, the Gram matrix is singular and the algorithm from above cannot be applied. Instead, one can apply the procedure **PQR** from (2.30) to determine r linearly independent columns of B. Let their indices form the subset $J' \subset J$. The reduction from J to J' does not introduce any error: $\varepsilon(J') = 0$. A further reduction can be performed by Algorithm 2.43, since the Gram matrix corresponding to J' is regular.

[22] The index set J is ordered such that ι is the last index.

Chapter 3
Algebraic Foundations of Tensor Spaces

Abstract Since tensor spaces are in particular vector spaces, we start in *Sect. 3.1* with vector spaces. Here, we introduce the free vector space (§3.1.2) and the quotient vector space (§3.1.3) which are needed later. Furthermore, the spaces of linear mappings and dual mappings are discussed in §3.1.4. The core of this chapter is *Sect. 3.2* containing the definition of the tensor space. *Section 3.3* is devoted to linear and multilinear mappings as well as to tensor spaces of linear mappings. Algebra structures are discussed in *Sect. 3.4*. Finally, symmetric and antisymmetric tensors are defined in *Sect. 3.5*.

3.1 Vector Spaces

3.1.1 Basic Facts

We recall that V is a *vector space* (also named '*linear space*') over the field $\mathbb{K}$, if $V \neq \emptyset$ is a commutative group (where the group operation is written as addition) and if a multiplication

$$\cdot : \mathbb{K} \times V \to V$$

is defined with the following properties:

$$\begin{aligned}
(\alpha\beta)\cdot v &= \alpha\cdot(\beta\cdot v) && \text{for } \alpha,\beta\in\mathbb{K},\ v\in V,\\
(\alpha+\beta)\cdot v &= \alpha\cdot v+\beta\cdot v && \text{for } \alpha,\beta\in\mathbb{K},\ v\in V,\\
\alpha\cdot(v+w) &= \alpha\cdot v+\alpha\cdot w && \text{for } \alpha\in\mathbb{K},\ v,w\in V,\\
1\cdot v &= v && \text{for } v\in V,\\
0\cdot v &= 0 && \text{for } v\in V,
\end{aligned} \tag{3.1}$$

where on the left-hand side 1 and 0 are the respective multiplicative and additive unit elements of the field $\mathbb{K}$, while on the right-hand side 0 is the zero element of the group V.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_3,

The sign '$\cdot$' for the multiplication $\cdot : \mathbb{K} \times V \to V$ is usually omitted, i.e., αv is written instead of $\alpha \cdot v$. Only when $\alpha\,(v+w)$ may be misunderstood as a function α evaluated at $v+w$, we prefer the original notation $\alpha \cdot (v+w)$.

Any vector space V has a *basis* $\{v_i : i \in B\} \subset V$ with the property that it is linearly independent and spans $V = \operatorname{span}\{v_i : i \in B\}$. In the infinite case of $\#B = \infty$, *linear independence* means that all *finite* sums $\sum_i a_i v_i$ vanish if and only if $a_i = 0$. Analogously, $\operatorname{span}\{v_i : i \in B\}$ consists of all *finite* sums $\sum_i a_i v_i$. Here 'finite sum' means a sum with finitely many terms or equivalently a sum, where only finitely many terms do not vanish.

The cardinality $\#B$ is independent of the choice of the basis and called *dimension*, denoted by $\dim(V)$. Note that there are many infinite cardinalities. Equality $\dim(V) = \dim(W)$ holds if and only if there is a bijection $B_V \leftrightarrow B_W$ between the corresponding index sets of the bases. Vector spaces of identical dimension are isomorphic. For finite dimension $n \in \mathbb{N}_0$ the model vector space is $\mathbb{K}^n$. The isomorphism between a general vector space V with basis $\{v_1, \ldots, v_n\}$ and $\mathbb{K}^n$ is given by $v = \sum \alpha_\nu v_\nu \mapsto (\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$.

Example 3.1. Let I be an infinite, but countable index set, i.e., $\#I = \aleph_0 := \#\mathbb{N}$. Then $\ell(I) = \mathbb{K}^I$ denotes the set of all sequences $(a_i)_{i \in I}$. The set $\ell(I)$ may be also be viewed as the set of all mappings $I \to \mathbb{K}$. A subset of $\ell(I)$ is

$$\ell_0(I) := \{a \in \ell(I) : a_i = 0 \text{ for almost all } i \in I\}\,. \tag{3.2}$$

The unit vectors $\{e^{(i)} : i \in I\}$ from (2.2) form a basis of $\ell_0(I)$, so that the dimension equals $\dim(\ell_0(I)) = \#I = \aleph_0$. However, the vector space $\ell(I)$ has a much larger basis: $\dim(\ell(I)) > \aleph_0 = \dim(\ell_0(I))$.

3.1.2 Free Vector Space over a Set

The aim of the following construction is a vector space of all linear combinations of elements of a set S such that S is a basis.

Let S be any non-empty set and $\mathbb{K}$ a field. Consider a mapping $\varphi : S \to \mathbb{K}$. Its support is defined by

$$\operatorname{supp}(\varphi) := \{s \in S : \varphi(s) \neq 0\} \subset S,$$

where 0 is the zero element in $\mathbb{K}$. Requiring $\#\operatorname{supp}(\varphi) < \infty$ means that $\varphi = 0$ holds for almost all $s \in S$. This property defines the set

$$V := \{\varphi : S \to \mathbb{K} : \#\operatorname{supp}(\varphi) < \infty\}\,.$$

We introduce an addition in V. For $\varphi, \psi \in V$, the sum $\sigma := \varphi + \psi$ is the mapping $\sigma : S \to \mathbb{K}$ defined by their images $\sigma(s) := \varphi(s) + \psi(s)$ for all $s \in S$. Note that the support of σ is contained in $\operatorname{supp}(\varphi) \cup \operatorname{supp}(\psi)$, which again has finite cardinality, so that $\sigma \in V$. Since $\varphi(s) + \psi(s)$ is the addition in $\mathbb{K}$, the operation is commutative:

$\varphi + \psi = \psi + \varphi$. Obviously, the zero function $0_V \in V$ with $0_V(s) = 0 \in \mathbb{K}$ for all $s \in S$ satisfies $\varphi + 0_V = 0_V + \varphi = \varphi$. Furthermore, $\varphi^- : S \to \mathbb{K}$ defined by $\varphi^-(s) = -\varphi(s)$ is the inverse of φ, i.e., $\varphi^- + \varphi = \varphi + \varphi^- = 0_V$. Altogether, $(V, +)$ is a commutative group.

The scalar multiplication $\cdot : \mathbb{K} \times V \to V$ maps $\alpha \in \mathbb{K}$ and $\varphi \in V$ into the mapping $\psi := \alpha\varphi$ defined by $\psi(s) = \alpha\varphi(s)$ for all $s \in S$.

Thereby, all axioms in (3.1) are satisfied, so that V represents a vector space over the field $\mathbb{K}$.

The characteristic functions χ_s are of particular interest for an element $s \in S$:

$$\chi_s(t) = \begin{cases} 1 & \text{if } t = s \in S, \\ 0 & \text{if } t \in S\backslash\{s\}. \end{cases}$$

Every $\varphi \in V$ may be written as a linear combination of such χ_s:

$$\varphi = \sum_{s \in \operatorname{supp}(\varphi)} \varphi(s)\chi_s = \sum_{s \in S} \varphi(s)\chi_s.$$

Here, two different notations are used: the first sum is finite, while the second one is infinite, but contains only finitely many nonzero terms.

Note that any finite subset of $\{\chi_s : s \in S\}$ is linearly independent. Assuming $\sum_{s \in S_0} \alpha_s \chi_s = 0_V$ for some $S_0 \subset S$ with $\#S_0 < \infty$ and $\alpha_s \in \mathbb{K}$, the evaluation at $t \in S_0$ yields

$$\alpha_t = \left(\sum\nolimits_{s \in S_0} \alpha_s \chi_s\right)(t) = 0_V(t) = 0$$

proving linear independence. Vice versa, any finite linear combination

$$\sum_{s \in S_0} \alpha_s \chi_s \qquad \text{with } \alpha_s \in \mathbb{K},\ S_0 \subset S,\ \#S_0 < \infty \tag{3.3}$$

belongs to V.

Let $\Phi_\chi : \chi_s \mapsto s$ be the one-to-one correspondence between the sets $\{\chi_s : s \in S\}$ and S. We can extend Φ_χ to Φ_V defined on V such that $v = \sum_{s \in S_0} \alpha_s \chi_s$ from (3.3) is mapped onto the formal linear combination $\sum_{s \in S_0} \alpha_s s$ of elements of S.

The image $\mathcal{V}_{\text{free}}(S) := \Phi_V(V)$ is called the *free vector space* over the set S.

3.1.3 Quotient Vector Space

Let V be a vector space and $V_0 \subset V$ a subspace. V_0 defines an equivalence relation on V:

$$v \sim w \qquad \text{if and only if} \quad v - w \in V_0.$$

Any $v \in V$ can be associated with an equivalence class

$$c_v := \{w \in V : w \sim v\}. \tag{3.4}$$

Here, v is called a representative of the class c_v, which is also written as $v + V_0$. Because of the definition of an equivalence relation, the classes are either equal or disjoint. Their union equals V. The set $\{c_v : v \in V\}$ of all equivalence classes is denoted by the quotient

$$V \,/\, V_0.$$

One may define $c' + c'' := \{v' + v'' : v' \in c', v'' \in c''\}$ for two classes $c', c'' \in V/V_0$ and can check that the resulting set is again an equivalence class, i.e., an element in V/V_0. Similarly, one defines $\lambda \cdot c \in V/V_0$ for $c \in V/V_0$. Using the notation c_v for the classes generated by $v \in V$, one finds the relations $c_v + c_w = c_{v+w}$ and $\lambda \cdot c_v = c_{\lambda v}$. In particular, c_0 is the zero element. Altogether, V/V_0 is again a vector space over the same field, called the *quotient vector space*.

Exercise 3.2. Prove the identity $\dim(V) = \dim(V/V_0)\dim(V_0)$ and the particular cases $V/V = \{0\}$ and $V/\{0\} = V$.

A mapping $\varphi : V/V_0 \to X$ (X any set) may possibly be induced by a mapping $\Phi : V \to X$ via

$$\varphi(c_v) := \Phi(v) \qquad (c_v \text{ from } (3.4)). \tag{3.5}$$

Whether (3.5) is a well-defined formulation, hinges upon the following consistency condition.

Lemma 3.3. *(a) Let $\Phi : V \to X$ be a general mapping. Then (3.5) for all $v \in V$ defines a mapping $\varphi : V/V_0 \to X$ if and only if Φ is constant on each equivalence class, i.e., $v \sim w$ implies $\Phi(v) = \Phi(w)$.*
(b) If $\Phi : V \to X$ (X a vector space) is a linear mapping, the necessary and sufficient condition reads $\Phi(v) = 0$ for all $v \in V_0$.

3.1.4 Linear and Multilinear Mappings, Algebraic Dual

Let X, Y be two vector spaces. $\varphi : X \to Y$ is a *linear mapping*, if $\varphi(\lambda x' + x'') = \lambda\varphi(x') + \varphi(x'')$ for all $\lambda \in \mathbb{K}$, $x', x'' \in X$. The set of linear mappings φ is denoted by

$$L(X,Y) := \{\varphi : X \to Y \text{ is linear}\}. \tag{3.6}$$

Let X_j ($1 \le j \le d$) and Y be vector spaces. A mapping $\varphi : \times_{j=1}^{d} X_j \to Y$ is called *multilinear* (or d-linear), if φ is linear in all d arguments:

$$\begin{aligned}&\varphi(x_1, \ldots, x_{j-1}, x_j' + \lambda x_j'', x_{j+1}, \ldots, x_d)\\ &= \varphi(x_1, \ldots, x_{j-1}, x_j', x_{j+1}, \ldots, x_d) + \lambda\varphi(x_1, \ldots, x_{j-1}, x_j'', x_{j+1}, \ldots, x_d)\end{aligned}$$

for all $x_i \in X_i$, $x_j', x_j'' \in X_j$, $1 \le j \le d$, $\lambda \in \mathbb{K}$.

In the case of $d = 2$, the term '*bilinear mapping*' is used.

Definition 3.4. $\Phi \in L(X,X)$ is called a *projection*, if $\Phi^2 = \Phi$. It is called a projection *onto* Y, if $Y = \mathrm{range}(\Phi)$.

Note that no topology is defined and therefore no continuity is required.

Remark 3.5. Let $\{x_i : i \in B\}$ be a basis of X. $\varphi \in L(X,Y)$ is uniquely determined by the values $\varphi(x_i)$, $i \in B$.

In the particular case of $Y = \mathbb{K}$, linear mappings $\varphi : X \to \mathbb{K}$ are called *linear forms*. They are elements of the vector space

$$X' := L(X, \mathbb{K}), \tag{3.7}$$

which is called the *algebraic dual* of X.

Definition 3.6. Let $S := \{x_i : i \in B\} \subset X$ be a system of linearly independent vectors. A *dual system* $\{\varphi_i : i \in B\} \subset X'$ is defined by $\varphi_i(x_j) = \delta_{ij}$ for $i, j \in B$ (cf. (2.1)). If $\{x_i : i \in B\}$ is a basis, $\{\varphi_i : i \in B\}$ is called *dual basis*.

Remark 3.7. The dual basis allows us to determine the coefficients α_i in the basis representation

$$x = \sum_{i \in B} \alpha_i x_i \in X \qquad \text{by } \alpha_i = \varphi_i(x).$$

A multilinear [bilinear] map into $Y = \mathbb{K}$ is called '*multilinear form*' ['*bilinear form*'].

3.2 Tensor Product

3.2.1 Formal Definition

There are various ways to define the tensor product of two vector spaces. We follow the quotient space formulation (cf. [195]). Other constructions yield isomorphic objects (see comment after Proposition 3.22).

Given two vector spaces V and W over some field $\mathbb{K}$, we start with the free vector space $\mathcal{V}_{\text{free}}(S)$ over the pair set $S := V \times W$ as introduced in §3.1.2. Note that $\mathcal{V}_{\text{free}}(V \times W)$ does not make use of the vector space properties of V or W. We recall that elements of $\mathcal{V}_{\text{free}}(V \times W)$ are linear combinations of pairs from $V \times W$:

$$\sum_{i=1}^{m} \lambda_i \ (v_i, w_i) \qquad \text{for any} \begin{cases} (\lambda_i, v_i, w_i) \in \mathbb{K} \times V \times W, \\ m \in \mathbb{N}_0. \end{cases} \tag{3.8}$$

A particular subspace of $\mathcal{V}_{\text{free}}(V \times W)$ is

$$N := \mathrm{span}\left\{ \begin{array}{l} \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} \alpha_i \beta_j \ (v_i, w_j) - \left(\sum\limits_{i=1}^{m} \alpha_i v_i, \sum\limits_{j=1}^{n} \beta_j w_j \right) \\ \text{for } m, n \in \mathbb{N}, \quad \alpha_i, \beta_j \in \mathbb{K}, \quad v_i \in V, \quad w_j \in W \end{array} \right\}. \tag{3.9}$$

The *algebraic tensor space* is defined by the quotient vector space

$$V \otimes_a W := \mathcal{V}_{\text{free}}(V \times W) \,/\, N \tag{3.10}$$

(cf. §3.1.3). The equivalence class $c_{(v,w)} \in V \otimes_a W$ generated by a pair $(v, w) \in V \times W$ is denoted by

$$v \otimes w.$$

Note that the tensor symbol $\otimes$ is used for two different purposes:[1]

(i) In the tensor space notation, the symbol $\otimes$ connects vector spaces and may carry the suffix 'a' (meaning 'algebraic') or a norm symbol in the later case of Banach tensor spaces (cf. (3.12) and §4).

(ii) In $v \otimes w$, the quantities v, w, $v \otimes w$ are vectors, i.e., elements of the respective vector spaces V, W, $V \otimes_a W$.

As $\mathcal{V}_{\text{free}}(V \times W)$ is the set of linear combinations of (v_i, w_i) , the quotient space $\mathcal{V}_{\text{free}}(V \times W)/N$ consists of the linear combinations of $v_i \otimes w_i$:

$$V \otimes_a W = \operatorname{span}\left\{v \otimes w : \ v \in V, \ w \in W\right\}. \tag{3.11}$$

If a norm topology is given, the completion with respect to the given norm $\|\cdot\|$ yields the topological tensor space

$$V \otimes_{\|\cdot\|} W := \overline{V \otimes_a W}. \tag{3.12}$$

In §4 we discuss the properties of the tensor product for Banach spaces. This includes the Hilbert spaces, which are considered in §4.4.

Notation 3.8. (a) For finite dimensional vector spaces V and W, the algebraic tensor space $V \otimes_a W$ is already complete with respect to any norm and therefore coincides with the topological tensor space $V \otimes_{\|\cdot\|} W$. In this case, we omit the suffices and simply write $V \otimes W$.

(b) Furthermore, the notation $V \otimes W$ is used, when both choices $V \otimes_a W$ and $V \otimes_{\|\cdot\|} W$ are possible or if the distinction between $\otimes_a$ and $\otimes_{\|\cdot\|}$ is irrelevant.

(c) The suffices of $\otimes_a$ and $\otimes_{\|\cdot\|}$ will be moved to the left side, when indices appear at the right side as in ${}_a\bigotimes_{j=1}^{d} V_j$ and ${}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$.

Definition 3.9 (tensor space, tensor). (a) $V \otimes_a W$ (or $V \otimes_{\|\cdot\|} W$) is again a vector space, which is now called '*tensor space*'.

(b) The explicit term '*algebraic tensor space*' emphasises that $V \otimes_a W$, and not $V \otimes_{\|\cdot\|} W$, is meant.

(c) Elements of $V \otimes_a W$ or $V \otimes_{\|\cdot\|} W$ are called *tensors*, in particular, $\mathbf{x} \in V \otimes_a W$ is an *algebraic tensor*, while $\mathbf{x} \in V \otimes_{\|\cdot\|} W$ is a *topological tensor*.

(d) Any product $v \otimes w$ $(v \in V, \ w \in W)$ is called '*elementary tensor*'.

[1] Similarly, the sum $v + w$ of vectors and the sum $V + W := \operatorname{span}\{v + w : v \in V, w \in W\}$ of vector spaces use the same symbol.

3.2.2 Characteristic Properties

Lemma 3.10. *The characteristic algebraic properties of the tensor space $V \otimes_a W$ is the bilinearity:*

$$\begin{array}{ll} (\lambda v) \otimes w = v \otimes (\lambda w) = \lambda \cdot (v \otimes w) & \textit{for } \lambda \in \mathbb{K},\ v \in V,\ w \in W, \\ (v' + v'') \otimes w = v' \otimes w + v'' \otimes w & \textit{for } v', v'' \in V,\ w \in W, \\ v \otimes (w' + w'') = v \otimes w' + v \otimes w'' & \textit{for } v \in V,\ w', w'' \in W, \\ 0 \otimes w = v \otimes 0 = 0 & \textit{for } v \in V,\ w \in W. \end{array} \tag{3.13}$$

Proof. The first equality in (3.13) follows from $\lambda (v, w) - (\lambda v, w) \in N$, i.e., $\lambda \cdot (v \otimes w) - \lambda v \otimes w = 0$ in the quotient space. The other identities are derived similarly. □

Here, the standard notational convention holds: the multiplication $\otimes$ has priority over the addition $+$, i.e., $a \otimes b + c \otimes d$ means $(a \otimes b) + (c \otimes d)$. The multiplication by a scalar needs no bracket, since the interpretation of $\lambda v \otimes w$ by $(\lambda v) \otimes w$ or $\lambda \cdot (v \otimes w)$ does not change the result (see first identity in (3.13)). In the last line[2] of (3.13), the three zeros belong to the different spaces V, W, and $V \otimes_a W$.

The following statements also hold for infinite dimensional spaces. Note that in this case the dimensions have to be understood as set theoretical cardinal numbers.

Lemma 3.11. *(a) Let $\{v_i : i \in B_V\}$ be a basis of V and $\{w_j : j \in B_W\}$ a basis of W. Then*

$$\mathfrak{B} := \{v_i \otimes w_j : i \in B_V,\ j \in B_W\} \tag{3.14}$$

is a basis of $V \otimes_a W$.
(b) $\dim (V \otimes_a W) = \dim(V) \cdot \dim(W)$.

Proof. Assume $\sum_{i,j} a_{ij} v_i \otimes w_j = 0$. For the linear independence of all $v_i \otimes w_j \in \mathfrak{B}$ we have to show $a_{ij} = 0$. The properties (3.13) show

$$\sum_{i \in I} v_i \otimes w_i' = 0 \qquad \text{for } w_i' := \sum_{j \in J} a_{ij} w_j. \tag{3.15}$$

Let $\varphi_i \in V'$ (cf. §3.1.4) be the linear form on V with $\varphi_i(v_j) = \delta_{ij}$ (cf. (2.1)). Define $\Phi_i : V \otimes_a W \to W$ by $\Phi_i (v \otimes w) = \varphi_i(v) w$. Application of Φ_i to (3.15) yields $w_i' = 0$. Since $\{w_j : j \in B_W\}$ is a basis, $a_{ij} = 0$ follows for all j. As i is chosen arbitrarily, we have shown $a_{ij} = 0$ for all coefficients in $\sum_{i,j} a_{ij} v_i \otimes w_j = 0$. Hence, $\mathfrak{B}$ is a system of linearly independent vectors.

By definition, a general tensor $\mathbf{x} \in V \otimes_a W$ has the form $\mathbf{x} = \sum_\nu v^{(\nu)} \otimes w^{(\nu)}$. Each $v^{(\nu)}$ can be expressed by the basis vectors: $v^{(\nu)} = \sum_i \alpha_i^{(\nu)} v_i$, and similarly, $w^{(\nu)} = \sum_j \beta_j^{(\nu)} w_j$. Note that all sums have finitely many terms. The resulting sum

$$\mathbf{x} = \sum_\nu \left(\sum_i \alpha_i^{(\nu)} v_i \right) \otimes \left(\sum_j \beta_j^{(\nu)} w_j \right) \underset{(3.13)}{=} \sum_{i,j} \left(\sum_\nu \alpha_i^{(\nu)} \beta_j^{(\nu)} \right) v_i \otimes w_j$$

[2] The last line can be derived from the first one by setting $\lambda = 0$.

is again finite and shows that $\operatorname{span}\{\mathfrak{B}\} = V \otimes_a W$, i.e., $\mathfrak{B}$ is a basis.

Since $\#\mathfrak{B} = \#B_V \cdot \#B_W$, we obtain the dimension identity of Part (b). □

The last two statements characterise the tensor space structure.

Proposition 3.12. *Let V, W, and T be vector spaces over the field $\mathbb{K}$. A product $\otimes : V \times W \to T$ is a tensor product and T a tensor space, i.e., it is isomorphic to $V \otimes_a W$, if the following properties hold:*

(i) *span property:* $T = \operatorname{span}\{v \otimes w : v \in V, w \in W\}$*;*

(ii) *bilinearity (3.13);*

(iii) *linearly independent vectors $\{v_i : i \in B_V\} \subset V$ and $\{w_j : j \in B_W\} \subset W$ lead to independent vectors $\{v_i \otimes w_j : i \in B_V, j \in B_W\}$ in T.*

Proof. Properties (i)-(iii) imply that $\mathfrak{B}$ from (3.14) is again a basis. □

Lemma 3.13. *For any tensor $\mathbf{x} \in V \otimes_a W$ there is an $r \in \mathbb{N}_0$ and a representation*

$$\mathbf{x} = \sum_{i=1}^{r} v_i \otimes w_i \tag{3.16}$$

with linearly independent vectors $\{v_i : 1 \le i \le r\} \subset V$ and $\{w_i : 1 \le i \le r\} \subset W$.

Proof. Take any representation $\mathbf{x} = \sum_{i=1}^{n} v_i \otimes w_i$. If, e.g., the system of vectors $\{v_i : 1 \le i \le n\}$ is not linearly independent, one v_i can be expressed by the others. Without loss of generality assume

$$v_n = \sum_{i=1}^{n-1} \alpha_i v_i \, .$$

Then

$$v_n \otimes w_n = \left(\sum_{i=1}^{n-1} \alpha_i v_i \right) \otimes w_n = \sum_{i=1}^{n-1} v_i \otimes (\alpha_i w_n)$$

shows that $\mathbf{x}$ possesses a representation with only $n-1$ terms:

$$\mathbf{x} = \left(\sum_{i=1}^{n-1} v_i \otimes w_i \right) + v_n \otimes w_n = \sum_{i=1}^{n-1} v_i \otimes w_i' \qquad \text{with} \quad w_i' := w_i + \alpha_i w_n .$$

Since each reduction step decreases the number of terms by one, this process terminates at a certain number r of terms, i.e., we obtain a representation with r linearly independent v_i and w_i. □

The number r appearing in Lemma 3.13 will be called the *rank* of the tensor $\mathbf{x}$ (cf. §3.2.6.2). This is in accordance with the usual matrix rank as seen in §3.2.3.

3.2.3 Isomorphism to Matrices for $\mathbf{d} = \mathbf{2}$

In the following, the index sets I and J are assumed to be finite. In the same way as $\mathbb{K}^I$ is the model vector space for vector spaces of dimension $\#I$, we obtain the model tensor space $\mathbb{K}^I \otimes \mathbb{K}^J \cong \mathbb{K}^{I\times J}$.

Proposition 3.14. *Let $\{v_i : i \in I\}$ be a basis of V and $\{w_j : j \in J\}$ a basis of W, where $\dim(V) < \infty$ and $\dim(W) < \infty$. Let $\Phi : \mathbb{K}^I \to V$ and $\Psi : \mathbb{K}^J \to W$ denote the isomorphisms $\Phi : (\alpha_i)_{i\in I} \mapsto \sum_{i\in I} \alpha_i v_i$ and $\Psi : (\beta_j)_{j\in J} \mapsto \sum_{j\in J} \beta_j w_j$.*
(a) Then the corresponding canonical isomorphism of the tensor spaces is given by

$$\Xi : \mathbb{K}^I \otimes \mathbb{K}^J \to V \otimes W \qquad \textit{with } (\alpha_i)_{i\in I} \otimes (\beta_j)_{j\in J} \mapsto \sum_{i\in I}\sum_{j\in J} \alpha_i\beta_j v_i \otimes w_j.$$

(b) Together with the identification of $\mathbb{K}^I \otimes \mathbb{K}^J$ with the matrix space $\mathbb{K}^{I\times J}$ (see Remark 1.3), we obtain an isomorphism between matrices from $\mathbb{K}^{I\times J}$ and tensors from $V \otimes W$:

$$\Xi : \mathbb{K}^{I\times J} \to V \otimes W \qquad \textit{with } (a_{ij})_{i\in I, j\in J} \mapsto \sum_{i\in I}\sum_{j\in J} a_{ij}\, v_i \otimes w_j.$$

Although the isomorphism looks identical to the usual isomorphism between matrices from $\mathbb{K}^{I\times J}$ and linear mappings $W \to V$, there is a small difference which will be discussed in Proposition 3.57.

Remark 3.15 (basis transformation). If we change the bases $\{v_i : i \in I\}$ and $\{w_j : j \in J\}$ from Proposition 3.14 by transformations S and T:

$$v_i = \sum_{n\in I} S_{ni}\hat{v}_n \quad \text{and} \quad w_j = \sum_{m\in J} T_{mj}\hat{w}_m,$$

then

$$\sum_{i\in I}\sum_{j\in J} a_{ij} v_i \otimes w_j = \sum_{n\in I}\sum_{m\in J} \hat{a}_{nm}\hat{v}_n \otimes \hat{w}_m$$

shows that $A = (a_{ij})$ and $\hat{A} = (\hat{a}_{ij})$ are related by

$$\hat{A} = S\,A\,T^{\mathsf{T}}.$$

On the side of the tensors, this transformation takes the form

$$(S \otimes T)\,(a \otimes b) \qquad \text{with } a := (\alpha_i)_{i\in I} \text{ and } b := (\beta_j)_{j\in J},$$

where $S \otimes T$ is the Kronecker product.

Corollary 3.16. If $V = \mathbb{K}^I$ and $W = \mathbb{K}^J$, the isomorphism of the tensor space $\mathbb{K}^I \otimes \mathbb{K}^J$ and the matrix space $\mathbb{K}^{I\times J}$ is even more direct, since it does not need a choice of bases.

Remark 3.17. Suppose $\dim(V) = 1$. Then the vector space V may be identified with the field $\mathbb{K}$. $V \otimes_a W$ is isomorphic to $\mathbb{K} \otimes_a W$ and to W. In the latter case, one identifies $\lambda \otimes w$ ($\lambda \in \mathbb{K}$, $w \in W$) with λw.

Lemma 3.18 (reduced singular value decomposition). *Let* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$*, and suppose* $\dim(V) < \infty$ *and* $\dim(W) < \infty$*. Then for any* $\mathbf{x} \in V \otimes W$ *there is a number* $r \leq \min\{\#I, \#J\}$ *and two families* $(x_i)_{i=1,\dots,r}$ *and* $(y_i)_{i=1,\dots,r}$ *of linearly independent vectors such that*

$$\mathbf{x} = \sum_{i=1}^{r} \sigma_i \, x_i \otimes y_i$$

with singular values $\sigma_1 \geq \ldots \geq \sigma_r > 0$.

Proof. The isomorphism $\varXi : \mathbb{K}^{I \times J} \to V \otimes W$ defines the matrix $A := \varXi^{-1}\mathbf{x}$, for which the reduced singular value decomposition $A = \sum_{i=1}^{r} \sigma_i \; a_i b_i^{\mathsf{T}}$ can be determined (cf. (2.21)). Note that $\varXi(a_i b_i^{\mathsf{T}}) = a_i \otimes b_i$ (cf. (1.3)). Backtransformation yields $\mathbf{x} = \varXi A = \sum_{i=1}^{r} \sigma_i \; \varXi(a_i \otimes b_i) = \sum_{i=1}^{r} \sigma_i \; \varPhi(a_i) \otimes \varPsi(b_i)$. The statement follows by setting $x_i := \varPhi(a_i)$ and $y_i := \varPsi(b_i)$. Note that linearly independent a_i yield linearly independent $\varPhi(a_i)$. □

We remark that the vectors x_i (as well as y_i) are not orthonormal, since such properties are not (yet) defined for V (and W). Lemma 3.18 yields a second proof of Lemma 3.13, but restricted to $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Remark 3.19. The tensor spaces $V \otimes_a W$ and $W \otimes_a V$ are isomorphic vector spaces via the (bijective) *transposition*

$$\begin{aligned} T: \quad & V \otimes_a W \to W \otimes_a V, \\ & \mathbf{x} = v \otimes w \mapsto \mathbf{x}^{\mathsf{T}} = w \otimes v. \end{aligned}$$

If $\mathbf{x} \in V \otimes W$ is related to a matrix $M = \varXi^{-1}(\mathbf{x})$ (cf. Proposition 3.14b), then $\mathbf{x}^{\mathsf{T}} \in W \otimes V$ is related to the transposed matrix M^{T}.

3.2.4 Tensors of Order $d \geq 3$

In principle, one can extend the construction from §3.2.1 to the case of more than two factors. However, this is not necessary as the next lemma shows.

Lemma 3.20. *(a) The tensor product is associative:*

$$U \otimes_a (V \otimes_a W) = (U \otimes_a V) \otimes_a W,$$

i.e., they are isomorphic vector spaces as detailed in the proof. We identify both notations and use the neutral notation $U \otimes_a V \otimes_a W$ *instead.*
(b) If U, V, W *are finite dimensional with* $\dim(U) = n_1$*,* $\dim(V) = n_2$*, and* $\dim(W) = n_3$*, the isomorphic model tensor space is* $\mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \mathbb{K}^{n_3}$.

Proof. Let u_i $(i \in B_U)$, v_j $(j \in B_V)$, w_k $(k \in B_W)$ be bases of U, V, W. As seen in Lemma 3.11, $V \otimes_a W$ has the basis $v_j \otimes w_k$ $\big((j,k) \in B_V \times B_W\big)$, while

$U \otimes_a (V \otimes_a W)$ has the basis $u_i \otimes (v_j \otimes w_k)$ with $(i,(j,k)) \in B_U \times (B_V \times B_W)$. Similarly, $(U \otimes_a V) \otimes_a W$ has the basis

$$(u_i \otimes v_j) \otimes w_k \qquad \text{with} \qquad ((i,j),k) \in (B_U \times B_V) \times B_W.$$

By the obvious bijection between $B_U \times (B_V \times B_W)$ and $(B_U \times B_V) \times B_W$, the isomorphism $U \otimes_a (V \otimes_a W) \cong (U \otimes_a V) \otimes_a W$ follows. This proves Part (a). For Part (b) see Remark 3.29. □

Repeating the product construction $(d-1)$-times, we get the generalisation of the previous results to the algebraic tensor product ${}_a\bigotimes_{j=1}^d V_j$ (cf. Notation 3.8).

Proposition 3.21. *Let V_j $(1 \le j \le d,\ d \ge 2)$ be vector spaces over $\mathbb{K}$.*

(a) The algebraic tensor space[3]

$$\mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j$$

is independent of the order in which the pairwise construction (3.10) is performed (more precisely, the resulting spaces are isomorphic and can be identified).

(b) T is the algebraic tensor space $\mathbf{V}$ if the following properties hold:

(i) *span property:* $T = \operatorname{span}\left\{\bigotimes_{j=1}^d v^{(j)} : v^{(j)} \in V_j\right\}$;

(ii) *multilinearity, i.e., for all $\lambda \in \mathbb{K}$, $v^{(j)}, w^{(j)} \in V_j$, and $j \in \{1,\dots,d\}$:*
$v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes \left(\lambda v^{(j)} + w^{(j)}\right) \otimes \ldots \otimes v^{(d)} =$
$\lambda v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes v^{(j)} \otimes \ldots \otimes v^{(d)} + v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes w^{(j)} \otimes \ldots \otimes v^{(d)}$;

(iii) *linearly independent vectors $\{v_i^{(j)} : i \in B_j\} \subset V_j$ $(1 \le j \le d)$ lead to linearly independent vectors $\{\bigotimes_{j=1}^d v_{i_j}^{(j)} : i_j \in B_j\}$ in T.*

(c) The dimension is given by

$$\dim(\mathbf{V}) = \prod_{j=1}^{d} \dim(V_j). \tag{3.17}$$

If $\{v_i^{(j)} : i \in B_j\}$ are bases of V_j $(1 \le j \le d)$, then $\{\mathbf{v_i} : \mathbf{i} \in \mathbf{B}\}$ is a basis of $\mathbf{V}$, where

$$\mathbf{v_i} = \bigotimes_{j=1}^{d} v_{i_j}^{(j)} \qquad \textit{for } \mathbf{i} = (i_1,\dots,i_d) \in \mathbf{B} := B_1 \times \ldots \times B_d.$$

An alternative definition of ${}_a\bigotimes_{j=1}^d V_j$ follows the construction from §3.2.1 with pairs (v,w) replaced by d-tuples and an appropriately defined subspace N. The following expression of the multilinear mapping φ by the linear mapping Φ is also called 'linearisation' of φ.

[3] The product $\bigotimes_{j=1}^d V_j$ is to be formed in the order of the indices j, i.e., $V_1 \otimes V_2 \otimes \ldots$ If we write $\bigotimes_{j\in K} V_j$ for an ordered index set K, the ordering of K determines the order of the factors.

Proposition 3.22 (universality of the tensor product). *Let V_j $(1 \le j \le d)$ and U be vector spaces over $\mathbb{K}$. Then, for any multilinear mapping $\varphi : V_1 \times \ldots \times V_d \to U$, i.e.,*

$$\begin{aligned} &\varphi(v^{(1)}, \ldots, \lambda v^{(j)} + w^{(j)}, \ldots, v^{(d)}) \\ &= \lambda\, \varphi(v^{(1)}, \ldots, v^{(j)}, \ldots, v^{(d)}) + \varphi(v^{(1)}, \ldots, w^{(j)}, \ldots, v^{(d)}) \qquad (3.18\text{a}) \\ &\textit{for all } v^{(j)}, w^{(j)} \in V_j,\ \lambda \in \mathbb{K},\ 1 \le j \le d, \end{aligned}$$

there is a unique linear mapping $\Phi : {}_a\bigotimes_{j=1}^d V_j \to U$ such that

$$\varphi(v^{(1)}, v^{(2)}, \ldots, v^{(d)}) = \Phi\big(v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes v^{(d)}\big) \qquad (3.18\text{b})$$

for all $v^{(j)} \in V_j$, $1 \le j \le d$.

Proof. Let $\{v_i^{(j)} : i \in B_j\}$ be a basis of V_j for $1 \le j \le d$. As stated in Proposition 3.21c, $\{\mathbf{v_i} : \mathbf{i} \in \mathbf{B}\}$ is a basis of $\mathbf{V} := {}_a\bigotimes_{j=1}^d V_j$. Define $\Phi(\mathbf{v_i}) := \varphi(v_{i_1}^{(1)}, \ldots, v_{i_d}^{(d)})$. This determines $\Phi : \mathbf{V} \to U$ uniquely (cf. Remark 3.5). Analogously, the multilinear mapping φ is uniquely determined by $\varphi(v_{i_1}^{(1)}, \ldots, v_{i_d}^{(d)})$. The multilinear nature of φ and $\mathbf{V}$ yields $\varphi\big(v^{(1)}, v^{(2)}, \ldots, v^{(d)}\big) = \Phi\big(v^{(1)} \otimes \ldots \otimes v^{(d)}\big)$ for all $v^{(j)} \in V_j$. □

The statement of Proposition 3.22 may also be used as an equivalent definition of ${}_a\bigotimes_{j=1}^d V_j$ (cf. Greub [76, Chap. I, §2]). The content of Proposition 3.22 is visualised by the commutative diagram to the right.

$$\begin{tikzcd} V_1 \times \ldots \times V_d \arrow[r, "\varphi"'] \arrow[d, "\otimes"'] & U \\ {}_a\bigotimes_{j=1}^d V_j \arrow[ur, "\Phi"] & \end{tikzcd}$$

Notation 3.23. If all $V_j = V$ are identical vector spaces, the notation $\bigotimes_{j=1}^d V_j$ is simplified by $\otimes^d V$. For a vector $v \in V$, we set $\otimes^d v := \bigotimes_{j=1}^d v$.

To avoid trivial situations, we introduce the notation of non-degenerate tensor spaces.

Definition 3.24. A tensor space $\mathbf{V} = {}_a\bigotimes_{j=1}^d V_j$ is called *non-degenerate*, if $d > 0$ and $\dim(V_j) \ge 2$ for all $1 \le j \le d$. Otherwise, $\mathbf{V}$ is called *degenerate*.

This definition is justified by the following remarks. If $\dim(V_j) = 0$ for one j, also $\mathbf{V} = \{0\}$ is the trivial vector space. If $d = 0$, the empty product is defined by $\mathbf{V} = \mathbb{K}$, which is also a trivial case. In the case of $\dim(V_j) = 1$ for some j, the formulation by ${}_a\bigotimes_{j=1}^d V_j$ can be reduced (see next remark).

Remark 3.25. (a) If $\dim(V_k) = 1$ for some k, the isomorphism $\mathbf{V} = {}_a\bigotimes_{j=1}^d V_j \cong {}_a\bigotimes_{j \in \{1,\ldots,d\}\backslash\{k\}} V_j$ allows us to omit the factor V_k.
(b) After eliminating all factors V_j with $\dim(V_j) = 1$ and renaming the remaining vector spaces, we obtain $\mathbf{V} \cong \mathbf{V}_{\mathrm{red}} = {}_a\bigotimes_{j=1}^{d_{\mathrm{red}}} V_j$. If still $d_{\mathrm{red}} > 0$, the representation is non-degenerate. Otherwise the tensor space is degenerate because of $d_{\mathrm{red}} = 0$.

As an illustration we may consider a matrix space (i.e., a tensor space of order $d=2$). If $\dim(V_2)=1$, the matrices consist of only one column. Hence, they may be considered as vectors (tensor space with $d=1$). If even $\dim(V_1)=\dim(V_2)=1$, the 1×1-matrices may be identified with scalars from the field $\mathbb{K}$.

Finally, we mention an isomorphism between the space of tuples of tensors and an extended tensor space.

Lemma 3.26. *Let* $\mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$ *be a tensor space over the field* $\mathbb{K}$ *and* $m\in\mathbb{N}$. *The vector space of* m*-tuples* $(\mathbf{v}_1,\ldots,\mathbf{v}_m)$ *with* $\mathbf{v}_i\in\mathbf{V}$ *is denoted by* $\mathbf{V}^m$. *Then the following vector space isomorphism is valid:*

$$\mathbf{V}^m \cong {}_a\bigotimes_{j=1}^{d+1} V_j = \mathbf{V}\otimes V_{d+1} \qquad \text{with } V_{d+1} := \mathbb{K}^m. \tag{3.19}$$

Proof. $(\mathbf{v}_1,\ldots,\mathbf{v}_m)\in\mathbf{V}^m$ corresponds to $\sum_{i=1}^m \mathbf{v}_i\otimes e^{(i)}$, where $e^{(i)}\in\mathbb{K}^m$ is the i-th unit vector. The opposite direction of the isomorphism is described by $\mathbf{v}\otimes x^{(d+1)}\cong(x_1\mathbf{v},\, x_2\mathbf{v},\,\ldots,x_m\mathbf{v})$ with $x^{(d+1)}=(x_i)_{i=1,\ldots,m}\in\mathbb{K}^m$. □

3.2.5 Different Types of Isomorphisms

For algebraic objects it is common to identify isomorphic ones. All characteristic properties of an algebraic structure should be invariant under an isomorphism. The question is what algebraic structures are meant. All previous isomorphisms were *vector space isomorphisms*. As it is well-known, two vector spaces are isomorphic if and only if the dimensions coincide (note that in the infinite dimensional case the cardinalities of the bases are decisive).

Any tensor space is a vector space, but not any vector space isomorphism preserves the tensor structure. A part of the tensor structure is the d-tuple of vector spaces $(V_1,\ldots,V_d)$ together with the dimensions of V_j (cf. Proposition 3.22). In fact, each space V_j can be regained from $\mathbf{V}:={}_a\bigotimes_{j=1}^{d}V_j$ as the range of the mapping $\boldsymbol{\Phi}=\bigotimes_{k=1}^{d}\phi_k$ (cf. (1.4a)) with $\phi_j=id$, while $0\neq\phi_k\in V_k'$ for $k\neq j$. Therefore, a *tensor space isomorphism* must satisfy that $\mathbf{V}={}_a\bigotimes_{j=1}^{d}V_j\cong\mathbf{W}$ implies that $\mathbf{W}={}_a\bigotimes_{j=1}^{d}W_j$ holds with isomorphic vector spaces $W_j\cong V_j$ for all $1\leq j\leq d$. In particular, the order d of the tensor spaces must coincide. This requirement is equivalent to the following definition.

Definition 3.27 (tensor space isomorphism). A *tensor space isomorphism*

$$\boldsymbol{\Phi}:\mathbf{V}:={}_a\bigotimes_{j=1}^{d}V_j\to\mathbf{W}={}_a\bigotimes_{j=1}^{d}W_j$$

is any bijection of the form $\boldsymbol{\Phi}=\bigotimes_{j=1}^{d}\phi_j$ (cf. (3.34b)), where $\phi_j:V_j\to W_j$ for $1\leq j\leq d$ are vector space isomorphisms.

For instance, $\mathbb{K}^2 \otimes \mathbb{K}^8$, $\mathbb{K}^4 \otimes \mathbb{K}^4$, and $\mathbb{K}^2 \otimes \mathbb{K}^2 \otimes \mathbb{K}^4$ are isomorphic vector spaces (since all have dimension 16), but their tensor structures are not identical. For the definition of $U \otimes_a V \otimes_a W$ we use in Lemma 3.20 that $U \otimes_a (V \otimes_a W) \cong (U \otimes_a V) \otimes_a W$. Note that these three spaces are isomorphic only in the sense of vector spaces, whereas their tensor structures are different. In particular, the latter spaces are tensor spaces of order two, while $U \otimes_a V \otimes_a W$ is of order three. We see the difference, when we consider the elementary tensors as in the next example.

Example 3.28. Let both $\{v_1, v_2\} \subset V$ and $\{w_1, w_2\} \subset W$ be linearly independent. Then $u \otimes (v_1 \otimes w_1 + v_2 \otimes w_2)$ is an elementary tensor in $U \otimes_a (V \otimes_a W)$ (since $u \in U$ and $v_1 \otimes w_1 + v_2 \otimes w_2 \in V \otimes_a W$), but it is not an elementary tensor of $U \otimes_a V \otimes_a W$.

Remark 3.29. In the finite dimensional case of $n_j := \dim(V_j) < \infty$, the isomorphic model tensor space is $\mathbf{W} := \bigotimes_{j=1}^d \mathbb{K}^{n_j}$. Choose some bases $\{b_\nu^{(j)} : 1 \le \nu \le n_j\}$ of V_j. Any $\mathbf{v} \in \mathbf{V} := \bigotimes_{j=1}^d V_j$ has a unique representation of the form

$$\mathbf{v} = \sum_{i_1 \cdots i_d} \mathbf{a_i}\, b_{i_1}^{(1)} \otimes \ldots \otimes b_{i_d}^{(d)}. \tag{3.20}$$

The coefficients $\mathbf{a_i}$ define the 'coefficient tensor' $\mathbf{a} \in \mathbf{W}$. The mapping $\Phi : \mathbf{V} \to \mathbf{W}$ by $\Phi(\mathbf{v}) = \mathbf{a}$ is a tensor space isomorphism.

The isomorphism sign $\cong$ is an equivalence relation in the set of the respective structure. Let $\cong_{\mathsf{vec}}$ denote the vector space isomorphism, while $\cong_{\mathsf{ten}}$ is the tensor space isomorphism. Then $\cong_{\mathsf{ten}}$ is the finer equivalence relation, since $\mathbf{V} \cong_{\mathsf{ten}} \mathbf{W}$ implies $\mathbf{V} \cong_{\mathsf{vec}} \mathbf{W}$, but not vice versa. There are further equivalence relations $\cong$, which are between $\cong_{\mathsf{ten}}$ and $\cong_{\mathsf{vec}}$, i.e., $\mathbf{V} \cong_{\mathsf{ten}} \mathbf{W} \Rightarrow \mathbf{V} \cong \mathbf{W} \Rightarrow \mathbf{V} \cong_{\mathsf{vec}} \mathbf{W}$. We give three examples.

1) We may not insist upon a strict ordering of the vector spaces V_j. Let $\pi : \{1, \ldots, d\} \to \{1, \ldots, d\}$ be a permutation. Then $\mathbf{V} := V_1 \otimes V_2 \otimes \ldots \otimes V_d$ and $\mathbf{V}^\pi := V_{\pi(1)} \otimes V_{\pi(2)} \otimes \ldots \otimes V_{\pi(d)}$ are considered as isomorphic. In a second step, each V_j may be replaced by an isomorphic vector space W_j.

2) In Remark 3.25 we have omitted vector spaces $V_j \cong \mathbb{K}$ of dimension one. This leads to an isomorphism $V_1 \otimes \mathbb{K} \cong V_1$ or $\mathbb{K} \otimes V_2 \cong V_2$, where V_j may be further replaced by an isomorphic vector space W_j. Note that the order of the tensor spaces is changed, but the nontrivial vector spaces are still pairwise isomorphic.

3) The third example will be of broader importance. Fix some $k \in \{1, \ldots, d\}$. Using the isomorphism from Item 1, we may state that

$$V_1 \otimes \ldots \otimes V_k \otimes \ldots \otimes V_d \;\cong\; V_k \otimes V_1 \otimes \ldots \otimes V_{k-1} \otimes V_{k+1} \otimes \ldots \otimes V_d$$

using the permutation $1 \leftrightarrow k$. Next, we make use of the argument of Lemma 3.20: associativity allows the vector space isomorphism

$$V_k \otimes V_1 \otimes \ldots \otimes V_{k-1} \otimes V_{k+1} \otimes \ldots \otimes V_d \cong V_k \otimes \Big[V_1 \otimes \ldots \otimes V_{k-1} \otimes V_{k+1} \otimes \ldots \otimes V_d \Big].$$

The tensor space in parentheses will be abbreviated by

$$\mathbf{V}_{[k]} := {}_a\bigotimes_{j\neq k} V_j \ , \tag{3.21a}$$

where

$$\bigotimes_{j\neq k} \quad \text{means} \quad \bigotimes_{j\in\{1,\ldots,d\}\backslash\{k\}} \ . \tag{3.21b}$$

Altogether, we have the vector space isomorphism

$$\mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j \cong V_k \otimes \mathbf{V}_{[k]}. \tag{3.21c}$$

We notice that $\mathbf{V}$ is a tensor space of order d, whereas $V_k \otimes \mathbf{V}_{[k]}$ has order 2. Nevertheless, a part of the tensor structure (the space V_k) is preserved. The importance of the isomorphism (3.21c) is already obvious from Lemma 3.20, since this allows a reduction to tensor spaces of order two. We shall employ (3.21c) to introduce the matricisation in §5.2.

In order to simplify the notation, we shall often replace the $\cong$ sign by equality: $\mathbf{V} = V_k \otimes \mathbf{V}_{[k]}$. This allows to write $\mathbf{v} \in \mathbf{V}$ as well as $\mathbf{v} \in V_k \otimes \mathbf{V}_{[k]}$, whereas the more exact notation is $\mathbf{v} \in \mathbf{V}$ and $\tilde{\mathbf{v}} = \Phi(\mathbf{v}) \in V_k \otimes \mathbf{V}_{[k]}$ with the vector space isomorphism $\Phi : \mathbf{V} \to V_k \otimes \mathbf{V}_{[k]}$. In fact, we shall see in Remark 3.33 that $\mathbf{v}$ and $\tilde{\mathbf{v}}$ have different properties. For elementary tensors of $\mathbf{V}$ we write

$$\mathbf{v} = \bigotimes_{j=1}^{d} v^{(j)} = v^{(k)} \otimes \mathbf{v}^{[k]}, \quad \text{where } \mathbf{v}^{[k]} := \bigotimes_{j\neq k} v^{(j)} \in \mathbf{V}_{[k]}. \tag{3.21d}$$

For a general (algebraic) tensor, the corresponding notation is

$$\mathbf{v} = \sum_i \bigotimes_{j=1}^{d} v_i^{(j)} = \sum_i v_i^{(k)} \otimes \mathbf{v}_i^{[k]} \quad \text{with } \mathbf{v}_i^{[k]} := \bigotimes_{j\neq k} v_i^{(j)} \in \mathbf{V}_{[k]}.$$

3.2.6 $\mathcal{R}_r$ and Tensor Rank

3.2.6.1 The set $\mathcal{R}_r$

Let V_j $(1 \le j \le d)$ be vector spaces generating $\mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j$. All linear combinations of r elementary tensors are contained in

$$\mathcal{R}_r := \mathcal{R}_r(\mathbf{V}) := \left\{ \sum_{\nu=1}^{r} v_\nu^{(1)} \otimes \ldots \otimes v_\nu^{(d)} : v_\nu^{(j)} \in V_j \right\} \qquad (r \in \mathbb{N}_0). \tag{3.22}$$

Deliberately, we use the same symbol $\mathcal{R}_r$ as in (2.6) as justified by Remark 3.35a.

Remark 3.30. $\mathbf{V} = \bigcup_{r\in\mathbb{N}_0} \mathcal{R}_r$ holds for the algebraic tensor space $\mathbf{V}$.

Proof. By definition (3.11), $\mathbf{v} \in \mathbf{V}$ is a finite linear combination of elementary tensors, i.e., $\mathbf{v} = \sum_{\nu=1}^{s} \alpha_\nu \mathbf{e}_\nu$ for some $s \in \mathbb{N}_0$ and suitable elementary tensors $\mathbf{e}_\nu$. The factor α_ν can be absorbed by the elementary tensor: $\alpha_\nu \mathbf{e}_\nu =: v_\nu^{(1)} \otimes \ldots \otimes v_\nu^{(d)}$. Hence, $\mathbf{v} \in \mathcal{R}_s \subset \bigcup_{r\in\mathbb{N}_0} \mathcal{R}_r$ proves $\mathbf{V} \subset \bigcup_{r\in\mathbb{N}_0} \mathcal{R}_r \subset \mathbf{V}$. □

Remark 3.31. The sets $\mathcal{R}_r$, which in general are not subspaces, are nested:

$$\{0\} = \mathcal{R}_0 \subset \mathcal{R}_1 \subset \ldots \subset \mathcal{R}_{r-1} \subset \mathcal{R}_r \subset \ldots \subset \mathbf{V} \qquad \text{for all } r \in \mathbb{N}, \tag{3.23a}$$

and satisfy the additive property

$$\mathcal{R}_r + \mathcal{R}_s = \mathcal{R}_{r+s}. \tag{3.23b}$$

Proof. Note that $\mathcal{R}_0 = \{0\}$ (empty sum convention). Since we may choose $v_r^{(j)} = 0$ in $\sum_{\nu=1}^{r} v_\nu^{(1)} \otimes \ldots \otimes v_\nu^{(d)}$, all sums of $r-1$ terms are included in $\mathcal{R}_r$. □

3.2.6.2 Tensor Rank

A non-vanishing elementary tensor $v_\nu^{(1)} \otimes \ldots \otimes v_\nu^{(d)}$ in (3.22) becomes a rank-1 matrix in the case of $d = 2$ (cf. §2.2). Remark 2.1 states that the rank r is the smallest integer such that a representation $M = \sum_{\nu=1}^{r} v_\nu^{(1)} \otimes v_\nu^{(2)}$ is valid (cf. (1.3)). This leads to the following generalisation.

Definition 3.32 (tensor rank). The *tensor rank* of $\mathbf{v} \in {}_a\bigotimes_{j=1}^{d} V_j$ is defined by

$$\operatorname{rank}(\mathbf{v}) := \min\{r : \mathbf{v} \in \mathcal{R}_r\} \in \mathbb{N}_0. \tag{3.24}$$

The definition makes sense since subsets of $\mathbb{N}_0$ have always a minimum. As in (2.6), we can characterise the set $\mathcal{R}_r$ by

$$\mathcal{R}_r = \{\mathbf{v} \in \mathbf{V} : \operatorname{rank}(\mathbf{v}) \le r\}.$$

We shall use the shorter 'rank' instead of 'tensor rank'. Note that there is an ambiguity if $\mathbf{v}$ is a matrix as well as a Kronecker tensor (see Remark 3.35b). If necessary, we use the explicit terms 'matrix rank' and 'tensor rank'.

Another ambiguity is caused by isomorphisms discussed in §3.2.5.

Remark 3.33. Consider the isomorphic vector spaces $U \otimes_a V \otimes_a W$ and $U \otimes_a X$ with $X := V \otimes_a W$ from Example 3.28. Let $\Phi : U \otimes_a V \otimes_a W \to U \otimes_a X$ be the vector space isomorphism. Then $\operatorname{rank}(\mathbf{v})$ and $\operatorname{rank}(\Phi(\mathbf{v}))$ are in general different. If we identify $\mathbf{v}$ and $\Phi(\mathbf{v})$, the tensor structure should be explicitly mentioned, e.g., by writing $\operatorname{rank}_{U\otimes V\otimes W}(\mathbf{v})$ or $\operatorname{rank}_{U\otimes X}(\mathbf{v})$. The same statement holds for $\mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j$ and $V_k \otimes_a \mathbf{V}_{[k]}$ from (3.21a).

The latter remark shows that the rank depends on the tensor structure ($U \otimes_a X$ versus $U \otimes_a V \otimes_a W$). However, Lemma 3.36 will prove invariance with respect to tensor space isomorphisms.

Practically, it may be hard to determine the rank. It is not only that the rank is a discontinuous function so that any numerical rounding error may change the rank (as for the matrix rank), even with exact arithmetic the computation of the rank is, in general, not feasible for large-size tensors because of the next statement.

Proposition 3.34 (Håstad [97]). *In general, the determination of the tensor rank is an NP-hard problem.*

If $V_j = \mathbb{K}^{I_j \times J_j}$ are matrix spaces, the tensor rank is also called *Kronecker rank*.

Remark 3.35. (a) For $d = 2$, the rank of $\mathbf{v} \in V_1 \otimes_a V_2$ is given by r from (3.16) and can be constructed as in the proof of Lemma 3.13. If, in addition, the spaces V_j are finite dimensional, Proposition 3.14 yields an isomorphism between $V_1 \otimes_a V_2$ and matrices of size $\dim(V_1) \times \dim(V_2)$. Independently of the choice of bases in Proposition 3.14, the matrix rank of the associated matrix coincides with the tensor rank.
(b) For $V_j = \mathbb{K}^{I_j \times J_j}$ the (Kronecker) tensors $\mathbf{A} \in \mathbf{V} := \bigotimes_{j=1}^{d} V_j$ are matrices. In this case, the matrix rank of $\mathbf{A}$ is completely unrelated to the tensor rank of $\mathbf{A}$. For instance, the identity matrix $\mathbf{I} \in \mathbf{V}$ has (full) matrix rank $\prod_{j=1}^{d} \# I_j$, whereas the tensor rank of the elementary tensor $\mathbf{I} = \bigotimes_{j=1}^{d} I_j$ ($I_j = id \in \mathbb{K}^{I_j \times I_j}$) equals 1.
(c) For tensors of order $d \in \{0, 1\}$, the rank is trivial: $\operatorname{rank}(\mathbf{v}) = \left\{ \begin{smallmatrix} 0 \text{ for } \mathbf{v}=0 \\ 1 \text{ otherwise} \end{smallmatrix} \right\}$.

So far, we have considered only algebraic tensor spaces $\mathbf{V}_{\text{alg}} := {}_a\bigotimes_{j=1}^{d} V_j$. A Banach tensor space $\mathbf{V}_{\text{top}} := {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ (cf. (3.12)) is the closure (completion) of $\bigcup_{r \in \mathbb{N}_0} \mathcal{R}_r$ (cf. Remark 3.30). We can extend the definition of the tensor rank by[4]

$$\operatorname{rank}(\mathbf{v}) := \infty \qquad \text{if} \quad \mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j \setminus {}_a\bigotimes_{j=1}^{d} V_j \; . \tag{3.25}$$

Lemma 3.36 (rank invariance). *Let* $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ *and* $\mathbf{W} = \bigotimes_{j=1}^{d} W_j$ *be either algebraic or topological tensor spaces.*
(a) Assume that $\mathbf{V}$ *and* $\mathbf{W}$ *are isomorphic tensor spaces, i.e., the vector spaces* $V_j \cong W_j$ *are isomorphic (cf. Definition 3.27). Let* $\boldsymbol{\Phi} = \bigotimes_{j=1}^{d} \phi^{(j)} : \mathbf{V} \to \mathbf{W}$ *be an isomorphism. Then the tensor rank of* $\mathbf{v} \in \mathbf{V}$ *is invariant under* $\boldsymbol{\Phi}$*:*

$$\operatorname{rank}(\mathbf{v}) = \operatorname{rank}(\boldsymbol{\Phi}(\mathbf{v})) \qquad \textit{for all } \mathbf{v} \in \mathbf{V}.$$

(b) Let $\mathbf{A} = \bigotimes_{j=1}^{d} A^{(j)} : \mathbf{V} \to \mathbf{W}$ *with* $A^{(j)} \in L(V_j, W_j)$*. Then,*

$$\operatorname{rank}(\mathbf{A}\mathbf{v}) \leq \operatorname{rank}(\mathbf{v}) \qquad \textit{for all } \mathbf{v} \in \mathbf{V}.$$

[4] This does not mean, in general, that $\mathbf{v}$ can be written as an infinite sum (but see §4.2.6 and Theorem 4.110).

Proof. For Part (b) consider $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} b_\nu^{(j)}$. Since the number of terms in $\mathbf{Av} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} (A^{(j)} b_\nu^{(j)})$ is unchanged, $\operatorname{rank}(\mathbf{v}) \geq \operatorname{rank}(\mathbf{Av})$ follows. For Part (a) we use this inequality twice for $\boldsymbol{\Phi}$ and $\boldsymbol{\Phi}^{-1}$: $\operatorname{rank}(\mathbf{v}) \geq \operatorname{rank}(\boldsymbol{\Phi}\mathbf{v}) \geq \operatorname{rank}(\boldsymbol{\Phi}^{-1}\boldsymbol{\Phi}\mathbf{v}) = \operatorname{rank}(\mathbf{v})$. □

Corollary 3.37. Remark 3.29 states a tensor space isomorphism Φ between a finite dimensional tensor space $\mathbf{V} := \bigotimes_{j=1}^{d} V_j$ with $n_j := \dim(V_j)$ and its coefficient tensor space $\mathbf{W} := \bigotimes_{j=1}^{d} \mathbb{K}^{n_j}$. Let $\mathbf{a}$ be the coefficient tensor of $\mathbf{v}$. Then $\operatorname{rank}(\mathbf{v}) = \operatorname{rank}(\mathbf{a})$. As a consequence, $\boldsymbol{\Phi}$ is a bijection between $\mathcal{R}_r(\mathbf{V})$ and $\mathcal{R}_r(\mathbf{W})$.

By definition of the rank of algebraic tensors, there is a representation

$$\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)} \qquad \text{with } r := \operatorname{rank}(\mathbf{v}) < \infty. \tag{3.26}$$

The following lemma offers a necessary condition for (3.26). The proof is known from Lemma 3.13.

Lemma 3.38. *Assume* $r = \operatorname{rank}(\mathbf{v})$. *Using* $v_\nu^{(j)}$ *from (3.26), define the elementary tensors*

$$\mathbf{v}_\nu^{[j]} := \bigotimes_{k \in \{1,\ldots,d\} \setminus \{j\}} v_\nu^{(k)} \in {}_a\bigotimes_{k \in \{1,\ldots,d\} \setminus \{j\}} V_k$$

(cf. (3.21d)). Then (3.26) implies that the tensors $\{\mathbf{v}_\nu^{[j]} : 1 \leq \nu \leq r\}$ *are linearly independent for all* $1 \leq j \leq d$, *while* $v_\nu^{(j)} \neq 0$ *for all* ν *and* j.

Proof. Let $j = 1$. Assume that the elementary tensors $\{\mathbf{v}_\nu^{[1]} : 1 \leq \nu \leq r\}$ are linearly dependent. Without loss of generality, suppose that $\mathbf{v}_r^{[1]}$ may be expressed by the other tensors: $\mathbf{v}_r^{[1]} = \sum_{\nu=1}^{r-1} \beta_\nu \mathbf{v}_\nu^{[1]}$. Then

$$\begin{aligned}\mathbf{v} &= \sum_{\nu=1}^{r} v_\nu^{(1)} \otimes \mathbf{v}_\nu^{[1]} = \sum_{\nu=1}^{r-1} \left(v_\nu^{(1)} + \beta_\nu v_r^{(1)} \right) \otimes \mathbf{v}_\nu^{[1]} \\ &= \sum_{\nu=1}^{r-1} \left(v_\nu^{(1)} + \beta_\nu v_r^{(1)} \right) \otimes v_\nu^{(2)} \otimes \ldots \otimes v_\nu^{(d)}\end{aligned}$$

implies $\operatorname{rank}(\mathbf{v}) < r$. Similarly, if $v_\nu^{(j)} = 0$, the j-th term can be omitted implying again the contradiction $\operatorname{rank}(\mathbf{v}) < r$. Analogously, $j > 1$ is treated. □

Remark 3.39. Note that Lemma 3.38 states linear independence only for the tensors $\mathbf{v}_\nu^{[1]}$, $1 \leq \nu \leq r$. The vectors $v_\nu^{(1)}$ are nonzero, but may be linearly dependent. An example is the tensor from (3.28), which has rank 3, while all subspaces $U_j = \operatorname{span}\{v_\nu^{(j)} : 1 \leq \nu \leq 3\}$ have only dimension 2.

Finally, we mention two extensions of the term 'rank'. Bergman [13] defines a *rank of a subspace* $\mathbf{U} \subset \bigotimes_{j=1}^{d} V_j$ by

$$\operatorname{rank}(\mathbf{U}) := \min\{\operatorname{rank}(\mathbf{x}) : 0 \neq \mathbf{x} \in \mathbf{U}\}.$$

For symmetric tensors $\mathbf{s} \in \mathfrak{S}_d(V)$ (cf. §3.5), a specific *symmetric rank* can be introduced:

$$\operatorname{rank}_{\mathrm{sym}}(\mathbf{s}) := \min\left\{ r \in \mathbb{N}_0 : \mathbf{s} = \sum_{i=1}^{r} \otimes^d v_i \text{ with } v_i \in V \right\} \quad \text{for } \mathbf{s} \in \mathfrak{S}_d(V).$$

Note that each term $\otimes^d v_i = v_i \otimes \ldots \otimes v_i$ is already symmetric.

3.2.6.3 Dependence on the Field

So far, the statements hold for any fixed choice of the field $\mathbb{K}$. Note that the 'real' tensor space $\mathbf{V}_\mathbb{R} := \bigotimes_{j=1}^{d} \mathbb{R}^{n_j}$ can be embedded into the 'complex' tensor space $\bigotimes_{j=1}^{d} \mathbb{C}^{n_j}$. On the other hand, $\mathbf{V}_\mathbb{C} := \mathbf{V}_\mathbb{R} + i\mathbf{V}_\mathbb{R}$ may be considered as a vector space over $\mathbb{R}$ with dimension $2\dim(\mathbf{V}_\mathbb{R})$. Concerning the tensor rank, the following problem arises. Let $\mathbf{v} \in \mathbf{V}_\mathbb{R}$ be a 'real' tensor. The tensor rank is the minimal number $r = r_\mathbb{R}$ of terms in (3.26) with $v_\nu^{(j)} \in \mathbb{R}^{n_j}$. We may also ask for the minimal number $r = r_\mathbb{C}$ of terms in (3.26) under the condition that $v_\nu^{(j)} \in \mathbb{C}^{n_j}$. Since $\mathbb{R}^{I_j} \subset \mathbb{C}^{I_j}$, the inequality $r_\mathbb{C} \leq r_\mathbb{R}$ is obvious, which already proves statement (a) below.

Proposition 3.40. *Let $\mathbf{V}_\mathbb{R} = \bigotimes_{j=1}^{d} V_j$ be a tensor space over the field $\mathbb{R}$. Define $\mathbf{V}_\mathbb{C} = \bigotimes_{j=1}^{d} V_{j,\mathbb{C}}$ as the corresponding complex version over $\mathbb{C}$. Let $r_\mathbb{R}(\mathbf{v})$ be the (real) tensor rank within $\mathbf{V}_\mathbb{R}$, while $r_\mathbb{C}(\mathbf{v})$ is the (complex) tensor rank within $\mathbf{V}_\mathbb{C}$.*

(a) For any $\mathbf{v} \in \mathbf{V}_\mathbb{R}$, the inequality $r_\mathbb{C}(\mathbf{v}) \leq r_\mathbb{R}(\mathbf{v})$ holds.

(b) (vector and matrix case) *If $0 \leq d \leq 2$, $r_\mathbb{C}(\mathbf{v}) = r_\mathbb{R}(\mathbf{v})$ holds for all $\mathbf{v} \in \mathbf{V}_\mathbb{R}$.*

(c) (proper tensor case) *If $d \geq 3$ and $\mathbf{V}_\mathbb{R} = \bigotimes_{j=1}^{d} V_j$ is a non-degenerate tensor space (cf. Definition 3.24), there are $\mathbf{v} \in \mathbf{V}_\mathbb{R}$ with strict inequality $r_\mathbb{C}(\mathbf{v}) < r_\mathbb{R}(\mathbf{v})$.*

Proof. a) Part (a) is already proved above.

b) If $d = 1$, Remark 3.35c shows $r_\mathbb{C} = r_\mathbb{R}$. If $d = 2$, $\mathbf{v} \in \mathbf{V}_\mathbb{R}$ may be interpreted as a real-valued matrix $M \in \mathbb{R}^{I_1 \times I_2}$. By Remark 2.2, the matrix rank is independent of the field: $r_\mathbb{C}(\mathbf{v}) = r_\mathbb{R}(\mathbf{v})$.

c) Example 3.44 below presents a counterexample for which $r_\mathbb{C}(\mathbf{v}) < r_\mathbb{R}(\mathbf{v})$ in the case of $d = 3$. It may be easily embedded into tensor spaces with larger d by setting $\mathbf{v}' := \mathbf{v} \otimes a_4 \otimes a_5 \otimes \ldots \otimes a_d$ with arbitrary $0 \neq a_j \in V_j$ $(4 \leq j \leq d)$. □

Another dependence on the field will be mentioned in §3.2.6.4.

3.2.6.4 Maximal Rank and Typical Ranks

The sequence $\mathcal{R}_0 \subset \mathcal{R}_1 \subset \ldots \subset \mathbf{V} := {}_a\bigotimes_{j=1}^d V_j$ from (3.23a) is properly increasing for infinite dimensional tensor spaces. On the other hand, for finite dimensional tensor spaces there must be a smallest $r_{\max}$ so that $\mathcal{R}_r = \mathcal{R}_{r_{\max}}$ for all $r \geq r_{\max}$. As a consequence,

$$\mathbf{V} = \mathcal{R}_{r_{\max}}, \text{ while } \mathcal{R}_{r_{\max}-1} \subsetneqq \mathbf{V}. \tag{3.27}$$

This $r_{\max}$ is called the *maximal rank* in $\mathbf{V}$ (cf. (2.5) for the matrix case).

Lemma 3.41. *Let $n_j := \dim(V_j) < \infty$ for $1 \leq j \leq d$. Then*

$$r_{\max} \leq \left(\prod_{j=1}^d n_j\right) / \max_{1\leq i\leq d} n_i = \min_{1\leq i\leq d} \prod_{j\in\{1,\ldots,d\}\backslash\{i\}} n_j$$

describes an upper bound of the maximal rank. For equal dimensions $n_j = n$, this is $r_{\max} \leq n^{d-1}$.

Proof. After a permutation of the factors we may assume that $n_d = \max_{1\leq i\leq d} n_i$. Consider the full representation (3.20) of any $\mathbf{v} \in \mathbf{V}$:

$$\begin{aligned}\mathbf{v} &= \sum_{i_1,\ldots,i_{d-1},i_d} \mathbf{a}[i_1,\ldots,i_d]\, b_{i_1}^{(1)} \otimes \ldots \otimes b_{i_d}^{(d)} \\ &= \sum_{i_1,\ldots,i_{d-1}} b_{i_1}^{(1)} \otimes \ldots \otimes b_{i_{d-1}}^{(d-1)} \otimes \left(\sum_{i_d} \mathbf{a}[i_1,\ldots,i_d]\, b_{i_d}^{(d)}\right).\end{aligned}$$

The sum in the last line is taken over $\bar{r} := \prod_{j=1}^{d-1} n_j$ elementary tensors. Hence, $\mathcal{R}_{\bar{r}} = \mathbf{V}$ proves $r_{\max} \leq \bar{r}$. □

The true value $r_{\max}$ may be clearly smaller than the bound from above. For instance, Kruskal [134] proves

$$\begin{aligned}&r_{\max} = \min\{n_1, n_2\} + \min\left\{n_1, n_2, \frac{\max\{n_1, n_2\}}{2}\right\} && \text{for } \mathbf{V} = \mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2} \otimes \mathbb{R}^2, \\ &r_{\max} = 5 && \text{for } \mathbf{V} = \mathbb{R}^3 \otimes \mathbb{R}^3 \otimes \mathbb{R}^3.\end{aligned}$$

Concerning the maximal rank, there is a remarkable difference to the matrix case. Random matrices and their rank are described in Remark 2.5. Random tensors may attain more than one rank with positive probability. Such ranks are called *typical ranks*. Kruskal [134] proves that $\{2, 3\}$ are the typical ranks of $\mathbf{V} = \mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$, while 3 is the maximal rank. Note that such results also depend on the field. For algebraically closed fields like $\mathbb{C}$ there is only one typical rank (cf. Strassen [180], Comon-Golub-Lim-Mourrain [39]). More details are given by Comon et al. [38].

3.2.6.5 Examples

As illustration we consider the tensor $\mathbf{v} \in V \otimes V$ defined by

$$\mathbf{v} = a \otimes a + b \otimes a + a \otimes b + b \otimes a,$$

where $a, b \in V$ are linearly independent. The given representation proves $\mathbf{v} \in \mathcal{R}_4$ and $\operatorname{rank}(\mathbf{v}) \leq 4$. The fact that all four terms are linearly independent is no indication for $\operatorname{rank}(\mathbf{v}) = 4$. In fact, another representation is

$$\mathbf{v} = (a + b) \otimes (a + b)$$

proving $\operatorname{rank}(\mathbf{v}) = 1$, since $\mathbf{v} \neq 0$ excludes $\operatorname{rank}(\mathbf{v}) = 0$.

For later use we exercise the determination of the rank for a special tensor.

Lemma 3.42. *Let V_j $(1 \leq j \leq 3)$ be vector spaces of dimension ≥ 2 and consider the tensor space $\mathbf{V} := V_1 \otimes V_2 \otimes V_3$. For linearly independent vectors $v_j, w_j \in V_j$ define*

$$\mathbf{v} := v_1 \otimes v_2 \otimes w_3 + v_1 \otimes w_2 \otimes v_3 + w_1 \otimes v_2 \otimes v_3. \tag{3.28}$$

Then $\operatorname{rank}(\mathbf{v}) = 3$ holds, i.e., the given representation is already the shortest one.

Proof. For $r = 0, 1, 2$ we show below that $\operatorname{rank}(\mathbf{v}) = r$ cannot be valid. Then the given representation proves $\operatorname{rank}(\mathbf{v}) = 3$.

1) $\operatorname{rank}(\mathbf{v}) = 0$ implies $\mathbf{v} = 0$. But the terms on the right-hand side of (3.28) are linearly independent and therefore their sum cannot vanish.

2) Assume $\operatorname{rank}(\mathbf{v}) = 1$, i.e., $\mathbf{v} = u \otimes v \otimes w$ with non-vanishing $u, v, w \in V$. There is a linear functional $\varphi \in V_1'$ with $\varphi(v_1) = 1$. Applying $\varphi \otimes id \otimes id : V_1 \otimes V_2 \otimes V_3 \to V_2 \otimes V_3$ to both representations of $\mathbf{v}$, we obtain

$$\varphi(u) v \otimes w = v_2 \otimes w_3 + w_2 \otimes v_3 + \varphi(w_1) v_2 \otimes v_3.$$

The matrix on the left-hand side has rank ≤ 1, while the matrix on the right-hand side has rank 2. Hence $\operatorname{rank}(\mathbf{v}) = 1$ cannot hold.

2) Assume $\operatorname{rank}(\mathbf{v}) = 2$, i.e., $\mathbf{v} = u \otimes v \otimes w + u' \otimes v' \otimes w'$. If u and u' are linearly dependent, there is a functional φ with $\varphi(u) = \varphi(u') = 0$, while either $\varphi(v_1) \neq 0$ or $\varphi(w_1) \neq 0$. Then

$$0 = (\varphi \otimes id \otimes id)(\mathbf{v}) = \varphi(v_1)(v_2 \otimes w_3 + w_2 \otimes v_3) + \varphi(w_1) v_2 \otimes v_3.$$

Since $v_2 \otimes w_3 + w_2 \otimes v_3$ and $v_2 \otimes v_3$ are linearly independent, this is a contradiction. Hence u and u' are linearly independent and one of the vectors u or u' must be linearly independent of v_1, say u' and v_1 are linearly independent. Choose $\varphi \in V_1'$ with $\varphi(v_1) = 1$ and $\varphi(u') = 0$. Then

$$\varphi(u) v \otimes w = (\varphi \otimes id \otimes id)(\mathbf{v}) = (v_2 \otimes w_3 + w_2 \otimes v_3) + \varphi(w_1) v_2 \otimes v_3.$$

The matrix on the left-hand side has rank ≤ 1, while the matrix on the right-hand side has rank 2. This contradiction completes the proof. □

Exercise 3.43. Consider $\mathbf{v} = \bigotimes_{j=1}^{d} v_j + \bigotimes_{j=1}^{d} w_j$ with non-vanishing vectors v_j and w_j. Show that $\mathrm{rank}(\mathbf{v}) \leq 1$ holds if and only if v_j and w_j are linearly dependent for at least $d-1$ indices $j \in \{1, \ldots, d\}$. Otherwise, $\mathrm{rank}(\mathbf{v}) = 2$.

Concerning the distinction of $\mathrm{rank}_{\mathbb{R}}$ and $\mathrm{rank}_{\mathbb{C}}$ we give the following example.

Example 3.44. Let $a, b, c, a', b', c' \in \mathbb{R}^n$ with $n \geq 2$ such that (a, a'), (b, b'), (c, c') are pairs of linearly independent vectors. The real part of the complex tensor $(a + ia') \otimes (b + ib') \otimes (c + ic') \in \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ has a representation

$$\mathbf{v} = \tfrac{1}{2}\,(a + ia')\otimes(b + ib')\otimes(c + ic') + \tfrac{1}{2}\,(a - ia')\otimes(b - ib')\otimes(c - ic') \in \mathcal{R}_2$$

in $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$. Exercise 3.43 proves that $\mathrm{rank}_{\mathbb{C}}(\mathbf{v}) = 2$ in $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$. Multilinearity yields the representation

$$\mathbf{v} = a \otimes b \otimes c - a' \otimes b' \otimes c - a' \otimes b \otimes c' - a \otimes b' \otimes c'$$

within $\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n$. One verifies that also

$$\mathbf{v} = (a - a') \otimes (b + b') \otimes c + a' \otimes b \otimes (c - c') - a \otimes b' \otimes (c + c') \tag{3.29}$$

holds. A further reduction is not possible so that $\mathrm{rank}_{\mathbb{R}}(\mathbf{v}) = 3 > 2 = \mathrm{rank}_{\mathbb{C}}(\mathbf{v})$ is valid.

Proof. Assume that $\mathbf{v} = A \otimes B \otimes C + A' \otimes B' \otimes C'$. Applying suitable functionals to the first two components, one sees that $C, C' \in \mathrm{span}\{c, c'\}$. If C and C' are linearly dependent, this leads to a quick contradiction. So assume that they are linearly independent and choose functionals $\gamma \in (\mathbb{R}^n)'$ with $\gamma(C) = 1$ and $\gamma(C') = 0$. Note that at least two of the numbers $\gamma(c)$, $\gamma(c - c')$, $\gamma(c + c')$ are nonzero. Hence application of $id \otimes id \otimes \gamma$ to $\mathbf{v} = A \otimes B \otimes C + A' \otimes B' \otimes C'$ yields $A \otimes B$ with matrix rank equal to 1, while the result for $\mathbf{v}$ from (3.29) is a linear combination of $(a - a') \otimes (b + b')$, $a' \otimes b$, $a \otimes b'$, where at least two terms are present. One verifies that the matrix rank is 2. This contradiction excludes $\mathrm{rank}_{\mathbb{R}}(\mathbf{v}) = 2$. It is even easier to exclude the smaller ranks 1 and 0. □

The next example of different n-term representations over the real or complex field, which is of practical interest, is taken from Mohlenkamp-Monzón [151] and Beylkin-Mohlenkamp [15].

Example 3.45. Consider the function $f(x_1, \ldots, x_d) := \sin\big(\sum_{j=1}^{d} x_j\big) \in \otimes^d V$ for $V = C(\mathbb{R})$. If $C(\mathbb{R})$ is regarded as vector space over $\mathbb{K} = \mathbb{C}$, $\mathrm{rank}(f) = 2$ holds and is realised by

$$\sin\left(\sum_{j=1}^{d} x_j\right) = \frac{1}{2\mathrm{i}}\mathrm{e}^{\mathrm{i}\sum_{j=1}^{d} x_j} - \frac{1}{2\mathrm{i}}\mathrm{e}^{-\mathrm{i}\sum_{j=1}^{d} x_j} = \frac{1}{2\mathrm{i}}\bigotimes_{j=1}^{d}\mathrm{e}^{\mathrm{i}x_j} - \frac{1}{2i}\bigotimes_{j=1}^{d}\mathrm{e}^{-\mathrm{i}x_j}.$$

If $C(\mathbb{R})$ is considered as vector space over $\mathbb{K} = \mathbb{R}$, the following representation needs d terms:

$$\sin\left(\sum_{j=1}^{d} x_j\right) = \sum_{\nu=1}^{d}\left(\bigotimes_{j=1}^{\nu-1}\frac{\sin(x_j+\alpha_j-\alpha_\nu)}{\sin(\alpha_j-\alpha_\nu)}\right)\otimes\sin(x_\nu)\otimes\left(\bigotimes_{j=\nu+1}^{d}\frac{\sin(x_j+\alpha_j-\alpha_\nu)}{\sin(\alpha_j-\alpha_\nu)}\right)$$

with arbitrary $\alpha_j \in \mathbb{R}$ satisfying $\sin(\alpha_j - \alpha_\nu) \neq 0$ for all $j \neq \nu$.

In the case of $d = 2$, the representation (3.26) corresponds to a matrix $M = \sum_{i=1}^{r} a_i b_i^{\mathsf{T}} \in \mathbb{K}^{I\times J}$ with vectors $a_i \in \mathbb{K}^I$ and $b_i \in \mathbb{K}^J$. Here, the minimal (tensor and matrix) rank r is attained if the vectors a_i and the vectors b_i are linearly independent. Moreover, the singular value decomposition yields the particular form $M = \sum_{i=1}^{r} \sigma_i a_i b_i^{\mathsf{T}}$ with $\sigma_i > 0$ and orthonormal a_i and b_i. Generalisations of these properties to $d \geq 3$ are not valid.

Remark 3.46. (a) A true generalisation of the singular value decomposition to d dimensions would be $\mathbf{v} = \sum_{\nu=1}^{r}\sigma_\nu \bigotimes_{j=1}^{d} v_\nu^{(j)} \in \mathbf{V} := \bigotimes_{j=1}^{d}\mathbb{K}^{n_j}$ with $r = \operatorname{rank}(\mathbf{v})$, orthonormal vectors $\{v_\nu^{(j)} : 1 \leq \nu \leq r\}$ for all $1 \leq j \leq d$, and $\sigma_\nu > 0$. Unfortunately, such tensors form only a small subset of $\mathbf{V}$, i.e., in general, $\mathbf{v} \in \mathbf{V}$ does not possess such a representation.
(b) Even the requirement that the vectors $\{v_\nu^{(j)} : 1 \leq \nu \leq r\}$ are linearly independent cannot be satisfied in general.

Proof. The tensors $a \otimes a \otimes a + a \otimes b \otimes b$ cannot be reduced to rank ≤ 1, although the first factors are equal. This proves Part (b), while (b) implies (a). □

3.2.6.6 Application: Strassen's Algorithm

The standard matrix-matrix multiplication of two $n \times n$ matrices costs $2n^3$ operations. A reduction to $4.7n^{\log_2 7} = 4.7n^{2.8074}$ proposed by Strassen [179] is based on the fact that two 2×2 block matrices can be multiplied as follows:

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}\begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}, \qquad a_i, b_i, c_i \text{ submatrices with} \tag{3.30}$$

$$\begin{aligned} c_1 &= m_1+m_4-m_5+m_7,\; c_2=m_2+m_4,\; c_3=m_3+m_5,\; c_4=m_1+m_3-m_2+m_6,\\ m_1 &= (a_1+a_4)(b_1+b_4),\; m_2=(a_3+a_4)b_1,\; m_3=a_1(b_2-b_4),\; m_4=a_4(b_3-b_1),\\ m_5 &= (a_1+a_2)b_4,\; m_6=(a_3-a_1)(b_1+b_2),\; m_7=(a_2-a_4)(b_3+b_4), \end{aligned}$$

where only 7 multiplications of block matrices are needed.

The entries of a tensor $\mathbf{v} \in \mathbb{K}^{4\times4\times4}$ are involved in

$$c_\nu = \sum_{\mu,\lambda=1}^{4} \mathbf{v}_{\nu\mu\lambda}\, a_\mu\, b_\lambda \qquad (1 \leq \nu \leq 4). \tag{3.31a}$$

For instance for $\nu = 1$, the identity $c_1 = a_1 b_1 + a_2 b_3$ shows that $\mathbf{v}_{111} = \mathbf{v}_{123} = 1$ and $\mathbf{v}_{1\mu\lambda} = 0$, otherwise. Assume a representation of $\mathbf{v}$ by r terms:

$$\mathbf{v} = \sum\nolimits_{i=1}^{r} \bigotimes\nolimits_{j=1}^{3} v_i^{(j)}. \tag{3.31b}$$

Insertion into (3.31a) yields

$$\begin{aligned} c_\nu &= \sum_{i=1}^{r} \sum_{\mu,\lambda=1}^{4} v_i^{(1)}[\nu]\, v_i^{(2)}[\mu]\, v_i^{(3)}[\lambda]\, a_\mu\, b_\lambda \\ &= \sum_{i=1}^{r} v_i^{(1)}[\nu] \left(\sum_{\mu=1}^{4} v_i^{(2)}[\mu]\, a_\mu \right) \left(\sum_{\lambda=1}^{4} v_i^{(3)}[\lambda]\, b_\lambda \right), \end{aligned} \tag{3.31c}$$

i.e., only r multiplications are needed. Algorithm (3.30) corresponds to a representation (3.31b) with $r = 7$.

3.3 Linear and Multilinear Mappings

Now, we consider linear mappings defined on $V \otimes_a W$ or, more generally, on ${}_a\bigotimes_{j=1}^{d} V_j$. The image space might be the field $\mathbb{K}$—then the linear mappings are called *linear forms* or *functionals*—or another tensor space.

In §3.3.1 we justify that it suffices to define a mapping by its values for elementary tensors. Often a linear mapping $\varphi_k : V_k \to W_k$ for a fixed k is extended to a linear mapping defined on $\mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j$. This leads to an embedding explained in §3.3.2. Functionals are a special kind of linear maps. Nevertheless, there are special properties which are addressed in §3.3.2.2.

3.3.1 Definition on the Set of Tuples

If a linear mapping ϕ is to be defined on a vector space V spanned by a basis $\{v_j\}$, it suffices to describe the images $\phi(v_j)$ (cf. Remark 3.5).

In the case of a linear mapping

$$\phi : \mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j \to X$$

we know that $\mathbf{V}$ is spanned by elementary tensors $\bigotimes_{j=1}^{d} v^{(j)}$. Hence, it is sufficient to know the image $\phi\big(\bigotimes_{j=1}^{d} v^{(j)}\big)$. In fact, such values are often given by means of a mapping $\varPhi : V_1 \times \ldots \times V_d \to X$:

$$\phi\left(\bigotimes_{j=1}^{d} v^{(j)} \right) = \varPhi(v^{(1)}, \ldots, v^{(d)}) \quad \text{for all } v^{(j)} \in V_j. \tag{3.32}$$

Since the elementary tensors are not linearly independent, it is not obvious whether these values are not contradictory.

The answer follows from the 'universality of the tensor product' formulated in Proposition 3.22: If

$$\Phi : V_1 \times \ldots \times V_d \to U \quad \text{is multilinear,}$$

(3.32) defines a unique linear mapping. Multilinearity of Φ is equivalent to

$$\Phi(n) = 0 \qquad \text{for all } n \in N,$$

where N is the analogue of (3.9) for general d.

The next lemma shows a particular case.

Lemma 3.47. *Let $\varphi \colon V \to U$ be a linear mapping. Then the definition*

$$\phi\,(v \otimes w) := \varphi(v) \otimes w \quad \textit{for all } v \in V, w \in W$$

defines a unique linear mapping from $V \otimes_a W$ into $U \otimes_a W$.

Proof. This is the case of $X = U \otimes_a W$ and $\Phi(v^{(1)}, v^{(2)}) = \varphi(v^{(1)}) \otimes v^{(2)}$. Linearity of φ shows multilinearity of Φ. □

Of course, the same statement holds for a linear mapping $\psi : W \to U$. Then $\phi\,(v \otimes w) := v \otimes \psi(w)$ defines a unique linear mapping from $V \otimes_a W$ to $V \otimes_a U$.

Also the generalisation to tensor spaces of order d is obvious. Let V_j $(1 \le j \le d)$ and W be vector spaces and fix an index $k \in \{1, \ldots, d\}$. Given a linear mapping $\varphi : V_k \to W$, define

$$\phi : \bigotimes_{j=1}^{d} v^{(j)} \mapsto v^{(1)} \otimes \ldots \otimes v^{(k-1)} \otimes \varphi(v^{(k)}) \otimes v^{(k+1)} \otimes \ldots \otimes v^{(d)}. \tag{3.33}$$

Then, there is a unique extension to

$$\phi \in L\left(\bigotimes_{j=1}^{d} V_j,\ \left(\bigotimes_{j=1}^{k-1} V_j \right) \otimes W \otimes \left(\bigotimes_{j=k+1}^{d} V_j \right) \right).$$

Another generalisation concerns bilinear mappings.

Remark 3.48. Let $\varphi_i : V_i \times W_i \to \mathbb{K}$ $(1 \le i \le d)$ be bilinear [sesquilinear] forms (cf. §3.1.4). Then

$$\phi\left(\bigotimes_{j=1}^{d} v^{(j)}, \bigotimes_{j=1}^{d} w^{(j)} \right) = \prod_{j=1}^{d} \varphi_i(v^{(j)}, w^{(j)}) \quad (v^{(j)} \in V_j,\ w^{(j)} \in W_j)$$

defines a unique bilinear [sesquilinear] form $\phi : \left(\bigotimes_{j=1}^{d} V_j \right) \times \left(\bigotimes_{j=1}^{d} W_j \right) \to \mathbb{K}$.

3.3.2 Embeddings

3.3.2.1 Embedding of Spaces of Linear Maps

In the following, we consider two d-tuples $(V_1,...,V_d)$ and $(W_1,...,W_d)$ of vector spaces and the corresponding tensor spaces $\mathbf{V} := {}_a\bigotimes_{j=1}^d V_j$ and $\mathbf{W} := {}_a\bigotimes_{j=1}^d W_j$. Since $L(V_j, W_j)$ for $1 \le j \le d$ are again vector spaces, we can build the tensor space

$$\mathbf{L} := {}_a\bigotimes_{j=1}^{d} L(V_j, W_j)\,. \tag{3.34a}$$

Elementary tensors from $\mathbf{L}$ are of the form $\boldsymbol{\Phi} = \bigotimes_{j=1}^d \varphi^{(j)}$ with $\varphi^{(j)} \in L(V_j, W_j)$. In §1.1.2, we have called $\boldsymbol{\Phi}$ the Kronecker product[5] of the mappings $\varphi^{(j)}$. They have a natural interpretation as mappings of $L(\mathbf{V}, \mathbf{W})$ via (1.4b):

$$\boldsymbol{\Phi}\left(\bigotimes_{j=1}^{d} v^{(j)}\right) = \bigotimes_{j=1}^{d} \varphi^{(j)}(v^{(j)}) \in \mathbf{W} \qquad \text{for any } \bigotimes_{j=1}^{d} v^{(j)} \in \mathbf{V}. \tag{3.34b}$$

Note that (3.34b) defines $\boldsymbol{\Phi}$ for all elementary tensors of $\mathbf{V}$. By the considerations from §3.3.1, $\boldsymbol{\Phi}$ can be uniquely extended to $\boldsymbol{\Phi} \in L(\mathbf{V},\mathbf{W})$. Linear combinations of such elementary tensors $\boldsymbol{\Phi}$ are again elements of $L(\mathbf{V},\mathbf{W})$. This leads to the embedding described below.

Proposition 3.49. *Let V_j, W_j, $\mathbf{V}$, and $\mathbf{W}$ as above. We identify ${}_a\bigotimes_{j=1}^d L(V_j, W_j)$ with a subspace of $L(\mathbf{V}, \mathbf{W})$ via (3.34b):*

$$\mathbf{L} = {}_a\bigotimes_{j=1}^{d} L(V_j, W_j) \subset L(\mathbf{V}, \mathbf{W}). \tag{3.34c}$$

In general, $\mathbf{L}$ is a proper subspace. If, however, the vector spaces V_j are finite dimensional, the spaces coincide:

$${}_a\bigotimes_{j=1}^{d} L(V_j, W_j) = L(\mathbf{V}, \mathbf{W}). \tag{3.34d}$$

Proof. a) Definition (3.34b) yields a linear mapping $\Upsilon : \mathbf{L} \to L(\mathbf{V}, \mathbf{W})$. It describes an embedding if and only if Υ is injective. For this purpose, we use induction over d and start with $d = 2$. We have to disprove $\Upsilon(\boldsymbol{\Lambda}) = 0$ for $0 \neq \boldsymbol{\Lambda} \in \mathbf{L} = L(V_1, W_1) \otimes_a L(V_1, W_1)$. If $\Upsilon(\boldsymbol{\Lambda}) = 0$, the interpretation (3.34b) of $\boldsymbol{\Lambda}$ produces the zero mapping $\Upsilon(\boldsymbol{\Lambda})$ in $L(\mathbf{V}, \mathbf{W})$. By Lemma 3.13, there

[5] At least, this term is used for matrix spaces $L(V_j, W_j)$ with $V_j = \mathbb{K}^{n_j}$ and $W_j = \mathbb{K}^{m_j}$. As mentioned in §1.6, the attribution to Kronecker is questionable.

is a representation $\boldsymbol{\Lambda} = \sum_{\nu=1}^{r} \varphi_\nu^{(1)} \otimes \varphi_\nu^{(2)}$ with linearly independent $\varphi_\nu^{(2)}$ and[6] $\varphi_1^{(1)} \neq 0$. Application to any $\mathbf{v} := v^{(1)} \otimes v^{(2)}$ yields

$$0 = \boldsymbol{\Lambda}(\mathbf{v}) = \sum_{\nu=1}^{r} \varphi_\nu^{(1)}(v^{(1)}) \otimes \varphi_\nu^{(2)}(v^{(2)}) \in \mathbf{W} = W_1 \otimes_a W_2.$$

Fix $v^{(1)}$ such that $\varphi_1^{(1)}(v^{(1)}) \neq 0$. Then there is some functional $\chi \in W_1'$ with $\chi(\varphi_1^{(1)}(v^{(1)})) \neq 0$. Application of $\chi \otimes id$ to $\boldsymbol{\Lambda}(\mathbf{v})$ yields

$$0 = (\chi \otimes id)\,(\boldsymbol{\Lambda}(\mathbf{v})) = \sum_{\nu=1}^{r} \alpha_\nu \, \varphi_\nu^{(2)}(v^{(2)}) \in W_2$$

(cf. Remark 3.54) with $\alpha_\nu := \chi(\varphi_\nu^{(1)}(v^{(1)}))$. The choice of $v^{(1)}$ and χ ensures that $\alpha_1 \neq 0$. Linear independence of $\varphi_\nu^{(2)}$ implies $\sum_{\nu=1}^{r} \alpha_\nu \varphi_\nu^{(2)} \neq 0$. Hence, there exists some $v^{(2)} \in V_2$ with $\sum_{\nu=1}^{r} \alpha_\nu \, \varphi_\nu^{(2)}(v^{(2)}) \neq 0$ in contradiction to $0 = (\chi \otimes id)\left(\boldsymbol{\Lambda}(\mathbf{v})(v^{(1)} \otimes v^{(2)})\right)$. This proves injectivity of Υ for $d = 2$.

Let the assertion be valid for $d - 1$. Represent $\boldsymbol{\Lambda}$ in the form

$$\boldsymbol{\Lambda} = \sum_{\nu=1}^{r} \varphi_\nu^{(1)} \otimes \varphi_\nu^{[1]} \qquad \text{with } r = \operatorname{rank}_{V_1 \otimes V_{[1]}}(\boldsymbol{\Lambda})$$

(cf. Remark 3.33). As stated in Lemma 3.38, $\{\varphi_\nu^{[1]}\}$ is linearly independent, while $\varphi_\nu^{(1)} \neq 0$. By induction, $\varphi_\nu^{[1]} \in L(\mathbf{V}_{[1]}, \mathbf{W}_{[1]})$ holds with $\mathbf{V}_{[1]} := {}_a\bigotimes_{j=2}^{d} V_j$ and $\mathbf{W}_{[1]} := {}_a\bigotimes_{j=2}^{d} W_j$. Now, all arguments from the inductive start $d = 2$ can be repeated.

b) The equality in (3.34d) holds, if Υ is surjective. For any $\boldsymbol{\Phi} \in L(\mathbf{V}, \mathbf{W})$, we shall construct $\boldsymbol{\phi} \in {}_a\bigotimes_{j=1}^{d} L(V_j, W_j)$ with $\Upsilon(\boldsymbol{\phi}) = \boldsymbol{\Phi}$, provided V_j are finite dimensional. Again, considerations for $d = 2$ are sufficient. Since $\mathbf{V}$ is finite dimensional, also the image $\boldsymbol{\Phi}(\mathbf{V}) \subset \mathbf{W}$ is finite dimensional. In fact, $\boldsymbol{\Phi}(\mathbf{V}) \subset \hat{W}_1 \otimes_a \hat{W}_2 \subset \mathbf{W}$ holds with finite dimensional subspaces $\hat{W}_j := U_j^{\min}(\boldsymbol{\Phi}(\mathbf{V})) \subset W_j$ introduced later in §6.2.3 (see also Exercise 6.14b). Let $\{b_{i,v}^{(j)} : 1 \leq i \leq \dim(V_j)\}$ be a basis of V_j, and $\{b_{i,w}^{(j)} : 1 \leq i \leq \dim(\hat{W}_j)\}$ a basis of $\hat{W}_j$. The dual basis $\{\chi_i^{(j)} : 1 \leq i \leq \dim(V_j)\}$ of V_j' satisfies $\chi_i^{(j)}(b_{i',v}^{(j)}) = \delta_{i,i'}$ (cf. Definition 3.6). Each image $\mathbf{w}_{ij} := \boldsymbol{\Phi}(b_{i,v}^{(1)} \otimes b_{j,v}^{(2)}) \in \hat{W}_1 \otimes_a \hat{W}_2$ has a representation $\mathbf{w}_{ij} = \sum_{\nu\mu} \alpha_{ij,\nu\mu} b_{\nu,w}^{(1)} \otimes b_{\mu,w}^{(2)}$. Set

$$\boldsymbol{\phi} := \sum_{ij\nu\mu} \alpha_{ij,\nu\mu} \varphi_{i,\nu}^{(1)} \otimes \varphi_{j,\mu}^{(2)}, \quad \text{where} \begin{cases} \varphi_{i,\nu}^{(1)}(v^{(1)}) := \chi_i^{(1)}(v^{(1)}) \cdot b_{\nu,w}^{(1)}\,, \\ \varphi_{j,\mu}^{(2)}(v^{(2)}) := \chi_j^{(2)}(v^{(2)}) \cdot b_{\mu,w}^{(2)}\,. \end{cases}$$

Since $\Upsilon(\boldsymbol{\phi})$ and $\boldsymbol{\Phi}$ coincide on all basis vectors $b_{i,v}^{(1)} \otimes b_{j,v}^{(2)}$, $\Upsilon(\boldsymbol{\phi}) = \boldsymbol{\Phi}$ holds.

c) A counterexample for infinite dimensional V_j will follow in Example 3.53. □

[6] In fact, $\varphi_\nu^{(1)}$ are linearly independent, but only $\varphi_1^{(1)} \neq 0$ is needed.

In Lemma 3.47 we use the mapping $v \otimes w \mapsto \varphi(v) \otimes w$, which is generalised to $\boldsymbol{\Phi} : \bigotimes_{j=1}^{d} v^{(j)} \mapsto v^{(1)} \otimes \ldots \otimes v^{(k-1)} \otimes \varphi\left(v^{(k)}\right) \otimes v^{(k+1)} \otimes \ldots \otimes v^{(d)}$ in (3.33). The latter mapping can be formulated as

$$\boldsymbol{\Phi} = \bigotimes_{j=1}^{d} \varphi^{(j)} \quad \text{with} \quad \begin{cases} \varphi^{(j)} = \varphi, \; W_j := W & \text{for } j = k, \\ \varphi^{(j)} = id, \; W_j := V_j & \text{for } j \neq k, \end{cases} \tag{3.35a}$$

or $\boldsymbol{\Phi} = id \otimes \ldots \otimes id \otimes \varphi \otimes id \otimes \ldots \otimes id$. Since such a notation is rather cumbersome, we identify[7] φ and $\boldsymbol{\Phi}$ as stated below.

Notation 3.50. (a) A mapping $\varphi \in L(V_k, W_k)$ for some $k \in \{1, \ldots, d\}$ is synonymously interpreted as $\boldsymbol{\Phi}$ from (3.35a). This defines the embedding

$$L(V_k, W_k) \subset L(\mathbf{V}, \mathbf{W}) \tag{3.35b}$$

where $\mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$ and $\mathbf{W} = {}_a\bigotimes_{j=1}^{d} W_j$ with $W_j = V_j$ for $j \neq k$.
(b) Let $\alpha \subset \{1, \ldots, d\}$ be a non-empty subset. Then the embedding

$${}_a\bigotimes_{k \in \alpha} L(V_k, W_k) \subset L(\mathbf{V}, \mathbf{W}) \tag{3.35c}$$

is defined analogously by inserting identity maps for all $j \in \{1, \ldots, d\} \backslash \alpha$.

Finally, we repeat the composition rule for Kronecker products from §4.6.3, where they are formulated for Kronecker matrices.

Remark 3.51. (a) Let $\boldsymbol{\Psi} = \bigotimes_{j=1}^{d} \psi^{(j)} \in L(\mathbf{U}, \mathbf{V})$ and $\boldsymbol{\Phi} = \bigotimes_{j=1}^{d} \varphi^{(j)} \in L(\mathbf{V}, \mathbf{W})$ be elementary tensors. Then the composition of the mappings satisfies

$$\boldsymbol{\Phi} \circ \boldsymbol{\Psi} = \bigotimes_{j=1}^{d} \left(\varphi^{(j)} \circ \psi^{(j)} \right) \in L(\mathbf{U}, \mathbf{W}).$$

(b) Let $\varphi \in L(V_k, W_k)$ and $\psi \in L(V_\ell, W_\ell)$ with $k \neq \ell$. Then $\varphi \circ \psi = \psi \circ \varphi$ holds. Moreover, $\varphi \circ \psi = \varphi \otimes \psi$ is valid, where the embedding (3.35b) is used on the left-hand side, while (3.35c) with $\alpha = \{k, \ell\}$ is used on the right-hand side.
(c) Let $\varphi^{(j)} \in L(V_j, W_j)$ for all $1 \leq j \leq d$. Then $\varphi^{(1)} \circ \varphi^{(2)} \circ \ldots \circ \varphi^{(d)}$ interpreted by (3.35b) equals $\bigotimes_{j=1}^{d} \varphi^{(j)}$.

3.3.2.2 Embedding of Linear Functionals

The previous linear mappings become functionals, if the image space W_j is the trivial vector space $\mathbb{K}$. However, the difference is seen from the following example. Consider the mapping $\varphi : V \to U$ from Lemma 3.47 and the induced mapping

[7] The identifications (3.35a-c) are standard in other fields. If we use the multi-index notation $\partial^{\mathbf{n}} f = \partial_x^{n_1} \partial_y^{n_2} \partial_z^{n_3} f$ for the partial derivative, the identities are expressed by $\partial_y^{n_2} = \partial_z^{n_3} = id$ if, e.g., $n_2 = n_3 = 0$. However, more often $\frac{\partial}{\partial x} f(x, y, z)$ is written omitting the identities in $\partial / \partial x \otimes id \otimes id$.

$\boldsymbol{\Phi} : v \otimes w \mapsto \varphi(v) \otimes w \in U \otimes_a W$. For $U = \mathbb{K}$, the mapping is a functional and the image belongs to $\mathbb{K} \otimes_a W$. As $\mathbb{K} \otimes_a W$ is isomorphic[8] to W, it is standard to simplify $U \otimes_a W$ to W (cf. Remark 3.25a). This means that $\varphi(v) \in \mathbb{K}$ is considered as scalar factor: $\varphi(v) \otimes w = \varphi(v) \cdot w \in W$.

When we want to reproduce the notations from §3.3.2.1, we have to modify $\mathbf{W}$ by omitting all factors $W_j = \mathbb{K}$.

The counterpart of Proposition 3.49 is the statement below. Here, $\mathbf{W} = \bigotimes_{j=1}^{d} \mathbb{K}$ degenerates to $\mathbb{K}$, i.e., $L(\mathbf{V}, \mathbf{W}) = \mathbf{V}'$.

Proposition 3.52. *Let V_j $(1 \le j \le d)$ be vector spaces generating $\mathbf{V} := {}_a\bigotimes_{j=1}^d V_j$. Elementary tensors of ${}_a\bigotimes_{j=1}^d V_j'$ are $\boldsymbol{\Phi} = \bigotimes_{j=1}^d \varphi^{(j)}$, $\varphi^{(j)} \in V_j'$. Their application to tensors from $\mathbf{V}$ is defined via*

$$\boldsymbol{\Phi}\left(\bigotimes_{j=1}^{d} v^{(j)}\right) = \prod_{j=1}^{d} \varphi^{(j)}(v^{(j)}) \in \mathbb{K}. \tag{3.36a}$$

This defines the embedding

$$ {}_a\bigotimes_{j=1}^{d} V_j' \subset \mathbf{V}', \quad \textit{and} \tag{3.36b}$$

$$ {}_a\bigotimes_{j=1}^{d} V_j' = \mathbf{V}', \quad \textit{if } \dim(V_j) < \infty \textit{ for } 1 \le j \le d. \tag{3.36c}$$

The next example shows that, in general, (3.36b) holds with proper inclusion.

Example 3.53. Choose $\mathbf{V} := \ell_0 \otimes_a \ell_0$ with ℓ_0 from (3.2) and consider the functional $\boldsymbol{\Phi} \in \mathbf{V}'$ with $\boldsymbol{\Phi}(v \otimes w) := \sum_{i \in \mathbb{Z}} v_i w_i$. Note that this infinite sum is well-defined, since $v_i w_i = 0$ for almost all $i \in \mathbb{Z}$. Since $\boldsymbol{\Phi}$ does not belong to $\ell_0' \otimes_a \ell_0'$, the latter space is a *proper subspace* of $(\ell_0 \otimes_a \ell_0)'$.

Proof. For an indirect proof assume that $\boldsymbol{\Phi} = \sum_{\nu=1}^{k} \varphi_\nu \otimes \psi_\nu \in \ell_0' \otimes_a \ell_0'$ for some $\varphi_\nu, \psi_\nu \in \ell_0'$ and $k \in \mathbb{N}$. Choose any integer $m > k$. Let $e^{(i)} \in \ell_0$ be the i-th unit vector. The assumed identity $\boldsymbol{\Phi} = \sum_{\nu=1}^{k} \varphi_\nu \otimes \psi_\nu$ tested for all $e^{(i)} \otimes e^{(j)} \in \ell_0 \otimes_a \ell_0$ with $1 \le i, j \le m$ yields m^2 equations

$$\delta_{jk} = \boldsymbol{\Phi}(e^{(i)} \otimes e^{(j)}) = \sum_{\nu=1}^{k} \varphi_\nu(e^{(i)}) \cdot \psi_\nu(e^{(j)}) \qquad (1 \le i, j \le m). \tag{3.37}$$

Define matrices $A, B \in \mathbb{K}^{m \times k}$ by $A_{i\nu} := \varphi_\nu(e^{(i)})$ and $B_{j\nu} := \psi_\nu(e^{(j)})$. Then equation (3.37) becomes $I = AB^{\mathsf{T}}$. Since $\operatorname{rank}(A) \le \min\{m, k\} = k$, also the rank of the product AB^{T} is bounded by k, contradicting $\operatorname{rank}(I) = m > k$. □

[8] Concerning isomorphisms compare the last paragraph in §3.2.5.

A very important case is the counterpart of (3.35b) from Notation 3.50a. In the following, we use the notations $\bigotimes_{j\neq k}$ and $\mathbf{V}_{[k]}$ from (3.21a,b).

Remark 3.54. (a) Let V_j $(1\leq j\leq d)$ be vector spaces generating $\mathbf{V} := {}_a\bigotimes_{j=1}^d V_j$. For a fixed index $k\in\{1,\ldots,d\}$ let $\varphi^{(k)}\in V_k'$ be a linear functional. Then $\varphi^{(k)}$ induces the definition of $\boldsymbol{\Phi}\in L(\mathbf{V},\mathbf{V}_{[k]})$ by

$$\boldsymbol{\Phi}\left(\bigotimes_{j=1}^d v^{(j)}\right) := \varphi^{(k)}\left(v^{(k)}\right)\cdot\bigotimes_{j\neq k} v^{(j)}. \tag{3.38a}$$

(b) According to (3.35b), we identify $\boldsymbol{\Phi} = id\otimes\ldots\otimes\varphi^{(k)}\otimes\ldots\otimes id$ and $\varphi^{(k)}$ and write $\varphi^{(k)}\left(\bigotimes_{j=1}^d v^{(j)}\right) = \varphi^{(k)}\left(v^{(k)}\right)\cdot\bigotimes_{j\neq k} v^{(j)}$. This leads to the embedding

$$V_k'\subset L(\mathbf{V},\mathbf{V}_{[k]}). \tag{3.38b}$$

Another extreme case occurs for $\mathbf{V}_{[k]}'$. Here, the elementary tensors are $\boldsymbol{\varphi}^{[k]} = \bigotimes_{j\neq k}\varphi^{(j)}$. In this case, the image space is $\mathbb{K}\otimes\ldots\otimes\mathbb{K}\otimes V_k\otimes\mathbb{K}\otimes\ldots\otimes\mathbb{K}\cong V_k$.

Remark 3.55. Let V_j $(1\leq j\leq d)$ be vector spaces generating $\mathbf{V} := {}_a\bigotimes_{j=1}^d V_j$. For a fixed index $k\in\{1,\ldots,d\}$ define $\mathbf{V}_{[k]}$ by (3.21a). Then elementary tensors $\boldsymbol{\varphi}^{[k]} = \bigotimes_{j\neq k}\varphi^{(j)}\in\mathbf{V}_{[k]}'$ $(\varphi^{(j)}\in V_j')$ are considered as mappings from $L(\mathbf{V},V_k)$ via

$$\boldsymbol{\varphi}^{[k]}\left(\bigotimes_{j=1}^d v^{(j)}\right) := \left(\prod_{j\neq k}\varphi^{(j)}(v^{(j)})\right)\cdot v^{(k)}. \tag{3.39a}$$

This justifies the embedding

$$\mathbf{V}_{[k]}'\subset L(\mathbf{V},V_k). \tag{3.39b}$$

The generalisation of the above case is as follows. Let $\alpha\subset\{1,...,d\}$ be any nonempty subset and define the complement $\alpha^c := \{1,\ldots,d\}\backslash\alpha$. Then the embedding

$$\mathbf{V}_\alpha'\subset L(\mathbf{V},\mathbf{V}_{\alpha^c})\qquad\text{with } \mathbf{V}_\alpha := {}_a\bigotimes_{j\in\alpha} V_j \tag{3.39c}$$

is defined via

$$\left(\bigotimes_{j\in\alpha}\varphi^{(j)}\right)\left(\bigotimes_{j=1}^d v^{(j)}\right) := \left(\prod_{j\in\alpha}\varphi^{(j)}(v^{(j)})\right)\cdot\bigotimes_{j\in\alpha^c} v^{(j)}. \tag{3.39d}$$

Functionals from V_k' or $\mathbf{V}_{[k]}'$ are often used in proofs. As a demonstration we prove a statement about linear independence of the vectors representing the tensor.

Lemma 3.56. *Let* $\mathbf{V} = {}_a\bigotimes_{j=1}^d V_j$ *and* $\mathbf{v} := \sum_{i=1}^r \bigotimes_{j=1}^d v_i^{(j)}$. *If, for some index* $k \in \{1,\ldots,d\}$, *the vectors* $\{v_i^{(k)} : 1 \le i \le r\} \subset V_k$ *are linearly independent,* $\mathbf{v} = 0$ *implies*

$$\mathbf{v}_i^{[k]} = 0 \qquad \textit{for all } \mathbf{v}_i^{[k]} := \bigotimes_{j\neq k} v_i^{(j)} \quad (1 \le i \le r).$$

Vice versa, for linearly dependent vectors $\{v_i^{(k)}: 1 \le i \le r\}$, *there are* $\mathbf{v}_i^{[k]} \in \mathbf{V}_{[k]}$, *not all vanishing, with* $\mathbf{v} = \sum_{i=1}^r \mathbf{v}_i^{[k]} \otimes v_i^{(k)} = 0$.

Proof. 1) We give two proofs. The first one makes use of V_k'. For linearly independent $v_i^{(k)}$ there is a dual system of functionals $\varphi_i \in V_k'$ with $\varphi_\nu(v_\mu^{(k)}) = \delta_{\nu\mu}$ (cf. Definition 3.6). Using the embedding $V_k' \subset L(\mathbf{V},\mathbf{V}_{[k]})$ from (3.38b), we derive from $\mathbf{v} = 0$ that

$$0 = \varphi_\nu(\mathbf{v}) = \sum_{i=1}^r \varphi_\nu(v_i^{(k)})\mathbf{v}_i^{[k]} = \mathbf{v}_\nu^{[k]},$$

proving the first statement. Linearly dependent $\{v_i^{(k)} : 1 \le i \le r\}$ allow a nontrivial linear combination $\sum_{i=1}^r c_i v_i^{(k)} = 0$. Choose any $0 \neq \mathbf{v}^{[k]} \in \mathbf{V}_{[k]}$. Then $\mathbf{v}_i^{[k]} := c_i \mathbf{v}^{[k]}$ are not vanishing for all $1 \le i \le r$, but $\mathbf{v} = \sum_{i=1}^r \mathbf{v}_i^{[k]} \otimes v_i^{(k)} = \mathbf{v}^{[k]} \otimes \sum_{i=1}^r c_i v_i^{(k)} = 0$.

2) The second proof uses a functional $\boldsymbol{\varphi}^{[k]} \in \mathbf{V}_{[k]}'$. Assume that not all $\mathbf{v}_i^{[k]}$ vanish, say, $\mathbf{v}_1^{[k]} \neq 0$. Then there is some $\boldsymbol{\varphi}^{[k]} \in \mathbf{V}_{[k]}'$ with $\boldsymbol{\varphi}^{[k]}(\mathbf{v}_1^{[k]}) \neq 0$. Using the embedding $\mathbf{V}_{[k]}' \subset L(\mathbf{V}, V_k)$, we obtain $0 = \boldsymbol{\varphi}^{[k]}(\mathbf{v}) = \sum_{i=1}^r c_i v_i^{(k)}$ with $c_i := \boldsymbol{\varphi}^{[k]}(\mathbf{v}_i^{[k]})$. Since $c_1 \neq 0$, the $v_i^{(k)}$ cannot be linearly independent. □

3.3.2.3 Further Embeddings

Proposition 3.57. *(a) The algebraic tensor space* $V \otimes_a W'$ *can be embedded into* $L(W,V)$ *via*

$$w \in W \mapsto (v \otimes w')\,(w) := w'(w) \cdot v \in V. \tag{3.40a}$$

(b) Similarly, $V \otimes_a W$ *can be embedded into* $L(W',V)$ *via*

$$w' \in W' \mapsto (v \otimes w)\,(w') := w'(w) \cdot v \in W. \tag{3.40b}$$

(c) The embeddings from above show that

$$V \otimes_a W' \subset L(W,V) \qquad \textit{and} \qquad V \otimes_a W \subset L(W',V). \tag{3.40c}$$

Corollary 3.58. If $\dim(W) < \infty$, $V \otimes_a W \cong V \otimes_a W' \cong L(W,V)$ are isomorphic.

Proof. $\dim(W) < \infty$ implies $W' \cong W$ and therefore also $V \otimes_a W \cong V \otimes_a W'$. Thanks to (3.40c), $V \otimes_a W'$ is isomorphic to a *subspace* of $L(W, V)$. In order to prove $V \otimes_a W \cong L(W,V)$, we have to demonstrate that any $\varphi \in L(W,V)$ can be realised by some $\mathbf{x} \in V \otimes_a W'$. Let $\{w_i\}$ be a basis of W and $\{\omega_i\}$ a dual basis with $\omega_i(w_j) = \delta_{ij}$. Set $\mathbf{x} := \sum_i \varphi(w_i) \otimes \omega_i$. One easily verifies that $\mathbf{x}(w_i) = \varphi(w_i)$ in the sense of the embedding (3.40b). □

Remark 3.59. Let $M(V_1, \ldots, V_d) := \{\varphi : \times_{j=1}^d V_j \to \mathbb{K}$ multilinear$\}$ denote the set of multilinear mappings from $\times_{j=1}^d V_j$ to $\mathbb{K}$. If the spaces V_j are finite dimensional, the following inclusions become equalities.

(a) The inclusion ${}_a\bigotimes_{j=1}^d V_j \subset M(V_1', \ldots, V_d')$ is interpreted by

$$\left(\bigotimes_{j=1}^d v_j\right)(v_1', \ldots, v_d') := \prod_{j=1}^d v_j'(v_j) \in \mathbb{K}.$$

(b) The inclusion ${}_a\bigotimes_{j=1}^d V_j' \subset M(V_1, \ldots, V_d)$ is interpreted by

$$\left(\bigotimes_{j=1}^d v_j'\right)(v_1, \ldots, v_d) := \prod_{j=1}^d v_j'(v_j) \in \mathbb{K}.$$

3.4 Tensor Spaces with Algebra Structure

Throughout this section, all tensor spaces are *algebraic* tensor spaces. Therefore, we omit the index 'a' in $\otimes_a$.

So far, we have considered tensor products of vector spaces A_j $(1 \le j \le d)$. Now, we suppose that A_j possesses a further operation[9]

$$\circ : A_j \times A_j \to A_j,$$

which we call multiplication (to be quite precise, we should introduce individual symbols $\circ_j$ for each A_j). We require that

$$\begin{array}{ll} (a+b)\circ c = a\circ c + b\circ c & \text{for all } a,b,c \in A_j, \\ a\circ(b+c) = a\circ b + a\circ c & \text{for all } a,b,c \in A_j, \\ (\lambda a)\circ b = a\circ(\lambda b) = \lambda\cdot(a\circ b) & \text{for all } \lambda\in\mathbb{K} \text{ and all } a,b\in A_j, \\ 1\circ a = a\circ 1 = a & \text{for some } 1 \in V_j \text{ and all } a \in A_j. \end{array} \tag{3.41}$$

These rules define a (non-commutative) *algebra with unit element* 1. Ignoring the algebra structure, we define the tensor space $\mathbf{A} := {}_a\bigotimes_{j=1}^d A_j$ as before and establish an operation $\circ : \mathbf{A} \times \mathbf{A} \to \mathbf{A}$ by means of

[9] $\mu : A_j \times A_j \to A_j$ defined by $(a, b) \mapsto a \circ b$ is called the *structure map* of the algebra A_j.

$$\left(\bigotimes_{j=1}^{d} a_j\right) \circ \left(\bigotimes_{j=1}^{d} b_j\right) = \bigotimes_{j=1}^{d} (a_j \circ b_j), \qquad \mathbf{1} := \bigotimes_{j=1}^{d} 1$$

for elementary tensors. The first two axioms in (3.41) are used to define the multiplication for tensors of ${}_a\bigotimes_{j=1}^{d} A_j$.

Example 3.60. (a) Consider the matrix spaces $A_j := \mathbb{K}^{I_j \times I_j}$ (I_j: finite index sets). Here, $\circ$ is the matrix-matrix multiplication in $\mathbb{K}^{I_j \times I_j}$. The unit element 1 is the identity matrix I. Then $\mathbf{A} := \bigotimes_{j=1}^{d} A_j = \mathbb{K}^{\mathbf{I} \times \mathbf{I}}$ ($\mathbf{I} = I_1 \times \ldots \times I_d$) is the space containing the Kronecker matrices.

(b) The vector spaces A_j and $\mathbf{A}$ from Part (a) yield another algebra, if $\circ$ is defined by the Hadamard product (entry-wise product, cf. (4.72a)): $(a \circ b)_i = a_i \cdot b_i$ for $i \in I_j \times I_j$ and $a, b \in A_j$. The unit element 1 is the matrix with all entries being one.

(c) Let $A_j := C(I_j)$ be the set of continuous functions on the interval $I_j \subset \mathbb{R}$. A_j becomes an algebra with $\circ$ being the pointwise multiplication. The unit element is the function with constant value $1 \in \mathbb{K}$. Then $\mathbf{A} := \bigotimes_{j=1}^{d} A_j \subset C(\mathbf{I})$ contains multivariate functions on the product domain $\mathbf{I} = I_1 \times \ldots \times I_d$.

(d) Let $A_j := \ell_0(\mathbb{Z})$ (cf. (3.2)). The multiplication $\circ$ in A_j may be defined by the convolution $\star$ from (4.73a). The unit element 1 is the sequence with $1_i = \delta_{i0}$ for all $i \in \mathbb{Z}$ (δ_{i0} from (2.1)). This defines the d-dimensional convolution $\star$ in $\mathbf{A} := \bigotimes_{j=1}^{d} A_j = \ell_0(\mathbb{Z}^d)$.

The term '*tensor algebra*' is used for another algebraic construction. Let V be a vector space and consider the tensor spaces

$$\otimes^d V := \bigotimes_{j=1}^{d} V \tag{3.42}$$

(cf. Notation 3.23) and define the direct sum of $\otimes^d V$ for all $d \in \mathbb{N}_0$:

$$\mathcal{A}(V) := \sum_{d \in \mathbb{N}_0} \otimes^d V. \tag{3.43}$$

Elements of the tensor algebra are finite sums $\sum_{d \in \mathbb{N}_0} v_d$ with $v_d \in \otimes^d V$. The multiplicative structure is given by $\circ = \otimes$:

$$v_n \in \otimes^d V,\ v_m \in \otimes^m V \quad \mapsto \quad v_n \otimes v_m \in \otimes^{n+m} V$$

(cf. [76, Chap. III]). The algebra $\mathcal{A}(V)$ has the unit element $1 \in \mathbb{K} = \otimes^0 V \subset \mathcal{A}(V)$. The algebra $\mathcal{A}(\ell_0)$ will be used in §14.3.3.

If we replace $\mathbb{N}_0$ in (3.43) by $\mathbb{N}$,

$$\mathcal{A}(V) := \sum_{d \in \mathbb{N}} \otimes^d V$$

is an algebra *without* unit element.

3.5 Symmetric and Antisymmetric Tensor Spaces

3.5.1 Basic Definitions

In the following, all vector spaces V_j coincide and are denoted by V:

$$V := V_1 = V_2 = \ldots = V_d.$$

The d-fold tensor product is now denoted by $\mathbf{V} = \otimes^d V$, where $d \geq 2$ is required. Here we refer to the *algebraic* tensor space, i.e., $\mathbf{V} = \otimes_a^d V$. The completion to a Banach or Hilbert tensor space will be considered in §4.7.2.

A bijection $\pi : D \to D$ of the set $D := \{1, \ldots, d\}$ is called *permutation*. Let

$$P := \{\pi : D \to D \text{ bijective}\}$$

be the set of all permutations. Note that its cardinality

$$\#P = d!$$

increases fast with increasing d. $(P, \circ)$ is a group, where $\circ$ is defined by composition: $(\tau \circ \pi)(j) = \tau(\pi(j))$. The inverse of π is denoted by π^{-1}. As known from the description of determinants, $\mathrm{sign} : P \to \{-1, +1\}$ can be defined such that transpositions (pairwise permutations) have the sign -1, while the function sign is multiplicative: $\mathrm{sign}(\tau \circ \pi) = \mathrm{sign}(\tau) \cdot \mathrm{sign}(\pi)$.

A permutation $\pi \in P$ gives rise to a mapping $\mathbf{V} \to \mathbf{V}$ denoted again by π:

$$\pi : \bigotimes_{j=1}^{d} v^{(j)} \mapsto \bigotimes_{j=1}^{d} v^{(\pi^{-1}(j))}. \tag{3.44}$$

Exercise 3.61. For $\mathbf{v} \in \otimes^d \mathbb{K}^n$ show that $(\pi(\mathbf{v}))_{\mathbf{i}} = \mathbf{v}_{\pi(\mathbf{i})}$ for all $\mathbf{i} \in \{1,...,n\}^d$, where

$$\pi(i_1, \ldots, i_d) := \left(i_{\pi(1)}, \ldots, i_{\pi(d)}\right)$$

is the action of $\pi \in P$ onto a tuple from $\{1, \ldots, n\}^d$.

Definition 3.62. (a) $\mathbf{v} \in \mathbf{V} = \otimes^d V$ is *symmetric*, if $\pi(\mathbf{v}) = \mathbf{v}$ for all[10] $\pi \in P$.

(b) The symmetric tensor space is defined by

$$\mathfrak{S} := \mathfrak{S}(V) := \mathfrak{S}_d(V) := \otimes_{\mathrm{sym}}^d V := \{\mathbf{v} \in \mathbf{V} : \mathbf{v} \text{ symmetric}\}.$$

(c) A tensor $\mathbf{v} \in \mathbf{V} = \otimes^d V$ is *antisymmetric* (synonymously: 'skew symmetric'), if $\pi(\mathbf{v}) = \mathrm{sign}(\pi)\mathbf{v}$ for all $\pi \in P$.

(d) The antisymmetric tensor space is defined by

[10] Replacing all $\pi \in P$ by a subgroup of permutations (e.g., only permutations of certain positions) one can define tensor spaces with partial symmetry.

$$\mathfrak{A} := \mathfrak{A}(V) := \mathfrak{A}_d(V) := \otimes_{\mathrm{anti}}^d V := \{\mathbf{v} \in \mathbf{V} : \mathbf{v} \text{ antisymmetric}\}\,.$$

An equivalent definition of a symmetric [antisymmetric] tensor $\mathbf{v}$ is $\mathbf{v} = \pi(\mathbf{v})$ [$\mathbf{v} = -\pi(\mathbf{v})$] for all pair interchanges

$$\pi : (1, \ldots, i, \ldots, j, \ldots, d) \mapsto (1, \ldots, j, \ldots, i, \ldots, d)\,,$$

since all permutations from P are products of pairwise permutations.

For $d=2$ and $V=\mathbb{K}^n$, tensors from $\mathfrak{S}$ and $\mathfrak{A}$ correspond to symmetric matrices $(M_{ij} = M_{ji})$ and antisymmetric matrices $(M_{ij} = -M_{ji})$, respectively.

Proposition 3.63. *(a)* $\mathfrak{S}$ *and* $\mathfrak{A}$ *are subspaces of* $\mathbf{V} = \otimes^d V$.

(b) The projections $P_\mathfrak{S}$ *and* $P_\mathfrak{A}$ *from* $\mathbf{V}$ *onto* $\mathfrak{S}$ *and* $\mathfrak{A}$*, respectively, are given by*

$$P_\mathfrak{S}(\mathbf{v}) := \frac{1}{d!}\sum_{\pi\in P}\pi(\mathbf{v}), \qquad P_\mathfrak{A}(\mathbf{v}) := \frac{1}{d!}\sum_{\pi\in P}\mathrm{sign}(\pi)\pi(\mathbf{v}). \tag{3.45}$$

$P_\mathfrak{S}$ *is called* symmetriser*,* $P_\mathfrak{A}$ alternator.

Remark 3.64. (a) Exercise 3.61 shows that $(P_\mathfrak{S}(\mathbf{v}))_\mathbf{i} = \frac{1}{d!}\sum_{\pi\in P}\mathbf{v}_{\pi(\mathbf{i})}$ and $(P_\mathfrak{A}(\mathbf{v}))_\mathbf{i} = \frac{1}{d!}\sum_{\pi\in P}\mathrm{sign}(\pi)\mathbf{v}_{\pi(\mathbf{i})}$ for $\mathbf{v} \in \otimes^d\mathbb{K}^n$ and all $\mathbf{i} \in \{1, \ldots, n\}^d$.

(b) Let $V = \mathbb{K}^n$, $\mathbf{I}_n := \{1, \ldots, n\}^d$, and

$$\mathbf{I}_n^{\mathrm{sym}} := \{\mathbf{i} \in \mathbf{I}_n : 1 \le i_1 \le i_2 \le \ldots \le i_d \le n\}\,. \tag{3.46a}$$

Then, $\mathbf{v} \in \otimes_{\mathrm{sym}}^d\mathbb{K}^n$ is completely determined by the entries $\mathbf{v}_\mathbf{i}$ for $\mathbf{i} \in \mathbf{I}_n^{\mathrm{sym}}$. All other entries $\mathbf{v}_\mathbf{i}$ coincide with $\mathbf{v}_{\pi(\mathbf{i})}$, where $\pi \in P$ is chosen such that $\pi(\mathbf{i}) \in \mathbf{I}_n^{\mathrm{sym}}$.

(c) With V and $\mathbf{I}_n$ from Part (b) let

$$\mathbf{I}_n^{\mathrm{anti}} := \{\mathbf{i} \in \mathbf{I}_n : 1 \le i_1 < i_2 < \ldots < i_d \le n\}\,. \tag{3.46b}$$

Then $\mathbf{v} \in \otimes_{\mathrm{anti}}^d\mathbb{K}^n$ is completely determined by the entries $\mathbf{v}_\mathbf{i}$ with $\mathbf{i} \in \mathbf{I}_n^{\mathrm{anti}}$. For $\mathbf{i} \in \mathbf{I}_n\backslash\mathbf{I}_n^{\mathrm{anti}}$ two cases have to be distinguished. If $\mathbf{i}$ contains two equal elements, i.e., $i_j = i_k$ for some $j \ne k$, then $\mathbf{v}_\mathbf{i} = 0$. Otherwise, there is some $\pi \in P$ with $\pi(\mathbf{i}) \in \mathbf{I}_n^{\mathrm{anti}}$ and $\mathbf{v}_\mathbf{i} = \mathrm{sign}(\pi)\mathbf{v}_{\pi(\mathbf{i})}$.

Conclusion 3.65. *(a)* $\mathfrak{A}_d(\mathbb{K}^n) = \{0\}$ *for* $d > n$*, since* $\mathbf{I}_n^{\mathrm{anti}} = \emptyset$.

(b) For $n = d$*,* $\mathfrak{A}_d(\mathbb{K}^d) = \mathrm{span}\{P_\mathfrak{A}(\bigotimes_{j=1}^d e^{(j)})\}$ *is one-dimensional* ($e^{(j)}$*: unit vectors, cf. (2.2)), since* $\#\mathbf{I}_n^{\mathrm{anti}} = 1$.

Proposition 3.66. *For* $\dim(V) = n < \infty$ *and* $d \ge 2$ *the dimensions of* $\mathfrak{S}$ *and* $\mathfrak{A}$ *satisfy*

$$\dim(\mathfrak{A}_d(V)) = \tbinom{n}{d} \;<\; n^d/d! \;<\; \dim(\mathfrak{S}_d(V)) = \tbinom{n+d-1}{d}.$$

Bounds are $\dim(\mathfrak{A}_d(V)) \le \left(n - \frac{d-1}{2}\right)^d/d!$ *and* $\dim(\mathfrak{S}_d(V)) \le \left(n + \frac{d-1}{2}\right)^d/d!\,.$

Proof. $\mathfrak{S}$ is isomorphic to $\mathfrak{S}(\mathbb{K}^n)$. Remark 3.64c shows $\dim(\mathfrak{S}) = \#\mathbf{I}_n^{\mathrm{sym}}$. By induction one shows that $\#\mathbf{I}_n^{\mathrm{sym}} = \binom{n}{d}$. The proof for $\mathfrak{A}$ is analogous. □

As Proposition 3.66 shows, $\mathfrak{A}(V)$ has a smaller dimension than $\mathfrak{S}(V)$, but the other side of the coin is that V must be higher dimensional to form $\mathfrak{A}_d(V)$ of a certain dimension. As long as $n = \dim(V) < d$, $\mathfrak{A}_d(V) = \{0\}$ is zero-dimensional.

Let $N_{\mathfrak{A}} = \ker(P_{\mathfrak{A}})$, $N_{\mathfrak{S}} = \ker(P_{\mathfrak{S}})$ be the kernels and note that $\mathfrak{A} = \mathrm{range}(P_{\mathfrak{A}})$ and $\mathfrak{S} = \mathrm{range}(P_{\mathfrak{S}})$ are the images. Then the tensor space $\mathbf{V} = \bigotimes^d V$ admits the direct decomposition

$$\mathbf{V} = N_{\mathfrak{A}} \oplus \mathfrak{A} = N_{\mathfrak{S}} \oplus \mathfrak{S},$$

i.e., any $\mathbf{v} \in \mathbf{T}$ has a unique decomposition into $\mathbf{v} = \mathbf{v}_N + \mathbf{v}_X$ with either $\mathbf{v}_N \in N_{\mathfrak{A}}$, $\mathbf{v}_X \in \mathfrak{A}$ or $\mathbf{v}_N \in N_{\mathfrak{S}}$, $\mathbf{v}_X \in \mathfrak{S}$. Consequently, $\mathfrak{S}$ and $\mathfrak{A}$ are isomorphic to the quotient spaces (cf. §3.1.3):

$$\mathfrak{S} \cong \mathbf{V}/N_{\mathfrak{S}}, \qquad \mathfrak{A} \cong \mathbf{V}/N_{\mathfrak{A}}. \tag{3.47}$$

$\mathbf{V}/N_{\mathfrak{A}} = \wedge^d V$ is called the d-th exterior power of V, and $v^{(1)} \wedge v^{(2)} \wedge \ldots \wedge v^{(d)}$ is the isomorphic image of $P_{\mathfrak{A}}(v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes v^{(d)})$. The operation $\wedge$ is called the *exterior product*. Analogously, $\mathbf{V}/N_{\mathfrak{S}} = \vee^d V$ is called the d-th symmetric power of V, and $v^{(1)} \vee v^{(2)} \vee \ldots \vee v^{(d)}$ is the isomorphic image of $P_{\mathfrak{S}}(v^{(1)} \otimes v^{(2)} \otimes \ldots \otimes v^{(d)})$.

In the context of (anti)symmetric tensor spaces, a mapping $\mathbf{A} \in L(\mathbf{V}, \mathbf{V})$ is called *symmetric*, if $\mathbf{A}$ commutes with all $\pi \in P$, i.e., $\mathbf{A}\pi = \pi\mathbf{A}$. This property implies that $\mathbf{A}P_{\mathfrak{S}} = P_{\mathfrak{S}}\mathbf{A}$ and $\mathbf{A}P_{\mathfrak{A}} = P_{\mathfrak{A}}\mathbf{A}$, and proves the following result.

Remark 3.67. If $\mathbf{A} \in L(\mathbf{V}, \mathbf{V})$ is symmetric, $\mathfrak{S}$ and $\mathfrak{A}$ are invariant under $\mathbf{A}$, i.e., the restrictions of $\mathbf{A}$ to $\mathfrak{S}$ and $\mathfrak{A}$ belong to $L(\mathfrak{S}, \mathfrak{S})$ and $L(\mathfrak{A}, \mathfrak{A})$, respectively.

The Hadamard product $\odot$ will be explained in (4.72a). In the case of functions, it is the usual pointwise product.

Exercise 3.68. Let $\mathbf{s}, \mathbf{s}' \in \mathfrak{S}$ and $\mathbf{a}, \mathbf{a}' \in \mathfrak{A}$. Show that $\mathbf{s} \odot \mathbf{s}', \mathbf{a} \odot \mathbf{a}' \in \mathfrak{S}$, whereas $\mathbf{s} \odot \mathbf{a}, \mathbf{a} \odot \mathbf{s} \in \mathfrak{A}$.

3.5.2 Quantics

The term ‘quantics’ introduced by Arthur Cayley [33] in 1854 is used for homogeneous polynomials in multiple variables, i.e., polynomials in the variables x_i $(i \in B)$ which are a finite sum of terms $a_{\boldsymbol{\nu}} \mathbf{x}^{\boldsymbol{\nu}} = a_{\boldsymbol{\nu}} \prod_{i \in B} x_i^{\nu_i}$ with multi-indices $\boldsymbol{\nu}$ of length $|\boldsymbol{\nu}| := \sum_{i \in B} \nu_i = d \in \mathbb{N}_0$ and $a_{\boldsymbol{\nu}} \in \mathbb{K}$. Such quantics have the property

$$p(\lambda \mathbf{x}) = \lambda^d p(\mathbf{x}).$$

Proposition 3.69. *Let V be a vector space with the (algebraic) basis $\{b_i : i \in B\}$. Then the algebraic symmetric tensor space $\mathfrak{S}_d(V)$ is isomorphic to the set of quantics in the variables $\{x_i : i \in B\}$.*

This statement holds also for $\dim(V) = \#B = \infty$. Note that the infinite product $\mathbf{x}^{\boldsymbol{\nu}} = \prod_{i\in B} x_i^{\nu_i}$ makes sense, since at most d exponents ν_i are different from zero.

For the proof of Proposition 3.69 consider a general tensor $\mathbf{v} \in \bigotimes_a^d V$. Using the basis $\{b_i : i \in B\}$, we may write $\mathbf{v} = \sum_{\mathbf{i}\in B^d} \mathbf{a_i} \bigotimes_{j=1}^d b_{i_j}$ (almost all $\mathbf{a_i}$ vanish). Therefore, any symmetric tensor can be written as $P_{\mathfrak{S}}\mathbf{v} = \sum_{\mathbf{i}\in B^d} P_{\mathfrak{S}}(\mathbf{a_i} \bigotimes_{j=1}^d b_{i_j})$. The isomorphism into the quantics of degree d is given by

$$\Phi : \; P_{\mathfrak{S}}\left(\mathbf{a_i} \bigotimes_{j=1}^{d} b_{i_j}\right) \quad \mapsto \quad \mathbf{a_i} \prod_{j=1}^{d} x_{i_j}.$$

Note that

$$\prod_{j=1}^{d} x_{i_j} = \prod_{i\in B} x_i^{\nu_i} \quad \text{with } \nu_i := \#\{j \in \{1,\ldots,d\} : i_j = i\} \quad \text{for all } i \in B.$$

The symmetry of $P_{\mathfrak{S}}\mathbf{v}$ corresponds to the fact that all factors x_i commute.

Above we have used the range of $P_{\mathfrak{S}}$ to define $\mathfrak{S}_d(V)$. In the case of $d = 2$, e.g., $\mathbf{v} := P_{\mathfrak{S}}(a \otimes b) = a \otimes b + b \otimes a$ represents a symmetric tensor. Instead, one may use $\mathbf{v} = \frac{1}{2} \otimes^2 (a + b) - \frac{1}{2} \otimes^2 (a - b)$. In the latter representation, each term itself is symmetric. In general, $\mathfrak{S}_d(V)$ can be generated by d-fold tensor products $\otimes^d v$:

$$\mathfrak{S}(V) = \operatorname{span}\{\otimes^d v : v \in V\}.$$

In the language of quantics this means that the polynomial is written as sum of d-th powers $\left(\sum_{i=1}^n a_i x_i\right)^d$ of linear forms. The decomposition of a homogeneous polynomial into this special form is addressed by Brachat - Comon - Mourrain - Tsigaridas [24] (see $\operatorname{rank}_{\text{sym}}$ on page 65).

3.5.3 Determinants

Since tensor products are related to multilinear forms (cf. Proposition 3.22) and the determinant is a special antisymmetric multilinear form, it is not surprising to find relations between antisymmetric tensors and determinants. We recall that the *determinant* $\det(A)$ of a matrix $A \in \mathbb{K}^{d\times d}$ equals

$$\det(A) = \sum_{\pi\in P} \operatorname{sign}(\pi) \prod_{j=1}^{d} a_{j,\pi(j)}. \tag{3.48}$$

The building block of usual tensors are the elementary tensors $\bigotimes_{j=1}^d u^{(j)}$. For antisymmetric tensors we have to use their antisymmetrisation $P_{\mathfrak{A}}\big(\bigotimes_{j=1}^d u^{(j)}\big)$.

Lemma 3.70. *An antisymmetrised elementary tensor* $\mathbf{v} := P_{\mathfrak{A}}\big(\bigotimes_{j=1}^{d} v^{(j)}\big)$ *with* $v^{(j)} \in V := \mathbb{K}^n$ *has the entries*

$$\mathbf{v}[i_1, \ldots, i_d] = \frac{1}{d!} \det \begin{bmatrix} v_{i_1}^{(1)} & v_{i_1}^{(2)} & \cdots & v_{i_1}^{(d)} \\ v_{i_2}^{(1)} & v_{i_2}^{(2)} & \cdots & v_{i_2}^{(d)} \\ \vdots & \vdots & \ddots & \vdots \\ v_{i_d}^{(1)} & v_{i_d}^{(2)} & \cdots & v_{i_d}^{(d)} \end{bmatrix}.$$

Proof. Set $\mathbf{e} := \bigotimes_{j=1}^{d} v^{(j)}$. By definition (3.48), the right-hand side equals

$$\frac{1}{d!} \sum_{\pi \in P} \operatorname{sign}(\pi) \prod_{j=1}^{d} v_{i_{\pi(j)}}^{(j)} = \frac{1}{d!} \sum_{\pi \in P} \operatorname{sign}(\pi) \mathbf{e}_{\pi(\mathbf{i})} = (P_{\mathfrak{A}}(\mathbf{e}))_{\mathbf{i}} = \mathbf{v}_{\mathbf{i}} = \mathbf{v}[i_1, ..., i_d],$$

proving the assertion. □

For $V = \mathbb{K}^d$ the antisymmetric space $\mathfrak{A}_d(\mathbb{K}^d)$ is one-dimensional (cf. Conclusion 3.65b). Therefore, any transformation by $\mathbf{A} = \otimes^d A$ is identical to a multiple of the identity.

Remark 3.71. Let $V = \mathbb{K}^d$. For any $A \in \mathbb{K}^{d \times d}$ and $\mathbf{v} = P_{\mathfrak{A}}(\bigotimes_{j=1}^{d} v^{(j)}) \in \mathfrak{A}_d(\mathbb{K}^d)$ one has

$$A(\mathbf{v}) := P_{\mathfrak{A}}\left(\bigotimes_{j=1}^{d} A v^{(j)}\right) = \det(A)\, P_{\mathfrak{A}}\left(\bigotimes_{j=1}^{d} v^{(j)}\right).$$

Proof. By Conclusion 3.65, tensors $\mathbf{u}$ from $\mathfrak{A}(\mathbb{K}^d)$ are determined by $\mathbf{u}[1, \ldots, d]$. Lemma 3.70 shows that $\mathbf{v}[1,..., d] = \frac{1}{d!} \det(M)$, where $M = [v^{(1)},..., v^{(d)}] \in \mathbb{K}^{d \times d}$, and that $A(\mathbf{v})_{1,...,d} = \frac{1}{d!} \det(AM) = \det(A) \cdot \frac{1}{d!} \det(M) = \det(A) \cdot \mathbf{v}[1,..., d]$. □

For function spaces V (e.g., $V = L^2(\mathbb{R})$, $V = C[0, 1]$, etc.) there is an analogue of Lemma 3.70.

Lemma 3.72. *For functions* $f_1,..., f_d \in V$ *the antisymmetrisation of the elementary tensor*

$$F = \bigotimes_{j=1}^{d} f_j, \qquad \text{i.e., } F(x_1, \ldots, x_d) = \prod_{j=1}^{d} f_j(x_j),$$

yields $G := P_{\mathfrak{A}}(F)$ *with*

$$G(x_1, \ldots, x_d) = \frac{1}{d!} \det \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_d(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_d(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_d) & f_2(x_d) & \cdots & f_d(x_d) \end{bmatrix}.$$

G is also called the Slater determinant *of* $f_1, \ldots, f_d$.

Part II
Functional Analysis of Tensor Spaces

Algebraic tensor spaces yield a suitable fundament for the finite dimensional case. But even in the finite dimensional case we want to formulate approximation problems, which require the introduction of a topology. Topological tensor spaces are a subject of functional analysis.

Standard examples of infinite dimensional tensor spaces are function spaces, since multivariate functions can be regarded as tensor products of univariate ones. To obtain a Banach tensor space, we need the completion with respect to a norm, which is not fixed by the normed spaces generating the tensor space. The scale of norms is a particular topic of the discussion of Banach tensor spaces in *Chap. 4*. A particular, but important case are Hilbert tensor spaces.

Chapter 5 has a stronger connection to algebraic tensor spaces than to topological ones. But, in particular, the technique of matricisation is a prerequisite required in Chap. 6.

In *Chap. 6* we discuss the so-called minimal subspaces which are important for the analysis of the later tensor representations in Part III.

Chapter 4
Banach Tensor Spaces

Abstract The discussion of topological tensor spaces has been started by Schatten [167] and Grothendieck [79, 80]. In *Sect. 4.2* we discuss the question how the norms of V and W are related to the norm of $V \otimes W$.
From the viewpoint of functional analysis, tensor spaces of order 2 are of particular interest, since they are related to certain operator spaces (cf. §4.2.13). However, for our applications we are more interested in tensor spaces of order ≥ 3. These spaces are considered in *Sect. 4.3*.
As preparation for the aforementioned sections and later ones, we need more or less well-known results from Banach space theory, which we provide in *Sect. 4.1*.
Section 4.4 discusses the case of Hilbert spaces. This is important, since many applications are of this kind. Many of the numerical methods require scalar products. The reason is that, unfortunately, the solution of approximation problems with respect to general Banach norms is much more involved than those with respect to a scalar product.

4.1 Banach Spaces

4.1.1 Norms

In the following we consider vector spaces X over one of the fields $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The topological structure will be generated by norms. We recall the axioms of a *norm* on X:

$$\begin{array}{ll} \|\cdot\| : X \to [0,\infty)\,, & \\ \|x\| = 0 & \text{if and only if } x = 0, \\ \|\lambda x\| = |\lambda|\,\|x\| & \text{for all } \lambda \in \mathbb{K} \text{ and } x \in X, \\ \|x+y\| \leq \|x\| + \|y\| & \text{for all } x, y \in X \quad \textit{(triangle inequality)}. \end{array} \tag{4.1}$$

The map $\|\cdot\| : X \to [0,\infty)$ is continuous because of the *inverse triangle inequality*

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_4,

$$\big|\, \|x\| - \|y\| \,\big| \le \|x - y\| \qquad \text{for all } x, y \in X. \tag{4.2}$$

Combining a vector space X with a norm defined on X, we obtain a *normed vector space* denoted by the pair $(X, \|\cdot\|)$. If there is no doubt about the choice of norm, the notation $(X, \|\cdot\|)$ is shortened by X.

A vector space X may be equipped with two different norms, i.e., $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ may be two different normed spaces although the set X is the same in both cases. For the set of all norms on X we can define a *semi-ordering*. A norm $\|\cdot\|_1$ is called *weaker* (or *not stronger*) than $\|\cdot\|_2$, in symbolic notation

$$\|\cdot\|_1 \lesssim \|\cdot\|_2$$

(equivalently: $\|\cdot\|_2$ *stronger* than $\|\cdot\|_1$, or $\|\cdot\|_2 \gtrsim \|\cdot\|_1$), if there is a constant C such that

$$\|x\|_1 \le C\, \|x\|_2 \quad \text{for all } x \in X.$$

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are *equivalent,* in symbolic notation

$$\|\cdot\|_1 \sim \|\cdot\|_2\,,$$

if $\|\cdot\|_1 \lesssim \|\cdot\|_2 \lesssim \|\cdot\|_1$, or equivalently, if there are $C_1, C_2 \in (0, \infty)$ with

$$\frac{1}{C_1}\, \|x\|_1 \le \|x\|_2 \le C_2\, \|x\|_1 \qquad \text{for all } x \in X.$$

4.1.2 Basic Facts about Banach Spaces

A sequence $x_i \in X$ in a normed vector space $(X, \|\cdot\|)$ is called a *Cauchy sequence* (with respect to $\|\cdot\|$) if

$$\sup_{i,j \ge n} \|x_i - x_j\| \to 0 \qquad \text{as } n \to \infty.$$

A normed vector space $(X, \|\cdot\|)$ is called a *Banach space* if it is also *complete* (with respect to $\|\cdot\|$). Completeness means that any Cauchy sequence $x_i \in X$ has a *limit* $x := \lim_{i\to\infty} x_i \in X$ (i.e., $\|x - x_i\| \to 0$).

A subset $X_0 \subset X$ of a Banach space $(X, \|\cdot\|)$ is *dense*, if for any $x \in X$ there is a sequence $x_i \in X_0$ with $\|x - x_i\| \to 0$. An equivalent criterion is that for any $\varepsilon > 0$ and any $x \in X$ there is some $x_\varepsilon \in X_0$ with $\|x - x_\varepsilon\| \le \varepsilon$. A dense subset may be, in particular, a dense subspace. An important property of dense subsets is noted in the next remark.

Remark 4.1. Let $\Phi : X_0 \to Y$ be a continuous mapping, where X_0 is dense in the Banach space $(X, \|\cdot\|)$ and $(Y, \|\cdot\|_Y)$ is some Banach space. Then there is a unique continuous extension $\overline{\Phi} : X \to Y$ with $\Phi(x) = \overline{\Phi}(x)$ for all $x \in X$.

If a normed vector space $(X, \|\cdot\|)$ is not complete, it has a unique completion $\big(\overline{X}, |\!|\!| \cdot |\!|\!|\big)$—up to isomorphisms—such that X is a dense subspace of the Banach space $\overline{X}$ and $|||\cdot|||$ is the continuous extension of $\|\cdot\|$ (note that $\varPhi := \|\cdot\| : X \to \mathbb{R}$ is a particular continuous function into the Banach space $(Y, \|\cdot\|_Y) = (\mathbb{R}, |\cdot|)$ with the extension $\overline{\varPhi} = |||\cdot|||$ being again a norm, cf. Remark 4.1). In the following, we shall use the same symbol $\|\cdot\|$ for the norm on the closure $\overline{X}$ (instead of $|||\cdot|||$). Furthermore, we write again $\varPhi$ instead of $\overline{\varPhi}$.

Remark 4.2. Let $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ be normed vector spaces with identical sets X, but different norms. Then we have to distinguish between the completions $(X_1, \|\cdot\|_1)$ and $(X_2, \|\cdot\|_2)$ with respect to the corresponding norms. Identity $X_1 = X_2$ holds if and only if $\|\cdot\|_1 \sim \|\cdot\|_2$. If $\|\cdot\|_1 \lesssim \|\cdot\|_2$, then[1] $X_1 \supset X_2 \supset X$.

Definition 4.3. A Banach space is *separable*, if there is a countable dense subset.

Definition 4.4. A closed subspace U of a Banach space X is called *direct* or *complemented*, if there is a subspace W such that $X = U \oplus W$ is a direct sum (cf. [108, p. 4]). The *Grassmannian*[2] $\mathbb{G}(X)$ is the set of all direct subspaces of X (cf. [110]).

Closedness of U in $X = U \oplus W$ implies that also W is closed.

4.1.3 Examples

Let I be a (possibly infinite) index set with $\#I \le \#\mathbb{N}$ (i.e., I finite or countable). Examples for I are $I = \{1, \ldots, n\}$, $I = \mathbb{N}$ or $I = \mathbb{Z}$ or products of these sets. The vector spaces $\ell(I) = \mathbb{K}^{I}$ and $\ell_0(I) \subset \ell(I)$ are already explained in Example 3.1.

Example 4.5. $\ell^p(I)$ consists of all $a \in \ell(I)$ with bounded norm

$$\begin{aligned} \|a\|_{\ell^p(I)} &:= \|a\|_p := \Big(\textstyle\sum_{i\in I} |a_i|^p\Big)^{1/p} && \text{for } p \in [1,\infty) \text{ or} \\ \|a\|_{\ell^\infty(I)} &:= \|a\|_\infty := \sup\{|a_i| : i \in I\} && \text{for } p = \infty. \end{aligned} \tag{4.3}$$

$(\ell^p(I), \|\cdot\|_p)$ is a Banach space for all $1 \le p \le \infty$.

Remark 4.6. (a) For $p < \infty$, $\ell_0(I)$ is a dense subspace of $\ell^p(I)$.
(b) For an infinite set I, the completion of $\ell_0(I)$ under the norm $\|\cdot\|_\infty$ yields the proper subset $(c_0(I), \|\cdot\|_\infty) \subsetneqq \ell^\infty(I)$ of zero sequences:

$$c_0(I) = \{a \in \ell^\infty(I) : \lim_{\nu\to\infty} a_{i_\nu} = 0\}, \tag{4.4}$$

where i_ν describes any enumeration of the countable set I.

Proof. 1) For finite I, $\ell_0(I) = \ell(I)$ holds and nothing is to be proved. In the following, I is assumed to be infinite and countable. One may choose any enumeration $I = \{i_\nu : \nu \in \mathbb{N}\}$, but for simplicity we write ν instead of i_ν.

[1] More precisely, the completion X_1 can be constructed such that $X_1 \supset X_2$.

[2] The correct spelling of the name is Graßmann, Hermann Günther.

2) Case $p\in[1,\infty)$. Since $\sum_{\nu=1}^{\infty}|a_\nu|^p<\infty$, for any $\varepsilon>0$, there is a ν_ε such that $\sum_{\nu>\nu_\varepsilon}|a_\nu|^p\le\varepsilon^p$, which proves $\|a-a'\|_p\le\varepsilon$ for $a'=(a'_i)_{i\in I}$ with $a'_\nu:=a_\nu$ for the finitely many $1\le\nu\le\nu_\varepsilon$ and with $a'_\nu:=0$ otherwise. Since $a'\in\ell_0(I)$, $\ell_0(I)$ is dense in $\ell^p(I)$.

3a) Case $p=\infty$. Assume $\ell_0(I)\ni a^{(n)}\to a\in\ell^\infty(I)$ with respect to $\|\cdot\|_\infty$. For any $n\in\mathbb{N}$ there is a ν_n such that $a_\nu^{(n)}=0$ for $\nu>\nu_n$. For any $\varepsilon>0$, there is an n_ε such that $\|a^{(n)}-a\|_\infty\le\varepsilon$ for $n\ge n_\varepsilon$. Hence, $|a_\nu|=|a_\nu^{(n_\varepsilon)}-a_\nu|\le\|a^{(n_\varepsilon)}-a\|_\infty\le\varepsilon$ for $\nu>\nu_{n_\varepsilon}$ proves $\lim_\nu a_\nu=0$, i.e., $a\in c_0(I)$. This proves $\overline{\ell_0(I)}\subset c_0(I)$.

3b) Assume $a\in c_0(I)$. For any $n\in\mathbb{N}$, there is a ν_n such that $|a_\nu|\le 1/n$ for $\nu>\nu_n$. Define $a^{(n)}\in\ell_0(I)$ by $a_\nu^{(n)}=a_\nu$ for $1\le\nu\le\nu_n$ and $a_\nu^{(n)}=0$ otherwise. Then $\|a^{(n)}-a\|_\infty\le 1/n$ hold, i.e., $a^{(n)}\to a$ and the reverse inclusion $c_0(I)\subset\overline{\ell_0(I)}$ holds. □

Example 4.7. Let $D\subset\mathbb{R}^m$ be a domain. (a) Assume $1\le p<\infty$. Then[3]

$$L^p(D):=\left\{f:D\to\mathbb{K}\text{ measurable and }\int_D|f(x)|^p\,dx<\infty\right\}$$

defines a Banach space with the norm $\|f\|_{L^p(D)}=\|f\|_p=\left(\int_D|f(x)|^p\,dx\right)^{1/p}$.
(b) For $p=\infty$, $L^\infty(D):=\{f:D\to\mathbb{K}\text{ measurable and }\|f\|_\infty<\infty\}$ equipped with the norm $\|f\|_\infty:=\operatorname*{ess\,sup}_{x\in D}|f(x)|$ is a Banach space.

Example 4.8. Let $D\subset\mathbb{R}^m$ be a domain. The following sets of continuous functions and n-times continuously differentiable functions are Banach spaces.
(a) $C(D)=C^0(D):=\{f:D\to\mathbb{K}\text{ with }\|f\|_{C(D)}<\infty\}$ with

$$\|f\|_{C(D)}=\sup\{|f(x)|:x\in D\}.$$

(b) Let $n\in\mathbb{N}_0$ and define $C^n(D):=\{f:I\to\mathbb{K}\text{ with }\|f\|_{C^n(D)}<\infty\}$ with the norm $\|f\|_{C^n(D)}=\max_{|\nu|\le n}\|\partial^{\boldsymbol{\nu}}f\|_{C(D)}$, where the maximum is taken over all multi-indices $\boldsymbol{\nu}=(\nu_1,\ldots,\nu_m)\in\mathbb{N}_0^m$ with $|\boldsymbol{\nu}|:=\sum_{i=1}^m\nu_i$. The mixed partial derivatives are abbreviated by

$$\partial^{\boldsymbol{\nu}}:=\prod\nolimits_{i=1}^m\left(\frac{\partial}{\partial x_i}\right)^{\nu_i}. \tag{4.5}$$

$\partial^{\boldsymbol{\nu}}f$ denotes the weak derivative (considered as a distribution).

Example 4.9. Let $D\subset\mathbb{R}^m$ be a domain. The Sobolev space

$$H^{1,2}(D):=\left\{f:D\to\mathbb{K}\text{ with }\|f\|_{H^{1,2}(D)}<\infty\right\}$$

is a Banach space, where

$$\|f\|_{H^{1,2}(D)}:=\sqrt{\sum_{|\boldsymbol{\nu}|\le1}\|\partial^{\boldsymbol{\nu}}f\|^2_{L^2(D)}}=\sqrt{\|f\|^2_{L^2(D)}+\sum_{i=1}^m\left\|\frac{\partial}{\partial x_i}f\right\|^2_{L^2(D)}}.$$

[3] To be precise, we have to form the quotient space $\{\ldots\}/N$ with $N:=\{f:f=0\text{ on }D\backslash S$ for all S with measure $\mu(S)=0\}$.

4.1.4 Operators

We recall the definition (3.6) of the set $L(X,Y)$ of linear mappings. The following remark states basic facts about *continuous* linear mappings, which are also called *operators*. The set of operators will be denoted by $\mathcal{L}(X,Y)$.

Remark 4.10. Let $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$ be normed spaces, and $\Phi \in L(X,Y)$.
(a) The following conditions (i) and (ii) are equivalent:
(i) Φ is continuous;
(ii) Φ is bounded, i.e., $\sup\{\|\Phi(x)\|_Y : x \in X \text{ with } \|x\|_X \leq 1\} < \infty$.
(b) The supremum from (ii) defines the operator norm[4]

$$\|\Phi\|_{Y\leftarrow X} := \sup\{\|\Phi(x)\|_Y : x \in X \text{ with } \|x\|_X \leq 1\}. \tag{4.6a}$$

A simple consequence is

$$\|\Phi(x)\|_Y \leq \|\Phi\|_{Y\leftarrow X}\,\|x\|_X \qquad \text{for all } x \in X. \tag{4.6b}$$

Another notation for the boundedness of Φ reads as follows: there is a constant $C < \infty$ such that

$$\|\Phi(x)\|_Y \leq C\,\|x\|_X \qquad \text{for all } x \in X. \tag{4.6c}$$

Then the minimal possible C in (4.6c) coincides with $\|\Phi\|_{Y\leftarrow X}$ (cf. (4.6b)).
(c) Let $\overline{X}$ and $\overline{Y}$ be the completions of X, Y so that $\left(\overline{X},\|\cdot\|_X\right)$ and $\left(\overline{Y},\|\cdot\|_Y\right)$ are Banach spaces. Then the continuation $\overline{\Phi} : \overline{X} \to \overline{Y}$ discussed in Remark 4.1 has an identical operator norm: $\|\overline{\Phi}\|_{\overline{Y}\leftarrow\overline{X}} = \|\Phi\|_{Y\leftarrow X}$. Because of this equality, we shall not distinguish between $\|\cdot\|_{\overline{Y}\leftarrow\overline{X}}$ and $\|\cdot\|_{Y\leftarrow X}$.
(d) The set of continuous linear mappings (operators) from X into Y is denoted by $\mathcal{L}(X,Y)$. Together with (4.6a),

$$(\mathcal{L}(X,Y),\|\cdot\|_{Y\leftarrow X}) \tag{4.6d}$$

forms a normed space. If $(Y,\|\cdot\|_Y)$ is a Banach space, also $(\mathcal{L}(X,Y),\|\cdot\|_{Y\leftarrow X})$ is a Banach space.

Proof. 1) If the boundedness (ii) holds, the definition of $\|\Phi\|_{Y\leftarrow X}$ makes sense and (4.6b) follows by linearity of Φ.

2) (boundedness ⇒ continuity) For any $\varepsilon > 0$ set $\delta := \varepsilon / \|\Phi\|_{Y\leftarrow X}$. Whenever $\|x' - x''\|_X \leq \delta$, we conclude that

$$\|\Phi(x') - \Phi(x'')\|_Y = \|\Phi(x'-x'')\|_Y \leq \|\Phi\|_{Y\leftarrow X}\,\|x'-x''\|_X \leq \|\Phi\|_{Y\leftarrow X}\,\delta = \varepsilon,$$

i.e., Φ is continuous.

[4] In (4.6a) one may replace $\|x\|_X \leq 1$ by $\|x\|_X = 1$, as long as the vector space is not the trivial space $X = \{0\}$ containing no x with $\|x\|_X = 1$. Since this trivial case is of minor interest, we will often use $\|x\|_X = 1$ instead.

3) The direction ‘continuity ⇒ boundedness’ is proved indirectly. Assume that (ii) is not valid. Then there are x_i with $\|x_i\|_X \leq 1$ and $\alpha_i := \|\Phi(x_i)\|_Y \to \infty$. The scaled vectors $x_i' := \frac{1}{1+\alpha_i} x_i$ satisfy $x_i' \to 0$, whereas $\|\Phi(x_i')\|_Y = \frac{\alpha_i}{1+\alpha_i} \to 1 \neq 0 = \|0\|_Y = \|\Phi(0)\|_Y$. Hence Φ is not continuous. Steps 2) and 3) prove Part (a).

4) For the last part of (d) let $\Phi_i \in \mathcal{L}(X,Y)$ be a Cauchy sequence, i.e.,

$$\sup_{i,j \geq n} \|\Phi_i - \Phi_j\|_{Y \leftarrow X} \to 0 \qquad \text{as } n \to \infty.$$

For any $x \in X$, also the images $y_i := \Phi_i(x)$ form a Cauchy sequence, as seen from $\|y_i - y_j\|_Y \leq \|\Phi_i - \Phi_j\|_{Y \leftarrow X} \|x\|_X \to 0$. By the Banach space property, $y := \lim y_i \in Y$ exists uniquely giving rise to a mapping $\Phi(x) := y$. One verifies that $\Phi \colon X \to Y$ is linear and bounded with $\|\Phi_i - \Phi\|_{Y \leftarrow X} \to 0$, i.e., $\Phi \in \mathcal{L}(X,Y)$. □

In the following, ‘*operator*’ is used as a shorter name for ‘continuous linear map’.

In later proofs we shall use the following property of the supremum in (4.6a).

Remark 4.11. For all operators $\Phi \in \mathcal{L}(X,Y)$ and all $\varepsilon > 0$, there is an $x_\varepsilon \in X$ with $\|x_\varepsilon\|_X \leq 1$ such that

$$\|\Phi\|_{Y \leftarrow X} \leq (1+\varepsilon)\, \|\Phi(x_\varepsilon)\|_Y \quad \text{and} \quad \|\Phi(x_\varepsilon)\|_Y \geq (1-\varepsilon)\, \|\Phi\|_{Y \leftarrow X}.$$

A subset K of a normed space is called *compact*, if any sequence $x_\nu \in K$ possesses a convergent subsequence with limit in K.

Definition 4.12. An operator $\Phi \in \mathcal{L}(X,Y)$ is called *compact*, if the unit ball $B := \{x \in X : \|x\|_X \leq 1\}$ is mapped onto $\Phi(B) := \{\Phi(x) : x \in B\} \subset Y$ and the closure $\overline{\Phi(B)}$ is a compact subset of Y. $\mathcal{K}(X,Y)$ denotes the set of *compact* operators.

The *approximation property* of a Banach space, which will be explained in Definition 4.81, is also related to compactness.

Finally, we add results about projections (cf. Definition 3.4). The next statement follows from Banach’s closed map theorem.

Lemma 4.13. *If $X = U \oplus W$ is the direct sum of closed subspaces, the decomposition $x = u + w$ ($u \in U$, $w \in W$) of any $x \in X$ defines projections $P_1, P_2 \in \mathcal{L}(X,X)$ onto these subspaces by $P_1 x = u$ and $P_2 x = w$.*

Theorem 4.14. *Let $Y \subset X$ be a subspace of a Banach space X with* $\dim(Y) \leq n$. *Then there exists a projection $\Phi \in \mathcal{L}(X,X)$ onto Y such that*

$$\|\Phi\|_{X \leftarrow X} \leq \sqrt{n}.$$

The proof can be found in DeVore-Lorentz [46, Chap. 9, §7] or [145, Proposition 12.14]. The bound is sharp for general Banach spaces, but can be improved to

$$\|\Phi\|_{X \leftarrow X} \leq n^{\left|\frac{1}{2} - \frac{1}{p}\right|} \quad \text{for } X = L^p. \tag{4.7}$$

4.1.5 Dual Spaces

A trivial example of a Banach space $(Y, \|\cdot\|_Y)$ is the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ with the absolute value $|\cdot|$ as norm. The operators $X \to \mathbb{K}$ are called *continuous functionals* or *continuous forms*. The Banach space $\mathcal{L}(X, \mathbb{K})$ is called the (continuous) dual of X and denoted by[5]

$$X^* := \mathcal{L}(X, \mathbb{K}).$$

The norm $\|\cdot\|_{X^*}$ of X^* follows from the general definition (4.6a):

$$\begin{aligned} \|\varphi\|_{X^*} &= \sup\{|\varphi(x)| : x \in X \text{ with } \|x\|_X \le 1\} \\ &= \sup\{|\varphi(x)| \,/\, \|x\|_X : 0 \ne x \in X\}. \end{aligned} \tag{4.8}$$

Instead of $\|\cdot\|_{X^*}$, we also use the notation $\|\cdot\|_X^*$ (meaning the dual norm corresponding to $\|\cdot\|_X$).

The next statement is one of the many versions of the *Hahn-Banach Theorem* (cf. Yosida [198, §IV.6]).

Theorem 4.15. *Let $(X, \|\cdot\|_X)$ be a normed linear space and $U \subset X$ a subspace. If a linear form is bounded on U, it can be extended to a continuous functional on X with the same bound. In particular, for $x_0 \in X$ there is some $\varphi \in X^*$ such that*

$$\varphi(x_0) = \|x_0\|_X \quad \textit{and} \quad \|\varphi\|_{X^*} = 1. \tag{4.9}$$

This implies that we recover the norm $\|\cdot\|_X$ from the dual norm $\|\cdot\|_{X^*}$ via the following maximum (no supremum is needed!):

$$\|x\|_X = \max\{|\varphi(x)| : \|\varphi\|_{X^*} = 1\} = \max\left\{\frac{|\varphi(x)|}{\|\varphi\|_{X^*}} : 0 \ne \varphi \in X^*\right\}. \tag{4.10}$$

Corollary 4.16. Let $\{x_\nu \in X : 1 \le \nu \le n\}$ be linearly independent. Then there are functionals $\varphi_\nu \in X^*$ such that $\varphi_\nu(x_\mu) = \delta_{\nu\mu}$ (cf. (2.1)). The functionals $(\varphi_\nu)_{\nu=1}^n$ are called *dual* to $(x_\nu)_{\nu=1}^n$.

Proof. Let $U := \operatorname{span}\{x_\nu : 1 \le \nu \le n\}$, and define the dual system $\{\varphi_\nu\}$ as in Definition 3.6. Since U is finite dimensional, the functionals φ_ν are bounded on U. By Theorem 4.15 there exists an extension of all φ_ν to X^* with the same bound. □

The following *Lemma of Auerbach* is proved, e.g., in [145, Lemma 10.5].

Lemma 4.17. *For any n-dimensional subspace of a Banach space X, there exists a basis $\{x_\nu : 1 \le \nu \le n\}$ and a corresponding dual system $\{\varphi_\nu : 1 \le \nu \le n\}$ such that $\|x_\nu\| = \|\varphi_\nu\|^* = 1$ $(1 \le \nu \le n)$.*

Lemma 4.18. *(a) Let $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ be two normed vector spaces with $\|\cdot\|_1 \le C\|\cdot\|_2$. By Remark 4.2, completion yields Banach spaces $(X_1, \|\cdot\|_1)$ and*

[5] Note that X^* is a subset of X', the space of the *algebraic* duals.

$(X_2, \|\cdot\|_2)$ *with* $X_2 \subset X_1$*. The corresponding duals* $\left(X_1^*, \|\cdot\|_1^*\right)$ *and* $\left(X_2^*, \|\cdot\|_2^*\right)$ *satisfy* $X_1^* \subset X_2^*$*. The dual norms fulfil* $\|\varphi\|_2^* \leq C\,\|\varphi\|_1^*$ *for all* $\varphi \in X_1^*$ *with the same constant* C*. If* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *are equivalent,* $X_1^* = X_2^*$ *holds.*

(b) If $\left(X^*, \|\cdot\|_1^*\right)$ *and* $\left(X^*, \|\cdot\|_2^*\right)$ *are identical sets with equivalent dual norms* $\|\cdot\|_1^*$ *and* $\|\cdot\|_2^*$ *generated by normed vector spaces* $(X, \|\cdot\|_1)$ *and* $(X, \|\cdot\|_2)$*, then also* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *are equivalent.*

Proof. 1) Let $\varphi \in X_1^*$. Since $X_2 \subset X_1$, $\varphi(x_2)$ is well-defined for any $x_2 \in X_2$ with $\|x_2\|_2 = 1$ and we estimate $|\varphi(x_2)| \leq \|\varphi\|_1^*\,\|x_2\|_1 \leq C\,\|\varphi\|_1^*\,\|x_2\|_2 = C\,\|\varphi\|_1^*$. Taking the supremum over all $x_2 \in X_2$ with $\|x_2\|_2 = 1$, we get $\|\varphi\|_2^* \leq C\,\|\varphi\|_1^*$. Again, by Remark 4.2, $X_1^* \subset X_2^*$ follows.

2) For equivalent norms both inclusions $X_1^* \subset X_2^* \subset X_1^*$ prove the assertion.

3) The identity $\|x\|_1 = \max\left\{\, |\varphi(x)| : \|\varphi\|_1^* = 1 \right\} = \max_{\varphi \neq 0}\left\{\, |\varphi(x)| \,/\, \|\varphi\|_1^* \,\right\}$ follows from (4.10). By equivalence, $\|\varphi\|_1^* \leq C\,\|\varphi\|_2^*$ follows, so that

$$\max_{\varphi \neq 0}\left\{|\varphi(x)| \,/\, \|\varphi\|_1^*\right\} \geq \tfrac{1}{C} \max_{\varphi \neq 0}\left\{|\varphi(x)| \,/\, \|\varphi\|_2^*\right\} = \tfrac{1}{C}\,\|x\|_2$$

and vice versa, proving Part (b). □

The Banach space $(X^*, \|\cdot\|_{X^*})$ has again a dual $(X^{**}, \|\cdot\|_{X^{**}})$ called the *bidual* of X. The embedding $X \subset X^{**}$ is to be understood as identification of $x \in X$ with the bidual mapping $\chi_x \in X^{**}$ defined by $\chi_x(\varphi) := \varphi(x)$ for all $\varphi \in X^*$. If $X = X^{**}$, the Banach space X is called *reflexive*.

Lemma 4.19. *Let* $\varphi \in X^*$*. Then*

$$\|\varphi\|_{X^*} = \sup_{0 \neq x \in X} \frac{|\varphi(x)|}{\|x\|_X} = \max_{0 \neq \Phi \in X^{**}} \frac{|\Phi(\varphi)|}{\|\Phi\|_{X^{**}}}.$$

If X *is reflexive,* $\|\varphi\|_{X^*} = \max_{0 \neq x \in X} |\varphi(x)| \,/\, \|x\|_X$ *holds* (max *instead of* sup)*.*

Proof. The left equality holds by definition of $\|\cdot\|_{X^*}$. The right equality is the identity (4.10) with x, X, φ replaced by φ, X^*, Φ. In the reflexive case, $\Phi(\varphi) = \varphi(x_\Phi)$ holds for some $x_\Phi \in X$ and proves the second part. □

Definition 4.20. If $\Phi \in \mathcal{L}(X, Y)$, the dual operator $\Phi^* \in \mathcal{L}(Y^*, X^*)$ is defined via $\Phi^* : \eta \mapsto \xi := \Phi^*\eta$ with $\xi(x) := \eta(\Phi x)$ for all $x \in X$.

Lemma 4.21. $\|\Phi^*\|_{X^* \leftarrow Y^*} = \|\Phi\|_{Y \leftarrow X}$.

4.1.6 Examples

In §4.1.3, examples of Banach spaces are given. Some of the corresponding dual spaces are easy to describe.

Example 4.22. (a) The dual of $\ell^p(I)$ for $1 \leq p < \infty$ is (isomorphic to) $\ell^q(I)$, where the conjugate q is defined by $\frac{1}{p} + \frac{1}{q} = 1$ and the embedding $\ell^q(I) \hookrightarrow (\ell^p(I))^*$ is defined by

$$\varphi(a) := \sum_{i\in I} a_i\varphi_i \in \mathbb{K} \qquad \text{for } a = (a_i)_{i\in I} \in \ell^p(I) \text{ and } \varphi = (\varphi_i)_{i\in I} \in \ell^q(I).$$

The subspace $c_0(I) \subset \ell^\infty(I)$ (cf. (4.4)) has the dual $\ell^1(I)$. If $\#I < \infty$, equality $(\ell^\infty(I))^* = \ell^1(I)$ holds; otherwise, $(\ell^\infty(I))^* \supsetneqq \ell^1(I)$.

(b) Similarly, $(L^p(D))^* \cong L^q(D)$ is valid for $1 \leq p < \infty$ with the embedding $g \in L^q(D) \mapsto g(f) := \int_D f(x)g(x)\mathrm{d}x$ for all $f \in L^p(D)$.

(c) Let $I = [a,b] \subset \mathbb{R}$ be an interval. Any functional $\varphi \in (C(I))^*$ corresponds to a function g of bounded variation such that $\varphi(f) = \int_I f(x)\mathrm{d}g(x)$ exists as Stieljes integral for all $f \in C(I)$. The latter integral with g chosen as the step function $g_s(x) := \left\{\begin{smallmatrix} 0 \text{ for } x \leq s \\ 1 \text{ for } x > s \end{smallmatrix}\right\}$ leads to the *Dirac functional* δ_s with the property

$$\delta_s(f) = f(s) \qquad \text{for all } f \in C(I) \text{ and } s \in I. \tag{4.11}$$

4.1.7 Weak Convergence

Let $(X, \|\cdot\|)$ be a Banach space. We say that $(x_n)_{n\in\mathbb{N}}$ *converges weakly to* $x \in X$, if $\lim \varphi(x_n) = \varphi(x)$ for all $\varphi \in X^*$. In this case, we write $x_n \rightharpoonup x$. Standard (strong) convergence $x_n \to x$ implies $x_n \rightharpoonup x$.

Lemma 4.23. *If $x_n \rightharpoonup x$, then $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$.*

Proof. Choose $\varphi \in X^*$ with $\|\varphi\|^* = 1$ and $|\varphi(x)| = \|x\|$ (cf. (4.9)) and note that $\|x\| \leftarrow |\varphi(x_n)| \leq \|x_n\|$. □

Lemma 4.24. *Let $N \in \mathbb{N}$. Assume that the sequences $(x_n^{(i)})_{n\in\mathbb{N}}$ for $1 \leq i \leq N$ converge weakly to linearly independent limits $x^{(i)} \in X$ (i.e., $x_n^{(i)} \rightharpoonup x^{(i)}$). Then there is an n_0 such that for all $n \geq n_0$ the N-tuples $(x_n^{(i)} : 1 \leq i \leq N)$ are linearly independent.*

Proof. There are functionals $\varphi^{(j)} \in X^*$ $(1 \leq j \leq N)$ with $\varphi^{(j)}(x^{(i)}) = \delta_{ij}$ (cf. Corollary 4.16). Set

$$\Delta_n := \det\left((\varphi^{(j)}(x_n^{(i)}))_{i,j=1}^N\right).$$

$x_n^{(i)} \rightharpoonup x^{(i)}$ implies $\varphi^{(j)}(x_n^{(i)}) \to \varphi^{(j)}(x^{(i)})$. Continuity of the determinant proves $\Delta_n \to \Delta_\infty := \det((\delta_{ij})_{i,j=1}^N) = 1$. Hence, there is an n_0 such that $\Delta_n > 0$ for all $n \geq n_0$, but $\Delta_n > 0$ proves linear independence of $\{x_n^{(i)} : 1 \leq i \leq N\}$. □

For a proof of the *local sequential weak compactness*, stated next, we refer to Yosida [198, Chap. V.2].

Lemma 4.25. *If X is a reflexive Banach space, any bounded sequence $x_n \in X$ has a subsequence x_{n_ν} converging weakly to some $x \in X$.*

Corollary 4.26. Let X be a reflexive Banach space, $x_n \in X$ a bounded sequence with $x_n = \sum_{i=1}^{r} \xi_{n,i}$, $\xi_{n,i} \in X$, and $\|\xi_{n,i}\| \leq C \|x_n\|$. Then there are $\xi_i \in X$ and a subsequence such that $\xi_{n_\nu,i} \rightharpoonup \xi_i$ and, in particular, $x_{n_\nu} \rightharpoonup x := \sum_{i=1}^{r} \xi_i$.

Proof. By Lemma 4.25, weak convergence $x_n \rightharpoonup x$ holds for $n \in \mathbb{N}^{(0)} \subset \mathbb{N}$, where $\mathbb{N}^{(0)}$ is an infinite subset of $\mathbb{N}$. Because of $\|\xi_{n,1}\| \leq C \|x_n\|$, also $\xi_{n,1}$ is a bounded sequence for $n \in \mathbb{N}^{(0)}$. A second infinite subset $\mathbb{N}^{(1)} \subset \mathbb{N}^{(0)}$ exists with the property $\xi_{n,1} \rightharpoonup \xi_1$ $(n \in \mathbb{N}^{(1)})$ for some $\xi_1 \in X$. Next, $\xi_{n,2} \rightharpoonup \xi_2 \in X$ can be shown for $n \in \mathbb{N}^{(2)} \subset \mathbb{N}^{(1)}$, etc. Finally, for $\mathbb{N}^{(r)} \ni n \to \infty$ all sequences $\xi_{n,i}$ converge weakly to ξ_i and summation over i yields $x_n \rightharpoonup x := \sum_{i=1}^{r} \xi_i$. □

Definition 4.27. A subset $M \subset X$ is called *weakly closed*, if $x_n \in M$ and $x_n \rightharpoonup x$ imply $x \in M$.

Note the implication 'M weakly closed $\Rightarrow$ M closed', i.e., 'weakly closed' is stronger than 'closed'.

Theorem 4.28. *Let $(X, \|\cdot\|)$ be a reflexive Banach space with a weakly closed subset $\emptyset \neq M \subset X$. Then the following minimisation problem has a solution: For any $x \in X$ find $v \in M$ with*

$$\|x - v\| = \inf\{\|x - w\| : w \in M\}.$$

Proof. Choose any sequence $w_n \in M$ with $\|x - w_n\| \searrow \inf\{\|x - w\| : w \in M\}$. Since $(w_n)_{n\in\mathbb{N}}$ is a bounded sequence in X, Lemma 4.25 ensures the existence of a weakly convergent subsequence $w_{n_i} \rightharpoonup v \in X$. v belongs to M because $w_{n_i} \in M$ and M is weakly closed. Since also $x - w_{n_i} \rightharpoonup x - v$ is valid, Lemma 4.23 shows $\|x - v\| \leq \liminf \|x - w_{n_i}\| \leq \inf\{\|x - w\| : w \in M\}$. □

Since the assumption of reflexivity excludes important spaces, we add some remarks on this subject. The existence of a minimiser or 'nearest point' v in a certain set $A \subset X$ to some $v \in V \backslash A$ is a well-studied subject. A set A is called 'proximinal' if for all $x \in X \backslash A$ the best approximation problem $\|x - v\| = \inf_{w\in A} \|x - w\|$ has at least one solution $v \in A$. Without the assumption of reflexivity, there are statements ensuring under certain conditions that the set of points $x \in X \backslash A$ possessing nearest points in A are dense (e.g., Edelstein [51]). However, in order to conclude from the weak closedness[6] of the minimal subspaces that they are proximinal, requires reflexivity as the following statement elucidates.

Theorem 4.29 ([64, p. 61]). *For a Banach space X the following is equivalent:*
(a) X is reflexive,
(b) All closed subspaces are proximinal.
(c) All weakly closed non-empty subsets are proximinal.

[6] On the other hand, if X coincides with the dual Y^* of another Banach space Y, every weak* closed set in X is proximinal (cf. Holmes [101, p. 123]). However, the later proofs of weak closedness of $U_j^{\min}(\mathbf{v})$ in §6 do not allow to conclude also weak* closedness.

4.1.8 Continuous Multilinear Mappings

Multilinearity is defined in (3.18a). The equivalence of continuity and boundedness shown in Remark 4.10a also holds for bilinear, or generally, for multilinear mappings. The proof follows by the same arguments.

Lemma 4.30. *(a) A bilinear mapping* $B : (V, \|\cdot\|_V) \times (W, \|\cdot\|_W) \to (X, \|\cdot\|)$ *is continuous if and only if there is some* $C \in \mathbb{R}$ *such that*

$$\|B(v,w)\| \le C \, \|v\|_V \, \|w\|_W \qquad \textit{for all } v \in V \textit{ and } w \in W.$$

(b) A multilinear mapping $A : \times_{i=1}^{d} \left(V_j, \|\cdot\|_j\right) \to (X, \|\cdot\|)$ *is continuous if and only if there is some* $C \in \mathbb{R}$ *such that*

$$\|A(v^{(1)}, \ldots, v^{(d)})\| \le C \prod\nolimits_{j=1}^{d} \|v^{(j)}\|_j \qquad \textit{for all } v^{(j)} \in V_j \, .$$

4.2 Topological Tensor Spaces

4.2.1 Notations

The algebraic tensor product $V \otimes_a W$ has been defined in (3.11) by the span of all elementary tensors $v \otimes w$ for $v \in V$ and $w \in W$. In pure algebraic constructions such a span is always a *finite* linear combination. Infinite sums as well as limits of sequences cannot be defined without topology. In the finite dimensional case, the algebraic tensor product $V \otimes_a W$ is already complete. Corollary 4.61 below will even show a similar case with one factor being infinite dimensional.

As already announced in (3.12), the completion $X := \overline{X_0}$ of $X_0 := V \otimes_a W$ with respect to some norm $\|\cdot\|$ yields a Banach space $(X, \|\cdot\|)$, which is denoted by

$$V \otimes_{\|\cdot\|} W := V \underset{\|\cdot\|}{\otimes} W := \overline{V \otimes_a W}^{\|\cdot\|}$$

and now called *Banach tensor space*. Note that the result of the completion depends on the norm $\|\cdot\|$ as already discussed in §4.1.2. A tensor $\mathbf{x} \in V \otimes_{\|\cdot\|} W$ is defined as limit $\mathbf{x} = \lim_{n\to\infty} \mathbf{x}_n$ of some $\mathbf{x}_n \in V \otimes_a W$ from the algebraic tensor space, e.g., $\mathbf{x}_n$ is the sum of say n elementary tensors. In general, such a limit of a sequence cannot be written as an infinite sum (but see §4.2.6). Furthermore, the convergence $\mathbf{x}_n \to \mathbf{x}$ may be arbitrarily slow. In practical applications, however, one is interested in fast convergence, in order to approximate $\mathbf{x}$ by $\mathbf{x}_n$ with reasonable n (if n is the number of involved elementary tensors, the storage will be related to n). In that case, the statement $\mathbf{x}_n \to \mathbf{x}$ should be replaced by a quantified error estimate:

$$\begin{aligned} &\|\mathbf{x}_n - \mathbf{x}\| \le O(\varphi(n)) && \text{with } \varphi(n) \to 0 \text{ as } n \to \infty, \\ \text{or}\quad &\|\mathbf{x}_n - \mathbf{x}\| \le o(\psi(n)) && \text{with } \sup \psi(n) < \infty, \end{aligned} \tag{4.12}$$

i.e., $\|\mathbf{x}_n - \mathbf{x}\| \le C\varphi(n)$ for some constant C or $\|\mathbf{x}_n - \mathbf{x}\| / \psi(n) \to 0$ as $n \to \infty$.

The notation $\otimes_{\|\cdot\|}$ becomes a bit cumbersome if the norm sign $\|\cdot\|$ carries further suffices, e.g., $\|\cdot\|_{C^n(I)}$. To shorten the notation, we often copy the (shortened) suffix[7] of the norm to the tensor sign, e.g., the association $p \leftrightarrow \|\cdot\|_{\ell^p(\mathbb{Z})}$ or $\wedge \leftrightarrow \|\cdot\|_{\wedge(V,W)}$ is used in

$$V \underset{p}{\otimes} W = V \otimes_p W, \qquad V \underset{\wedge}{\otimes} W = V \otimes_\wedge W, \qquad \text{etc.}$$

The neutral notation

$$V \otimes W$$

is used only if

(i) there is not doubt about the choice of norm of the Banach tensor space or

(ii) a statement holds both for the algebraic tensor product and the topological one.

In the finite dimensional case, where no completion is necessary, we shall use $V \otimes W$ for the algebraic product space (without any norm) as well as for any normed tensor space $V \otimes W$.

4.2.2 Continuity of the Tensor Product and Crossnorms

We start from Banach spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$. The question arises whether the norms $\|\cdot\|_V$ and $\|\cdot\|_W$ define a norm $\|\cdot\|$ on $V \otimes W$ in a canonical way. A suitable choice seems to be the definition

$$\|v \otimes w\| = \|v\|_V \, \|w\|_W \qquad \text{for all } v \in V \text{ and } w \in W \tag{4.13}$$

for the norm of elementary tensors. However, differently from linear maps, the definition of a norm on the set of elementary tensors does not determine $\|\cdot\|$ on the whole of $V \otimes_a W$. Hence, in contrast to algebraic tensor spaces, the topological tensor space $(V \otimes W, \|\cdot\|)$ is not uniquely determined by the components $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$.

Definition 4.31. Any norm $\|\cdot\|$ on $V \otimes_a W$ satisfying (4.13) is called a *crossnorm.*

A necessary condition for $\|\cdot\|$ is the continuity of the tensor product, i.e., the mapping $(v, w) \in V \times W \mapsto v \otimes w \in V \otimes W$ must be continuous. Since $\otimes$ is bilinear (cf. Lemma 3.10), we may apply Lemma 4.30a.

Remark 4.32. Continuity of the tensor product $\otimes : (V, \|\cdot\|_V) \times (W, \|\cdot\|_W) \to (V \otimes W, \|\cdot\|)$ is equivalent to the existence of some $C < \infty$ such that

$$\|v \otimes w\| \le C \, \|v\|_V \, \|w\|_W \qquad \text{for all } v \in V \text{ and } w \in W. \tag{4.14}$$

Here, $(V \otimes W, \|\cdot\|)$ may be the algebraic (possibly incomplete) tensor space $V \otimes_a W$ equipped with norm $\|\cdot\|$ or the Banach tensor space $(V \otimes_{\|\cdot\|} W, \|\cdot\|)$. Inequality (4.14) is, in particular, satisfied by a crossnorm.

[7] The letter a must be avoided, since this denotes the algebraic product $\otimes_a$.

Continuity of $\otimes$ is necessary, since otherwise there are pairs $v \in V$ and $w \in W$ such that $v \otimes w \in V \otimes_a W$ has *no* finite norm $\|v \otimes w\|$. Hence, $\|\cdot\|$ cannot be defined on all of $V \otimes_a W$ and $(V \otimes_a W, \|\cdot\|)$ is *not* a normed vector space. For a proof we use the following lemma (cf. [45, I.1.2]).

Lemma 4.33. *Let V be a Banach space (W may be not complete). Assume separate continuity: $\|v \otimes w\| \le C_w \|v\|_V$ and $\|v \otimes w\| \le C_v \|w\|_W$ hold for all $v \in V$ and $w \in W$ with C_v [C_w] depending on v [w]. Then (4.14) follows for some $C < \infty$.*

Assuming that (4.14) is not true, we conclude from Lemma 4.33 that there is some $w \in W$ such that $\sup_{0 \ne v \in V} \|v \otimes w\| / \|v\|_V = \infty$. Since for fixed w a norm $|||v|||_w := \|v \otimes w\|$ is defined which is strictly stronger than $\|\cdot\|_V$, there is $v_0 \in V$ which has no finite $|||\cdot|||_w$ norm (cf. Remark 4.2), i.e., $\|v_0 \otimes w\|$ is not defined.

In the definition of $V \otimes_{\|\cdot\|} W$ we have not fixed, whether V and W are complete or not. The next lemma shows that in any case the same Banach tensor space $V \otimes_{\|\cdot\|} W$ results. Any non-complete normed space $(V, \|\cdot\|_V)$ may be regarded as a dense subset of the Banach space $(\overline{V}, \|\cdot\|_V)$. Replacing the notations $\overline{V}, V$ by V, V_0, we can apply the following lemma. There we replace V and W in $V \otimes W$ by dense subspaces V_0 and W_0. Three norms are involved: $\|\cdot\|_V$ for V_0 and V, $\|\cdot\|_W$ for W_0 and W, and the norm $\|\cdot\|$ for the tensor spaces $V_0 \otimes_a W_0$, $V \otimes_a W$, $V \otimes_{\|\cdot\|} W$.

Lemma 4.34. *Let V_0 be dense in $(V, \|\cdot\|_V)$ and W_0 be dense in $(W, \|\cdot\|_W)$. Assume $\otimes : (V, \|\cdot\|_V) \times (W, \|\cdot\|_W) \to (V \otimes_{\|\cdot\|} W, \|\cdot\|)$ to be continuous, i.e., (4.14) holds. Then $V_0 \otimes_a W_0$ is dense in $V \otimes_{\|\cdot\|} W$, i.e.,*

$$\overline{V_0 \otimes_a W_0} = V \otimes_{\|\cdot\|} W.$$

Proof. For any $\varepsilon > 0$ and any $\mathbf{x} \in V \otimes_{\|\cdot\|} W$ we have to show that there is an $\mathbf{x}_\varepsilon \in V_0 \otimes_a W_0$ with $\|\mathbf{x} - \mathbf{x}_\varepsilon\| \le \varepsilon$. By definition of $V \otimes_{\|\cdot\|} W$, there is an $\mathbf{x}' \in V \otimes_a W$ with $\|\mathbf{x} - \mathbf{x}'\| \le \varepsilon/2$ and a finite sum representation $\mathbf{x}' = \sum_{i=1}^n v_i' \otimes w_i'$ with $v_i' \in V$ and $w_i' \in W$. We set

$$C_{\max} := \max \left\{ \|v_i'\|_V , \|w_i'\|_W : 1 \le i \le n \right\} ,$$

and choose δ so small that

$$nC\delta \left(2C_{\max} + \delta\right) \le \varepsilon/2$$

with C being the equally named constant from (4.14). We select $v_i \in V_0$ and $w_i \in W_0$ with $\|v_i' - v_i\|_V \le \delta$ and $\|w_i' - w_i\|_W \le \delta$ and set $\mathbf{x}_\varepsilon := \sum_{i=1}^n v_i \otimes w_i$. Then

$$\begin{aligned} \|\mathbf{x}' - \mathbf{x}_\varepsilon\| &= \left\| \sum_{i=1}^n \left(v_i' \otimes w_i' - v_i \otimes w_i\right) \right\| \\ &= \left\| \sum_{i=1}^n \left\{ (v_i' - v_i) \otimes w_i' + v_i' \otimes (w_i' - w_i) + (v_i - v_i') \otimes (w_i' - w_i) \right\} \right\| \end{aligned}$$

$$\leq \sum_{i=1}^{n} \{\|(v_i' - v_i)\otimes w_i'\| + \|v_i' \otimes (w_i' - w_i)\| + \|(v_i - v_i')\otimes (w_i' - w_i)\|\} \underset{(4.14)}{\leq}$$
$$\leq \sum_{i=1}^{n} \{C\,\|v_i' - v_i\|_V\,\|w_i'\|_W + C\,\|v_i'\|_V\,\|w_i' - w_i\|_W + C\,\|v_i - v_i'\|\,\|w_i' - w_i\|\}$$
$$\leq nC\delta\,\{2C_{\max} + \delta\} \leq \varepsilon/2$$

proves $\|\mathbf{x} - \mathbf{x}_\varepsilon\| \leq \varepsilon$. □

Remark 4.35. If $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ are separable (cf. Definition 4.3) and continuity (4.14) holds, also $(V \otimes_{\|\cdot\|} W, \|\cdot\|)$ is separable.

Proof. Let $\{v_i : i \in \mathbb{N}\}$ and $\{w_j : j \in \mathbb{N}\}$ be dense subsets of V and W, respectively. The set $E := \{v_i \otimes w_j : i, j \in \mathbb{N}\} \subset V \otimes_a W$ is again countable. By continuity of $\otimes$, the closure $\overline{E}$ contains all elementary tensors $v \otimes w \in V \otimes_a W$. Let $B := \operatorname{span}\{E\}$ be the set of finite linear combinations of $v_i \otimes w_j \in E$. The previous result shows $V \otimes_{\|\cdot\|} W \supset \overline{B} \supset V \otimes_a W$, which proves $\overline{B} = V \otimes_{\|\cdot\|} W$. □

4.2.3 Examples

Example 4.36 (ℓ^p). Let $V := (\ell^p(I), \|\cdot\|_{\ell^p(I)})$ and $W := (\ell^p(J), \|\cdot\|_{\ell^p(J)})$ for $1 \leq p \leq \infty$ and for some finite or countable I and J. Then the tensor product of the vectors $a = (a_\nu)_{\nu \in I} \in \ell^p(I)$ and $b = (b_\mu)_{\mu \in J} \in \ell^p(J)$ may be considered as the (infinite) matrix $a \otimes b =: c = (c_{\nu\mu})_{\nu \in I, \mu \in J}$ with entries $c_{\nu\mu} := a_\nu b_\mu$. Any elementary tensor $a \otimes b$ belongs to[8] $\ell^p(I \times J)$, and $\|\cdot\|_{\ell^p(I \times J)}$ is a crossnorm:

$$\|a \otimes b\|_{\ell^p(I \times J)} = \|a\|_{\ell^p(I)}\,\|b\|_{\ell^p(J)} \qquad \text{for } a \in \ell^p(I),\ b \in \ell^p(J). \tag{4.15}$$

Proof. For $p < \infty$, (4.15) follows from

$$\|a \otimes b\|^p_{\ell^p(I \times J)} = \sum_{\nu \in I}\sum_{\mu \in J} |a_\nu b_\mu|^p = \sum_{\nu \in I}\sum_{\mu \in J} |a_\nu|^p\,|b_\mu|^p$$
$$= \Big(\sum_{\nu \in I} |a_\nu|^p\Big)\Big(\sum_{\mu \in J} |b_\mu|^p\Big) = \Big(\|a\|_{\ell^p(I)}\,\|b\|_{\ell^p(J)}\Big)^p.$$

For $p = \infty$, use $|c_{\nu\mu}| = |a_\nu b_\mu| = |a_\nu|\,|b_\mu| \leq \|a\|_{\ell^\infty(I)}\,\|b\|_{\ell^\infty(J)}$ to show that $\|a \otimes b\|_{\ell^\infty(I \times J)} \leq \|a\|_{\ell^\infty(I)}\,\|b\|_{\ell^\infty(J)}$. For the reverse inequality use that there are indices ν^* and μ^* with $|a_{\nu^*}| \geq (1 - \varepsilon)\,\|a\|_{\ell^\infty(I)}$ and $|b_{\mu^*}| \geq (1 - \varepsilon)\,\|b\|_{\ell^\infty(J)}$. □

Since linear combinations of finitely many elementary tensors again belong to $\ell^p(I \times J)$, we obtain $\ell^p(I) \otimes_a \ell^p(J) \subset \ell^p(I \times J)$. Hence, the completion with respect to $\|\cdot\|_{\ell^p(I \times J)}$ yields a Banach tensor space $\ell^p(I) \otimes_p \ell^p(J) \subset \ell^p(I \times J)$. The next statement shows that, except for $p = \infty$, even equality holds.

[8] Note that $I \times J$ is again countable.

Remark 4.37. $\ell^p(I) \underset{p}{\otimes} \ell^p(J) = \ell^p(I \times J)$ holds for $1 \leq p < \infty$.

Proof. Let $1 \leq p < \infty$. It is sufficient to show that $\ell^p(I) \otimes_a \ell^p(J)$ is dense in $\ell^p(I \times J)$. Let i_ν and j_μ $(\nu, \mu \in \mathbb{N})$ be any enumerations of the index sets I and J, respectively. Note that $c \in \ell^p(I \times J)$ has entries indexed by $(i,j) \in I \times J$. Then for each $c \in \ell^p(I \times J)$, $\lim_{n\to\infty} \sum_{1\leq\nu,\mu\leq n} |c_{i_\nu,j_\mu}|^p = \|c\|^p_{\ell^p(I\times J)}$ holds. Define $c^{(n)} \in \ell^p(I \times J)$ by

$$c^{(n)}_{i_\nu,j_\mu} = \begin{cases} c_{i_\nu,j_\mu} & \text{for } 1 \leq \nu, \mu \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

The previous limit expresses that $c^{(n)} \to c$ as $n \to \infty$. Let $e^I_\nu \in \ell^p(I)$ and $e^J_\mu \in \ell^1(J)$ be the unit vectors (cf. (2.2)). Then the identity $c^{(n)} = \sum_{1\leq\nu,\mu\leq n} c^{(n)}_{i_\nu,j_\mu} e^I_\nu \otimes e^J_\mu$ shows that $c^{(n)} \in \ell^p(I) \otimes_a \ell^p(J)$. Since $c = \lim c^{(n)} \in \ell^p(I \times J)$ is an arbitrary element, $\ell^p(I) \otimes_a \ell^p(J)$ is dense in $\ell^p(I \times J)$. □

For $p = \infty$ we consider the proper subspace $c_0(I) \subsetneqq \ell^\infty(I)$ endowed with the same norm $\|\cdot\|_{\ell^\infty(I)}$ (cf. (4.4)). The previous proof idea can be used to show

$$c_0(I) \underset{\infty}{\otimes} c_0(J) = c_0(I \times J).$$

Next, we consider the particular case of $p = 2$, i.e., $V := (\ell^2(I), \|\cdot\|_{\ell^2(I)})$ and $W := (\ell^2(J), \|\cdot\|_{\ell^2(J)})$ for some finite or countable I and J. Again, $c := a \otimes b$ with $a \in \ell^2(I)$ and $b \in \ell^2(J)$ as well as any linear combination from $V \otimes_a W$ may be considered as a (possibly) infinite matrix $c = (c_{\nu\mu})_{\nu\in I,\mu\in J}$. Theorem 4.114 will provide an infinite singular value decomposition of c :

$$c = \sum_{i=1}^{\infty} \sigma_i \, v_i \otimes w_i \qquad \begin{array}{l} \text{with } \sigma_1 \geq \sigma_2 \geq \ldots \geq 0 \\ \text{and orthonormal systems } \{v_i\} \text{ and } \{w_i\}. \end{array} \tag{4.16}$$

Algebraic tensors $\mathbf{v} = \sum_{i=1}^n x_i \otimes y_i$ lead to $\sigma_i = 0$ for all $i > n$. Therefore, the sequence $\sigma := (\sigma_i)_{i\in\mathbb{N}}$ belongs to $\ell_0(\mathbb{N})$. Only for topological tensors sequences $(\sigma_i)_{i\in\mathbb{N}}$ with infinitely many nonzero entries can appear.

Definition 4.38 (Schatten norms). If the singular values $\boldsymbol{\sigma} = (\sigma_\nu)_{\nu=1}^\infty$ of c from (4.16) have a finite ℓ^p norm $\|\boldsymbol{\sigma}\|_p$, we set

$$\|c\|_{\mathrm{SVD},p} := \|\boldsymbol{\sigma}\|_p := \left(\sum_{\nu=1}^{\infty} |\sigma_\nu|^p \right)^{1/p} \qquad \text{for } 1 \leq p < \infty. \tag{4.17}$$

As already seen in (4.15), $\|\cdot\|_{\ell^2(I\times J)}$ is a crossnorm. The next example shows that there is more than one crossnorm.

Example 4.39. Consider Example 4.36 for $p = 2$, i.e., $V := (\ell^2(I), \|\cdot\|_{\ell^2(I)})$ and $W := (\ell^2(J), \|\cdot\|_{\ell^2(J)})$. For any $1 \leq p \leq \infty$, the Schatten[9] norm $\|\cdot\|_{\mathrm{SVD},p}$ is a

[9] Brief remarks about Robert Schatten including his publication list can be found in [139, p. 138].

crossnorm on $V \otimes_a W$:

$$\|a \otimes b\|_{\mathrm{SVD},p} = \|a\|_{\ell^2(I)} \|b\|_{\ell^2(J)} \qquad \text{for } a \in \ell^2(I),\ b \in \ell^2(J), 1 \leq p \leq \infty. \tag{4.18}$$

As a consequence, $\otimes : V \times W \to V \otimes_{\mathrm{SVD},p} W$ is continuous for all $1 \leq p \leq \infty$. In particular, $\|\cdot\|_{\mathrm{SVD},2} = \|\cdot\|_{\ell^2(I \times J)}$ holds.

Proof. The rank-1 matrix $c := a \otimes b$ has the singular values $\sigma_1 = \|a\|_{\ell^2(I)} \|b\|_{\ell^2(J)}$ and $\sigma_i = 0$ for $i \geq 2$. Since the sequence $\boldsymbol{\sigma} := (\sigma_i)$ has at most one nonzero entry, $\|\boldsymbol{\sigma}\|_p = \sigma_1$ holds for all $1 \leq p \leq \infty$ and implies (4.18). Concerning the last statement compare (2.19c). □

Example 4.40. (a) Let $V = C(I)$ and $W = C(J)$ be the spaces of continuous functions on some domains I and J with supremum norms $\|\cdot\|_V = \|\cdot\|_{C(I)}$ and $\|\cdot\|_W = \|\cdot\|_{C(J)}$ (cf. Example 4.8). Then the norm

$$\|\cdot\|_\infty = \|\cdot\|_{C(I \times J)} \qquad \text{on } V \otimes_a W = C(I) \otimes_a C(J)$$

satisfies the crossnorm property (4.13).
(b) If I and J are compact sets, the completion of $V \otimes_a W$ with respect to $\|\cdot\|_\infty = \|\cdot\|_{C(I \times J)}$ yields

$$C(I) \underset{\infty}{\otimes} C(J) = C(I \times J). \tag{4.19}$$

(c) Also for $V = L^p(I)$ and $W = L^p(J)$ $(1 \leq p < \infty)$, the crossnorm property (4.13) and the following identity hold:

$$L^p(I) \underset{p}{\otimes} L^p(J) = L^p(I \times J) \qquad \text{for } 1 \leq p < \infty.$$

Proof. 1) For Part (a), one proves (4.13) as in Example 4.36.

2) Obviously, polynomials in $C(I \times J)$ belong to $C(I) \otimes_a C(J)$. Since I and J and, therefore, also $I \times J$ are compact, the Stone-Weierstraß Theorem (cf. Yosida [198]) states that polynomials are dense in $C(I \times J)$. This implies that $C(I) \otimes_a C(J)$ is dense in $C(I \times J)$ and (4.19) follows.

3) In the case of L^p, one may replace polynomials by step functions. □

An obvious generalisation of L^p is the Sobolev space

$$H^{1,p}(D) := \left\{ f \in L^p(D) : \frac{\partial}{\partial x_j} f \in L^p(D) \text{ for } 1 \leq j \leq d \right\} \qquad \text{for } D \subset \mathbb{R}^d$$

with the norm

$$\|f\|_{1,p} := \|f\|_{H^{1,p}(D)} := \left(\|f\|_p^p + \sum_{j=1}^{d} \left\| \frac{\partial f}{\partial x_j} \right\|_p^p \right)^{1/p} \qquad \text{for } 1 \leq p < \infty.$$

Let I_1, I_2 be intervals. Then the algebraic tensor space $H^{1,p}(I_j) \otimes_a H^{1,p}(I_2)$ is a dense subset of $H^{1,p}(I_1 \times I_2)$. Hence, the completion with respect to $\|\cdot\|_{H^{1,p}(I_1 \times I_2)}$ yields

$$H^{1,p}(I_1 \times I_2) = H^{1,p}(I_1) \otimes_{1,p} H^{1,p}(I_2). \tag{4.20}$$

Example 4.41 **($H^{1,p}$).** The tensor space $(H^{1,p}(I_1 \times I_2), \|\cdot\|_{H^{1,p}(I_1 \times I_2)})$ from (4.20) for[10] $1 \le p \le \infty$ satisfies the continuity inequality (4.14) with $C = 1$, but the norm is *not* a crossnorm.

Proof. For $f = g \otimes h$, i.e., $f(x, y) = g(x)h(y)$, we have

$$\|g \otimes h\|^p_{H^{1,p}(I_1 \times I_2)} = \|g\|_p^p \|h\|_p^p + \|g'\|_p^p \|h\|_p^p + \|g\|_p^p \|h'\|_p^p \le \|g\|_{1,p}^p \|h\|_{1,p}^p ,$$

where equality holds if and only if $\|g'\|_p \|h'\|_p = 0$. □

For later use we introduce the *anisotropic* Sobolev space

$$H^{(1,0),p}(I_1 \times I_2) := \{f \in L^p(I_1 \times I_2) : \partial f/\partial x_1 \in L^p(I_1 \times I_2)\}$$

with the norm

$$\|f\|_{(1,0),p} := \|f\|_{H^{(1,0),p}(I_1 \times I_2)} := \left(\|f\|_p^p + \left\| \frac{\partial f}{\partial x_1} \right\|_p^p \right)^{1/p} \quad \text{for } 1 \le p < \infty.$$

In this case, $\|g \otimes h\|^p_{H^{(1,0),p}(I_1 \times I_2)} = \|g\|_p^p \|h\|_p^p + \|g'\|_p^p \|h\|_p^p = \|g\|_{1,p}^p \|h\|_p^p$ proves the following result.

Example 4.42. Let $1 \le p < \infty$. The tensor space

$$\left(H^{(1,0),p}(I_1 \times I_2), \|\cdot\|_{H^{(1,0),p}(I_1 \times I_2)} \right) = H^{1,p}(I_1) \otimes_{(1,0),p} L^p(I_2)$$

satisfies (4.13), i.e., $\|\cdot\|_{H^{(1,0),p}(I_1 \times I_2)}$ is a crossnorm.

4.2.4 Projective Norm $\|\cdot\|_{\wedge(V,W)}$

Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $V \otimes_a W$ and denote the corresponding completions by $V \otimes_1 W$ and $V \otimes_2 W$. If $\|\cdot\|_1 \lesssim \|\cdot\|_2$, we have already stated that $V \otimes_1 W \supset V \otimes_2 W$ (cf. Remark 4.2).

If the mapping $(v, w) \mapsto v \otimes w$ is continuous with respect to the three norms $\|\cdot\|_V$, $\|\cdot\|_W$, $\|\cdot\|_2$ of V, W, $V \otimes_2 W$, then it is also continuous for any weaker norm $\|\cdot\|_1$ of $V \otimes_1 W$. For a proof combine $\|v \otimes w\|_2 \le C \|v\|_V \|w\|_W$ from (4.14) with $\|\cdot\|_1 \le C' \|\cdot\|_2$ to obtain boundedness $\|v \otimes w\|_1 \le C'' \|v\|_V \|w\|_W$ with the constant $C'' := CC'$.

[10] To include $p = \infty$, define $\|f\|_{1,\infty} := \max\{\|f\|_\infty , \|\partial f/\partial x_1\|_\infty , \|\partial f/\partial x_2\|_\infty\}$.

On the other hand, if $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$, continuity may fail. Therefore, one may ask for the *strongest possible norm* still ensuring continuity. Note that the strongest possible norm yields the smallest possible Banach tensor space containing $V \otimes_a W$ (with continuous $\otimes$). Since $\lesssim$ is only a semi-ordering of the norms, it is not trivial that there exists indeed a strongest norm. The answer is given in the following exercise.

Exercise 4.43. Let $\mathcal{N}$ be the set of all norms α on $V \otimes_a W$ satisfying $\alpha(v \otimes w) \leq C_\alpha \|v\|_V \|w\|_W$ $(v \in V,\ w \in W)$. Replacing α by the equivalent norm α/C_α, we obtain the set $\mathcal{N}' \subset \mathcal{N}$ of norms with $\alpha(v \otimes w) \leq \|v\|_V \|w\|_W$. Define $\|\mathbf{x}\| := \sup\{\alpha(\mathbf{x}) : \alpha \in \mathcal{N}'\}$ for $\mathbf{x} \in V \otimes_a W$ and show that $\|\cdot\| \in \mathcal{N}'$ satisfies $\alpha \lesssim \|\cdot\|$ for all $\alpha \in \mathcal{N}$.

The norm defined next will be the candidate for the strongest possible norm.

Definition 4.44 (projective norm). The normed spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ induce the *projective norm*[11] $\|\cdot\|_{\wedge(V,W)} = \|\cdot\|_\wedge$ on $V \otimes_a W$ defined by

$$\begin{aligned} \|\mathbf{x}\|_{\wedge(V,W)} &:= \|\mathbf{x}\|_\wedge \\ &:= \inf\left\{ \sum_{i=1}^{n} \|v_i\|_V \|w_i\|_W : \mathbf{x} = \sum_{i=1}^{n} v_i \otimes w_i \right\} \quad \text{for } \mathbf{x} \in V \otimes_a W. \end{aligned} \tag{4.21}$$

Completion of $V \otimes_a W$ with respect to $\|\cdot\|_{\wedge(V,W)}$ defines the Banach tensor space

$$\left(V \underset{\wedge}{\otimes} W,\ \|\cdot\|_{\wedge(V,W)} \right).$$

Note that the infimum in (4.21) is taken over all representations of $\mathbf{x}$.

Lemma 4.45. $\|\cdot\|_{\wedge(V,W)}$ *is not only a norm, but also a crossnorm.*

Proof. 1) The first and third norm axioms from (4.1) are trivial. For the proof of the triangle inequality $\|\mathbf{x}' + \mathbf{x}''\|_\wedge \leq \|\mathbf{x}'\|_\wedge + \|\mathbf{x}''\|_\wedge$ choose any $\varepsilon > 0$. By definition of the infimum in (4.21), there are representations $\mathbf{x}' = \sum_{i=1}^{n'} v_i' \otimes w_i'$ and $\mathbf{x}'' = \sum_{i=1}^{n''} v_i'' \otimes w_i''$ with

$$\sum_{i=1}^{n'} \|v_i'\|_V \|w_i'\|_W \leq \|\mathbf{x}'\|_\wedge + \frac{\varepsilon}{2} \quad \text{and} \quad \sum_{i=1}^{n''} \|v_i''\|_V \|w_i''\|_W \leq \|\mathbf{x}''\|_\wedge + \frac{\varepsilon}{2}.$$

A possible representation of $\mathbf{x} = \mathbf{x}' + \mathbf{x}''$ is $\mathbf{x} = \sum_{i=1}^{n'} v_i' \otimes w_i' + \sum_{i=1}^{n''} v_i'' \otimes w_i''$. Hence, a bound of the infimum involved in $\|\mathbf{x}\|_\wedge$ is

$$\|\mathbf{x}\|_\wedge \leq \sum_{i=1}^{n'} \|v_i'\|_V \|w_i'\|_W + \sum_{i=1}^{n''} \|v_i''\|_V \|w_i''\|_W \leq \|\mathbf{x}'\|_\wedge + \|\mathbf{x}''\|_\wedge + \varepsilon.$$

[11] Grothendieck [80] introduced the notations $\|\cdot\|_\wedge$ for the projective norm and $\|\cdot\|_\vee$ for the injective norm from §4.2.7. The older notations by Schatten [167] are $\gamma(\cdot)$ for $\|\cdot\|_{\wedge(V,W)}$ and $\lambda(\cdot)$ for $\|\cdot\|_{\vee(V,W)}$.

Since $\varepsilon > 0$ is arbitrary, $\|\mathbf{x}\|_\wedge \leq \|\mathbf{x}'\|_\wedge + \|\mathbf{x}''\|_\wedge$ follows. It remains to prove $\|\mathbf{x}\|_\wedge > 0$ for $\mathbf{x} \neq 0$. In §4.2.7 we shall introduce another norm $\|\mathbf{x}\|_\vee$ for which $\|\mathbf{x}\|_\vee \leq \|\mathbf{x}\|_\wedge$ will be shown in Lemma 4.56. Hence, the norm property of $\|\cdot\|_\vee$ proves that $\mathbf{x} \neq 0$ implies $0 < \|\mathbf{x}\|_\vee \leq \|\mathbf{x}\|_\wedge$.

2) Definition (4.21) implies the inequality $\|v \otimes w\|_\wedge \leq \|v\|_V \|w\|_W$.

3) For the reverse inequality consider any representation $v \otimes w = \sum_{i=1}^n v_i \otimes w_i$. By Theorem 4.15 there is a continuous functional $\Lambda \in V^*$ with $\Lambda(v) = \|v\|_V$ and $\|\Lambda\|_{V^*} \leq 1$. The latter estimate implies $|\Lambda(v_i)| \leq \|v_i\|_V$. Applying Λ to the first component in $v \otimes w = \sum_{i=1}^n v_i \otimes w_i$, we conclude that

$$\|v\|_V \, w = \sum_{i=1}^n \Lambda(v_i) w_i$$

(cf. Remark 3.54). The triangle inequality of $\|\cdot\|_W$ yields

$$\|v\|_V \|w\|_W \leq \sum_{i=1}^n |\Lambda(v_i)| \, \|w_i\|_W \leq \sum_{i=1}^n \|v_i\|_V \|w_i\|_W \, .$$

The infimum over all representations yields $\|v\|_V \|w\|_W \leq \|v \otimes w\|_\wedge$. □

Proposition 4.46. *Given $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$, the norm (4.21) is the strongest one ensuring continuity of $(v, w) \mapsto v \otimes w$. More precisely, if some norm $\|\cdot\|$ satisfies (4.14) with a constant C, then $\|\cdot\| \leq C \|\cdot\|_\wedge$ holds with the same constant.*

Proof. Let $\|\cdot\|$ be a norm on $V \otimes_a W$ such that $(v, w) \mapsto v \otimes w$ is continuous. Then (4.14), i.e., $\|v \otimes w\| \leq C \|v\|_V \|w\|_W$ holds. Let $\mathbf{x} = \sum_{i=1}^n v_i \otimes w_i \in V \otimes_a W$. The triangle inequality yields $\|\mathbf{x}\| \leq \sum_{i=1}^n \|v_i \otimes w_i\|$. Together with (4.14), $\|\mathbf{x}\| \leq C \sum_{i=1}^n \|v_i\|_V \|w_i\|_W$ follows. Taking the infimum over all representations $\mathbf{x} = \sum_{i=1}^n v_i \otimes w_i$, $\|\mathbf{x}\| \leq C \|\mathbf{x}\|_\wedge$ follows, i.e., $\|\cdot\|_\wedge$ is stronger than $\|\cdot\|$. Since $\|\cdot\|$ is arbitrary under the continuity side condition, $\|\cdot\|_\wedge$ is the strongest norm. □

The property of $\|\cdot\|_\wedge$ as the strongest possible norm ensuring continuity justifies calling $\|\cdot\|_{\wedge(V,W)}$ a canonical norm of $V \otimes_a W$ induced by $\|\cdot\|_V$ and $\|\cdot\|_W$. The completion $V \otimes_\wedge W$ with respect to $\|\cdot\|_\wedge$ is the smallest one. Any other norm $\|\cdot\|$ with (4.14) leads to a larger Banach space $V \otimes_{\|\cdot\|} W$.

4.2.5 Examples

Example 4.47 (ℓ^1). Consider $V = (\ell^1(I), \|\cdot\|_{\ell^1(I)})$ and $W = (\ell^1(J), \|\cdot\|_{\ell^1(J)})$ for some finite or countable sets I and J (cf. Example 4.36). The projective norm and the resulting Banach tensor space are

$$\|\cdot\|_{\wedge(V,W)} = \|\cdot\|_{\ell^1(I \times J)} \quad \text{and} \quad \ell^1(I) \underset{\wedge}{\otimes} \ell^1(J) = \ell^1(I \times J).$$

Proof. 1) $c \in V \otimes_a W \subset \ell^1(I \times J)$ has entries $c_{\nu\mu}$ for $\nu \in I$ and $\mu \in J$. A possible representation of c is

$$c = \sum_{\nu\in I} \sum_{\mu\in J} c_{\nu\mu}\, e_V^{(\nu)} \otimes e_W^{(\mu)}$$

with unit vectors $e_V^{(\nu)}$ and $e_W^{(\mu)}$ (cf. (2.2)). By $\|e_V^{(\nu)}\|_{\ell^1(I)} = \|e_W^{(\mu)}\|_{\ell^1(J)} = 1$ and the definition of $\|\cdot\|_{\wedge(V,W)}$, we obtain the inequality

$$\|c\|_{\wedge(V,W)} \le \sum_{\nu\in I} \sum_{\mu\in J} |c_{\nu\mu}| = \|c\|_{\ell^1(I\times J)} \,.$$

2) $\|\cdot\|_{\ell^1(I\times J)}$ is a crossnorm (cf. (4.15)), i.e., (4.14) holds with $C = 1$. Hence Proposition 4.46 shows $\|c\|_{\ell^1(I\times J)} \le \|c\|_{\wedge(V,W)}$. Together with the previous part, $\|\cdot\|_{\wedge(V,W)} = \|\cdot\|_{\ell^1(I\times J)}$ follows.

3) The latter identity together with Remark 4.37 (for $p = 1$) implies that the tensor product equals $\ell^1(I) \underset{\wedge}{\otimes} \ell^1(J) = \ell^1(I \times J)$. □

Example 4.48 (ℓ^2). Let $V := (\ell^2(I), \|\cdot\|_{\ell^2(I)})$ and $W := (\ell^2(J), \|\cdot\|_{\ell^2(J)})$ for some finite or countable sets I and J. Then

$$\|\cdot\|_{\wedge(V,W)} = \|\cdot\|_{\mathrm{SVD},1} \qquad \text{(cf. (4.17) for } \|\cdot\|_{\mathrm{SVD},p}).$$

Note that $\|\cdot\|_{\mathrm{SVD},1} \gneqq \|\cdot\|_{\mathrm{SVD},2} = \|\cdot\|_{\ell^2(I\times J)}$ for $\#I, \#J > 1$.

Proof. 1) Let $c \in V \otimes_a W$. Its singular value decomposition $c = \sum_i \sigma_i\, v_i \otimes w_i$ has only finitely many nonzero singular values σ_i. By definition of $\|\cdot\|_{\wedge(V,W)}$ we have

$$\|c\|_{\wedge(V,W)} \le \sum_i \sigma_i \, \|v_i\|_V \, \|w_i\|_W \,.$$

$\|v_i\|_V = \|w_i\|_W = 1$ (v_i, w_i orthonormal) yields $\|c\|_{\wedge(V,W)} \le \|\sigma\|_1 = \|c\|_{\mathrm{SVD},1}$.

2) $\|\cdot\|_{\mathrm{SVD},1}$ satisfies (4.14) with $C = 1$ (cf. (4.18)), so that Proposition 4.46 implies the opposite inequality $\|c\|_{\mathrm{SVD},1} \le \|c\|_{\wedge(V,W)}$. Together, the assertion $\|c\|_{\mathrm{SVD},1} = \|c\|_{\wedge(V,W)}$ is proved. □

4.2.6 Absolutely Convergent Series

By definition, any topological tensor $\mathbf{x} \in V \otimes_{\|\cdot\|} W$ is the limit of some sequence $\mathbf{x}^{(\nu)} = \sum_{i=1}^{n_\nu} v_i^{(\nu)} \otimes w_i^{(\nu)} \in V \otimes_a W$. If, in particular, $v_i^{(\nu)} = v_i$ and $w_i^{(\nu)} = w_i$ are independent of ν, the partial sums $\mathbf{x}_\nu$ define the series $\mathbf{x} = \sum_{i=1}^{\infty} v_i \otimes w_i$. Below we state that there is an absolutely convergent series $\mathbf{x} = \sum_{i=1}^{\infty} v_i \otimes w_i$, whose representation is almost optimal compared with the infimum $\|\mathbf{x}\|_\wedge$.

Proposition 4.49. *For any $\varepsilon > 0$ and any $\mathbf{x} \in V \otimes_\wedge W$, there is an absolutely convergent infinite sum*

$$\mathbf{x} = \sum_{i=1}^{\infty} v_i \otimes w_i \qquad (v_i \in V,\, w_i \in W) \tag{4.22a}$$

with

$$\sum_{i=1}^{\infty} \|v_i\|_V \, \|w_i\|_W \le (1+\varepsilon) \, \|\mathbf{x}\|_\wedge \, . \tag{4.22b}$$

Proof. We abbreviate $\|\cdot\|_\wedge$ by $\|\cdot\|$ and set $\varepsilon_\nu := \frac{\varepsilon}{3} \|\mathbf{x}\| / 2^\nu$. If $\mathbf{x} = 0$, nothing is to be done. Otherwise, choose some $s_1 \in V \otimes_a W$ with

$$\|\mathbf{x} - s_1\| \le \varepsilon_1. \tag{4.23a}$$

Hence, $\|s_1\| \le \|\mathbf{x}\| + \varepsilon_1$ follows. By definition of the norm, there is a representation $s_1 = \sum_{i=1}^{n_1} v_i \otimes w_i$ with

$$\sum_{i=1}^{n_1} \|v_i\|_V \, \|w_i\|_W \le \|s_1\| + \varepsilon_1 \le \|\mathbf{x}\| + 2\varepsilon_1. \tag{4.23b}$$

Set $d_1 := \mathbf{x} - s_1$ and approximate d_1 by $s_2 \in V \otimes_a W$ such that

$$\|d_1 - s_2\| \le \varepsilon_2, \; s_2 = \sum_{i=n_1+1}^{n_2} v_i \otimes w_i, \; \sum_{i=n_1+1}^{n_2} \|v_i\|_V \, \|w_i\|_W \le \|s_2\| + \varepsilon_2 \le \varepsilon_1 + 2\varepsilon_2 \tag{4.23c}$$

(here we use $\|s_2\| \le \|d_1\| + \varepsilon_2$ and $\|d_1\| \le \varepsilon_1$; cf. (4.23a)). Analogously, we set $d_2 := d_1 - s_2$ and choose s_3 and its representation such that

$$\|d_2 - s_3\| \le \varepsilon_3, \; s_3 = \sum_{i=n_2+1}^{n_3} v_i \otimes w_i, \; \sum_{i=n_2+1}^{n_3} \|v_i\|_V \, \|w_i\|_W \le \|s_3\| + \varepsilon_3 \le \varepsilon_2 + 2\varepsilon_3. \tag{4.23d}$$

By induction, using (4.23a,c,d) one obtains statement (4.22a) in the form of $\mathbf{x} = \sum_{\nu=1}^{\infty} s_\nu = \sum_{i=1}^{\infty} v_i \otimes w_i$. The estimates of the partial sums in (4.23b,c,d) show that $\sum_{i=1}^{\infty} \|v_i\|_V \, \|w_i\|_W \le \|\mathbf{x}\| + 3 \sum_{\nu=1}^{\infty} \varepsilon_\nu = \|\mathbf{x}\| + \varepsilon \, \|\mathbf{x}\|$ proving (4.22b).□

4.2.7 Duals and Injective Norm $\|\cdot\|_{\vee(V,W)}$

The normed spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ give rise to the dual spaces V^* and W^* endowed with the dual norms $\|\cdot\|_{V^*}$ and $\|\cdot\|_{W^*}$ described in (4.8). Consider the tensor space $V^* \otimes_a W^*$. Elementary tensors $\varphi \otimes \psi$ from $V^* \otimes_a W^*$ may be viewed as linear forms on $V \otimes_a W$ via the definition

$$(\varphi \otimes \psi)\,(v \otimes w) := \varphi(v) \cdot \psi(w) \in \mathbb{K} \, .$$

As discussed in §3.3.2.2, any $\mathbf{x}^* \in V^* \otimes_a W^*$ is a linear form on $V \otimes_a W$. Hence,

$$V^* \otimes_a W^* \subset (V \otimes_a W)' . \tag{4.24}$$

Note that $(V \otimes_a W)'$ is the *algebraic* dual, since continuity is not yet ensured.

A norm $\|\cdot\|$ on $V \otimes_a W$ leads to a dual space $(V \otimes_a W)^*$ with a dual norm denoted by $\|\cdot\|^*$. We would like to have

$$V^* \otimes_a W^* \subset (V \otimes_a W)^* \tag{4.25}$$

instead of (4.24). Therefore the requirement on the dual norm $\|\cdot\|^*$ (and indirectly on $\|\cdot\|$) is that $\otimes : (V^*, \|\cdot\|_{V^*}) \times (W^*, \|\cdot\|_{W^*}) \to \left(V^* \otimes_a W^*, \|\cdot\|^*\right)$ is continuous. The latter property, as seen in §4.2.2, is expressed by

$$\|\varphi \otimes \psi\|^* \leq C \, \|\varphi\|_{V^*} \, \|\psi\|_{W^*} \qquad \text{for all } \varphi \in V^* \text{ and } \psi \in W^*. \tag{4.26}$$

Lemma 4.50. *Inequality (4.26) implies*

$$\|v\|_V \, \|w\|_W \leq C \, \|v \otimes w\| \qquad \textit{for all } v \in V, w \in W, \tag{4.27}$$

which coincides with (4.14) up to the direction of the inequality sign. Furthermore, (4.26) implies the inclusion (4.25).

Proof. Given $v \otimes w \in V \otimes_a W$, choose φ and ψ according to Theorem 4.15: $\|\varphi\|_{V^*} = \|\psi\|_{W^*} = 1$ and $\varphi(v) = \|v\|_V$, $\psi(w) = \|w\|_W$. Then

$$\begin{aligned}\|v\|_V \, \|w\|_W &= \varphi(v)\psi(w) = |(\varphi \otimes \psi)\,(v \otimes w)| \leq \|\varphi \otimes \psi\|^* \, \|v \otimes w\| \underset{(4.26)}{\leq} \\ &\leq C \, \|\varphi\|_{V^*} \, \|\psi\|_{W^*} \, \|v \otimes w\| = C \, \|v \otimes w\|\end{aligned}$$

proves (4.27) with the same constant C as in (4.26). □

A similar result with a 'wrong' inequality sign for $\|\cdot\|^*$ is stated next.

Lemma 4.51. *Let $\|\cdot\|$ satisfy the continuity condition (4.14) with constant C. Then $C \, \|\varphi \otimes \psi\|^* \geq \|\varphi\|_{V^*} \|\psi\|_{W^*}$ holds for all $\varphi \in V^*$ and $\psi \in W^*$.*

Proof. The desired inequality follows from

$$\frac{|\varphi(v)|}{\|v\|_V} \frac{|\psi(w)|}{\|w\|_W} = \frac{|(\varphi \otimes \psi)\,(v \otimes w)|}{\|v\|_V \, \|w\|_W} \leq \frac{\|\varphi \otimes \psi\| \, \|v \otimes w\|}{\|v\|_V \, \|w\|_W} \underset{(4.14)}{\leq} C \, \|\varphi \otimes \psi\|$$

for all $0 \neq v \in V$ and $0 \neq w \in W$. □

As in §4.2.2, we may ask for the strongest dual norm satisfying inequality (4.26). As seen in Lemma 4.18, weaker norms $\|\cdot\|$ correspond to stronger dual norms $\|\cdot\|^*$. Therefore, the following two questions are equivalent:

- Which norm $\|\cdot\|$ on $V \otimes_a W$ yields the strongest dual norm $\|\cdot\|^*$ satisfying (4.26)?
- What is the weakest norm $\|\cdot\|$ on $V \otimes_a W$ such that the corresponding dual norm $\|\cdot\|^*$ satisfies (4.26)?

Exercise 4.52. Prove analogously to Exercise 4.43 that there exists a unique weakest norm $V \otimes_a W$ satisfying (4.26).

A candidate will be defined below. Since this will be again a norm determined only by $\|\cdot\|_V$ and $\|\cdot\|_W$, it may also be called an induced norm.

Definition 4.53 (injective norm). Normed spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ induce the *injective norm* $\|\cdot\|_{\vee(V,W)}$ on $V \otimes_a W$ defined by

$$\|\mathbf{x}\|_{\vee(V,W)} := \|\mathbf{x}\|_{\vee} := \sup_{\substack{\varphi\in V^*,\ \|\varphi\|_{V^*}=1 \\ \psi\in W^*,\ \|\psi\|_{W^*}=1}} \left|(\varphi\otimes\psi)(\mathbf{x})\right|. \tag{4.28}$$

The completion of $V \otimes_a W$ with respect to $\|\cdot\|_\vee$ defines $(V \otimes_\vee W, \|\cdot\|_\vee)$.

Lemma 4.54. *(a) $\|\cdot\|_{\vee(V,W)}$ defined in (4.28) is a crossnorm on $V \otimes_a W$, i.e., (4.13) holds implying (4.14) and (4.27).*
(b) The dual norm $\|\varphi\otimes\psi\|^*_{\vee(V,W)}$ is a crossnorm on $V^* \otimes_a W^*$, i.e.,

$$\|\varphi\otimes\psi\|^*_{\vee(V,W)} = \|\varphi\|_{V^*}\|\psi\|_{W^*} \qquad \textit{for all } \varphi\in V^*, \psi\in W^* \tag{4.29}$$

holds, implying (4.26).

Proof. 1) The norm axiom $\|\lambda\mathbf{x}\|_\vee = |\lambda|\,\|\mathbf{x}\|_\vee$ and the triangle inequality are standard. To show positivity $\|\mathbf{x}\|_\vee > 0$ for $0 \neq \mathbf{x} \in V \otimes_a W$, apply Lemma 3.13: $\mathbf{x}$ has a representation $\mathbf{x} = \sum_{i=1}^r v_i \otimes w_i$ with linearly independent v_i and w_i. Note that $r \geq 1$ because of $\mathbf{x} \neq 0$. Then there are normalised functionals $\varphi \in V^*$ and $\psi \in W^*$ with $\varphi(v_1) \neq 0$ and $\psi(w_1) \neq 0$, while $\varphi(v_i) = \psi(w_i) = 0$ for $i \geq 2$. This leads to

$$|(\varphi\otimes\psi)(\mathbf{x})| = \left|(\varphi\otimes\psi)\left(\sum_{i=1}^r v_i\otimes w_i\right)\right| = \left|\sum_{i=1}^r \varphi(v_i)\psi(w_i)\right| = |\varphi(v_1)\psi(w_1)| > 0.$$

Hence also $\|\mathbf{x}\|_\vee \geq |(\varphi\otimes\psi)(\mathbf{x})|$ is positive.

2) Application of (4.28) to an elementary tensor $v \otimes w$ yields

$$\begin{aligned}\|v\otimes w\|_{\vee(V,W)} &= \sup_{\substack{\|\varphi\|_{V^*}=1\\ \|\psi\|_{W^*}=1}} |(\varphi\otimes\psi)(v\otimes w)| = \sup_{\substack{\|\varphi\|_{V^*}=1\\ \|\psi\|_{W^*}=1}} |\varphi(v)|\,|\psi(w)| \\ &= \left(\sup_{\|\varphi\|_{V^*}=1} |\varphi(v)|\right)\left(\sup_{\|\psi\|_{W^*}=1} |\psi(w)|\right) \underset{(4.10)}{=} \|v\|_V\,\|w\|_W .\end{aligned}$$

3) For $0 \neq \varphi\otimes\psi \in V^* \otimes_a W^*$ introduce the normalised continuous functionals $\hat\varphi := \varphi/\|\varphi\|_{V^*}$ and $\hat\psi := \psi/\|\psi\|_{W^*}$. Then for all $\mathbf{x} \in V \otimes_a W$, the inequality

$$\begin{aligned}&|(\varphi\otimes\psi)(\mathbf{x})| = \|\varphi\|_{V^*}\|\psi\|_{W^*}\left|(\hat\varphi\otimes\hat\psi)(\mathbf{x})\right| \\ &\leq \|\varphi\|_{V^*}\|\psi\|_{W^*} \sup_{\substack{\varphi'\in V^*,\ \|\varphi'\|_{V^*}=1\\ \psi'\in W^*,\ \|\psi'\|_{W^*}=1}} |(\varphi'\otimes\psi')(\mathbf{x})| = \|\varphi\|_{V^*}\|\psi\|_{W^*}\|\mathbf{x}\|_\vee\end{aligned}$$

follows. The supremum over all $\mathbf{x} \in V \otimes_a W$ with $\|\mathbf{x}\|_\vee = 1$ yields the dual norm so that $\|\varphi \otimes \psi\|^*_{\vee(V,W)} \leq \|\varphi\|_{V^*} \|\psi\|_{W^*}$. This is already (4.26) with $C = 1$.

4) Let $\varepsilon > 0$ and $\varphi \otimes \psi \in V^* \otimes_a W^*$ be arbitrary. According to Remark 4.11, there are $v_\varepsilon \in V$ and $w_\varepsilon \in W$ with $\|v_\varepsilon\|_V = \|w_\varepsilon\|_W = 1$ and

$$|\varphi(v_\varepsilon)| \geq (1-\varepsilon) \|\varphi\|_{V^*} \quad \text{and} \quad |\psi(w_\varepsilon)| \geq (1-\varepsilon) \|\psi\|_{W^*}.$$

Note that by (4.13) $\mathbf{x}_\varepsilon := v_\varepsilon \otimes w_\varepsilon$ satisfies $\|\mathbf{x}_\varepsilon\|_\vee = \|v_\varepsilon\|_V \|w_\varepsilon\|_W = 1$. Hence

$$\begin{aligned} \|\varphi \otimes \psi\|^*_{\vee(V,W)} &= \sup_{\|\mathbf{x}\|_\vee = 1} |(\varphi \otimes \psi)(\mathbf{x})| \geq |(\varphi \otimes \psi)(\mathbf{x}_\varepsilon)| = \\ = |(\varphi \otimes \psi)(v_\varepsilon \otimes w_\varepsilon)| &= \quad |\varphi(v_\varepsilon)| \, |\psi(w_\varepsilon)| \quad \geq (1-\varepsilon)^2 \|\varphi\|_{V^*} \|\psi\|_{W^*}. \end{aligned}$$

As $\varepsilon > 0$ is arbitrary, the reverse inequality $\|\varphi \otimes \psi\|^*_{\vee(V,W)} \geq \|\varphi\|_{V^*} \|\psi\|_{W^*}$ follows. Together with Step 3), we have proved (4.29). □

Proposition 4.55. $\|\cdot\| = \|\cdot\|_{\vee(V,W)}$ *is the weakest norm on* $V \otimes_a W$ *subject to the additional condition that the dual norm* $\|\cdot\|^*$ *satisfies (4.26).*

Proof. Let $\|\cdot\|$ be a weaker norm, i.e., $\|\cdot\| \leq C \|\cdot\|_{\vee(V,W)}$. The dual norms satisfy $\|\cdot\|^*_{\vee(V,W)} \leq C \|\cdot\|^*$ (cf. Lemma 4.18). Choose any $0 \neq \varphi \otimes \psi \in V^* \otimes_a W^*$. Again, we apply Remark 4.11. Given any $\varepsilon > 0$, there is some $\mathbf{x}_\varepsilon \in V \otimes_a W$ with $\|\mathbf{x}_\varepsilon\| = 1$ and

$$\begin{aligned} \|\varphi \otimes \psi\|^* &\leq (1+\varepsilon) |(\varphi \otimes \psi)(\mathbf{x}_\varepsilon)| \leq (1+\varepsilon) \|(\varphi \otimes \psi)\|^* \overbrace{\|\mathbf{x}_\varepsilon\|}^{=1} \underset{(4.26)}{\leq} \\ &\leq C (1+\varepsilon) \|\varphi\|_{V^*} \|\psi\|_{W^*} \underset{(4.29)}{=} C (1+\varepsilon) \|\varphi \otimes \psi\|^*_{\vee(V,W)}, \end{aligned}$$

i.e., $\|\cdot\|^* \leq C \|\cdot\|^*_{\vee(V,W)}$. Together with assumption $\|\cdot\|^*_{\vee(V,W)} \leq C \|\cdot\|^*$, equivalence of both norms follows. Thus, also $\|\cdot\|$ and $\|\cdot\|_{\vee(V,W)}$ are equivalent (cf. Lemma 4.18). This proves that there is no weaker norm than $\|\cdot\|_{\vee(V,W)}$. □

Lemma 4.56. $\|\cdot\|_{\vee(V,W)} \leq \|\cdot\|_{\wedge(V,W)}$ *holds on* $V \otimes_a W$.

Proof. Choose any $\varepsilon > 0$. Let $\mathbf{x} = \sum_i v_i \otimes w_i \in V \otimes_a W$ be some representation with $\sum_i \|v_i\|_V \|w_i\|_W \leq \|\mathbf{x}\|_{\wedge(V,W)} + \varepsilon$. Choose normalised functionals φ and ψ with $\|\mathbf{x}\|_{\vee(V,W)} \leq |(\varphi \otimes \psi)(\mathbf{x})| + \varepsilon$. Then

$$\begin{aligned} \|\mathbf{x}\|_{\vee(V,W)} &\leq \left|(\varphi \otimes \psi)\left(\sum\nolimits_i v_i \otimes w_i\right)\right| + \varepsilon = \left|\sum\nolimits_i \varphi(v_i)\psi(w_i)\right| + \varepsilon \\ &\leq \sum_i |\varphi(v_i)| \, |\psi(w_i)| + \varepsilon \leq \sum_i \|v_i\|_V \|w_i\|_W + \varepsilon \leq \|\mathbf{x}\|_{\wedge(V,W)} + 2\varepsilon. \end{aligned}$$

As $\varepsilon > 0$ is arbitrary, $\|\mathbf{x}\|_{\vee(V,W)} \leq \|\mathbf{x}\|_{\wedge(V,W)}$ holds for all $\mathbf{x} \in V \otimes_a W$. □

Exercise 4.57. For any $a_1, \ldots, a_n \geq 0$ and $b_1, \ldots, b_n > 0$ show that

$$\min_{1\leq i\leq n} \frac{a_i}{b_i} \leq \frac{a_1 + \ldots + a_n}{b_1 + \ldots + b_n} \leq \max_{1\leq i\leq n} \frac{a_i}{b_i}.$$

So far, we have considered the norm $\|\mathbf{x}\|_{\vee(V,W)}$ on $V \otimes_a W$. Analogously, we can define $\|\mathbf{x}\|_{\vee(V^*,W^*)}$ on $V^* \otimes_a W^*$. For the latter norm we shall establish a connection with $\|\cdot\|_{\wedge(V,W)}$ from §4.2.4, which states that in a certain sense the injective norm is dual to the projective norm.[12]

Proposition 4.58. $\|\cdot\|_{\vee(V^*,W^*)} = \|\cdot\|^*_{\wedge(V,W)}$ *on* $V^* \otimes_a W^*$.

Proof. 1) The norm $\|\boldsymbol{\Phi}\|^*_{\wedge(V,W)}$ of $\boldsymbol{\Phi} \in V^* \otimes_a W^*$ is bounded by

$$\begin{aligned}\|\boldsymbol{\Phi}\|^*_{\wedge(V,W)} &= \sup_{0\neq\mathbf{x}\in V\otimes_a W} \frac{|\boldsymbol{\Phi}(\mathbf{x})|}{\|\mathbf{x}\|_{\wedge(V,W)}} = \sup_{0\neq\mathbf{x}=\sum_i v_i\otimes w_i\in V\otimes_a W} \frac{|\boldsymbol{\Phi}(\sum_i v_i \otimes w_i)|}{\sum_i \|v_i\|_V \|w_i\|_W} \\ &\leq \sup_{0\neq\mathbf{x}=\sum_i v_i\otimes w_i\in V\otimes_a W} \frac{\sum_i |\boldsymbol{\Phi}(v_i \otimes w_i)|}{\sum_i \|v_i\|_V \|w_i\|_W} \underset{\text{Exercise 4.57}}{\leq} \sup_{\substack{0\neq v\in V\\0\neq w\in W}} \frac{|\boldsymbol{\Phi}(v \otimes w)|}{\|v\|_V \|w\|_W}.\end{aligned}$$

On the other hand, the elementary tensor $\mathbf{x} = v \otimes w$ appearing in the last expression is only a subset of those $\mathbf{x}$ used in $\sup_{0\neq\mathbf{x}=\sum_i v_i\otimes w_i} \frac{|\boldsymbol{\Phi}(\sum_i v_i\otimes w_i)|}{\sum_i\|v_i\|_V\|w_i\|_W} = \|\boldsymbol{\Phi}\|^*_{\wedge(V,W)}$, so that $\|\boldsymbol{\Phi}\|^*_{\wedge(V,W)}$ must be an upper bound. Together, we arrive at

$$\|\boldsymbol{\Phi}\|^*_{\wedge(V,W)} = \sup_{\substack{0\neq v\in V\\0\neq w\in W}} \frac{|\boldsymbol{\Phi}(v \otimes w)|}{\|v\|_V \|w\|_W}. \tag{4.30a}$$

2) Let $\boldsymbol{\Phi} = \sum_i \varphi_i \otimes \psi_i$ and set $\varphi := \sum_i \psi_i(w)\varphi_i \in V^*$ for some fixed $w \in W$. Then

$$\begin{aligned}\sup_{0\neq v\in V} \frac{|\sum_i \varphi_i(v)\psi_i(w)|}{\|v\|} &= \sup_{0\neq v\in V} \frac{|\varphi(v)|}{\|v\|} \underset{\text{Lemma 4.19}}{=} \sup_{0\neq v^{**}\in V^{**}} \frac{|v^{**}(\varphi)|}{\|v^{**}\|^{**}} \\ &= \sup_{0\neq v^{**}\in V^{**}} \frac{|\sum_i v^{**}(\varphi_i)\psi_i(w))|}{\|v^{**}\|^{**}}.\end{aligned}$$

Similarly, the supremum over w can be replaced by a supremum over w^{**} :

$$\sup_{\substack{0\neq v\in V\\0\neq w\in W}} \frac{|\sum_i \varphi_i(v)\psi_i(w)|}{\|v\|\,\|w\|} = \sup_{\substack{0\neq v^{**}\in V^{**}\\0\neq w^{**}\in W^{**}}} \frac{|(v^{**}\otimes w^{**})\,(\boldsymbol{\Phi})|}{\|v^{**}\|^{**}\,\|w^{**}\|^{**}} \quad \text{for } \boldsymbol{\Phi} = \sum_i \varphi_i \otimes \psi_i. \tag{4.30b}$$

3) The left-hand side of (4.30b) coincides with the right-hand side of (4.30a), since $|\boldsymbol{\Phi}(v \otimes w)| = |\sum_i \varphi_i(v)\psi_i(w)|$. The right-hand side of (4.30b) is the definition of $\|\boldsymbol{\Phi}\|_{\vee(V^*,W^*)}$. Together, $\|\boldsymbol{\Phi}\|^*_{\wedge(V,W)} = \|\boldsymbol{\Phi}\|_{\vee(V^*,W^*)}$ is shown. □

[12] The reverse statement is not true, but nearly (see [45, §I.6]).

Corollary 4.59. The norm $\|\cdot\|_{\wedge(V,W)}$ satisfies not only (4.26), but also

$$\|\varphi\otimes\psi\|^*_{\wedge(V,W)} = \|\varphi\|_{V^*}\,\|\psi\|_{W^*} \qquad \text{for all } \varphi\in V^*,\psi\in W^* \quad \text{(cf. (4.29))}.$$

Proof. This is the crossnorm property for $\|\cdot\|_{\vee(V^*,W^*)} = \|\cdot\|^*_{\wedge(V,W)}$ stated in Lemma 4.54. □

According to Remark 3.54, we consider $V^*\subset V'$ as a subspace of $L(V\otimes_a W,W)$ via $\varphi\in V^*\mapsto\varphi\big(\sum_i v_i\otimes w_i\big)=\sum_i\varphi\,(v_i\otimes w_i)=\sum_i\varphi(v_i)\,w_i\in W$. The crucial question is, whether the map $\mathbf{x}\mapsto\varphi(\mathbf{x})$ is continuous, i.e., whether $V^*\subset\mathcal{L}\,(V\otimes_a W,W)$. The supposition $\|\cdot\|\gtrsim\|\cdot\|_{\vee(V,W)}$ of the next proposition is satisfied for all reasonable crossnorms (cf. §4.2.9 and Proposition 4.55).

Proposition 4.60. *If $V\otimes_a W$ is equipped with a norm $\|\cdot\|\gtrsim\|\cdot\|_{\vee(V,W)}$, the embedding $V^*\subset\mathcal{L}\,(V\otimes_a W,W)$ is valid. In particular,*

$$\|\varphi(\mathbf{x})\|_W\le\|\varphi\|_{V^*}\,\|\mathbf{x}\|_{\vee(V,W)} \qquad \textit{for all } \varphi\in V^* \textit{ and } \mathbf{x}\in V\otimes_a W. \tag{4.31}$$

An analogous result holds for $W^*\subset\mathcal{L}\,(V\otimes_a W,V)$.

Proof. For $w:=\varphi(\mathbf{x})\in W$ choose $\psi\in W^*$ with $\psi(w)=\|w\|_W$ and $\|\psi\|_{W^*}=1$ (cf. Theorem 4.15). Then

$$\begin{aligned}\|\varphi(\mathbf{x})\|_W &= |\psi(\varphi(\mathbf{x}))| \underset{\mathbf{x}=\sum_i v_i\otimes w_i}{=} \Big|\psi\Big(\sum\nolimits_i\varphi(v_i)w_i\Big)\Big|\\ &= \Big|\sum\nolimits_i\varphi(v_i)\psi(w_i)\Big| = |(\varphi\otimes\psi)\,(\mathbf{x})| \le \|\varphi\otimes\psi\|^*_{\vee(V,W)}\,\|\mathbf{x}\|_{\vee(V,W)} \underset{(4.29)}{=}\\ &= \|\varphi\|_{V^*}\,\|\psi\|_{W^*}\,\|\mathbf{x}\|_{\vee(V,W)} = \|\varphi\|_{V^*}\,\|\mathbf{x}\|_{\vee(V,W)}\end{aligned}$$

proves (4.31). □

Corollary 4.61. Let V and W be two Banach spaces, where either V or W are finite dimensional. Equip $V\otimes_a W$ with a norm $\|\cdot\|\gtrsim\|\cdot\|_{\vee(V,W)}$. Then $V\otimes_a W$ is already complete.

Proof. Let $\dim(V)=n$ and choose a basis $\{v_1,\ldots,v_n\}$ of V. Then it is easy to see that all tensors from $V\otimes_a W$ may be written as $\sum_{i=1}^n v_i\otimes w_i$ with some $w_i\in W$. Let $\mathbf{x}_k=\sum_{i=1}^n v_i\otimes w_i^k\in V\otimes_a W$ be a Cauchy sequence. V^* has a dual basis $\{\varphi_1,\ldots,\varphi_n\}$ with $\varphi_\nu(v_\mu)=\delta_{\nu\mu}$ (cf. (2.1)). The embedding $V^*\subset\mathcal{L}\,(V\otimes_a W,W)$ discussed above, yields $\varphi_i(\mathbf{x}_k)=w_i^k$. Since, by Proposition 4.60, $\varphi_i:\mathbf{x}_k\mapsto w_i^k$ is continuous, also $\big(w_i^k\big)_{k\in\mathbb{N}}$ is a Cauchy sequence. Since W is a Banach space, $w_i^k\to\psi_i\in W$ proves $\lim\mathbf{x}_k=\sum_{i=1}^n\varphi_i\otimes\psi\in V\otimes_a W$. □

Exercise 4.62. Let $U\subset V$ be a closed subspace. Show that the norm $\|\cdot\|=\|\cdot\|_{\vee(U,W)}$ (involving functionals $\psi\in U^*$) and the restriction of $\|\cdot\|=\|\cdot\|_{\vee(V,W)}$ to $U\otimes W$ (involving $\varphi\in V^*$) lead to the same closed subspace $U\otimes_{\|\cdot\|}W$.

4.2.8 Examples

Again, we consider the spaces $\ell^p(I)$ and $\ell^p(J)$. To simplify the reasoning, we first restrict the analysis to finite index sets $I = J = \{1, \ldots, n\}$. The duals of $\ell^p(I)$ and $\ell^p(J)$ for $1 \le p < \infty$ are $\ell^q(I)$ and $\ell^q(J)$ with $\frac{1}{p} + \frac{1}{q} = 1$ (cf. Example 4.22). Let $\varphi \in \ell^q(I)$ and $\psi \in \ell^q(J)$. The definition of $\|\cdot\|_\vee$ makes use of $(\varphi \otimes \psi)(\mathbf{x})$, where for $\mathbf{x} = v \otimes w$ the definition $(\varphi \otimes \psi)(v \otimes w) = \varphi(v) \cdot \psi(w)$ holds. The interpretation of $\varphi(v)$ for a vector $v \in \ell^p(I)$ and a (dual) vector $\varphi \in \ell^q(I)$ is $\varphi(v) := \varphi^\mathsf{T} v$. Similarly, $\psi(w) = \psi^\mathsf{T} w$. Elements from $\ell^p(I) \otimes \ell^p(J) = \ell^p(I \times J)$ are standard $n \times n$ matrices, which we shall denote by M. We recall that $v \otimes w$ ($v \in \ell^p(I)$, $w \in \ell^p(J)$) corresponds to the matrix vw^T. Hence, with $M = vw^\mathsf{T}$, the definition of $(\varphi \otimes \psi)(v \otimes w)$ becomes $\varphi^\mathsf{T} M \psi \in \mathbb{K}$. This leads to the interpretation of $\|\mathbf{x}\|_{\vee(V,W)}$ in (4.28) by

$$\|M\|_{\vee(\ell^p(I),\ell^p(J))} = \sup_{\|\varphi\|_q = \|\psi\|_q = 1} \left|\varphi^\mathsf{T} M \psi\right| = \sup_{\varphi,\psi \neq 0} \frac{|\varphi^\mathsf{T} M \psi|}{\|\varphi\|_q \|\psi\|_q}.$$

Remark 4.63. (a) Let $1 \le p < \infty$ and assume $\#I > 1$ and $\#J > 1$. Then inequality $\|\cdot\|_{\vee(\ell^p(I),\ell^p(J))} \le \|\cdot\|_{\ell^p(I \times J)}$ holds, but the corresponding equality is not valid.
(b) For $p = 2$, $\|\cdot\|_{\vee(\ell^2(I),\ell^2(J))} = \|\cdot\|_{\mathrm{SVD},\infty} \neq \|\cdot\|_{\mathrm{SVD},2} = \|\cdot\|_{\ell^2(I \times J)}$ holds with $\|\cdot\|_{\mathrm{SVD},p}$ defined in (4.17).

Proof. To prove $\|\cdot\|_{\vee(\ell^p(I),\ell^p(J))} \neq \|\cdot\|_{\ell^p(I \times J)}$, choose $M = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ for the case[13] $I = J = \{1, 2\}$. For instance for $p = 1$, $\|M\|_{\ell^1(I \times J)} = 4$, while an elementary analysis of $\frac{|\varphi^\mathsf{T} M \psi|}{\|\varphi\|_\infty \|\psi\|_\infty}$ shows that $\|M\|_{\vee(\ell^1(I),\ell^1(J))} = 2$. □

In the case of the projective norm $\|\cdot\|_{\wedge(\ell^p(I),\ell^p(J))}$ from §4.2.4, we have seen in §4.2.5 that the norms $\|\cdot\|_{\wedge(\ell^p(I),\ell^p(J))}$ and $\|\cdot\|_{\ell^p(I \times J)}$ coincide for $p = 1$. Now, coincidence happens for $p = \infty$.

Remark 4.64. $\|\cdot\|_{\vee(\ell^\infty(I),\ell^\infty(J))} = \|\cdot\|_{\ell^\infty(I \times J)}$.

Proof. Choose unit vectors $e_I^{(i)} \in \ell^1(I)$, $e_J^{(j)} \in \ell^1(J)$. Then $(e_I^{(i)})^\mathsf{T} M e_J^{(j)} = M_{ij}$ and $\|e_I^{(i)}\|_1 = \|e_J^{(j)}\|_1 = 1$ show that $|(e_I^{(i)})^\mathsf{T} M e_j^J| / \big(\|e_I^{(i)}\|_1 \|e_J^{(j)}\|_1\big) = |M_{ij}|$ and

$$\|M\|_{\vee(\ell^\infty(I),\ell^\infty(J))} = \sup_{0 \neq \varphi \in (\ell^\infty(I))^*} \; \sup_{0 \neq \psi \in (\ell^\infty(J))^*} \frac{|\varphi^\mathsf{T} M \psi|}{\|\varphi\|_\infty^* \|\psi\|_\infty^*}.$$

A subset of $(\ell^\infty(I))^*$ is $\ell^1(I)$. The particular choice of the unit vectors $e_I^{(i)}$ $(i \in I)$ and $e_J^{(j)}$ $(j \in J)$ yields

$$\|M\|_{\vee(\ell^\infty(I),\ell^\infty(J))} \ge \sup_{i \in I, j \in J} \frac{|(e_I^{(i)})^\mathsf{T} M e_J^{(j)}|}{\|e_I^{(i)}\|_1 \|e_J^{(j)}\|_1} = \sup_{i \in I, j \in J} |M_{ij}| = \|M\|_{\ell^\infty(I \times J)}.$$

[13] This 2×2 example can be embedded in any model with $\#I, \#J \ge 2$.

On the other hand, for all $i \in I$ we have

$$|(M\psi)_i| = \Big|\sum_j M_{ij}\psi_j\Big| \le \Big(\sup_j |M_{ij}|\Big)\sum_j |\psi_j| \le \|M\|_{\ell^\infty(I\times J)}\,\|\psi\|_1$$

implying $\|M\psi\|_\infty \le \|M\|_{\ell^\infty(I\times J)}\,\|\psi\|_1$. Finally, $|\varphi^{\mathsf{T}}M\psi| \le \|\varphi\|_1\,\|M\psi\|_\infty$ proves the reverse inequality $\|M\|_{\vee(\ell^\infty(I),\ell^\infty(J))} \le \|M\|_{\ell^\infty(I\times J)}$. □

Among the function spaces, $C(I)$ is of interest, since again the supremum norm $\|\cdot\|_\infty$ is involved.

Remark 4.65. Let $V = (C(I), \|\cdot\|_{C(I)})$ and $W = (C(J), \|\cdot\|_{C(J)})$ with certain domains I and J (cf. Example 4.8). Then

$$\|\cdot\|_{\vee(C(I),C(J))} = \|\cdot\|_{C(I\times J)}\,.$$

Proof. 1) Let $f(\cdot,\cdot) \in C(I\times J)$. The duals $C(I)^*$ and $C(J)^*$ contain the delta functionals δ_x, δ_y ($x \in I$, $y \in J$) with $\|\delta_x\|_{C(I)^*} = 1$ (cf. (4.11)). Hence,

$$\begin{aligned}\|f\|_{\vee(C(I),C(J))} &= \sup_{\varphi,\psi\neq 0}\frac{|\,(\varphi\otimes\psi)\,f|}{\|\varphi\|_{C(I)^*}\,\|\psi\|_{C(J)^*}}\\ &\ge \sup_{x\in I,y\in J}|\,(\delta_x\otimes\delta_y)\,f| = \sup_{x\in I,y\in J}|f(x,y)| = \|f\|_{C(I\times J)}\,.\end{aligned}$$

2) For the reverse inequality, consider the function $f_y := f(\cdot,y)$ for fixed $y \in J$. Then $f_y \in C(I)$ has the norm $\|f_y\|_{C(I)} \le \|f\|_{C(I\times J)}$ for all $y \in J$. Application of φ yields $g(y) := \varphi(f_y)$ and $|g(y)| = |\varphi(f_y)| \le \|\varphi\|_{C(I)^*}\,\|f\|_{C(I\times J)}$ for all $y \in J$; hence, $\|g\|_{C(J)} \le \|\varphi\|_{C(I)^*}\,\|f\|_{C(I\times J)}$.

Application of ψ to g gives $|\,(\varphi\otimes\psi)\,f| = |\psi(g)| \le \|\psi\|_{C(J)^*}\,\|g\|_{C(J)} \le \|f\|_{C(I\times J)}\,\|\varphi\|_{C(I)^*}\,\|\psi\|_{C(J)^*}$ implying $\|f\|_{\vee(C(I),C(J))} \le \|f\|_{C(I\times J)}$. □

4.2.9 Reasonable Crossnorms

Now, we combine the inequalities (4.14) and (4.26) (with $C = 1$), i.e., we require, simultaneously, continuity of $\otimes : (V, \|\cdot\|_V) \times (W, \|\cdot\|_W) \to (V \otimes_a W, \|\cdot\|)$ and $\otimes : (V^*, \|\cdot\|_{V^*}) \times (W^*, \|\cdot\|_{W^*}) \to (V^* \otimes_a W^*, \|\cdot\|^*)$. Note that $\|\cdot\|^*$ is the dual norm of $\|\cdot\|$.

Definition 4.66. A norm $\|\cdot\|$ on $V \otimes_a W$ is a *reasonable crossnorm*,[14] if $\|\cdot\|$ satisfies

$$\|v\otimes w\| \le \|v\|_V\,\|w\|_W \qquad \text{for all } v \in V \text{ and } w \in W, \tag{4.32a}$$

$$\|\varphi\otimes\psi\|^* \le \|\varphi\|_{V^*}\,\|\psi\|_{W^*} \qquad \text{for all } \varphi \in V^* \text{ and } \psi \in W^*. \tag{4.32b}$$

[14] Also the name ‘dualisable crossnorm’ has been used (cf. [172]). Schatten [167] used the term ‘crossnorm whose associate is a crossnorm’ (‘associate norm’ means dual norm).

Lemma 4.67. *If $\|\cdot\|$ is a reasonable crossnorm, then (4.32c,d) holds:*

$$\|v\otimes w\| \;= \|v\|_V\,\|w\|_W \qquad \text{for all } v\in V \text{ and } w\in W, \tag{4.32c}$$
$$\|\varphi\otimes\psi\|^* = \|\varphi\|_{V^*}\,\|\psi\|_{W^*} \qquad \text{for all } \varphi\in V^* \text{ and } \psi\in W^*. \tag{4.32d}$$

Proof. 1) Note that (4.32b) is (4.26) with $C=1$. By Lemma 4.50, inequality (4.27) holds with $C=1$, i.e., $\|v\otimes w\|\ge\|v\|_V\,\|w\|_W$. Together with (4.32a) we obtain (4.32c).

2) Similarly, (4.32a) is (4.14) with $C=1$. Lemma 4.51 proves $\|\varphi\otimes\psi\|^*\ge\|\varphi\|_{V^*}\,\|\psi\|_{W^*}$ so that, together with (4.32b), identity (4.32d) follows. □

By the previous lemma, an equivalent definition of a reasonable crossnorm $\|\cdot\|$ is: $\|\cdot\|$ and $\|\cdot\|^*$ are crossnorms.

Lemma 4.45, Corollary 4.59, and Lemma 4.54 prove that the norms $\|\cdot\|_{\wedge(V,W)}$ and $\|\cdot\|_{\vee(V,W)}$ are particular reasonable crossnorms. Furthermore, Lemma 4.56, together with Propositions 4.46 and 4.55, shows the next statement.

Proposition 4.68. *$\|\cdot\|_{\vee(V,W)}$ is the weakest and $\|\cdot\|_{\wedge(V,W)}$ is the strongest reasonable crossnorm, i.e., any reasonable crossnorm $\|\cdot\|$ satisfies*

$$\|\cdot\|_{\vee(V,W)}\lesssim\|\cdot\|\lesssim\|\cdot\|_{\wedge(V,W)}\,. \tag{4.33}$$

Proposition 4.69. *If $\|\cdot\|$ is a reasonable crossnorm on $V\otimes W$, then also $\|\cdot\|^*$ is a reasonable crossnorm on $V^*\otimes W^*$.*

Proof. 1) We have to show that $\|\cdot\|^*$ satisfies (4.32a,b) with $\|\cdot\|$, V, W replaced by $\|\cdot\|^*$, V^*, W^*. The reformulated inequality (4.32a) is (4.32b). Hence, this condition is satisfied by assumption. It remains to show the reformulated version of (4.32b):

$$\|v^{**}\otimes w^{**}\|^{**}\le\|v^{**}\|_{V^{**}}\,\|w^{**}\|_{W^{**}} \quad \text{for all } v^{**}\in V^{**},\ w^{**}\in W^{**}. \tag{4.34a}$$

2) By Proposition 4.68, $\|\cdot\|_{\vee(V,W)}\le\|\cdot\|\le\|\cdot\|_{\wedge(V,W)}$ is valid. Applying Lemma 4.18 twice, we see that $\|\cdot\|^{**}_{\vee(V,W)}\le\|\cdot\|^{**}\le\|\cdot\|^{**}_{\wedge(V,W)}$ holds for the bidual norm; in particular,

$$\|v^{**}\otimes w^{**}\|^{**}\le\|v^{**}\otimes w^{**}\|^{**}_{\wedge(V,W)}\,. \tag{4.34b}$$

Proposition 4.58 states that $\|\cdot\|^*_{\wedge(V,W)}=\|\cdot\|_{\vee(V^*,W^*)}$, which implies

$$\|\cdot\|^{**}_{\wedge(V,W)}=\|\cdot\|^*_{\vee(V^*,W^*)}\,. \tag{4.34c}$$

Since $\|\cdot\|_{\vee(V^*,W^*)}$ is a reasonable crossnorm on $V^*\otimes W^*$ (cf. Proposition 4.68), it satisfies the corresponding inequality (4.32b):

$$\|v^{**}\otimes w^{**}\|^*_{\vee(V^*,W^*)}\le\|v^{**}\|_{V^{**}}\,\|w^{**}\|_{W^{**}} \quad \text{for all } v^{**}\in V^{**},\ w^{**}\in W^{**}. \tag{4.34d}$$

Now, the equations (4.34b-d) prove (4.34a). □

4.2.10 Examples and Counterexamples

Example 4.70 (ℓ^2). Let $V = \ell^2(I)$ and $W = \ell^2(J)$ for finite or countable index sets I, J with norms $\|\cdot\|_{\ell^2(I)}$ and $\|\cdot\|_{\ell^2(J)}$. Then all norms $\|\cdot\|_{\mathrm{SVD},p}$ for $1 \le p \le \infty$ are reasonable crossnorms on $\ell^2(I) \otimes_a \ell^2(J)$. In particular,

$$\begin{aligned} \|\cdot\|_{\vee(\ell^2(I),\ell^2(J))} &= \|\cdot\|_{\mathrm{SVD},\infty} \le \|\cdot\|_{\mathrm{SVD},p} \le \|\cdot\|_{\mathrm{SVD},q} \\ &\le \|\cdot\|_{\mathrm{SVD},1} \;\; = \|\cdot\|_{\wedge(\ell^\infty(I),\ell^\infty(J))} \qquad \text{for all } p \ge q. \end{aligned}$$

Example 4.71 (ℓ^p). $\|\cdot\|_{\ell^p(I\times J)}$ is a reasonable crossnorm for $1 \le p < \infty$.

Proof. (4.15) proves that $\|\cdot\|_{\ell^p(I\times J)}$ satisfies (4.32a). The same statement (4.15) for p replaced by q (defined by $\frac{1}{p} + \frac{1}{q} = 1$) shows (4.32b). □

The next example can be shown analogously.

Example 4.72 (L^p). Let $V = L^p(I)$ and $W = L^p(J)$ for intervals I and J. Then $\|\cdot\|_{L^p(I\times J)}$ is a reasonable norm on $V \otimes_a W = L^p(I \times J)$ for $1 \le p < \infty$.

For the next example we recall that

$$\|f\|_{C^1(I)} = \max_{x\in I} \{|f(x)|, |f'(x)|\}$$

is the norm of continuously differentiable functions in one variable $x \in I \subset \mathbb{R}$. The name $\|\cdot\|_{1,\mathrm{mix}}$ of the following norm is derived from the mixed derivative involved.

Example 4.73. Let I and J be compact intervals in $\mathbb{R}$ and set $V = (C^1(I), \|\cdot\|_{C^1(I)})$, $W = (C^1(J), \|\cdot\|_{C^1(J)})$. For the tensor space $V \otimes_a W$ we introduce the *mixed norm*

$$\begin{aligned} \|\varphi\|_{C^1_{\mathrm{mix}}(I\times J)} &:= \|\varphi\|_{1,\mathrm{mix}} \qquad (4.35) \\ &:= \max_{(x,y)\in I\times J} \left\{ |\varphi(x,y)|, \left|\frac{\partial\varphi(x,y)}{\partial x}\right|, \left|\frac{\partial\varphi(x,y)}{\partial y}\right|, \left|\frac{\partial^2\varphi(x,y)}{\partial x \partial y}\right| \right\}. \end{aligned}$$

Then, $\|\cdot\|_{1,\mathrm{mix}}$ is a reasonable crossnorm.

Proof. $\|f \otimes g\|_{C^1_{\mathrm{mix}}(I\times J)} \le \|f\|_{C^1(I)} \|g\|_{C^1(J)}$ is easy to verify. The proof of (4.32b) uses similar ideas as those from the proof of Remark 4.65. □

However, the standard norm for $C^1(I \times J)$ is

$$\|\varphi\|_{C^1(I\times J)} = \max_{(x,y)\in I\times J} \left\{ |\varphi(x,y)|, \left|\frac{\partial}{\partial x}\varphi(x,y)\right|, \left|\frac{\partial}{\partial y}\varphi(x,y)\right| \right\}.$$

As $\|\cdot\|_{C^1(I\times J)} \le \|\cdot\|_{C^1_{\mathrm{mix}}(I\times J)}$, inequality $\|f \otimes g\|_{C^1(I\times J)} \le \|f\|_{C^1(I)} \|g\|_{C^1(J)}$ proves (4.32a). However, the second inequality (4.32b) characterising a reasonable crossnorm cannot be satisfied. For a counterexample, choose the continuous functionals $\delta'_{x_0} \in V^*$ ($x_0 \in I$) and $\delta'_{y_0} \in V^*$ ($y_0 \in J$) defined by $\delta'_{x_0}(f) = -f'(x_0)$

and $\delta'_{y_0}(g) = -g'(y_0)$. Then $\left(\delta'_{x_0} \otimes \delta'_{y_0}\right)(f \otimes g) = f'(x_0)g'(y_0)$ cannot be bounded by $\|f \otimes g\|_{C^1(I\times J)}$, since the term $|\frac{\partial^2}{\partial x \partial y}\varphi(x,y)|$ from (4.35) is missing. For the treatment of this situation we refer to §4.3.6.

An analogous situation happens for $H^{1,p}(I \times J) = H^{1,p}(I) \otimes_{1,p} H^{1,p}(J)$ from Example 4.41. As mentioned in Example 4.41, $\|\cdot\|_{H^{1,p}(I\times J)}$ is no crossnorm. Furthermore, it does not satisfy (4.26). On the other hand, $H^{1,p}_{\mathrm{mix}}(I \times J)$ allows the crossnorm $(\|f\|_p^p + \|\partial f/\partial x\|_p^p + \|\partial f/\partial y\|_p^p + \left\|\partial^2 f/\partial x \partial y\right\|_p^p)^{1/p}$.

The anisotropic Sobolev space $H^{(1,0),p}(I_1 \times I_2) = H^{1,p}(I_1) \otimes_{(1,0),p} L^p(I_2)$ is introduced in Example 4.42.

Remark 4.74. The norm $\|\cdot\|_{(1,0),p} = \|\cdot\|_{H^{(1,0),p}(I_1\times I_2)}$ on $H^{(1,0),p}(I_1 \times I_2)$ is a reasonable crossnorm for $1 \le p < \infty$.

Proof. As already stated in Example 4.42, the norm satisfies (4.32c). It remains to prove (4.32b). The functionals $\varphi \in V^* := (H^{1,p}(I_1))^*$ and $\psi \in W^* := L^p(I_2)^*$ may be normalised: $\|\varphi\|_{V^*} = \|\psi\|_{W^*} = 1$. Then $\|\varphi \otimes \psi\|^* \le 1$ is to be proved. By definition of $\|\cdot\|^*$, it is sufficient to show

$$|(\varphi \otimes \psi)(f)| \le \|f\|_{(1,0),p} \qquad \text{for all } f \in H^{(1,0),p}(I_1 \times I_2).$$

Next, we may restrict $f \in H^{(1,0),p}(I_1 \times I_2)$ to the dense subset $f \in C^\infty(I_1 \times I_2)$. As stated in Example 4.22b, the dual space W^* can be identified with $L^q(I_2)$, i.e., $\psi \in L^q(I_2)$ and $\|\psi\|_q = 1$. Application of $\psi \in W^*$ to f yield the following function of $x \in I_1$:

$$F(x) := \int_{I_2} f(x,y)\psi(y)\mathrm{d}y \in C^\infty(I_1).$$

The functional $\varphi \in V^*$ acts with respect to the x-variable:

$$(\varphi \otimes \psi)(f) = \varphi(F) = \int_{I_2} \varphi\left[f(\cdot,y)\right]\psi(y)\mathrm{d}y.$$

For a fixed $y \in I_2$, the estimate $|\varphi\left[f(\cdot,y)\right]| \le \|\varphi\|_{V^*}\|f(\cdot,y)\|_{1,p} = \|f(\cdot,y)\|_{1,p}$ implies that

$$\begin{aligned}
|(\varphi \otimes \psi)(f)| &\le \int_{I_2} \|f(\cdot,y)\|_{1,p}\,\psi(y)\mathrm{d}y \\
&\le \left(\int_{I_2} \|f(\cdot,y)\|_{1,p}^p\,\mathrm{d}y\right)^{1/p}\left(\int_{I_2} |\psi(y)|^q\,\mathrm{d}y\right)^{1/q} = \underbrace{\|\psi\|_q}_{=1}\sqrt[p]{\int_{I_2} \|f(\cdot,y)\|_{1,p}^p\,\mathrm{d}y} \\
&= \sqrt[p]{\int_{I_2}\left(\int_{I_1} |f(x,y)|^p + \left|\frac{\partial}{\partial x} f(x,y)\right|^p \mathrm{d}x\right)\mathrm{d}y} = \|f\|_{(1,0),p}\,. \qquad \square
\end{aligned}$$

Although the solution of elliptic partial differential equations is usually a function of the standard Sobolev spaces and not of mixed spaces like $H^{1,p}_{\mathrm{mix}}$, there are important exceptions. As proved by Yserentant [199], the solutions of the electronic Schrödinger equation have mixed regularity because of the additional Pauli principle (i.e., the solutions must be antisymmetric).

4.2.11 Reflexivity

Let $\|\cdot\|$ be a crossnorm norm on $V \otimes_a W$. The dual space of $V \otimes_{\|\cdot\|} W$ or[15] $V \otimes_a W$ is $\left(V \otimes_{\|\cdot\|} W\right)^*$. From (4.25) we derive that

$$V^* \otimes_{\|\cdot\|^*} W^* \subset \left(V \otimes_{\|\cdot\|} W\right)^* . \tag{4.36}$$

Lemma 4.75. *Assume that $V \otimes_{\|\cdot\|} W$ is equipped with a reasonable crossnorm $\|\cdot\|$ and is reflexive. Then the identity*

$$V^* \otimes_{\|\cdot\|^*} W^* = \left(V \otimes_{\|\cdot\|} W\right)^* \tag{4.37}$$

holds. Furthermore, the spaces V and W are reflexive.

Proof. 1) For an indirect proof assume that (4.37) is not valid. Then there is some $\phi \in (V \otimes_{\|\cdot\|} W)^*$ with $\phi \notin V^* \otimes_{\|\cdot\|^*} W^*$. By Hahn-Banach there is some bidual $\Phi \in (V \otimes_{\|\cdot\|} W)^{**}$ such that $\Phi(\phi) \neq 0$, while $\Phi(\omega) = 0$ for all $\omega \in V^* \otimes_{\|\cdot\|^*} W^*$. Because of reflexivity, $\Phi(\phi)$ has a representation as $\phi(x_\Phi)$ for some $0 \neq x_\Phi \in V \otimes_{\|\cdot\|} W$. As $0 \neq x_\Phi$ implies $\|x_\Phi\|_{\vee(V,W)} > 0$, there is some $\omega = \varphi \otimes \psi \in V^* \otimes_a W^*$ such that $|\omega(x_\Phi)| > 0$. This is in contradiction to $0 = \Phi(\omega) = \omega(x_\Phi)$ for all $\omega \in V^* \otimes_a W^*$. Hence, identity (4.37) must hold.

2) The statement analogous to (4.36) for the dual spaces is

$$V^{**} \otimes_{\|\cdot\|^{**}} W^{**} \subset \left(V^* \otimes_{\|\cdot\|^*} W^*\right)^* \underset{(4.37)}{=} \left(V \otimes_{\|\cdot\|} W\right)^{**} = V \otimes_{\|\cdot\|} W,$$

implying $V^{**} \subset V$ and $W^{**} \subset W$. Together with the general property $V \subset V^{**}$ and $W \subset W^{**}$, we obtain reflexivity of V and W. □

The last lemma shows that reflexivity of the Banach spaces V and W is necessary for $V \otimes_{\|\cdot\|} W$ to be reflexive. One might expect that reflexivity of V and W is also sufficient, i.e., the tensor product of reflexive spaces is again reflexive. This is wrong as the next example shows (for a proof see Schatten [167, p. 139]; note that the Banach spaces $\ell^p(\mathbb{N})$ are reflexive for $1 < p < \infty$).

Example 4.76. $\ell^p(\mathbb{N}) \underset{\vee}{\otimes} \ell^q(\mathbb{N})$ for $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ is non-reflexive.

4.2.12 Uniform Crossnorms

Let $(V \otimes_a W, \|\cdot\|)$ be a tensor space with crossnorm $\|\cdot\|$ and consider operators $A \in \mathcal{L}(V, V)$ and $B \in \mathcal{L}(W, W)$ with operator norms $\|A\|_{V \leftarrow V}$ and $\|B\|_{W \leftarrow W}$. As discussed in §3.3.2.1, $A \otimes B$ is defined on elementary tensors $v \otimes w$ via

$$(A \otimes B)(v \otimes w) := (Av) \otimes (Bw) \in V \otimes_a W.$$

[15] A Banach space X and any dense subspace $X_0 \subset X$ yield the same dual space $X^* = X_0^*$.

While $A \otimes B$ is well-defined on finite linear combinations from $V \otimes_a W$, the question is, whether $A \otimes B : V \otimes_a W \to V \otimes_a W$ is (uniformly) bounded, i.e., $A \otimes B \in \mathcal{L}(V \otimes_a W, V \otimes_a W)$. In the positive case, $A \otimes B$ also belongs to $\mathcal{L}(V \otimes_{\|\cdot\|} W, V \otimes_{\|\cdot\|} W)$. For elementary tensors, the estimate

$$\begin{aligned}&\|(A \otimes B)\,(v \otimes w)\| = \|(Av) \otimes (Bw)\| = \|Av\|\,\|Bw\| \\ &\leq \|A\|_{V\leftarrow V}\,\|B\|_{W\leftarrow W}\,\|v\|_V\,\|w\|_W = \|A\|_{V\leftarrow V}\,\|B\|_{W\leftarrow W}\,\|v \otimes w\|\end{aligned} \tag{4.38}$$

follows by the crossnorm property. However, this inequality does not automatically extend to general tensors from $V \otimes_a W$. Instead, the desired estimate is subject of the next definition (cf. [167]).

Definition 4.77. A crossnorm on $V \otimes_a W$ is called *uniform*, if $A \otimes B$ belongs to $\mathcal{L}(V \otimes_a W, V \otimes_a W)$ with the operator norm

$$\|A \otimes B\|_{V\otimes_a W\leftarrow V\otimes_a W} \leq \|A\|_{V\leftarrow V}\,\|B\|_{W\leftarrow W}\,. \tag{4.39}$$

Taking the supremum over all $v \otimes w$ with $\|v \otimes w\| = 1$, one concludes from (4.38) that $\|A \otimes B\|_{V\otimes_a W\leftarrow V\otimes_a W} \geq \|A\|_{V\leftarrow V}\,\|B\|_{W\leftarrow W}$. Therefore, one may replace inequality (4.39) by

$$\|A \otimes B\|_{V\otimes_a W\leftarrow V\otimes_a W} = \|A\|_{V\leftarrow V}\,\|B\|_{W\leftarrow W}\,.$$

Proposition 4.78. *$\|\cdot\|_{\wedge(V,W)}$ and $\|\cdot\|_{\vee(V,W)}$ are uniform crossnorms.*

Proof. 1) Let $\mathbf{x} = \sum_i v_i \otimes w_i \in V \otimes_a W$ and $A \in \mathcal{L}(V,V)$, $B \in \mathcal{L}(W,W)$. Then

$$\begin{aligned}\|(A \otimes B)\,(\mathbf{x})\|_{\wedge(V,W)} &= \left\|\sum\nolimits_i (Av_i)\otimes(Bw_i)\right\|_{\wedge(V,W)} \leq \sum\nolimits_i \|Av_i\|_V\|Bw_i\|_W \\ &\leq \|A\|_{V\leftarrow V}\,\|B\|_{W\leftarrow W}\sum\nolimits_i \|v_i\|_V\|w_i\|_W\end{aligned}$$

holds for all representations $\mathbf{x} = \sum_i v_i \otimes w_i$. The infimum over all representations yields

$$\|(A \otimes B)\,(\mathbf{x})\|_{\wedge(V,W)} \leq \|A\|_{V\leftarrow V}\,\|B\|_{W\leftarrow W}\,\|\mathbf{x}\|_{\wedge(V,W)}\,,$$

i.e., (4.39) holds for $\left(V \otimes_a W, \|\cdot\|_{\wedge(V,W)}\right)$.

2) Let $\mathbf{x} = \sum_i v_i \otimes w_i \in V \otimes_a W$ and note that

$$\begin{aligned}\|(A \otimes B)\,\mathbf{x}\|_{\vee(V,W)} &= \sup_{0\neq\varphi\in V^*,0\neq\psi\in W^*}\frac{|(\varphi\otimes\psi)\,((A\otimes B)\,\mathbf{x})|}{\|\varphi\|_{V^*}\,\|\psi\|_{W^*}} \\ &= \sup_{0\neq\varphi\in V^*,0\neq\psi\in W^*}\frac{|\sum_i (\varphi\otimes\psi)\,((Av_i)\otimes(Bw_i))|}{\|\varphi\|_{V^*}\,\|\psi\|_{W^*}} \\ &= \sup_{0\neq\varphi\in V^*,0\neq\psi\in W^*}\frac{|\sum_i(\varphi(Av_i)\cdot\psi(Bw_i)|}{\|\varphi\|_{V^*}\,\|\psi\|_{W^*}}.\end{aligned}$$

By Definition 4.20, $A^* \in \mathcal{L}(V^*, V^*)$ and $B^* \in \mathcal{L}(W^*, W^*)$ satisfy

$$\begin{aligned}\left|\sum\nolimits_i \varphi(Av_i) \cdot \psi(Bw_i)\right| &= \left|\sum\nolimits_i \left(A^*\varphi\right)(v_i) \cdot \left(B^*\psi\right)(w_i)\right| \\ &= \left|\left(\left(A^*\varphi\right) \otimes \left(B^*\psi\right)\right)(\mathbf{x})\right| .\end{aligned}$$

We continue:

$$\begin{aligned}\|(A \otimes B)\,\mathbf{x}\|_{\vee(V,W)} &= \sup_{0\neq\varphi\in V^*, 0\neq\psi\in W^*} \frac{\left|\left(\left(A^*\varphi\right) \otimes \left(B^*\psi\right)\right)(\mathbf{x})\right|}{\|\varphi\|_{V^*}\,\|\psi\|_{W^*}} \\ &= \sup_{0\neq\varphi, 0\neq\psi} \frac{\|A^*\varphi\|_{V^*}}{\|\varphi\|_{V^*}} \frac{\|B^*\psi\|_{W^*}}{\|\psi\|_{W^*}} \frac{\left|\left(\left(A^*\varphi\right) \otimes \left(B^*\psi\right)\right)(\mathbf{x})\right|}{\|A^*\varphi\|_{V^*}\,\|B^*\psi\|_{W^*}} .\end{aligned}$$

By Lemma 4.21, the inequalities $\frac{\|A^*\varphi\|_{V^*}}{\|\varphi\|_{V^*}} \leq \|A^*\|_{V^*\leftarrow V^*} = \|A\|_{V\leftarrow V}$ and $\frac{\|B^*\psi\|_{W^*}}{\|\psi\|_{W^*}} \leq \|B^*\|_{W^*\leftarrow W^*} = \|B\|_{W\leftarrow W}$ hold, while

$$\frac{\left|\left(\left(A^*\varphi\right) \otimes \left(B^*\psi\right)\right)(\mathbf{x})\right|}{\|A^*\varphi\|_{V^*}\,\|B^*\psi\|_{W^*}} \leq \|\mathbf{x}\|_{\vee(V,W)} .$$

Together, $\|(A \otimes B)\,\mathbf{x}\|_{\vee(V,W)} \leq \|A\|_{V\leftarrow V}\,\|B\|_{W\leftarrow W}\,\|\mathbf{x}\|_{\vee(V,W)}$ proves that also $\|\cdot\|_{\vee(V,W)}$ is uniform. □

By definition, a uniform crossnorm is a crossnorm. As shown in Simon [172], it is also a reasonable crossnorm.

Lemma 4.79. *A uniform crossnorm is a reasonable crossnorm.*

Proof. Let $\varphi \in V^*$ and $\psi \in W^*$ and choose some $0 \neq v \in V$ and $0 \neq w \in W$. Define the operator $\Phi \in \mathcal{L}(V, V)$ by $\Phi = v\varphi$ (i.e., $\Phi(x) = \varphi(x) \cdot v$) and, similarly, $\Psi \in \mathcal{L}(W, W)$ by $\Psi := w\psi$. The identities $\|\Phi\|_{V\leftarrow V} = \|v\|_V\,\|\varphi\|_{V^*}$ and $\|\Psi\|_{W\leftarrow W} = \|w\|_W\,\|\psi\|_{W^*}$ are valid as well as

$$(\Phi \otimes \Psi)(\mathbf{x}) = \left((\varphi \otimes \psi)(\mathbf{x})\right) \cdot (v \otimes w) \qquad \text{for all } \mathbf{x} \in \mathbf{X} := V \otimes_a W.$$

The crossnorm property yields

$$\|(\Phi \otimes \Psi)(\mathbf{x})\| = |(\varphi \otimes \psi)(\mathbf{x})|\,\|v \otimes w\| = |(\varphi \otimes \psi)(\mathbf{x})|\,\|v\|_V\,\|w\|_W ,$$

while the uniform crossnorm property allows the estimate

$$\begin{aligned}|(\varphi \otimes \psi)(\mathbf{x})|\,\|v\|_V\,\|w\|_W = \|(\Phi \otimes \Psi)(\mathbf{x})\| &\leq \|\Phi\|_{V\leftarrow V}\,\|\Psi\|_{W\leftarrow W}\,\|\mathbf{x}\| \\ &= \|v\|_V\,\|\varphi\|_{V^*}\,\|w\|_W\,\|\psi\|_{W^*}\,\|\mathbf{x}\| .\end{aligned}$$

Dividing by $\|v\|_V\,\|w\|_W \neq 0$, we obtain $|(\varphi \otimes \psi)(\mathbf{x})| \leq \|\varphi\|_{V^*}\,\|\psi\|_{W^*}\,\|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbf{X}$. Hence, $\|\cdot\|$ is a reasonable crossnorm. □

Proposition 4.80. *Suppose that the Banach spaces V and W are reflexive. If $\|\cdot\|$ is a uniform crossnorm on $V \otimes_a W$, then also $\|\cdot\|^*$ is a uniform and reasonable crossnorm on $V^* \otimes_a W^*$.*

Proof. By Lemma 4.79, $\|\cdot\|$ is a reasonable crossnorm, while by Proposition 4.69 also $\|\cdot\|^*$ is a reasonable crossnorm. To prove uniformity, let $A^* \in \mathcal{L}(V^*, V^*)$ and $B^* \in \mathcal{L}(W^*, W^*)$. Because of reflexivity, the adjoint operators of A^* and B^* are $A^{**} = A \in \mathcal{L}(V, V)$ and $B \in \mathcal{L}(W, W)$. For $\mathbf{x}^* = \sum_i \varphi_i \otimes \psi_i \in V^* \otimes_a W^*$ and $\mathbf{x} = \sum_j v_j \otimes w_j \in V \otimes_a W$ we have

$$\begin{aligned}
\left|\left((A^* \otimes B^*)(\mathbf{x}^*)\right)(\mathbf{x})\right| &= \left|\sum_i \sum_j (A^*\varphi_i)(v_j) \cdot (B^*\psi_i)(w_j)\right| \\
&= \left|\sum_i \sum_j \varphi_i(Av_j) \cdot \psi_i(Bw_j)\right| = \left|\mathbf{x}^*\left((A \otimes B)(\mathbf{x})\right)\right| \\
&\leq \|\mathbf{x}^*\|^* \, \|(A \otimes B)(\mathbf{x})\| \underset{\|\cdot\| \text{ uniform}}{\leq} \|\mathbf{x}^*\|^* \, \|A\|_{V \leftarrow V} \, \|B\|_{W \leftarrow W} \, \|\mathbf{x}\| .
\end{aligned}$$

From $\|(A^* \otimes B^*)(\mathbf{x}^*)\|^* = \sup_{\mathbf{x} \neq 0} \frac{|((A^* \otimes B^*)(\mathbf{x}^*))(\mathbf{x})|}{\|\mathbf{x}\|} \leq \|A\|_{V \leftarrow V} \|B\|_{W \leftarrow W} \|\mathbf{x}^*\|^*$, $\|A\|_{V \leftarrow V} = \|A^*\|_{V^* \leftarrow V^*}$, and $\|B\|_{W \leftarrow W} = \|B^*\|_{W^* \leftarrow W^*}$ we derive that the dual norm $\|\cdot\|^*$ is uniform. □

4.2.13 Nuclear and Compact Operators

Suppose that V and W are Banach spaces and consider the tensor space $V \otimes_a W^*$. The inclusion

$$V \otimes_a W^* \subset \mathcal{L}(W, V)$$

is defined via

$$(v \otimes \psi)(w) := \psi(w)v \in V \qquad \text{for all } v \in V,\ \psi \in W^*,\ w \in W.$$

Similarly as in Proposition 3.57a, $V \otimes_a W^*$ is interpreted as a subspace of $\mathcal{L}(W, V)$ and denoted by $\mathcal{F}(W, V)$. Elements $\Phi \in \mathcal{F}(W, V)$ are called *finite rank operators*. We recall Definition 4.12: $\mathcal{K}(W, V)$ is the set of compact operators.

Definition 4.81. A Banach space X has the *approximation property*, if for any compact set $K \subset X$ and $\varepsilon > 0$ there is $\Phi_{K,\varepsilon} \in \mathcal{F}(X, X)$ with $\sup_{x \in K} \|\Phi_{K,\varepsilon} x - x\|_V \leq \varepsilon$.

Proposition 4.82. *(a) The completion with respect to the operator norm $\|\cdot\|_{V \leftarrow W}$ from (4.6a) yields*

$$\overline{\mathcal{F}(W, V)} \subset \mathcal{K}(W, V).$$

(b) Sufficient for $\overline{\mathcal{F}(W, V)} = \mathcal{K}(W, V)$ is the approximation property of W^*.

Proof. $\Phi \in \mathcal{F}(W,V)$ is compact since its range is finite dimensional. Part (a) follows, because limits of compact operators are compact. For Part (b) see [139, p. 17]. □

Next, we relate the operator norm $\|\cdot\|_{V\leftarrow W}$ with the crossnorms of $V \otimes_{\|\cdot\|} W^*$.

Lemma 4.83. *$\|\Phi\|_{V\leftarrow W} \leq \|\Phi\|_{\vee(V,W^*)}$ holds for all $\Phi \in V \otimes_{\vee} W^*$. Reflexivity of W implies the equality $\|\cdot\|_{V\leftarrow W} = \|\cdot\|_{\vee(V,W^*)}$.*

Proof. $\|\Phi\|_{\vee(V,W^*)}$ is the supremum of $|(\varphi \otimes w^{**})(\Phi)|$ over all normalised $\varphi \in V^*$ and $w^{**} \in W^{**}$. Replacing W^{**} by its subspace W, we get a lower bound:

$$\begin{aligned}\|\Phi\|_{\vee(V,W^*)} &\geq \sup_{\|\varphi\|_{V^*}=\|w\|_W=1} |(\varphi \otimes w)(\Phi)| = \sup_{\|\varphi\|_{V^*}=\|w\|_W=1} |\varphi(\Phi(w))| \\ &= \sup_{\|w\|_W=1} \|\Phi(w)\|_V = \|\Phi\|_{V\leftarrow W}\,.\end{aligned}$$

If $W = W^{**}$, equality follows. □

Corollary 4.84. As all reasonable crossnorms $\|\cdot\|$ are stronger than $\|\cdot\|_{\vee(V,W^*)}$, we have $V \otimes_{\|\cdot\|} W^* \subset \overline{\mathcal{F}(W,V)} \subset \mathcal{K}(W,V)$. This holds in particular for $\|\cdot\|_{\wedge(V,W^*)}$.

The definition of nuclear operators can be found in Grothendieck [79].

Definition 4.85. $\mathcal{N}(W,V) := V \otimes_{\wedge} W^*$ is the space of *nuclear operators*.

If V and W are assumed to be Hilbert spaces, the infinite singular value decomposition enables further conclusions which will be given in §4.4.3.

Exercise 4.86. Show that for Banach spaces V and W, the dual $(V \otimes_{\wedge} W)^*$ is isomorphic to $\mathcal{L}(V, W^*)$.

4.3 Tensor Spaces of Order *d*

4.3.1 Continuity, Crossnorms

In the following, $\|\cdot\|_j$ are the norms associated with the vector spaces V_j, while $\|\cdot\|$ is the norm of the tensor space ${}_a\bigotimes_{j=1}^d V_j$ and the Banach tensor space ${}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$. Lemma 4.30b implies the following result.

Remark 4.87. Let $(V_j, \|\cdot\|_j)$ be normed vector spaces for $1 \leq j \leq d$. The d-fold tensor product

$$\bigotimes_{j=1}^{d} : \; V_1 \times \ldots \times V_d \quad \rightarrow \quad V_1 \otimes_a \ldots \otimes_a V_d$$

is continuous, if and only if there is some constant C such that

$$\left\| \bigotimes_{j=1}^{d} v^{(j)} \right\| \leq C \prod_{j=1}^{d} \|v^{(j)}\|_j \qquad \text{for all } v^{(j)} \in V_j \quad (1 \leq j \leq d)\,.$$

Again, we call $\|\cdot\|$ a *crossnorm*, if

$$\left\| \bigotimes_{j=1}^{d} v^{(j)} \right\| = \prod_{j=1}^{d} \|v^{(j)}\|_j \qquad \text{for all } v^{(j)} \in V_j \quad (1 \leq j \leq d) \tag{4.40}$$

holds for elementary tensors.

Similarly, we may consider the d-fold tensor product

$$\bigotimes_{j=1}^{d} : V_1^* \times \ldots \times V_d^* \to {}_a\bigotimes_{j=1}^{d} V_j^*$$

of the dual spaces. We recall that the normed space $\left({}_a\bigotimes_{j=1}^{d} V_j\,, \|\cdot\|\right)$ has a dual equipped with the dual norm $\left(({}_a\bigotimes_{j=1}^{d} V_j\,)^*, \|\cdot\|^*\right)$. We interpret $\varphi_1 \otimes \ldots \otimes \varphi_d \in {}_a\bigotimes_{j=1}^{d} V_j^*$ as functional on ${}_a\bigotimes_{j=1}^{d} V_j$, i.e., as an element of $({}_a\bigotimes_{j=1}^{d} V_j\,)^*$, via

$$(\varphi_1 \otimes \ldots \otimes \varphi_d) \left(v^{(1)} \otimes \ldots \otimes v^{(d)}\right) := \varphi_1(v^{(1)}) \cdot \varphi_2(v^{(2)}) \cdot \ldots \cdot \varphi_d(v^{(d)}).$$

Then, continuity of $\bigotimes_{j=1}^{d} : V_1^* \times \ldots \times V_d^* \to {}_a\bigotimes_{j=1}^{d} V_j^*$ is equivalent to

$$\left\| \bigotimes_{j=1}^{d} \varphi_j \right\|^* \leq C \prod_{j=1}^{d} \|\varphi_j\|_j^* \qquad \text{for all } \varphi_j \in V_j^* \ (1 \leq j \leq d)\,.$$

A crossnorm $\|\cdot\|$ on $\bigotimes_{j=1}^{d} V_j$ is called a *reasonable crossnorm*, if

$$\left\| \bigotimes_{j=1}^{d} \varphi_j \right\|^* = \prod_{j=1}^{d} \|\varphi_j\|_j^* \qquad \text{for all } \varphi_j \in V_j^* \ (1 \leq j \leq d)\,. \tag{4.41}$$

A crossnorm $\|\cdot\|$ on $\mathbf{V} := {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ is called *uniform crossnorm*, if elementary tensors $\mathbf{A} := \bigotimes_{j=1}^{d} A^{(j)}$ have the operator norm

$$\|\mathbf{A}\|_{\mathbf{V}\leftarrow\mathbf{V}} = \prod_{j=1}^{d} \|A^{(j)}\|_{V_j\leftarrow V_j} \qquad \left(A^{(j)} \in \mathcal{L}(V_j,V_j),\, 1 \leq j \leq d\right) \tag{4.42}$$

(we may write $\leq$ instead, but equality follows, compare Definition 4.77 and the following comment on page 119).

The proofs of Lemma 4.79 and Proposition 4.80 can easily be extended to d factors yielding the following result.

Lemma 4.88. *(a) A uniform crossnorm on $\bigotimes_{j=1}^{d} V_j$ is a reasonable crossnorm.*
(b) Let $\|\cdot\|$ be a uniform crossnorm on $\bigotimes_{j=1}^{d} V_j$ with reflexive Banach spaces V_j. Then $\|\cdot\|^*$ is a uniform and reasonable crossnorm on $\bigotimes_{j=1}^{d} V_j^*$.

4.3.2 Recursive Definition of the Topological Tensor Space

As mentioned in §3.2.4, the algebraic tensor space $\mathbf{V}_{\mathrm{alg}} := {}_a\bigotimes_{j=1}^d V_j$ can be constructed recursively by pairwise products:

$$\mathbf{X}_2^{\mathrm{alg}} := V_1 \otimes_a V_2, \quad \mathbf{X}_3^{\mathrm{alg}} := \mathbf{X}_2^{\mathrm{alg}} \otimes_a V_3, \ \ldots, \ \mathbf{V}_{\mathrm{alg}} := \mathbf{X}_d^{\mathrm{alg}} := \mathbf{X}_{d-1}^{\mathrm{alg}} \otimes_a V_d\,.$$

For a similar construction of the topological tensor space $\mathbf{V} := {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$, we need in addition suitable norms $\|\cdot\|_{\mathbf{X}_k}$ on $\mathbf{X}_k$ so that

$$\mathbf{X}_k := \mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k \qquad \text{for } k = 2, \ldots, d \text{ with } \mathbf{X}_1 := V_1$$

yielding $\mathbf{V} = \mathbf{X}_d$ with $\|\cdot\| = \|\cdot\|_{\mathbf{X}_d}$. In the case of a (reasonable) crossnorm, it is natural to require that also $\|\cdot\|_{\mathbf{X}_k}$ is a (reasonable) crossnorm.

The crossnorm property is not a property of $\|\cdot\|$ alone, but describes its relation to the norms of the generating normed spaces. For $d \geq 3$, different situations are possible as explained below.

Remark 4.89. There are two interpretations of the crossnorm property of $\|\cdot\|_{\mathbf{X}_k}$:

(i) A crossnorm on $\mathbf{X}_k = {}_{\|\cdot\|_{\mathbf{X}_k}}\bigotimes_{j=1}^k V_j$ requires

$$\left\| \bigotimes_{j=1}^k v^{(j)} \right\|_{\mathbf{X}_k} = \prod_{j=1}^k \|v^{(j)}\|_j \qquad (v^{(j)} \in V_j),$$

(ii) whereas a crossnorm on $\mathbf{X}_k = \mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k$ requires the stronger condition

$$\|\mathbf{x} \otimes v^{(k)}\|_{\mathbf{X}_k} = \|\mathbf{x}\|_{\mathbf{X}_{k-1}} \|v^{(k)}\|_k \quad \text{for } \mathbf{x} \in \mathbf{X}_{k-1} \text{ and } v^{(k)} \in V_k\,.$$

If, for $2 \leq k \leq d$, $\|\cdot\|_{\mathbf{X}_k}$ are crossnorms in the sense of Item (ii), $\|\cdot\|_{\mathbf{X}_k}$ is uniquely defined by the norm $\|\cdot\|$ of $\mathbf{V}$ via

$$\|\mathbf{x}\|_{\mathbf{X}_k} = \left\| \mathbf{x} \otimes \bigotimes_{j=k+1}^d v^{(j)} \right\| \quad \text{for arbitrary } v^{(j)} \in V_j \text{ with } \|v^{(j)}\|_j = 1. \tag{4.43}$$

Proof. The induction starts with $k = d-1$. The crossnorm property (ii) states that $\|\mathbf{x} \otimes v^{(d)}\| = \|\mathbf{x}\|_{\mathbf{X}_{d-1}} \|v^{(d)}\|_d = \|\mathbf{x}\|_{\mathbf{X}_{d-1}}$ for any normalised vector $v^{(d)} \in V_d$. The cases $k = d-2, \ldots, 2$ follow by recursion. □

Note that (4.43) requires that a crossnorm $\|\cdot\|_{\mathbf{X}_k}$ exists in the sense of (ii). Under the assumption that $\|\cdot\|$ is a uniform crossnorm on $\mathbf{V}$, we now show the opposite direction: The intermediate norms $\|\cdot\|_{\mathbf{X}_k}$ from (4.43) are well-defined and have the desired properties.

Proposition 4.90. *Assume that $\|\cdot\|$ is a uniform crossnorm on $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$. Then the definition (4.43) does not depend on the choice of $v^{(j)} \in V_j$ and the resulting norm $\|\cdot\|_{\mathbf{X}_k}$ is a uniform and reasonable crossnorm on ${}_{\|\cdot\|_{\mathbf{X}_k}}\bigotimes_{j=1}^k V_j$. Furthermore, $\|\cdot\|_{\mathbf{X}_k}$ is a reasonable crossnorm on $\mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k$.*

Proof. 1) It suffices to consider the case $k = d - 1$, so that definition (4.43) becomes $\|\mathbf{x}\|_{\mathbf{X}_{d-1}} := \|\mathbf{x} \otimes v^{(d)}\|$ with $\|v^{(d)}\|_d = 1$. There is some $\varphi^{(d)} \in V_d^*$ with $\|\varphi^{(d)}\|_d^* = 1$ and $\varphi^{(d)}(v^{(d)}) = 1$. Let $w^{(d)} \in V_d$ with $\|w^{(d)}\|_d = 1$ be another choice. Set $A^{(d)} := w^{(d)}\varphi^{(d)} \in \mathcal{L}(V_d, V_d)$, i.e., $A^{(d)}v = \varphi^{(d)}(v)w^{(d)}$. Because of $\|A^{(d)}\|_{V_d \leftarrow V_d} = \|\varphi^{(d)}\|_d^* \|w^{(d)}\|_d = 1$, the uniform crossnorm property (4.42) with $\mathbf{A} := \bigotimes_{j=1}^d A^{(j)}$, where $A^{(j)} = I$ for $1 \le j \le d-1$, implies

$$\|\mathbf{x} \otimes w^{(d)}\| = \|\mathbf{A}(\mathbf{x} \otimes v^{(d)})\| \le \|\mathbf{x} \otimes v^{(d)}\|.$$

Interchanging the rôles of $w^{(d)}$ and $v^{(d)}$, we obtain $\|\mathbf{x} \otimes v^{(d)}\| = \|\mathbf{x} \otimes w^{(d)}\|$. Obviously, $\|\cdot\|_{\mathbf{X}_{d-1}} = \|\cdot \otimes v^{(d)}\|$ is a norm on $\mathbf{X}_{d-1}$.

2) For $\mathbf{x} := \bigotimes_{j=1}^{d-1} v^{(j)}$ we form $\mathbf{x} \otimes v^{(d)} = \bigotimes_{j=1}^{d} v^{(j)}$ with some $\|v^{(d)}\|_d = 1$. The crossnorm property of $\|\cdot\|$ implies the crossnorm property of $\|\cdot\|_{\mathbf{X}_{d-1}}$ on $_{\|\cdot\|_{\mathbf{X}_{d-1}}}\bigotimes_{j=1}^{d-1} V_j$:

$$\Big\|\bigotimes_{j=1}^{d-1} v^{(j)}\Big\|_{\mathbf{X}_{d-1}} = \|\mathbf{x} \otimes v^{(d)}\| = \Big\|\bigotimes_{j=1}^{d} v^{(j)}\Big\|$$
$$= \prod_{j=1}^{d} \|v^{(j)}\|_j \underset{\|v^{(d)}\|_d = 1}{=} \prod_{j=1}^{d-1} \|v^{(j)}\|_{V_j \leftarrow V_j}.$$

Similarly, the uniform crossnorm property can be shown:

$$\left\|\left(\bigotimes_{j=1}^{d-1} A^{(j)}\right)\mathbf{x}\right\|_{\mathbf{X}_{d-1}} = \left\|\left(\left(A^{(1)} \otimes \ldots \otimes A^{(d-1)}\right)\mathbf{x}\right) \otimes v^{(d)}\right\|$$
$$= \left\|\left(A^{(1)} \otimes \ldots \otimes A^{(d-1)} \otimes I\right)\left(\mathbf{x} \otimes v^{(d)}\right)\right\|$$
$$\le \left(\prod_{j=1}^{d-1} \|A^{(j)}\|_{V_j \leftarrow V_j}\right) \underbrace{\|I\|_{V_d \leftarrow V_d}}_{=1} \|\mathbf{x} \otimes v^{(d)}\| = \prod_{j=1}^{d-1} \|A^{(j)}\|_{V_j \leftarrow V_j} \|\mathbf{x}\|_{\mathbf{X}_{d-1}}.$$

As a consequence, by Lemma 4.79, $\|\cdot\|_{\mathbf{X}_{d-1}}$ is also a reasonable crossnorm.

3) Now we consider $\mathbf{V}$ as the tensor space $\mathbf{X}_{d-1} \otimes_a V_d$ (interpretation (ii) of Remark 4.89). Let $\mathbf{v} := \mathbf{x} \otimes w_d$ with $\mathbf{x} \in \mathbf{X}_{d-1}$ and $0 \ne w^{(d)} \in V_d$. Set $v^{(d)} := w^{(d)}/\|w^{(d)}\|_d$. Then

$$\|\mathbf{v}\| = \|\mathbf{x} \otimes w^{(d)}\| = \|w^{(d)}\|_d \|\mathbf{x} \otimes v^{(d)}\|_{\mathbf{V}} = \|\mathbf{x}\|_{\mathbf{X}_{d-1}} \|w^{(d)}\|_d \tag{4.44a}$$

follows by definition (4.43) of $\|\mathbf{x}\|_{\mathbf{X}_{d-1}}$. This proves that $\|\cdot\|$ is a crossnorm on $\mathbf{X}_{d-1} \otimes_a V_d$.

Since $\|\cdot\|$ is not necessarily uniform on $\mathbf{X}_{d-1} \otimes_a V_d$, we need another argument to prove that $\|\cdot\|$ is a reasonable crossnorm on $\mathbf{X}_{d-1} \otimes V_d$. Let $\boldsymbol{\psi} \in \mathbf{X}_{d-1}^*$ and $\varphi^{(d)} \in V_d^*$. We need to prove that $\|\boldsymbol{\psi} \otimes \varphi^{(d)}\|^* = \|\boldsymbol{\psi}\|_{\mathbf{X}_{d-1}}^* \|\varphi^{(d)}\|_d^*$. Using the crossnorm property for an elementary tensor $\mathbf{v}$, we get

$$\frac{|(\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{v})|}{\|\mathbf{v}\|} = \frac{|\boldsymbol{\psi}(\mathbf{x})|}{\|\mathbf{x}\|_{\mathbf{X}_{d-1}}} \frac{|\varphi^{(d)}(v^{(d)})|}{\|v^{(d)}\|_d} \le \|\boldsymbol{\psi}\|_{\mathbf{X}_{d-1}}^* \|\varphi^{(d)}\|_d^* \quad \text{if } \mathbf{v} = \mathbf{x} \otimes v^{(d)} \ne 0. \tag{4.44b}$$

Taking the supremum over all $\mathbf{v} = \mathbf{x} \otimes v^{(d)} \neq 0$ ($\mathbf{x} \in \mathbf{X}_{d-1}$), we obtain

$$\|\boldsymbol{\psi} \otimes \varphi^{(d)}\|^* = \sup_{\mathbf{v}\in\mathbf{X}_d} \frac{|(\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{v})|}{\|\mathbf{v}\|} \geq \sup_{\mathbf{v}=\mathbf{x}\otimes v^{(d)}} \frac{|(\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{v})|}{\|\mathbf{v}\|} = \|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} \|\varphi^{(d)}\|^*_d.$$

Define the operator $\mathbf{A} := \bigotimes_{j=1}^d A^{(j)} \in \mathcal{L}(\mathbf{V}, \mathbf{V})$ by $A^{(j)} = I$ ($1 \leq j \leq d-1$) and $A^{(d)} = \hat{v}^{(d)}\varphi^{(d)}$ with $0 \neq \hat{v}^{(d)} \in V_d$. Then $\mathbf{Av}$ is an elementary vector of the form $\mathbf{x} \otimes \hat{v}^{(d)}$ ($\mathbf{x} \in \mathbf{X}_{d-1}$), and $\|A^{(d)}\|_{V_d \leftarrow V_d} = \|\hat{v}^{(d)}\|_d \|\varphi^{(d)}\|^*_d$ holds. This fact and the crossnorm property $\|\mathbf{Av}\| \leq \|\hat{v}^{(d)}\|_d \|\varphi^{(d)}\|^*_d \|\mathbf{v}\|$ lead us to

$$\|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} \|\varphi^{(d)}\|^*_d \underset{(4.44b)}{\geq} \frac{|(\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{Av})|}{\|\mathbf{Av}\|} \geq \frac{|(\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{Av})|}{\|\hat{v}^{(d)}\|_d \, \|\varphi^{(d)}\|^*_d \, \|\mathbf{v}\|}.$$

Since $(\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{Av}) = (\boldsymbol{\psi} \otimes (\varphi^{(d)} A^{(d)}))(\mathbf{v}) = \varphi^{(d)}(\hat{v}^{(d)}) \cdot (\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{v})$, the estimate can be continued by

$$\|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} \|\varphi^{(d)}\|^*_d \geq \frac{|\varphi^{(d)}(\hat{v}^{(d)})|}{\|\hat{v}^{(d)}\|_d \|\varphi^{(d)}\|^*_d} \frac{|(\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{v})|}{\|\mathbf{v}\|} \quad \text{for all } 0 \neq \hat{v}^{(d)} \in V_d.$$

Since $\sup_{\hat{v}^{(d)}\neq 0} \left|\varphi^{(d)}(\hat{v}^{(d)})\right| / \|\hat{v}^{(d)}\|_d = \|\varphi^{(d)}\|^*_d$, it follows that

$$\frac{\left|\left(\boldsymbol{\psi} \otimes \varphi^{(d)}\right)(\mathbf{v})\right|}{\|\mathbf{v}\|} \leq \|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} \|\varphi^{(d)}\|^*_d \qquad \text{for all } \mathbf{v} \in \mathbf{V},$$

so that $\|\boldsymbol{\psi} \otimes \varphi^{(d)}\|^* \leq \|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} \|\varphi^{(d)}\|^*_d$. Together with the opposite inequality from above, we have proved $\|\boldsymbol{\psi} \otimes \varphi^{(d)}\|^* = \|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} \|\varphi^{(d)}\|^*_d$. □

Corollary 4.91. For $\varphi^{(d)} \in V_d^*$ and $\boldsymbol{\psi} \in \mathbf{X}^*_{d-1}$, where $\mathbf{X}_{d-1} = {}_{\|\cdot\|_{\mathbf{X}_{d-1}}}\bigotimes_{j=1}^{d-1} V_j$ is equipped with the norm $\|\cdot\|_{\mathbf{X}_{d-1}}$ defined in Proposition 4.90, the following two inequalities hold:

$$\begin{aligned} \left\|\left(I \otimes \ldots \otimes I \otimes \varphi^{(d)}\right)(\mathbf{v})\right\|_{\mathbf{X}_{d-1}} &\leq \|\varphi^{(d)}\|^*_d \, \|\mathbf{v}\| \quad \text{and} \\ \|(\boldsymbol{\psi} \otimes I)(\mathbf{v})\|_d &\leq \|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} \|\mathbf{v}\|. \end{aligned} \tag{4.45}$$

Proof. Any $\boldsymbol{\psi} \in \mathbf{X}^*_{d-1}$ satisfies

$$\boldsymbol{\psi} \otimes \varphi^{(d)} = \boldsymbol{\psi}\left(I \otimes \ldots \otimes I \otimes \varphi^{(d)}\right).$$

For $\mathbf{v}_{[d]} := \left(I \otimes \ldots \otimes I \otimes \varphi^{(d)}\right)(\mathbf{v})$ there is a $\boldsymbol{\psi} \in \mathbf{X}^*_{d-1}$ with $\|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} = 1$ and $\left|\boldsymbol{\psi}\left(\mathbf{v}_{[d]}\right)\right| = \|\mathbf{v}_{[d]}\|_{\mathbf{X}_{d-1}}$ (cf. (4.10)). Hence,

$$\begin{aligned} &\left\|(I \otimes \ldots \otimes I \otimes \varphi^{(d)})(\mathbf{v})\right\|_{\mathbf{X}_{d-1}} = \left|\boldsymbol{\psi}\left((I \otimes \ldots \otimes I \otimes \varphi^{(d)})(\mathbf{v})\right)\right| \\ &= |(\boldsymbol{\psi} \otimes \varphi^{(d)})(\mathbf{v})| \leq \|\boldsymbol{\psi} \otimes \varphi^{(d)}\|^* \, \|\mathbf{v}\| = \|\boldsymbol{\psi}\|^*_{\mathbf{X}_{d-1}} \|\varphi^{(d)}\|^*_d \, \|\mathbf{v}\| = \|\varphi^{(d)}\|^*_d \, \|\mathbf{v}\| \end{aligned}$$

proves the first inequality in (4.45). The second one can be proved analogously. □

4.3.3 Projective Norm $\|\cdot\|_\wedge$

First we discuss the generalisation of $\|\cdot\|_\wedge$ to the d-fold tensor product. As for $d = 2$ (cf. §4.2.4), there is a norm $\|\mathbf{x}\|_{\wedge(V_1,\ldots V_d)}$ on ${}_a\bigotimes_{j=1}^d V_j$ induced by the norms $\|\cdot\|_j = \|\cdot\|_{V_j}$ for ($1 \le j \le d$), which can be defined as follows.

Remark 4.92. (a) For $\mathbf{x} \in {}_a\bigotimes_{j=1}^d V_j$ define $\|\cdot\|_{\wedge(V_1,\ldots V_d)}$ by

$$\|\mathbf{x}\|_{\wedge(V_1,\ldots V_d)} := \|\mathbf{x}\|_\wedge := \inf\left\{\sum_{i=1}^n \prod_{j=1}^d \|v_i^{(j)}\|_j : \mathbf{x} = \sum_{i=1}^n \bigotimes_{j=1}^d v_i^{(j)}\right\}.$$

(b) $\|\cdot\|_\wedge$ satisfies the crossnorm property (4.40).

(c) $\|\cdot\|_\wedge$ is the strongest norm for which the map $\bigotimes_{j=1}^d : V_1 \times \ldots \times V_d \to {}_a\bigotimes_{j=1}^d V_j$ is continuous.

The parts (b) and (c) are proved analogously to the case of $d = 2$ in Lemma 4.45 and Proposition 4.46

Proposition 4.93. $\|\cdot\|_\wedge$ *is a uniform and reasonable crossnorm on* ${}_a\bigotimes_{j=1}^d V_j$.

Proof. 1) The proof of the uniform crossnorm property in Proposition 4.78 can easily be extended from $d = 2$ to $d \ge 3$.

2) The result of Part 1) together with Lemma 4.79 implies that $\|\cdot\|_\wedge$ is a reasonable crossnorm. □

According to Proposition 4.90, the tensor space $\mathbf{V} := {}_\wedge\bigotimes_{j=1}^d V_j$ is generated recursively by $\mathbf{X}_k := \mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k$ for $k = 2, \ldots, d$ starting with $\mathbf{X}_2 := V_2$ and producing $\mathbf{V} = \mathbf{X}_d$. The concrete form of the norm $\|\cdot\|_{\mathbf{X}_k}$ constructed in (4.43) is described below.

Lemma 4.94. *The norm* $\|\cdot\|_{\mathbf{X}_k}$ *on* $\bigotimes_{j=1}^k V_j$ *(*$2 \le k \le d$*) which leads to the projective norm* $\|\cdot\|_{\mathbf{X}_d} = \|\cdot\|_\wedge = \|\cdot\|_{\wedge(V_1,\ldots,V_d)}$ *, is the projective norm of* $V_1, \ldots, V_k$ *:*

$$\|\cdot\|_{\mathbf{X}_k} = \|\cdot\|_{\wedge(V_1,\ldots,V_k)} . \tag{4.46a}$$

Considering $\mathbf{X}_k$ *as the tensor space* $\mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k$*, the construction of* $\|\cdot\|_{\mathbf{X}_k}$ *from* $\|\cdot\|_{\mathbf{X}_{k-1}}$ *and* $\|\cdot\|_k$ *is given by*

$$\|\cdot\|_{\wedge(V_1,\ldots,V_k)} = \|\cdot\|_{\wedge(X_{k-1},V_k)} . \tag{4.46b}$$

$\|\cdot\|_{\wedge(V_1,\ldots,V_k)}$ *is not only a uniform and reasonable crossnorm on* ${}_{\|\cdot\|_{\mathbf{X}_k}}\bigotimes_{j=1}^k V_j$*, but also*[16] *on* $\mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k$.

[16] Proposition 4.90 does not state the uniform crossnorm property on $\mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k$.

Proof. 1) For $\mathbf{v} := \mathbf{x} \otimes v^{(d)} \in {}_a\bigotimes_{j=1}^{d} V_j$ with $\mathbf{x} \in \mathbf{X}_{d-1}$ and $\|v^{(d)}\|_d = 1$ we have $\|\mathbf{x}\|_{\mathbf{X}_{d-1}} = \|\mathbf{v}\|_\wedge$ (cf. (4.43)). Let $\mathbf{v} = \sum_i \bigotimes_{j=1}^{d} v_i^{(j)}$ be any representation so that $\|\mathbf{v}\|_\wedge \le \sum_i \prod_{j=1}^{d} \|v_i^{(j)}\|_j$. A particular representation with $v_i^{(d)} = v^{(d)}$ for all i can be obtained as follows. Let $\psi \in V_d^*$ be the functional with $\|\psi\|_d^* = 1$ and $\psi(v^{(d)}) = 1$ (cf. Theorem 4.15) and set $\varPsi := v^{(d)}\psi$. Since $(I \otimes \varPsi)\,\mathbf{v} = \mathbf{v}$, another representation is

$$\mathbf{v} = (I \otimes \varPsi) \sum_i \bigotimes_{j=1}^{d} v_i^{(j)} = \sum_i \left(\bigotimes_{j=1}^{d-1} v_i^{(j)} \right) \otimes \varPsi(v_i^{(d)})$$
$$= \left(\sum_i \psi(v_i^{(d)}) \bigotimes_{j=1}^{d-1} v_i^{(j)} \right) \otimes v^{(d)}$$

leading to the same estimate $\|\mathbf{v}\|_\wedge \le \sum_i \prod_{j=1}^{d} \|v_i^{(j)}\|_j$ because of $\|v^{(d)}\|_d = 1$ and $|\psi(v_i^{(d)})| \le \|v_i^{(d)}\|_d$. Hence, the infimum $\|\mathbf{v}\|_\wedge$ can be obtained by all representations $\mathbf{v} = \sum_i \bigotimes_{j=1}^{d} v_i^{(j)}$ with $v_i^{(d)} = v^{(d)}$. Because of $\|v^{(d)}\|_d = 1$ we obtain

$$\|\mathbf{v}\|_\wedge = \inf \left\{ \sum_i \prod_{j=1}^{d-1} \|v_i^{(j)}\|_j : \mathbf{v} = \sum_i \bigotimes_{j=1}^{d} v_i^{(j)} \text{ and } v_i^{(d)} = v^{(d)} \right\}.$$

In the latter case, $\mathbf{v} = \left(\sum_i \bigotimes_{j=1}^{d-1} v_i^{(j)} \right) \otimes v^{(d)}$ implies that $\sum_i \bigotimes_{j=1}^{d-1} v_i^{(j)}$ is a representation of $\mathbf{x} \in \mathbf{X}_{d-1}$, since $\mathbf{v} = \mathbf{x} \otimes v^{(d)}$. Therefore, the right-hand side in the last formula is $\|\mathbf{x}\|_{\wedge(V_1,\ldots,V_{d-1})}$. This proves (4.46a) for $k = d-1$. Recursion yields (4.46a) for $k = d-2, \ldots, 2$.

2a) Let $\varepsilon > 0$. For $\mathbf{v} \in \mathbf{X}_{d-1} \otimes_a V_d$ there is a representation $\mathbf{v} = \sum_\nu \mathbf{x}_\nu \otimes v_\nu^{(d)}$ so that $\|\mathbf{v}\|_{\wedge(\mathbf{X}_{d-1},V_d)} \ge \sum_\nu \|\mathbf{x}_\nu\|_X \, \|v_\nu^{(d)}\|_d - \varepsilon$ ($\mathbf{x}_\nu \in \mathbf{X}_{d-1}$, $v_\nu^{(d)} \in V_d$). For each $\mathbf{x}_\nu \in \mathbf{X}_{d-1} = \bigotimes_{j=1}^{d-1} V_j$ ($\nu \ge 1$) choose representations $\mathbf{x}_\nu = \sum_\mu \bigotimes_{j=1}^{d-1} v_{\nu,\mu}^{(j)}$ such that $\|\mathbf{x}_\nu\|_{\mathbf{X}_{d-1}} \|v_\nu^{(d)}\|_d \ge \left(\sum_\mu \prod_{j=1}^{d-1} \|v_{\nu,\mu}^{(j)}\|_j \right) \|v_\nu^{(d)}\|_d - 2^{-\nu}\varepsilon$. Altogether, it follows that

$$\|\mathbf{v}\|_{\wedge(\mathbf{X}_{d-1},V_d)} \ge \sum_\nu \left(\sum_\mu \prod_{j=1}^{d-1} \|v_{\nu,\mu}^{(j)}\|_j \right) \|v_\nu^{(i)}\|_d - 2\varepsilon.$$

A possible representation of $\mathbf{v} \in \bigotimes_{j=1}^{d} V_j$ is $\mathbf{v} = \sum_\nu \sum_\mu \bigotimes_{j=1}^{d} v_{\nu,\mu}^{(j)}$ with $v_{\nu,\mu}^{(d)} := v_\nu^{(d)}$ (independent of μ); hence, $\sum_{\nu,\mu} \left(\prod_{j=1}^{d-1} \|v_{\nu,\mu}^{(j)}\|_j \right) \|v_\nu^{(d)}\|_d \ge \|\mathbf{v}\|_{\wedge(V_1,\ldots,V_d)}$. As $\varepsilon > 0$ is arbitrary, $\|\mathbf{v}\|_{\wedge(\mathbf{X}_{d-1},V_d)} \ge \|\mathbf{v}\|_{\wedge(V_1,\ldots,V_d)}$ is proved.

2b) For the reverse inequality choose a representation $\mathbf{v} = \sum_\nu \bigotimes_{j=1}^{d} v_\nu^{(j)}$ with

$$\|z\|_{\wedge(V_1,\ldots,V_d)} \ge \sum_\nu \prod_{j=1}^{d} \|v_\nu^{(j)}\|_j - \varepsilon.$$

Define $\mathbf{x}_\nu := \bigotimes_{j=1}^{d-1} v_\nu^{(j)}$. Then $\mathbf{v} = \sum_\nu \mathbf{x}_\nu \otimes v_\nu^{(d)}$ and $\prod_{j=1}^{d-1} \|v_\nu^{(j)}\|_j = \|\mathbf{x}_\nu\|_{\mathbf{X}_{d-1}}$ (crossnorm property) are valid, and one concludes that

$$\|\mathbf{v}\|_{\wedge(V_1,\ldots,V_d)} \geq \sum_\nu \|\mathbf{x}_\nu\|_{\mathbf{X}_{d-1}} \|v_\nu^{(d)}\|_d - \varepsilon \geq \|\mathbf{v}\|_{\wedge(\mathbf{X}_{d-1},V_d)} - \varepsilon$$

for all $\varepsilon > 0$, which proves $\|\mathbf{v}\|_{\wedge(V_1,\ldots,V_d)} \geq \|\mathbf{v}\|_{\wedge(\mathbf{X}_{d-1},V_d)}$.

Again, induction for $k = d-2, \ldots, 2$ shows (4.46b) for all k. We remark that equality of the norms implies that the completion yields identical Banach spaces ${}_\wedge\bigotimes_{j=1}^{k} V_j = \mathbf{X}_{k-1} \otimes_\wedge V_k$.

3) Since the projective norm is a uniform crossnorm, Proposition 4.78 states that $\|\cdot\|_{\mathbf{X}_k}$ is a uniform crossnorm on $\mathbf{X}_{k-1} \otimes_{\wedge(\mathbf{X}_{k-1},V_k)} V_k$. □

Equations (4.40) and (4.41) together with Remark 4.92c show that $\|\cdot\|_{\wedge(V_1,\ldots,V_d)}$ is the strongest reasonable crossnorm on ${}_a\bigotimes_{j=1}^{d} V_j$.

4.3.4 Injective Norm $\|\cdot\|_\vee$

The analogue of the definition of $\|\cdot\|_\vee$ in §4.2.7 for d factors yields the following formulation. Let $\varphi_1 \otimes \varphi_2 \otimes \ldots \otimes \varphi_d$ be an elementary tensor of the tensor space ${}_a\bigotimes_{j=1}^{d} V_j^*$ involving the dual spaces. The proof of the next remark uses the same arguments as used in Lemma 4.54 and Proposition 4.55.

Remark 4.95. (a) For $\mathbf{v} \in {}_a\bigotimes_{j=1}^{d} V_j$ define $\|\cdot\|_{\vee(V_1,\ldots V_d)}$ by[17]

$$\|\mathbf{v}\|_{\vee(V_1,\ldots,V_d)} := \|\mathbf{v}\|_\vee := \sup_{\substack{0\neq\varphi_j\in V_j^* \\ 1\leq j\leq d}} \frac{|(\varphi_1 \otimes \varphi_2 \otimes \ldots \otimes \varphi_d)(\mathbf{v})|}{\prod_{j=1}^{d} \|\varphi_j\|_j^*}. \tag{4.47}$$

(b) For elementary tensors the crossnorm property (4.40) holds:

$$\left\| \bigotimes_{j=1}^{d} v^{(j)} \right\|_{\vee(V_1,\ldots,V_d)} = \prod_{j=1}^{d} \|v^{(j)}\|_j \quad \text{for all } v^{(j)} \in V_j \ (1 \leq j \leq d).$$

(c) $\|\cdot\|_{\vee(V_1,\ldots V_d)}$ is the weakest norm with $\bigotimes : \bigtimes_{j=1}^{d} V_j^* \to {}_a\bigotimes_{j=1}^{d} V_j^*$ being continuous.

Lemma 4.94 can be repeated with $\|\cdot\|_\wedge$ replaced by $\|\cdot\|_\vee$.

Lemma 4.96. *The norms $\|\cdot\|_{\mathbf{X}_k}$ on $\bigotimes_{j=1}^{k} V_j$ for $2 \leq k \leq d$ leading to the injective norm $\|\cdot\|_{\mathbf{X}_d} = \|\cdot\|_\vee = \|\cdot\|_{\vee(V_1,\ldots,V_d)}$, are the injective norms of $V_1, \ldots, V_k$:*

$$\|\cdot\|_{\mathbf{X}_k} = \|\cdot\|_{\vee(V_1,\ldots,V_k)} .$$

[17] In [201, Def. 1.2], the expression $(\varphi_1 \otimes \varphi_2 \otimes \ldots \otimes \varphi_d)(\mathbf{v}) / \prod_{j=1}^{d} \|\varphi_j\|_j^*$ is introduced as the *generalised Rayleigh quotient*.

Considering $\mathbf{X}_k$ as the tensor space $\mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k$, the construction of $\|\cdot\|_{\mathbf{X}_k}$ from $\|\cdot\|_{\mathbf{X}_{k-1}}$ and $\|\cdot\|_k$ is given by

$$\|\cdot\|_{\vee(V_1,\ldots,V_k)} = \|\cdot\|_{\vee(\mathbf{X}_{k-1},V_k)} \,.$$

$\|\cdot\|_{\vee(V_1,\ldots,V_k)}$ is not only a uniform and reasonable crossnorm on ${}_{\|\cdot\|_{\mathbf{X}_k}}\bigotimes_{j=1}^k V_j$, but also on $\mathbf{X}_{k-1} \otimes_{\|\cdot\|_{\mathbf{X}_k}} V_k$.

Proof. 1) By (4.43), $\|\mathbf{x}\|_{\mathbf{X}_{d-1}} = \|\mathbf{x} \otimes v^{(d)}\|_\vee$ holds for any normalised $v^{(d)} \in V_d$. Then the right-hand side in (4.47) can be simplified by

$$\sup_{0\neq\varphi_d\in V_d^*} \frac{|(\varphi_1 \otimes \ldots \otimes \varphi_d)(\mathbf{v})|}{\|\varphi_d\|_d^*} = \left|\left(\bigotimes_{j=1}^d \varphi_j\right)(\mathbf{x})\right| \|v^{(d)}\|_d = \left|\left(\bigotimes_{j=1}^d \varphi_j\right)(\mathbf{x})\right|$$

to

$$\|\mathbf{x} \otimes v^{(d)}\|_\vee = \sup_{\substack{0\neq\varphi_j\in V_j^* \\ 1\leq j\leq d-1}} \left|\left(\bigotimes_{j=1}^d \varphi_j\right)(\mathbf{x})\right| \Big/ \prod_{j=1}^d \|\varphi_j\|_j^* = \|\mathbf{x}\|_{\vee(V_1,\ldots,V_{d-1})}$$

proving the first assertion for $k = d-1$ (other k by recursion).

2) For the proof of the second part, let $\mathbf{v} = \sum_i \mathbf{x}_i \otimes v_i^{(k)} \in \mathbf{X}_{k-1} \otimes_a V_k$. By definition,

$$\|\mathbf{v}\|_{\vee(\mathbf{X}_{k-1},V_k)} = \sup_{\substack{0\neq\mathbf{\Phi}\in\mathbf{X}_{k-1}^* \\ 0\neq\varphi_k\in V_k^*}} \frac{|(\mathbf{\Phi} \otimes \varphi_k)(\mathbf{v})|}{\|\mathbf{\Phi}\|_{\mathbf{X}_{k-1}}^* \, \|\varphi_k\|_k^*}$$

holds. We perform the supremum over $0 \neq \boldsymbol{\Phi} \in \mathbf{X}_{k-1}^*,\ 0 \neq \varphi_k \in V_k^*$ sequentially by

$$\sup_{0\neq v_k^*\in V_k^*} \left\{ \frac{1}{\|\varphi_k\|_k^*} \sup_{0\neq\boldsymbol{\Phi}\in\mathbf{X}_{k-1}^*} \frac{|(\boldsymbol{\Phi} \otimes \varphi_k)(\mathbf{v})|}{\|\boldsymbol{\Phi}\|_{\mathbf{X}_{k-1}}^*} \right\}.$$

The nominator $|(\boldsymbol{\Phi} \otimes \varphi_k)(\mathbf{v})|$ becomes $\sum_i \boldsymbol{\Phi}(\mathbf{x}_i)\varphi_k(v_i^{(k)})$. We introduce the abbreviation $\lambda_i := \varphi_k(v_i^{(k)}) \in \mathbb{K}$ and obtain $\sum_i \mathbf{\Phi}(\mathbf{x}_i)\varphi_k(v_i^{(k)}) = \sum_i \boldsymbol{\Phi}(\mathbf{x}_i)\lambda_i = \boldsymbol{\Phi}\left(\sum_i \lambda_i \mathbf{x}_i\right)$. The inner supremum yields

$$\sup_{0\neq\boldsymbol{\Phi}\in\mathbf{X}_{k-1}^*} \frac{|(\boldsymbol{\Phi} \otimes \varphi_k)(\mathbf{v})|}{\|\boldsymbol{\Phi}\|_{\mathbf{X}_{k-1}}^*} = \sup_{0\neq\boldsymbol{\Phi}\in\mathbf{X}_{k-1}^*} \frac{|\boldsymbol{\Phi}\left(\sum_i \lambda_i \mathbf{x}_i\right)|}{\|\boldsymbol{\Phi}\|_{\mathbf{X}_{k-1}}^*} \underset{(4.10)}{=} \left\|\sum\nolimits_i \lambda_i \mathbf{x}_i\right\|_{\mathbf{X}_{k-1}}.$$

By definition of $\|\cdot\|_{\mathbf{X}_{k-1}}$ this is

$$\left\|\sum\nolimits_i \lambda_i \mathbf{x}_i\right\|_{\mathbf{X}_{k-1}} = \sup_{\substack{0\neq\varphi_j\in V_j^* \\ 1\leq j\leq k-1}} \frac{|(\varphi_1 \otimes \ldots \otimes \varphi_{k-1})\left(\sum_i \lambda_i \mathbf{x}_i\right)|}{\prod_{j=1}^{k-1} \|\varphi_j\|_j^*}.$$

Now, we re-insert $\lambda_i = \varphi_k(v_i^{(k)})$ and perform the outer supremum:

$$\sup_{0\neq\varphi_k\in V_k^*}\frac{\displaystyle\sup_{0\neq\mathbf{x}^*\in\mathbf{X}_{k-1}^*}\frac{\left|\left(\boldsymbol{\Phi}\otimes\varphi_k\right)(\mathbf{v})\right|}{\|\boldsymbol{\Phi}\|_{\mathbf{X}_{k-1}}^*}}{\|\varphi_k\|_k^*}=\sup_{\substack{0\neq\varphi_j\in V_j^*\\1\leq j\leq k}}\frac{\left|\left(\bigotimes_{j=1}^{k-1}\varphi_j\right)\left(\sum_i\varphi_k(v_i^{(k)})\mathbf{x}_i\right)\right|}{\prod_{j=1}^k\|\varphi_j\|_j^*}$$

$$=\sup_{\substack{0\neq\varphi_j\in V_j^*\\1\leq j\leq k}}\frac{\left|\left(\varphi_1\otimes\ldots\otimes\varphi_k\right)\left(\sum_i\mathbf{x}_i\otimes v_i^{(k)}\right)\right|}{\prod_{j=1}^k\|\varphi_j\|_j^*}=\sup_{\substack{0\neq\varphi_j\in V_j^*\\1\leq j\leq k}}\frac{\left|\left(\varphi_1\otimes\ldots\otimes\varphi_k\right)(\mathbf{v})\right|}{\prod_{j=1}^k\|\varphi_j\|_j^*}$$

$$=\|\mathbf{x}\|_{\vee(V_1,\ldots,V_k)}\,,$$

which finishes the proof of $\|\cdot\|_{\vee(V_1,\ldots,V_k)}=\|\cdot\|_{\vee(\mathbf{X}_{k-1},V_k)}\,.$

3) By Proposition 4.78, $\|\cdot\|_{\vee(\mathbf{X}_{k-1},V_k)}$ is a uniform crossnorm on $\mathbf{X}_{k-1}\otimes V_k$. □

Lemma 4.97. *For fixed $j\in\{1,\ldots,d\}$, the mapping*

$$\boldsymbol{\Phi}=\bigotimes_{k\in\{1,\ldots,d\}\backslash\{j\}}\varphi_k\ \in\ {}_a\bigotimes_{k\in\{1,\ldots,d\}\backslash\{j\}}V_k^*$$

is also understood as a mapping from $\left({}_a\bigotimes_{k=1}^d V_k\ ,\|\cdot\|_\vee\right)$ onto V_j :

$$\boldsymbol{\Phi}\left(\bigotimes_{k=1}^d v^{(k)}\right):=\left(\prod_{k\in\{1,\ldots,d\}\backslash\{j\}}\varphi_k(v^{(k)})\right)\cdot v^{(j)}.\tag{4.48}$$

The more precise notation for φ in the sense of (4.48) is

$$\boldsymbol{\Phi}=\varphi_1\otimes\ldots\otimes\varphi_{j-1}\otimes id\otimes\varphi_{j+1}\otimes\ldots\otimes\varphi_d.$$

Then $\boldsymbol{\Phi}$ is continuous, i.e., $\boldsymbol{\Phi}\in\mathcal{L}\left({}_\vee\bigotimes_{k=1}^d V_k,V_j\right)$. Its norm is

$$\|\boldsymbol{\Phi}\|_{V_j\leftarrow\,_\vee\bigotimes_{k=1}^d V_k}=\prod_{k\in\{1,\ldots,d\}\backslash\{j\}}\|\varphi_k\|_k^*\,.$$

Proof. Let $\varphi_j\in V_j^*$ and note that the composition $\varphi_j\circ\boldsymbol{\Phi}$ equals $\bigotimes_{k=1}^d\varphi_k$. Hence,

$$\|\boldsymbol{\Phi}(\mathbf{v})\|_j\underset{(4.10)}{=}\max_{\varphi_j\in V_j^*,\,\|\varphi_j\|_j^*=1}|\varphi_j(\boldsymbol{\Phi}(\mathbf{v}))|=\max_{\|\varphi_j\|_j^*=1}|(\varphi_j\circ\boldsymbol{\Phi})(\mathbf{v})|$$

$$=\max_{\|\varphi_j\|_j^*=1}\left|\left(\bigotimes_{k=1}^d\varphi_k\right)(\mathbf{v})\right|\leq\left(\prod\nolimits_{k\in\{1,\ldots,d\}\backslash\{j\}}\|\varphi_k\|_k^*\right)\|\mathbf{v}\|_\vee\,.$$

Equation (4.8) shows $\sup\{\|\boldsymbol{\Phi}(\mathbf{v})\|_j:\|\mathbf{v}\|_\vee=1\}=\prod_{k\in\{1,\ldots,d\}\backslash\{j\}}\|\varphi_k\|_k^*\,.$ □

Corollary 4.98. For any norm $\|\cdot\|$ on ${}_a\bigotimes_{k=1}^d V_k$ not weaker than $\|\mathbf{v}\|_\vee$ (in particular, for all reasonable crossnorms) $\boldsymbol{\Phi}\in\mathcal{L}\left({}_{\|\cdot\|}\bigotimes_{k=1}^d V_k,V_j\right)$ is valid.

4.3.5 Examples

Example 4.47 can be generalised to tensors of order d.

Example 4.99. Let $V_j := \ell^1(I_j)$ with finite or countable index set I_j for $1 \leq j \leq d$. The induced norm $\|\cdot\|_{\wedge(V_1,\ldots,V_d)}$ of ${}_a\bigotimes_{j=1}^{j} V_j$ coincides with $\|\cdot\|_{\ell^1(I_1\times I_2\times\ldots\times I_d)}$.

Remark 4.65 leads to the following generalisation.

Example 4.100. Let $V_j = (C(I_j), \|\cdot\|_{C(I_j)})$ with certain domains I_j (e.g., intervals). Then

$$\|\cdot\|_{\vee(C(I_1),\ldots,C(I_d))} = \|\cdot\|_{C(I_1\times I_2\times\ldots\times I_d)}.$$

Proof. We perform the product sequentially. Remark 4.65 shows $\|\cdot\|_{\vee(C(I_1),C(I_2))} = \|\cdot\|_{C(I_1\times I_2)}$. Iteration yields $\|\cdot\|_{\vee(C(I_1),C(I_2),C(I_3))} = \|\cdot\|_{\vee(C(I_1\times I_2),C(I_3))} = \|\cdot\|_{C(I_1\times I_2\times I_3)}$. Induction completes the proof. □

Since $\|c\|_{\mathrm{SVD},2} = \|\cdot\|_{\ell^2(I_1\times I_2)}$ is not equivalent to $\|\cdot\|_{\mathrm{SVD},1}$, the explicit interpretation of $\|\cdot\|_{\wedge(V_1,\ldots,V_d)}$ for $(V_j, \|\cdot\|_{\ell^2(I_j)})$ and $d \geq 3$ is not obvious.

4.3.6 Intersections of Banach Tensor Spaces

If two Banach spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ have a non-empty intersection $Z := X \cap Y$, the *intersection norm* $\|\cdot\|_Z$ is given by $\|z\|_Z := \max\{\|z\|_X, \|z\|_Y\}$ or equivalent ones. Below, we shall make use of this construction.

At the end of §4.2.10 we have studied the example $C^1(I \times J)$. This space can be obtained as the closure of $C^1(I) \otimes_a C^1(J)$ with respect to the norm $\|\cdot\|_{C^1(I\times J)}$, however, this norm is not a reasonable crossnorm, it satisfies (4.32a), but not (4.32b). Instead, the mixed norm $\|\cdot\|_{1,\mathrm{mix}}$ from (4.35) is a reasonable crossnorm, but the resulting space $C^1_{\mathrm{mix}}(I \times J)$ is a proper subspace of $C^1(I \times J)$.

There is another way to obtain $C^1(I\times J)$. First we consider the anisotropic spaces

$$C^{(1,0)}(I\times J) := \{f : f, \tfrac{\partial}{\partial x} f \in C(I\times J)\}, \quad \|f\|_{(1,0)} := \max\{\|f\|_\infty, \|f_x\|_\infty\},$$
$$C^{(0,1)}(I\times J) := \{f : f, \tfrac{\partial}{\partial y} f \in C(I\times J)\}, \quad \|f\|_{(0,1)} := \max\{\|f\|_\infty, \|f_y\|_\infty\},$$

with $\|f\|_\infty := \sup_{(x,y)\in I\times J} |f(x,y)|$. Then we obtain $C^1(I \times J)$ and its norm by

$$C^1(I\times J) = C^{(1,0)}(I\times J) \cap C^{(0,1)}(I\times J), \quad \|\cdot\|_{C^1(I\times J)} = \max\{\|f\|_{(1,0)}, \|f\|_{(0,1)}\}.$$

The proof of Remark 4.65 can be extended to show that $\|\cdot\|_{(1,0)}$ $[\|\cdot\|_{(0,1)}]$ is a reasonable crossnorm of $C^{(1,0)}(I \times J)$ $[C^{(0,1)}(I \times J)]$.

We give another important example. Here, $N \in \mathbb{N}$ is a fixed degree.

Example 4.101. For $I_j \subset \mathbb{R}$ $(1 \le j \le d)$ and $1 \le p < \infty$, the Sobolev space $H^{N,p}(I_j)$ consists of all functions f from $L^p(I_j)$ with bounded norm [18]

$$\|f\|_{N,p;I_j} := \left(\sum_{n=0}^{N} \int_{I_j} \left| \frac{\mathrm{d}^n}{\mathrm{d}x^n} f \right|^p \mathrm{d}x \right)^{1/p}, \tag{4.49a}$$

whereas $H^{N,p}(\mathbf{I})$ for $\mathbf{I} = I_1 \times \ldots \times I_d \subset \mathbb{R}^d$ is endowed with the norm

$$\|f\|_{N,p} := \left(\sum_{0 \le |\mathbf{n}| \le N} \int_{\mathbf{I}} |\partial^{\mathbf{n}} f|^p \, \mathrm{d}x \right)^{1/p} \tag{4.49b}$$

with $\mathbf{n} \in \mathbb{N}_0^d$ being a multi-index of length $|\mathbf{n}| := \sum_{j=1}^{d} n_j$, and $\partial^{\mathbf{n}}$ as in (4.5).

Again, the norm $\|\cdot\|_{N,p}$ satisfies (4.32a), but not (4.32b), in particular, it is not a reasonable crossnorm. Instead, for each $\mathbf{n} \in \mathbb{N}_0^d$ with $|\mathbf{n}| \le N$ we define the space $H^{\mathbf{n},p}(\mathbf{I}) := \{f \in L^p(\mathbf{I}) : \partial^{\mathbf{n}} f \in L^p(\mathbf{I})\}$ with the reasonable crossnorm

$$\|f\|_{\mathbf{n},p} := (\|f\|_{0,p}^p + \|\partial^{\mathbf{n}} f\|_{0,p}^p)^{1/p}.$$

Then, the Sobolev space $H^{N,p}(\mathbf{I})$ is equal to the intersection $\bigcap_{0 \le |\mathbf{n}| \le N} H^{\mathbf{n},p}(\mathbf{I})$, and its norm (4.49b) is equivalent to $\max_{0 \le |\mathbf{n}| \le N} \|\cdot\|_{\mathbf{n},p}$.

Note that $H^{\mathbf{n},p}$ for $\mathbf{n} = (1,0)$ is considered in Example 4.42. If $\mathbf{n} \in \mathbb{N}_0^d$ is a multiple of a unit vector, i.e., $\mathbf{n}_i = 0$ except for one i, the proof of Remark 4.74 can be used to show that $\|\cdot\|_{\mathbf{n},p}$ is a reasonable crossnorm for $1 \le p < \infty$.

The Sobolev spaces $H^{m,p}(I_j)$ for $m = 0, 1, \ldots, N$ are an example for a *scale* of Banach spaces. In the following, we fix integers N_j and denote the j-th scale by

$$V_j = V_j^{(0)} \supset V_j^{(1)} \supset \ldots \supset V_j^{(N_j)} \text{ with dense embeddings,} \tag{4.50}$$

i.e., $V_j^{(n)}$ is a dense subspace of $(V_j^{(n-1)}, \|\cdot\|_{j,n-1})$ for $1 \le n \le N_j$. This fact implies that the corresponding norms satisfy $\|\cdot\|_{j,n} \gtrsim \|\cdot\|_{j,m}$ for $N_j \ge n \ge m \ge 0$ on $V_j^{(N_j)}$.

Lemma 4.102. *By (4.50), all $V_j^{(n)}$ $(1 \le n \le N_j)$ are dense in $(V_j^{(0)}, \|\cdot\|_{j,0})$.*

Let numbers $N_j \in \mathbb{N}_0$ be given and define $\mathcal{N} \subset \mathbb{N}_0^d$ as a subset of d-tuples satisfying

$$\mathbf{n} \in \mathcal{N} \Rightarrow 0 \le n_j \le N_j, \tag{4.51a}$$

$$\mathbf{0} := (0, \ldots, 0) \in \mathcal{N}, \tag{4.51b}$$

$$\mathbf{N}_j := (\underbrace{0, \ldots, 0}_{j-1}, N_j, \underbrace{0, \ldots, 0}_{d-j}) \in \mathcal{N}. \tag{4.51c}$$

The standard choice of $\mathcal{N}$ is

$$\mathcal{N} := \left\{ \mathbf{n} \in \mathbb{N}_0^d \text{ with } |\mathbf{n}| \le N \right\}, \quad \text{where } N_j = N \text{ for all } 1 \le j \le d. \tag{4.51d}$$

[18] It suffices to have the terms for $n = 0$ and $n = N$ in (4.49a). The derivatives are to be understood as weak derivatives (cf. [82, §6.2.1]).

For each $\mathbf{n} \in \mathcal{N}$ we define the tensor space

$$\mathbf{V}^{(\mathbf{n})} := {}_a\bigotimes_{j=1}^{d} V_j^{(n_j)} . \tag{4.52a}$$

Then we can choose a reasonable crossnorm $\|\cdot\|_{\mathbf{n}}$ on $\mathbf{V}^{(\mathbf{n})}$ or an equivalent one. The *intersection Banach tensor space* is defined by

$$\mathbf{V} := \overline{\bigcap_{\mathbf{n}\in\mathcal{N}} \mathbf{V}^{(\mathbf{n})}} = \bigcap_{\mathbf{n}\in\mathcal{N}} \overline{\mathbf{V}^{(\mathbf{n})}} \text{ with intersection norm } \|\mathbf{v}\| := \max_{\mathbf{n}\in\mathcal{N}} \|\mathbf{v}\|_{\mathbf{n}} \tag{4.52b}$$

or an equivalent norm.

Remark 4.103. Assume $V_j^{(0)} \supsetneqq V_j^{(N_j)}$ (this excludes the finite dimensional case) and let $\mathbf{V}$ be defined by (4.52b).

(a) Only if $(N_1, \ldots, N_d) \in \mathcal{N}$, $\mathbf{V} = \mathbf{V}_{\mathrm{mix}} := \overline{\mathbf{V}^{(N_1,N_2,\ldots,N_d)}}$ holds (cf. (4.35)).

(b) Otherwise, $\mathbf{V}_{\mathrm{mix}} \subsetneqq \mathbf{V} \subsetneqq \mathbf{V}^{(\mathbf{0})}$ and $\mathbf{V}_{\mathrm{mix}}$ is dense in $\mathbf{V}$.

(c) In Case (a), a reasonable crossnorm $\|\cdot\|_{\mathrm{mix}}$ may exist, whereas in Case (b) condition (4.32b) required for a reasonable crossnorm cannot be satisfied.

Proof. For Part (c) note that $\varphi \in \bigotimes_{j=1}^d (V_j^{(N_j)})^*$ are continuous functionals on $\mathbf{V}_{\mathrm{mix}}$, but not necessarily on $\mathbf{V} \subsetneqq \mathbf{V}_{\mathrm{mix}}$ endowed with a strictly weaker norm. □

Proposition 4.104. *Under the conditions (4.51a-c), the Banach tensor space* $\mathbf{V}$ *from (4.52b) satisfies the inclusion*

$$\left({}_a\bigotimes_{j=1}^{d} V_j \right) \cap \mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j^{(N_j)} \subset \mathbf{V}_{\mathrm{mix}} := {}_{\|\cdot\|_{(N_1,\ldots,N_d)}}\bigotimes_{j=1}^{d} V_j^{(N_j)} ,$$

i.e., an algebraic tensor in $\mathbf{V}$ *does not differ from an algebraic tensor in* $\mathbf{V}_{\mathrm{mix}}$*. Each* $\mathbf{v} \in {}_a\bigotimes_{j=1}^d V_j \cap \mathbf{V}$ *has a representation* $\mathbf{v} = \sum_{i=1}^r \bigotimes_{j=1}^d v_i^{(j)}$ *with* $v_i^{(j)} \in V_j^{(N_j)}$.

Proof. By definition (4.52a),

$$\left({}_a\bigotimes_{j=1}^{d} V_j \right) \cap \mathbf{V} = \bigcap_{\mathbf{n}\in\mathcal{N}} \left[\left({}_a\bigotimes_{j=1}^{d} V_j \right) \cap \overline{\mathbf{V}^{(\mathbf{n})}} \right]$$

holds. Since $\mathbf{v} \in \left({}_a\bigotimes_{j=1}^d V_j \right) \cap \overline{\mathbf{V}^{(\mathbf{n})}}$ is an algebraic tensor, it belongs to the space $\left({}_a\bigotimes_{j=1}^d V_j \right) \cap \mathbf{V}^{(\mathbf{n})} = \mathbf{V}^{(\mathbf{n})}$. Lemma 6.11 will show that

$$\mathbf{v} \in \bigcap_{\mathbf{n}\in\mathcal{N}} \mathbf{V}^{(\mathbf{n})} = {}_a\bigotimes_{j=1}^{d} \left[\bigcap_{\mathbf{n}\in\mathcal{N}} V_j^{(n_j)} \right].$$

By condition (4.51c), $\mathbf{v} \in {}_a\bigotimes_{j=1}^d \left(\bigcap_{\mathbf{n}\in\mathcal{N}} V_j^{(n_j)} \right) = {}_a\bigotimes_{j=1}^d V_j^{(N_j)}$ can be concluded from the fact that one of the n_j equals N_j. □

Application to $V_j^{(0)} = C^0(I_j)$ and $V_j^{(1)} = C^1(I_j)$ yields that all functions from the algebraic tensor space $\left({}_a\bigotimes_{j=1}^d C^0(I_j) \right) \cap C^1(\mathbf{I})$ are already in $\mathbf{V}_{\mathrm{mix}} = C^1_{\mathrm{mix}}(\mathbf{I})$ (cf. (4.35)), which is a proper subspace of $C^1(\mathbf{I})$.

The dual space $\mathbf{V}^*$ is the sum (span) of the duals of $\mathbf{V}^{(\mathbf{n})}$: $\mathbf{V}^* = \sum_{\mathbf{n}\in\mathcal{N}} \left(\mathbf{V}^{(\mathbf{n})}\right)^*$.

4.3.7 Tensor Space of Operators

Let $\mathbf{V} = {}_a\bigotimes_{j=1}^d V_j$ and $\mathbf{W} = {}_a\bigotimes_{j=1}^d W_j$ be two Banach tensor spaces with the respective norms $\|\cdot\|_{\mathbf{V}}$ and $\|\cdot\|_{\mathbf{W}}$, while $\|\cdot\|_{V_j}$ and $\|\cdot\|_{W_j}$ are the norms of V_j and W_j. The space $\mathcal{L}(V_j, W_j)$ is endowed with the operator norm $\|\cdot\|_{W_j\leftarrow V_j}$. Their algebraic tensor space is

$$\mathbf{L} := {}_a\bigotimes_{j=1}^d \mathcal{L}(V_j, W_j) \,.$$

The obvious action of an elementary tensor $\mathbf{A} = \bigotimes_{j=1}^d A^{(j)} \in \mathbf{L}$ on $\mathbf{v} = \bigotimes_{j=1}^d v^{(j)} \in \mathbf{V}$ yields the following tensor from $\mathbf{W}$:

$$\mathbf{A}\mathbf{v} = \left(\bigotimes_{j=1}^d A^{(j)}\right)\left(\bigotimes_{j=1}^d v^{(j)}\right) = \bigotimes_{j=1}^d A^{(j)}v^{(j)} \in \mathbf{W}.$$

If $\|\cdot\|_{\mathbf{V}}$ and $\|\cdot\|_{\mathbf{W}}$ are crossnorms, we estimate $\|\mathbf{A}\mathbf{v}\|_{\mathbf{W}}$ by

$$\prod_{j=1}^d \|A^{(j)}v^{(j)}\|_{W_j} \le \prod_{j=1}^d \left[\|A^{(j)}\|_{W_j\leftarrow V_j} \|v^{(j)}\|_{V_j} \right] = \left(\prod_{j=1}^d \|A^{(j)}\|_{W_j\leftarrow V_j} \right) \|\mathbf{v}\|_{\mathbf{V}} \,.$$

Hence, $\|\mathbf{A}\mathbf{v}\|_{\mathbf{W}} \le \|\mathbf{A}\|\|\mathbf{v}\|_{\mathbf{V}}$ holds for all elementary tensors. However, we cannot expect that all crossnorms $\|\cdot\|_{\mathbf{V}}$ and $\|\cdot\|_{\mathbf{W}}$ satisfy the estimate $\|\mathbf{A}\mathbf{v}\|_{\mathbf{W}} \le \|\mathbf{A}\|\|\mathbf{v}\|_{\mathbf{V}}$ for *general* tensors $\mathbf{v} \in \mathbf{V}$. In the special case of $\mathbf{V} = \mathbf{W}$, we have called crossnorms uniform if they satisfy this estimate (cf. §4.2.12). We show that the induced norms are uniform crossnorms.

Proposition 4.105. *(a) If* $\|\cdot\|_{\mathbf{V}} = \|\cdot\|_{\wedge(V_1,\ldots,V_d)}$ *and* $\|\cdot\|_{\mathbf{W}} = \|\cdot\|_{\wedge(W_1,\ldots,W_d)}$, *then* $\mathbf{A} = \bigotimes_{j=1}^d A^{(j)} \in \mathbf{L}$ *has the operator norm*

$$\|\mathbf{A}\|_{\mathbf{W}\leftarrow\mathbf{V}} = \prod_{j=1}^d \|A^{(j)}\|_{W_j\leftarrow V_j}. \tag{4.53}$$

(b) If $\|\cdot\|_{\mathbf{V}} = \|\cdot\|_{\vee(V_1,\ldots,V_d)}$ *and* $\|\cdot\|_{\mathbf{W}} = \|\cdot\|_{\vee(W_1,\ldots,W_d)}$, *(4.53) holds again with respect to the corresponding operator norm.*

Proof. The same arguments as in the proof of Proposition 4.78 can be applied. □

Since $\|\mathbf{A}\|_{\mathbf{W}\leftarrow\mathbf{V}}$ is finite for elementary tensors $\mathbf{A}=\bigotimes_{j=1}^{d}A^{(j)}$, boundedness holds for all $\mathbf{A}\in\mathbf{L}:={}_a\bigotimes_{j=1}^{d}\mathcal{L}(V_j,W_j)$. The completion of $(\mathbf{L},\|\cdot\|_{\mathbf{W}\leftarrow\mathbf{V}})$ yields the tensor space

$$_{\|\cdot\|_{\mathbf{W}\leftarrow\mathbf{V}}}\bigotimes_{j=1}^{d}\mathcal{L}(V_j,W_j)\quad\subset\quad\mathcal{L}(\mathbf{V},\mathbf{W}).$$

4.4 Hilbert Spaces

4.4.1 Scalar Product

Again, we restrict the field to either $\mathbb{K}=\mathbb{R}$ or $\mathbb{K}=\mathbb{C}$. A normed vector space $(V,\|\cdot\|)$ is a pre-Hilbert space if the norm is defined by

$$\|v\|=\sqrt{\langle v,v\rangle}<\infty\qquad\text{for all } v\in V,\tag{4.54}$$

where $\langle\cdot,\cdot\rangle:V\times V\to\mathbb{K}$ is a *scalar product* on V. In the case of $\mathbb{K}=\mathbb{R}$, a scalar product is a *bilinear form*, which, in addition, must be symmetric and positive:

$$\langle v,w\rangle=\langle w,v\rangle\qquad\text{for } v,w\in V,\tag{4.55a}$$

$$\langle v,v\rangle>0\qquad\text{for } v\neq 0.\tag{4.55b}$$

In the complex case $\mathbb{K}=\mathbb{C}$, the form must be *sesquilinear*, i.e., bilinearity and (4.55a) is replaced by[19]

$$\begin{aligned}
&\langle v,w\rangle=\overline{\langle w,v\rangle} &&\text{for } v,w\in V,\\
&\langle u+\lambda v,w\rangle=\langle u,w\rangle+\lambda\langle v,w\rangle &&\text{for all } u,v,w\in V,\ \lambda\in\mathbb{C},\\
&\langle w,u+\lambda v\rangle=\langle w,u\rangle+\overline{\lambda}\langle w,v\rangle &&\text{for all } u,v,w\in V,\ \lambda\in\mathbb{C}.
\end{aligned}$$

The triangle inequality of the norm (4.54) follows from the Schwarz inequality

$$|\langle v,w\rangle|\leq\|v\|\,\|w\|\qquad\text{for } v,w\in V.$$

We describe a pre-Hilbert space by $(V,\langle\cdot,\cdot\rangle)$ and note that this defines uniquely a normed space $(V,\|\cdot\|)$ via (4.54).

If $(V,\langle\cdot,\cdot\rangle)$ is complete, i.e., if $(V,\|\cdot\|)$ is a Banach space, we call $(V,\langle\cdot,\cdot\rangle)$ a *Hilbert space*.

Example 4.106. The Euclidean scalar product on $\mathbb{K}^I$ is defined by

$$\langle v,w\rangle=\sum_{i\in I}v_i\,\overline{w_i}.$$

[19] In physics, the opposite ordering is common: the scalar product is antilinear in the first and linear in the second argument.

4.4.2 Basic Facts about Hilbert Spaces

Vectors $u, v \in V$ are *orthogonal*, if $\langle v,w \rangle = 0$. A subset $S \subset V$ is an *orthogonal system*, if all pairs of different $v, w \in S$ are orthogonal. If an orthogonal system is a basis, it is called an *orthogonal basis*. If, in addition, $\|v\| = \|w\| = 1$ holds, we have *orthonormal* vectors, an *orthonormal system*, and an *orthonormal basis*, respectively. In the infinite dimensional Hilbert case, the term 'orthonormal basis' is to be understood as 'complete basis', which is different from the algebraic basis: $\mathfrak{b} = \{b_\nu : \nu \in B\}$ is a *complete basis* of V, if any $v \in V$ can uniquely be written as unconditionally[20] convergent series $v = \sum_{\nu \in B} \alpha_\nu b_\nu$ ($\alpha_\nu \in \mathbb{K}$). If V is separable, B is (at most) countable; otherwise, B is not countable, but for each $v \in V$ the series $\sum_{\nu \in B} \alpha_\nu b_\nu$ contains only countably many nonzero coefficients.

The orthogonal complement of a subset $S \subset V$ is

$$S^\perp = \{v \in V : \langle v,w \rangle = 0 \text{ for all } w \in S\}.$$

Remark 4.107. (a) Any orthogonal complement is closed.
(b) If $S \subset V$ is a closed subset, $V = S \oplus S^\perp$ is a direct sum, i.e., every $v \in V$ has a unique decomposition $v = s + t$ with $s \in S$ and $t \in S^\perp$.

An unpleasant feature of general Banach spaces is the possible non-reflexivity $X^{**} \supsetneqq X$. This does not happen for Hilbert spaces as stated next.

Remark 4.108. (a) All Hilbert spaces satisfy $V = V^{**}$.
(b) The dual space V^* is isomorphic to V: For any $\varphi \in V^*$ there is exactly one $v_\varphi \in V$ with

$$\varphi(v) = \langle v,v_\varphi \rangle \qquad \text{for all } v \in V \tag{4.56}$$

(theorem of Fréchet-Riesz', cf. Riesz [163, §II.30]). Vice versa, every element $v_\varphi \in V$ generates a functional $\varphi \in V^*$ via (4.56). This defines the Fréchet-Riesz isomorphism $J : V \to V^*$ with $\langle v,w \rangle_V = \langle Jv,Jw \rangle_{V^*}$.

Notation 4.109. (a) For $v \in V$ we shall denote $Jv \in V^*$ by v^*, i.e., $v^*(\cdot) = \langle \cdot, v \rangle$. For finite dimensional vector spaces, v^* equals v^{H} (cf. §2.1).
(b) It is possible (but not necessary) to identify V with V^* by setting $v = v^*$.
(c) Let $v \in V$ and $w \in W$. Then $wv^* \in \mathcal{L}(V, W)$ denotes the operator

$$(wv^*)(x) := v^*(x) \cdot w \in W \qquad \text{for all } x \in V. \tag{4.57}$$

Theorem 4.110. *For every Hilbert space V there is an orthonormal basis* $\{\phi_i : i \in S\}$ *with the property that*

$$v = \sum_{i \in S} \langle v, \phi_i \rangle \phi_i, \quad \|v\|^2 = \sum_{i \in S} |\langle v, \phi_i \rangle|^2 \qquad \textit{for all } v \in V. \tag{4.58}$$

The second identity in (4.58) is the Parseval equality.

[20] An *unconditionally convergent* series gives the same finite value for any ordering of the terms.

Exercise 4.111. Let $v, w \in V$. Show that

$$\langle v, w \rangle = \sum_{i \in S} \langle v, \phi_i \rangle \langle \phi_i, w \rangle$$

for any orthonormal basis $\{\phi_i : i \in S\}$ of V.

4.4.3 Operators on Hilbert Spaces

Throughout this subsection, V and W are Hilbert spaces.

Exercise 4.112. The operator norm $\|\Phi\|_{W \leftarrow V}$ of $\Phi \in \mathcal{L}(V, W)$ defined in (4.6a) coincides with the definition

$$\|\Phi\|_{W \leftarrow V} = \sup_{0 \neq v \in V,\ 0 \neq w \in W} \frac{|\langle \Phi v, w \rangle_W|}{\sqrt{\langle v, v \rangle_V \langle w, w \rangle_W}}.$$

Definition 4.113. (a) The operator $\Phi \in \mathcal{L}(V, W)$ gives rise to the *adjoint* operator[21] $\Phi^* \in \mathcal{L}(W, V)$ defined by

$$\langle \Phi v, w \rangle_W = \langle v, \Phi^* w \rangle_V \ .$$

(b) If $V = W$ and $\Phi = \Phi^* \in \mathcal{L}(V, V)$, the operator is called *self-adjoint*.

Next, we consider the subspace $\mathcal{K}(W, V) \subset \mathcal{L}(W, V)$ of compact operators (cf. Definition 4.12 and §4.2.13). We recall that $W \otimes V$ can be interpreted as a subspace of $\mathcal{K}(W, V)$ (cf. Corollary 4.84).

The (finite) singular value decomposition from Lemma 2.20 can be generalised to the infinite dimensional case.

Theorem 4.114 (infinite singular value decomposition). *(a) For $\Phi \in \mathcal{K}(V, W)$ there are singular values $\sigma_1 \geq \sigma_2 \geq \ldots$ with $\sigma_\nu \searrow 0$ and orthonormal systems $\{w_\nu \in W : \nu \in \mathbb{N}\}$ and $\{v_\nu \in V : \nu \in \mathbb{N}\}$ such that*

$$\Phi = \sum_{\nu=1}^{\infty} \sigma_\nu w_\nu v_\nu^* \qquad \textit{(cf. (4.57))}, \tag{4.59}$$

where the sum converges with respect to the operator norm $\|\cdot\|_{W \leftarrow V}$:

$$\|\Phi - \Phi^{(k)}\|_{W \leftarrow V} = \sigma_{k+1} \searrow 0 \quad \textit{for} \ \ \Phi^{(k)} := \sum_{\nu=1}^{k} \sigma_\nu w_\nu v_\nu^*.$$

(b) Vice versa, any Φ defined by (4.59) with $\sigma_k \searrow 0$ belongs to $\mathcal{K}(V, W)$.

[21] There is a slight difference between the adjoint operator defined here and the dual operator from Definition 4.20, since the latter belongs to $\mathcal{L}(W^*, V^*)$. As we may identify $V = V^*$ and $W = W^*$, this difference is not essential.

Proof. Set $\Psi := \Phi^*\Phi \in \mathcal{L}(V,V)$. As product of compact operators, Ψ is compact. The Riesz-Schauder theory (cf. [82, Theorem 6.4.12]) states that Ψ has eigenvalues λ_ν with $\lambda_\nu \to 0$. Since Ψ is self-adjoint, there are corresponding eigenfunctions w_ν which can be chosen orthonormally defining an orthonormal system $\{v_\nu : \nu \in \mathbb{N}\}$. As Ψ is positive semidefinite, i.e., $\langle \Psi v, v\rangle_V \geq 0$ for all $v \in V$, one concludes that $\lambda_\nu \geq 0$. Hence, the singular values $\sigma_\nu := \sqrt[+]{\lambda_\nu}$ are well-defined. Finally, set $w_\nu := \Phi v_\nu / \|\Phi v_\nu\| = \frac{1}{\sigma_\nu}\Phi v_\nu$ (the latter equality follows from $\|\Phi v_\nu\|^2 = \langle \Phi v_\nu, \Phi v_\nu\rangle = \langle v_\nu, \Phi^*\Phi v_\nu\rangle = \langle v_\nu, \Psi v_\nu\rangle = \lambda_\nu \langle v_\nu, v_\nu\rangle = \lambda_\nu$). The vectors w_ν are already normalised. Since

$$\langle w_\nu, w_\mu\rangle \|\Phi v_\nu\| \|\Phi v_\mu\| = \langle \Phi v_\nu, \Phi v_\mu\rangle = \langle v_\nu, \Phi^*\Phi v_\mu\rangle = \lambda_\mu \langle v_\nu, v_\mu\rangle = 0$$

for $\nu \neq \mu$, $\{w_\nu : \nu \in \mathbb{N}\}$ is an orthonormal system in W.

Besides $\Phi v_\nu = \sigma_\nu w_\nu$ (by definition of w_ν) also $\Phi^{(k)} v_\nu = \sigma_\nu w_\nu$ holds for $\nu \leq k$, since $v_\mu^*(v_\nu) = \langle v_\nu, v_\mu\rangle = \delta_{\nu\mu}$ and

$$\left(\sum_{\mu=1}^{k} \sigma_\mu w_\mu v_\mu^*\right)(v_\nu) = \sum_{\mu=1}^{k} \sigma_\mu w_\mu \delta_{\nu\mu} = \sigma_\nu w_\nu.$$

One concludes that $\left(\Phi - \Phi^{(k)}\right)(v_\nu) = 0$ for $\nu \leq k$, while $\left(\Phi - \Phi^{(k)}\right)(v_\nu) = \Phi(v_\nu)$ for $\nu > k$. Hence, $\left(\Phi - \Phi^{(k)}\right)^*\left(\Phi - \Phi^{(k)}\right)$ has the eigenvalues $\sigma_{k+1}^2 \geq \sigma_{k+2}^2 \geq \ldots$ This implies that

$$\|\Phi - \Phi^{(k)}\|_{W \leftarrow V} = \sigma_{k+1}.$$

Convergence follows by $\sigma_\nu \searrow 0$.

For the opposite direction use that $\Phi^{(k)}$ is compact because of the finite dimensional range and that limits of compact operators are again compact. □

Corollary 4.115. If $\kappa(\cdot,\cdot)$ is the Schwartz kernel of Φ, i.e.,

$$\Phi(v) := \int_\Omega \kappa(\cdot, y) v(y) \mathrm{d}y \qquad \text{for } v \in V,$$

we may write (4.59) as

$$\kappa(x,y) = \sum_{\nu=1}^{\infty} \sigma_\nu w_\nu(x) v_\nu(y).$$

Representation (4.59) allows us to define a scale of norms $\|\cdot\|_{\mathrm{SVD},p}$ (cf. (4.17)), which use the ℓ^p norm of the sequence $\boldsymbol{\sigma} = (\sigma_\nu)_{\nu=1}^\infty$ of singular values.[22]

Remark 4.116. (a) $\|\Phi\|_{\mathrm{SVD},\infty} = \|\Phi\|_{\vee(W,V)} = \|\Phi\|_{V \leftarrow W}$ is the operator norm.

(b) $\|\Phi\|_{\mathrm{SVD},2} = \|\Phi\|_{\mathrm{HS}}$ is the Hilbert-Schmidt norm.

(c) $\|\Phi\|_{\mathrm{SVD},1} = \|\Phi\|_{\wedge(W,V)}$ determines the *nuclear* operators.

[22] In physics, in particular quantum information, the entropy $-\sum_\nu \sigma_\nu \ln(\sigma_\nu)$ is of interest.

In the context of Hilbert spaces, it is of interest that the *Hilbert-Schmidt operators* form again a Hilbert space, where the scalar product is defined via the trace, which for the finite dimensional case is already defined in (2.8). In the infinite dimensional case, this definition is generalised by

$$\operatorname{trace}(\Phi) := \sum_{i\in S} \langle \phi_i, \Phi\phi_i\rangle \quad \text{for any orthonormal basis } \{\phi_i : i \in S\}\,. \tag{4.60}$$

To show that this definition makes sense, one has to prove that the right-hand side does not depend on the particular basis. Let $\{\psi_j : j \in T\}$ be another orthonormal basis. Then Exercise 4.111 shows

$$\begin{aligned}\sum_{i\in S} \langle \phi_i, \Phi\phi_i\rangle &= \sum_{i\in S}\sum_{j\in T} \langle \phi_i, \psi_j\rangle \langle \psi_j, \Phi\phi_i\rangle = \sum_{j\in T} \left\langle \psi_j, \Phi\Big(\sum_{i\in S} \langle \phi_i, \psi_j\rangle\, \phi_i\Big)\right\rangle \\ &= \sum_{j\in T} \langle \psi_j, \Phi\psi_j\rangle\,.\end{aligned}$$

Definition 4.117 (Hilbert-Schmidt space). The Hilbert-Schmidt scalar product of $\Phi, \Psi \in \mathcal{L}(V,W)$ is defined by $\langle\Phi,\Psi\rangle_{\mathrm{HS}} := \operatorname{trace}(\Psi^*\Phi)$ and defines the norm

$$\|\Phi\|_{\mathrm{HS}} := \sqrt{\langle\Phi,\Phi\rangle_{\mathrm{HS}}} = \sqrt{\operatorname{trace}(\Phi^*\Phi)}.$$

The operators $\Phi \in \mathcal{L}(V,W)$ with $\|\Phi\|_{\mathrm{HS}} < \infty$ form the Hilbert-Schmidt space $HS(V,W)$.

As stated in Remark 4.116b, the norms $\|\Phi\|_{\mathrm{SVD},2} = \|\Phi\|_{\mathrm{HS}}$ coincide. Since finiteness of $\sum_{\nu=1}^{\infty}\sigma_\nu^2$ implies $\sigma_\nu \searrow 0$, Theorem 4.114b proves the next result.

Remark 4.118. $HS(V,W) \subset \mathcal{K}(V,W)$.

A tensor from $V \otimes W'$ may be interpreted as map $(v\otimes w') : w \mapsto \langle w, w'\rangle_W \cdot v$ from $\mathcal{L}(W,V)$. In §4.2.13 this approach has led to the nuclear operator equipped with the norm $\|\Phi\|_{\mathrm{SVD},1}$. For Hilbert spaces, the norm $\|\Phi\|_{\mathrm{SVD},2} = \|\Phi\|_{\mathrm{HS}}$ is more natural.

Lemma 4.119. *Let $V \otimes_{\|\cdot\|} W = V \otimes_{\|\cdot\|} W'$ be the Hilbert tensor space generated by the Hilbert spaces V and W. Interpreting $\Phi = v \otimes w \in V \otimes_{\|\cdot\|} W$ as a mapping from $\mathcal{L}(W,V)$, the tensor norm $\|v\otimes w\|$ coincides with the Hilbert-Schmidt norm $\|\Phi\|_{\mathrm{HS}}$.*

Proof. By Theorem 4.114, there is a representation $\Phi = \sum_{\nu=1}^{\infty}\sigma_\nu v_\nu w_\nu^* \in \mathcal{L}(W,V)$ with orthonormal v_ν and w_ν. The Hilbert-Schmidt norm equals $\sqrt{\sum_{\nu=1}^{\infty}\sigma_\nu^2}$ (cf. Remark 4.116b). The interpretation of Φ as a tensor from $V\otimes_{\|\cdot\|}W$ uses the notation $\Phi = \sum\limits_{\nu=1}^{\infty}\sigma_\nu v_\nu \otimes w_\nu$. By orthonormality, $\|v\otimes w\|^2 = \sum\limits_{\nu=1}^{\infty}\sigma_\nu^2$ leads to the same norm. □

Combining this result with Remark 4.118, we derive the next statement.

Remark 4.120. $\Phi \in V \otimes_{\|\cdot\|} W$ interpreted as mapping from W into V is compact.

4.4.4 Orthogonal Projections

A (general) projection is already defined in Definition 3.4.

Definition 4.121. $\varPhi \in \mathcal{L}(V, V)$ is called an *orthogonal projection*, if it is a projection and self-adjoint.

Remark 4.122. (a) Set $R := \mathrm{range}(\varPhi) := \{\varPhi v : v \in V\}$ for a projection $\varPhi \in \mathcal{L}(V, V)$. Then $\varPhi$ is called a *projection onto* R. $v = \varPhi(v)$ holds if and only if $v \in R$.

(b) Let $\varPhi \in \mathcal{L}(V, V)$ be an orthogonal projection onto R. Then R is closed and $\varPhi$ is characterised by

$$\varPhi v = \left\{ \begin{matrix} v & \text{for } v \in R \\ 0 & \text{for } v \in R^{\perp} \end{matrix} \right\}, \qquad \text{where } V = R \oplus R^{\perp} \text{ (cf. Remark 4.107b).} \tag{4.61}$$

(c) Let a closed subspace $R \subset V$ and $w \in V$ be given. Then the *best approximation problem*

$$\text{find a minimiser } v_{\mathrm{best}} \in R \text{ of} \qquad \|w - v_{\mathrm{best}}\| = \min_{v \in R} \|w - v\|$$

has the unique solution $v_{\mathrm{best}} = \varPhi w$, where $\varPhi$ is the projection onto R from (4.61).

(d) An orthogonal projection $0 \neq \varPhi \in \mathcal{L}(V, V)$ has the norm $\|\varPhi\|_{V \leftarrow V} = 1$.

(e) If $\varPhi$ is the orthogonal projection onto $R \subset V$, then $I - \varPhi$ is the orthogonal projection onto $R^{\perp}$.

(f) Let $\{b_1, \ldots, b_r\}$ be an orthonormal basis of a subspace $R \subset V$. Then the orthogonal projection onto R is explicitly given by

$$\varPhi = \sum_{\nu=1}^{r} b_\nu b_\nu^*, \qquad \text{i.e.,} \quad \varPhi v = \sum_{\nu=1}^{r} \langle v, b_\nu \rangle \, b_\nu .$$

In the particular case of $V = \mathbb{K}^n$ with the Euclidean scalar product, one forms the orthogonal matrix $U := [b_1, \ldots, b_r]$. Then

$$\varPhi = U U^{\mathsf{H}} \in \mathbb{K}^{n \times n}$$

is the orthogonal projection onto $R = \mathrm{range}\{U\}$.

Lemma 4.123. *(a) Let $P_1, P_2 \in \mathcal{L}(V, V)$ be two orthogonal projections. Then*

$$\| (I - P_1 P_2) \, v\|_V^2 \leq \| (I - P_1) \, v\|_V^2 + \| (I - P_2) \, v\|_V^2 \quad \textit{for any } v \in V.$$

(b) Let $P_j \in \mathcal{L}(V, V)$ be orthogonal projections for $1 \leq j \leq d$. Then

$$\left\| \left(I - \prod_{j=1}^{d} P_j \right) v \right\|_V^2 \leq \sum_{j=1}^{d} \| (I - P_j) \, v\|_V^2 \qquad \textit{for any } v \in V.$$

holds for any ordering of the factors P_j in the product.

Proof. In $(I - P_1 P_2)\, v = (I - P_1)\, v + P_1\, (I - P_2)\, v$, the two terms on the right-hand side are orthogonal. Therefore

$$\begin{aligned}\|\,(I - P_1 P_2)\, v\|_V^2 &= \|\,(I - P_1)\, v\|_V^2 + \|P_1\,(I - P_2)\, v\|_V^2 \\ &\le \|\,(I - P_1)\, v\|_V^2 + \|P_1\|_{V\leftarrow V}\,\|\,(I - P_2)\, v\|_V^2 \\ &\le \|\,(I - P_1)\, v\|_V^2 + \|\,(I - P_2)\, v\|_V^2\end{aligned}$$

using $\|P_1\|_{V\leftarrow V} \le 1$ from Remark 4.122d proves Part (a). Part (b) follows by induction: replace P_2 in Part (a) by $\prod_{j=2}^{d} P_j$. □

4.5 Tensor Products of Hilbert Spaces

4.5.1 Induced Scalar Product

Let $\langle\cdot,\cdot\rangle_j$ be a scalar product defined on V_j $(1 \le j \le d)$, i.e., V_j is a pre-Hilbert space. Then $\mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j$ is again a pre-Hilbert space with a scalar product $\langle\cdot,\cdot\rangle$ which is defined for elementary tensors $\mathbf{v} = \bigotimes_{j=1}^{d} v$ and $\mathbf{w} = \bigotimes_{j=1}^{d} w^{(j)}$ by

$$\left\langle \bigotimes_{j=1}^{d} v^{(j)}, \bigotimes_{j=1}^{d} w^{(j)} \right\rangle := \prod_{j=1}^{d} \langle v^{(j)}, w^{(j)}\rangle_j \quad \text{for all } v^{(j)}, w^{(j)} \in V_j. \tag{4.62}$$

In the case of norms we have seen that a norm defined on elementary tensors does not determine the norm on the whole tensor space. This is different for a scalar product. One verifies that $\langle \mathbf{v}, \mathbf{w}\rangle$ is a sesquilinear form. Hence, its definition on elementary tensors extends to $\mathbf{V} \times \mathbf{V}$. Also the symmetry $\langle \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{v}\rangle$ follows immediately from the symmetry of $\langle\cdot,\cdot\rangle_j$. It remains to prove the positivity (4.55b).

Lemma 4.124. *Equation (4.62) defines a unique scalar product on* ${}_a\bigotimes_{j=1}^{d} V_j$*, which is called the* induced scalar product.

Proof. 1) Consider $d = 2$, i.e., a scalar product on $V \otimes_a W$. Let $\langle\cdot,\cdot\rangle_V$, $\langle\cdot,\cdot\rangle_W$ be scalar products of V, W, and $\mathbf{x} = \sum_{i=1}^{n} v_i \otimes w_i \neq 0$. Without loss of generality we may assume that the v_i and w_i are linearly independent (cf. Lemma 3.13). Consequently, the Gram matrices $G_v = \left(\langle v_i, v_j\rangle_V\right)_{i,j=1}^{n}$ and $G_w = \left(\langle w_i, w_j\rangle_W\right)_{i,j=1}^{n}$ are positive definite (cf. Exercise 2.16b). The scalar product $\langle \mathbf{x}, \mathbf{x}\rangle$ equals

$$\sum_{i,j=1}^{n} \langle v_i, v_j\rangle_V \langle w_i, w_j\rangle_W = \sum_{i,j=1}^{n} G_{v,ij} G_{w,ij} = \operatorname{trace}(G_v G_w^{\mathsf{T}}).$$

Exercise 2.7a with $A := G_v^{1/2}$ and $B := G_v^{1/2} G_w^{\mathsf{T}}$ (cf. Remark 2.13a) yields $\operatorname{trace}(G_v G_w^{\mathsf{T}}) = \operatorname{trace}(G_v^{1/2} G_w^{\mathsf{T}} G_v^{1/2})$. The positive definite matrix $G_v^{1/2} G_w^{\mathsf{T}} G_v^{1/2}$ has positive diagonal elements (cf. Remark 2.13b), proving $\langle \mathbf{x}, \mathbf{x}\rangle > 0$.

2) For $d \geq 3$ the assertion follows by induction: ${}_a\bigotimes_{j=1}^{d} V_j = \left({}_a\bigotimes_{j=1}^{d-1} V_j\right) \otimes_a V_d$ with the scalar product of ${}_a\bigotimes_{j=1}^{d-1} V_j$ as in (4.62), but with d replaced by $d-1$. □

Definition (4.62) implies that elementary tensors $\mathbf{v}$ and $\mathbf{w}$ are orthogonal if and only if $v_j \perp w_j$ for at least one index j. A simple observation is stated next.

Remark 4.125. Orthogonal [orthonormal] systems $\{\phi_i^{(j)} : i \in B_j\} \subset V_j$ for $1 \leq j \leq d$ induce the orthogonal [orthonormal] system in $\mathbf{V}$ consisting of

$$\boldsymbol{\phi}_{\mathbf{i}} := \bigotimes_{j=1}^{d} \phi_{i_j}^{(j)} \qquad \text{for all } \mathbf{i} = (i_1, \ldots, i_d) \in B := B_1 \times \ldots \times B_d.$$

If $\{\phi_i^{(j)} : i \in B_j\}$ are orthonormal bases, $\{\boldsymbol{\phi}_{\mathbf{i}} : \mathbf{i} \in B\}$ is an orthonormal basis of $\mathbf{V}$.

Example 4.126. Consider $V_j = \mathbb{K}^{I_j}$ endowed with the Euclidean scalar product from Example 4.106. Then the induced scalar product of $\mathbf{v}, \mathbf{w} \in \mathbf{V} = \bigotimes_{j=1}^{d} V_j$ is given by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{\mathbf{i} \in \mathbf{I}} \mathbf{v}_{\mathbf{i}} \overline{\mathbf{w}_{\mathbf{i}}} = \sum_{i_1 \in I_1} \cdots \sum_{i_d \in I_d} \mathbf{v}[i_1 \cdots i_d] \, \overline{\mathbf{w}[i_1 \cdots i_d]}.$$

The corresponding (Euclidean) norm is denoted by $\|\cdot\|$ or more specifically by $\|\cdot\|_2$.

There is a slight mismatch in the matrix case $d = 2$: the previously defined norm $\|v\|_2 = \sqrt{\sum_{i,j} |v_{ij}|^2}$ is introduced for matrices as Frobenius norm $\|\cdot\|_{\mathsf{F}}$.

The standard Sobolev space H^N is a Hilbert space corresponding to $p = 2$ in Example 4.101. As seen in §4.3.6, H^N is an intersection space with a particular intersection norm. Therefore we cannot define $H^N = \bigotimes_{j=1}^{d} V_j$ by the induced scalar product (4.62). Let $\left(V_j^{(n)}, \langle \cdot, \cdot \rangle_{j,n}\right)$, the space $\mathbf{V}^{(\mathbf{n})}$, and the set $\mathcal{N}$ be defined as in §4.3.6. Then the canonical scalar product on $\mathbf{V} = \bigcap_{\mathbf{n} \in \mathcal{N}} \mathbf{V}^{(\mathbf{n})}$ is defined by

$$\left\langle \bigotimes_{j=1}^{d} v^{(j)}, \bigotimes_{j=1}^{d} w^{(j)} \right\rangle := \sum_{\mathbf{n} \in \mathcal{N}} \prod_{j=1}^{d} \langle v^{(j)}, w^{(j)} \rangle_{j,n_j} \text{ for all } v^{(j)}, w^{(j)} \in V_j^{(N_j)}. \tag{4.63a}$$

In this definition, $\mathbf{v}$ and $\mathbf{w}$ are elementary tensors of the space $\mathbf{V}_{\mathrm{mix}}$, which by Remark 4.103b is dense in $\mathbf{V}$. The bilinear (sesquilinear) form defined in (4.63a) is positive, since a convex combination of positive forms is again positive. The corresponding norm

$$\|\mathbf{v}\| = \sqrt{\sum_{\mathbf{n} \in \mathcal{N}} \|\mathbf{v}\|_{\mathbf{n}}^2} \tag{4.63b}$$

is equivalent to $\max_{\mathbf{n} \in \mathcal{N}} \|\mathbf{v}\|_{\mathbf{n}}$ from (4.52b).

4.5.2 Crossnorms

Proposition 4.127. *The norm derived from the scalar product (4.62) is a reasonable crossnorm. Furthermore, it is a uniform crossnorm, i.e.,*

$$\left\| \bigotimes_{j=1}^{d} A^{(j)} \right\|_{\mathbf{V}\leftarrow\mathbf{V}} = \prod_{j=1}^{d} \|A^{(j)}\|_{V_j\leftarrow V_j} \quad \textit{for all } A^{(j)} \in \mathcal{L}(V_j, V_j). \tag{4.64}$$

Proof. 1) Taking $\mathbf{v} = \mathbf{w}$ in (4.62) shows $\|\mathbf{v}\| = \prod_{i=1}^{d} \|v_i\|$ for all $\mathbf{v} \in \mathbf{V}$.

2) Since the dual spaces V_i^* and $\mathbf{V}^*$ may be identified with V_i and $\mathbf{V}$, part 1) shows also the crossnorm property (4.41) for $\mathbf{V}^*$.

3) First we consider the *finite* dimensional case. Let $\mathbf{A} = \bigotimes_{j=1}^{d} A^{(j)}$ with $A^{(j)} \in \mathcal{L}(V_j, V_j)$ and $\mathbf{v} \in \mathbf{V}$. Diagonalisation yields $A^{(j)*}A^{(j)} = U_j^* D_j U_j$ (U_j unitary, D_j diagonal). The columns $\{\phi_i^{(j)} : 1 \le i \le \dim(V_j)\}$ of U_j form an orthonormal bases of V_j. Define the orthonormal basis $\{\boldsymbol{\phi}_\mathbf{i}\}$ according to Remark 4.125 and represent $\mathbf{v}$ as $\mathbf{v} = \sum_\mathbf{i} c_\mathbf{i} \boldsymbol{\phi}_\mathbf{i}$. Note that $\|\mathbf{v}\|^2 = \sum_\mathbf{i} |c_\mathbf{i}|^2$ (cf. (4.58)). Then

$$\|\mathbf{A}\mathbf{v}\|^2 = \left\| \sum_\mathbf{i} c_\mathbf{i} \bigotimes_{j=1}^{d} A^{(j)}(\phi_{i_j}^{(j)}) \right\|^2 = \sum_{\mathbf{i},\mathbf{k}} \left\langle c_\mathbf{i} \bigotimes_{j=1}^{d} A^{(j)}(\phi_{i_j}^{(j)}), c_\mathbf{k} \bigotimes_{j=1}^{d} A^{(j)}(\phi_{k_j}^{(j)}) \right\rangle$$

$$= \sum_{\mathbf{i},\mathbf{k}} c_\mathbf{i} \overline{c_\mathbf{k}} \prod_{j=1}^{d} \left\langle A^{(j)}\phi_{i_j}^{(j)}, A^{(j)}\phi_{k_j}^{(j)} \right\rangle_j = \sum_{\mathbf{i},\mathbf{k}} c_\mathbf{i} \overline{c_\mathbf{k}} \prod_{j=1}^{d} \left\langle \phi_{i_j}^{(j)}, A^{(j)*}A^{(j)}\phi_{k_j}^{(j)} \right\rangle.$$

Since $\phi_{k_j}^{(j)}$ are eigenvectors of $A^{(j)*}A^{(j)}$, the products $\langle \phi_{i_j}^{(j)}, A^{(j)*}A^{(j)}\phi_{k_j}^{(j)} \rangle$ vanish for $i_j \ne k_j$. Hence,

$$\|\mathbf{A}\mathbf{v}\|^2 = \sum_\mathbf{i} |c_\mathbf{i}|^2 \prod_{j=1}^{d} \left\langle \phi_{i_j}^{(j)}, A^{(j)*}A^{(j)}\phi_{i_j}^{(j)} \right\rangle = \sum_\mathbf{i} |c_\mathbf{i}|^2 \prod_{j=1}^{d} \left\| A^{(j)}\phi_{i_j}^{(j)} \right\|_j^2$$

$$\le \left(\prod_{j=1}^{d} \|A^{(j)}\|_{V_j\leftarrow V_j} \right)^2 \sum_\mathbf{i} |c_\mathbf{i}|^2 \prod_{j=1}^{d} \underbrace{\|\phi_{i_j}^{(j)}\|_j^2}_{=1}$$

$$= \left(\prod_{j=1}^{d} \|A^{(j)}\|_{V_j\leftarrow V_j} \right)^2 \sum_\mathbf{i} |c_\mathbf{i}|^2 = \left(\prod_{j=1}^{d} \|A^{(j)}\|_{V_j\leftarrow V_j} \right)^2 \|\mathbf{v}\|^2$$

proves that the crossnorm is uniform (the equality in (4.64) is trivial).

4) Next we consider the *infinite* dimensional case. The tensor $\mathbf{v} \in \mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j$ has some representation $\mathbf{v} = \sum_{i=1}^{n} \bigotimes_{j=1}^{d} v_i^{(j)}$, therefore $\mathbf{v} \in \mathbf{V}_0 := \bigotimes_{j=1}^{d} V_{0,j}$ with the finite dimensional subspaces $V_{0,j} := \operatorname{span}\{v_i^{(j)} : 1 \le i \le n\}$. Let $\Phi_j = \Phi_j^* \in \mathcal{L}(V_j, V_j)$ be the orthogonal projection onto $V_{0,j}$. An easy exercise shows

$$\|\mathbf{A}\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{A}^*\mathbf{A}\mathbf{v} \rangle$$
$$= \left\langle \mathbf{v}, \left(\bigotimes_{j=1}^{d} A^{(j)*}A^{(j)} \right) \mathbf{v} \right\rangle = \left\langle \mathbf{v}, \left(\bigotimes_{j=1}^{d} \Phi_j^* A^{(j)*}A^{(j)}\Phi_j \right) \mathbf{v} \right\rangle.$$

Set $C_j := \Phi_j^* A^{(j)*} A^{(j)} \Phi_j = (A^{(j)}\Phi_j)^*(A^{(j)}\Phi_j) = B^{(j)*}B^{(j)}$ for the well-defined square root $B^{(j)} := C_j^{1/2}$. Since the operator acts in the finite dimensional subspace $V_{0,j}$ only, Part 3) applies. The desired estimate follows from

$$\begin{aligned}\|B^{(j)}\|_{V_j\leftarrow V_j}^2 &= \|B^{(j)*}B^{(j)}\|_{V_j\leftarrow V_j} = \|(A^{(j)}\Phi_j)^*(A^{(j)}\Phi_j)\|_{V_j\leftarrow V_j} \\ &= \|A^{(j)}\Phi_j\|_{V_j\leftarrow V_j}^2 \le \|A^{(j)}\|_{V_j\leftarrow V_j}^2 \|\Phi_j\|_{V_j\leftarrow V_j}^2 = \|A^{(j)}\|_{V_j\leftarrow V_j}^2\end{aligned}$$

(cf. Remark 4.122). □

The projective crossnorm $\|\cdot\|_\wedge$ for $\ell^2(I) \times \ell^2(J)$ is discussed in Example 4.48. The result shows that the generalisation for $d \ge 3$ does not lead to a standard norm.

The injective crossnorm $\|\cdot\|_\vee$ of ${}_a\bigotimes_{j=1}^d V_j$ is defined in (4.47). For instance, $V_j = \ell^2(I_j)$ endowed with the Euclidean scalar product leads to

$$\|\mathbf{v}\|_{\vee(\ell^2,\ldots,\ell^2)} = \sup_{\substack{0\neq w^{(j)}\in V_j\\ 1\le j\le d}} \frac{\left|\sum_{i_1\in I_1}\cdots\sum_{i_d\in I_d} \mathbf{v}[i_1\cdots i_d]\cdot w_{i_1}^{(1)}\cdot\ldots\cdot w_{i_d}^{(d)}\right|}{\|w^{(1)}\|_2\cdot\ldots\cdot\|w^{(d)}\|_d}.$$

If $d=1$, $\|\mathbf{v}\|_\vee$ coincides with $\|\mathbf{v}\|_2$. For $d=2$, $\|\mathbf{v}\|_\vee$ is the spectral norm $\|\mathbf{v}\|_2$ for $\mathbf{v}$ interpreted as matrix (cf. (2.13)).

4.5.3 Tensor Products of $\mathcal{L}(V_j, V_j)$

The just proved uniformity shows that the Banach spaces $\left(\mathcal{L}(V_j,V_j), \|\cdot\|_{V_j\leftarrow V_j}\right)$ form the tensor space

$${}_a\bigotimes_{j=1}^d \mathcal{L}(V_j,V_j) \subset \mathcal{L}(\mathbf{V},\mathbf{V})$$

and that the operator norm $\|\cdot\|_{\mathbf{V}\leftarrow\mathbf{V}}$ is a crossnorm (cf. (4.64)).

Note that $(\mathcal{L}(\mathbf{V},\mathbf{V}), \|\cdot\|_{\mathbf{V}\leftarrow\mathbf{V}})$ is a Banach space, but not a Hilbert space. To obtain a Hilbert space, we have to consider the space $HS(V_j,V_j)$ of the Hilbert-Schmidt operators with the scalar product $\langle\cdot,\cdot\rangle_{j,\mathrm{HS}}$ (cf. Definition 4.117). The scalar products $\langle\cdot,\cdot\rangle_{j,\mathrm{HS}}$ induce the scalar product $\langle\cdot,\cdot\rangle_{\mathrm{HS}}$ on $\mathbf{H} := {}_a\bigotimes_{j=1}^d HS(V_j,V_j)$. Equation (4.71) shows that $\langle\cdot,\cdot\rangle_{\mathrm{HS}}$ is defined by the trace on $\mathbf{H}$.

Exercise 4.128. For $A^{(j)}v^{(j)} = \lambda_j v^{(j)}$ ($v^{(j)} \neq 0$) and $\mathbf{A} := \bigotimes_{j=1}^d A^{(j)}$ prove:

(a) The elementary tensor $\mathbf{v} := \bigotimes_{j=1}^d v^{(j)}$ is an eigenvector of $\mathbf{A}$ with eigenvalue $\lambda := \prod_{j=1}^d \lambda_j$, i.e., $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

(b) Assume that $A^{(j)} \in \mathcal{L}(V_j,V_j)$ has $\dim(V_j) < \infty$ eigenpairs $(\lambda_j, v^{(j)})$. Then all eigenpairs $(\lambda,\mathbf{v})$ constructed in Part (a) yield the complete set of eigenpairs of $\mathbf{A}$.

Exercise 4.128b requires that all $A^{(j)}$ are diagonalisable. The next lemma considers the general case.

Lemma 4.129. *Let $A^{(j)} \in \mathbb{C}^{I_j \times I_j}$ be a matrix with $\#I_j < \infty$ for $1 \leq j \leq d$ and form the Kronecker product $\mathbf{A} := \bigotimes_{j=1}^{d} A^{(j)} \in \mathbb{C}^{\mathbf{I}\times\mathbf{I}}$. Let $(\lambda_{j,k})_{k\in I_j}$ be the tuple of eigenvalues of $A^{(j)}$ corresponding to their multiplicity. Then*

$$(\lambda_{\mathbf{k}})_{\mathbf{k}\in\mathbf{I}} \qquad \textit{with } \lambda_{\mathbf{k}} := \prod_{j=1}^{d} \lambda_{j,k_j} \tag{4.65}$$

represents all eigenvalues of $\mathbf{A}$ together with their multiplicity. Note that $\lambda_{\mathbf{k}}$ might be a multiple eigenvalue by two reasons: (a) $\lambda_{j,k_j} = \lambda_{j,k_j+1} = \ldots = \lambda_{j,k_j+\mu-1}$ is a μ-fold eigenvalue of $A^{(j)}$ with $\mu > 1$, (b) different factors $\lambda_{j,k_j} \neq \lambda_{j,k'_j}$ may produce the same product $\lambda_{\mathbf{k}} = \lambda_{\mathbf{k}'}$.

Proof. For each matrix $A^{(j)}$ there is a unitary similarity transformation $R^{(j)} = U^{(j)}A^{(j)}U^{(j)\mathsf{H}}$ ($U^{(j)}$ unitary) into an upper triangular matrix $R^{(j)}$ (Schur normal form; cf. [81, Theorem 2.8.1]). Hence, $A^{(j)}$ and $R^{(j)}$ have identical eigenvalues including their multiplicity. Set $\mathbf{U} := \bigotimes_{j=1}^{d} U^{(j)}$ and $\mathbf{R} := \bigotimes_{j=1}^{d} R^{(j)}$. $\mathbf{U}$ is again unitary (cf. (4.70a,b)), while $\mathbf{R} \in \mathbb{C}^{\mathbf{I}\times\mathbf{I}}$ is of upper triangular form with $\lambda_{\mathbf{k}}$ from (4.65) as diagonal entries. Since the eigenvalues of triangular matrices are given by the diagonal elements of $R^{(j)}$ (including the multiplicity), the assertion follows. □

4.5.4 Partial Scalar Products

Let

$$\mathbf{X} := V_1 \otimes_a W \quad \text{and} \quad \mathbf{Y} := V_2 \otimes_a W$$

be two tensor spaces sharing a pre-Hilbert space $(W, \langle\cdot,\cdot\rangle_W)$. We define a sesquilinear mapping (again denoted by $\langle\cdot,\cdot\rangle_W$) via

$$\langle\cdot,\cdot\rangle_W : \mathbf{X}\times\mathbf{Y} \to V_1 \otimes_a V_2,$$
$$\langle v_1 \otimes w_1, v_2 \otimes w_2\rangle_W := \langle w_1, w_2\rangle_W \cdot v_1 \otimes \overline{v_2} \quad \text{for } v_1 \in V_1,\ v_2 \in V_2,\ w_1, w_2 \in W.$$

We call this operation a *partial scalar product*, since it acts on the W part only.

In the following, we assume $V_1 = V_2$ so that $\mathbf{X} = \mathbf{Y}$. We rename $\mathbf{X}$ by $\mathbf{V}$ with the usual structure $\mathbf{V} = {}_a\bigotimes_{j\in D} V_j$, where, e.g., $D = \{1,\ldots,d\}$. In this case, W from above corresponds to $\mathbf{V}_\alpha = {}_a\bigotimes_{j\in\alpha} V_j$ for a non-empty subset $\alpha \subset D$. The notation $\langle\cdot,\cdot\rangle_W$ is replaced by $\langle\cdot,\cdot\rangle_\alpha$:

$$\langle\cdot,\cdot\rangle_\alpha : \mathbf{V}\times\mathbf{V} \to \mathbf{V}_{D\setminus\alpha} \otimes_a \mathbf{V}_{D\setminus\alpha}, \tag{4.66}$$
$$\left\langle \bigotimes_{j=1}^{d} v_j, \bigotimes_{j=1}^{d} w_j \right\rangle_\alpha := \left[\prod_{j\in\alpha} \langle v_j, w_j\rangle\right] \cdot \left(\bigotimes_{j\in D\setminus\alpha} v_j\right) \otimes \overline{\left(\bigotimes_{j\in D\setminus\alpha} w_j\right)}.$$

The partial scalar product $\langle\cdot,\cdot\rangle_\alpha : \mathbf{V}\times\mathbf{V}\to\mathbf{V}_{\alpha^c}\otimes_a\mathbf{V}_{\alpha^c}$ can be constructed as composition of the following two mappings:

1) sesquilinear concatenation $(\mathbf{v},\mathbf{w})\mapsto\mathbf{v}\otimes\overline{\mathbf{w}}\in\mathbf{V}\otimes_a\mathbf{V}$ followed by

2) *contractions*[23] explained below.

Definition 4.130. For a non-empty, finite index set D let $\mathbf{V}=\mathbf{V}_D={}_a\bigotimes_{j\in D}V_j$ be a pre-Hilbert space with induced scalar product. For any $j\in D$, the contraction $\mathfrak{C}_j:\mathbf{V}\otimes_a\mathbf{V}\to\mathbf{V}_{D\backslash\{j\}}\otimes_a\mathbf{V}_{D\backslash\{j\}}$ is defined by

$$\mathfrak{C}_j\left(\left(\bigotimes_{k\in D}v_k\right)\otimes\left(\bigotimes_{k\in D}w_k\right)\right):=\langle v_j,\overline{w_j}\rangle\left(\bigotimes_{k\in D\backslash\{j\}}v_k\right)\otimes\left(\bigotimes_{k\in D\backslash\{j\}}w_k\right).$$

For a subset $\alpha\subset D$, the contraction $\mathfrak{C}_\alpha:\mathbf{V}\otimes_a\mathbf{V}\to\mathbf{V}_{D\backslash\alpha}\otimes_a\mathbf{V}_{D\backslash\alpha}$ is the product $\mathfrak{C}_\alpha=\prod_{j\in\alpha}\mathfrak{C}_j$ with the action

$$\mathfrak{C}_\alpha\left(\left(\bigotimes_{j\in D}v_j\right)\otimes\left(\bigotimes_{j\in D}w_j\right)\right)=\left[\prod_{j\in\alpha}\langle v_j,\overline{w_j}\rangle\right]\cdot\left(\bigotimes_{j\in D\backslash\alpha}v_j\right)\otimes\left(\bigotimes_{j\in D\backslash\alpha}w_j\right).$$

§5.2 will show further matrix interpretations of these partial scalar products. The definition allows to compute partial scalar product recursively. Formally, we may define $\mathfrak{C}_\emptyset(\mathbf{v}):=\mathbf{v}$ and $\langle\mathbf{v},\mathbf{w}\rangle_\emptyset:=\mathbf{v}\otimes\overline{\mathbf{w}}$.

Corollary 4.131. If $\emptyset\subsetneqq\alpha\subsetneqq\beta\subset D$, then $\langle\mathbf{v},\mathbf{w}\rangle_\beta=\mathfrak{C}_{\beta\backslash\alpha}(\langle\mathbf{v},\mathbf{w}\rangle_\alpha)$.

4.6 Tensor Operations

In the following we enumerate operations which later are to be realised numerically in the various formats. With regard to practical applications, we mainly focus to a finite dimensional setting.

4.6.1 Vector Operations

The trivial vector space operations are the *scalar multiplication* $\lambda\cdot\mathbf{v}$ ($\lambda\in\mathbb{K}$, $\mathbf{v}\in{}_a\bigotimes_{j=1}^d V_j$) and the *addition* $\mathbf{v}+\mathbf{w}$. By definition, $\mathbf{v}$ and $\mathbf{w}\in{}_a\bigotimes_{j=1}^d V_j$ have representations as finite linear combinations. Obviously, the sum might have a representation with even more terms. This will become a source of trouble.

[23] In tensor algebras, contractions are applied to tensors from $\bigotimes_j V_j$, where V_j is either the space V or its dual V'. If, e.g., $V_1=V'$ and $V_2=V$, the corresponding contraction is defined by $\bigotimes_j v^{(j)}\mapsto v^{(1)}(v^{(2)})\cdot\bigotimes_{j\geq3}v^{(j)}$ (cf. Greub [76, p. 72]).

The *scalar product* of two elementary tensors $\mathbf{v} = \bigotimes_{j=1}^{d} v_j$ and $\mathbf{w} = \bigotimes_{j=1}^{d} w_j$ reduces by definition to the scalar product of the simple vectors $v_j, w_j \in V_j$:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \prod_{j=1}^{d} \langle v_j, w_j \rangle . \tag{4.67}$$

The naive computation of the scalar product by $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{\mathbf{i} \in I^d} \mathbf{v_i w_i}$ would be much too costly. Therefore, the reduction to scalar products in V_j is very helpful.

General vectors $\mathbf{v}, \mathbf{w} \in {}_a\bigotimes_{j=1}^{d} V_j$ are sums of elementary tensors. Assume that the (minimal) number of terms is n_v and n_w, respectively. Then $n_v n_w$ scalar products (4.67) must be performed and added. Again, it becomes obvious that large numbers n_v and n_w cause problems.

Note that the evaluation of $\langle \mathbf{v}, \mathbf{w} \rangle$ is not restricted to the discrete setting $V_j = \mathbb{R}^{I_j}$ with finite index sets I_j. Assume the infinite dimensional case of continuous functions from $V_j = C([0,1])$. As long as v_j, w_j belong to a (possibly infinite) family of functions for which the scalar product $\langle v_j, w_j \rangle = \int_0^1 v_j(x) w_j(x) \mathrm{d}x$ is exactly known, the evaluation of $\langle \mathbf{v}, \mathbf{w} \rangle$ can be realised.

4.6.2 Matrix-Vector Multiplication

Again, we consider $V_j = \mathbb{K}^{I_j}$ and $\mathbf{V} = \mathbb{K}^{\mathbf{I}} = \bigotimes_{j=1}^{d} V_j$ with $\mathbf{I} = I_1 \times \ldots \times I_d$. Matrices from $\mathbb{K}^{\mathbf{I} \times \mathbf{I}}$ are described by Kronecker products in $\bigotimes_{j=1}^{d} \mathbb{K}^{I_j \times I_j}$. For elementary tensors $\mathbf{A} = \bigotimes_{j=1}^{d} A^{(j)} \in \bigotimes_{j=1}^{d} \mathbb{K}^{I_j \times I_j}$ and $\mathbf{v} = \bigotimes_{j=1}^{d} v_j \in \bigotimes_{j=1}^{d} \mathbb{K}^{I_j}$ the evaluation of

$$\mathbf{A}\mathbf{v} = \bigotimes_{j=1}^{d} \left(A^{(j)} v^{(j)} \right) \tag{4.68}$$

requires d simple matrix-vector multiplications, while the naive evaluation of $\mathbf{Av}$ may be beyond the computer capacities.

The same holds for a rectangular matrix $\mathbf{A} \in \mathbb{K}^{\mathbf{I} \times \mathbf{J}} = \bigotimes_{j=1}^{d} \mathbb{K}^{I_j \times J_j}$.

4.6.3 Matrix-Matrix Operations

Concerning the addition of Kronecker tensors $\mathbf{A}, \mathbf{B} \in \bigotimes_{j=1}^{d} \mathbb{K}^{I_j \times I_j}$ the same statement holds as for the addition of vectors.

The multiplication rule for elementary Kronecker tensors is

$$\left(\bigotimes_{j=1}^{d} A^{(j)} \right) \left(\bigotimes_{j=1}^{d} B^{(j)} \right) = \bigotimes_{j=1}^{d} A^{(j)} B^{(j)} \quad \text{for all } A^{(j)}, B^{(j)} \in \mathbb{K}^{I_j \times I_j}. \tag{4.69}$$

Similarly for $A^{(j)} \in \mathbb{K}^{I_j \times J_j}$, $B^{(j)} \in \mathbb{K}^{J_j \times K_j}$. If $\mathbf{A}$ $(\mathbf{B})$ is a linear combination of n_A (n_B) elementary Kronecker tensors, $n_A n_B$ evaluations of (4.69) are needed.

Further rules for elementary Kronecker tensors are:

$$\left(\bigotimes_{j=1}^{d} A^{(j)}\right)^{-1} = \bigotimes_{j=1}^{d} \left(A^{(j)}\right)^{-1} \qquad \text{for all invertible } A^{(j)} \in \mathbb{K}^{I_j \times I_j}, \tag{4.70a}$$

$$\left(\bigotimes_{j=1}^{d} A^{(j)}\right)^{\mathsf{T}} = \bigotimes_{j=1}^{d} \left(A^{(j)}\right)^{\mathsf{T}} \qquad \text{for all } A^{(j)} \in \mathbb{K}^{I_j \times J_j}. \tag{4.70b}$$

Exercise 4.132. Assume that all matrices $A^{(j)} \in \mathbb{K}^{I_j \times J_j}$ have one of the properties $\{$regular, symmetric, Hermitean, positive definite, diagonal, lower triangular, upper triangular, orthogonal, unitary, positive, permutation matrix$\}$. Show that the Kronecker matrix $\bigotimes_{j=1}^{d} A^{(j)}$ possesses the same property. What statements hold for negative, negative definite, or antisymmetric matrices $A^{(j)}$?

Exercise 4.133. Let $\mathbf{A} := \bigotimes_{j=1}^{d} A^{(j)}$. Assume that one of the decompositions $A^{(j)} = Q^{(j)}R^{(j)}$ (QR), $A^{(j)} = L^{(j)}L^{(j)\mathsf{H}}$ (Cholesky), or $A^{(j)} = U^{(j)}\Sigma^{(j)}V^{(j)\mathsf{T}}$ (SVD) is given for all $1 \leq j \leq d$. Prove that $\mathbf{A}$ possesses the respective decomposition $\mathbf{QR}$ (QR), $\mathbf{LL}^{\mathsf{H}}$ (Cholesky), $\mathbf{U\Sigma V}^{\mathsf{T}}$ (SVD) with the Kronecker matrices $\mathbf{Q} := \bigotimes_{j=1}^{d} Q^{(j)}$, $\mathbf{R} := \bigotimes_{j=1}^{d} R^{(j)}$, etc.

Exercise 4.134. Prove the following statements about the *matrix rank* (cf. Remark 2.1):

$$\operatorname{rank}\left(\bigotimes_{j=1}^{d} A^{(j)}\right) = \prod_{j=1}^{d} \operatorname{rank}(A^{(j)}),$$

and the *trace* of a matrix (cf. (2.8)):

$$\operatorname{trace}\left(\bigotimes_{j=1}^{d} A^{(j)}\right) = \prod_{j=1}^{d} \operatorname{trace}(A^{(j)}). \tag{4.71}$$

The *determinant* involving $A^{(j)} \in \mathbb{K}^{I_j \times I_j}$ equals

$$\det\left(\bigotimes_{j=1}^{d} A^{(j)}\right) = \prod_{j=1}^{d} \left(\det(A^{(j)})\right)^{p_j} \qquad \text{with } p_j := \prod_{k \in \{1,\ldots,d\} \setminus \{j\}} \# I_k.$$

The latter identity for $d = 2$ is treated in the historical paper by Zehfuss [200] (cf. §1.6): matrices $A \in \mathbb{K}^{p \times p}$ and $B \in \mathbb{K}^{q \times q}$ lead to the determinant

$$\det(A \otimes B) = (\det A)^{q} \, (\det B)^{p}\,.$$

Further statements about elementary Kronecker products can be found in Langville-Stewart [137] and Van Loan-Pitsianis [189].

4.6.4 Hadamard Multiplication

As seen in §1.1.3, univariate functions may be subject of a tensor product producing multivariate functions. Given two functions $f(\mathbf{x})$ and $g(\mathbf{x})$ with $\mathbf{x} = (x_1, \ldots, x_d) \in [0,1]^d$, the (pointwise) multiplication $f \cdot g$ is a standard operation. Replace $[0,1]^d$ by a finite grid

$$G_n := \{\mathbf{x_i} : \mathbf{i} \in \mathbf{I}\} \subset [0,1]^d, \quad \text{where}$$
$$\mathbf{I} = \{\mathbf{i}=(i_1, \ldots, i_d) \colon 0 \le i_j \le n\}, \; \mathbf{x_i} = (x_{i_1}, \ldots, x_{i_d}) \in [0,1]^d, \; x_\nu = \nu/n.$$

Then the entries $\mathbf{a_i} := f(\mathbf{x_i})$ and $\mathbf{b_i} := g(\mathbf{x_i})$ define tensors in $\mathbb{K}^{\mathbf{I}} = \bigotimes_{j=1}^d \mathbb{K}^{I_j}$, where $I_j = \{0, \ldots, n\}$. The pointwise multiplication $f \cdot g$ corresponds to the entry-wise multiplication of $\mathbf{a}$ and $\mathbf{b}$, which is called *Hadamard product*:[24]

$$\mathbf{a} \odot \mathbf{b} \in \mathbb{K}^{\mathbf{I}} \quad \text{with entries} \quad (\mathbf{a} \odot \mathbf{b})_{\mathbf{i}} = \mathbf{a_i b_i} \qquad \text{for all } \mathbf{i} \in \mathbf{I}. \tag{4.72a}$$

Performing the multiplication for all entries would be too costly. For elementary tensors it is much cheaper to use

$$\left(\bigotimes_{j=1}^d a^{(j)} \right) \odot \left(\bigotimes_{j=1}^d b^{(j)} \right) = \bigotimes_{j=1}^d \left(a^{(j)} \odot b^{(j)} \right). \tag{4.72b}$$

The following rules are valid:

$$\begin{aligned} \mathbf{a} \odot \mathbf{b} &= \mathbf{b} \odot \mathbf{a}, \\ (\mathbf{a}' + \mathbf{a}'') \odot \mathbf{b} &= \mathbf{a}' \odot \mathbf{b} + \mathbf{a}'' \odot \mathbf{b}, \\ \mathbf{a} \odot (\mathbf{b}' + \mathbf{b}'') &= \mathbf{a} \odot \mathbf{b}' + \mathbf{a} \odot \mathbf{b}''. \end{aligned} \tag{4.72c}$$

4.6.5 Convolution

There are various versions of a convolution $a \star b$. First, we consider sequences from $\ell_0(\mathbb{Z})$ (cf. Example 3.1). The convolution in $\mathbb{Z}$ is defined by

$$c := a \star b \quad \text{with} \quad c_\nu = \sum_{\mu \in \mathbb{Z}} a_\mu b_{\nu - \mu} \qquad (a, b, c \in \ell_0(\mathbb{Z})). \tag{4.73a}$$

Sequences $a \in \ell_0(\mathbb{N}_0)$ can be embedded into $a^0 \in \ell_0(\mathbb{Z})$ by setting $a_i^0 = a_i$ for $i \in \mathbb{N}_0$ and $a_i^0 = 0$ for $i < 0$. Omitting the zero terms in (4.73a) yields the convolution in $\mathbb{N}_0$:

$$c := a \star b \quad \text{with} \quad c_\nu = \sum_{\mu=0}^{\nu} a_\mu b_{\nu - \mu} \qquad (a, b, c \in \ell_0(\mathbb{N}_0)). \tag{4.73b}$$

[24] Although the name ‘Hadamard product’ for this product is widely used, it does not go back to Hadamard. However, Issai Schur mentions this product in his paper [169] from 1911. In this sense, the term ‘Schur product’ would be more correct.

The convolution of two vectors $a = (a_0, \ldots, a_{n-1})$ and $b = (b_0, \ldots, b_{m-1})$, with possibly $n \neq m$, yields

$$c := a \star b \quad \text{with} \quad c_\nu = \sum_{\mu=\max\{0,\nu-m+1\}}^{\min\{n-1,\nu\}} a_\mu b_{\nu-\mu} \quad \text{for } 0 \leq \nu \leq n+m-2. \tag{4.73c}$$

Note that the resulting vector has increased length: $c = (c_0, \ldots, c_{n+m-2})$.

For finite $a = (a_0, a_1, \ldots, a_{n-1}) \in \ell(\{0, 1, \ldots, n-1\}) = \mathbb{K}^n$, the periodic convolution (with period n) is explained by

$$c := a \star b \quad \text{with } c_\nu = \sum_{\mu=0}^{n-1} a_\mu b_{[\nu-\mu]} \qquad (a, b, c \in \mathbb{K}^n), \tag{4.73d}$$

where $[\nu - \mu]$ is the rest class modulo n, i.e., $[m] \in \{0, 1, \ldots, n-1\}$ with $[m] - m$ being a multiple of n.

Remark 4.135. For $a, b \in \mathbb{K}^n$ let $c \in \mathbb{K}^{2n-1}$ be the result of (4.73c) and define $c^{\text{per}} \in \mathbb{K}^n$ by $c^{\text{per}}_\nu := c_\nu + c_{\nu+n}$ for $0 \leq \nu \leq n-2$ and $c^{\text{per}}_{n-1} := c_{n-1}$. Then c^{per} is the periodic convolution result from (4.73d).

The index sets $\mathbb{Z}$, $\mathbb{N}$, $I_n := \{0, 1, \ldots, n-1\}$ may be replaced by the d-fold products $\mathbb{Z}^d$, $\mathbb{N}^d$, I_n^d. For instance, (4.73a) becomes

$$c := a \star b \quad \text{with} \quad c_{\boldsymbol{\nu}} = \sum_{\boldsymbol{\mu} \in \mathbb{Z}^d} a_{\boldsymbol{\mu}} b_{\boldsymbol{\nu}-\boldsymbol{\mu}} \qquad \left(a, b, c \in \ell_0(\mathbb{Z}^d)\right). \tag{4.73e}$$

For any $I \in \{\mathbb{Z}, \mathbb{N}, I_n\}$, the space $\ell_0(I^d)$ is isomorphic to $\otimes_a^d \ell_0(I)$. For elementary tensors $a, b \in \otimes_a^d \ell_0(I)$, we may apply the following rule:

$$\left(\bigotimes_{j=1}^{d} a^{(j)}\right) \star \left(\bigotimes_{j=1}^{d} b^{(j)}\right) = \bigotimes_{j=1}^{d} a^{(j)} \star b^{(j)}, \quad a^{(j)}, b^{(j)} \in \ell_0(I). \tag{4.74}$$

Note that $a \star b$ is again an elementary tensor.

Since almost all entries of $a \in \ell_0$ are zero, the sums in (4.73a) and (4.74) contain only finitely many nonzero terms. If we replace ℓ_0 by some Banach space ℓ^p, the latter sums may contain infinitely many terms and one has to check its convergence.

Lemma 4.136. *For $a \in \ell^p(\mathbb{Z})$ and $b \in \ell^1(\mathbb{Z})$, the sum in (4.73a) is finite and produces $a \star b \in \ell^p(\mathbb{Z})$ for all $1 \leq p \leq \infty$; furthermore,*

$$\|a \star b\|_{\ell^p(\mathbb{Z})} \leq \|a\|_{\ell^p(\mathbb{Z})} \|b\|_{\ell^1(\mathbb{Z})}.$$

Proof. Choose any $d \in \ell^q(\mathbb{Z})$ with $\|d\|_{\ell^q(\mathbb{Z})} = 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then the scalar product $\langle d, a \star b\rangle = \sum_{\nu,\mu\in\mathbb{Z}} a_\mu b_{\nu-\mu} d_\nu$ can be written as $\sum_{\alpha\in\mathbb{Z}} b_\alpha \sum_{\nu\in\mathbb{Z}} a_{\nu-\alpha} d_\nu$. Since the shifted sequence $(a_{\nu-\alpha})_{\nu\in\mathbb{Z}}$ has the norm $\|a\|_{\ell^p(\mathbb{Z})}$, we obtain

$$\left|\sum\nolimits_{\nu\in\mathbb{Z}} a_{\nu-\alpha} d_\nu\right| \leq \|a\|_{\ell^p(\mathbb{Z})} \|d\|_{\ell^q(\mathbb{Z})} = \|a\|_{\ell^p(\mathbb{Z})}.$$

$|\langle d, a \star b\rangle|$ can be estimated by $\sum_{\alpha\in\mathbb{Z}} |b_\alpha| \|a\|_{\ell^p(\mathbb{Z})} = \|a\|_{\ell^p(\mathbb{Z})} \|b\|_{\ell^1(\mathbb{Z})}$. Since ℓ^q is isomorphic to $(\ell^p)'$ for $1 \leq p < \infty$, the assertion is proved except for $m = \infty$. The latter case is an easy conclusion from $|c_{\boldsymbol{\nu}}| \leq \|a\|_{\ell^\infty(\mathbb{Z})} \|b\|_{\ell^1(\mathbb{Z})}$ (cf. (4.73a)). □

So far, discrete convolutions have been described. Analogous integral versions for univariate functions are

$$(f \star g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\mathrm{d}t, \quad (f \star g)(x) = \int_{0}^{x} f(t)g(x-t)\mathrm{d}t, \tag{4.75a}$$

$$(f \star g)(x) = \int_{0}^{1} f(t)g([x-t])\mathrm{d}t, \quad \text{where } [x] \in [0,1),\ [x]-x \in \mathbb{Z}. \tag{4.75b}$$

The multivariate analogue of (4.75a) is

$$(f \star g)(\mathbf{x}) = \int_{\mathbb{R}^d} f(t_1,\ldots,t_d)\, g(x_1-t_1,\ldots,x_d-t_d)\, \mathrm{d}t_1 \ldots \mathrm{d}t_d.$$

Again, elementary tensors $f(\mathbf{x}) = \prod_{j=1}^{d} f^{(j)}(x_j)$ and $g(\mathbf{x}) = \prod_{j=1}^{d} g^{(j)}(x_j)$ satisfy the counterpart of (4.74):

$$\left(\bigotimes_{j=1}^{d} f^{(j)}\right) \star \left(\bigotimes_{j=1}^{d} g^{(j)}\right) = \bigotimes_{j=1}^{d} \left(f^{(j)} \star g^{(j)}\right), \tag{4.75c}$$

i.e., the d-dimensional convolution can be reduced to d one-dimensional ones.

4.6.6 Function of a Matrix

A square matrix of size $n \times n$ has n eigenvalues $\lambda_i \in \mathbb{C}$ ($1 \leq i \leq n$, counted according to their multiplicity). They form the *spectrum*

$$\sigma(M) := \{\lambda \in \mathbb{C} : \lambda \text{ eigenvalue of } M\}.$$

The *spectral radius* is defined by

$$\rho(M) := \max\{|\lambda| : \lambda \in \sigma(M)\}. \tag{4.76}$$

Let $f : \Omega \subset \mathbb{C} \to \mathbb{C}$ be a holomorphic function[25] with open domain Ω. The application of f to a matrix M is possible if

$$\sigma(M) \subset \Omega.$$

Proposition 4.137. *(a) Assume $M \in \mathbb{C}^{I\times I}$ and let D be an (open) domain with $\sigma(M) \subset D \subset \overline{D} \subset \Omega$. Then a holomorphic function on Ω gives rise to a matrix $f(M) \in \mathbb{C}^{I\times I}$ defined by*

$$f(M) := \frac{1}{2\pi i} \int_{\partial D} (\zeta I - M)^{-1} f(\zeta)\, \mathrm{d}\zeta. \tag{4.77a}$$

(b) Assume that $f(z) = \sum_{\nu=0}^{\infty} a_\nu z^\nu$ converges for $|z| < R$ with $R > \rho(M)$. Then an equivalent definition of $f(M)$ is

[25] Functions with other smoothness properties can be considered too. Compare §13.1 in [86].

$$f(M) = \sum_{\nu=0}^{\infty} a_\nu M^\nu. \tag{4.77b}$$

Important functions are, e.g.,

$$\begin{aligned} &f(z) = \exp(z), \quad f(z) = \exp(\sqrt{z}), && \text{if } \sigma(M) \subset \{z \in \mathbb{C} : \Re e(z) > 0\}, \\ &f(z) = 1/z && \text{if } 0 \notin \sigma(M). \end{aligned}$$

Lemma 4.138. *If $f(M)$ is defined, then*

$$f\left(I \otimes \ldots \otimes I \otimes M \otimes I \otimes \ldots \otimes I\right) = I \otimes \ldots \otimes I \otimes f(M) \otimes I \otimes \ldots \otimes I.$$

Proof. Set $\mathbf{M} := I \otimes \ldots \otimes I \otimes M \otimes I \otimes \ldots \otimes I$ and $\mathbf{I} := \bigotimes_{j=1}^{d} I$. As $\sigma(\mathbf{M}) = \sigma(M)$ (cf. Lemma 4.129), $f(M)$ can be defined by (4.77a) if and only if $f(\mathbf{M})$ is well-defined. (4.77a) yields $f(\mathbf{M}) := \frac{1}{2\pi i} \int_{\partial D} (\zeta \mathbf{I} - \mathbf{M})^{-1} f(\zeta) \mathrm{d}\zeta$. Use

$$\begin{aligned} \zeta \mathbf{I} - \mathbf{M} &= I \otimes \ldots \otimes I \otimes (\zeta I) \otimes I \otimes \ldots \otimes I - I \otimes \ldots \otimes I \otimes M \otimes I \otimes \ldots \otimes I \\ &= I \otimes \ldots \otimes I \otimes (\zeta I - M) \otimes I \otimes \ldots \otimes I \end{aligned}$$

and (4.70a) and proceed by

$$\begin{aligned} f(\mathbf{M}) &= \frac{1}{2\pi i} \int_{\partial D} \left(I \otimes \ldots \otimes I \otimes (\zeta I - M)^{-1} \otimes I \otimes \ldots \otimes I \right) f(\zeta)\, \mathrm{d}\zeta \\ &= I \otimes \ldots \otimes I \otimes \left(\frac{1}{2\pi i} \int_{\partial D} (\zeta I - M)^{-1} f(\zeta) \mathrm{d}\zeta \right) \otimes I \otimes \ldots \otimes I \\ &= I \otimes \ldots \otimes I \otimes f(M) \otimes I \otimes \ldots \otimes I\,. \qquad \square \end{aligned}$$

For later use, we add rules about the exponential function.

Lemma 4.139. *(a) If $A, B \in \mathbb{C}^{I \times I}$ are commutative matrices (i.e., $AB = BA$), then*

$$\exp(A) \exp(B) = \exp(A + B). \tag{4.78a}$$

(b) Let $A^{(j)} \in \mathbb{K}^{I_j \times I_j}$ and

$$\begin{aligned} \mathbf{A} = A^{(1)} \otimes I \otimes \ldots \otimes I + I \otimes A^{(2)} \otimes \ldots \otimes I + \ldots \\ + I \otimes \ldots \otimes I \otimes A^{(d-1)} \otimes I + I \otimes \ldots \otimes I \otimes A^{(d)} \in \mathbb{K}^{\mathbf{I} \times \mathbf{I}}. \end{aligned} \tag{4.78b}$$

Then

$$\exp(t\mathbf{A}) = \bigotimes_{j=1}^{d} \exp(tA^{(j)}) \qquad (t \in \mathbb{K}). \tag{4.78c}$$

Proof. The d terms in (4.78b) are pairwise commutative, therefore (4.78a) proves $\exp(\mathbf{A}) = \prod_{j=1}^{d} \exp(\mathbf{A}^{(j)})$ for $\mathbf{A}^{(j)} := I \otimes \ldots \otimes I \otimes A^{(j)} \otimes I \otimes \ldots \otimes I$. Lemma 4.138 shows $\exp(\mathbf{A}^{(j)}) = I \otimes \ldots \otimes I \otimes \exp(A^{(j)}) \otimes I \otimes \ldots \otimes I$. Thanks to (4.69), their product yields $\bigotimes_{j=1}^{d} \exp(A^{(j)})$. Replacing $A^{(j)}$ by $tA^{(j)}$, (4.78c) can be concluded. □

Finally, we mention quite a different kind of a function application to a tensor. Let $\mathbf{v} \in \mathbb{K}^{\mathbf{I}}$ with $\mathbf{I} = I_1 \times \ldots \times I_d$. Then the entry-wise application of a function $f : \mathbb{K} \to \mathbb{K}$ yields

$$f(\mathbf{v}) \in \mathbb{K}^{\mathbf{I}} \quad \text{with} \quad f(\mathbf{v})_{\mathbf{i}} := f(\mathbf{v}_{\mathbf{i}}) \text{ for all } \mathbf{i} \in \mathbf{I}.$$

For a matrix $\mathbf{v} \in \mathbb{K}^{n \times m}$ this is a rather unusual operation. It becomes more natural, when we consider multivariate functions $C(\mathbf{I})$ defined on $\mathbf{I} = I_1 \times \ldots \times I_d$ (product of intervals). Let $\varphi \in C(\mathbf{I}) = {}_{\|\cdot\|_\infty}\bigotimes_{j=1}^{d} C(I_j)$. Then the definition

$$f(\varphi) \in C(\mathbf{I}) \quad \text{with} \quad (f(\varphi))\,(x) = f\,(\varphi(x)) \text{ for all } x \in \mathbf{I}$$

shows that $f(\varphi) = f \circ \varphi$ is nothing than the usual composition of mappings.

If f is a polynomial, one can use the fact that the power function $f(x) = x^n$ applied to $\mathbf{v}$ coincides with the n-fold Hadamard product (4.72a). But, in general, not even for elementary tensors $\mathbf{v}$ the result $f(\mathbf{v})$ has an easy representation.

4.7 Symmetric and Antisymmetric Tensor Spaces

4.7.1 Hilbert Structure

Given a Hilbert space $(V, \langle\cdot,\cdot\rangle_V)$, define $\langle\cdot,\cdot\rangle$ on $\mathbf{V}$ by the induced scalar product (4.62). We recall the set P of permutations and the projections $P_{\mathfrak{S}}$, $P_{\mathfrak{A}}$ (cf. §3.5.1): $P_{\mathfrak{S}}$ and $P_{\mathfrak{A}}$ are orthogonal projections from $\mathbf{V}$ onto the symmetric tensor space $\mathfrak{S}$ and the antisymmetric tensor space $\mathfrak{A}$, respectively (cf. Proposition 3.63).

As a consequence, e.g., the identities

$$\langle P_{\mathfrak{A}}(\mathbf{u}), P_{\mathfrak{A}}(\mathbf{v})\rangle = \langle P_{\mathfrak{A}}(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, P_{\mathfrak{A}}(\mathbf{v})\rangle\,, \tag{4.79a}$$

$$\langle P_{\mathfrak{A}}(\mathbf{u}), \mathbf{A}P_{\mathfrak{A}}(\mathbf{v})\rangle = \langle P_{\mathfrak{A}}(\mathbf{u}), P_{\mathfrak{A}}(\mathbf{A}\mathbf{v})\rangle = \langle P_{\mathfrak{A}}(\mathbf{u}), \mathbf{A}\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{A}P_{\mathfrak{A}}(\mathbf{v})\rangle \tag{4.79b}$$

hold for all $\mathbf{u}, \mathbf{v} \in \mathbf{V}$ and symmetric $\mathbf{A} \in \mathcal{L}(\mathbf{V}, \mathbf{V})$.

By definition of the induced scalar product (4.62), the scalar product $\langle \mathbf{u}, \mathbf{v}\rangle$ of elementary tensors $\mathbf{u}$ and $\mathbf{v}$ reduces to products of scalar products in V. In $\mathfrak{A}$, elementary tensors $\mathbf{u} = \bigotimes_{j=1}^{d} u^{(j)}$ are to be replaced by $P_{\mathfrak{A}}\big(\bigotimes_{j=1}^{d} u^{(j)}\big)$. Their scalar product reduces to determinants of scalar products in V.

Lemma 4.140. *Antisymmetrised elementary tensors satisfy the product rule*

$$\left\langle P_{\mathfrak{A}}\Big(\bigotimes_{j=1}^{d} u^{(j)}\Big), P_{\mathfrak{A}}\Big(\bigotimes_{j=1}^{d} v^{(j)}\Big)\right\rangle = \frac{1}{d!}\det\Big(\langle u^{(i)}, v^{(j)}\rangle_V\Big)_{i,j=1,\ldots,d}. \tag{4.80}$$

Proof. The left-hand side equals $\big\langle \bigotimes_{j=1}^{d} u^{(j)}, P_{\mathfrak{A}}(\bigotimes_{j=1}^{d} v^{(j)})\big\rangle$ because of (4.79a). Definitions (3.44) and (3.45) show that

$$\left\langle \bigotimes_{j=1}^{d} u^{(j)}, P_{\mathfrak{A}}\Big(\bigotimes_{j=1}^{d} v^{(j)}\Big)\right\rangle = \frac{1}{d!}\left\langle \bigotimes_{j=1}^{d} u^{(j)}, \sum_{\pi\in P} \operatorname{sign}(\pi)\pi\Big(\bigotimes_{j=1}^{d} v^{(j)}\Big)\right\rangle$$

$$= \frac{1}{d!}\left\langle \bigotimes_{j=1}^{d} u^{(j)}, \sum_{\pi\in P} \operatorname{sign}(\pi) \bigotimes_{j=1}^{d} v^{(\pi(j))}\right\rangle = \frac{1}{d!}\sum_{\pi\in P} \operatorname{sign}(\pi) \prod_{j=1}^{d} \left\langle u^{(j)}, v^{(\pi(j))}\right\rangle_V.$$

A comparison with (3.48) finishes the proof. □

Corollary 4.141. For *biorthonormal* systems $\{u^{(j)}\}$ and $\{v^{(j)}\}$, i.e., $\langle u^{(i)}, v^{(j)}\rangle_V = \delta_{ij}$, the right-hand side in (4.80) becomes $1/d!$. The systems are in particular biorthonormal, if $u^{(j)} = v^{(j)}$ forms an orthonormal system.

Let $\{b_i : i \in I\}$ be an orthonormal system in V with $\#I \geq d$. For $(i_1, \ldots, i_d) \in I^d$ define the elementary tensor

$$\mathbf{e}^{(i_1,\ldots,i_d)} := \bigotimes_{j=1}^{d} b_{i_j}.$$

If $(i_1, \ldots, i_d)$ contains two identical indices, $P_{\mathfrak{A}}(\mathbf{e}^{(i_1,\ldots,i_d)}) = 0$ follows.

Remark 4.142. For two tuples $(i_1, \ldots, i_d)$ and $(j_1, \ldots, j_d)$ consisting of d different indices, the following identity holds:

$$\left\langle P_{\mathfrak{A}}(\mathbf{e}^{(i_1,\ldots,i_d)}), P_{\mathfrak{A}}(\mathbf{e}^{(j_1,\ldots,j_d)})\right\rangle$$
$$= \begin{cases} \operatorname{sign}(\pi)/d! & \text{if } \pi\,(i_1, \ldots, i_d) = (j_1, \ldots, j_d) \text{ for some } \pi \in P, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. $\langle P_{\mathfrak{A}}(\mathbf{e}^{(i_1,\ldots,i_d)}), P_{\mathfrak{A}}(\mathbf{e}^{(j_1,\ldots,j_d)})\rangle = \langle \mathbf{e}^{(i_1,\ldots,i_d)}, P_{\mathfrak{A}}(\mathbf{e}^{(j_1,\ldots,j_d)})\rangle$ follows from (4.79a). If $\{i_1, \ldots, i_d\} = \{j_1, \ldots, j_d\}$, there is $\pi \in P$ with $\pi(i_1, \ldots, i_d) = (j_1, \ldots, j_d)$, and $P_{\mathfrak{A}}(\mathbf{e}^{(j_1,\ldots,j_d)})$ contains a term $\frac{\operatorname{sign}(\pi)}{d!}\mathbf{e}^{(i_1,\ldots,i_d)}$. All other terms are orthogonal to $\mathbf{e}^{(i_1,\ldots,i_d)}$. □

4.7.2 Banach Spaces and Dual Spaces

Let V be a Banach space (possibly, a Hilbert space) with norm $\|\cdot\|_V$. The norm of the algebraic tensor space $\mathbf{V}_{\text{alg}} = \otimes_a^d V$ is denoted by $\|\cdot\|$. We require that $\|\cdot\|$ is invariant with respect to permutations, i.e.,

$$\|\mathbf{v}\| = \|\pi\,(\mathbf{v})\| \qquad \text{for all } \pi \in P \text{ and } \mathbf{v} \in \otimes_a^d V. \tag{4.81}$$

Conclusion 4.143. *Assume (4.81). The mapping $\pi : \otimes_a^d V \to \otimes_a^d V$ corresponding to $\pi \in P$ as well as the mappings $P_{\mathfrak{S}}$ and $P_{\mathfrak{A}}$ are bounded by* 1.

Proof. The bound for π is obvious. For $P_{\mathfrak{S}}$ use that $\|P_{\mathfrak{S}}\| = \left\|\frac{1}{d!}\sum_{\pi\in P}\pi\right\| \leq \frac{1}{d!}\sum_{\pi\in P}\|\pi\| = \frac{1}{d!}d! = 1$ holds for the operator norm. Similarly for $P_{\mathfrak{A}}$. □

$\mathbf{V}_{\|\cdot\|} := \otimes^d_{\|\cdot\|} V$ is defined by completion with respect to $\|\cdot\|$.

Lemma 4.144. *Assume (4.81). Denote the algebraic symmetric and antisymmetric tensor spaces by $\mathfrak{S}_{alg}(V)$ and $\mathfrak{A}_{alg}(V)$. Both are subspaces of $\mathbf{V}_{alg}$. The completion of $\mathfrak{S}_{alg}(V)$ and $\mathfrak{A}_{alg}(V)$ with respect to $\|\cdot\|$ yields subspaces $\mathfrak{S}_{\|\cdot\|}(V)$ and $\mathfrak{A}_{\|\cdot\|}(V)$ of $\mathbf{V}_{\|\cdot\|}$. An equivalent description of $\mathfrak{S}_{\|\cdot\|}(V)$ and $\mathfrak{A}_{\|\cdot\|}(V)$ is*

$$\mathfrak{S}_{\|\cdot\|}(V) = \left\{\mathbf{v} \in \mathbf{V}_{\|\cdot\|} : \mathbf{v} = \pi(\mathbf{v}) \text{ for all } \pi \in P\right\},$$
$$\mathfrak{A}_{\|\cdot\|}(V) = \left\{\mathbf{v} \in \mathbf{V}_{\|\cdot\|} : \mathbf{v} = \operatorname{sign}(\pi)\pi(\mathbf{v}) \text{ for all } \pi \in P\right\}.$$

Proof. 1) By Conclusion 4.143, π is continuous. For any sequence $\mathbf{v}_n \in \mathfrak{S}_{\mathrm{alg}}(V)$ with $\mathbf{v}_n \to \mathbf{v} \in \mathfrak{S}_{\|\cdot\|}(V)$ the property $\mathbf{v}_n = \pi(\mathbf{v}_n)$ is inherited by $\mathbf{v} \in \mathfrak{S}_{\|\cdot\|}(V)$.

2) Vice versa, let $\mathbf{v} \in \mathbf{V}_{\|\cdot\|}$ with $\mathbf{v} = \pi(\mathbf{v})$ for all $\pi \in P$. This is equivalent to $\mathbf{v} = P_{\mathfrak{S}}(\mathbf{v})$. Let $\mathbf{v}_n \to \mathbf{v}$ for some $\mathbf{v}_n \in \mathbf{V}_{\mathrm{alg}}$ and construct

$$\mathbf{u}_n := P_{\mathfrak{S}}(\mathbf{v}_n) \in \mathfrak{S}_{\mathrm{alg}}(V).$$

Continuity of $P_{\mathfrak{S}}$ (cf. Conclusion 4.143) implies

$$\mathbf{u} := \lim \mathbf{u}_n = P_{\mathfrak{S}}(\lim \mathbf{v}_n) = P_{\mathfrak{S}}(\mathbf{v}) \in \mathbf{V}_{\|\cdot\|};$$

hence $\mathbf{v} = \mathbf{u}$ lies in the completion $\mathfrak{S}_{\|\cdot\|}(V)$ of $\mathfrak{S}_{\mathrm{alg}}(V)$. Analogously for the space $\mathfrak{A}_{\|\cdot\|}(V)$. □

Any dual form $\varphi \in \otimes^d_a V'$ is also a dual form on the subspaces $\mathfrak{S}_{\mathrm{alg}}(V)$ and $\mathfrak{A}_{\mathrm{alg}}(V)$. For $\pi \in P$ let π' be the dual mapping, i.e., $\pi'(\bigotimes_{j=1}^d \varphi_j) \in \otimes^d_a V'$ acts as $\left(\pi' \bigotimes_{j=1}^d \varphi_j\right)(\mathbf{v}) = \varphi(\pi(\mathbf{v}))$. One concludes that $\pi' \bigotimes_{j=1}^d \varphi_j = \bigotimes_{j=1}^d \varphi_{\pi(j)}$ and that all $\varphi \in \otimes^d_a V'$ with $P'_{\mathfrak{S}}\varphi = 0$ represent the zero mapping on $\mathfrak{S}_{\mathrm{alg}}(V)$. Thus, $\otimes^d_a V'$ reduces to the quotient space $\otimes^d_a V' / \ker P'_{\mathfrak{S}}$. A comparison with (3.47) shows that $\otimes^d_a V' / \ker P'_{\mathfrak{S}}$ can be viewed as the symmetric tensor space $\mathfrak{A}_{\mathrm{alg}}(V')$ derived from V'. Similarly, $\mathfrak{A}_{\mathrm{alg}}(V') \cong \otimes^d_a V' / \ker P'_{\mathfrak{A}}$.

The same statements hold for the continuous functionals:

$$\mathfrak{S}_{\|\cdot\|^*}(V^*) \cong \left(\otimes^d_{\|\cdot\|^*} V^*\right) / \ker P^*_{\mathfrak{S}} \quad \text{and}$$
$$\mathfrak{A}_{\|\cdot\|^*}(V^*) \cong \left(\otimes^d_{\|\cdot\|^*} V^*\right) / \ker P^*_{\mathfrak{A}},$$

where by the previous considerations $\ker P^*_{\mathfrak{S}}$ and $\ker P^*_{\mathfrak{A}}$ are closed subspaces.

Chapter 5
General Techniques

Abstract In this chapter, isomorphisms between the tensor space of order d and vector spaces or other tensor spaces are considered. The *vectorisation* from *Sect. 5.1* ignores the tensor structure and treats the tensor space as a usual vector space. In finite dimensional implementations this means that multivariate arrays are organised as linear arrays. After vectorisation, linear operations between tensor spaces become matrices expressed by Kronecker products (cf. §5.1.2).

While vectorisation ignores the tensor structure completely, *matricisation* keeps one of the spaces and leads to a tensor space of order two (cf. *Sect. 5.2*). In the finite dimensional case, this space is isomorphic to a matrix space. The interpretation as matrix allows to formulate typical matrix properties like the rank leading to the j-rank for a direction j and the α-rank for a subset α of the directions $1, \ldots, d$. In the finite dimensional or Hilbert case, the singular value decomposition can be applied to the matricised tensor.

In *Sect. 5.3*, the *tensorisation* is introduced, which maps a vector space (usually without any tensor structure) into an isomorphic tensor space. The artificially constructed tensor structure allows interesting applications. While Sect. 5.3 gives only an introduction into this subject, details about tensorisation will follow in Chap. 14.

5.1 Vectorisation

5.1.1 Tensors as Vectors

In program languages, matrices or multi-dimensional arrays are mapped internally into a linear array (vector) containing the entries in a lexicographical ordering. Note that the ordering is not uniquely determined (even different program languages may use different lexicographical orderings). Without further data, it is impossible to restore a matrix or even its format from the vector. This fact expresses that structural data are omitted.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_5,

Vectorisation is implicitly expressed by the notation (1.5) for Kronecker matrices. Applying this notation to $n \times 1$ matrices which are regarded as (column) vectors, (1.5) becomes

$$a \otimes b = \begin{bmatrix} a_1 b \\ a_2 b \\ \vdots \end{bmatrix} \in \mathbb{K}^{n \cdot m} \text{ for } a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \end{bmatrix} \in \mathbb{K}^n \text{ and } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \end{bmatrix} \in \mathbb{K}^m.$$

Hence, the resulting tensor is immediately expressed as vector in $\mathbb{K}^{n \cdot m}$. Only with this vectorisation, Kronecker products of matrices $A \in \mathbb{K}^{n \times n}$ and $B \in \mathbb{K}^{m \times m}$ can be interpreted as matrices from $\mathbb{K}^{n \cdot m \times n \cdot m}$.

For a mathematical formulation of the vectorisation assume $V_j = \mathbb{K}^{I_j}$ with $n_j := \#I_j < \infty$. Choose any index set J with $\#J = \prod_{j=1}^d n_j$ together with a bijection $\phi : I_1 \times \ldots \times I_d \to J$. This defines the isomorphism $\Phi : \mathbf{V} = \bigotimes_{j=1}^d V_j \to \mathbb{K}^J$ between the tensor space $\mathbf{V}$ and the vector space $\mathbb{K}^J$. Tensor entries $\mathbf{v}[i_1, \ldots, i_d]$ are mapped into vector entries $v[\phi(i_1, \ldots, i_d)]$ and vice versa. Note that Φ is a vector space isomorphism in the sense of §3.2.5.

In the case of a linear system with n equations and unknowns, we are used to deal with vectors x, b and a matrix M:

$$Mx = b \qquad \left(M \in \mathbb{K}^{J \times J},\ x, b \in \mathbb{K}^J\right). \tag{5.1a}$$

In particular, LU and Cholesky decompositions require an ordered index set J, e.g., $J = \{1, \ldots, N\}$ with $N := \#J$.

An example, where such a system is described differently, are matrix equations, e.g., the *Lyapunov matrix equation*

$$AX + XA^\top = B, \tag{5.1b}$$

where matrices $A, B \in \mathbb{K}^{I \times I}$ are given and the solution $X \in \mathbb{K}^{I \times I}$ is sought. Let $n := \#I$. The number of unknown entries X_{ij} is n^2. Furthermore, Eq. (5.1b) is linear in all unknowns X_{ij}, i.e., (5.1b) turns out to represent a linear system of n^2 equations for n^2 unknowns. Lemma 5.1 from below allows us to translate the Lyapunov equation (5.1b) into

$$\mathbf{Ax} = \mathbf{b}, \tag{5.1c}$$

where $\mathbf{x}, \mathbf{b} \in \mathbf{V} := \mathbb{K}^I \otimes \mathbb{K}^I$ are tensors and $\mathbf{A} \in L(\mathbf{V}, \mathbf{V})$ is the following Kronecker product:

$$\mathbf{A} = A \otimes I + I \otimes A \in L(\mathbf{V}, \mathbf{V}). \tag{5.1d}$$

Using the vectorisation isomorphism $\Phi : \mathbf{V} \to \mathbb{K}^J$ from above, we obtain the linear system (5.1a) with $M = \Phi \mathbf{A} \Phi^{-1}$, $x = \Phi \mathbf{x}$, and $b = \Phi \mathbf{b}$.

Lemma 5.1. *The matrices $U, V \in \mathbb{K}^{I \times I}$ define Kronecker products $\mathbf{U} = U \otimes I$, $\mathbf{V} = I \otimes V$, $\mathbf{W} = U \otimes V \in L(\mathbb{K}^I \otimes \mathbb{K}^I, \mathbb{K}^I \otimes \mathbb{K}^I)$. The products $\mathbf{Ux}$, $\mathbf{Vx}$, and $\mathbf{Wx}$ correspond to UX, $XV^\top$, and $UXV^\top$, where $X \in \mathbb{K}^{I \times I}$ is the matrix interpretation of the tensor $\mathbf{x} \in \mathbb{K}^I \otimes \mathbb{K}^I$.*

Proof. $W := UXV^\top$ has the matrix coefficients $W_{i,j} = \sum_{k,\ell\in I} U_{i,k}\, X_{k,\ell}\, V_{j,\ell}$. The Kronecker matrix $\mathbf{W} = U \otimes V$ has the entries $\mathbf{W}_{(i,j),(k,\ell)} = U_{i,k}\, V_{j,\ell}$. Hence,

$$(\mathbf{W}\mathbf{x})_{(i,j)} = \sum_{(k,\ell)\in I\times I} \mathbf{W}_{(i,j),(k,\ell)}\, \mathbf{x}_{(k,\ell)} = W_{i,j}.$$

The special cases $U = I$ or $V = I$ yield the first two statements. □

From (5.1d) we easily conclude that a positive definite matrix A in (5.1b) leads to a positive definite matrix M and, therefore, enables a Cholesky decomposition.

5.1.2 Kronecker Tensors

As already mentioned in the previous section, the interpretation of a Kronecker tensor product as a matrix is based on vectorisation. However, there is a second possibility for vectorisation. Matrices, which may be seen as tensors of order two, can be mapped isomorphically into vectors. This will be done by the mappings ϕ_j from below.

Let I_j and J_j be the index sets of the matrix space $M_j := \mathbb{K}^{I_j\times J_j}$. An isomorphic vector space is $V_j := \mathbb{K}^{K_j}$ with $K_j = I_j \times J_j$. The following isomorphism ϕ_j describes the vectorisation:

$$\phi_j : M_j \to V_j, \quad A^{(j)} = \left(A^{(j)}_{\ell,m}\right)_{\ell\in I_j, m\in J_j} \mapsto a^{(j)} := \phi_j(A^{(j)}) = \left(a^{(j)}_i\right)_{i\in K_j}.$$

We identify $\mathbf{M} := \bigotimes_{j=1}^d M_j$ with the matrix space $\mathbb{K}^{\mathbf{I}\times\mathbf{J}}$, where $\mathbf{I} := \times_{j=1}^d I_j$ and $\mathbf{J} := \times_{j=1}^d J_j$, while $\mathbf{V} := \bigotimes_{j=1}^d V_j$ is identified with the vector space $\mathbb{K}^{\mathbf{K}}$ for $\mathbf{K} := \times_{j=1}^d K_j = \times_{j=1}^d (I_j \times J_j)$. Note that the matrix-vector multiplication $\mathbf{y} = \mathbf{M}\mathbf{x}$ is written as $\mathbf{y}_\mathbf{i} = \sum_{\mathbf{j}\in\mathbf{J}} \mathbf{M}_{\mathbf{ij}}\mathbf{x}_\mathbf{j}$ for $\mathbf{i} \in \mathbf{I}$.

Elementary tensors $\mathbf{A} = \bigotimes_{j=1}^d A^{(j)} \in \mathbf{M}$ and $\mathbf{a} = \bigotimes_{j=1}^d a^{(j)} \in \mathbf{V}$ have the entries

$$\mathbf{A}[(\ell_1,\ldots,\ell_d),(m_1,\ldots,m_d)] = \prod_{j=1}^d A^{(j)}_{\ell_j,m_j} \quad \text{and} \quad \mathbf{a}[i_1\ldots,i_d] = \prod_{j=1}^d a^{(j)}_{i_j}.$$

Define $a^{(j)}$ by $\phi_j(A^{(j)})$. Then $\mathbf{a} = \left(\bigotimes_{j=1}^d \phi_j\right)(\mathbf{A})$ holds and gives rise to the following definition:

$$\boldsymbol{\Phi} = \bigotimes_{j=1}^d \phi_j : \mathbf{A} \in \mathbf{M} \mapsto \mathbf{a} \in \mathbf{V} \qquad \text{with}$$
$$\mathbf{A}[(\ell_1,\ldots,\ell_d),(m_1,\ldots,m_d)] \mapsto \mathbf{a}[(\ell_1,m_1),\ldots,(\ell_d,m_d)] \text{ for } (\ell_j,m_j)\in K_j.$$

$\boldsymbol{\Phi}$ can be regarded as vectorisation of the Kronecker matrix space $\mathbf{M}$. For $d \geq 3$, we have the clear distinction that $\mathbf{M}$ is a matrix space, whereas $\mathbf{V}$ is a tensor space of order $d \geq 3$. For $d = 2$, however, also $\mathbf{V}$ can be viewed as a matrix space (cf. Van Loan-Pitsianis [189]).

Remark 5.2. Suppose that $d = 2$.

(a) The matrix $\mathbf{A}$ with entries $\mathbf{A}[(\ell_1, \ell_2), (m_1, m_2)]$ is mapped by $\boldsymbol{\Phi}$ into $\mathbf{a}$ with entries $\mathbf{a}[(\ell_1, m_1), (\ell_2, m_2)]$. Since $d = 2$, the tensor $\mathbf{a}$ can again be viewed as a matrix from $\mathbb{K}^{K_1 \times K_2}$. Note that, in general, $\mathbf{a} \in \mathbb{K}^{K_1 \times K_2}$ is of another format than $\mathbf{A} \in \mathbb{K}^{\mathbf{I} \times \mathbf{J}}$. However, $\#(K_1 \times K_2) = \#(\mathbf{I} \times \mathbf{J})$ holds, and $\mathbf{A}$ and $\mathbf{a}$ have the same Frobenius norm, i.e., $\boldsymbol{\Phi}$ is also isometric (cf. Remark 2.8).

(b) The singular value decomposition of $\mathbf{A}$ in the sense of Lemma 3.18 can be applied as follows. Apply Lemma 3.18 to $\mathbf{a} = \boldsymbol{\Phi}(\mathbf{A})$ resulting in $\mathbf{a} = \sum_{i=1}^{r} \sigma_i \, x_i \otimes y_i$ ($x_i \in V_1$, $y_i \in V_2$). Then application of $\boldsymbol{\Phi}^{-1}$ yields

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i X_i \otimes Y_i \qquad \text{with } X_i = \phi_1^{-1}(x_i),\ Y_i = \phi_2^{-1}(y_i). \tag{5.2}$$

Note that (5.2) is not the singular value decomposition of the *matrix* $\mathbf{A}$.

As an illustration of Part (a) consider the identity matrix $\mathbf{A} = \mathbf{I}$ for the index sets $I_1 = J_1 = \{1, 2\}$ and $I_2 = J_2 = \{a, b, c\}$. The matrices $\mathbf{A}$ and $\mathbf{a} = \boldsymbol{\Phi}(\mathbf{A})$ are given below together with the indices for the rows and columns (only nonzero entries are indicated):

$$\mathbf{A} = \begin{array}{c|cccccc|} & \text{1a} & \text{1b} & \text{1c} & \text{2a} & \text{2b} & \text{2c} \\ \hline \text{1a} & 1 & & & & & \\ \text{1b} & & 1 & & & & \\ \text{1c} & & & 1 & & & \\ \text{2a} & & & & 1 & & \\ \text{2b} & & & & & 1 & \\ \text{2c} & & & & & & 1 \\ \hline \end{array} \quad \underset{\boldsymbol{\Phi}}{\mapsto} \quad \mathbf{a} = \begin{array}{c|ccccccccc|} & \text{aa} & \text{ab} & \text{ac} & \text{ba} & \text{bb} & \text{bc} & \text{ca} & \text{cb} & \text{cc} \\ \hline 11 & 1 & & & & 1 & & & & 1 \\ 12 & & & & & & & & & \\ 21 & & & & & & & & & \\ 22 & 1 & & & & 1 & & & & 1 \\ \hline \end{array}$$

5.2 Matricisation

Synonyms for ‘*matricisation*’ are ‘*matrix unfolding*’ or ‘*flattening*’. We prefer the first term, since it clearly indicates the correspondence to matrices (at least in the finite dimensional case).

We recall the two types of isomorphisms discussed in §3.2.5. The strongest form is the tensor space isomorphism which preserves the tensor structure (cf. Definition 3.27). The weakest form is the vector space isomorphism which identifies all tensor spaces $\mathbf{V}$ and $\mathbf{W}$ of same dimension not regarding the tensor structure. An intermediate form groups the d spaces V_j from $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ such that the order is reduced. For instance, $\bigotimes_{j=1}^{5} V_j$ is isomorphic to the rearrangement

$\mathbf{V}_{\text{new}} = (V_1 \otimes V_5) \otimes (V_2 \otimes V_3) \otimes V_4$, which is a tensor space of order $d_{\text{new}} = 3$. In this setting, vectorisation results from $d_{\text{new}} = 1$, while matricisation corresponds to $d_{\text{new}} = 2$. Since tensor spaces of order two are close to matrix spaces, matricisation tries to exploit all features of matrices.

In the following we use the sign $\otimes$ without subscripts $\otimes_a$ or $\otimes_{\|\cdot\|}$, since both cases are allowed.

5.2.1 General Case

To get $d_{\text{new}} = 2$, we have to divide the whole index set $\{1, \ldots, d\}$ into two (disjoint) subsets. For a systematic approach we introduce the set

$$D = \{1, \ldots, d\} \tag{5.3a}$$

and consider proper subsets

$$\emptyset \subsetneqq \alpha \subsetneqq D. \tag{5.3b}$$

The complement of α is denoted by

$$\alpha^c := D \backslash \alpha. \tag{5.3c}$$

We define the (partial) tensor spaces

$$\mathbf{V}_\alpha = \bigotimes_{j \in \alpha} V_j \qquad \text{for } \alpha \subset D, \tag{5.3d}$$

which include the cases $\mathbf{V}_\emptyset = \mathbb{K}$ for $\alpha = \emptyset$ and $\mathbf{V}_D = \mathbf{V}$ for $\alpha = D$. For singletons we have the synonymous notation

$$\mathbf{V}_{\{j\}} = V_j \qquad (j \in D). \tag{5.3e}$$

Instead of $\mathbf{V}_{\{j\}^c}$ˆ for the complement $\{j\}^c$ we have already introduced the symbol

$$\mathbf{V}_{[j]} = \bigotimes_{k \in D \backslash \{j\}} V_k$$

(cf. (3.21a)). Depending on the context, the spaces $\mathbf{V}_\alpha$ and $\mathbf{V}_{[j]}$ may be algebraic or topological tensor spaces. Concerning the (choice of the) norm of the partial tensor space $\mathbf{V}_\alpha$ in the case of $\emptyset \subsetneqq \alpha \subsetneqq D$, we refer to §4.3.2.

Below, we introduce the isomorphism $\mathcal{M}_\alpha$ from $\mathbf{V}$ onto the binary tensor space $\mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$:

$$\mathbf{V} = \bigotimes_{j \in D} V_j \cong \mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}. \tag{5.4}$$

Often, α is a singleton $\{j\}$, i.e., $\mathbf{V}_\alpha = \mathbf{V}_{\{j\}} = V_j$. In this case, (5.4) becomes $\mathbf{V} \cong \mathbf{V}_j \otimes \mathbf{V}_{[j]}$ and the isomorphism is denoted by $\mathcal{M}_j$.

Definition 5.3 **($\mathcal{M}_\alpha, \mathcal{M}_j$).** The matricisation $\mathcal{M}_\alpha$ with α from[1] (5.3b) is the isomorphism[2]

$$\begin{aligned} \mathcal{M}_\alpha : \bigotimes_{k\in D} V_k &\to \mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c} \\ \bigotimes_{k\in D} v^{(k)} &\mapsto \mathbf{v}^{(\alpha)} \otimes \mathbf{v}^{(\alpha^c)} \quad \text{with } \mathbf{v}^{(\alpha)} = \bigotimes_{k\in\alpha} v^{(k)},\ \mathbf{v}^{(\alpha^c)} = \bigotimes_{k\in\alpha^c} v^{(k)}. \end{aligned}$$

In particular, for $j \in D$, $\mathcal{M}_j$ is the isomorphism

$$\begin{aligned} \mathcal{M}_j : \bigotimes_{k\in D} V_k &\to V_j \otimes \mathbf{V}_{[j]} \\ \bigotimes_{k\in D} v^{(k)} &\mapsto v^{(j)} \otimes \mathbf{v}^{[j]} \quad \text{with } \mathbf{v}^{[j]} = \bigotimes_{k\in D\setminus\{j\}} v^{(k)}. \end{aligned}$$

Next, we check how $\mathcal{M}_\alpha(\mathbf{v})$ behaves when we apply an elementary Kronecker product $\bigotimes_{j=1}^d A^{(j)} : \mathbf{V} \to \mathbf{W}$ to $\mathbf{v}$. This includes the case of a tensor space isomorphism $\boldsymbol{\Phi} : \mathbf{V} \to \mathbf{W}$ (cf. Definition 3.27).

Remark 5.4. For $\mathbf{V} = \bigotimes_{j=1}^d V_j$ and $\mathbf{W} = \bigotimes_{j=1}^d W_j$ let $\mathbf{A} := \bigotimes_{j=1}^d A^{(j)} : \mathbf{V} \to \mathbf{W}$ be an elementary Kronecker product. For α from (5.3b) set $\mathbf{A}^{(\alpha)} := \bigotimes_{j\in\alpha} A^{(j)}$ and $\mathbf{A}^{(\alpha^c)} := \bigotimes_{j\in\alpha^c} A^{(j)}$. Then,

$$\mathcal{M}_\alpha(\mathbf{A}\mathbf{v}) = \left(\mathbf{A}^{(\alpha)} \otimes \mathbf{A}^{(\alpha^c)}\right) \mathcal{M}_\alpha(\mathbf{v}) \qquad \text{for all } \mathbf{v} \in \mathbf{V}.$$

If $A^{(j)} : V_j \to W_j$ are isomorphisms, $\mathbf{A}^{(\alpha)} \otimes \mathbf{A}^{(\alpha^c)}$ describes the isomorphism between $\mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$ and $\mathbf{W}_\alpha \otimes \mathbf{W}_{\alpha^c}$.

5.2.2 *Finite Dimensional Case*

5.2.2.1 Example

For finite dimensions, the binary tensor space $\mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$ resulting from the matricisation may be interpreted as matrix space (cf. §3.2.3). If, e.g., $V_j = \mathbb{K}^{I_j}$, then $\mathcal{M}_\alpha$ maps into[3] $\mathbb{K}^{\mathbf{I}_\alpha \times \mathbf{I}_{\alpha^c}}$, where $\mathbf{I}_\alpha = \times_{k\in\alpha} I_k$ and $\mathbf{I}_{\alpha^c} = \times_{k\in\alpha^c} I_k$. Hence, a tensor $\mathbf{v}$ with entries $\mathbf{v}[(i_\kappa)_{\kappa\in D}]$ becomes a matrix $M = \mathcal{M}_\alpha(\mathbf{v})$ with entries $M[(i_\kappa)_{\kappa\in\alpha}, (i_\lambda)_{\lambda\in\alpha^c}]$.

To demonstrate the matricisations, we illustrate all $\mathcal{M}_\alpha$ for a small example.

[1] By condition (5.3b) we have avoided the empty set $\alpha = \emptyset$ and $\alpha = D$ ($\Rightarrow \alpha^c = \emptyset$). Since the empty tensor product is interpreted as the field $\mathbb{K}$, one may view $\mathcal{M}_D : \mathbf{V} \to \mathbf{V} \otimes \mathbb{K}$ as the vectorisation (column vector) and $\mathcal{M}_\emptyset : \mathbf{V} \to \mathbb{K} \otimes \mathbf{V}$ as mapping into a row vector.

[2] In the case of Banach tensor spaces, the isomorphism must also be *isometric*.

[3] This means that $\mathcal{M}_\alpha$ is replaced by $\Xi^{-1} \circ \mathcal{M}_\alpha$, where Ξ is the isomorphism from the matrix space $\mathbb{K}^{\mathbf{I}_\alpha \times \mathbf{I}_{\alpha^c}}$ onto the tensor space $\mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$ (see Proposition 3.14). For simplicity, we write $\mathcal{M}_\alpha$ instead of $\Xi^{-1} \circ \mathcal{M}_\alpha$.

Example 5.5. Below, all matricisations are given for the tensor

$$\mathbf{v} \in \mathbb{K}^{I_1} \otimes \mathbb{K}^{I_2} \otimes \mathbb{K}^{I_3} \otimes \mathbb{K}^{I_4} \quad \text{with } I_1 = I_2 = I_3 = I_4 = \{1, 2\}.$$

The matrix $\mathcal{M}_1(\mathbf{v})$ belongs to $\mathbb{K}^{I_1 \times J}$ with $J = I_2 \times I_3 \times I_4$. For the sake of the following notation we introduce the lexicographical ordering of the triples from $I_2 \times I_3 \times I_4 : (1,1,1), (1,1,2), (1,2,1), \ldots, (2,2,2)$. Under these assumptions, $\mathbb{K}^{I_1 \times J}$ becomes $\mathbb{K}^{2\times 8}$:[4]

$$\mathcal{M}_1(\mathbf{v}) = \begin{pmatrix} \mathbf{v}_{\mathbf{1}111} & \mathbf{v}_{\mathbf{1}112} & \mathbf{v}_{\mathbf{1}121} & \mathbf{v}_{\mathbf{1}122} & \mathbf{v}_{\mathbf{1}211} & \mathbf{v}_{\mathbf{1}212} & \mathbf{v}_{\mathbf{1}221} & \mathbf{v}_{\mathbf{1}222} \\ \mathbf{v}_{\mathbf{2}111} & \mathbf{v}_{\mathbf{2}112} & \mathbf{v}_{\mathbf{2}121} & \mathbf{v}_{\mathbf{2}122} & \mathbf{v}_{\mathbf{2}211} & \mathbf{v}_{\mathbf{2}212} & \mathbf{v}_{\mathbf{2}221} & \mathbf{v}_{\mathbf{2}222} \end{pmatrix}.$$

$\mathcal{M}_2(\mathbf{v})$ belongs to $\mathbb{K}^{I_2 \times J}$ with $J = I_1 \times I_3 \times I_4$. Together with the lexicographical ordering in J we get

$$\mathcal{M}_2(\mathbf{v}) = \begin{pmatrix} \mathbf{v}_{1\mathbf{1}11} & \mathbf{v}_{1\mathbf{1}12} & \mathbf{v}_{1\mathbf{1}21} & \mathbf{v}_{1\mathbf{1}22} & \mathbf{v}_{2\mathbf{1}11} & \mathbf{v}_{2\mathbf{1}12} & \mathbf{v}_{2\mathbf{1}21} & \mathbf{v}_{2\mathbf{1}22} \\ \mathbf{v}_{1\mathbf{2}11} & \mathbf{v}_{1\mathbf{2}12} & \mathbf{v}_{1\mathbf{2}21} & \mathbf{v}_{1\mathbf{2}22} & \mathbf{v}_{2\mathbf{2}11} & \mathbf{v}_{2\mathbf{2}12} & \mathbf{v}_{2\mathbf{2}21} & \mathbf{v}_{2\mathbf{2}22} \end{pmatrix}.$$

Similarly,

$$\mathcal{M}_3(\mathbf{v}) = \begin{pmatrix} \mathbf{v}_{11\mathbf{1}1} & \mathbf{v}_{11\mathbf{1}2} & \mathbf{v}_{12\mathbf{1}1} & \mathbf{v}_{12\mathbf{1}2} & \mathbf{v}_{21\mathbf{1}1} & \mathbf{v}_{21\mathbf{1}2} & \mathbf{v}_{22\mathbf{1}1} & \mathbf{v}_{22\mathbf{1}2} \\ \mathbf{v}_{11\mathbf{2}1} & \mathbf{v}_{11\mathbf{2}2} & \mathbf{v}_{12\mathbf{2}1} & \mathbf{v}_{12\mathbf{2}2} & \mathbf{v}_{21\mathbf{2}1} & \mathbf{v}_{21\mathbf{2}2} & \mathbf{v}_{22\mathbf{2}1} & \mathbf{v}_{22\mathbf{2}2} \end{pmatrix},$$

$$\mathcal{M}_4(\mathbf{v}) = \begin{pmatrix} \mathbf{v}_{111\mathbf{1}} & \mathbf{v}_{112\mathbf{1}} & \mathbf{v}_{121\mathbf{1}} & \mathbf{v}_{122\mathbf{1}} & \mathbf{v}_{211\mathbf{1}} & \mathbf{v}_{212\mathbf{1}} & \mathbf{v}_{221\mathbf{1}} & \mathbf{v}_{222\mathbf{1}} \\ \mathbf{v}_{111\mathbf{2}} & \mathbf{v}_{112\mathbf{2}} & \mathbf{v}_{121\mathbf{2}} & \mathbf{v}_{122\mathbf{2}} & \mathbf{v}_{211\mathbf{2}} & \mathbf{v}_{212\mathbf{2}} & \mathbf{v}_{221\mathbf{2}} & \mathbf{v}_{222\mathbf{2}} \end{pmatrix}.$$

Next, we consider $\alpha = \{1, 2\}$. $\mathcal{M}_{\{1,2\}}(\mathbf{v})$ belongs to $\mathbb{K}^{I \times J}$ with $I = I_1 \times I_2$ and $J = I_3 \times I_4$. Lexicographical ordering of I and J yields a matrix from $\mathbb{K}^{4\times 4}$:

$$\mathcal{M}_{\{1,2\}}(\mathbf{v}) = \begin{pmatrix} \mathbf{v}_{\mathbf{11}11} & \mathbf{v}_{\mathbf{11}12} & \mathbf{v}_{\mathbf{11}21} & \mathbf{v}_{\mathbf{11}22} \\ \mathbf{v}_{\mathbf{12}11} & \mathbf{v}_{\mathbf{12}12} & \mathbf{v}_{\mathbf{12}21} & \mathbf{v}_{\mathbf{12}22} \\ \mathbf{v}_{\mathbf{21}11} & \mathbf{v}_{\mathbf{21}12} & \mathbf{v}_{\mathbf{21}21} & \mathbf{v}_{\mathbf{21}22} \\ \mathbf{v}_{\mathbf{22}11} & \mathbf{v}_{\mathbf{22}12} & \mathbf{v}_{\mathbf{22}21} & \mathbf{v}_{\mathbf{22}22} \end{pmatrix}.$$

Similarly,

$$\mathcal{M}_{\{1,3\}}(\mathbf{v}) = \begin{pmatrix} \mathbf{v}_{\mathbf{1}1\mathbf{1}1} & \mathbf{v}_{\mathbf{1}1\mathbf{1}2} & \mathbf{v}_{\mathbf{1}1\mathbf{2}1} & \mathbf{v}_{\mathbf{1}2\mathbf{1}2} \\ \mathbf{v}_{\mathbf{1}1\mathbf{2}1} & \mathbf{v}_{\mathbf{1}1\mathbf{2}2} & \mathbf{v}_{\mathbf{1}2\mathbf{2}1} & \mathbf{v}_{\mathbf{1}2\mathbf{2}2} \\ \mathbf{v}_{\mathbf{2}1\mathbf{1}1} & \mathbf{v}_{\mathbf{2}1\mathbf{1}2} & \mathbf{v}_{\mathbf{2}2\mathbf{1}1} & \mathbf{v}_{\mathbf{2}2\mathbf{1}2} \\ \mathbf{v}_{\mathbf{2}1\mathbf{2}1} & \mathbf{v}_{\mathbf{2}1\mathbf{2}2} & \mathbf{v}_{\mathbf{2}2\mathbf{2}1} & \mathbf{v}_{\mathbf{2}2\mathbf{2}2} \end{pmatrix},$$

$$\mathcal{M}_{\{1,4\}}(\mathbf{v}) = \begin{pmatrix} \mathbf{v}_{\mathbf{1}11\mathbf{1}} & \mathbf{v}_{\mathbf{1}12\mathbf{1}} & \mathbf{v}_{\mathbf{1}11\mathbf{2}} & \mathbf{v}_{\mathbf{1}22\mathbf{1}} \\ \mathbf{v}_{\mathbf{1}11\mathbf{2}} & \mathbf{v}_{\mathbf{1}12\mathbf{2}} & \mathbf{v}_{\mathbf{1}21\mathbf{2}} & \mathbf{v}_{\mathbf{1}22\mathbf{2}} \\ \mathbf{v}_{\mathbf{2}11\mathbf{1}} & \mathbf{v}_{\mathbf{2}12\mathbf{1}} & \mathbf{v}_{\mathbf{2}21\mathbf{1}} & \mathbf{v}_{\mathbf{2}22\mathbf{1}} \\ \mathbf{v}_{\mathbf{2}11\mathbf{2}} & \mathbf{v}_{\mathbf{2}12\mathbf{2}} & \mathbf{v}_{\mathbf{2}21\mathbf{2}} & \mathbf{v}_{\mathbf{2}22\mathbf{2}} \end{pmatrix}.$$

The further $\mathcal{M}_\alpha(\mathbf{v})$ are transposed versions of the already described matrices: $\mathcal{M}_{\{2,3\}} = \mathcal{M}_{\{1,4\}}^\mathsf{T}$, $\mathcal{M}_{\{2,4\}} = \mathcal{M}_{\{1,3\}}^\mathsf{T}$, $\mathcal{M}_{\{3,4\}} = \mathcal{M}_{\{1,2\}}^\mathsf{T}$, $\mathcal{M}_{\{1,2,3\}} = \mathcal{M}_4^\mathsf{T}$, $\mathcal{M}_{\{1,2,4\}} = \mathcal{M}_3^\mathsf{T}$, $\mathcal{M}_{\{1,3,4\}} = \mathcal{M}_2^\mathsf{T}$, $\mathcal{M}_{\{2,3,4\}} = \mathcal{M}_2^\mathsf{T}$.

[4] Bold face indices correspond to the row numbers.

5.2.2.2 Invariant Properties and α-Rank

The interpretation of tensors $\mathbf{v}$ as matrices M enables us

(i) to transfer the matrix terminology from M to $\mathbf{v}$,
(ii) to apply all matrix techniques to M.

In Remark 3.15 we have considered an isomorphism $\mathbf{v} \cong M$ and stated that the multiplication of a tensor by a Kronecker product $A \otimes B$ has the isomorphic expression $(A \otimes B)\,\mathbf{v} \cong AMB^{\mathsf{T}}$. More generally, the following statement holds, which is the matrix interpretation of Remark 5.4.

Lemma 5.6. *Let* $\mathbf{v} \in \mathbf{V} = \bigotimes_{j\in D} \mathbb{K}^{I_j}$ *and* $\mathbf{A} = \bigotimes_{j\in D} A^{(j)} \in \bigotimes_{j\in D} L(\mathbb{K}^{I_j}, \mathbb{K}^{J_j})$. *The product* $\mathbf{Av} \in \mathbf{W} = \bigotimes_{j\in D} \mathbb{K}^{J_j}$ *satisfies*

$$\mathcal{M}_\alpha(\mathbf{Av}) = \mathbf{A}^{(\alpha)} \mathcal{M}_\alpha(\mathbf{v}) \mathbf{A}^{(\alpha^c)\mathsf{T}} \textit{ with } \mathbf{A}^{(\alpha)} = \bigotimes_{j\in\alpha} A^{(j)}, \; \mathbf{A}^{(\alpha^c)} = \bigotimes_{j\in\alpha^c} A^{(j)}. \tag{5.5}$$

In particular, if all $A^{(j)}$ *are regular matrices, the matrix ranks of* $\mathcal{M}_\alpha(\mathbf{Av})$ *and* $\mathcal{M}_\alpha(\mathbf{v})$ *coincide.*

Proof. Define the index sets $\mathbf{I}_\alpha := \times_{j\in\alpha} I_j$ and $\mathbf{I}_{\alpha^c} := \times_{j\in\alpha^c} I_j$, and similarly $\mathbf{J}_\alpha$ and $\mathbf{J}_{\alpha^c}$. In the following, the indices $\mathbf{i} \in \mathbf{I} := \times_{j\in D} I_j$ are written as $(\mathbf{i}', \mathbf{i}'')$ with $\mathbf{i}' \in \mathbf{I}_\alpha$ and $\mathbf{i}'' \in \mathbf{I}_{\alpha^c}$. Similarly for $\mathbf{j} = (\mathbf{j}', \mathbf{j}'') \in \mathbf{J}$. Note that $\bullet_{\mathbf{j}',\mathbf{j}''}$ denotes a matrix entry, while $\bullet_{\mathbf{j}} = \bullet_{(\mathbf{j}',\mathbf{j}'')}$ is a tensor entry. The identity

$$\begin{aligned}\mathcal{M}_\alpha(\mathbf{Av})_{\mathbf{j}',\mathbf{j}''} &= (\mathbf{Av})_{(\mathbf{j}',\mathbf{j}'')} = \sum_{\mathbf{i}\in\mathbf{I}} \mathbf{A}_{(\mathbf{j}',\mathbf{j}''),\mathbf{i}} \mathbf{v}_{\mathbf{i}} = \sum_{\mathbf{i}'\in\mathbf{I}_\alpha} \sum_{\mathbf{i}''\in\mathbf{I}_{\alpha^c}} \mathbf{A}_{(\mathbf{j}',\mathbf{j}''),(\mathbf{i}',\mathbf{i}'')} \mathbf{v}_{(\mathbf{i}',\mathbf{i}'')} \\ &= \sum_{\mathbf{i}'\in\mathbf{I}_\alpha} \sum_{\mathbf{i}''\in\mathbf{I}_{\alpha^c}} \mathbf{A}^{(\alpha)}_{\mathbf{j}',\mathbf{i}'} \mathbf{v}_{(\mathbf{i}',\mathbf{i}'')} \mathbf{A}^{(\alpha^c)}_{\mathbf{j}'',\mathbf{i}''} = \sum_{\mathbf{i}'\in\mathbf{I}_\alpha} \sum_{\mathbf{i}''\in\mathbf{I}_{\alpha^c}} \mathbf{A}^{(\alpha)}_{\mathbf{j}',\mathbf{i}'} \mathcal{M}_\alpha(\mathbf{v})_{\mathbf{i}',\mathbf{i}''} \mathbf{A}^{(\alpha^c)}_{\mathbf{j}'',\mathbf{i}''}\end{aligned}$$

proves (5.5). □

According to item (i), we may define the matrix rank of $\mathcal{M}_\alpha(\mathbf{v})$ as a property of $\mathbf{v}$. By Lemma 5.6, the rank of $\mathcal{M}_\alpha(\mathbf{v})$ is invariant under tensor space isomorphisms.

Definition 5.7 (rank_α)**.** For any[5] $\alpha \subset D$ from (5.3b) and all $j \in D$ we define

$$\mathrm{rank}_\alpha(\mathbf{v}) := \mathrm{rank}\left(\mathcal{M}_\alpha(\mathbf{v})\right), \tag{5.6a}$$
$$\mathrm{rank}_j(\mathbf{v}) := \mathrm{rank}_{\{j\}}(\mathbf{v}) = \mathrm{rank}\left(\mathcal{M}_j(\mathbf{v})\right). \tag{5.6b}$$

In 1927, Hitchcock [100, p. 170] has introduced $\mathrm{rank}_j(\mathbf{v})$ as 'the rank on the j^{th} index'. Also $\mathrm{rank}_\alpha(\mathbf{v})$ is defined by him as the 'α-plex rank'. We shall call it j-rank or α-rank, respectively. Further properties of the α-rank will follow in Lemma 6.19 and Corollary 6.20.

[5] Usually, we avoid $\alpha = \emptyset$ and $\alpha = D$. Formally, the definition of $\mathcal{M}_\emptyset$, $\mathcal{M}_D$ from Footnote 1 yields $\mathrm{rank}_\emptyset(\mathbf{v}) = \mathrm{rank}_D(\mathbf{v}) = 1$ for $\mathbf{v} \neq 0$ and $\mathrm{rank}_\emptyset(0) = \mathrm{rank}_D(0) = 0$, otherwise.

The tensor rank defined in (3.24) is not directly related to the family of ranks $\{\mathrm{rank}_\alpha(\mathbf{v}) : \emptyset \subsetneqq \alpha \subsetneqq D\}$ from (5.6a) or the ranks $\{\mathrm{rank}_j(\mathbf{v}) : j \in D\}$ from (5.6b). Later, in Remark 6.21, we shall prove

$$\mathrm{rank}_\alpha(\mathbf{v}) \leq \mathrm{rank}(\mathbf{v}) \qquad \text{for all } \alpha \subset D.$$

Remark 5.8. Let $\|\cdot\|$ be the Euclidean norm in $\mathbf{V} = \bigotimes_{j=1}^d \mathbb{K}^{I_j}$ (cf. Example 4.126), while $\|\cdot\|_\mathsf{F}$ is the Frobenius norm for matrices (cf. (2.9)). Then, the norms coincide:

$$\|\mathbf{v}\| = \|\mathcal{M}_\alpha(\mathbf{v})\|_\mathsf{F} \qquad \text{for all } \emptyset \subsetneqq \alpha \subsetneqq \{1,\ldots,d\} \text{ and all } \mathbf{v} \in \mathbf{V}.$$

Proof. Use Remark 2.8. □

5.2.2.3 Singular Value Decomposition

Later, singular values of $\mathcal{M}_\alpha(\mathbf{v})$ will be important. The next proposition compares the singular values of $\mathcal{M}_\alpha(\mathbf{v})$ and $\mathcal{M}_\alpha(\mathbf{A}\mathbf{v})$.

Proposition 5.9. *Let $\mathbf{v} \in \mathbf{V} = \bigotimes_{\kappa\in D} \mathbb{K}^{I_\kappa}$. The Kronecker matrix $\mathbf{A} = \mathbf{A}^{(\alpha)} \otimes \mathbf{A}^{(\alpha^c)}$ is assumed to be composed of $\mathbf{A}^{(\alpha)} \in L(\mathbf{V}_\alpha, \mathbf{V}_\alpha)$ and $\mathbf{A}^{(\alpha^c)} \in L(\mathbf{V}_{\alpha^c}, \mathbf{V}_{\alpha^c})$ with the properties $\mathbf{A}^{(\alpha)\mathsf{H}}\mathbf{A}^{(\alpha)} \leq \mathbf{I}$ and $\mathbf{A}^{(\alpha^c)\mathsf{H}}\mathbf{A}^{(\alpha^c)} \leq \mathbf{I}$. Then the singular values fulfil*

$$\sigma_k(\mathcal{M}_\alpha(\mathbf{A}\mathbf{v})) \leq \sigma_k(\mathcal{M}_\alpha(\mathbf{v})) \qquad \textit{for all } k \in \mathbb{N}.$$

Proof. Combine $\mathcal{M}_\alpha(\mathbf{A}\mathbf{v}) = \mathbf{A}^{(\alpha)}\mathcal{M}_\alpha(\mathbf{v})\mathbf{A}^{(\alpha^c)\mathsf{T}}$ from (5.5) and Lemma 2.27c. □

Corollary 5.10. The assumptions of Proposition 5.9 are in particular satisfied, if $\mathbf{A}^{(\alpha)}$ and $\mathbf{A}^{(\alpha^c)}$ are orthogonal projections.

Remark 5.11. The reduced singular value decomposition

$$\mathcal{M}_\alpha(\mathbf{v}) = U\Sigma V^\mathsf{T} = \sum_{i=1}^{r_\alpha} \sigma_i^{(\alpha)} u_i v_i^\mathsf{T}$$

($\sigma_i > 0,\ u_i, v_i$ columns of U and V, $r_\alpha = \mathrm{rank}_\alpha(\mathbf{v})$) translates into

$$\mathbf{v} = \sum_{i=1}^{r_\alpha} \sigma_i^{(\alpha)}\, \mathbf{u}_i \otimes \mathbf{v}_i \qquad (\mathbf{u}_i \in \mathbf{V}_\alpha,\ \mathbf{v}_i \in \mathbf{V}_{\alpha^c}). \tag{5.7}$$

Here, u_i and v_i are the isomorphic vector interpretations of the tensors $\mathbf{u}_i$, $\mathbf{v}_i$.

Remark 5.12. (a) In the case of matrices (i.e., $D = \{1,2\}$), the ranks are equal:[6] $\mathrm{rank}_1(\mathbf{v}) = \mathrm{rank}_2(\mathbf{v})$. Further, the singular values of $\mathcal{M}_\alpha(\mathbf{v})$ coincide: $\sigma_i^{(1)} = \sigma_i^{(2)}$.
(b) If $d \geq 3$, the values $\mathrm{rank}_k(\mathbf{v})$ $(k \in D)$ may not coincide. Furthermore, the

[6] The true generalisation of this property for general d is Eq. (6.17a).

singular values $\sigma_i^{(k)}$ of $\mathcal{M}_k(\mathbf{v})$ may be different; however, the following quantity is invariant:

$$\sum_{i=1}^{\operatorname{rank}_k(\mathbf{v})} \left(\sigma_i^{(k)}\right)^2 = \|\mathbf{v}\|_2^2 \quad \text{for all } k \in D \qquad (\|\cdot\|_2 \text{ from (4.126)}).$$

Proof. 1) Let $\mathbf{v}_{(i,j)} = M_{i,j}$ be the isomorphism between $\mathbf{v} \in \mathbb{K}^{\mathbf{I}} = \mathbb{K}^{I_1} \otimes \mathbb{K}^{I_2}$ and the matrix $M \in \mathbb{K}^{I_1 \times I_2}$. Then $\mathcal{M}_1(\mathbf{v}) = M$, while $\mathcal{M}_2(\mathbf{v}) = M^{\mathsf{T}}$. Since M and M^{T} have identical rank and identical singular values, Part (a) follows.

2) Consider the tensor $\mathbf{v} = a_1 \otimes a_2 \otimes a_3 + a_1 \otimes b_2 \otimes b_3 \in \bigotimes_{j=1}^3 \mathbb{K}^2$ with $a_i = \binom{1}{0}$ $(i = 1, 2, 3)$ and $b_i = \binom{0}{1}$ $(i = 2, 3)$. We have

$$\mathcal{M}_1(\mathbf{v}) = a_1 \otimes \mathbf{c} \in \mathbb{K}^2 \otimes \left(\mathbb{K}^2 \otimes \mathbb{K}^2\right) \cong \mathbb{K}^2 \otimes \mathbb{K}^4$$

with $\mathbf{c} := a_2 \otimes a_3 + b_2 \otimes b_3 \cong (1\,0\,0\,1) \in \mathbb{K}^4$ (i.e., $\mathcal{M}_1(\mathbf{v}) \cong \left(\begin{smallmatrix} 1\,0\,0\,1 \\ 0\,0\,0\,0 \end{smallmatrix}\right)$ has rank 1), whereas

$$\mathcal{M}_2(\mathbf{v}) = a_2 \otimes \mathbf{c} + b_2 \otimes \mathbf{d} \cong \begin{pmatrix} c \\ d \end{pmatrix} \quad \text{with} \begin{cases} \mathbf{c} := a_1 \otimes a_3 \cong c := (1\,0\,0\,0), \\ \mathbf{d} := a_1 \otimes b_3 \cong d := (0\,1\,0\,0) \end{cases}$$

has two linearly independent rows and therefore rank 2.

3) Since the rank is also the number of positive singular values, they must be different for the given example. The sum of the squared singular values is the squared Frobenius norm of the corresponding matrix: $\|\mathcal{M}_k(\mathbf{v})\|_{\mathsf{F}}^2$. Since the matrix entries of $\mathcal{M}_k(\mathbf{v})$ are only a permutation of the entries of $\mathbf{v}$ (cf. Example 5.5), the sum of their squares equals $\|\mathbf{v}\|_2^2$. □

5.2.2.4 Infinite Dimensional Spaces

For infinite dimensional vector spaces V_j, these quantities generalise as follows. In the finite dimensional case, $\operatorname{rank}(\mathcal{M}_\alpha(\mathbf{v}))$ is equal to the dimension of the range of $\mathcal{M}_\alpha(\mathbf{v})$ (cf. Remark 2.1), where

$$\operatorname{range}(\mathcal{M}_\alpha(\mathbf{v})) = \{\mathcal{M}_\alpha(\mathbf{v})\mathbf{z} : \mathbf{z} \in \mathbf{V}_{\alpha^c}\}.$$

Since $\mathcal{M}_\alpha(\mathbf{v}) \in \mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$ has the form $\sum_\nu \mathbf{x}_\nu \otimes \mathbf{y}_\nu$ (cf. Definition 5.3), the matrix-vector multiplication $\mathcal{M}_\alpha(\mathbf{v})\mathbf{z}$ can be considered as $\sum_\nu \mathbf{z}(\mathbf{y}_\nu) \cdot \mathbf{x}_\nu \in \mathbf{V}_\alpha$, where $\mathbf{z} \in V'_{\alpha^c}$ is considered as an element of the dual vector space (for $\dim(\mathbf{V}_{\alpha^c}) < \infty$, $\mathbf{V}'_{\alpha^c}$ may be identified with $\mathbf{V}_{\alpha^c}$). The mapping $\sum_\nu \mathbf{x}_\nu \otimes \mathbf{y}_\nu \mapsto \sum_\nu \mathbf{z}(\mathbf{y}_\nu) \cdot \mathbf{x}_\nu$ is denoted by $id \otimes \mathbf{z}$. Then the matrix-vector multiplication $\mathcal{M}_\alpha(\mathbf{v})\mathbf{z}$ may be rewritten as $(id \otimes \mathbf{z})\, \mathcal{M}_\alpha(\mathbf{v})$. This leads to the notation

$$\operatorname{rank}_\alpha(\mathbf{v}) := \dim \left\{(id \otimes \mathbf{z})\, \mathcal{M}_\alpha(\mathbf{v}) : \mathbf{z} \in \mathbf{V}'_{\alpha^c}\right\}.$$

The transition to the dual space $\mathbf{V}'_{\alpha^c}$ (or $\mathbf{V}^*_{\alpha^c}$) is necessary, since, in the infinite dimensional case, $\mathcal{M}_\alpha(\mathbf{v})$ cannot be interpreted as a mapping from $\mathbf{V}_{\alpha^c}$ into $\mathbf{V}_\alpha$, but as a mapping from $\mathbf{V}'_{\alpha^c}$ into $\mathbf{V}_\alpha$. The set $\{(id \otimes \mathbf{z})\, \mathcal{M}_\alpha(\mathbf{v}) : \mathbf{z} \in \mathbf{V}'_{\alpha^c}\}$ on the right-hand side will be defined in §6 as the minimal subspace $\mathbf{U}_\alpha^{\min}(\mathbf{v})$, so that

$$\operatorname{rank}_\alpha(\mathbf{v}) := \dim(\mathbf{U}_\alpha^{\min}(\mathbf{v})) \tag{5.8}$$

is the generalisation to infinite dimensional (algebraic) vector spaces as well as to Banach spaces.

An identification of $\mathbf{V}^*_{\alpha^c}$ with $\mathbf{V}_{\alpha^c}$ becomes possible, if all V_j are Hilbert spaces. This case is discussed next.

5.2.3 Hilbert Structure

Next, we consider a pre-Hilbert space $\mathbf{V} = {}_a\bigotimes_{j\in D} V_j$ and the left-sided singular value decomposition problem of $\mathcal{M}_\alpha(\mathbf{v})$ for some $\mathbf{v} \in \mathbf{V}$. The standard singular value decomposition is

$$\mathcal{M}_\alpha(\mathbf{v}) = \sum_{i=1}^{r} \sigma_i^{(\alpha)}\, \mathbf{u}_i \otimes \mathbf{v}_i \quad \Bigg(\mathbf{u}_i \in \mathbf{V}_\alpha := {}_a\bigotimes_{j\in\alpha} V_j\,,\ \mathbf{v}_i \in \mathbf{V}_{\alpha^c} := {}_a\bigotimes_{j\in\alpha^c} V_j \Bigg) \tag{5.9}$$

with two orthonormal families $\{\mathbf{u}_i\}$, $\{\mathbf{v}_i\}$, and $\sigma_1^{(\alpha)} \geq \ldots \geq \sigma_r^{(\alpha)} > 0$. The left-sided singular value decomposition problem asks for $\{\mathbf{u}_i\}$ and $\{\sigma_i^{(\alpha)}\}$. If $V_j = \mathbb{K}^{I_j}$ allows us to interpret $\mathcal{M}_\alpha(\mathbf{v})$ as a matrix, the data $\{u_i\}$, $\{\sigma_i^{(\alpha)}\}$ are determined by $\mathbf{LSVD}(\mathbf{I}_\alpha, \mathbf{I}_{\alpha^c}, r, \mathcal{M}_\alpha(\mathbf{v}), U, \Sigma)$ (cf. (2.32)). We recall that its computation may use the diagonalisation of $\mathcal{M}_\alpha(\mathbf{v})\mathcal{M}_\alpha(\mathbf{v})^{\mathsf{H}} = \sum_{i=1}^{r} (\sigma_i^{(\alpha)})^2\, u_i u_i^{\mathsf{H}}$. In the infinite dimensional setting, the latter expression can be expressed by the partial scalar product[7] from §4.5.4:

$$\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v}) \rangle_{\alpha^c} = \langle \mathbf{v}, \mathbf{v} \rangle_{\alpha^c} \in \mathbf{V}_\alpha \otimes \mathbf{V}_\alpha$$

Assuming the singular value decomposition $\mathcal{M}_\alpha(\mathbf{v}) = \sum_{i=1}^{r} \sigma_i^{(\alpha)}\, \mathbf{u}_i \otimes \mathbf{v}_i$ (possibly with $r = \infty$, cf. (4.16)), the partial scalar product yields the diagonalisation

$$\begin{aligned}\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v}) \rangle_{\alpha^c} &= \left\langle \sum_{i=1}^{r} \sigma_i^{(\alpha)}\, \mathbf{u}_i \otimes \mathbf{v}_i, \sum_{j=1}^{r} \sigma_j^{(\alpha)}\, \mathbf{u}_j \otimes \mathbf{v}_j \right\rangle_{\alpha^c} \\ &= \sum_{i=1}^{r} \sum_{j=1}^{r} \sigma_i^{(\alpha)} \sigma_j^{(\alpha)} \underbrace{\langle \mathbf{v}_i, \mathbf{v}_j \rangle_{\alpha^c}}_{=\delta_{ij}} \mathbf{u}_i \otimes \overline{\mathbf{u}_j} = \sum_{i=1}^{r} (\sigma_i^{(\alpha)})^2\, \mathbf{u}_i \otimes \overline{\mathbf{u}_i}.\end{aligned}$$

[7] If the image $\mathcal{M}_\alpha(\mathbf{v}) = \mathbf{v}^{(\alpha)} \otimes \mathbf{v}^{(\alpha^c)}$ under the isomorphism $\mathcal{M}_\alpha : \mathbf{V} \to \mathbf{V}^{(\alpha)} \otimes \mathbf{V}^{(\alpha^c)}$ is an elementary tensors, the partial scalar product is defined by $\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v}) \rangle_{\alpha^c} = \left\langle \mathbf{v}^{(\alpha^c)}, \mathbf{v}^{(\alpha^c)} \right\rangle_{\alpha^c} \cdot \mathbf{v}^{(\alpha)} \otimes \overline{\mathbf{v}^{(\alpha)}} \in \mathbf{V}_\alpha \otimes \mathbf{V}_\alpha$. The expression $\langle \mathbf{v}, \mathbf{v} \rangle_{\alpha^c}$ has the same meaning.

We summarise.

Lemma 5.13. *Let* $\mathbf{V} = {}_a\bigotimes_{j\in D} V_j \cong \mathbf{V}_\alpha \otimes_a \mathbf{V}_{\alpha^c}$ *with* $\mathbf{V}_\alpha, \mathbf{V}_{\alpha^c}$ *as in (5.9). The left singular vectors* $\mathbf{u}_i^{(\alpha)} \in \mathbf{V}_\alpha$ *and singular values* $\sigma_i^{(\alpha)}$ *are obtainable from the diagonalisation*

$$\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha^c} = \sum_{i=1}^{r} \left(\sigma_i^{(\alpha)}\right)^2 \mathbf{u}_i^{(\alpha)} \otimes \overline{\mathbf{u}_i^{(\alpha)}}. \tag{5.10a}$$

Analogously, the right singular vectors $\mathbf{v}_i^{(\alpha)} \in \mathbf{V}_{\alpha^c}$ *and singular values* $\sigma_i^{(\alpha)}$ *are obtainable from the diagonalisation*

$$\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha} = \sum_{i=1}^{r} \left(\sigma_i^{(\alpha)}\right)^2 \mathbf{v}_i^{(\alpha)} \otimes \overline{\mathbf{v}_i^{(\alpha)}}. \tag{5.10b}$$

Corollary 4.131 allows us to determine the partial scalar product $\langle\cdot,\cdot\rangle_\beta$ from $\langle\cdot,\cdot\rangle_\alpha$ if $\alpha \subsetneqq \beta \subset D$. As a consequence, $\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha^c}$ can be obtained from $\langle \mathcal{M}_\beta(\mathbf{v}), \mathcal{M}_\beta(\mathbf{v})\rangle_{\beta^c}$:

$$\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha^c} = \mathfrak{C}_{\beta\backslash\alpha}\left(\langle \mathcal{M}_\beta(\mathbf{v}), \mathcal{M}_\beta(\mathbf{v})\rangle_{\beta^c}\right) \tag{5.11}$$

with the contraction $\mathfrak{C}_{\beta\backslash\alpha}$ from Definition 4.130. In order to apply $\mathfrak{C}_{\beta\backslash\alpha}$, the tensor $\langle \mathcal{M}_\beta(\mathbf{v}), \mathcal{M}_\beta(\mathbf{v})\rangle_{\beta^c} \in \mathbf{V}_\beta \otimes \mathbf{V}_\beta$ is interpreted as $\mathbf{V}_\alpha \otimes \mathbf{V}_{\beta\backslash\alpha} \otimes \mathbf{V}_\alpha \otimes \mathbf{V}_{\beta\backslash\alpha}$. Using basis representations, we obtain the following result.

Theorem 5.14. *Assume* $\emptyset \subsetneqq \alpha_1 \subsetneqq \alpha \subset D$ *and*

$$\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha^c} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)}\, \mathbf{b}_i^{(\alpha)} \otimes \overline{\mathbf{b}_j^{(\alpha)}} \in \mathbf{V}_\alpha \otimes \mathbf{V}_\alpha. \tag{5.12a}$$

Set $\alpha_2 := \alpha\backslash\alpha_1$. *Then* $\alpha = \alpha_1 \dot\cup \alpha_2$ *holds. Consider* $\mathbf{b}_i^{(\alpha)} \in \mathbf{V}_\alpha$ *as elements of* $\mathbf{V}_{\alpha_1} \otimes \mathbf{V}_{\alpha_2}$ *with the representation*

$$\mathbf{b}_i^{(\alpha)} = \sum_{\nu=1}^{r_{\alpha_1}} \sum_{\mu=1}^{r_{\alpha_2}} c_{\nu\mu}^{(i)}\, \mathbf{b}_\nu^{(\alpha_1)} \otimes \mathbf{b}_\mu^{(\alpha_2)}. \tag{5.12b}$$

We introduce the matrices $C_i := \left(c_{\nu\mu}^{(i)}\right) \in \mathbb{K}^{r_{\alpha_1}\times r_{\alpha_2}}$ *for* $1 \le i \le r_\alpha$. *Then*

$$\langle \mathcal{M}_{\alpha_k}(\mathbf{v}), \mathcal{M}_{\alpha_k}(\mathbf{v})\rangle_{\alpha_k^c} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha_k)}\, \mathbf{b}_i^{(\alpha_k)} \otimes \overline{\mathbf{b}_j^{(\alpha_k)}} \in \mathbf{V}_{\alpha_k} \otimes \mathbf{V}_{\alpha_k} \qquad (k = 1, 2)$$

holds with coefficient matrices $E_{\alpha_k} = \left(e_{ij}^{(\alpha_k)}\right) \in \mathbb{K}^{r_{\alpha_k}\times r_{\alpha_k}}$ *defined by*

$$E_{\alpha_1} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)} C_i G_{\alpha_2}^{\mathsf{T}} C_j^{\mathsf{H}}, \quad E_{\alpha_2} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)} C_i^{\mathsf{T}} G_{\alpha_1}^{\mathsf{T}} \overline{C_j}, \tag{5.12c}$$

where $G_{\alpha_k} = \left(g_{\nu\mu}^{(\alpha_k)}\right)$ *is the Gram matrix with entries* $g_{\nu\mu}^{(\alpha_k)} = \left\langle \mathbf{b}_\mu^{(\alpha_k)}, \mathbf{b}_\nu^{(\alpha_k)}\right\rangle$.

Proof. Insertion of (5.12b) in (5.12a) yields

$$\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha^c} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)} \sum_{\nu,\mu,\sigma,\tau} c_{\nu\mu}^{(i)} \overline{c_{\sigma\tau}^{(j)}}\, \mathbf{b}_\nu^{(\alpha_1)} \otimes \mathbf{b}_\mu^{(\alpha_2)} \otimes \overline{\mathbf{b}_\sigma^{(\alpha_1)}} \otimes \overline{\mathbf{b}_\tau^{(\alpha_2)}}.$$

Applying (5.11) with $\beta\backslash\alpha$ replaced by $\alpha_2 = \alpha\backslash\alpha_1$ yields

$$\langle \mathcal{M}_{\alpha_1}(\mathbf{v}), \mathcal{M}_{\alpha_1}(\mathbf{v})\rangle_{\alpha_1^c} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)} \sum_{\nu,\mu,\sigma,\tau} c_{\nu\mu}^{(i)} \overline{c_{\sigma\tau}^{(j)}} \left\langle \mathbf{b}_\mu^{(\alpha_2)}, \mathbf{b}_\tau^{(\alpha_2)} \right\rangle \mathbf{b}_\nu^{(\alpha_1)} \otimes \overline{\mathbf{b}_\sigma^{(\alpha_1)}}$$

proving $e_{\nu\sigma}^{(\alpha_1)} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)} \sum_{\mu,\tau} c_{\nu\mu}^{(i)} \overline{c_{\sigma\tau}^{(j)}}\, g_{\tau\mu}^{(\alpha_2)}$, i.e., $E_{\alpha_1} = \sum_{i,j=1} e_{ij}^{(\alpha)} C_i G_{\alpha_2}^{\mathsf{T}} C_j^{\mathsf{H}}$. The case of E_{α_2} is analogous. □

Corollary 5.15. Assume the finite dimensional case with orthonormal basis $\{\mathbf{b}_i^{(\alpha)}\}$. Form the matrix $\mathbf{B}_\alpha = [\mathbf{b}_1^{(\alpha)}\ \mathbf{b}_2^{(\alpha)}\ \cdots]$. Then

$$\mathcal{M}_\alpha(\mathbf{v})\mathcal{M}_\alpha(\mathbf{v})^{\mathsf{H}} = \mathbf{B}_\alpha E_\alpha \mathbf{B}_\alpha^{\mathsf{H}}$$

holds in the matrix interpretation. In particular, $(\sigma_i^{(\alpha)})^2 = \lambda_i^{(\alpha)}$ is valid for the singular values $\sigma_i^{(\alpha)}$ of $\mathcal{M}_\alpha(\mathbf{v})$ and the eigenvalues $\lambda_i^{(\alpha)}$ of E_α.

In the following example, we apply the matricisation to a topological Hilbert tensor space $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j\in D} V_j$ with induced scalar product.

Example 5.16. Let $V_j = L^2(I_j)$ with $I_j \subset \mathbb{R}$ for $1 \le j \le d$. Then $L^2(I) = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ holds for $I = \times_{j=1}^d I_j$. Consider a function $\mathbf{f} \in L^2(I)$. To obtain the left singular vectors $u_i^{(j)} \in L^2(I_j)$, we have to form the operator

$$\mathcal{K}_j := \mathcal{M}_j(\mathbf{f})\mathcal{M}_j^*(\mathbf{f}) = \langle \mathcal{M}_j(\mathbf{f}), \mathcal{M}_j(\mathbf{f})\rangle_{[j]} \in \mathcal{L}(V_j, V_j).$$

The application of $\mathcal{K}_j$ to $g \in L^2(I_j)$ is given by

$$\mathcal{K}_j(g)(\xi) = \int_{I_j} k_j(\xi,\xi')g(\xi')\mathrm{d}\xi' \quad \text{with } I_{[j]} = \bigtimes_{k\neq j} I_k,\ \mathrm{d}x_{[j]} = \prod_{k\neq j} \mathrm{d}x_k \text{ in}$$

$$k_j(\xi,\xi') := \int_{I_{[j]}} \mathbf{f}(\ldots,x_{j-1},\xi,x_{j+1},\ldots)\, \overline{\mathbf{f}(\ldots,x_{j-1},\xi',x_{j+1},\ldots)}\, \mathrm{d}x_{[j]}.$$

The singular vectors $u_i^{(j)}$ and singular values $\sigma_i^{(j)}$ can be obtained from the eigenvalue problem

$$\mathcal{K}_j(u_i^{(j)}) = \left(\sigma_i^{(j)}\right)^2 u_i^{(j)} \qquad (i \in \mathbb{N}).$$

If $\mathbf{f} \in {}_a\bigotimes_{j=1}^d V_j$ is an algebraic tensor, $\mathcal{K}_j$ has finite rank and delivers only finitely many singular vectors $u_i^{(j)}$ with positive singular values $\sigma_i^{(j)}$.

5.2.4 Matricisation of a Family of Tensors

Let $\mathbf{F} = (\mathbf{v}_i)_{i\in I}$ be a family of tensors $\mathbf{v}_i \in \mathbf{V}_D = \bigotimes_{j\in D} V_j$. According to Lemma 3.26, the tuple $(\mathbf{v}_i)_{i\in I}$ may be considered as an element of the tensor space $\mathbf{V}_D \otimes \mathbb{K}^I$. If $D = \{1,\ldots,d\}$, define the *extended index set* $D_{\mathrm{ex}} := D \cup \{d+1\}$ and the extended tensor space $\mathbf{V}_{\mathrm{ex}} = \bigotimes_{j\in D_{\mathrm{ex}}} V_j$, where $V_{d+1} := \mathbb{K}^I$. Using the identification described in Lemma 3.26, we may view $\mathbf{F}$ as an element of $\mathbf{V}_{\mathrm{ex}}$. This allows us to define $\mathcal{M}_\alpha(\mathbf{F})$ for all $\alpha \subset D_{\mathrm{ex}}$. For instance, $\alpha = D$ yields

$$\mathcal{M}_D(\mathbf{F}) = \sum_{i\in I} \mathbf{v}_i \otimes e_i \qquad \left(\mathbf{v}_i \in \mathbf{V}_D, e_i \in \mathbb{K}^I : i\text{-th unit vector}\right).$$

From this representation one concludes the following result about the left-sided singular value decomposition (cf. (5.10a)).

Remark 5.17. $\langle \mathcal{M}_\alpha(\mathbf{F}), \mathcal{M}_\alpha(\mathbf{F})\rangle_{D_{\mathrm{ex}}\backslash\alpha} = \sum_{i\in I} \langle \mathcal{M}_\alpha(\mathbf{v}_i), \mathcal{M}_\alpha(\mathbf{v}_i)\rangle_{D\backslash\alpha}$ for $\alpha \subset D$.

5.3 Tensorisation

Tensorisation is the opposite of vectorisation: a vector is isomorphically transformed into a tensor, even if the tensor structure is not given beforehand.

One example of tensorisation has been presented in §1.2.4.1. There, $\mathbf{u}$ and $\mathbf{f}$ are grid functions and are usually considered as vectors. Because of the special shape of the grid G_n, the entries of $\mathbf{u}$ and $\mathbf{f}$ are of the form $\mathbf{u}_{ijk}$ and $\mathbf{f}_{ijk}$ $(1 \le i,j,k \le n)$ and $\mathbf{u}$ and $\mathbf{f}$ can be regarded as tensors from $\mathbb{K}^n \otimes \mathbb{K}^n \otimes \mathbb{K}^n$.

However, the tensorisation may also be rather artificial. Consider, e.g., any vector from $\mathbb{K}^I$ with $I := \{0,\ldots,n-1\}$ and assume that n is not a prime, so that a factorisation $n = n_1 n_2$ $(n_1, n_2 \ge 2)$ exists. Choose index sets $J_1 := \{0,\ldots,n_1-1\}$ and $J_2 := \{0,\ldots,n_2-1\}$ and $\mathbf{J} := J_1 \times J_2$. Since $\#\mathbf{J} = \#I$, there is a bijection $\alpha : \mathbf{J} \to I$ leading to an isomorphism

$$\begin{aligned} &x \in \mathbb{K}^I \longleftrightarrow \mathbf{v} \in \mathbb{K}^{\mathbf{J}} = \mathbb{K}^{J_1} \otimes \mathbb{K}^{J_2} \qquad \text{with} \\ &\mathbf{v}[j_1, j_2] = x[\alpha(j_1, j_2)] \text{ and } \alpha(j_1, j_2) = j_2 n_2 + j_1. \end{aligned} \tag{5.13}$$

In the latter case, $\mathbf{v} \in \mathbb{K}^{\mathbf{J}}$ is an $n_1 \times n_2$ matrix or a tensor of order two. Obviously, tensors of higher order can be obtained by exploiting a factorisation $n = n_1 n_2 \cdot \ldots \cdot n_d$ (assuming $n_j \ge 2$ to avoid trivial cases):

$$\mathbb{K}^I \cong \bigotimes_{j=1}^{d} \mathbb{K}^{J_j} \qquad \text{for } \#I = \prod_{j=1}^{d} \#J_j.$$

An extreme case is $\mathbb{K}^I$ with the dimension $\#I = 2^d$:

$$\mathbb{K}^{2^d} \cong \bigotimes_{j=1}^{d} \mathbb{K}^2. \tag{5.14}$$

The advantage of the representation as tensor is the fact that vectors (of length n) rewritten as tensors may require less storage. In the following example, the corresponding tensors are elementary tensors (cf. Khoromskij [118]). Consider

$$x \in \mathbb{K}^{\{0\ldots,n-1\}} \qquad \text{with } x_\nu = \zeta^\nu \quad \text{for } \ 0 \leq \nu \leq n-1, \tag{5.15}$$

where $\zeta \in \mathbb{K}$ is arbitrary. Such vectors appear in exponential sum approximations as well as in Fourier representations. The isomorphism (5.13) based on $n = n_1 n_2$ yields the matrix

$$M = \begin{bmatrix} \zeta^0 & \zeta^{n_1} & \ldots & \zeta^{(n_2-1)n_1} \\ \zeta^1 & \zeta^{n_1+1} & \ldots & \zeta^{(n_2-1)n_1+1} \\ \vdots & \vdots & \ddots & \vdots \\ \zeta^{n_1-1} & \zeta^{2n_1-1} & \ldots & \zeta^{n_2 n_1-1} \end{bmatrix} = \begin{bmatrix} \zeta^0 \\ \zeta^1 \\ \vdots \\ \zeta^{n_1-1} \end{bmatrix} \begin{bmatrix} \zeta^0 \ \zeta^{n_1} \ \cdots \ \zeta^{(n_2-1)n_1} \end{bmatrix},$$

which corresponds to the tensor product

$$x \in \mathbb{K}^{\{0\ldots,n-1\}} \longleftrightarrow \begin{bmatrix} \zeta^0 \\ \zeta^1 \\ \vdots \\ \zeta^{n_1-1} \end{bmatrix} \otimes \begin{bmatrix} \zeta^0 \\ \zeta^{n_1} \\ \vdots \\ \zeta^{(n_2-1)n_1} \end{bmatrix} \in \mathbb{K}^{\{0\ldots,n_1-1\}} \otimes \mathbb{K}^{\{0\ldots,n_2-1\}}.$$

The decomposition can be repeated for the first vector, provided that $n_1 = n'n''$ $(n', n'' \geq 2)$ and yields (after renaming n', n'', n_2 by n_1, n_2, n_3)

$$x \in \mathbb{K}^{\{0\ldots,n-1\}} \longleftrightarrow \begin{bmatrix} \zeta^0 \\ \zeta^1 \\ \vdots \\ \zeta^{n_1-1} \end{bmatrix} \otimes \begin{bmatrix} \zeta^0 \\ \zeta^{n_1} \\ \vdots \\ \zeta^{(n_2-1)n_1} \end{bmatrix} \otimes \begin{bmatrix} \zeta^0 \\ \zeta^{n_1 n_2} \\ \vdots \\ \zeta^{(n_3-1)n_1 n_2} \end{bmatrix}.$$

By induction, this proves the next statement.

Remark 5.18. Let $n = n_1 n_2 \cdots n_d$ with $n_j \geq 2$. The vector x from (5.15) corresponds to the elementary tensor $v^{(1)} \otimes \ldots \otimes v^{(d)}$ with $v^{(j)} \in \mathbb{K}^{\{0\ldots,n_j-1\}}$ defined by

$$v^{(j)} = \begin{bmatrix} \zeta^0 \\ \zeta^{p_j} \\ \vdots \\ \zeta^{(n_j-1)p_j} \end{bmatrix} \qquad \text{with } p_j := \prod_{k=1}^{j-1} n_k.$$

In the case of $n = 2^d$, i.e., $n_j = 2$, $p_j = 2^{j-1}$, and $v^{(j)} = \begin{bmatrix} 1 \\ \zeta^{2^{j-1}} \end{bmatrix}$, the data size is reduced from n to $2d = 2\log_2 n$.

Interestingly, this representation does not only save storage, but also provides a more stable representation. As an example consider the integral involving the oscillatory function $f(t) = \exp(-\alpha t)$ for $\alpha = 2\pi \mathrm{i} k + \beta$, $k \in \mathbb{N}$, $\beta > 0$, and $g(t) = 1$. For large k, the value

$$\int_0^1 f(t)g(t)\mathrm{d}t = \frac{1 - \exp(-\beta)}{\beta + 2\pi \mathrm{i} k}$$

is small compared with $\int_0^1 |f(t)g(t)| \, \mathrm{d}t = (1 - \exp(-\beta))\,/\beta$. For usual numerical integration one has to expect a cancellation error with the amplification factor

$$\varkappa := \int_0^1 |f(t)g(t)| \, \mathrm{d}t \,/\, \left| \int_0^1 f(t)g(t)\mathrm{d}t \right| = \sqrt{1 + (2\pi k/\beta)^2} \sim \frac{2\pi k}{\beta},$$

which is large for large k and small β. If we approximate the integral by[8]

$$S := \frac{1}{n} \sum_{\nu=0}^{n-1} f\left(\frac{\nu}{n}\right) g\left(\frac{\nu}{n}\right) \qquad \text{for } n = 2^d,$$

the floating point errors are amplified by $\varkappa$ from above. Using the tensorised grid functions $\mathbf{f} = \bigotimes_{j=1}^d f^{(j)}$ and $\mathbf{g} = \bigotimes_{j=1}^d g^{(j)}$ with

$$f^{(j)} = \begin{bmatrix} 1 \\ \exp(-(2\pi \mathrm{i} k + \beta)2^{j-1-d}) \end{bmatrix} \quad \text{and} \quad g^{(j)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

according to Remark 5.18, we rewrite[9] the sum (scalar product) as $\frac{1}{n} \langle \mathbf{f}, \mathbf{g} \rangle = \frac{1}{n} \prod_{j=1}^d \left\langle f^{(j)}, g^{(j)} \right\rangle$:

$$S = \frac{1}{n} \prod_{j=1}^d \left[1 + \exp(-(2\pi \mathrm{i} k + \beta)2^{j-1-d}) \right].$$

In this case, the amplification factor for the floating point errors is $O(d + 1)$ and does not deteriorate for $k \to \infty$ and $\beta \to 0$.

More details about tensorisation will follow in Chap. 14.

[8] Because of other quadrature weights for $\nu = 0$ and n, the sum S is not exactly the trapezoidal rule. For $\beta = 1/10$, $k = 1000$, and $d = 20$ (i.e., $n = 2^{20} = 1\,048\,576$), the value of S is $4.562_{10}\text{-}8 - 1.515_{10}\text{-}5\,\mathrm{i}$ (exact integral value: $2.410_{10}\text{-}10 - 1.515_{10}\text{-}5\,\mathrm{i}$).

[9] Note that $\prod_{j=1}^d \left(1 + x^{2^{j-1}}\right) = \sum_{\nu=0}^{2^d - 1} x^\nu$.

Chapter 6
Minimal Subspaces

Abstract The notion of minimal subspaces is closely connected with the representations of tensors, provided these representations can be characterised by (dimensions of) subspaces. A separate description of the theory of minimal subspaces can be found in Falcó-Hackbusch [57].

The tensor representations discussed in the later *Chapters 8, 11, 12* will lead to subsets $\mathcal{T}_{\mathbf{r}}$, $\mathcal{H}_{\mathfrak{r}}$, $\mathbb{T}_{\boldsymbol{\rho}}$ of a tensor space. The results of this chapter will prove weak closedness of these sets. Another result concerns the question of a best approximation: is the infimum also a minimum? In the positive case, it is guaranteed that the best approximation can be found in the same set.

For tensors $\mathbf{v} \in {}_a\bigotimes_{j=1}^d V_j$ we shall define 'minimal subspaces' $U_j^{\min}(\mathbf{v}) \subset V_j$ in *Sects. 6.1-6.4*. In *Sect. 6.5* we consider weakly convergent sequences $\mathbf{v}_n \rightharpoonup \mathbf{v}$ and analyse the connection between $U_j^{\min}(\mathbf{v}_n)$ and $U_j^{\min}(\mathbf{v})$. The main result will be presented in Theorem 6.24. While *Sects. 6.1-6.5* discuss minimal subspaces of algebraic tensors $\mathbf{v} \in {}_a\bigotimes_{j=1}^d V_j$, *Sect. 6.6* investigates $U_j^{\min}(\mathbf{v})$ for topological tensors $\mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$. The final *Sect. 6.7* is concerned with intersection spaces.

6.1 Statement of the Problem, Notations

Consider an algebraic tensor space $\mathbf{V} = {}_a\bigotimes_{j=1}^d V_j$ and a fixed tensor $\mathbf{v} \in \mathbf{V}$. Among the subspaces $U_j \subset V_j$ with

$$\mathbf{v} \in \mathbf{U} := {}_a\bigotimes_{j=1}^d U_j \tag{6.1}$$

we are looking for the smallest ones. We have to show that minimal subspaces U_j exist and that these minimal subspaces can be obtained simultaneously in (6.1) for all $1 \le j \le d$. Since it will turn out that the minimal subspaces are uniquely determined by $\mathbf{v}$, we use the notation

$$U_j^{\min}(\mathbf{v}) \subset V_j.$$

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_6,

The determination of $U_j^{\min}(\mathbf{v})$ will be given in (6.6) and (6.10). We shall characterise the features of $U_j^{\min}(\mathbf{v})$, e.g., the dimension

$$r_j := \dim(U_j^{\min}(\mathbf{v})).$$

Furthermore, the properties of $U_j^{\min}(\mathbf{v})$ for varying $\mathbf{v}$ are of interest. In particular, we consider $U_j^{\min}(\mathbf{v}_n)$ and its dimension for a sequence $\mathbf{v}_n \rightharpoonup \mathbf{v}$.

First, in §6.2, we explore the matrix case $d = 2$. In §6.6 we replace the algebraic tensor space by a Banach tensor space.

An obvious advantage of (6.1) is the fact that the subspaces U_j can be of finite dimension even if $\dim(V_j) = \infty$, as stated next.

Remark 6.1. For $\mathbf{v} \in {}_a\bigotimes_{j=1}^d V_j$ there are always finite dimensional subspaces $U_j \subset V_j$ satisfying (6.1). More precisely, $\dim(U_j) \leq \operatorname{rank}(\mathbf{v})$ can be achieved.

Proof. By definition of the algebraic tensor space, $\mathbf{v} \in {}_a\bigotimes_{j=1}^d V_j$ means that there is a *finite* linear combination

$$\mathbf{v} = \sum_{\nu=1}^{n} \bigotimes_{j=1}^{d} v_\nu^{(j)} \tag{6.2a}$$

with some integer $n \in \mathbb{N}_0$ and certain vectors $v_\nu^{(j)} \in V_j$. Define

$$U_j := \operatorname{span}\{v_\nu^{(j)} : 1 \leq \nu \leq n\} \qquad \text{for } 1 \leq j \leq d. \tag{6.2b}$$

Then $\mathbf{v} \in \mathbf{U} := {}_a\bigotimes_{j=1}^d U_j$ proves (6.1) with subspaces of dimension $\dim(U_j) \leq n$.

By definition of the tensor rank, the smallest n in (6.2a) is $n := \operatorname{rank}(\mathbf{v})$. □

6.2 Tensors of Order Two

6.2.1 Existence of Minimal Subspaces

First, we consider the matrix case $d = 2$ and admit any field $\mathbb{K}$. To ensure the existence of minimal subspaces, we need the *lattice property*

$$(X_1 \otimes_a X_2) \cap (Y_1 \otimes_a Y_2) \;=\; (X_1 \cap Y_1) \otimes_a (X_2 \cap Y_2)\,, \tag{6.3}$$

which is formulated in the next lemma more generally.

Lemma 6.2. *Let A be an index set of possibly infinite cardinality. Then*

$$\bigcap_{\alpha\in A} (U_{1,\alpha} \otimes_a U_{2,\alpha}) = \Big(\bigcap_{\alpha\in A} U_{1,\alpha}\Big) \otimes_a \Big(\bigcap_{\alpha\in A} U_{2,\alpha}\Big)$$

holds for any choice of subspaces $U_{j,\alpha} \subset V_j$.

Proof. The inclusion $\left(\bigcap_{\alpha\in A} U_{1,\alpha}\right) \otimes_a \left(\bigcap_{\alpha\in A} U_{2,\alpha}\right) \subset \bigcap_{\alpha\in A} \left(U_{1,\alpha} \otimes_a U_{2,\alpha}\right)$ is obvious. It remains to show that $\mathbf{v} \in U_{1,\beta} \otimes_a U_{2,\beta}$ for all $\beta \in A$ implies that $\mathbf{v} \in \left(\bigcap_{\alpha\in A} U_{1,\alpha}\right) \otimes_a \left(\bigcap_{\alpha\in A} U_{2,\alpha}\right)$. Choose some $\beta \in A$ and let $\gamma \in A$ be arbitrary. By assumption, $\mathbf{v}$ has representations

$$\mathbf{v} = \sum_{\nu=1}^{n_\beta} u_{\nu,\beta}^{(1)} \otimes u_{\nu,\beta}^{(2)} = \sum_{\nu=1}^{n_\gamma} u_{\nu,\gamma}^{(1)} \otimes u_{\nu,\gamma}^{(2)} \qquad \text{with } u_{\nu,\beta}^{(j)} \in U_{j,\beta},\ u_{\nu,\gamma}^{(j)} \in U_{j,\gamma}.$$

Thanks to Lemma 3.13, we may assume that $\{u_{\nu,\beta}^{(1)}\}$ and $\{u_{\nu,\beta}^{(2)}\}$ are linearly independent. A dual system $\varphi_\mu \in V_2'$ of $\{u_{\nu,\beta}^{(2)}\}$ satisfies $\varphi_\mu(u_{\nu,\beta}^{(2)}) = \delta_{\nu\mu}$ (cf. Definition 3.6). Application of $id \otimes \varphi_\mu$ to the first representation yields $(id \otimes \varphi_\mu)(\mathbf{v}) = u_{\mu,\beta}^{(1)}$, while the second representation leads to $\sum_{\nu=1}^{n_\gamma} \varphi_\mu(u_{\nu,\gamma}^{(2)})\, u_{\nu,\gamma}^{(1)}$. The resulting equation $u_{\mu,\beta}^{(1)} = \sum_{\nu=1}^{n_\gamma} \varphi_\mu(u_{\nu,\gamma}^{(2)})\, u_{\nu,\gamma}^{(1)}$ shows that $u_{\mu,\beta}^{(1)}$ is a linear combination of vectors $u_{\nu,\gamma}^{(1)} \in U_{1,\gamma}$, i.e., $u_{\mu,\beta}^{(1)} \in U_{1,\gamma}$. Since $\gamma \in A$ is arbitrary, $u_{\mu,\beta}^{(1)} \in \bigcap_{\alpha\in A} U_{1,\alpha}$ follows.

Analogously, using the dual system of $\{u_{\nu,\beta}^{(1)}\}$, one proves $u_{\nu,\gamma}^{(2)} \in \bigcap_{\alpha\in A} U_{2,\alpha}$. Hence, $\mathbf{v} \in \left(\bigcap_{\alpha\in A} U_{1,\alpha}\right) \otimes_a \left(\bigcap_{\alpha\in A} U_{2,\alpha}\right)$. □

Definition 6.3. For an algebraic tensor $\mathbf{v} \in V_1 \otimes_a V_2$, subspaces $U_1^{\min}(\mathbf{v}) \subset V_1$ and $U_2^{\min}(\mathbf{v}) \subset V_2$ are called *minimal subspaces*, if they satisfy

$$\mathbf{v} \in U_1^{\min}(\mathbf{v}) \otimes_a U_2^{\min}(\mathbf{v}), \tag{6.4a}$$

$$\mathbf{v} \in U_1 \otimes_a U_2 \quad \Rightarrow \quad U_1^{\min}(\mathbf{v}) \subset U_1 \text{ and } U_2^{\min}(\mathbf{v}) \subset U_2. \tag{6.4b}$$

Proposition 6.4. *All* $\mathbf{v} \in V_1 \otimes_a V_2$ *possess unique minimal subspaces* $U_j^{\min}(\mathbf{v})$ *for* $j = 1, 2$.

Proof. To prove existence and uniqueness of minimal subspaces, define the set

$$\mathcal{F} := \mathcal{F}(\mathbf{v}) := \{(U_1, U_2) : \mathbf{v} \in U_1 \otimes_a U_2 \text{ for subspaces } U_j \subset V_j\}.$$

$\mathcal{F}$ is non-empty, since $(V_1, V_2) \in \mathcal{F}$. Then $U_j^{\min}(\mathbf{v}) := \bigcap_{(U_1,U_2)\in\mathcal{F}} U_j$ holds for $j = 1, 2$. In fact, by Lemma 6.2, $\mathbf{v} \in U_1^{\min}(\mathbf{v}) \otimes_a U_2^{\min}(\mathbf{v})$ holds and proves (6.4a), while (6.4b) is a consequence of the construction by $\bigcap_{(U_1,U_2)\in\mathcal{F}} U_j$. □

Lemma 6.5. *Assume (3.16), i.e.,* $\mathbf{v} = \sum_{\nu=1}^{r} u_\nu^{(1)} \otimes u_\nu^{(2)}$ *holds with linearly independent* $\{u_\nu^{(j)} : 1 \le \nu \le r\}$ *for* $j = 1, 2$. *Then these vectors span the minimal subspaces:*

$$U_j^{\min}(\mathbf{v}) = \operatorname{span}\left\{u_\nu^{(j)} : 1 \le \nu \le r\right\} \qquad \text{for } j = 1, 2. \tag{6.5}$$

Proof. Apply the proof of Lemma 6.2 to the set $A := \{\beta, \gamma\}$ and the subspaces $U_{j,\beta} := \operatorname{span}\{u_\nu^{(j)} : 1 \le \nu \le r\}$, and $U_{j,\gamma} := U_j^{\min}(\mathbf{v})$. It shows that $U_{j,\beta} \subset U_j^{\min}(\mathbf{v})$. Since a strict inclusion is excluded, $U_{j,\beta} = U_j^{\min}(\mathbf{v})$ proves the assertion. □

As a consequence of (6.5), $\dim(U_j^{\min}(\mathbf{v}))=r$ holds for $j=1$ and $j=2$, proving the following result.

Corollary 6.6. $U_1^{\min}(\mathbf{v})$ and $U_2^{\min}(\mathbf{v})$ have identical finite dimensions.

A constructive algorithm for the determination of $U_j^{\min}(\mathbf{v})$ is already given in the proof of Lemma 3.13: As long as the vectors v_ν or w_ν in $\mathbf{v}=\sum_{\nu=1}^{n} v_\nu \otimes w_\nu$ are linearly dependent, one can reduce the number of terms by one. This process has to terminate after at most n steps. By Lemma 6.5, the resulting vectors $\{v_\nu\}$ and $\{w_\nu\}$ span $U_1^{\min}(\mathbf{v})$ and $U_2^{\min}(\mathbf{v})$.

In the proof of Lemma 6.5 we have already made indirect use of the following characterisation of $U_j^{\min}(\mathbf{v})$. The tensor $id\otimes\varphi_2 \in L(V_1,V_1)\otimes V_2'$ can be considered as a mapping from $L(V_1 \otimes_a V_2, V_1)$ (cf. §3.3.2.2):

$$v \otimes w \in V_1 \otimes_a V_2 \;\mapsto\; (id \otimes \varphi_2)\,(v \otimes w) := \varphi_2(w)\cdot v \in V_1 .$$

The action of $\varphi_1 \otimes id \in V_1' \otimes L(V_2,V_2) \subset L(V_1 \otimes_a V_2, V_2)$ is analogous.

Proposition 6.7. *For* $\mathbf{v} \in V_1 \otimes_a V_2$ *the minimal subspaces are characterised by*

$$U_1^{\min}(\mathbf{v}) = \{(id \otimes \varphi_2)\,(\mathbf{v}) : \varphi_2 \in V_2'\}\,, \tag{6.6a}$$

$$U_2^{\min}(\mathbf{v}) = \{(\varphi_1 \otimes id)\,(\mathbf{v}) : \varphi_1 \in V_1'\}\,. \tag{6.6b}$$

Proof. Repeat the proof of Lemma 6.2: there are maps $id \otimes \varphi_2$ yielding $u_\nu^{(1)}$. By Lemma 6.5, the vectors $u_\nu^{(1)}$ span $U_1^{\min}(\mathbf{v})$. Similarly for (6.6b). Note that the right-hand sides in (6.6a,b) are linear subspaces. □

For $V_1 = \mathbb{K}^{n_1}$ and $V_2 = \mathbb{K}^{n_2}$, tensors from $V_1 \otimes V_2$ are isomorphic to matrices from $\mathbb{K}^{n_1\times n_2}$. Then definition (6.6a) may be interpreted as

$$U_1^{\min}(\mathbf{v}) = \mathrm{range}\{M\} = \{Mx : x \in V_2\},$$

where $M = \mathcal{M}_1(\mathbf{v})$ is the matrix corresponding to $\mathbf{v}$. Similarly, (6.6b) becomes $U_2^{\min}(\mathbf{v}) = \mathrm{range}\{M^{\mathsf{T}}\}$.

Corollary 6.8. (a) Once $U_1^{\min}(\mathbf{v})$ and $U_2^{\min}(\mathbf{v})$ are given, one may select any basis $\{u_\nu^{(1)} : 1 \le \nu \le r\}$ of $U_1^{\min}(\mathbf{v})$ and find a representation $\mathbf{v} = \sum_{\nu=1}^{r} u_\nu^{(1)} \otimes u_\nu^{(2)}$ (cf. (3.16)) with the given $u_\nu^{(1)}$ and some basis $\{u_\nu^{(2)}\}$ of $U_2^{\min}(\mathbf{v})$. Vice versa, one may select a basis $\{u_\nu^{(2)} : 1 \le \nu \le r\}$ of $U_2^{\min}(\mathbf{v})$, and obtains $\mathbf{v} = \sum_{\nu=1}^{r} u_\nu^{(1)} \otimes u_\nu^{(2)}$ with the given $u_\nu^{(2)}$ and some basis $\{u_\nu^{(1)}\}$ of $U_1^{\min}(\mathbf{v})$.

(b) If $\{u_\nu^{(1)} : 1 \le \nu \le s\}$ is a basis of a larger subspace $U_1 \supsetneqq U_1^{\min}(\mathbf{v})$, a representation $\mathbf{v} = \sum_{\nu=1}^{s} u_\nu^{(1)} \otimes u_\nu^{(2)}$ still exists, but the vectors $u_\nu^{(2)}$ are linearly dependent.

(c) If we fix a basis $\{u_\nu^{(2)} : 1 \le \nu \le r\}$ of some subspace $U_2 \subset V_2$, there are mappings $\{\boldsymbol{\psi}_\nu : 1 \le \nu \le r\} \subset L(V_1 \otimes_a U_2, V_1)$ such that $\boldsymbol{\psi}_\nu(\mathbf{w}) \in U_1^{\min}(\mathbf{w})$ and

$$\mathbf{w} = \sum_{\nu=1}^{r} \boldsymbol{\psi}_\nu(\mathbf{w}) \otimes u_\nu^{(2)} \qquad \text{for all } \mathbf{w} \in V_1 \otimes U_2. \tag{6.7}$$

Proof. 1) Assume $\mathbf{v}=\sum_{\nu=1}^{r}\hat{u}_\nu^{(1)}\otimes\hat{u}_\nu^{(2)}$ and choose another basis $\{u_\nu^{(1)}: 1\leq\nu\leq r\}$. Inserting the transformation $\hat{u}_\nu^{(1)}=\sum_\mu a_{\nu\mu}u_\mu^{(1)}$ $(1\leq\nu\leq r)$, we get

$$\mathbf{v}=\sum_{\nu=1}^{r}\hat{u}_\nu^{(1)}\otimes\hat{u}_\nu^{(2)}=\sum_{\nu=1}^{r}\sum_{\mu=1}^{r}a_{\nu\mu}u_\mu^{(1)}\otimes\hat{u}_\nu^{(2)}=\sum_{\mu=1}^{r}u_\mu^{(1)}\otimes\sum_{\nu=1}^{r}a_{\nu\mu}\hat{u}_\nu^{(2)}=\sum_{\mu=1}^{r}u_\mu^{(1)}\otimes u_\mu^{(2)}$$

with the new basis $u_\mu^{(2)}:=\sum_{\nu=1}^{r}a_{\nu\mu}\hat{u}_\nu^{(2)}$. If r is minimal [Part (a)], the vectors $u_\mu^{(2)}$ are linearly independent (cf. Lemma 6.5); otherwise [Part (b)], they are linearly dependent.

2) Consider $\mathbf{w}\in V_1\otimes U_2$ and a basis $\{u_\nu^{(2)}: 1\leq\nu\leq r\}$ of $U_2\subset V_2$. By Part (b), there is a representation $\mathbf{w}=\sum_{\mu=1}^{r}u_\mu^{(1)}\otimes u_\mu^{(2)}$ with suitable $u_\mu^{(1)}\in V_1$. Let $\varphi_\nu\in V_2'$ be a dual system to $\{u_\nu^{(2)}: 1\leq\nu\leq r\}$ (cf. Definition 3.6) and set $\boldsymbol{\psi}_\nu:=id\otimes\varphi_\nu$. Application of $\boldsymbol{\psi}_\nu$ to $\mathbf{w}$ yields

$$\boldsymbol{\psi}_\nu(\mathbf{w})=\boldsymbol{\psi}_\nu\left(\sum_{\mu=1}^{r}u_\mu^{(1)}\otimes u_\mu^{(2)}\right)=\sum_{\mu=1}^{r}\varphi_\nu(u_\mu^{(2)})\cdot u_\mu^{(1)}=\sum_{\mu=1}^{r}\delta_{\nu\mu}u_\mu^{(1)}=u_\nu^{(1)},$$

proving assertion (6.7). □

6.2.2 Use of the Singular Value Decomposition

If $\mathbf{v}\in U_1\otimes_a U_2$ holds for subspaces $U_j\subset V_j$ of not too large dimension, the singular value decomposition offers a practical construction of $U_1^{\min}(\mathbf{v})$ and $U_2^{\min}(\mathbf{v})$. Although the singular value decomposition produces orthonormal bases, no Hilbert structure is required for V_1 and V_2. The approach is restricted to the fields $\mathbb{R}$ and $\mathbb{C}$.

Remark 6.9. Let $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$. Suppose that a representation of $\mathbf{v}\in U_1\otimes_a U_2$ by $\mathbf{v}=\sum_{\nu=1}^{n}v_\nu^{(1)}\otimes v_\nu^{(2)}$ with $v_\nu^{(j)}\in U_j$ and $\dim(U_j)<\infty$ is given.
1) Choose bases $\{u_i^{(j)}: 1\leq i\leq n_j\}$ of U_j $(j=1,2)$ and determine the coefficients of $v_\nu^{(j)}$:

$$v_\nu^{(j)}=\sum\nolimits_{i=1}^{n_j}c_{\nu i}^{(j)}\,u_i^{(j)}\qquad(j=1,2).$$

Hence, $\mathbf{v}=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}M_{ij}\,u_i^{(1)}\otimes u_j^{(2)}$ has the coefficients $M_{ij}:=\sum_{\nu=1}^{n}c_{\nu i}^{(1)}c_{\nu j}^{(2)}$.
2) Determine the reduced singular value decomposition of the matrix $M\in\mathbb{K}^{n_1\times n_2}$ by calling the procedure $\mathbf{RSVD}(n_1,n_2,r,M,U,\varSigma,V)$, i.e.,

$$M=U\varSigma V^\mathsf{T}=\sum\nolimits_{\nu=1}^{r}\sigma_\nu a_\nu b_\nu^\mathsf{T}.$$

3) Define $\{\hat{a}_\nu: 1\leq\nu\leq r\}\subset U_1$ and $\{\hat{b}_\nu: 1\leq\nu\leq r\}\subset U_2$ by[1]

[1] $a_\nu[i]$ is the i-th entry of $a_\nu\in\mathbb{K}^{n_1}$, etc.

$$\hat{a}_\nu := \sum_{i=1}^{n_1} \sigma_\nu \, a_\nu[i] \, u_i^{(1)} \quad \text{and} \quad \hat{b}_\nu := \sum_{j=1}^{n_2} b_\nu[j] \, u_j^{(2)} \qquad \text{for } 1 \le \nu \le r.$$

They span the minimal subspaces $U_1^{\min}(\mathbf{v}) := \operatorname{span}\{\hat{a}_\nu : 1 \le \nu \le r\} \subset U_1$ and $U_2^{\min}(\mathbf{v}) := \operatorname{span}\{\hat{b}_\nu : 1 \le \nu \le r\} \subset U_2$, and $\mathbf{v} = \sum_{\nu=1}^{r} \hat{a}_\nu \otimes \hat{b}_\nu$ holds.

4) $\dim(U_1^{\min}(\mathbf{v})) = \dim(U_2^{\min}(\mathbf{v})) = \operatorname{rank}(M) = r$.

Proof. Singular value decomposition yields

$$\begin{aligned} \mathbf{v} &= \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} M_{ij} \, u_i^{(1)} \otimes u_j^{(2)} = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left(\sum_{\nu=1}^{r} \sigma_\nu a_\nu[i] b_\nu[j] \right) u_i^{(1)} \otimes u_j^{(2)} \\ &= \sum_{\nu=1}^{r} \left(\sum_{i=1}^{n_v} \sigma_\nu a_\nu[i] \, u_i^{(1)} \right) \otimes \left(\sum_{j=1}^{n_w} b_\nu[j] \, u_j^{(2)} \right) = \sum_{\nu=1}^{r} \hat{a}_\nu \otimes \hat{b}_\nu . \end{aligned}$$

Since the vectors a_ν are linearly independent, also the $\hat{a}_\nu$ are linearly independent. Similarly, $\{\hat{b}_\nu\}$ forms a basis. □

6.2.3 Minimal Subspaces for a Family of Tensors

The minimal subspaces $U_j^{\min}(\mathbf{v})$ serve for representing a *single* tensor $\mathbf{v} \in V_1 \otimes_a V_2$. Now we replace the tensor $\mathbf{v}$ by a subset $F \subset V_1 \otimes_a V_2$ and ask for minimal subspaces $U_1^{\min}(F)$ and $U_2^{\min}(F)$ so that $\mathbf{v} \in U_1^{\min}(F) \otimes_a U_2^{\min}(F)$ holds for all $\mathbf{v} \in F$.

The obvious result is summarised in the next remark.

Proposition 6.10. *Let $F \subset V_1 \otimes_a V_2$ be a non-empty subset. Then the minimal subspaces $U_1^{\min}(F)$ and $U_2^{\min}(F)$ are* [2]

$$U_1^{\min}(F) := \sum_{\mathbf{v} \in F} U_1^{\min}(\mathbf{v}) \quad \textit{and} \quad U_2^{\min}(F) := \sum_{\mathbf{v} \in F} U_2^{\min}(\mathbf{v}). \tag{6.8a}$$

Another characterisation is

$$\begin{aligned} U_1^{\min}(F) &= \operatorname{span}\{(id \otimes \varphi_2)(\mathbf{v}) : \varphi_2 \in V_2', \mathbf{v} \in F\}, \\ U_2^{\min}(F) &= \operatorname{span}\{(\varphi_1 \otimes id)(\mathbf{v}) : \varphi_1 \in V_1', \mathbf{v} \in F\}. \end{aligned} \tag{6.8b}$$

Proof. $\mathbf{v} \in F$ and $\mathbf{v} \in U_1^{\min}(F) \otimes_a U_2^{\min}(F)$ require $U_j^{\min}(\mathbf{v}) \subset U_j^{\min}(\mathbf{v})(F)$ for $j = 1, 2$ and all $\mathbf{v} \in F$; hence, $\bigcup_{\mathbf{v} \in F} U_j^{\min}(\mathbf{v}) \subset U_j^{\min}(\mathbf{v})(F)$. The smallest subspace containing $\bigcup_{\mathbf{v} \in F} U_j^{\min}(\mathbf{v})$ is the sum $\sum_{\mathbf{v} \in F} U_j^{\min}(\mathbf{v})$, implying (6.8a). Equations (6.6a,b) prove (6.8b). □

[2] The sum of subspaces is defined by the span of their union.

6.3 Minimal Subspaces of Higher Order Tensors

In the following we assume that $d \geq 3$ and generalise some of the features of tensors of second order.

By Remark 6.1, we may assume $\mathbf{v} \in \mathbf{U} := \bigotimes_{j=1}^{d} U_j$ with *finite* dimensional subspaces $U_j \subset V_j$. The lattice structure from Lemma 6.2 generalises to higher order.

Lemma 6.11. *Let $X_j, Y_j \subset V_j$ for $1 \leq j \leq d$. Then the identity*

$$\left({}_a\bigotimes_{j=1}^{d} X_j \right) \cap \left({}_a\bigotimes_{j=1}^{d} Y_j \right) = {}_a\bigotimes_{j=1}^{d} (X_j \cap Y_j)$$

holds and can be generalised to infinitely many intersections.

Proof. For the start of the induction at $d = 2$ use Lemma 6.2. Assume that the assertion holds for $d - 1$ and use ${}_a\bigotimes_{j=1}^{d} X_j = X_1 \otimes X_{[1]}$ with $X_{[1]} := {}_a\bigotimes_{j=2}^{d} X_j$ and ${}_a\bigotimes_{j=1}^{d} Y_j = Y_1 \otimes Y_{[1]}$. Lemma 6.2 states that $\mathbf{v} \in (X_1 \cap Y_1) \otimes \left(X_{[1]} \cap Y_{[1]}\right)$. By inductive hypothesis, $X_{[1]} \cap Y_{[1]} = {}_a\bigotimes_{j=2}^{d} (X_j \cap Y_j)$ holds proving the assertion. □

Again, the minimal subspaces $U_j^{\min}(\mathbf{v})$ can be defined by the intersection of all subspaces $U_j \subset V_j$ satisfying $\mathbf{v} \in {}_a\bigotimes_{j=1}^{d} U_j$.

The algebraic characterisation of $U_j^{\min}(\mathbf{v})$ is similar as for $d = 2$. Here we use the short notation $\bigotimes_{k \neq j}$ instead of $\bigotimes_{k \in \{1,\ldots,d\} \backslash \{j\}}$. Note that the following right-hand sides of (6.9a-d) involve the spaces ${}_a\bigotimes_{k \neq j} V_k'$, $({}_a\bigotimes_{k \neq j} V_k)'$, ${}_a\bigotimes_{k \neq j} V_k^*$, $(\bigotimes_{k \neq j} V_k)^*$, which may differ. Nevertheless, the image spaces are identical.

Lemma 6.12. *Let $\mathbf{v} \in \mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$. (a) The two spaces*

$$U_j^{I}(\mathbf{v}) := \left\{ \varphi(\mathbf{v}) : \varphi \in {}_a\bigotimes\nolimits_{k \neq j} V_k' \right\}, \tag{6.9a}$$

$$U_j^{II}(\mathbf{v}) := \left\{ \varphi(\mathbf{v}) : \varphi \in \left({}_a\bigotimes\nolimits_{k \neq j} V_k \right)' \right\} \tag{6.9b}$$

coincide: $U_j^{I}(\mathbf{v}) = U_j^{II}(\mathbf{v})$.

(b) If V_j are normed spaces, one may replace algebraic functionals by continuous functionals:

$$U_j^{III}(\mathbf{v}) := \left\{ \varphi(\mathbf{v}) : \varphi \in {}_a\bigotimes\nolimits_{k \neq j} V_k^* \right\}. \tag{6.9c}$$

Then $U_j^{I}(\mathbf{v}) = U_j^{II}(\mathbf{v}) = U_j^{III}(\mathbf{v})$ is valid.

(c) If ${}_a\bigotimes_{k \neq j} V_k$ is a normed space, one may define

$$U_j^{IV}(\mathbf{v}) := \left\{ \varphi(\mathbf{v}) : \varphi \in \left(\bigotimes\nolimits_{k \neq j} V_k \right)^* \right\}. \tag{6.9d}$$

Then $U_j^{I}(\mathbf{v}) = U_j^{II}(\mathbf{v}) = U_j^{IV}(\mathbf{v})$ holds.

Proof. 1) Since the mappings φ are applied to $\mathbf{v} \in \mathbf{U} := \bigotimes_{j=1}^{d} U_j$ (cf. Remark 6.1), one may replace $\varphi \in {}_a\bigotimes_{k\neq j} V_k'$ by $\varphi \in {}_a\bigotimes_{k\neq j} U_k'$ and $\varphi \in ({}_a\bigotimes_{k\neq j} V_k)'$ by $\varphi \in ({}_a\bigotimes_{k\neq j} U_k)'$ without changing $\varphi(\mathbf{v})$. Since $\dim(U_k) < \infty$, Proposition 3.52c states that ${}_a\bigotimes_{k\neq j} U_k' = ({}_a\bigotimes_{k\neq j} U_k)'$. This proves Part (a).

2) As in Part 1) we may restrict φ to ${}_a\bigotimes_{k\neq j} U_k' = ({}_a\bigotimes_{k\neq j} U_k)'$. Since $\dim(U_k) < \infty$, algebraic duals are continuous, i.e.,

$$\varphi \in {}_a\bigotimes_{k\neq j} U_k' = \Big({}_a\bigotimes\nolimits_{k\neq j} U_k\Big)' = {}_a\bigotimes_{k\neq j} U_k^* \,.$$

By Hahn-Banach (Theorem 4.15), such mappings can be extended to ${}_a\bigotimes_{k\neq j} V_k^*$. This proves Part (b), while Part (c) is analogous. □

Theorem 6.13. *(a) For any* $\mathbf{v} \in \mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$ *there exist minimal subspaces* $U_j^{\min}(\mathbf{v})$ $(1 \leq j \leq d)$. *An algebraic characterisation of* $U_j^{\min}(\mathbf{v})$ *is*

$$U_j^{\min}(\mathbf{v}) = \operatorname{span}\left\{ \begin{array}{c} (\varphi_1 \otimes \ldots \otimes \varphi_{j-1} \otimes id \otimes \varphi_{j+1} \otimes \ldots \otimes \varphi_d)\,(\mathbf{v}) \\ \text{with } \varphi_k \in V_k' \text{ for } k \neq j \end{array} \right\} \tag{6.10a}$$

or equivalently

$$U_j^{\min}(\mathbf{v}) = \Big\{ \varphi(\mathbf{v}) : \ \varphi \in {}_a\bigotimes\nolimits_{k\neq j} V_k' \Big\}, \tag{6.10b}$$

where the action of the functional φ *is understood as in (6.10a).* $U_j^{\min}(\mathbf{v})$ *coincides with the sets from (6.9a-d).*
(b) For a subset $F \subset \mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$ *of tensors, the minimal subspaces* $V_{j,F}$ *with* $F \subset {}_a\bigotimes_{j=1}^{d} V_{j,F}$ *are*

$$U_j^{\min}(F) = \sum_{\mathbf{v}\in F} U_j^{\min}(\mathbf{v}). \tag{6.10c}$$

(c) For finite dimensional V_j, *the* j*-rank is defined in (5.6b) and satisfies*

$$\operatorname{rank}_j(\mathbf{v}) = \dim(U_j^{\min}(\mathbf{v})), \tag{6.10d}$$

while, for the infinite dimensional case, Eq. (6.10d) is the true generalisation of the definition of rank_j .

Proof. 1) The equivalence of (6.10a) and (6.10b) is easily seen: Linear combinations of elementary tensors $\bigotimes_{k\neq j} \varphi_k$ are expressed by $\operatorname{span}\{\ldots\}$ in (6.10a) and by $\varphi \in {}_a\bigotimes_{k\neq j} V_k'$ in (6.10b).

2) We apply the matricisation from §5.2. The isomorphism $\mathcal{M}_j$ from Definition 5.3 maps ${}_a\bigotimes_{k=1}^{d} V_k$ into $V_j \otimes_a V_{[j]}$. Proposition 6.7 states that

$$U_j^{\min}(\mathbf{v}) = \Big\{ \varphi(\mathbf{v}) : \ \varphi \in V_{[j]}' \Big\} = \Big\{ \varphi(\mathbf{v}) : \ \varphi \in \Big({}_a\bigotimes\nolimits_{k\neq j} V_k \Big)' \Big\} \tag{6.11}$$

is the minimal subspace. The set on the right-hand side is $U_j^{II}(\mathbf{v})$ (cf. (6.9b)) and

Lemma 6.12 states that $U_j^{II}(\mathbf{v}) = U_j^{I}(\mathbf{v})$, where $U_j^{I}(\mathbf{v})$ coincides with the set on the right-hand side of (6.10b). So far, we have proved $\mathbf{v} \in U_j^{\min}(\mathbf{v}) \otimes_a V_{[j]}$. Thanks to Lemma 6.11, the intersection may be performed componentwise yielding $\mathbf{v} \in \bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$.

3) For families of tensors, the argument of Proposition 6.10 proves Part (b).

4) Concerning rank_j see the discussion in §6.4. □

The right-hand side in (6.10a) is the span of a subset. For $d = 2$, the symbol 'span' may be omitted, since the subset is already a subspace (cf. Proposition 6.7).

Exercise 6.14. (a) For a subset $F \subset \mathbf{V}$ let $U_F := \mathrm{span}\{F\} \supset F$. Show that $U_j^{\min}(F) = U_j^{\min}(U_F)$.
(b) Let $F \subset \mathbf{V}$ be a subspace of finite dimension. Show that $\dim(U_j^{\min}(F)) < \infty$.

The determination of $U_j^{\min}(\mathbf{v})$ by (6.10a,b) is not very constructive, since it requires the application of all dual mappings $\varphi_\nu \in V_\nu'$. Another approach is already used in the proof above.

Remark 6.15. For $j \in \{1, \ldots, d\}$ apply the matricisation

$$M_j := \mathcal{M}_j(\mathbf{v}) \in V_j \otimes_a V_{[j]} \qquad \text{with} \quad V_{[j]} := {}_a\bigotimes_{k \in \{1,\ldots,d\}\backslash\{j\}} V_k \, .$$

The techniques of §6.2.1 and §6.2.2 may be used to determine the minimal subspaces $U_j^{\min}(\mathbf{v})$ and $U_{[j]}^{\min}(\mathbf{v})$: $\mathcal{M}_j(\mathbf{v}) \in U_j^{\min}(\mathbf{v}) \otimes U_{[j]}^{\min}(\mathbf{v})$. In particular, if a singular value decomposition is required, one can make use of Remark 2.24, since only the first subspace $U_j^{\min}(\mathbf{v})$ is of interest.

Remark 6.16. While $\dim(U_1^{\min}(\mathbf{v})) = \dim(U_2^{\min}(\mathbf{v}))$ holds for $d=2$ (cf. Corollary 6.6), the dimensions of $U_j^{\min}(\mathbf{v})$ may be different for $d \geq 3$.

6.4 Hierarchies of Minimal Subspaces and rank_α

So far, we have defined minimal subspaces $U_j^{\min}(\mathbf{v})$ for a single index $j \in D := \{1, \ldots, d\}$. We can extend this definition to $\mathbf{U}_\alpha^{\min}(\mathbf{v})$, where $\emptyset \subsetneqq \alpha \subsetneqq D$ are subsets (cf. (5.3b)). For illustration we consider the example

$$\mathbf{v} \in \mathbf{V} = \bigotimes_{j=1}^{7} V_j = \underbrace{(V_1 \otimes V_2)}_{=\mathbf{V}_\alpha} \otimes \underbrace{(V_3 \otimes V_4)}_{=\mathbf{V}_\beta} \otimes \underbrace{(V_5 \otimes V_6 \otimes V_7)}_{=\mathbf{V}_\gamma} = \mathbf{V}_\alpha \otimes \mathbf{V}_\beta \otimes \mathbf{V}_\gamma,$$

in which we use the isomorphism between $\mathbf{V} = \bigotimes_{j=1}^{7} V_j$ and $\mathbf{V}_\alpha \otimes \mathbf{V}_\beta \otimes \mathbf{V}_\gamma$. Ignoring the tensor structure of $\mathbf{V}_\alpha, \mathbf{V}_\beta, \mathbf{V}_\gamma$, we regard $\mathbf{V} = \mathbf{V}_\alpha \otimes \mathbf{V}_\beta \otimes \mathbf{V}_\gamma$ as tensor space of order 3. Consequently, for $\mathbf{v} \in \mathbf{V}$ there are minimal subspaces $\mathbf{U}_\alpha^{\min}(\mathbf{v}) \subset \mathbf{V}_\alpha = V_1 \otimes V_2$, $\mathbf{U}_\beta^{\min}(\mathbf{v}) \subset \mathbf{V}_\beta = V_3 \otimes V_4$, and $\mathbf{U}_\gamma^{\min}(\mathbf{v}) \subset \mathbf{V}_\gamma = V_5 \otimes V_6 \otimes V_7$

such that $\mathbf{v}\in\mathbf{U}_\alpha^{\min}(\mathbf{v})\otimes\mathbf{U}_\beta^{\min}(\mathbf{v})\otimes\mathbf{U}_\gamma^{\min}(\mathbf{v})$. These minimal subspaces may be constructively determined from $\mathcal{M}_\alpha(\mathbf{v})$, $\mathcal{M}_\beta(\mathbf{v})$, $\mathcal{M}_\gamma(\mathbf{v})$.

As in the example, we use the notations (5.3a-d) for D, α, α^c, and $\mathbf{V}_\alpha$. By $\mathbf{V}=\mathbf{V}_\alpha\otimes\mathbf{V}_{\alpha^c}$, any $\mathbf{v}\in\mathbf{V}$ gives rise to minimal subspaces $\mathbf{U}_\alpha^{\min}(\mathbf{v})\subset\mathbf{V}_\alpha$ and $\mathbf{U}_{\alpha^c}^{\min}(\mathbf{v})\subset\mathbf{V}_{\alpha^c}$.

Proposition 6.17. *Let* $\mathbf{v}\in\mathbf{V}=\bigotimes_{j=1}^d V_j$*, and* $\emptyset\neq\alpha\subset D$*. Then the minimal subspace* $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ *and the minimal subspaces* $U_j^{\min}(\mathbf{v})$ *for* $j\in\alpha$ *are related by*

$$\mathbf{U}_\alpha^{\min}(\mathbf{v})\subset\bigotimes_{j\in\alpha}U_j^{\min}(\mathbf{v}). \tag{6.12}$$

Proof. We know that $\mathbf{v}\in\mathbf{U}=\bigotimes_{j=1}^d U_j^{\min}(\mathbf{v})$. Writing $\mathbf{U}$ as $\mathbf{U}_\alpha\otimes\mathbf{U}_{\alpha^c}$ with $\mathbf{U}_\alpha:=\bigotimes_{j\in\alpha}U_j^{\min}(\mathbf{v})$ and $\mathbf{U}_{\alpha^c}:=\bigotimes_{j\in\alpha^c}U_j^{\min}(\mathbf{v})$, we see that $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ must be contained in $\mathbf{U}_\alpha=\bigotimes_{j\in\alpha}U_j^{\min}(\mathbf{v})$. □

An obvious generalisation is the following.

Corollary 6.18. Let $\mathbf{v}\in\mathbf{V}=\bigotimes_{j=1}^d V_j$. Assume that $\emptyset\neq\alpha_1,\ldots,\alpha_m,\beta\subset D$ are subsets such that $\beta=\bigcup_{\mu=1}^m\alpha_\mu$ is a disjoint union. Then

$$\mathbf{U}_\beta^{\min}(\mathbf{v})\subset\bigotimes_{\mu=1}^m\mathbf{U}_{\alpha_\mu}^{\min}(\mathbf{v}).$$

In particular, if $\emptyset\neq\alpha,\alpha_1,\alpha_2\subset D$ satisfy $\alpha=\alpha_1\,\dot\cup\,\alpha_2$ (disjoint union), then

$$\mathbf{U}_\alpha^{\min}(\mathbf{v})\subset\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})\otimes\mathbf{U}_{\alpha_2}^{\min}(\mathbf{v}). \tag{6.13}$$

If

$$\prod_{\mu\in\alpha^c}\dim(V_\mu)\geq\dim(\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v}))\cdot\dim(\mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})),$$

there are tensors $\mathbf{v}\in\mathbf{V}$ such that (6.13) holds with equality sign.

Proof. For the last statement let $\{b_i^{(1)}\}$ be a basis of $\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})$ and $\{b_j^{(2)}\}$ a basis of $\mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})$. Then $\{b_i^{(1)}\otimes b_j^{(2)}\}$ is a basis of $\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})\otimes\mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})$. For all pairs (i,j) choose linearly independent tensors $w_{ij}\in\bigotimes_{\mu\in\alpha^c}V_\mu$ and set

$$\mathbf{v}:=\sum\nolimits_{i,j}w_{ij}\otimes b_i^{(1)}\otimes b_j^{(2)}\in\mathbf{V}.$$

One verifies that $\mathbf{U}_\alpha^{\min}(\mathbf{v})=\operatorname{span}\{b_i^{(1)}\otimes b_j^{(2)}\}=\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})\otimes\mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})$. □

The algebraic characterisation of $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ is analogous to (6.10a,b):

$$\mathbf{U}_\alpha^{\min}(\mathbf{v})=\operatorname{span}\left\{\varphi_{\alpha^c}(\mathbf{v}):\ \varphi_{\alpha^c}=\bigotimes\nolimits_{j\in\alpha^c}\varphi^{(j)},\ \varphi^{(j)}\in V_j'\right\}, \tag{6.14}$$

where $\varphi_{\alpha^c}\big(\bigotimes_{j=1}^d v^{(j)}\big):=\varphi_{\alpha^c}\big(\bigotimes_{j\in\alpha^c}v^{(j)}\big)\cdot\bigotimes_{j\in\alpha}v^{(j)}$.

In Definition 5.7, rank_α is introduced by $\mathrm{rank}_\alpha(\mathbf{v}) := \mathrm{rank}\left(\mathcal{M}_\alpha(\mathbf{v})\right)$, where $\mathcal{M}_\alpha(\mathbf{v})$ may be interpreted as a matrix. In the finite dimensional case, $\mathrm{rank}(\mathcal{M}_\alpha(\mathbf{v}))$ equals the dimension of $\mathrm{range}(\mathcal{M}_\alpha(\mathbf{v}))$. In general, $\mathcal{M}_\alpha(\mathbf{v})$ is a mapping from $\mathbf{V}'_{\alpha^c}$ into $\mathbf{V}_\alpha$, whereas the interpretation of the range of the matrix $\mathcal{M}_\alpha(\mathbf{v})$ considers $\mathcal{M}_\alpha(\mathbf{v})$ as a mapping from $\mathbf{V}_{\alpha^c}$ into $\mathbf{V}_\alpha$, which is true for the finite dimensional case, since then $\mathbf{V}'_{\alpha^c}$ and $\mathbf{V}_{\alpha^c}$ may be identified. As announced in (5.8), the true generalisation is

$$\mathrm{rank}_\alpha(\mathbf{v}) = \dim(\mathbf{U}_\alpha^{\min}(\mathbf{v})) \tag{6.15}$$

(cf. Theorem 6.13c), which includes the case of $\mathrm{rank}_\alpha(\mathbf{v}) = \infty$ for $\mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ with $\dim(\mathbf{U}_\alpha^{\min}(\mathbf{v})) = \infty$. For completeness, we define

$$\mathrm{rank}_\emptyset(\mathbf{v}) = \mathrm{rank}_D(\mathbf{v}) = \begin{cases} 1 & \text{if } \mathbf{v} \neq 0, \\ 0 & \text{if } \mathbf{v} = 0 \end{cases} \tag{6.16}$$

(cf. Footnote 5 on page 164). The α-ranks satisfy the following basic rules.

Lemma 6.19. *(a) The ranks for $\alpha \subset D$ and for the complement α^c coincide:*

$$\mathrm{rank}_\alpha(\mathbf{v}) = \mathrm{rank}_{\alpha^c}(\mathbf{v}). \tag{6.17a}$$

(b) If $\alpha \subset D$ is the disjoint union $\alpha = \beta \,\dot\cup\, \gamma$, then

$$\mathrm{rank}_\alpha(\mathbf{v}) \le \mathrm{rank}_\beta(\mathbf{v}) \cdot \mathrm{rank}_\gamma(\mathbf{v}) \tag{6.17b}$$

(c) If

$$\prod_{\mu \in \alpha^c} \dim(V_\mu) \ge \mathrm{rank}_\beta(\mathbf{v}) \cdot \mathrm{rank}_\gamma(\mathbf{v}), \tag{6.17c}$$

then there are $\mathbf{v}$ such that equality holds in (6.17b):

$$\mathrm{rank}_\alpha(\mathbf{v}) = \mathrm{rank}_\beta(\mathbf{v}) \cdot \mathrm{rank}_\gamma(\mathbf{v}) \tag{6.17d}$$

In particular, under condition (6.17c), random tensors satisfy (6.17d) with probability one.

Proof. 1) In the finite dimensional case, we can use $\mathcal{M}_{\alpha^c}(\mathbf{v}) = \mathcal{M}_\alpha(\mathbf{v})^\mathsf{T}$ to derive (6.17a) from Definition 5.7. In general, use $\mathbf{V} = \mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$ and Corollary 6.6.

2) Definition (6.15) together with (6.13) yields (6.17b).

3) The last statement in Corollary 6.18 yields Part (c). A random tensor yields a random matrix $\mathcal{M}_\alpha(\mathbf{v})$, so that Remark 2.5 applies. □

Corollary 6.20. (a) Decompose $D = \{1, \ldots, d\}$ disjointly into $D = \alpha \,\dot\cup\, \beta \,\dot\cup\, \gamma$. Then the following inequalities hold:

$$\begin{aligned} \mathrm{rank}_\alpha(\mathbf{v}) &\le \mathrm{rank}_\beta(\mathbf{v}) \cdot \mathrm{rank}_\gamma(\mathbf{v}), \\ \mathrm{rank}_\beta(\mathbf{v}) &\le \mathrm{rank}_\alpha(\mathbf{v}) \cdot \mathrm{rank}_\gamma(\mathbf{v}), \\ \mathrm{rank}_\gamma(\mathbf{v}) &\le \mathrm{rank}_\alpha(\mathbf{v}) \cdot \mathrm{rank}_\beta(\mathbf{v}). \end{aligned}$$

(b) Let $\alpha = \{\underline{j}, \underline{j}+1, \ldots, \overline{j}\}$, $\beta = \{1, \ldots, \underline{j}-1\}$, $\gamma = \{1, \ldots, \overline{j}\}$. Then (6.17b) holds again.

Proof. 1) Since $\alpha^c = \beta \,\dot{\cup}\, \gamma$, the combination of (6.17a,b) proves the first inequality of Part (a). Since α, β, γ are symmetric in their properties, the further two inequalities follow.

2) For Part (b) note that $D = \alpha \,\dot{\cup}\, \beta \,\dot{\cup}\, \gamma^c$. □

We conclude with a comparison of the α-rank and the tensor rank introduced in Definition 3.32.

Remark 6.21. $\mathrm{rank}_\alpha(\mathbf{v}) \le \mathrm{rank}(\mathbf{v})$ holds for $\mathbf{v} \in {}_a\bigotimes_{k=1}^d V_k$ and $\alpha \subset \{1, \ldots, d\}$. While $\mathrm{rank}(\cdot)$ may depend on the underlying field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ (cf. §3.2.6.3), the value of $\mathrm{rank}_\alpha(\cdot)$ is independent.

Proof. 1) Rewrite $\mathbf{v} = \sum_{i=1}^r \bigotimes_{j=1}^d u_i^{(j)}$ with $r = \mathrm{rank}(\mathbf{v})$ as

$$\sum_{i=1}^r \mathbf{u}_i^{(\alpha)} \otimes \mathbf{u}_i^{(\alpha^c)}, \quad \text{where } \mathbf{u}_i^{(\alpha)} := \bigotimes_{j\in\alpha} u_i^{(j)}.$$

The dimension of $\mathbf{U}_\alpha := \mathrm{span}\{\mathbf{u}_{\alpha,i}^{(j)} : 1 \le i \le r\}$ satisfies

$$\mathrm{rank}_\alpha(\mathbf{v}) = \dim(\mathbf{U}_\alpha^{\min}(\mathbf{v})) \le \dim(\mathbf{U}_\alpha) \le r = \mathrm{rank}(\mathbf{v}).$$

2) Because $\mathrm{rank}_\alpha(\mathbf{v})$ is the *matrix rank* of $\mathcal{M}_\alpha(\mathbf{v})$, Remark 2.2 proves independence of the field. □

6.5 Sequences of Minimal Subspaces

Let $\mathbf{V} := {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ be a Banach tensor space with norm $\|\cdot\|$ and assume that

$$\|\cdot\| \gtrsim \|\cdot\|_\vee\,. \tag{6.18}$$

We recall that all reasonable crossnorms satisfy (6.18) (cf. Proposition 4.68).

The following lemma allows us to define a subspace $U_j^{\min}(\mathbf{v})$ for *topological* tensors $\mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ not belonging to the algebraic tensor space ${}_a\bigotimes_{j=1}^d V_j$. Since we do not exclude infinite dimensional spaces $U_j^{\min}(\mathbf{v})$, we use the closure of the respective sets in (6.19) with respect to the norm of V_j. However, whether $U_j^{\min}(\mathbf{v})$ is closed or not is irrelevant as long as only the Banach subspace

$$\mathbf{U}(\mathbf{v}) := {}_{\|\cdot\|}\bigotimes_{j=1}^d U_j^{\min}(\mathbf{v})$$

is of interest (cf. Lemma 4.34). In §6.6 we shall discuss the meaning of $U_j^{\min}(\mathbf{v})$ and $\mathbf{U}(\mathbf{v})$ for non-algebraic tensors.

Lemma 6.22. *Assume (6.18). (a) Let* $\mathbf{v} \in {}_a\bigotimes_{j=1}^d V_j$*. Then the space of linear functionals* V_ν' *may be replaced in (6.10) by* continuous *linear functionals from* V_ν^**:*

$$\begin{aligned} U_j^{\min}(\mathbf{v}) &= \overline{\operatorname{span}\left\{ \begin{matrix} \left(\varphi^{(1)} \otimes \ldots \otimes \varphi^{(j-1)} \otimes id \otimes \varphi^{(j+1)} \otimes \ldots \otimes \varphi^{(d)}\right)(\mathbf{v}), \\ where \;\; \varphi^{(\nu)} \in V_\nu^* \;\; for \; \nu \in \{1,\ldots,d\}\backslash\{j\} \end{matrix} \right\}} \\ &= \overline{\left\{\varphi(\mathbf{v}) : \; \varphi \in {}_a\bigotimes\nolimits_{k\in\{1,\ldots,d\}\backslash\{j\}} V_k^* \;\right\}} \\ &= \overline{\left\{\varphi(\mathbf{v}) : \; \varphi \in \left({}_{\|\cdot\|}\bigotimes\nolimits_{k\in\{1,\ldots,d\}\backslash\{j\}} V_k\right)^*\right\}}. \end{aligned} \tag{6.19}$$

(b) If $\mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ *is not an algebraic tensor, take (6.19) as a definition of* $U_j^{\min}(\mathbf{v})$*. Note that in this case* $U_j^{\min}(\mathbf{v})$ *may be infinite dimensional.*

Proof. 1) By Lemma 4.97, the mapping $\varphi^{(1)}\otimes\ldots\otimes\varphi^{(j-1)}\otimes id\otimes\varphi^{(j+1)}\otimes\ldots\otimes\varphi^{(d)}$ is continuous on $({}_\vee\bigotimes_{j=1}^d V_j, \|\cdot\|_\vee)$, where $\|\cdot\|_\vee$ may also be replaced by a stronger norm. By Lemma 6.12b, V_ν' may be replaced by V_ν^*.

2) Let $U_j^{\min}(\mathbf{v})$ be defined by the first or second line in (6.19) and denote the subspace of the third line by $\hat{U}_j^{\min}(\mathbf{v})$. Since

$${}_a\bigotimes\nolimits_{k\in\{1,\ldots,d\}\backslash\{j\}} V_k^* \subset \left({}_{\|\cdot\|}\bigotimes\nolimits_{k\in\{1,\ldots,d\}\backslash\{j\}} V_k \right)^*,$$

$U_j^{\min}(\mathbf{v}) \subset \hat{U}_j^{\min}(\mathbf{v})$ holds. Assuming that $\hat{U}_j^{\min}(\mathbf{v})$ is strictly larger than $U_j^{\min}(\mathbf{v})$, there is some $\psi \in ({}_{\|\cdot\|}\bigotimes_{k\neq j} V_k)^*$ and $u := \psi(\mathbf{v}) \in \hat{U}_j^{\min}(\mathbf{v})$ such that $u \notin U_j^{\min}(\mathbf{v})$. By Hahn-Banach, there is a functional $\varphi^{(j)} \in V_j^*$ with $\varphi^{(j)}(u) \neq 0$ and $\varphi^{(j)}|_{U_j^{\min}(\mathbf{v})} = 0$. The tensor $\mathbf{w} := \varphi^{(j)}(\mathbf{v}) \in {}_{\|\cdot\|}\bigotimes_{k\neq j} V_k$ does not vanish, since

$$\psi(\mathbf{w}) = \left(\varphi^{(j)} \otimes \psi\right)(\mathbf{v}) = \varphi^{(j)}(\psi(\mathbf{v})) = \varphi^{(j)}(u) \neq 0.$$

Hence, the definition of $\|\mathbf{w}\|_\vee > 0$ implies that there is an elementary tensor $\varphi^{[j]} = \bigotimes_{k\neq j} \varphi^{(k)}$ with $|\varphi^{[j]}(\mathbf{w})| > 0$. Set $\varphi := \varphi^{(j)} \otimes \varphi^{[j]} = \bigotimes_{k=1}^d \varphi^{(k)}$. Now,

$$\varphi(\mathbf{v}) = \varphi^{[j]}(\varphi^{(j)}(\mathbf{v})) = \varphi^{[j]}(\mathbf{w}) \neq 0$$

is a contradiction to $\varphi(\mathbf{v}) = \varphi^{(j)}(\varphi^{[j]}(\mathbf{v})) = 0$ because of $\varphi^{[j]}(\mathbf{v}) \in U_j^{\min}(\mathbf{v})$ and $\varphi^{(j)}|_{U_j^{\min}(\mathbf{v})} = 0$. Hence, $U_j^{\min}(\mathbf{v}) = \hat{U}_j^{\min}(\mathbf{v})$ is valid.

3) Since $\varphi \in {}_a\bigotimes_{k\neq j} V_k^*$ is continuous, any $\mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$, defined by $\mathbf{v} = \lim \mathbf{v}_n$, has a well-defined limit $\varphi(\mathbf{v}) = \lim \varphi(\mathbf{v}_n)$. □

Lemma 6.23. *Assume (6.18). For all* $\mathbf{v}_n, \mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ *with* $\mathbf{v}_n \rightharpoonup \mathbf{v}$*, we have*

$$\varphi^{[j]}(\mathbf{v}_n) \rightharpoonup \varphi^{[j]}(\mathbf{v}) \textit{ in } V_j \qquad \textit{for all } \varphi^{[j]} \in {}_a\bigotimes_{k\neq j} V_k^* \;.$$

Proof. Let $\varphi^{[j]} = \bigotimes_{k\in\{1,\ldots,d\}\setminus\{j\}} \varphi^{(k)}$ with $\varphi^{(k)} \in V_k^*$ be an elementary tensor. In order to prove $\varphi^{[j]}(\mathbf{v}_n) \rightharpoonup \varphi^{[j]}(\mathbf{v})$, we have to show for all $\varphi^{(j)} \in V_j^*$ that

$$\varphi^{(j)}(\varphi^{[j]}(\mathbf{v}_n)) \to \varphi^{(j)}(\varphi^{[j]}(\mathbf{v})). \tag{6.20}$$

However, the composition $\varphi^{(j)} \circ \varphi^{[j]} = \bigotimes_{k=1}^d \varphi^{(k)}$ belongs to $({}_\vee\bigotimes_{k=1}^d V_k)^*$ and because of (6.18) also to $\mathbf{V}^* = ({}_{\|\cdot\|}\bigotimes_{k=1}^d V_k)^*$. Hence $\mathbf{v}_n \rightharpoonup \mathbf{v}$ implies (6.20) and proves the assertion for an elementary tensor $\varphi^{[j]}$. The result extends immediately to finite linear combinations $\varphi^{[j]} \in {}_a\bigotimes_{k\neq j} V_k^*$. □

Theorem 6.24. *Assume (6.18). If* $\mathbf{v}_n \in {}_a\bigotimes_{j=1}^d V_j$ *satisfies* $\mathbf{v}_n \rightharpoonup \mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$, *then*

$$\boxed{\dim(U_j^{\min}(\mathbf{v})) \le \liminf_{n\to\infty} \dim(U_j^{\min}(\mathbf{v}_n)) \qquad \text{for all } 1 \le j \le d.}$$

Proof. Choose a subsequence (again denoted by $\mathbf{v}_n$) such that $\dim(U_j^{\min}(\mathbf{v}_n))$ is weakly increasing. In the case of $\dim(U_j^{\min}(\mathbf{v}_n)) \to \infty$, nothing is to be proved. Therefore, let $\lim \dim(U_j^{\min}(\mathbf{v}_n)) = N < \infty$. For an indirect proof assume that $\dim(U_j^{\min}(\mathbf{v})) > N$. Since $\{\varphi(\mathbf{v}) : \varphi \in {}_a\bigotimes_{k\neq j} V_k^*\}$ is dense in $U_j^{\min}(\mathbf{v}_n)$ (cf. Lemma 6.22), there are $N+1$ linearly independent vectors

$$b^{(i)} = \varphi_i^{[j]}(\mathbf{v}) \qquad \text{with } \varphi_i^{[j]} \in {}_a\bigotimes_{k\neq j} V_k^* \qquad \text{for } 1 \le i \le N+1.$$

By Lemma 6.23, weak convergence $b_n^{(i)} := \varphi_i^{[j]}(\mathbf{v}_n) \rightharpoonup b^{(i)}$ holds. By Lemma 4.24, for large enough n, also $(b_n^{(i)} : 1 \le i \le N+1)$ is linearly independent. Because of $b_n^{(i)} = \varphi_i^{[j]}(\mathbf{v}_n) \in U_j^{\min}(\mathbf{v}_n)$, this contradicts $\dim(U_j^{\min}(\mathbf{v}_n)) \le N$. □

If the spaces $U_j^{\min}(\mathbf{v})$ and $U_j^{\min}(\mathbf{v}_n)$ are infinite dimensional, one may ask whether they have different (infinite) cardinalities. The proof of the next remark shows that $U_j^{\min}(\mathbf{v})$ is the completion of a space of dimension $\le \aleph_0 = \#\mathbb{N}$.

Remark 6.25. Even if the tensor space $\mathbf{V}$ is nonseparable, the minimal subspaces $U_j^{\min}(\mathbf{v})$ are separable.

Proof. Let $\mathbf{v}_i \to \mathbf{v}$ be a converging sequence with algebraic tensors $\mathbf{v}_i \in \mathbf{V}_{\mathrm{alg}}$. The subspace $U_i^{(j)} := U_j^{\min}(\mathbf{v}_i)$ is finite dimensional. There is a sequence of basis elements b_ν and integers $n_m^{(j)} \in \mathbb{N}$ such that

$$S_m^{(j)} := \sum_{i=1}^m U_i^{(j)} = \operatorname{span}\{b_\nu : 1 \le \nu \le n_m^{(j)}\}.$$

The spaces $S^{(j)} := \sum_{i=1}^\infty U_i^{(j)} = \operatorname{span}\{b_\nu : \nu \in \mathbb{N}\}$ are separable and satisfy $\mathbf{v} \in \mathbf{S} := {}_{\|\cdot\|}\bigotimes_{j=1}^d S^{(j)}$. By Remark 4.35, $\mathbf{S}$ is separable. Because of minimality, the inclusion $U_j^{\min}(\mathbf{v}) \subset \overline{S^{(j)}}$ holds and proves the assertion. □

6.6 Minimal Subspaces of Topological Tensors

6.6.1 Interpretation of $U_j^{\min}(\mathbf{v})$

In Lemma 6.22b, under condition (6.18), we have defined $U_j^{\min}(\mathbf{v})$ also for topological tensors from ${}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$. Accordingly, we may define the Banach subspace

$$\mathbf{U}(\mathbf{v}) := {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v}) . \tag{6.21}$$

For *algebraic* tensors we know that[3] $\mathbf{v} \in \mathbf{U}(\mathbf{v})$. However, the corresponding conjecture $\mathbf{v} \in \mathbf{U}(\mathbf{v})$ for *topological* tensors turns out to be not quite obvious.[4]

We discuss the property $\mathbf{v} \in \mathbf{U}(\mathbf{v})$ in three different cases.

1. The case of $\dim(U_j^{\min}(\mathbf{v})) < \infty$ is treated in §6.6.2. From the practical viewpoint, this is the most important case. If $\dim(U_j^{\min}(\mathbf{v})) < \infty$ results from Theorem 6.24, we would like to know whether the (weak) limit $\mathbf{v}$ satisfies[3] $\mathbf{v} \in \mathbf{U}(\mathbf{v}) = {}_a\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$, which implies that $\mathbf{v}$ is in fact an algebraic tensor.
2. The general Banach case is studied in §6.6.3. We give a proof for $\mathbf{v} = \lim \mathbf{v}_n$, provided that the convergence is fast enough (cf. (4.12)) or that the minimal subspaces are Grassmannian (cf. Definition 4.4).
3. In the Hilbert case, a positive answer can be given (see §6.6.4).

6.6.2 Case of $\dim(U_j^{\min}(\mathbf{v})) < \infty$

Theorem 6.26. *Assume (6.18) and* $\dim(U_j^{\min}(\mathbf{v})) < \infty$ *for* $\mathbf{v} \in \mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$. *Then* $\mathbf{v}$ *belongs to the* algebraic *tensor space* $\mathbf{U}(\mathbf{v}) = {}_a\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$.

[3] Since $\dim(U_j^{\min}(\mathbf{v})) < \infty$, ${}_a\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v}) = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v}) = \mathbf{U}(\mathbf{v})$ holds.

[4] To repeat the proof of the existence of minimal subspaces, we need the counterpart of (6.3) which would be

$$\left(X_1 \otimes_{\|\cdot\|} X_2\right) \cap \left(Y_1 \otimes_{\|\cdot\|} Y_2\right) = (X_1 \cap Y_1) \otimes_{\|\cdot\|} (X_2 \cap Y_2) .$$

Again, $(X_1 \cap Y_1) \otimes_{\|\cdot\|} (X_2 \cap Y_2) \subset \left(X_1 \otimes_{\|\cdot\|} X_2\right) \cap \left(Y_1 \otimes_{\|\cdot\|} Y_2\right)$ is a trivial statement. For the opposite direction one should have that

$$\left(X_1 \otimes_{\|\cdot\|} X_2\right) \cap \left(Y_1 \otimes_{\|\cdot\|} Y_2\right) = \left(\overline{X_1 \otimes_a X_2}\right) \cap \left(\overline{Y_1 \otimes_a Y_2}\right)$$

is a subset of

$$\overline{(X_1 \otimes_a X_2) \cap (Y_1 \otimes_a Y_2)} \underset{(6.3)}{=} \overline{(X_1 \cap Y_1) \otimes_a (X_2 \cap Y_2)} = (X_1 \cap Y_1) \otimes_{\|\cdot\|} (X_2 \cap Y_2) .$$

However, the closure and the intersection of sets satisfy the rule $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$, whereas the previous argument requires the reverse inclusion.

The underlying difficulty is that a topological tensor $\mathbf{v} \in \mathbf{V}$ is defined as limit of some sequence $\mathbf{v}_n \in {}_a\bigotimes_{j=1}^{d} V_j$ and that the statement $\mathbf{v} \in \mathbf{U}(\mathbf{v})$ requires to prove the existence of another sequence $\mathbf{u}_n \in {}_a\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$ with $\mathbf{u}_n \to \mathbf{v}$.

Proof. 1) Let $\{b_i^{(j)} : 1 \leq i \leq r_j\}$ be a basis of $U_j^{\min}(\mathbf{v})$. There is a dual system $\varphi_i^{(j)} \in V_j^*$ with the property $\varphi_i^{(j)}(b_k^{(j)}) = \delta_{ik}$. Define $\mathbf{a_i} := \bigotimes_{j=1}^d \varphi_{i_j}^{(j)} \in {}_a\bigotimes_{j=1}^d V_j^*$ and $\mathbf{b_i} := \bigotimes_{j=1}^d b_{i_j}^{(j)} \in \mathbf{U}$ for $\mathbf{i} = (i_1, \ldots, i_d)$ with $1 \leq i_j \leq r_j$. Any $\mathbf{u} \in \mathbf{U}$ is reproduced by

$$\mathbf{u} = \sum\nolimits_{\mathbf{i}} \mathbf{a_i}(\mathbf{u})\mathbf{b_i}.$$

We set

$$\mathbf{u_v} := \sum_{\mathbf{i}} \mathbf{a_i}(\mathbf{v})\mathbf{b_i} \in {}_a\bigotimes_{j=1}^d U_j^{\min}(\mathbf{v}) \tag{6.22a}$$

and want to prove that $\mathbf{v} = \mathbf{u_v} \in {}_a\bigotimes_{j=1}^d U_j^{\min}(\mathbf{v})$.

2) The norm $\|\mathbf{v} - \mathbf{u_v}\|_\vee$ is defined by means of $\boldsymbol{\alpha}(\mathbf{v} - \mathbf{u_v})$ with normalised functionals $\boldsymbol{\alpha} = \bigotimes_{j=1}^d \alpha^{(j)} \in {}_a\bigotimes_{j=1}^d V_j^*$ (cf. (4.47)). If we can show

$$\boldsymbol{\alpha}(\mathbf{v} - \mathbf{u_v}) = 0 \qquad \text{for all } \boldsymbol{\alpha} = \bigotimes_{j=1}^d \alpha^{(j)} \in {}_a\bigotimes_{j=1}^d V_j^* \,, \tag{6.22b}$$

the norm $\|\mathbf{v} - \mathbf{u_v}\|_\vee$ vanishes and $\mathbf{v} = \mathbf{u_v}$ is proved. The proof of (6.22b) is given in the next part.

3) Write $\alpha^{(j)} = \alpha_0^{(j)} + \sum_i c_i \varphi_i^{(j)}$ with $c_i := \alpha^{(j)}(b_i^{(j)})$ and $\alpha_0^{(j)} := \alpha^{(j)} - \sum_i c_i \varphi_i^{(j)}$. It follows that $\alpha_0^{(j)}(b_i^{(j)}) = 0$ for all i, i.e.,

$$\alpha_0^{(j)}(u^{(j)}) = 0 \qquad \text{for all } u^{(j)} \in U_j^{\min}(\mathbf{v}). \tag{6.22c}$$

We expand the product into

$$\boldsymbol{\alpha} = \bigotimes_{j=1}^d \alpha^{(j)} = \bigotimes_{j=1}^d \left(\alpha_0^{(j)} + \sum_i c_i \varphi_i^{(j)} \right) = \bigotimes_{j=1}^d \left(\sum_i c_i \varphi_i^{(j)} \right) + R,$$

where all products in R contain at least one factor $\alpha_0^{(j)}$. Consider such a product in R, where, without loss of generality, we assume that $\alpha_0^{(j)}$ appears for $j = 1$, i.e., $\alpha_0^{(1)} \otimes \boldsymbol{\gamma}^{[1]}$ with $\boldsymbol{\gamma}^{[1]} \in {}_a\bigotimes_{j=2}^d V_j^*$. We conclude that $(\alpha_1^{(0)} \otimes \boldsymbol{\gamma}^{[1]})(\mathbf{u_v}) = 0$, since $(\alpha_0^{(1)} \otimes id \otimes \ldots \otimes id)(\mathbf{u_v}) = 0$ and $\alpha_0^{(1)} \otimes \boldsymbol{\gamma}^{[1]} = \boldsymbol{\gamma}^{[1]} \circ (\alpha_0^{(1)} \otimes id \otimes \ldots \otimes id)$. Furthermore,

$$\left(\alpha_0^{(1)} \otimes \boldsymbol{\gamma}^{[1]} \right)(\mathbf{v}) = \alpha_0^{(1)}(w) \qquad \text{for } w := (id \otimes \boldsymbol{\gamma}^{[1]})(\mathbf{v}).$$

By definition of $U_1^{\min}(\mathbf{v})$, $w \in U_1^{\min}(\mathbf{v})$ holds and $\alpha_0^{(1)}(w) = (\alpha_0^{(1)} \otimes \boldsymbol{\gamma}^{[1]})(\mathbf{v}) = 0$ follows from (6.22c). Together, $(\alpha_0^{(1)} \otimes \boldsymbol{\gamma}^{[1]})(\mathbf{v} - \mathbf{u_v}) = 0$ is shown. Since this statement holds for all terms in R, we obtain $R(\mathbf{v} - \mathbf{u_v}) = 0$.

It remains to analyse $\left(\bigotimes_{j=1}^d \left(\sum_i c_i \varphi_i^{(j)} \right) \right)(\mathbf{v} - \mathbf{u_v}) = \left(\sum_{\mathbf{i}} \mathbf{c_i a_i} \right)(\mathbf{v} - \mathbf{u_v})$

with $\mathbf{c_i} := \prod_{j=1}^{d} c_{i_j}$. Application to $\mathbf{u_v}$ yields

$$\left(\sum\nolimits_{\mathbf{i}} \mathbf{c_i a_i}\right)(\mathbf{u_v}) = \sum\nolimits_{\mathbf{i}} \mathbf{c_i a_i}(\mathbf{v}) \in \mathbb{K}$$

(cf. (6.22a)). Since this value coincides with $\left(\sum_{\mathbf{i}} \mathbf{c_i a_i}\right)(\mathbf{v}) = \sum_{\mathbf{i}} \mathbf{c_i a_i}(\mathbf{v})$, we have demonstrated that

$$\left(\bigotimes_{j=1}^{d}\left(\sum_{i} c_i \varphi_i^{(j)}\right)\right)(\mathbf{v} - \mathbf{u_v}) = 0.$$

Altogether we have proved (6.22b), which implies the assertion of the theorem. □

6.6.3 The General Banach Space Case

For the next theorem we need a further assumption on the norm $\|\cdot\|$. A sufficient condition is that $\|\cdot\|$ is a uniform crossnorm (cf. §4.3.1). The uniform crossnorm property implies that $\|\cdot\|$ is a reasonable crossnorm (cf. Lemma 4.79). Hence, condition (6.18) is ensured (cf. Proposition 4.68).

The proof of the theorem requires that the speed of convergence is fast enough. Since there are several cases, where exponential convergence $\mathbf{v}_n \to \mathbf{v}$ holds, the condition from below is not too restrictive. Here, the index n in $\mathbf{v}_n$ refers to the tensor rank or, more weakly, to the j-rank for some $1 \le j \le d$. In Theorem 6.27 we take $j = d$. Another more abstract criterion used in Theorem 6.29 requires instead that the minimal subspaces are Grassmannian.

Theorem 6.27. *Assume that* $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ *is a Banach tensor space with a uniform crossnorm* $\|\cdot\|$. *If* $\mathbf{v} \in \mathbf{V}$ *is the limit of* $\mathbf{v}_n = \sum_{i=1}^{n} \mathbf{v}_{i,n}^{[d]} \otimes v_{i,n}^{(d)} \in {}_a\bigotimes_{j=1}^{d} V_j$ *with the rate*[5]

$$\|\mathbf{v}_n - \mathbf{v}\| \le o(n^{-3/2}),$$

then $\mathbf{v} \in \mathbf{U}(\mathbf{v}) = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$, *i.e.,* $\mathbf{v} = \lim \mathbf{u}_n$ *with* $\mathbf{u}_n \in {}_a\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$.

Proof. We use the setting $\mathbf{V}_{\mathrm{alg}} = \mathbf{X}_{d-1} \otimes_a V_d$ from Proposition[6] 4.90 and rewrite the norm on $\mathbf{X}_{d-1}$ by $\|\cdot\|_{[d]} := \|\cdot\|_{\mathbf{X}_{d-1}}$. Thus, each $\mathbf{v}_n \in \mathbf{V}_{\mathrm{alg}}$ has a representation in $U_{[d]}^{\min}(\mathbf{v}_n) \otimes U_d^{\min}(\mathbf{v}_n)$ with $U_{[d]}^{\min}(\mathbf{v}_n) \subset \mathbf{X}_{d-1}$, $U_d^{\min}(\mathbf{v}_n) \subset V_d$, and $r := \dim U_{[d]}^{\min}(\mathbf{v}_n) = \dim U_d^{\min}(\mathbf{v}_n) \le n$. Renaming r by n, we obtain the

[5] The condition can be weakened to $o(n^{-1-|1/2-1/p|})$ if (4.7) applies.

[6] Differently from the setting in Proposition 4.90, we define $\mathbf{X}_{d-1} = {}_a\bigotimes_{j=1}^{d} V_j$ as algebraic tensor space equipped with the norm $\|\cdot\|_{\mathbf{X}_{d-1}}$. This does not change the statement of the proposition because of Lemma 4.34.

representation $\mathbf{v}_n = \sum_{i=1}^{n} \mathbf{v}_{i,n}^{[d]} \otimes v_{i,n}^{(d)}$. According to Corollary 6.8c, we can fix any basis $\{v_i^{(d)}\}$ of $U_d^{\min}(\mathbf{v}_n)$ and recover

$$\mathbf{v}_n = \sum_{i=1}^{n} \psi_i^{(d)}(\mathbf{v}_n) \otimes v_i^{(d)}$$

from a dual basis $\{\psi_i^{(d)}\}$. Here we use Remark 3.54, i.e., $\psi_i^{(d)}$ is the abbreviation for $id \otimes \ldots \otimes id \otimes \psi_i^{(d)} \in \mathcal{L}(\mathbf{V}_{\text{alg}}, \mathbf{X}_{d-1})$. We choose $v_i^{(d)}$ and $\psi_i^{(d)}$ according to Lemma 4.17 with $\|v_i^{(d)}\|_d = \|\psi_i^{(d)}\|_d^* = 1$ and define

$$\mathbf{u}_n^I := \sum_{i=1}^{n} \psi_i^{(d)}(\mathbf{v}) \otimes v_i^{(d)} \in U_{[d]}^{\min}(\mathbf{v}) \otimes_a V_d.$$

The triangle inequality yields

$$\begin{aligned} \|\mathbf{u}_n^I - \mathbf{v}_n\| &= \left\| \sum_{i=1}^{n} \left(\psi_i^{(d)}(\mathbf{v}) - \psi_i^{(d)}(\mathbf{v}_n) \right) \otimes v_i^{(d)} \right\| \qquad (6.23a) \\ &= \left\| \sum_{i=1}^{n} \psi_i^{(d)}(\mathbf{v} - \mathbf{v}_n) \otimes v_i^{(d)} \right\| \le \sum_{i=1}^{n} \left\| \psi_i^{(d)}(\mathbf{v} - \mathbf{v}_n) \otimes v_i^{(d)} \right\| \\ &\underset{(4.44a)}{=} \sum_{i=1}^{n} \left\| \psi_i^{(d)}(\mathbf{v} - \mathbf{v}_n) \right\|_{[d]} \underbrace{\|v_i^{(d)}\|_d}_{=1} \\ &\underset{(4.45)}{\le} \sum_{i=1}^{n} \underbrace{\|\psi_i^{(d)}\|_d^*}_{=1} \|\mathbf{v} - \mathbf{v}_n\| = n\|\mathbf{v} - \mathbf{v}_n\|. \end{aligned}$$

Note that

$$\mathbf{u}_n^I \in \mathbf{U}_{[d],n} \otimes V_d \text{ with } \mathbf{U}_{[d],n} := \operatorname{span}\left\{\psi_i^{(d)}(\mathbf{v}) : 1 \le i \le n\right\} \subset \mathbf{U}_{[d]}^{\min}(\mathbf{v}),$$

where $\dim \mathbf{U}_{[d],n} \le n$.

Again by Lemma 4.17, we can choose a basis $\{\mathbf{v}_i^{[d]}\}_{i=1}^n$ of $\mathbf{U}_{[d]}^{\min}(\mathbf{v}_n) \subset \mathbf{X}_{d-1}$ and a corresponding dual system $\{\boldsymbol{\chi}_i^{[d]}\}_{i=1}^n \subset \mathbf{X}_{d-1}^*$. An analogous proof shows that

$$\mathbf{u}_n^{II} := \sum_{i=1}^{n} \mathbf{v}_i^{[d]} \otimes \boldsymbol{\chi}_i^{[d]}(\mathbf{v}) \in \mathbf{X}_{d-1} \otimes_a U_{d,n}$$

satisfies the estimate

$$\|\mathbf{u}_n^{II} - \mathbf{v}_n\| \le n\|\mathbf{v} - \mathbf{v}_n\|, \qquad (6.23b)$$

where $U_{d,n} := \operatorname{span}\{\boldsymbol{\chi}_i^{[d]}(\mathbf{v}) : 1 \le i \le n\} \subset U_d^{\min}(\mathbf{v})$. We choose the projection Φ_d onto the subspace $U_{d,n}$ according to Theorem 4.14. We denote $id \otimes \ldots \otimes id \otimes \Phi_d$

again by Φ_d and define

$$\mathbf{u}_n := \Phi_d(\mathbf{u}_n^I) \in U_{[d],n} \otimes_a U_{d,n} \subset U_{[d]}^{\min}(\mathbf{v}) \otimes_a U_d^{\min}(\mathbf{v}) \underset{(6.12)}{\subset} {}_a\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v}) \,.$$

The uniform crossnorm property (4.42) with $A_j = id$ $(1 \le j \le d-1)$ and $A_d = \Phi_d$ implies the estimate $\|\Phi_d\|_{\mathbf{V}\leftarrow\mathbf{V}} = \|\Phi_d\|_{V_d \leftarrow V_d} \le \sqrt{n}$, where the latter bound is given by Theorem 4.14 because of $\dim(U_{d,n}) \le n$. Since $\Phi_d\big(\mathbf{u}_n^{II}\big) = \mathbf{u}_n^{II}$, the estimates

$$\begin{aligned} \|\Phi_d(\mathbf{v}_n) - \mathbf{u}_n^{II}\| &= \|\Phi_d\left(\mathbf{v}_n - \mathbf{u}_n^{II}\right)\| \le \sqrt{n}\|\mathbf{v}_n - \mathbf{u}_n^{II}\| \underset{(6.23b)}{\le} n^{3/2}\|\mathbf{v} - \mathbf{v}_n\|, \\ \|\mathbf{u}_n - \Phi_d(\mathbf{v}_n)\| &= \|\Phi_d\left(\mathbf{u}_n^{I} - \mathbf{v}_n\right)\| \le \sqrt{n}\|\mathbf{u}_n^{I} - \mathbf{v}_n\| \underset{(6.23a)}{\le} n^{3/2}\|\mathbf{v} - \mathbf{v}_n\| \end{aligned}$$

are valid. Altogether, we get the estimate

$$\begin{aligned} \|\mathbf{u}_n - \mathbf{v}\| &= \|\,[\mathbf{u}_n - \Phi_d(\mathbf{v}_n)] + \left[\Phi_d(\mathbf{v}_n) - \mathbf{u}_n^{II}\right] + \left[\mathbf{u}_n^{II} - \mathbf{v}_n\right] + [\mathbf{v}_n - \mathbf{v}]\,\| \\ &\le \left(2n^{3/2} + n + 1\right)\|\mathbf{v} - \mathbf{v}_n\|. \end{aligned}$$

The assumption $\|\mathbf{v} - \mathbf{v}_n\| \le o(n^{-3/2})$ implies $\|\mathbf{u}_n - \mathbf{v}\| \to 0$. □

A second criterion[7] for $\mathbf{v} \in \mathbf{U}(\mathbf{v})$ makes use of the Grassmannian $\mathbb{G}(\cdot)$ from Definition 4.4.

Lemma 6.28. *Let $U_j \in \mathbb{G}(V_j)$ for $1 \le j \le d$. Assume that $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ is a Banach tensor space with a uniform crossnorm $\|\cdot\|$. Then the following intersection property holds:*

$$\bigcap_{1\le j\le d} \left(U_j \otimes_{\|\cdot\|} \mathbf{V}_{[j]}\right) = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j \;, \quad \textit{where } \mathbf{V}_{[j]} := {}_{\|\cdot\|}\bigotimes_{k\ne j} V_k$$

Proof. Since induction can be used (cf. Lemma 6.11), we consider only the case $d = 2$. The result for the algebraic tensor spaces implies that

$$\begin{aligned} U_1 \otimes_{\|\cdot\|} U_2 &= \overline{U_1 \otimes_a U_2} = \overline{(U_1 \otimes_a V_2) \cap (V_1 \otimes_a U_2)} \\ &\subset \overline{U_1 \otimes_a V_2} \cap \overline{V_1 \otimes_a U_2} = \left(U_1 \otimes_{\|\cdot\|} V_2\right) \cap \left(V_1 \otimes_{\|\cdot\|} U_2\right) \end{aligned}$$

because of the general rule $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$. It remains to prove the opposite inclusion

$$\left(U_1 \otimes_{\|\cdot\|} V_2\right) \cap \left(V_1 \otimes_{\|\cdot\|} U_2\right) \subset U_1 \otimes_{\|\cdot\|} U_2.$$

For any $\mathbf{v} \in U_1 \otimes_{\|\cdot\|} V_2$ there is a sequence $\mathbf{v}_n \in U_1 \otimes_a V_2$ with $\mathbf{v}_n \to \mathbf{v}$. The projection $P_1 \in \mathcal{L}(V_1, V_1)$ from Lemma 4.13 onto U_1 satisfies $(P_1 \otimes id)\,\mathbf{v}_n = \mathbf{v}_n$.

[7] This approach is communicated to the author by A. Falcó.

By the uniform crossnorm property, $P_1 \otimes id$ is continuous on $\mathbf{V} = V_1 \otimes_{\|\cdot\|} V_2$ so that

$$\mathbf{v} = \lim \mathbf{v}_n = \lim \left(P_1 \otimes id\right) \mathbf{v}_n = \left(P_1 \otimes id\right) \left(\lim \mathbf{v}_n\right) = \left(P_1 \otimes id\right) \mathbf{v}.$$

Analogously, $\mathbf{v} = (id \otimes P_2)\, \mathbf{v}$ holds. $(P_1 \otimes id)$ commutes with $(id \otimes P_2)$ and yields the product $P_1 \otimes P_2 = (P_1 \otimes id) \circ (id \otimes P_2)$. This proves $\mathbf{v} = (P_1 \otimes P_2)\, \mathbf{v}$. Note that $\mathbf{u}_n = (P_1 \otimes P_2)\, \mathbf{v}_n \in U_1 \otimes_a U_2$ and $\mathbf{v} = \lim \mathbf{u}_n$, i.e., $\mathbf{v} \in U_1 \otimes_{\|\cdot\|} U_2$, which implies the desired reverse inclusion. □

Theorem 6.29. *Assume that* $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ *is a Banach tensor space with a uniform crossnorm* $\|\cdot\|$. *For* $\mathbf{v} \in \mathbf{V}$ *assume that* $U_j^{\min}(\mathbf{v}) \in \mathbb{G}(V_j)$ *for* $1 \leq j \leq d$. *Then*

$$\mathbf{v} \in \mathbf{U}(\mathbf{v}) := {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$$

holds.

Proof. Set $U_j := U_j^{\min}(\mathbf{v})$. According to Lemma 6.28 we have to show that $\mathbf{v} \in U_j \otimes_{\|\cdot\|} \mathbf{V}_{[j]}$ for all $1 \leq j \leq d$. Let $I_j \in \mathcal{L}(V_j, V_j)$ and $\mathbf{I}_{[j]} \in \mathcal{L}(\mathbf{V}_{[j]}, \mathbf{V}_{[j]})$ be the identity mappings. We split $\mathbf{v} = \left(I_j \otimes \mathbf{I}_{[j]}\right) \mathbf{v}$ into

$$\mathbf{v} = \left(P_j \otimes \mathbf{I}_{[j]}\right) \mathbf{v} + \left((I_j - P_j) \otimes \mathbf{I}_{[j]}\right) \mathbf{v} \tag{6.24}$$

with $P_j \in \mathcal{L}(V_j, V_j)$ as in the proof of Lemma 6.28. For an indirect proof we assume that $\mathbf{d} := \left((I_j - P_j) \otimes \mathbf{I}_{[j]}\right) \mathbf{v} \neq 0$. By Lemma 4.79, $\|\cdot\|$ is a reasonable crossnorm. Therefore, the injective norm $\|\mathbf{d}\|_{\vee}$ is defined, which is the supremum of all $|(\varphi_j \otimes \boldsymbol{\varphi}_{[j]})\mathbf{d}|$ with normalised φ_j and $\boldsymbol{\varphi}_{[j]} = \bigotimes_{k \neq j} \varphi_k$. Write $\varphi_j \otimes \boldsymbol{\varphi}_{[j]}$ as $\varphi_j \circ (I_j \otimes \boldsymbol{\varphi}_{[j]})$ and note that

$$(I_j \otimes \boldsymbol{\varphi}_{[j]})\mathbf{d} = \Big((I_j \otimes \boldsymbol{\varphi}_{[j]}) \circ \left((I_j - P_j) \otimes \mathbf{I}_{[j]}\right)\Big)\, \mathbf{v} = \Big((I_j - P_j) \circ (I_j \otimes \boldsymbol{\varphi}_{[j]})\Big)\, \mathbf{v}.$$

By definition of $U_j^{\min}(\mathbf{v})$, $(I_j \otimes \boldsymbol{\varphi}_{[j]})\mathbf{v} \in U_j$ holds proving

$$(I_j - P_j)\left((I_j \otimes \boldsymbol{\varphi}_{[j]})\mathbf{v}\right) = 0.$$

This shows that $\|\mathbf{d}\|_{\vee} = 0$; hence, $\mathbf{d} = 0$. From (6.24), we conclude that $\mathbf{v} = \left(P_j \otimes \mathbf{I}_{[j]}\right) \mathbf{v} \in U_j \otimes_{\|\cdot\|} \mathbf{V}_{[j]}$. □

The assumption of a uniform crossnorm can be weakened. Instead, we require that tensor products of *projections* are uniformly bounded:

$$\|\mathbf{P}\mathbf{v}\| \leq C_P \, \|\mathbf{v}\| \qquad \begin{cases} \text{for all } \mathbf{v} \in \mathbf{V} \text{ and} \\ \text{all } \mathbf{P} = \bigotimes_{j=1}^{d} P_j, \; P_j \in \mathcal{L}(V_j, V_j) \text{ projection.} \end{cases} \tag{6.25}$$

The reason is that the proof from above involves only projections. Also in this case the norm is not weaker than the injective norm.

Remark 6.30. A crossnorm $\|\cdot\|$ satisfying condition (6.25) fulfils $\|\cdot\|_{\vee} \leq C_P \|\cdot\|$ (cf. (4.33)).

Proof. Note that the proof of Lemma 4.79 uses projections. □

6.6.4 Hilbert Spaces

Since in Hilbert spaces every closed subspace U is complemented: $V_j = U \oplus U^{\perp}$, Theorem 6.29 yields the following result.

Theorem 6.31. *For Hilbert spaces V_j let $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ be the Hilbert tensor space with the induced scalar product. Then $\mathbf{v} \in \mathbf{U}(\mathbf{v})$ holds for all $\mathbf{v} \in \mathbf{V}$ with $\mathbf{U}(\mathbf{v})$ from (6.21).*

6.7 Minimal Subspaces for Intersection Spaces

Banach spaces which are intersection spaces (see §4.3.6) do not satisfy the basic assumption (6.18). Therefore, we have to check whether the previous results can be extended to this case. We recall the general setting. For each $1 \leq j \leq d$ we have a scale of spaces $V_j^{(n)}$, $0 \leq n \leq N_j$, which leads to tensor spaces

$$\mathbf{V}^{(\mathbf{n})} = {}_{\|\cdot\|_{\mathbf{n}}}\bigotimes_{j=1}^{d} V_j^{(n_j)} \qquad \text{for multi-indices } \mathbf{n} \in \mathcal{N} \subset \mathbb{N}_0^d,$$

where the subset $\mathcal{N}$ satisfies the conditions (4.51a-c). The final tensor subspace is

$$\mathbf{V}_{\text{top}} := \bigcap_{\mathbf{n} \in \mathcal{N}} \mathbf{V}^{(\mathbf{n})}$$

endowed with the intersection norm (4.52b).

The algebraic counterparts are denoted by

$$\mathbf{V}_{\text{alg}}^{(\mathbf{n})} := {}_{a}\bigotimes_{j=1}^{d} V_j^{(n_j)} \qquad \text{and} \qquad \mathbf{V}_{\text{alg}} := \bigcap_{\mathbf{n} \in \mathcal{N}} \mathbf{V}_{\text{alg}}^{(\mathbf{n})}.$$

There are different conclusions for algebraic and topological tensor spaces, which are presented in the next subsections.

6.7.1 Algebraic Tensor Space

All spaces $\mathbf{V}^{(\mathbf{n})}$ are dense subspaces of

$$\mathbf{V}^{(\mathbf{0})} = {}_{\|\cdot\|_0}\bigotimes_{j=1}^{d} V_j\ , \quad \text{where } V_j = V_j^{(0)}$$

(cf. Lemma 4.102). Here, we consider algebraic tensors from ${}_a\bigotimes_{j=1}^d V_j$. Each multi-index $\mathbf{n} \in \mathcal{N}$ defines another space $\mathbf{V}_{[j]}^{(\mathbf{n})} = {}_a\bigotimes_{k\in\{1,\ldots,d\}\backslash\{j\}} V_k^{(n_k)}$. Nevertheless, each $\mathbf{n}$ yields the same minimal subspace $U_j^{\min}(\mathbf{v})$.

Remark 6.32. For all $\mathbf{n} \in \mathbb{N}_0^d$ with $n_k \leq N_k$ and $\mathbf{v} \in \mathbf{V}_{\text{alg}}$, the minimal subspace is given by

$$U_j^{\min}(\mathbf{v}) = \left\{\varphi(\mathbf{v}) : \varphi \in (\mathbf{V}_{[j]}^{(\mathbf{n})})'\right\} \subset V_j^{(N_j)}. \tag{6.26}$$

Proof. By Proposition 4.104, $\mathbf{v} \in \mathbf{V}_{\text{alg}}^{(N_1,\ldots,N_d)} \subset \mathbf{V}_{\text{alg}}^{(\mathbf{n})}$ holds for all $\mathbf{n}$ with $n_k \leq N_k$. From $\mathbf{v} \in \mathbf{V}_{\text{alg}}^{(\mathbf{n})}$ we derive (6.26) and $\mathbf{v} \in \bigotimes_{j=1}^d U_j^{\min}(\mathbf{v})$. By $\mathbf{v} \in \mathbf{V}_{\text{alg}}^{(N_1,\ldots,N_d)}$ and minimality of $U_j^{\min}(\mathbf{v})$, the inclusion $U_j^{\min}(\mathbf{v}) \subset V_j^{(N_j)}$ follows. □

6.7.2 Topological Tensor Space

Remark 6.32 does not hold for non-algebraic tensors. A simple counter-example is $f \in C^1(I \times J)$ with $f(x,y) = F(x+y)$ and $F \notin C^2$. Choose the functional $\varphi = \delta_\eta' \in C^1(J)^*$. Then $\varphi(f)(x) = -F'(x+\eta) \in C^0(I)$, but $\varphi(f)$ is not in $C^1(I)$ in contrast to Remark 6.32.

While in Remark 6.32 we could take functionals from $\mathbf{V}_{[j]}^{(\mathbf{n})}$ for any $\mathbf{n}$ bounded by $n_k \leq N_k$, we now have to restrict the functionals to $\mathbf{n} = \mathbf{0}$. Because of the notation $V_k^{(0)} = V_k$, the definition coincides with the one in Lemma 6.22:

$$\begin{aligned} U_j^{\min}(\mathbf{v}) &:= \overline{\left\{\varphi(\mathbf{v}) : \varphi \in \left({}_a\bigotimes_{k\in\{1,\ldots,d\}\backslash\{j\}} V_k\right)^*\right\}}^{\|\cdot\|_\mathbf{0}} \\ &= \overline{\operatorname{span}\left\{\varphi(\mathbf{v}) : \varphi = \bigotimes_{k\in\{1,\ldots,d\}\backslash\{j\}} \varphi^{(k)},\ \varphi^{(k)} \in V_k^*\right\}}^{\|\cdot\|_\mathbf{0}}, \end{aligned} \tag{6.27}$$

where the completion is performed with respect to the norm $\|\cdot\|_\mathbf{0}$ of $\mathbf{V}_{[j]}^{(\mathbf{0})}$.

In the following we show that the same results can be derived as in the standard case. Condition (6.18) used before has to be adapted to the situation of the intersection space. Consider the tuples $\mathbf{N}_j = (0,\ldots,0,N_j,0,\ldots,0) \in \mathcal{N}$ from (4.51c) and the corresponding topological tensor spaces

$$\mathbf{V}^{(\mathbf{N}_j)} = V_1 \otimes \ldots \otimes V_{j-1} \otimes V_j^{(N_j)} \otimes V_{j+1} \otimes \ldots \otimes V_{d+1}$$

endowed with the norm $\|\cdot\|_{\mathbf{N}_j}$. We require

$$\|\cdot\|_{\vee(V_1,\ldots,V_{j-1},V_j^{(N_j)},V_{j+1},\ldots,V_{d+1})} \lesssim \|\cdot\|_{\mathbf{N}_j} \qquad \text{for all } 1 \leq j \leq d. \tag{6.28}$$

Lemma 6.33. *Assume (6.28). Let* $\varphi^{[j]} \in {}_a\bigotimes_{k\neq j} V_k^*$ *and* $\mathbf{v}_m, \mathbf{v} \in \mathbf{V}$ *with* $\mathbf{v}_m \rightharpoonup \mathbf{v}$.
(a) Then $\varphi^{[j]}(\mathbf{v}_m) \rightharpoonup \varphi^{[j]}(\mathbf{v})$ *in* $V_j^{(N_j)}$.
(b) The estimate

$$\|\varphi^{[j]}(\mathbf{v} - \mathbf{v}_m)\|_{j,N_j} \leq C \, \|\mathbf{v} - \mathbf{v}_m\|_{\mathbf{N}_j}$$

holds for elementary tensors $\varphi^{[j]} = \bigotimes_{k\neq j} \varphi^{(k)} \in {}_a\bigotimes_{k\neq j} V_k^*$ *with* $\|\varphi^{(k)}\|^* = 1$, *where* C *is the norm constant involved in (6.28).*

Proof. Repeat the proof of Lemma 6.23 and note that a functional $\varphi^{(j)} \in (V_j^{(N_j)})^*$ composed with an elementary tensor $\varphi^{[j]} = \bigotimes_{k\neq j} \varphi^{(k)} \in {}_a\bigotimes_{k\neq j} V_k^*$ yields $\varphi = \bigotimes_{k=1}^{d} \varphi^{(k)} \in {}_a\bigotimes_{k=1}^{d} (V_k^{(n_j)})^*$, where n_j are the components of $\mathbf{n} = \mathbf{N}_j$. By (6.28), φ belongs to $(\mathbf{V}^{(\mathbf{N}_j)})^*$. □

Conclusion 6.34. *Under assumption (6.28),* $U_j^{\min}(\mathbf{v}) \subset V_j^{(N_j)}$ *holds for all* $\mathbf{v} \in \mathbf{V}$ *and all* $1 \leq j \leq d$.

Proof. Let $\mathbf{v}_m \in \mathbf{V}_{\text{alg}}$ be a sequence with $\mathbf{v}_m \to \mathbf{v} \in \mathbf{V}$. By definition (4.52b) of the intersection norm, $\|\mathbf{v}_m - \mathbf{v}\|_{\mathbf{N}_j} \to 0$ holds for all j. Then Lemma 6.33b shows that $\|\varphi^{[j]}(\mathbf{v} - \mathbf{v}_m)\|_{j,N_j} \to 0$. Since $\varphi^{[j]}(\mathbf{v}_m) \in V_j^{(N_j)}$ by Proposition 4.104, also the limit $\varphi^{[j]}(\mathbf{v})$ belongs to $V_j^{(N_j)}$. □

Theorem 6.35. *Assume (6.28) and* $\mathbf{v}_m \in \mathbf{V}_{\text{alg}}$ *with* $\mathbf{v}_m \rightharpoonup \mathbf{v} \in \mathbf{V}$. *Then*

$$\dim(U_j^{\min}(\mathbf{v})) \leq \liminf_{m\to\infty} \dim(U_j^{\min}(\mathbf{v}_m)) \qquad \textit{for all } 1 \leq j \leq d.$$

Proof. We can repeat the proof from Theorem 6.24. □

6.8 Linear Constraints and Regularity Properties

Let $\varphi_k \in V_k^*$ be a continuous linear functional. We say that a tensor $\mathbf{v} \in \bigotimes_{j=1}^{d} V_j$ satisfies the linear constraint φ_k if

$$(id \otimes \ldots \otimes id \otimes \varphi_k \otimes id \otimes \ldots \otimes id)\, \mathbf{v} = 0. \tag{6.29}$$

A single constraint can be replaced by a family $\Phi \subset V_k^*$: $\mathbf{v}$ satisfies the linear constraints $\Phi \subset V_k^*$, if (6.29) holds for all $\varphi_k \in \Phi$.

If, for instance, $V_k = \mathbb{K}^{n_k \times n_k}$ is a matrix space, the subset of symmetric matrices is characterised by $\varphi_{\nu\mu}(M) := M_{\nu\mu} - M_{\mu\nu} = 0$ for all $1 \leq \nu \leq \mu \leq n_k$. In the same way, tridiagonal matrices, sparse matrices with a fixed sparsity pattern, etc. can be defined by means of linear constraints. The next statement is mentioned by Tyrtyshnikov [185].

Remark 6.36. Under the assumption (6.18), statements (a) and (b) are equivalent:
(a) $\mathbf{v} \in \bigotimes_{j=1}^{d} V_j$ satisfies a linear constraint φ_k,
(b) the minimal subspace $U_k^{\min}(\mathbf{v})$ fulfils $\varphi_k(U_k^{\min}(\mathbf{v})) = 0$, i.e., $\varphi_k(u_k) = 0$ holds for all $u_k \in U_k^{\min}(\mathbf{v})$.

Proof. The direction (b)⇒(a) is trivial. Assume (a) and choose any $u_k \in U_k^{\min}(\mathbf{v})$. There is some $\varphi_{[k]} \in V_{[k]}^*$ with $u_k = \varphi_{[k]}(\mathbf{v})$. Since $\varphi_k \circ \varphi_{[k]} = \varphi_{[k]} \circ \varphi_k$, one concludes that $\varphi_k(u_k) = \varphi_k(\varphi_{[k]}(\mathbf{v})) = \varphi_{[k]}(\varphi_k(\mathbf{v})) = \varphi_{[k]}(0) = 0$. □

In the infinite dimensional case, when $\mathbf{v}$ is a multivariate function, regularity properties of $\mathbf{v}$ are characterised by the *boundedness* of certain functionals. For instance, $\mathbf{v} \in \mathbf{V} = H^1(I)$ defined on $I := I_1 \times \ldots \times I_d$ is differentiable (in the weak sense) with respect to x_k, if

$$\|(id \otimes \ldots \otimes id \otimes \varphi_k \otimes id \otimes \ldots \otimes id)\,\mathbf{v}\|_{L^2(I)} < \infty \quad \text{for } \varphi_k = \partial/\partial x_k.$$

The formulation corresponding to Remark 6.36 is: If $\varphi_k : (\mathbf{V}, \|\cdot\|_{H^1(I)}) \to L^2(I)$ is bounded, also $\varphi_k : (U_k^{\min}(\mathbf{v}), \|\cdot\|_{H^1(I_k)}) \to L^2(I_k)$ is bounded.

The more general formulation of this property is Conclusion 6.34: If

$$id \otimes \ldots \otimes id \otimes \varphi_k \otimes id \otimes \ldots \otimes id : \mathbf{V}^{(\mathbf{N})} \to \mathbf{V}^{(\mathbf{0})}$$

is bounded, $U_k^{\min}(\mathbf{v}) \subset V_k^{(N_k)}$ holds. Note that $U_k^{\min}(\mathbf{v})$ is defined via $\left(\mathbf{V}_{[k]}^{(\mathbf{0})}\right)^*$. Therefore, considering $\mathbf{v} \in \mathbf{V}$ as a function $\mathbf{v} \in \mathbf{V}^{(\mathbf{0})}$, we obtain the same minimal subspace $U_k^{\min}(\mathbf{v}) \subset V_k^{(0)}$. This leads to the next observation.

Remark 6.37. Let the Banach tensor space $\mathbf{V}^{(\mathbf{0})} = {}_{\|\cdot\|_{\mathbf{0}}}\bigotimes_{j=1}^{d} V_j$ with norm $\|\cdot\|_{\mathbf{0}}$ satisfy (6.18). Then the minimal subspace of $\mathbf{v} \in \mathbf{V}^{(\mathbf{n})}$ satisfies $U_j^{\min}(\mathbf{v}) \subset V_j^{(n_j)}$, while $\mathbf{v} \in \mathbf{V} = \bigcap_{\mathbf{n} \in \mathcal{N}} \overline{\mathbf{V}^{(\mathbf{n})}}$ leads to $U_j^{\min}(\mathbf{v}) \subset V_j^{(N_j)}$.

Part III
Numerical Treatment

The numerical treatment of tensors is based on a suitable *tensor representation*. The first four chapters are devoted to two well-known representations. *Chapter 7* describes the r-term format (also called canonical or CP format), while *Chap. 8* is concerned with the tensor subspace format (also called Tucker format). In both chapters, tensors are *exactly* represented. Quite another topic is the *approximation* of tensors. *Chapter 7* studies the approximation within the r-term format. Here, it becomes obvious that tensors of larger order than two have much less favourable properties than matrices, which are tensors of order two. Approximation within the tensor subspace format is addressed in *Chap. 8*. Here, the technique of *higher order singular value decomposition* (HOSVD; cf. §10.1) is very helpful, both theoretically and practically.

While the r-term format suffers from a possible numerical instability, the storage size of the tensor subspace format increases exponentially with the tensor order d. A format avoiding both drawbacks is the *hierarchical format* described in *Chap. 11*. Here, the storage is strictly bounded by the product of the maximal involved rank, the maximal dimension of the vector spaces V_j, and d, the order of the tensor. Again, HOSVD techniques can be used for a quasi-optimal truncation. Since the format is closed, numerical instability does not occur.

The hierarchical format is based on a dimension partition tree. A particular choice of the tree leads to the *matrix product representation* or TT format described in *Chap. 12*.

The essential part of the numerical tensor calculus is the performance of *tensor operations*. In *Chap. 13* we describe all operations, their realisation in the different formats, and the corresponding arithmetical cost.

Chapter 14 contains the details of the *tensorisation* technique. When applied to (grid) functions, tensorisation corresponds to a multiscale approach.

Chapter 15 is devoted to the *generalised cross approximation*, which has several important applications. If a tensor can be evaluated entry-wise, this method allows to construct a tensor approximation in the hierarchical format.

In *Chap. 16*, the application of the tensor calculus to elliptic boundary value problems and elliptic eigenvalue problems is discussed.

The final *Chap. 17* collects a number of further topics. Section 17.1 considers general minimisation problems. Another minimisation approach described in Sect. 17.2 applies directly to the parameters of the tensor representation. Dynamic problems are studied in Sect. 17.3, while the ANOVA method is mentioned in Sect. 17.4.

Chapter 7
r-Term Representation

Abstract The r-term representation $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$, i.e., a representation by sums of r elementary tensors, is already used in the algebraic definition (3.11) of tensors. In different fields, the r-term representation has different names: 'canonical decomposition' in psychometrics (cf. [30]), 'parallel factors model' (cf. [96]) in chemometrics.[1] The word 'representation' is often replaced by 'format'. The short form 'CP' is proposed by Comon [38] meaning 'canonical polyadic decomposition'. Here, the notation 'r-term representation' is used with 'r' considered as a variable from $\mathbb{N}_0$, which may be replaced by other variable names or numbers.

Before we discuss the r-term representation in *Sect. 7.3*, we consider representations in general (*Sect. 7.1*) and the full representation (*Sect. 7.2*). The sensitivity of the r-term representation is analysed in *Sect. 7.4*. *Section 7.5* discusses possible representations of the vectors $v_\nu^{(j)} \in V_j$. We briefly mention the conversion from full format to r-term representation (cf. §7.6.1) and modifications (cf. *Sect. 7.7*).

The discussion of arithmetical operations with tensors in r-term representation is postponed to Chap. 13. In this chapter we restrict our considerations to the exact representation in the r-term format. Approximations, which are of greater interest in practice, will be discussed in Chap. 9.

7.1 Representations in General

7.1.1 Concept

For any practical implementation we have to distinguish between the mathematical objects and the way we represent them for the purpose of a computer implementation. To give a simple example: any rational number $r \in \mathbb{Q}$ may be represented by a pair (p, q) of integers. Here we need the semantic explanation that $r = p/q$.

[1] The combination of both names has led to the awful abbreviation 'CANDECOMP/PARAFAC decomposition'.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_7,

From this example we see that the representation is not necessarily unique. The representation concept may be iterated, e.g., in the previous example we need a representation of the integers by a bit sequence together with a coding of the sign.

Let S be a mathematical set. In general, a representation of $s \in S$ is based on a mapping

$$\rho_S : P_S \to S \tag{7.1}$$

where usually the set P_S consists of tuples $p = (p_1, \ldots, p_n)$ of parameters which are assumed to be realisable in computer language implementations. The mapping ρ_S is only the explanation of the meaning of $p \in P_S$ and is not a subject of implementation.

First we discuss *surjectivity* of ρ_S. Since most of the mathematical sets are infinite, whereas a real computer has only finite size, surjectivity cannot hold in general. There are two ways to overcome this problem.

Like in the concept of the Turing machine, we may base our considerations on a virtual computer with infinite storage. Then, e.g., all integers can be represented by bit sequences of arbitrary length.

The second remedy is the replacement of S by a finite subset $S_0 \subset S$ such that $\rho_S : P_S \to S_0$ becomes surjective. For instance, we may restrict the integers to an interval $S_0 = \mathbb{Z} \cap [-i_{\max}, i_{\max}]$. A consequence is that we have to expect problems when we try to perform the addition $i_{\max} + 1$. Another type of replacement $S_0 \subset S$ is known for the case $S = \mathbb{R}$. Here, the set S_0 of machine numbers satisfies a density property: Any real number[2] $x \in S = \mathbb{R}$ can be approximated by $x_0 \in S_0$ such that the relative error is smaller than the so-called *machine precision eps*.

From now on we assume that the mapping $\rho_S : P_S \to S$ is surjective (since infinite storage is assumed or/and S is replaced by a subset which again is called S).

There is no need to require *injectivity* of ρ_S. In general, the inverse $\rho_S^{-1}(s)$ of some $s \in S$ is set-valued. Any $p \in \rho_S^{-1}(s)$ may be used equally well to represent $s \in S$.

7.1.2 Computational and Memory Cost

An important property of $s \in S$ is the storage size needed for its representation. For instance, a natural number $n \in S = \mathbb{N}$ needs $1 + \lfloor \log_2 n \rfloor$ bits. In general, we have to deal with the storage needed for a parameter tuple $p = (p_1, \ldots, p_n)$. We denote the necessary storage by

$$N_{\mathrm{mem}}(p).$$

Since $s \in S$ may have many representations, we associate s with the memory size[3]

$$N_{\mathrm{mem}}(s) := \min \left\{ N_{\mathrm{mem}}(p) : p \in \rho_S^{-1}(s) \right\}.$$

[2] Here, we ignore the problems of overflow and underflow, which is a difficulty of the same kind as discussed above for integers.

[3] The memory size $N_{\mathrm{mem}}(p)$ is assumed to be a natural number. Any subset of $\mathbb{N}$ has a minimum.

Practically, when $\rho_S^{-1}(s)$ is a large set, it might be hard to find $p \in \rho_S^{-1}(s)$ with $N_{\mathrm{mem}}(p) = N_{\mathrm{mem}}(s)$.

Usually, we want to perform some operations between mathematical objects or we want to evaluate certain functions. Assume, e.g., a binary operation $\boxdot$ within the set S. The assignment $s := s_1 \boxdot s_2$ requires to find a representation p of s provided that representations p_i of s_i $(i = 1, 2)$ are given. The corresponding operation $\widehat{\boxdot}$ on the side of the parameter representations becomes

$$p := p_1 \mathbin{\widehat{\boxdot}} p_2 \qquad :\Longleftrightarrow \qquad \rho_S(p) = \rho_S(p_1) \boxdot \rho_S(p_2) \tag{7.2}$$

(note that $p \in \rho_S^{-1}(\rho_S(p_1) \boxdot \rho_S(p_2))$ is in general not unique). The right-hand side in (7.2) explains only the meaning of $\widehat{\boxdot}$. It cannot be used for the implementation since ρ_S and ρ_S^{-1} are not implementable. We assume that there is some algorithm mapping the arguments $p_1, p_2 \in P_S$ into some $p = p_1 \mathbin{\widehat{\boxdot}} p_2 \in P_S$ with finite $N_{\mathrm{mem}}(p)$ such that this computation requires a finite number of arithmetical operations. The latter number is denoted by $N_{\boxdot}$ and may be a function of the arguments. In standard considerations, where mainly the arithmetical operations $+, -, *, /$ of real numbers (machine numbers) appear, these form the unit of $N_{\boxdot}$.

The same setting holds for an n-variate function

$$\varphi : S_1 \times \ldots \times S_n \to S_0.$$

Assume representations $\rho_i : P_i \to S_i$ for $0 \le i \le n$. Then, on the level of representation, φ becomes

$$\hat{\varphi} : P_1 \times \ldots \times P_n \to P_0 \quad \text{with} \quad \rho_0\left(\hat{\varphi}(p_1, \ldots, p_n)\right) = \varphi(\rho_1(p_1), \ldots, \rho_n(p_n)).$$

The required number of arithmetical operations is denoted by N_φ.

7.1.3 Tensor Representation versus Tensor Decomposition

The term ‘decomposition’ is well-known, e.g., from the QR decomposition or singular value decomposition. One may define a decomposition as an (at least essentially) injective representation.

As an example we take the singular value decomposition. We may represent a matrix M by the three matrix-valued parameters $p_1 = U$ (unitary matrix), $p_2 = \Sigma$ (diagonal matrix), and V (unitary matrix). The semantic explanation is $\rho_{\mathrm{SVD}}(U, \Sigma, V) = M := U \Sigma V^{\mathsf{T}}$. However, the representation of M is not the purpose of SVD. Instead, the parameters U, Σ, V of $\rho_{\mathrm{SVD}}(U, \Sigma, V) = M$ are of interest, since they indicate important properties of M. Injectivity of ρ_{SVD} is necessary to speak about *the* singular vectors u_i, v_i and *the* singular values Σ_{ii}. We know from Corollary 2.21b that injectivity does not hold for multiple singular values. Therefore, the vague formulation ‘essentially injective’ has been used above.

In a certain way, 'representation' and 'decomposition' play opposite rôles like synthesis and analysis.

- 'representation': The parameters in $\rho_S(p_1, \ldots, p_n) = s$ are only of auxiliary nature. In the case of non-injectivity, any parameter tuple is as good as another. Only, if the data sizes are different, one may be interested in the cost-optimal choice. The representation of s is illustrated by the direction
$$p_1, \ldots, p_n \mapsto s.$$
- 'decomposition': For a given $s \in S$ one likes to obtain the parameters p_i in $\rho_S(p_1, \ldots, p_n) = s$. Therefore, the direction is
$$s \mapsto p_1, \ldots, p_n.$$

'Tensor decomposition' it is applied, when features of a concrete object should be characterised by parameters of tensor-valued data about this object. The r-term representation can be considered as decomposition, since often essential injectivity holds (cf. Remark 7.4b). The HOSVD decomposition from §8.3 is another example.

Since our main interest is the calculation with tensors, we are only interested in representations of tensors.

7.2 Full and Sparse Representation

Consider the tensor space
$$\mathbf{V} = \bigotimes_{j=1}^{d} V_j$$
of finite dimensional vector spaces V_j. As mentioned before, in the finite dimensional case we need not distinguish between algebraic and topological tensor spaces. After introducing bases $\{b_1^{(j)}, b_2^{(j)}, \ldots\}$ of V_j and index sets $I_j := \{1, \ldots, \dim(V_j)\}$ we reach the representation of elements from $\bigotimes_{j=1}^{d} V_j$ by elements from $\mathbb{K}^{\mathbf{I}} = \bigotimes_{j=1}^{d} \mathbb{K}^{I_j}$, where
$$\mathbf{I} = I_1 \times \ldots \times I_d.$$

This representation $\rho : P = \mathbb{K}^{\mathbf{I}} \to S = \bigotimes_{i=1}^{d} V_i$ (cf. (7.1)) is defined by
$$\rho_{\text{full}}(\mathbf{a}) = \sum_{\mathbf{i} \in \mathbf{I}} \mathbf{a}_{\mathbf{i}}\, b_{i_1}^{(1)} \otimes \ldots \otimes b_{i_d}^{(d)} \quad \text{with } \mathbf{a} \in \mathbb{K}^{\mathbf{I}} \text{ and } \mathbf{i} = (i_1, \ldots, i_d) \in \mathbf{I}. \tag{7.3}$$

In the case of $V_j = \mathbb{K}^{I_j}$, the bases are formed by the unit vectors.

Notation 7.1. The full representation uses the coefficients $\mathbf{a}_{\mathbf{i}} \in \mathbb{K}^{\mathbf{I}}$ with the interpretation (7.3). The data size is

$$N_{\text{mem}}^{\text{full}} = \#\mathbf{I} = \dim\left(\bigotimes_{j=1}^{d} V_j\right) = \prod_{j=1}^{d} \dim(V_j). \tag{7.4}$$

In the model case of $\dim(V_j) = n$ for all $1 \le j \le d$, the storage size is n^d. Unless n and d are very small numbers, the value of n^d is too huge for practical realisations. In particular when $d \to \infty$, the exponential growth of n^d is a severe hurdle.

In the case of matrices, the format of sparse matrices is very popular. For completeness, we formulate the sparse tensor format. A concrete example will follow in §7.6.5.

Remark 7.2. Given a subset $\mathring{\mathbf{I}} \subset \mathbf{I}$ and $\mathbf{a_i} \in \mathbb{K}$ for all $\mathbf{i} \in \mathring{\mathbf{I}}$, the sparse representation consists of the data $\mathring{\mathbf{I}}$ and $(\mathbf{a_i})_{\mathbf{i}\in\mathring{\mathbf{I}}}$ and represents the tensor

$$\rho_{\text{sparse}}\left(\mathring{\mathbf{I}}, (\mathbf{a_i})_{\mathbf{i}\in\mathring{\mathbf{I}}}\right) = \sum_{\mathbf{i}\in\mathring{\mathbf{I}}} \mathbf{a_i}\, b_{i_1}^{(1)} \otimes \ldots \otimes b_{i_d}^{(d)}. \tag{7.5}$$

The data size is $N_{\text{mem}}^{\text{sparse}} = 2\#\mathring{\mathbf{I}}$.

The traditional understanding in linear (and multilinear) algebra is that their objects are completely given. A vector $v \in \mathbb{K}^I$ needs knowledge of all v_i $(i \in I)$ and, correspondingly, all coefficients v_i should be stored simultaneously. In the field of analysis, the alternative concept of functions is dominating. Note that $\mathbb{K}^I$ for arbitrary, possibly infinite sets I is isomorphic to the set of functions $I \to \mathbb{K}$. Although a function f, say from $I = [0,1]$ to $\mathbb{K}$ is defined as the set $\{(x, f(x)) : x \in [0,1]\}$ of all pairs, implementations of f do not suffer from the fact that there are infinitely many function values. Instead of requiring all values to be present, one asks for the possibility to determine $f(x)$ only for a given $x \in I$.

Accordingly, the *full functional representation* of some $\mathbf{a} \in \mathbb{K}^{\mathbf{I}}$ needs the implementation of a function

$$\textit{function}\ a(i_1, i_2, \ldots, i_d) \tag{7.6}$$

which returns the entry $\mathbf{a_i} \in \mathbb{K}$ for any *single* index $\mathbf{i} = (i_1, \ldots, i_d)$.

Often tensors which can be represented by (7.6), are called *function related tensors*. This naming is a bit vague: In principle, any tensor $\mathbf{a} \in \mathbb{K}^{\mathbf{I}}$ with finite $\mathbf{I}$ can be implemented by (7.6). A typical function related tensor is

$$f(i_1 h, i_2 h, \ldots, i_d h) \qquad \text{for } i_j \in \{0, 1, \ldots, n\} \text{ and } h = 1/n,$$

which describes the restriction of a function $f : [0,1]^d \to \mathbb{K}$ to a uniform grid $G_h \subset [0,1]^d$ of grid size h (this restriction we call 'grid function'). The evaluation time of $f(x)$ for an $x \in \mathbb{K}^d$ is assumed to be independent of the grid size h. The uniform grid may equally well be replaced by $f(x_{i_1}^{(1)}, x_{i_2}^{(2)}, \ldots, x_{i_d}^{(d)})$, where $\{x_i^{(j)} : i \in I_j\}$ is a non-equidistant grid in the j-th direction.

The functional representation is of particular interest, if partial evaluations of $\mathbf{a} \in \mathbb{K}^{\mathbf{I}}$ are required (cf. §15).

7.3 r-Term Representation

The set $\mathcal{R}_r$ defined in (3.22) is fundamental for the r-term representation.

Definition 7.3. For variable $r \in \mathbb{N}_0$, the r-term representation is explained by the mapping

$$\rho_{\text{r-term}}\Big(r, (v_\nu^{(j)})_{\substack{1\le j\le d\\ 1\le\nu\le r}}\Big) := \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)} \in {}_a\bigotimes_{j=1}^{d} V_j \quad \text{with} \begin{cases} r \in \mathbb{N}_0\,, \\ v_\nu^{(j)} \in V_j\,. \end{cases} \tag{7.7a}$$

For fixed $r \in \mathbb{N}_0$, the r-term representation

$$\begin{aligned} &\rho_{\text{r-term}} : \mathbb{N}_0 \times \textstyle\bigcup_{r\in\mathbb{N}_0} (V_1 \times \ldots \times V_d)^r \to \mathcal{R}_r\,, \\ &\rho_{\text{r-term}}\Big(r, (v_\nu^{(j)})_{\substack{1\le j\le d\\ 1\le\nu\le r}}\Big) = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)} \in \mathcal{R}_r \end{aligned} \tag{7.7b}$$

refers to all tensors from $\mathcal{R}_r$. We call r the *representation rank*.

Note that the representation rank refers to the representation by the parameters $(v_\nu^{(j)})_{1\le j\le d, 1\le\nu\le r}$, not to the represented tensor. Because of $\mathcal{R}_r \subset \mathcal{R}_{r+1}$, a tensor expressed with representation rank r can also be expressed by any larger representation rank.

The relation between the representation rank and the tensor rank is as follows:
(i) If $\mathbf{v} \in \mathbf{V}$ is represented by a representation rank r, then $\operatorname{rank}(\mathbf{v}) \le r$.
(ii) Let $r := \operatorname{rank}(\mathbf{v})$. Then there exists a representation of $\mathbf{v}$ with representation rank r. However, finding this representation may be NP-hard (cf. Proposition 3.34).

Remark 7.4. (a) The r-term representation is by no means injective; e.g., $v_\nu^{(j)}$ may be replaced by $\lambda_{j,\nu} v_\nu^{(j)}$ with scalars $\lambda_{j,\nu} \in \mathbb{K}$ satisfying $\prod_{j=1}^{d} \lambda_{j,\nu} = 1$. To reduce this ambiguity, one may consider the modified representation

$$\mathbf{v} = \sum_{\nu=1}^{r} a_\nu \bigotimes_{j=1}^{d} v_\nu^{(j)} \qquad \text{with } \|v_\nu^{(j)}\|_{V_j} = 1 \text{ for all } 1 \le j \le d$$

with normalised $v_\nu^{(j)}$ and factors $a_\nu \in \mathbb{K}$. Still the sign of $v_\nu^{(j)}$ is not fixed (and in the case of $\mathbb{K}=\mathbb{C}$, $v_\nu^{(j)}$ may be replaced by $\lambda_{j,\nu} v_\nu^{(j)}$ with $|\lambda_{j,\nu}|=1$ and $\prod_{j=1}^{d} \lambda_{j,\nu}=1$). For large d, this representation is not the best choice for practical use, since over- or underflow of the floating point numbers a_ν may occur. A better normalisation[4] is

$$\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)} \qquad \text{with } \|v_\nu^{(j)}\|_{V_j} = \|v_\nu^{(k)}\|_{V_k} \text{ for all } 1 \le j, k \le d.$$

Another trivial ambiguity of the r-term representation is the ordering of the terms.
(b) The representation $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$ is called *essentially unique*, if the scalar factors and ordering of the terms mentioned in Part (a) are the only ambiguities. Under certain conditions, essential uniqueness holds (cf. Kruskal [133]). For a detailed discussion see Kolda-Bader [128, §3.2] and De Lathauwer [40].

[4] Cf. Mohlenkamp [149, Remark 2.1].

The second formula in Example 3.45 shows that there are tensors which allow more than countably many representations which are essentially different. As long as we are only interested in a cheap representation of tensors and cheap realisations of tensor operations, (essential) uniqueness is not relevant. This is different in applications, where the vectors $v_\nu^{(j)}$ are used for an interpretation of certain data $\mathbf{v}$.

Remark 7.5. (a) The storage size for the parameter $p = \left(r, (v_\nu^{(j)})_{1\le j\le d, 1\le\nu\le n}\right)$ is

$$N_{\mathrm{mem}}^{r\text{-term}}(p) = r \cdot \sum\nolimits_{j=1}^{d} size(v_\nu^{(j)}). \tag{7.8a}$$

Here, the representation is iterated: we need some representation of $v_\nu^{(j)} \in V_j$ and the related storage size is denoted by $size(v_\nu^{(j)})$. More details about $size(\cdot)$ will follow in §7.5.
(b) For the standard choice $V_j = \mathbb{K}^{I_j}$ with $n_j := \# I_j$ the full representation of a vector $v_\nu^{(j)} \in V_j$ requires $size(v_\nu^{(j)}) = n_j$. This results in

$$N_{\mathrm{mem}}^{r\text{-term}}(p) = r \cdot \sum\nolimits_{j=1}^{d} n_j. \tag{7.8b}$$

(c) Assume $n_j = n$ for all $1 \le j \le d$. Then the characteristic size is

$$N_{\mathrm{mem}}^{r\text{-term}}(p) = r \cdot d \cdot n. \tag{7.8c}$$

By Corollary 3.37, an isomorphism $\Phi : \mathbf{V} \to \mathbf{W}$ is a bijection $\mathcal{R}_r(\mathbf{V}) \rightleftarrows \mathcal{R}_r(\mathbf{W})$, i.e., the r-term format remains invariant: $\mathbf{v} \in \mathcal{R}_r(\mathbf{V}) \Leftrightarrow \mathbf{w} := \Phi(\mathbf{v}) \in \mathcal{R}_r(\mathbf{W})$.

The following remark is a reformulation of Remark 6.1.

Remark 7.6. $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$ is an element of $\bigotimes_{j=1}^{d} U_j$ with subspaces $U_j := \mathrm{span}\{v_\nu^{(j)} : 1 \le \nu \le r\}$ for $1 \le j \le d$. In particular, $U_j^{\min}(\mathbf{v}) \subset U_j$ holds.

We state two results involving the minimal subspaces $U_j^{\min}(\mathbf{v}) \subset V_j$ from §6. Given some r-term representation $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$, the first statement shows that the vectors $v_\nu^{(j)}$ may be projected to $\hat{v}_\nu^{(j)} \in U_j^{\min}(\mathbf{v})$ and $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} \hat{v}_\nu^{(j)}$ is still a correct representation. This implies that the search for candidates of $v_\nu^{(j)}$ may immediately be restricted to $U_j^{\min}(\mathbf{v})$.

Lemma 7.7. *For $\mathbf{v} \in \mathbf{V} := \bigotimes_{j=1}^{d} V_j$ let $P_j \in L(V_j, V_j)$ be a projection onto $U_j^{\min}(\mathbf{v})$.*
(a) The vectors $v_i^{(j)}$ in the r-term representation may be replaced by $P_j v_i^{(j)}$:

$$\mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} P_j v_i^{(j)}. \tag{7.9}$$

(b) Equation (7.9) is also valid, if some of the P_j are replaced by the identity.

Proof. Set $\mathbf{P} := \bigotimes_{j=1}^{d} P_j$ and use $\mathbf{v} = \mathbf{P}\mathbf{v}$. □

The second result concerns representations $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$ with minimal $r = \mathrm{rank}(\mathbf{v})$ and states that in this case equality $U_j^{\min}(\mathbf{v}) = U_j$ must hold. Hence, $v_\nu^{(j)} \in U_j^{\min}(\mathbf{v})$ is a necessary condition for a representation with minimal r.

Proposition 7.8. *If* $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$ *is a representation with* $r = \operatorname{rank}(\mathbf{v})$*, the subspaces from Remark 7.6 are* $U_j = U_j^{\min}(\mathbf{v})$ *(cf. (6.10a)).*

Proof. Let $j \in \{1, \ldots, d\}$. P_j is some projection onto $U_j^{\min}(\mathbf{v})$. By Lemma 3.38, $\mathbf{v} = \sum_{\nu=1}^{r} v_\nu^{(j)} \otimes \mathbf{v}_\nu^{[j]}$ holds with linearly independent $\mathbf{v}_\nu^{[j]} \in {}_a\bigotimes_{k \neq j} V_k$. Apply P_j (regarded as a map from $L(\mathbf{V}, \mathbf{V})$, cf. Notation 3.50) to $\mathbf{v}$. By Lemma 7.7b, $\mathbf{v} = P_j \mathbf{v} = \sum_{\nu=1}^{r} (P_j v_\nu^{(j)}) \otimes \mathbf{v}_\nu^{[j]}$ is valid implying $0 = \sum_{\nu=1}^{r} (v_\nu^{(j)} - P_j v_\nu^{(j)}) \otimes \mathbf{v}_\nu^{[j]}$. Linear independence of $\mathbf{v}_\nu^{[j]}$ proves $v_\nu^{(j)} - P_j v_\nu^{(j)} = 0$ (cf. Lemma 3.56), i.e., all $v_\nu^{(j)}$ belong to $U_j^{\min}(\mathbf{v})$. □

A consequence are the following conclusions from §6.8.

Remark 7.9. Suppose that $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$ with $r = \operatorname{rank}(\mathbf{v})$.
(a) If $\mathbf{v}$ satisfies a linear constraint φ_k (as defined in §6.8), then $\varphi_k(v_\nu^{(k)}) = 0$ holds for all $1 \leq \nu \leq r$ (cf. Tyrtyshnikov [185, Theorem 2.1]).
(b) Let $\mathbf{V} = \bigcap_{\mathbf{n} \in \mathcal{N}} \overline{\mathbf{V}^{(\mathbf{n})}}$ be the intersection Banach spaces from §4.3.6. Then $\mathbf{v} \in \mathbf{V}^{(\mathbf{n})}$ $[\mathbf{V}]$ implies $v_\nu^{(j)} \in V_j^{(n_j)}$ $[V_j^{(N_j)}]$ for all $1 \leq \nu \leq r$.

7.4 Sensitivity

We have started in §7.1 with general comments about representations $\rho_S(p_1, \ldots, p_n)$ by means of parameters p_j. From the numerical point of view it is important to know how ρ_S behaves under perturbations of p_j. The derivative $\partial\rho_S/\partial p_j$ may be called *sensitivity with respect to* p_j. There are several reasons why one is interested in these numbers. Since we are almost never working with exact data, the true parameter p_j may be perturbed by rounding or other effects. Another reason are approximations, where the parameters $p = (p_1, \ldots, p_n)$ are replaced by approximate ones. Whether such perturbations lead to dangerous effects for $\rho_S(p_1, \ldots, p_n)$ is detected by the sensitivity analysis.

In the case of the r-term representation

$$\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}, \tag{7.10a}$$

we use $v_\nu^{(j)}$ as parameters (cf. (7.7b)). For all $v_\nu^{(j)}$ we allow perturbations $d_\nu^{(j)}$:

$$\tilde{\mathbf{v}} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} \left(v_\nu^{(j)} + d_\nu^{(j)} \right). \tag{7.10b}$$

It is convenient to consider the relative error

$$\delta_\nu^{(j)} := \frac{\|d_\nu^{(j)}\|}{\|v_\nu^{(j)}\|}. \tag{7.10c}$$

Proposition 7.10. *Let* $\|\cdot\|$ *be any crossnorm on* $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ *(cf. Definition 4.31 and (4.40)). Write the tensor* $\mathbf{v} \in \mathbf{V}$ *from (7.10a) as* $\mathbf{v} = \sum_{\nu=1}^r \mathbf{v}_\nu$ *with the elementary tensors* $\mathbf{v}_\nu := \bigotimes_{j=1}^d v_\nu^{(j)}$*. Then the following error estimate holds:*

$$\|\tilde{\mathbf{v}} - \mathbf{v}\| \le \sum_{\nu=1}^r \delta_\nu \|\mathbf{v}_\nu\| \qquad \textit{with } \delta_\nu := \Big[\prod\nolimits_{j=1}^d (1 + \delta_\nu^{(j)})\Big] - 1. \tag{7.11}$$

For small $\delta_\nu^{(j)}$ *the first order approximation is* $\delta_\nu \approx \sum_{j=1}^d \delta_\nu^{(j)}$.

Proof. The term $\mathbf{v}_\nu$ is only effected by $\delta_\nu^{(j)}$ $(1 \le j \le d)$. Induction by d will prove

$$\left\| \bigotimes_{j=1}^d \left(v_\nu^{(j)} + d_\nu^{(j)}\right) - \bigotimes_{j=1}^d v_\nu^{(j)} \right\| \le \Big[\prod\nolimits_{j=1}^d (1 + \delta_\nu^{(j)}) - 1\Big] \prod\nolimits_{j=1}^d \|v_\nu^{(j)}\|.$$

For $d = 1$ the statement is trivial. Assume that the statement holds for $d-1$, i.e., for the products $\bigotimes_{j=2}^d \cdots$. Then

$$\begin{aligned}
&\left\| \bigotimes_{j=1}^d \left(v_\nu^{(j)} + d_\nu^{(j)}\right) - \bigotimes_{j=1}^d v_\nu^{(j)} \right\| \\
&= \left\| \left(v_\nu^{(1)} + d_\nu^{(1)}\right) \otimes \left\{ \bigotimes_{j=2}^d \left(v_\nu^{(j)} + d_\nu^{(j)}\right) - \bigotimes_{j=2}^d v_\nu^{(j)} \right\} + d_\nu^{(1)} \otimes \bigotimes_{j=2}^d v_\nu^{(j)} \right\| \underset{(4.40)}{\le} \\
&\le \|v_\nu^{(1)} + d_\nu^{(1)}\| \left\| \bigotimes_{j=2}^d \left(v_\nu^{(j)} + d_\nu^{(j)}\right) - \bigotimes_{j=2}^d v_\nu^{(j)} \right\| + \|d_\nu^{(1)}\| \prod_{j=2}^d \|v_\nu^{(j)}\| \underset{\substack{\text{inductive}\\\text{hypothesis}}}{\le} \\
&\le (1 + \delta_\nu^{(1)}) \|v_\nu^{(1)}\| \left[\prod_{j=2}^d (1 + \delta_\nu^{(j)}) - 1 \right] \prod\nolimits_{j=2}^d \|v_\nu^{(j)}\| + \delta_\nu^{(1)} \|v_\nu^{(1)}\| \prod_{j=2}^d \|v_\nu^{(j)}\| \\
&= \Big[\prod\nolimits_{j=1}^d \left(1 + \delta_\nu^{(j)}\right) - 1\Big] \prod\nolimits_{j=1}^d \|v_\nu^{(j)}\|
\end{aligned}$$

proves the statement. □

The error estimate (7.11) should be combined with the *stability estimate*

$$\sum_{i=1}^r \left\| \bigotimes\nolimits_{j=1}^d v_i^{(j)} \right\| \le \varkappa \left\| \sum_{i=1}^r \bigotimes\nolimits_{j=1}^d v_i^{(j)} \right\|, \tag{7.12}$$

which will be discussed in more detail in Definition 9.15. Note that the best (smallest) stability constant is $\varkappa = 1$. Together, we can estimate the relative error of $\tilde{\mathbf{v}}$:

$$\frac{\|\tilde{\mathbf{v}} - \mathbf{v}\|}{\|\mathbf{v}\|} \le \varkappa\,\delta \qquad \text{with } \delta := \max\{\delta_\nu : 1 \le \nu \le r\}.$$

Since the condition $\varkappa$ may be as large as possible (cf. §9.4), there is no guarantee that a small relative perturbation in $v_i^{(j)}$ leads to a similarly small relative error of $\mathbf{v}$.

Finally, we consider the ℓ^2 norm $\sqrt{\sum_{j=1}^d (\delta_\nu^{(j)})^2}$. In the case of general errors

$d_\nu^{(j)}$, $\delta_\nu = \prod_{j=1}^d (1+\delta_\nu^{(j)}) - 1 \approx \sum_{j=1}^d \delta_\nu^{(j)}$ is the best result, so that $\|\tilde{\mathbf{v}} - \mathbf{v}\| \leq \delta$ with $\delta \approx \sqrt{rd}\sqrt{\sum_{\nu=1}^r \|\mathbf{v}_\nu\|^2 \sum_{j=1}^d (\delta_\nu^{(j)})^2}$. The estimate improves a bit, if the error $d_\nu^{(j)}$ is a projection error.

Remark 7.11. Let $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ be a Hilbert tensor space with induced scalar product. Consider orthogonal projections $P_j : V_j \to V_j$ and the resulting errors $d_\nu^{(j)} := (P_j - I)v_\nu^{(j)}$ ($v_\nu^{(j)}$ from (7.10a)). Then the following error estimate holds:

$$\|\tilde{\mathbf{v}} - \mathbf{v}\| \leq \sqrt{r}\sqrt{\sum_{\nu=1}^r \|\mathbf{v}_\nu\|^2 \sum_{j=1}^d (\delta_\nu^{(j)})^2} \qquad \text{with } \|\mathbf{v}_\nu\| = \left\| \bigotimes_{j=1}^d v_\nu^{(j)} \right\|^2.$$

Proof. We repeat the inductive proof. The first and second lines from above are $\bigotimes_{j=1}^d (v_\nu^{(j)} + d_\nu^{(j)}) - \bigotimes_{j=1}^d v_\nu^{(j)} = (v_\nu^{(1)} + d_\nu^{(1)}) \otimes \cdots + d_\nu^{(1)} \otimes \cdots$. Since $v_\nu^{(1)} + d_\nu^{(1)} = P_1 v_\nu^{(1)}$ and $d_\nu^{(1)} = (P_j - I)v_\nu^{(j)}$ are orthogonal and $\|P_1 v_\nu^{(1)}\| \leq \|v_\nu^{(1)}\|$, the squared norm of $\bigotimes_{j=1}^d (v_\nu^{(j)} + d_\nu^{(j)}) - \bigotimes_{j=1}^d v_\nu^{(j)}$ is bounded by

$$\|v_\nu^{(1)}\|^2 \left\| \bigotimes_{j=2}^d (v_\nu^{(j)} + d_\nu^{(j)}) - \bigotimes_{j=2}^d v_\nu^{(j)} \right\|^2 + (\delta_\nu^{(j)})^2 \left\| \bigotimes_{j=1}^d v_\nu^{(j)} \right\|^2.$$

Induction leads us to $\|\tilde{\mathbf{v}}_\nu - \mathbf{v}_\nu\|^2 \leq \|\mathbf{v}_\nu\|^2 \sum_{j=1}^d (\delta_\nu^{(j)})^2$. Schwarz' inequality of the sum over ν proves the assertion. □

7.5 Representation of V_j

In (7.7a), $\mathbf{v} = \rho_{\text{r-term}}(r, (v_\nu^{(j)})_{j,\nu})$ is described as a representation of the tensor $\mathbf{v}$. However, representations may be become recursive if the parameters of the representation need again a representation. In this case, the involved vectors $v_i^{(j)} \in V_j$ must be implementable. If $V_j = \mathbb{K}^{I_j}$ with $n_j := \#I_j$, we might try to store the vector $v_i^{(j)}$ by full representation (i.e., as an array of length n_j). But depending on the size of n_j and the nature of the vectors $v_i^{(j)}$, there may be other solutions, e.g., representation as sparse vector if it contains mostly zero components. Another approach has been mentioned in §5.3 and will be continued in §14: usual vectors from $\mathbb{K}^{n_j}$ may be interpreted as higher order tensors. Under certain assumptions the storage of such tensor representations may be much cheaper than n_j (possibly, it becomes $O(\log n_j)$). These considerations are in particular of interest, if approximations are exceptable (see §9).

The spaces V_j may be matrix spaces: $V_j = \mathbb{K}^{I_j \times J_j}$. Full representation of large-scale matrices is usually avoided. Possibly, one can exploit the sparsity of $v_i^{(j)} \in \mathbb{K}^{I_j \times J_j}$. Another possibility is the representation of $v_i^{(j)}$ as hierarchical matrix (cf. Hackbusch [86]). In all these cases, the required storage size may strongly deviate from $\dim(V_j)$. We shall therefore use the notation

$$size(v^{(j)}) \qquad \text{for } v^{(j)} \in V_j$$

as already done in (7.8a).

Remark 7.12. Many computations require scalar products $\langle u, v\rangle_j$ for $u, v \in V_j$. We denote its computational cost by N_j. The standard Euclidean scalar product in $V_j = \mathbb{K}^{n_j}$ costs $N_j = 2n_j - 1$ arithmetical operations. N_j may be smaller than $2n_j - 1$ for certain representations of $u, v \in V_j$, while it may be larger for a scalar product $\langle u, v\rangle_j = v^{\mathsf{H}} A_j u$ involving some positive definite A_j.

Full representations are obviously impossible if $\dim(V_j) = \infty$. This happens, e.g., for $V_j = C([0,1])$. Really general functions cannot be represented in a finite way. A remedy is the approximation, e.g., by interpolation. Such an approach will be studied in §10.4. In this chapter we discuss exact representations. Often, the involved functions can be described by well-known function classes, e.g., $v_i^{(j)} \in C([0,1])$ are polynomials, trigonometric polynomials or sums of exponentials $\exp(\alpha x)$. Then we are led to the following situation:

$$v_i^{(j)} = \sum_{\nu \in B_j} \beta_{\nu,i}^{(j)} b_\nu^{(j)} \in U_j \subset V_j \qquad \text{with} \quad U_j = \operatorname{span}\{b_\nu^{(j)} : \nu \in B_j\}. \tag{7.13}$$

Now, $v_i^{(j)}$ is represented by its coefficients $(\beta_{\nu,i}^{(j)})_{\nu \in B_j}$, while the basis functions $b_\nu^{(j)}$ are fully characterised by $\nu \in B_j$ (e.g., the monomial x^n is completely described by the integer $n \in \mathbb{N}_0$). In §8.2.4 we shall obtain (7.13) in a different way and call it 'hybrid format'.

Representation (7.13) via a basis of a subspace U_j may even be of interest when $V_j = \mathbb{K}^{I_j}$ contains standard vectors (of large size $n_j = \#I_j$). If more than one tensor shares the same subspaces U_j, the costly storage of the basis vectors is needed only once, while the storage for the coefficients $\beta_{\nu,i}^{(j)}$ is of minor size. Having precomputed the Gram matrix $\Gamma^{(j)} \in \mathbb{K}^{B_j \times B_j}$ of the scalar products $\Gamma_{\nu,\mu}^{(j)} := \langle b_\nu^{(j)}, b_\mu^{(j)}\rangle$, we can reduce a scalar product $\langle v, u\rangle$ of $v, u \in U_j$ to the scalar product $\langle \Gamma\beta_v, \beta_u\rangle$ of the respective coefficients in $\mathbb{K}^{B_j}$. The resulting cost is $N_j = 2(\#B_j)^2 + \#B_j$. A similar situation is discussed in the next remark.

Remark 7.13. Assume that, according to (7.13), $\{v_i^{(j)} : 1 \le i \le r\} \subset V_j = \mathbb{K}^{I_j}$ is represented by coefficients $\beta_{\nu,i}^{(j)}$, where $r_j := \#B_j$ and $n_j = \dim(V_j)$. If the task is to compute the Gram matrix M_j with entries $M_{\nu\mu} := \langle v_\nu^{(j)}, v_\mu^{(j)}\rangle$, the direct approach would cost $\frac{1}{2}r(r+1)N_j \approx r^2 n_j$. The computation of $\Gamma^{(j)} \in \mathbb{K}^{B_j \times B_j}$ mentioned above requires $r_j^2 n_j$ operations. The Cholesky decomposition $\Gamma^{(j)} = L^{(j)} L^{(j)\mathsf{H}}$ can be determined by $\frac{1}{3}r_j^3$ operations. The coefficients $\beta_{\nu,i}^{(j)}$ define the vectors $\beta_i^{(j)} := (\beta_{\nu,i}^{(j)})_{\nu \in B_j}$ $(1 \le i \le r)$. Using

$$\langle v_\nu^{(j)}, v_\mu^{(j)}\rangle_{V_j} = \big\langle \Gamma^{(j)}\beta_\nu^{(j)}, \beta_\mu^{(j)}\big\rangle_{\mathbb{K}^{B_j}} = \big\langle L^{(j)\mathsf{H}}\beta_\nu^{(j)}, L^{(j)\mathsf{H}}\beta_\mu^{(j)}\big\rangle_{\mathbb{K}^{B_j}},$$

we compute all $M_{\nu\mu}$ by $r_j^2 r + r^2 r_j$ operations. The total cost of this approach is

$$r_j^2 n_j + \tfrac{1}{3} r_j^3 + r_j^2 r + r^2 r_j.$$

It is cheaper than the direct computation if $(r^2 - r_j^2)n_j > \frac{1}{3}r_j^3 + r_j^2 r + r^2 r_j$.

Remark 7.14. In the case of $V_j = \mathbb{K}^{I_j}$, the standard method for producing a basis $b_\nu^{(j)}$ and the coefficients $\beta_{\nu,i}^{(j)}$ from (7.13) is the reduced QR decomposition. Form the matrix $A := [v_1^{(j)} \cdots v_r^{(j)}]$ and decompose $A = QR$ with $Q \in \mathbb{K}^{I_j \times r_j}$, $r_j := \operatorname{rank}(A)$, $R \in \mathbb{K}^{r_j \times r}$ (cf. Lemma 2.19). Then $v_i^{(j)} = \sum_{\nu=1}^{r} Q_{\bullet,\nu} R_{\nu,i}$ holds, i.e., $b_\nu^{(j)} := Q_{\bullet,\nu}$, $\beta_{\nu,i}^{(j)} := R_{\nu,i}$, and $B_j := \{1, \ldots, r_j\}$. The computational cost is $N_{\mathrm{QR}}(n_j, r)$.

7.6 Conversions between Formats

In this section, we consider the tensor space $\mathbf{V} := \bigotimes_{j=1}^{d} \mathbb{K}^{I_j}$ with index sets of size $n_j := \#I_j$. The tensor index set is $\mathbf{I} := I_1 \times \ldots \times I_d$. An interesting number is

$$N := \left(\prod_{j=1}^{d} n_j \right) \Big/ \max_{1 \le i \le d} n_i, \tag{7.14}$$

which appears in Lemma 3.41 as bound of the maximal rank in $\mathbf{V}$. To simplify the notation, we assume that $n_1 = \max\{n_i : 1 \le i \le d\}$ and introduce the index set

$$\mathbf{I}' := I_2 \times \ldots \times I_d \tag{7.15}$$

of size $N = \#\mathbf{I}'$.

In particular if $r > N$, it can be interesting to convert a tensor from r-term format into another one. We discuss the conversion from full format (abbreviation: $\mathcal{F}$) or r-term format ($\mathcal{R}_r$) into other ones. At the right we give a summary of the costs using $n := \max_j n_j$.

conversion	arithmetical cost	
$\mathcal{F} \to \mathcal{R}_N$	0	
$\mathcal{R}_r \to \mathcal{F}$	$2rn^d$	for any r,
$\mathcal{R}_R \to \mathcal{R}_N$	$2Rn^d$	if $R > N$.

7.6.1 From Full Representation into r-Term Format

Assume that $\mathbf{v} \in \mathbf{V}$ is given in full representation, i.e., by all entries $\mathbf{v}[i_1 \ldots i_d]$ $(i_j \in I_j)$. The associated memory size is $N_{\mathrm{mem}}^{\mathrm{full}} = \prod_{j=1}^{d} n_j$. Theoretically, a shortest r-term representation $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$ exists with $r := \operatorname{rank}(\mathbf{v})$ and memory size $N_{\mathrm{mem}}^{\text{r-term}} = r \sum_{j=1}^{d} n_j$. However, its computation is usually far too difficult (cf. Proposition 3.34). On the other hand, we have constructively proved in Lemma 3.41 that the tensor rank is always bounded by N from (7.14).

Remark 7.15. Given $\mathbf{v} \in \mathbf{V}$ in full representation, the N-term representation with N from (7.14) is realised by

$$\mathbf{v} = \sum_{\mathbf{i}' \in \mathbf{I}'} \bigotimes_{j=1}^{d} v_{\mathbf{i}'}^{(j)} \text{ with } \begin{cases} v_{\mathbf{i}'}^{(1)} \in \mathbb{K}^{I_1},\ v_{\mathbf{i}'}^{(1)}[k] := \mathbf{v}[k, i_2', \ldots, , i_d'] & \text{for } j = 1, \\ v_{\mathbf{i}'}^{(j)} = e^{(j,i_j')} \in \mathbb{K}^{I_j} & \text{for } 2 \le j \le d, \end{cases} \tag{7.16}$$

where $e^{(j,i)}$ is the i-th unit vector in $\mathbb{K}^{I_j}$, i.e., $e^{(j,i)}[k] = \delta_{ik}$. Note that no arithmetical operations occur.

Unfortunately, conversion to N-term format increases the storage requirement. Even, if we do not associate any storage to the unit vectors $e^{(j,i)}$, $j \ge 2$, the vectors $v_{\mathbf{i}'}^{(1)}$, $\mathbf{i}' \in \mathbf{I}'$ (cf. (7.15)), have the same data size as the fully represented tensor $\mathbf{v}$.

Because of the large memory cost, the transfer described above, is restricted to $d = 3$ and moderate n_j. Further constructions leading to smaller ranks than N will be discussed in §7.6.5. Even smaller ranks can be reached, if we do not require an *exact* conversion, but allow for an approximation (cf. §9).

7.6.2 From r-Term Format into Full Representation

Conversion from r-term format into full representation can be of interest if $r > N$, since then the full format is cheaper. Given $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$, $v_\nu^{(j)} \in \mathbb{K}^{I_j}$, the expressions $\mathbf{v}[\mathbf{i}] = \sum_{\nu=1}^{r} \prod_{j=1}^{d} v_\nu^{(j)}[i_j]$ have to be evaluated for all $\mathbf{i} \in \mathbf{I}$.

Lemma 7.16. *Conversion of an r-term tensor into a full tensor requires $2r \prod_{j=1}^{d} n_j$ operations (plus lower order terms).*

Proof. Note that $v_\nu^{[1]}[\mathbf{i}'] := \prod_{j=2}^{d} v_\nu^{(j)}[i_j']$ can be obtained for all $\mathbf{i}' \in \mathbf{I}'$ (cf. (7.15)) by $r(d-1) \prod_{j=2}^{d} n_j$ operations. This is a lower order term compared with the operation count for $\sum_{\nu=1}^{r} v_\nu^{(1)}[i_1] \cdot v_\nu^{[1]}[\mathbf{i}']$. □

7.6.3 From r -Term into N-Term Format with r > N

Here, we assume that $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)} \in \mathbf{V}$ is given with $r > N$, where N from (7.14) is the upper bound N of the maximal rank. Such a situation may occur after operations, where the number of terms is the product of those of the operands (see, e.g., §13.5).

Remark 7.17. Let $v_{\mathbf{i}'}^{(j)}$ for $2 \le j \le d$ be the unit vectors from (7.16). Then the tensor $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$ is equal to

$$\mathbf{v} = \sum_{\mathbf{i}' \in \mathbf{I}'} \bigotimes_{j=1}^{d} v_{\mathbf{i}'}^{(j)} \in \mathcal{R}_N \qquad \text{with} \quad v_{\mathbf{i}'}^{(1)} := \sum_{\nu=1}^{r} \left(\prod_{j=2}^{d} v_\nu^{(j)}[i'_j] \right) v_\nu^{(1)}.$$

The number of terms is $\#\mathbf{I}' = N$. The computational cost is $2r \prod_{j=1}^{d} n_j$ plus lower order terms.

The performed operations are identical to those from §7.6.3, only the interpretation of the result as N-term format is different.

Another conversion for the particular case $d = 3$ is related to [119, Remark 2.7]. Below, $e^{(1,i)}$ ($i \in I_1$) is again the i-th unit vector in $\mathbb{K}^{I_1}$.

Remark 7.18. Without loss of generality, assume that $n_1 = \min_j n_j$. Write $\mathbf{v}$ as $\sum_{i \in I_1} e^{(1,i)} \otimes w_i^{[1]}$ with tensors $w_i^{[1]} \in V_2 \otimes V_3$ defined by $w_i^{[1]}[i_2, i_3] := \mathbf{v}[i, i_2, i_3]$. For $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{3} v_\nu^{(j)}$, all $w_i^{[1]}$ have the form

$$w_i^{[1]} = \sum_{\nu=1}^{r} v_\nu^{(1)}[i] \cdot v_\nu^{(2)} \otimes v_\nu^{(3)}.$$

Evaluation of $w_i^{[1]}$ for all i in full format costs $2rn_1n_2n_3 + (r - n_1)n_2n_3$ operations. Note that the leading term is the same as above. The reduced singular value decomposition $w_i^{[1]} = \sum_{\mu=1}^{m_i} x_{i,\mu} \otimes y_{i,\mu}$ is an m_i-term representation, where the (matrix and tensor) rank $m_i = \operatorname{rank}(w_i^{[1]})$ is bounded by $\min\{n_2, n_3\}$. The cost of the SVDs is $O(n_1 \cdot N_{\mathsf{SVD}}(n_2, n_3)) = O(n^4)$, where $n = \max_j n_j$. Because of $r > N = O(n^2)$, $O(n^4)$ is a lower order term compared with $2rn_1n_2n_3$. Thus, this method needs almost the same computational work, while the resulting N'-term representation is

$$\mathbf{v} = \sum_{(i,\mu)} e^{(1,i)} \otimes x_{i,\mu} \otimes y_{i,\mu} \quad \text{with } N' := \sum_{i \in I_1} m_i \le n_1 \min\{n_2, n_3\} = N.$$

7.6.4 Sparse Grid Approach

The sparse grid approach is used to interpolate functions in higher spatial dimensions or it serves as ansatz for discretising partial differential equations. For a review of sparse grids we refer to Bungartz-Griebel [29]. Here, we only sketch the main line and its relation to tensor representations. To simplify the notation, we assume that the tensor space $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ uses identical spaces $V = V_j$, which allow a nested sequence of subspaces:

$$V = V_{(\ell)} \supset V_{(\ell-1)} \supset \ldots \supset V_{(2)} \supset V_{(1)}. \tag{7.17}$$

A typical example are finite element spaces $V_{(\ell)}$ of functions, say, on the interval $[0, 1]$ using the step size $2^{-\ell}$. For sparse grids in Fourier space compare Sprengel [174]. While the usual uniform discretisation by $\mathbf{V} = \otimes^d V_{(\ell)}$ has a dimension of order $2^{-\ell d}$, the sparse grid approach uses the sum of tensor spaces

$$\mathbf{V}_{\mathrm{sg},\ell} = \sum_{\sum_{j=1}^{d} \ell_j = \ell+d-1} \bigotimes_{j=1}^{d} V_{(\ell_j)}. \tag{7.18}$$

The background is the estimation of the interpolation error[5] by $O(2^{-2\ell}\ell^{d-1})$ for functions of suitable regularity (cf. [29, Theorem 3.8]). This is to be compared with $\dim(\mathbf{V}_{\mathrm{sg}}) \approx 2^{\ell}\ell^{d-1}$ (cf. [29, (3.63)]). The basis vectors in $\mathbf{V}_{\mathrm{sg},\ell}$ are elementary tensors $\bigotimes_{j=1}^{d} b_{k,\ell_j}^{(j)}$, where ℓ_j denotes the level: $b_{k,\ell_j}^{(j)} \in V_{(\ell_j)}$. Since the number of terms is limited by the dimension of $\mathbf{V}_{\mathrm{sg},\ell}$, the tensor $\mathbf{v} \in \mathbf{V}_{\mathrm{sg},\ell}$ belongs to $\mathcal{R}_r$ with $r = \dim(\mathbf{V}_{\mathrm{sg},\ell}) \approx 2^{-\ell}\ell^{d-1}$.

For the practical implementation one uses hierarchical bases (cf. [29, §3]). In $V_{(1)}$ we choose, e.g., the standard hat function basis $\mathfrak{b}_1 := (b_i)_{1\le i\le n_1}$. The basis $\mathfrak{b}_2$ of $V_{(2)}$ is $\mathfrak{b}_1$ enriched by $n_2/2$ hat functions from $V_{(2)}$. The latter additional basis functions are indexed by $n_1 + 1 \le i \le n_1 + n_2/2 = n_2$. In general, the basis $\mathfrak{b}_\lambda \subset V_{(\lambda)}$ consists of $\mathfrak{b}_{\lambda-1}$ and additional $n_\lambda/2$ hat functions of $V_{(\lambda)}$. The index ℓ_j corresponds to the dimension $n_j = 2^{\ell_j}$ of $V_{(\ell_j)}$. The additive side condition $\sum_{j=1}^{d} \ell_j \le L := \ell + d - 1$ in (7.18) can be rewritten as $\prod_{j=1}^{d} n_j \le N := 2^L$. It follows from $i_j \le n_j$ that the involved indices of the basis functions $b_{i_j} \in V_{(\ell_j)}$ satisfy $\prod_{j=1}^{d} i_j \le N$. For ease of notation, we replace $\mathbf{V}_{\mathrm{sg},\ell}$ by

$$\mathbf{V}_{\mathrm{sg}} := \operatorname{span}\left\{\bigotimes_{j=1}^{d} b_{i_j} : \prod_{j=1}^{d} i_j \le N\right\}.$$

Since $\mathbf{V}_{\mathrm{sg}} \supset \mathbf{V}_{\mathrm{sg},\ell}$, the approximation is not worse, while $\dim(\mathbf{V}_{\mathrm{sg}})$ has the same asymptotic behaviour $2^{-\ell}\ell^{d-1}$ as $\dim(\mathbf{V}_{\mathrm{sg},\ell})$. The inequality $\prod_{j=1}^{d} i_j \le N$ gives rise to the name 'hyperbolic cross'.

Remark 7.19. The typical hyperbolic cross approach is the approximation of a function f with the (exact) series expansion $f = \sum_{\mathbf{i}\in\mathbb{N}^d} \mathbf{v}_{\mathbf{i}} \bigotimes_{j=1}^{d} \phi_{i_j}$ by

$$f_N := \sum_{\prod_{j=1}^{d} i_j \le N} \mathbf{v}_{\mathbf{i}} \bigotimes_{j=1}^{d} \phi_{i_j}.$$

The behaviour of the number $\sigma_d(N)$ of tuples $\mathbf{i}$ involved in the summation with respect to N is

$$\sigma_d(N) = O\big(N \log^{d-1}(N)\big). \tag{7.19}$$

In the previous example, $O(N^{-2}\log^{d-1}(N))$ is the accuracy of f_N. The following table shows the values of $\sigma_2(N)$ and $\sigma_{10}(N)$ for different N as well as values of $\sigma_d(10)$ for increasing d:

$d = 2$	N	2	4	8	16	32	64	128	256
	$\sigma_d(N)$	3	8	20	50	119	280	645	1466
$d = 10$	N	2	4	8	16	32	64	128	256
	$\sigma_d(N)$	11	76	416	2 056	9 533	41 788	172 643	675 355
$N = 10$	d	2	3	5	10	20	50	100	1000
	$\sigma_d(N)$	27	53	136	571	2 841	29 851	202 201	170 172 001

[5] Any L^p norm with $2 \le p \le \infty$ can be chosen.

7.6.5 From Sparse Format into r-Term Format

Finally, we consider the sparse format $\mathbf{v} = \rho_{\mathrm{sparse}}(\mathring{\mathbf{I}}, (\mathbf{v_i})_{\mathbf{i}\in\mathring{\mathbf{I}}})$ from (7.5). By definition, $\mathbf{v} = \sum_{\mathbf{i}\in\mathring{\mathbf{I}}} \mathbf{v_i} \bigotimes_{j=1}^{d} b_{i_j}^{(j)}$ holds. The latter expression is an r-term representation of $\mathbf{v}$ with $r := \#\mathring{\mathbf{I}}$ nonzero terms.

The function f from Remark 7.19 is a tensor isomorphically represented by the coefficient $\mathbf{v} \in \otimes^d \mathbb{K}^{\mathbb{N}}$, where the tensor space is to equipped with the suitable norm. The tensor $\mathbf{v}_{\mathrm{sg}}$ corresponding to the approximation f_N has sparse format: $\mathbf{v}_{\mathrm{sg}} = \rho_{\mathrm{sparse}}(\mathring{\mathbf{I}}, (\mathbf{v_i})_{\mathbf{i}\in\mathring{\mathbf{I}}})$ with $\mathring{\mathbf{I}} = \{\mathbf{i} \in \mathbb{N}^d : \prod_{j=1}^{d} i_j \leq N\}$. This ensures an r-term representation with $r := \#\mathring{\mathbf{I}}$. The representation rank r can be reduced because of the special structure of $\mathring{\mathbf{I}}$. Here, we follow the idea from the proof of Lemma 3.41: for fixed i_j with $j \in \{1, \ldots, d\}\backslash\{k\}$ we can collect all terms for $i_k \in \mathbb{N}$ in

$$\left(\bigotimes_{j=1}^{k-1} b_{i_j}^{(j)}\right) \otimes \left(\sum_{i_k} \mathbf{v_i} b_{i_k}^{(k)}\right) \otimes \left(\bigotimes_{j=k+1}^{d} b_{i_j}^{(j)}\right) \in \mathcal{R}_1. \tag{7.20}$$

The obvious choice of $(i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_d)$ are indices such that the sum $\sum_{i_k}$ contains as many nonzero terms as possible.

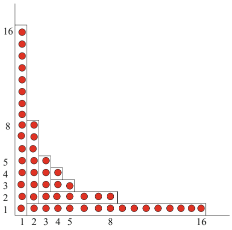


Fig. 7.1 Sparse grid indices

First, we discuss the situation for $d = 2$. Fig. 7.1 shows the pairs $(i_1, i_2) \in \mathring{\mathbf{I}}$ with $\prod_{j=1}^{d} i_j \leq N = 16$. For the first choice $k = 1$ and $i_2 = 1$, the indices $\mathbf{i}$ involved in (7.20) are contained in the first column of height 16. The second choice $k = 2$ and $i_1 = 1$ leads to the lower row. Here, the sum in (7.20) ranges from 2 to 16, since $\mathbf{v}[1, 1]$ belongs to the previous column. Similarly, two further columns and rows correspond to $i_2 = 2, 3$ and $i_1 = 2, 3$. Then we are left with a single index $\mathbf{i} = (4, 4)$ so that we have decomposed $\mathring{\mathbf{I}}$ into seven groups. Each group gives rise to one elementary tensor (7.20). This finishes the construction of a 7-term representation of $\mathbf{v}_{\mathrm{sg}}$. Obviously, for general N we can construct a representation in $\mathcal{R}_r$ with

$$r \leq 2\lfloor\sqrt{N}\rfloor \leq 2\sqrt{N}$$

and even $r \leq 2\sqrt{N} - 1$ if $\sqrt{N} \in \mathbb{N}$.

For general d, the decomposition of $\mathring{\mathbf{I}}$ can be achieved as follows. Let

$$T := \left\{(t_1, \ldots, t_{d-1}) \in \mathbb{N}^{d-1} : \max_{j=1}^{d-1}\{t_j\} \cdot \prod_{j=1}^{d-1} t_j \leq N\right\}$$

be a set of $(d-1)$-tuples. For each $t := (t_1, \ldots, t_{d-1}) \in T$ and $1 \leq k \leq d$ define

$$\mathring{\mathbf{I}}_{t,k} := \left\{(t_1, \ldots, t_{k-1}, i_k, t_k, \ldots, t_{d-1}) \in \mathring{\mathbf{I}} \quad \text{with } i_k \geq \max_{j=1}^{d-1}\{t_j\}\right\}.$$

We claim that $\bigcup_{t\in T} \bigcup_{k=1}^{d} \mathring{\mathbf{I}}_{t,k} = \mathring{\mathbf{I}}$. For a proof take any $\mathbf{i} = (i_1, \ldots, i_d) \in \mathring{\mathbf{I}}$ and let k and m be indices of the largest and second largest i_j, i.e.,

$$i_k \geq i_m \geq i_j \quad \text{for all } j \in \{1, \ldots, d\} \backslash \{k, m\} \text{ with } k \neq m.$$

Inequality $i_m \leq i_k$ and $\mathbf{i} \in \mathring{\mathbf{I}}$ imply that $i_m \cdot \prod_{j \neq k} i_j \leq \prod_{j=1}^d i_j \leq N$. Therefore the tuple $t := (i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_d)$ belongs to T and shows that $\mathbf{i} \in \mathring{\mathbf{I}}_{t,k}$. This proves $\mathring{\mathbf{I}} \subset \bigcup_{t \in T} \bigcup_{k=1}^d \mathring{\mathbf{I}}_{t,k}$, while direction '$\supset$' follows by definition of $\mathring{\mathbf{I}}_{t,k}$.

We conclude that $\{\mathring{\mathbf{I}}_{t,k} : t \in T, 1 \leq k \leq d\}$ is a (not necessarily disjoint) decomposition of $\mathring{\mathbf{I}}$, whose cardinality is denoted by $\tau_d(N)$. Each[6] set $\mathring{\mathbf{I}}_{t,k}$ gives rise to an elementary tensor (7.20) and proves $\mathbf{v}_{\text{sg}} \in \mathcal{R}_r$, where[7]

$$r := \tau_d(N) \leq d \cdot \#T.$$

It remains to estimate $\#T$. For $t := (t_1, \ldots, t_{d-1}) \in T$ let m be an index with $t_m = \max_{j=1}^{d-1}\{t_j\}$. From $t_m^2 \prod_{j \neq m} t_j \leq N$ we conclude that $1 \leq t_m \leq \sqrt{N}$. In the following, we distinguish the cases $t_m \leq N^{1/d}$ and $N^{1/d} < t_m \leq N^{1/2}$.

If $t_m \leq N^{1/d}$, also $t_j \leq N^{1/d}$ holds and all such values satisfy the condition $\max_{j=1}^{d-1}\{t_j\} \cdot \prod_{j=1}^{d-1} t_j \leq N$. The number of tuples $t \in T$ with $\max_j\{t_j\} \leq N^{1/d}$ is bounded by

$$N^{(d-1)/d}.$$

Now we consider the case $N^{1/d} < t_m \leq N^{1/2}$. The remaining components t_j $(j \neq k)$ satisfy $\prod_{j \neq m} t_j \leq N/t_m^2$. We ignore the condition $t_j \leq t_m$, and ask for all $(d-2)$-tuples $(t_j : j \in \{1, \ldots, d-1\} \backslash \{m\})$ with $\prod_{j \neq m} t_j \leq N/t_m^2$. Its number is $\sigma_{d-2}(N/t_m^2) = O\big(\frac{N}{t_m^2} \log^{d-3}(\frac{N}{t_m^2})\big)$ (cf. (7.19)). It remains to bound the sum $\sum_{N^{1/d} < t \leq N^{1/2}} \frac{N}{t^2} \log^{d-3}(\frac{N}{t^2})$. Instead, we consider the integral

$$\int_{N^{1/d}}^{N^{1/2}} \frac{N}{x^2} \log^{d-3}\left(\frac{N}{x^2}\right) \mathrm{d}x < N \log^{d-3}(N^{(d-2)/d}) \int_{N^{1/d}}^{N^{1/2}} \frac{\mathrm{d}x}{x^2}$$
$$< N \log^{d-3}(N^{(d-2)/d})[N^{-1/d} - N^{-1/2}] < N^{(d-1)/d} \log^{d-3}(N^{(d-2)/d}).$$

Therefore, any tensor $\mathbf{v} = \rho_{\text{sparse}}(\mathring{\mathbf{I}}, (\mathbf{v_i})_{\mathbf{i} \in \mathring{\mathbf{I}}}$ can be written as $\mathbf{v} \in \mathcal{R}_r$ with representation rank $r \leq O\big(N^{(d-1)/d} \log^{d-3}(N)\big)$. This proves that although $\#\mathring{\mathbf{I}}$ is not bounded by $O(N)$, the rank is strictly better than $O(N)$.

Proposition 7.20. *For an index set* $\mathring{\mathbf{I}} \subset \{\mathbf{i} \in \mathbb{N}^d : \prod_{j=1}^d i_j \leq N\}$, *any tensor* $\mathbf{v} = \rho_{\text{sparse}}(\mathring{\mathbf{I}}, (\mathbf{v_i})_{\mathbf{i} \in \mathring{\mathbf{I}}})$ *can be explicitly converted into* $\mathbf{v} \in \mathcal{R}_r$ *with a representation rank* $r = \tau_d(N) \leq O\big(N^{(d-1)/d} \log^{d-3}(N)\big)$.

The factor $\log^{d-3}$ may be an artifact of the rough estimate. Numerical tests show that the asymptotic behaviour appears rather late. The next table shows

$$\gamma_d(N) := \log_2(\tau_d(N)/\tau_d(N/2)) \qquad \text{for } N = 2^n.$$

[6] Since the sets $\mathring{\mathbf{I}}_{t,k}$ may overlap, one must take care that each $\mathbf{v_i}$ is associated to only one $\mathring{\mathbf{I}}_{t,k}$.

[7] $\tau_d(N) < d \cdot \#T$ may occur, since $\mathring{\mathbf{I}}_{t,k} = \mathring{\mathbf{I}}_{t,k'}$ may hold for $k \neq k'$. An example is the decomposition from Fig. 7.1, where $\tau_2(16) = 7$.

$\gamma_d(N)$ should converge to $(d-1)/d$.

n	3	6	9	12	15	18	21	24	27	30	
$\gamma_3(2^n)$	1.0	0.88	0.78	0.73	0.71	0.693	0.684	0.679	0.675	0.672	$\to \frac{2}{3}$
$\gamma_4(2^n)$	1.4	1.15	0.96	0.89	0.85	0.825	0.806	0.794	0.784	0.777	$\to \frac{3}{4}$

Finally, we compare the quantities $\sigma_d(N)$ (cardinality of the sparse grid) and $\tau_d(N)$ for $d = 3$ and $d = 6$:

N	10	100	1 000	10 000	100 000	1 000 000
$\sigma_3(N)$	53	1 471	29 425	496 623	7 518 850	106 030 594
$\tau_3(N)$	12	102	606	3265	16542	81 050
$\sigma_6(N)$	195	14 393	584 325	17 769 991	439 766 262	-
$\tau_6(N)$	51	2 047	36 018	502 669	5 812 401	59 730 405

7.7 Modifications

Restricting V_j to certain subsets $A_j \subset V_j$, we can define a modified r-term format $\mathcal{R}_r\left((A_j)_{j=1}^d\right)$ by

$$\mathcal{R}_r\left((A_j)_{j=1}^d\right) := \left\{ \sum_{\nu=1}^{r} v_\nu^{(1)} \otimes \ldots \otimes v_\nu^{(d)} \, : \, v_\nu^{(j)} \in A_j \right\} \qquad (r \in \mathbb{N}_0).$$

Examples of A_j are

- $V_j = \mathbb{R}^{n_j}$, $A_j = \{v \in V_j : v_i \geq 0\}$. Hence, A_j contains the non-negative vectors.
- $V_j = L^p(I_j)$, then $A_j = \{v \in V_j : v(x) \geq 0 \text{ for } x \in I_j\}$ describes the non-negative functions.
- $V_j = \mathbb{C}^{n_j \times n_j}$, $A_j = \{M \in V_j : M \text{ Hermitean matrix}\}$.
- $V_j = \mathbb{C}^{n_j \times n_j}$, $A_j = \{M \in V_j : M \text{ positive semidefinite}\}$

We shall refer to these modifications in §9.6, §9.7.1, and §9.7.2.5.1.

A challenging problem is the representation of (anti-)symmetric tensors. Assume, e.g., that $\mathbf{v} \in \mathfrak{A}(V) \subset \otimes^d V$ is an antisymmetric tensor (cf. §3.5). A possible representation uses a usual tensor $\mathbf{v}' \in \otimes^d V$ with the property $\mathbf{v} = P_{\mathfrak{A}}(\mathbf{v}')$. Note that an elementary tensor $\mathbf{v}'$ leads to the Slater determinant $\mathbf{v}$ requiring $d!$ terms (cf. Lemma 3.70). The difficulty of this approach comes with the operations. For instance, the scalar product $\langle \mathbf{v}, \mathbf{w} \rangle$ of two antisymmetric tensors $\mathbf{v}, \mathbf{w} \in \mathfrak{A}(V)$ is to be expressed by means of their representations $\mathbf{v}'$ and $\mathbf{w}'$. For a solution of this problem compare Beylkin-Mohlenkamp-Pérez [16].

Chapter 8
Tensor Subspace Representation

Abstract We use the term '*tensor subspace*'[1] for the tensor product $\mathbf{U} := {}_a\bigotimes_{j=1}^{d} U_j$ of subspaces $U_j \subset V_j$. Obviously, $\mathbf{U}$ is a subspace of $\mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j$, but not any subspace of $\mathbf{V}$ is a tensor subspace.[2] For $d = 2$, r-term and tensor subspace representations (also called *Tucker representation*) are identical. Therefore, both approaches can be viewed as extensions of the concept of rank-r matrices to the tensor case $d \geq 3$. The resulting set $\mathcal{T}_\mathbf{r}$ introduced in *Sect. 8.1* will be characterised by a vector-valued rank $\mathbf{r} = (r_1, \ldots, r_d)$. Since by definition, tensors $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ are closely related to subspaces, their descriptions by means of frames or bases is of interest (see *Sect. 8.2*). Differently from the r-term format, algebraic tools like the singular value decomposition can be applied and lead to a higher order singular value decomposition (HOSVD), which is a quite important feature of the tensor subspace representation (cf. *Sect. 8.3*). Moreover, HOSVD yields a connection to the minimal subspaces from Chap. 6. In *Sect. 8.5* we compare the formats discussed so far and describe conversions between the formats. In a natural way, a hybrid format appears using the r-term format for the coefficient tensor of the tensor subspace representation (cf. §8.2.4). *Section 8.6* deals with the problem of joining two representation systems, as it is needed when we add two tensors involving different tensor subspaces.

8.1 The Set $\mathcal{T}_\mathbf{r}$

Consider an algebraic tensor space

$$\mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$$

[1] 'Tensor subspace' is to be understood as 'subspace and tensor space'.

[2] For instance, choose two linearly independent vectors $a, b \in V$. Then $\mathbf{U} := \mathrm{span}\{a \otimes a, b \otimes b\}$ is a two-dimensional subspace of $V \otimes V$, but the smallest tensor subspace containing $\mathbf{U}$ is $V_0 \otimes V_0$ with $V_0 := \mathrm{span}\{a, b\}$ and has dimension four.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_8,

and a fixed tensor $\mathbf{v} \in \mathbf{V}$. We want to find possibly lower dimensional subspaces $U_j \subset V_j$ such that

$$\mathbf{v} \in \mathbf{U} := {}_a\bigotimes_{j=1}^{d} U_j \qquad \text{with } r_j := \dim(U_j) \tag{8.1}$$

holds. In §8.2 we shall represent $\mathbf{v}$ by means of chosen bases of U_j. A possible choice of $U_j \subset V_j$ for all $1 \leq j \leq d$ are the minimal subspaces $U_j^{\min}(\mathbf{v})$ described in §6. Note that in this chapter only *exact* representations are considered. Approximations in possibly even smaller subspaces will be discussed in §10.

Above, we have started with $\mathbf{v} \in \mathbf{V}$ and have been looking for a subspace family $\{U_j\}$ with property (8.1). Now, we reverse the viewpoint and fix the dimensions r_j of U_j,

$$\mathbf{r} := (r_1, \ldots, r_d) \in \mathbb{N}_0^d, \tag{8.2}$$

and ask for all $\mathbf{v}$ satisfying (8.1) for subspaces U_j which may depend on $\mathbf{v}$. The dimensions[3] r_j play a similar rôle as the parameter r of the r-term representation. Therefore, we refer to $\mathbf{r}$ as the *tensor subspace representation rank*. Since it is vector-valued, it cannot be confounded with the tensor rank from §3.2.6.

Definition 8.1 ($\mathcal{T}_\mathbf{r}$). Let $\mathbf{V} := {}_a\bigotimes_{j=1}^{d} V_j$, fix $\mathbf{r} := (r_1, \ldots, r_d) \in \mathbb{N}_0^d$, and set

$$\mathcal{T}_\mathbf{r} := \mathcal{T}_\mathbf{r}(\mathbf{V}) := \left\{ \mathbf{v} \in \mathbf{V} : \begin{array}{l} \text{there are subspaces } U_j \subset V_j \text{ such that} \\ \dim(U_j) = r_j \text{ and } \mathbf{v} \in \mathbf{U} := \bigotimes_{j=1}^{d} U_j \end{array} \right\}. \tag{8.3}$$

For $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ we say that $\mathbf{v}$ possesses an $(r_1, \ldots, r_d)$*-tensor subspace representation*.

The symbol $\mathcal{T}_\mathbf{r}$ is used if the reference to the underlying tensor space $\mathbf{V}$ is obvious; otherwise, $\mathcal{T}_\mathbf{r}(\mathbf{V})$ is preferred.

An equivalent definition is

$$\mathcal{T}_\mathbf{r} = \bigcup_{U_j \subset V_j \text{ subspaces with } \dim(U_j) = r_j \ (1 \leq j \leq d)} \ \bigotimes_{j=1}^{d} U_j.$$

The letter $\mathcal{T}$ may also be read as 'Tucker format' (cf. Tucker [184]).

Note that the subspaces U_j involved in (8.3) vary with $\mathbf{v}$. The set $\mathcal{T}_\mathbf{r}$ corresponds to $\mathcal{R}_r$ from (3.22). The parameters $r_1, \ldots, r_d$ are not necessarily the optimal ones, i.e., the subspaces U_j may be larger than necessary.

Exercise 8.2. (a) If $r_j = 0$ for some j, then $\mathcal{T}_\mathbf{r} = \mathcal{R}_0 = \{0\}$ is the trivial subspace.
(b) Coincidence $\mathcal{T}_\mathbf{1} = \mathcal{R}_1$ holds for $\mathbf{1} = (1, \ldots, 1) \in \mathbb{N}_0^d$.
(c) If $d = 2$, the identity $\mathcal{T}_\mathbf{r} = \mathcal{R}_r$ holds for $\mathbf{r} = (r, \ldots, r)$ and all $r \in \mathbb{N}_0$.

Except for the cases from Exercise 8.2, the sets $\mathcal{T}_\mathbf{r}$ and $\mathcal{R}_r$ do not coincide. Relations and conversions between both formats will be discussed in §8.5.

Assume that $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ holds with $\mathbf{v} \in \mathbf{U} := \bigotimes_{j=1}^{d} U_j$ and $\dim(U_j) = r_j$. For

[3] Obviously, only integer $r_j \leq \dim(V_j)$ are of interest.

any $s_j \geq r_j$ there are larger subspaces $W_j \supset U_j$ with $\dim(W_j) = s_j$. Obviously, $\mathbf{v} \in \mathbf{W} := \bigotimes_{j=1}^d W_j$ holds and shows that also $\mathbf{v} \in \mathcal{T}_\mathbf{s}$ is valid, i.e.,[4]

$$\mathbf{v} \in \mathcal{T}_\mathbf{r} \;\Rightarrow\; \mathbf{v} \in \mathcal{T}_\mathbf{s} \quad \text{for } \mathbf{s} \geq \mathbf{r}.$$

This proves the following statement.

Corollary 8.3. In definition (8.3) we may replace $\dim(U_j) = r_j$ by $\dim(U_j) \leq r_j$ without changing the set $\mathcal{T}_\mathbf{r}$.

$\mathcal{T}_\mathbf{r}$ satisfies similar properties as $\mathcal{R}_r$ in (3.23a):

$$\begin{array}{ll} \{0\} = \mathcal{T}_\mathbf{r} & \text{if } r_j = 0 \text{ for at least one } j, \\ \mathcal{T}_\mathbf{r} \subset \mathcal{T}_\mathbf{s} & \text{for } \mathbf{r} \leq \mathbf{s}, \\ \mathcal{T}_\mathbf{r} + \mathcal{T}_\mathbf{s} \subset \mathcal{T}_{\mathbf{r}+\mathbf{s}} & \text{for all } \mathbf{r}, \mathbf{s} \in \mathbb{N}_0^d. \end{array} \tag{8.4}$$

Note that $\mathcal{T}_\mathbf{r}$ is not a subspace! Two different tensors $\mathbf{v}, \mathbf{w} \in \mathcal{T}_\mathbf{r}$ may belong to different systems of subspaces: $\mathbf{v} \in \bigotimes_{j=1}^d U_j$ and $\mathbf{w} \in \bigotimes_{j=1}^d W_j$. In the worst case, $U_j \cap W_j = \{0\}$ holds and the sum $\mathbf{v} + \mathbf{w}$ requires the subspace $U_j + W_j$ with $\dim(U_j + W_j) = r_j + s_j$ for its tensor subspace representation, proving the last line of (8.4). Differently from (3.23b), $\mathcal{T}_\mathbf{r} + \mathcal{T}_\mathbf{s}$ is a proper subset of $\mathcal{T}_{\mathbf{r}+\mathbf{s}}$ if $\mathbf{r}, \mathbf{s} \neq 0$.

Summarising the results of §6.3, we can state:

Remark 8.4 (Tucker rank). Given $\mathbf{v} \in \mathbf{V}$, there is a *minimal* $\mathbf{r} = \mathbf{r}_{\min}(\mathbf{v}) \in \mathbb{N}_0^d$ with $\mathbf{v} \in \mathcal{T}_\mathbf{r}$. This $\mathbf{r}_{\min}(\mathbf{v})$ has the components

$$r_j = \operatorname{rank}_j(\mathbf{v}) = \dim\left(U_j^{\min}(\mathbf{v})\right)$$

(cf. (5.6b)). The corresponding subspaces from (8.3) are $U_j := U_j^{\min}(\mathbf{v})$. The vector $\mathbf{r}_{\min}(\mathbf{v})$ is called the *tensor subspace rank* or *'Tucker rank'* of $\mathbf{v}$ (although this rank is much earlier introduced by Hitchcock [100]).

Example 8.5. (a) Let $\mathcal{P}_{p_j} \subset V_j := L^2([0,1])$ be the subspace of polynomials of degree at most p_j. All multivariate polynomials $f(x_1, \ldots, x_d) \in \mathbf{V} := {}_a\bigotimes_{j=1}^d \mathbf{V}_j$ with polynomial degree $\leq p_j$ with respect to x_j belong to $\mathbf{U} := \bigotimes_{j=1}^d \mathcal{P}_{p_j} \subset \mathcal{T}_\mathbf{r}$ with $r_j = p_j + 1$.
(b) The particular polynomial $f(x, y, z) = xz + x^2 y$ belongs to $\mathcal{T}_{(2,2,2)}$ involving the subspaces $U_1 := \operatorname{span}\{x, x^2\}$, $U_2 := \operatorname{span}\{1, y\}$, $U_3 := \operatorname{span}\{1, z\}$.

The next property of $\mathcal{T}_\mathbf{r}$ will become important for approximation problems.

Lemma 8.6. *Let $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ be a Banach tensor space with a norm not weaker than $\|\cdot\|_\vee$ (cf. (6.18)). Then the subset $\mathcal{T}_\mathbf{r} \subset \mathbf{V}$ is weakly closed.*

Proof. Let $\mathbf{v}_n \in \mathcal{T}_\mathbf{r}$ be a weakly convergent sequence with $\mathbf{v}_n \rightharpoonup \mathbf{v} \in \mathbf{V}$. From $\mathbf{v}_n \in \mathcal{T}_\mathbf{r}$ we infer that $U_j^{\min}(\mathbf{v}_n)$ has a dimension not exceeding r_j. By Theorem 6.24, $\dim(U_j^{\min}(\mathbf{v})) \leq r_j$ follows, implying $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ (cf. Theorem 6.26). □

[4] Inequalities $\mathbf{s} \geq \mathbf{r}$ for vectors from $\mathbb{N}_0^d$ are understood componentwise: $s_j \geq r_j$ for all $1 \leq j \leq d$.

8.2 Tensor Subspace Formats

8.2.1 General Frame or Basis

The characterisation of a tensor by $\mathbf{v} \in \mathbf{U} := \bigotimes_{j=1}^{d} U_j \subset \mathbf{V} := \bigotimes_{j=1}^{d} V_j$ corresponds to the (theoretical) level of linear algebra. The numerical treatment requires a description of the subspaces by a frame[5] or basis. Even if a basis (in contrast to a frame) is the desired choice, there are intermediate situations, where frames cannot be avoided (cf. §8.6). By obvious reasons, we have to suppose that $\dim(U_j) < \infty$ (cf. Remark 6.1). Without loss of generality, we enumerate the frame vectors of U_j by $b_i^{(j)}$, $1 \le i \le r_j$, and form the r_j-tuple

$$B_j := \left[b_1^{(j)}, b_2^{(j)}, \ldots, b_{r_j}^{(j)}\right] \in (V_j)^{r_j}. \tag{8.5a}$$

Set

$$\mathbf{J} = J_1 \times \ldots \times J_d \quad \text{with} \quad J_j = \{1 \le i \le r_j\} \quad \text{for } 1 \le j \le d.$$

$B_j \in (V_j)^{r_j}$ can be considered as an element from the set $\mathcal{L}(\mathbb{K}^{J_j}, V_j)$. In the case of $V_j = \mathbb{K}^{I_j}$, B_j is a matrix:

$$B_j \in \mathbb{K}^{I_j \times J_j}. \tag{8.5b}$$

The elementary Kronecker product

$$\mathbf{B} := \bigotimes_{j=1}^{d} B_j \in \mathcal{L}(\mathbb{K}^{\mathbf{J}}, \mathbf{V}) \tag{8.5c}$$

becomes a matrix from $\mathbb{K}^{\mathbf{I}\times\mathbf{J}}$, if $V_j = \mathbb{K}^{I_j}$ and $\mathbf{V} = \mathbb{K}^{\mathbf{I}}$. In the following, the frame data will be described by $B_j \in (V_j)^{r_j}$ or $B_j \in \mathbb{K}^{I_j \times J_j}$ for $1 \le j \le d$, which includes the information about $r_j = \#J_j$. These quantities define $\mathbf{B}$ by (8.5c). A column of $\mathbf{B}$ corresponding to a multi-index $\mathbf{i} \in \mathbf{I}$ is $\mathbf{b_i} = \bigotimes_{j=1}^{d} b_{i_j}^{(j)}$. Hence, all columns of $\mathbf{B}$ form the frame [or basis] of $\mathbf{U} \subset \mathbf{V}$. For later use we add that for any index subset $\emptyset \subsetneqq \alpha \subsetneqq \{1, \ldots, d\}$ a frame of $\mathbf{U}_\alpha = \bigotimes_{j\in\alpha} U_j$ is denoted by

$$\mathbf{B}_\alpha := \left[\mathbf{b}_1^{(\alpha)}, \mathbf{b}_2^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}\right] \in (\mathbf{V}_\alpha)^{r_\alpha}. \tag{8.5d}$$

We specify the following data:

$$\begin{array}{ll} B_j \in (V_j)^{r_j} & \text{frame or basis of } U_j \text{ for } 1 \le j \le d,, \\ J_j := \{1 \le i \le r_j\} & \text{for } 1 \le j \le d, \\ \mathbf{a} \in \bigotimes_{j=1}^{d} \mathbb{K}^{J_j} = \mathbb{K}^{\mathbf{J}} & \text{for } \mathbf{J} = J_1 \times \ldots \times J_d \end{array} \tag{8.6a}$$

[5] The *frame* is a system of vectors generating the subspace without assuming linear independence. When the term ‘frame’ is used, this does not exclude the special case of a basis; otherwise, we use the term ‘proper frame’. Note that a frame cannot be described by a set $\{b_\nu^{(j)} : 1 \le \nu \le r_j\}$, since $b_\nu^{(j)} = b_\mu^{(j)}$ may hold for $\nu \ne \mu$.

so that

$$\mathbf{v} = \mathbf{B}\,\mathbf{a} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}_{\mathbf{i}} \bigotimes_{j=1}^{d} b_{i_j}^{(j)} \tag{8.6b}$$

$$= \sum_{i_1=1}^{r_1}\sum_{i_2=1}^{r_2}\cdots\sum_{i_d=1}^{r_d} \mathbf{a}[i_1 i_2 \cdots i_d]\, b_{i_1}^{(1)} \otimes b_{i_2}^{(2)} \otimes \ldots \otimes b_{i_d}^{(d)}.$$

Note that $r_j \geq \dim(U_j)$. Equality $r_j = \dim(U_j)$ holds if and only if B_j is a basis. According to §7.1, the representation is the mapping

$$\rho_{\mathrm{TS}}\big(\mathbf{a}, (B_j)_{j=1}^d\big) := \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}_{\mathbf{i}} \bigotimes_{j=1}^{d} b_{i_j}^{(j)} = \mathbf{B}\,\mathbf{a}. \tag{8.6c}$$

The coefficient tensor $\mathbf{a}$ is also called '*core tensor*' (Tucker [184, p. 287] uses the term 'core matrix').

Formally, representation (8.6c) looks very similar to the full representation (7.3). However, there are two important differences. First, the index set $\mathbf{J}$ is hopefully much smaller than the original index set $\mathbf{I}$. Second, the frame vectors $\{b_i^{(j)}\}$ are of different nature. In the case of the full representation (7.3), B_j is a fixed basis. For instance, for the space $\mathbf{V}$ of multivariate polynomials, $b_i^{(j)} = x_j^i$ may be the monomials, or for $\mathbf{V} = \mathbb{K}^{\mathbf{I}}$ the basis vectors $b_i^{(j)}$ are the unit vectors $e^{(i)} \in \mathbb{K}^{I_j}$ (cf. (2.2)). Because of the fixed (symbolic) meaning, these basis vectors need not be stored. The opposite is true for the representation (8.6c). Here, we have chosen a special frame B_j of U_j and must store the frame vectors $b_i^{(j)}$ explicitly.

A tensor $\mathbf{v} \in \mathcal{T}_{\mathbf{r}}$ may still be represented in different versions. A general one is given next, orthonormality is required in (8.8a), while a special orthonormal basis is used in Definition 8.23.

Remark 8.7 (general tensor subspace representation). (a) The storage requirements of the vectors $b_i^{(j)}$ depend on the nature of U_j (cf. §7.5). Denoting the storage of each frame vector by $size(U_j)$, the basis data require

$$N_{\mathrm{mem}}^{\mathrm{TSR}}\left((B_j)_{j=1}^d\right) = \sum\nolimits_{j=1}^{d} r_j \cdot size(U_j). \tag{8.6d}$$

(b) The coefficient tensor $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$ is given by its full representation (cf. §7.2) and requires a storage of size

$$N_{\mathrm{mem}}^{\mathrm{TSR}}(\mathbf{a}) = \prod_{j=1}^{d} r_j. \tag{8.6e}$$

(c) For the optimal choice $U_j = U_j^{\min}(\mathbf{v})$ together with bases B_j of U_j, the numbers r_j are given by $r_j = \mathrm{rank}_j(\mathbf{v})$ (cf. Remark 8.4).
(d) If, at least for one j, B_j is not a basis, the coefficient tensor $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$ is not uniquely defined.

The counterpart of Remark 7.9 reads as follows.

Remark 8.8. Suppose a representation of $\mathbf{v}$ with $r_j = \mathrm{rank}_j(\mathbf{v})$ for $1 \leq j \leq d$.
(a) If $\mathbf{v}$ satisfies a linear constraint φ_k (as defined in §6.8), then $\varphi_k(b_i^{(k)}) = 0$ holds for all basis vectors $b_i^{(k)}$ from B_k.

(b) Let $\mathbf{V}^{(\mathbf{n})}$ be the intersection Banach spaces from §4.3.6. Then $\mathbf{v} \in \mathbf{V}^{(\mathbf{n})}$ implies $b_i^{(k)} \in V_k^{(n_k)}$ for all $b_i^{(k)}$ from B_k.

Let $n := \max_j size(U_j)$ and $r := \max_j r_j$. Then the memory costs (8.6d,f) sum to $rdn + r^d$. How rdn and r^d compare, depends on the sizes of r and d. If r is small compared with n and if d is small (say $d = 3$), $r^d < rdn$ may hold. For medium sized d, the term r^d becomes easily larger than rdn. For really large d, this term makes the representation infeasible.

The frame B_j may be transformed using $B_j^{\text{new}} = \left[b_{1,\text{new}}^{(j)}, \ldots, b_{r_j^{\text{new}},\text{new}}^{(j)}\right]$ and an $r_j^{\text{new}} \times r_j$ matrix $T^{(j)}$:

$$B_j = B_j^{\text{new}} T^{(j)}, \text{ i.e., } \quad b_i^{(j)} = \sum_{k=1}^{r_j^{\text{new}}} T_{ki}^{(j)}\, b_{k,\text{new}}^{(j)} \qquad \text{for } 1 \le i \le r_j, \tag{8.7a}$$

or, more shortly, $\mathbf{B} = \mathbf{B}_{\text{new}}\mathbf{T}$ with $\mathbf{B} = \bigotimes_{j=1}^{d} B_j$, $\mathbf{B}_{\text{new}} = \bigotimes_{j=1}^{d} B_j^{\text{new}}$, $\mathbf{T} = \bigotimes_{j=1}^{d} T^{(j)}$.

In the case of *bases* B_j and B_j^{new} with $r_j^{\text{new}} = r_j$, the transformation matrix $T^{(j)}$ is regular and the inverse transformation is $b_{k,\text{new}}^{(j)} = \sum_{i=1}^{r_j} S_{ik}^{(j)}\, b_i^{(j)}$ with $S^{(j)} = (T^{(j)})^{-1}$. In general, the inclusion[6] $\operatorname{range}(B_j) \subset \operatorname{range}(B_j^{\text{new}})$ following from (8.7a) ensures that all $\mathbf{v} = \mathbf{B}\mathbf{a}$ can be expressed by means of $\mathbf{B}_{\text{new}}$.

Lemma 8.9 (frame transformation). *Let $\mathbf{v} \in \mathbf{U}$ be described by (8.6a,b) and consider the transformation of B_j^{new} to the frames B_j by means of (8.7a) with matrices $T^{(j)}$, i.e., $\mathbf{B} = \mathbf{B}_{\text{new}}\mathbf{T}$. The corresponding transformation of the coefficients is*

$$\mathbf{a}_{\text{new}} := \mathbf{T}\,\mathbf{a} \qquad \textit{with } \mathbf{T} = \bigotimes_{j=1}^{d} T^{(j)}. \tag{8.7b}$$

Then

$$\rho_{\text{TS}}\left(\mathbf{a}, (B_j)_{j=1}^d\right) = \rho_{\text{TS}}\left(\mathbf{a}_{\text{new}}, (B_j^{\text{new}})_{j=1}^d\right). \tag{8.7c}$$

Proof. $\mathbf{B}\mathbf{a} = (\mathbf{B}_{\text{new}}\mathbf{T})\mathbf{a} = \mathbf{B}_{\text{new}}(\mathbf{T}\mathbf{a}) = \mathbf{B}_{\text{new}}\mathbf{a}_{\text{new}}$ proves (8.7c). □

The elementwise formulation of (8.7b) reads as

$$\begin{aligned} &\mathbf{a}_{\text{new}}[i_1 i_2 \cdots i_d] \\ &= \sum_{k_1=1}^{r_1}\sum_{k_2=1}^{r_2}\cdots\sum_{k_d=1}^{r_d} T^{(1)}[i_1,k_1]T^{(2)}[i_2,k_2]\cdots T^{(d)}[i_d,k_d]\,\mathbf{a}[k_1k_2\cdots k_d]. \end{aligned} \tag{8.7d}$$

One can interpret (8.7c) also in the reverse direction (then, the affices ‘old’ and ‘new’ are to be exchanged).

Corollary 8.10. Let $\mathbf{v} = \rho_{\text{TS}}\left(\mathbf{a}, (B_j)_{j=1}^d\right)$ be given. If the coefficient tensor allows the formulation $\mathbf{a} = \mathbf{S}\,\mathbf{a}_{\text{new}}$, the tensor subspace format can be transformed into

$$\mathbf{v} = \rho_{\text{TS}}\left(\mathbf{a}_{\text{new}}, (B_j^{\text{new}})_{j=1}^d\right) \text{ with } \mathbf{B}_{\text{new}} := \mathbf{B}\,\mathbf{S}.$$

[6] This inclusion does imply that $r_j \le r_j^{\text{new}}$. If B_j is a proper frame and B_j^{new} a basis of the range of B_j, even $r_j^{\text{new}} < r_j$ holds.

8.2.2 Orthonormal Basis

Let $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ be a [pre-]Hilbert space with scalar product $\langle\cdot,\cdot\rangle$ induced by the scalar products $\langle\cdot,\cdot\rangle_j$ of V_j. Consider again the representation (8.6a,b) of $\mathbf{v}\in\mathbf{U}$ with a basis $(B_j)_{1\leq j\leq d}$. In the Hilbert space setting, an orthonormal basis is obviously the desirable choice. An *orthonormal* basis is characterised by $B_j \in \mathcal{L}(\mathbb{K}^{J_j}, V_j)$ with the property

$$B_j^* B_j = I \in \mathbb{K}^{J_j\times J_j} \qquad \text{for } 1 \leq j \leq d.$$

A consequence is $\mathbf{B}^*\mathbf{B} = \mathbf{id} \in \mathbb{K}^{\mathbf{J}\times\mathbf{J}}$. Note that in the matrix case B_j^* and $\mathbf{B}^*$ are written as B_j^{H} and $\mathbf{B}^{\mathsf{H}}$. This setting yields the next representation.

Definition 8.11 (orthonormal tensor subspace representation). (a) If the bases B_j of U_j are orthonormal, the representation (8.6a,b) of $\mathbf{v}\in\mathbf{U}=\bigotimes_{j=1}^{d} U_j$ is called an *orthonormal tensor subspace representation* in $\mathbf{U}$.

(b) The detailed parameters of the representation are

$$\begin{array}{ll} r_j := \dim(U_j) & \text{for } 1\leq j\leq d, \\ B_j \in (V_j)^{r_j} & \text{orthonormal basis of } U_j, \\ J_j := \{1\leq i\leq r_j\} & \text{for } 1\leq j\leq d, \\ \mathbf{a}\in\bigotimes_{j=1}^{d}\mathbb{K}^{J_j} = \mathbb{K}^{\mathbf{J}} & \text{for } \mathbf{J} = J_1\times\ldots\times J_d \end{array} \tag{8.8a}$$

with

$$\rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^{d}\big) := \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}_{\mathbf{i}} \bigotimes_{j=1}^{d} b_{i_j}^{(j)} = \mathbf{B}\mathbf{a}. \tag{8.8b}$$

In the following, $V_j = \mathbb{K}^{I_j}$ is assumed. If, starting from general frames B_j, we want to obtain orthonormal bases, we have to find transformations such that B_j^{new} is an orthogonal matrix. For this purpose, two standard approaches can be applied. We recall Exercise 4.133: a QR decomposition of $\mathbf{B}$ or a Cholesky decomposition of $\mathbf{B}^{\mathsf{H}}\mathbf{B}$ are equivalent to the respective decomposition of B_j or $B_j^{\mathsf{H}}B_j$.

Lemma 8.12. *Let* $\mathbf{v} = \mathbf{B}\mathbf{a}$ *be given. (a) The QR decomposition* $\mathbf{B} = \mathbf{QR}$ *yields*

$$\mathbf{v} = \mathbf{Q}\,\mathbf{a}_{\mathrm{new}} \qquad \textit{with } \mathbf{a}_{\mathrm{new}} := \mathbf{R}\,\mathbf{a}.$$

By definition, $\mathbf{Q}$ *is an orthogonal matrix representing an orthonormal basis.*
(b) Let $\mathbf{B}$ *represent a basis. The Cholesky decomposition* $\mathbf{B}^{\mathsf{H}}\mathbf{B} = \mathbf{L}\mathbf{L}^{\mathsf{H}} \in \mathbb{K}^{\mathbf{J}\times\mathbf{J}}$ *defines the transformation*

$$\mathbf{v} = \Big(\mathbf{B}\,\mathbf{L}^{-\mathsf{H}}\Big)\,\mathbf{a}_{\mathrm{new}} \qquad \textit{with } \mathbf{a}_{\mathrm{new}} := \mathbf{L}^{\mathsf{H}}\mathbf{a}.$$

$\mathbf{B}\,\mathbf{L}^{-\mathsf{H}}$ *is an orthogonal matrix.*

Proof. Under the assumption of Part (b), the Gram matrix $\mathbf{B}^{\mathsf{H}}\mathbf{B}$ is positive definite and a decomposition $\mathbf{L}\mathbf{L}^{\mathsf{H}}$ exists. Orthogonality follows from $(\mathbf{B}\,\mathbf{L}^{-\mathsf{H}})^{\mathsf{H}}(\mathbf{B}\,\mathbf{L}^{-\mathsf{H}}) = \mathbf{L}^{-1}(\mathbf{B}^{\mathsf{H}}\mathbf{B})\mathbf{L}^{-\mathsf{H}} = \mathbf{L}^{-1}(\mathbf{L}\mathbf{L}^{\mathsf{H}})\mathbf{L}^{-\mathsf{H}} = \mathbf{id}$. □

Corollary 8.10 can be supplemented with orthogonality conditions.

Corollary 8.13. Let $\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$ be given. Assume $\mathbf{a} = \mathbf{S}\,\mathbf{a}_{\mathrm{new}}$ with an orthogonal $\mathbf{S}$, i.e., $\mathbf{S}^{\mathsf{H}}\mathbf{S} = \mathbf{I}$. Then also the new tensor subspace representation is orthonormal:

$$\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}_{\mathrm{new}}, (B_j^{\mathrm{new}})_{j=1}^d\big) \qquad \text{with } \mathbf{B}_{\mathrm{new}} := \mathbf{B}\,\mathbf{S}.$$

Proof. Use $\mathbf{B}_{\mathrm{new}}^{\mathsf{H}}\mathbf{B}_{\mathrm{new}} = \mathbf{S}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{B}\mathbf{S} = \mathbf{S}^{\mathsf{H}}\mathbf{S} = \mathbf{I}$. □

In the case of Corollary 8.13, $\mathrm{range}(B_j^{\mathrm{new}}) \subset \mathrm{range}(B_j)$ holds. If both orthonormal bases span the same subspace, transformations must be *unitary*. Given unitary transformations $Q^{(j)}$ of B_j into B_j^{new}:

$$b_i^{(j)} = \sum_{k=1}^{r_j} Q_{ki}^{(j)} b_{k,\mathrm{new}}^{(j)}, \quad b_{k,\mathrm{new}}^{(j)} = \sum_{i=1}^{r_j} \overline{Q_{ik}^{(j)}}\, b_i^{(j)} \quad \text{for } 1 \le i \le r_j,\ 1 \le j \le d, \tag{8.9a}$$

the Kronecker product $\mathbf{Q} := \bigotimes_{j=1}^d Q^{(j)}$ is also unitary and the coefficients transform according to $\mathbf{a}_{\mathrm{new}} = \mathbf{Q}\mathbf{a}$, i.e.,

$$\rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big) = \rho_{\mathrm{orth}}\big(\mathbf{a}_{\mathrm{new}}, (B_j^{\mathrm{new}})_{j=1}^d\big) \tag{8.9b}$$

(cf. (8.7c) with $\mathbf{T} = \mathbf{Q}$ and $(Q^{(j)})^{-1} = Q^{(j)\mathsf{H}}$).

Above, the new coefficient tensor $\mathbf{a}_{\mathrm{new}}$ is obtained from $\mathbf{a}$ by some transformation $\mathbf{T}\mathbf{a}$. Alternatively, the coefficient tensor can be obtained directly from $\mathbf{v}$ via projection.

Lemma 8.14. *(a) Let $\mathbf{v} \in \mathbf{U}$ and orthonormal bases B_j $(1 \le j \le d)$ be given: $\mathbf{v} = \mathbf{B}\mathbf{a}$ with $\mathbf{B} := \bigotimes_{j=1}^d B_j$. Then the coefficient tensor $\mathbf{a}$ of $\mathbf{v}$ has the entries*

$$\mathbf{a}_{\mathbf{i}} := \Big\langle \mathbf{v}, \bigotimes_{j=1}^d b_{i_j}^{(j)} \Big\rangle, \qquad \text{i.e., } \mathbf{a} = \mathbf{B}^*\mathbf{v}. \tag{8.10}$$

(b) For a general basis, the coefficient tensor from (8.6b) equals $\mathbf{a} = \mathbf{G}^{-1}\mathbf{b}$ with $\mathbf{b}_{\mathbf{k}} := \big\langle \mathbf{v}, \bigotimes_{j=1}^d b_{k_j}^{(j)} \big\rangle$, $\mathbf{G} = \bigotimes_{j=1}^d G^{(j)}$, where the Gram matrix $G^{(j)}$ (cf. (2.16)) has the entries

$$G_{ik}^{(j)} := \big\langle b_k^{(j)}, b_i^{(j)} \big\rangle \qquad \text{for } 1 \le i, k \le r_j,\ 1 \le j \le d. \tag{8.11}$$

Exercise 8.15. (a) Prove that the orthonormal tensor subspace representation $\mathbf{v} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}_{\mathbf{i}} \bigotimes_{j=1}^d b_{i_j}^{(j)}$ (cf. (8.8b)) implies that

$$\|\mathbf{v}\| = \|\mathbf{a}\|_2\,,$$

where $\|\cdot\| : \mathbf{V} \to \mathbb{R}$ is the norm associated with the induced scalar product of $\mathbf{V}$, while $\|\cdot\|_2$ is the Euclidean norm of $\mathbb{K}^{\mathbf{J}}$ (cf. Example 4.126).

(b) If a second tensor $\mathbf{w} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{c}_{\mathbf{i}} \bigotimes_{j=1}^d b_{i_j}^{(j)}$ uses the same bases, the scalar products $\big($that is $\langle\cdot,\cdot\rangle$ in $\mathbf{V}$, $\langle\cdot,\cdot\rangle_2$ in $\mathbb{K}^{\mathbf{J}}\big)$ coincide:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{a}, \mathbf{c} \rangle_2\,.$$

More details about the computation of orthonormal bases in the case of $V_j = \mathbb{K}^{I_j}$ will follow in §8.2.3.2.

8.2.3 *Tensors in* $\mathbb{K}^{\mathbf{I}}$

8.2.3.1 Representations and Transformations

Here, we consider the tensor space $\mathbf{V} = \mathbb{K}^{\mathbf{I}} = \bigotimes_{j=1}^{d} \mathbb{K}^{I_j}$ with $\mathbf{I} = I_1 \times \ldots \times I_d$ and $\mathbf{U} = \bigotimes_{j=1}^{d} U_j$ with subspaces $U_j \subset \mathbb{K}^{I_j}$. According to (8.5b), the quantity B_j representing a frame or basis is the matrix

$$B_j := \left[b_1^{(j)}, b_2^{(j)}, \ldots, b_{r_j}^{(j)}\right] \in \mathbb{K}^{I_j \times J_j} \qquad (1 \leq j \leq d)\,, \tag{8.12a}$$

where the index sets $J_j := \{1, \ldots, r_j\}$ form the product $\mathbf{J} = J_1 \times \ldots \times J_d$. Note that $r_j = \dim(U_j)$ in the case of a basis; otherwise, $r_j > \dim(U_j)$.

For the sake of simplicity, we shall speak about 'the frame B_j or basis B_j', although B_j is a matrix and only the collection of its columns form the frame or basis. Note that an "orthonormal basis B_j" and an "orthogonal matrix B_j" are equivalent expressions (cf. (2.3)).

The matrices B_j generate the Kronecker product

$$\mathbf{B} := \bigotimes_{j=1}^{d} B_j \in \mathbb{K}^{\mathbf{I} \times \mathbf{J}} \tag{8.12b}$$

(cf. (8.5c)).

We repeat the formats (8.6a-c) and (8.8a,b) for $V_j = \mathbb{K}^{I_j}$ with the modification that the frames are expressed by matrices B_j.

Lemma 8.16 (general tensor subspace representation). *(a) The coefficient tensor*

$$\mathbf{a} \in \bigotimes_{j=1}^{d} \mathbb{K}^{J_j} = \mathbb{K}^{\mathbf{J}} \qquad \textit{for } \mathbf{J} = J_1 \times \ldots \times J_d, \tag{8.13a}$$

and the tuple $(B_j)_{1 \leq j \leq d}$ *of frames represent the tensor* $\mathbf{v} = \mathbf{Ba}$ *with the entries*

$$\begin{aligned}\mathbf{v}[i_1 \cdots i_d] &= \sum_{k_1=1}^{r_1} \sum_{k_2=1}^{r_2} \cdots \sum_{k_d=1}^{r_d} B_1[i_1, k_1] B_2[i_2, k_2] \cdots B_d[i_d, k_d]\, \mathbf{a}[k_1 k_2 \cdots k_d] \\ &= \sum_{k_1=1}^{r_1} \sum_{k_2=1}^{r_2} \cdots \sum_{k_d=1}^{r_d} b_{k_1}^{(1)}[i_1] b_{k_2}^{(2)}[i_2] \cdots b_{k_d}^{(d)}[i_d]\, \mathbf{a}[k_1 k_2 \cdots k_d],\end{aligned} \tag{8.13b}$$

using the columns $b_k^{(j)} = B_j[\bullet, k]$ *of* B_j. *Equation (8.13b) is equivalent to* $\mathbf{v} = \mathbf{Ba}$. *The representation by*

$$\rho_{\text{frame}}(\mathbf{a}, (B_j)_{j=1}^{d}) = \mathbf{Ba} \tag{8.13c}$$

is identical to (8.6c), but now the data B_j *are stored as (fully populated) matrices.*
(b) The storage required by $\mathbf{B}$ *and* $\mathbf{a}$ *is*

$$N_{\text{mem}}\left(\mathbf{B}\right) = N_{\text{mem}}\left((B_j)_{j=1}^{d}\right) = \sum_{j=1}^{d} r_j \cdot \# I_j, \quad N_{\text{mem}}\left(\mathbf{a}\right) = \prod_{j=1}^{d} r_j\,. \tag{8.13d}$$

Orthonormal bases are characterised by orthogonal matrices B_j (cf. §8.2.2): $B_j^{\mathsf{H}} B_j = I \in \mathbb{K}^{r_j \times r_j}$. This property holds for all $1 \le j \le d$, if and only if $\mathbf{B}^{\mathsf{H}}\mathbf{B} = \mathbf{I}$. Because of numerical stability, orthonormal bases are the standard choice for the tensor subspace representation.

Lemma 8.17 (orthonormal tensor subspace representation). *Assume that B_j are orthogonal matrices. B_j and the coefficient tensor (8.13a) are the data of the orthonormal tensor subspace representation:*

$$\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big) = \left(\bigotimes_{j=1}^d B_j\right)\mathbf{a} \quad \text{with } B_j^{\mathsf{H}} B_j = I. \tag{8.14a}$$

The required memory size is the same as in (8.13d). The coefficient tensor $\mathbf{a}$ can be obtained from v by

$$\mathbf{a} = \mathbf{B}^{\mathsf{H}}\mathbf{v}. \tag{8.14b}$$

Proof. Use (8.10). A direct proof is $\mathbf{a} \underset{\mathbf{B}^{\mathsf{H}}\mathbf{B}=\mathbf{I}}{=} \mathbf{B}^{\mathsf{H}}\mathbf{B}\mathbf{a} = \mathbf{B}^{\mathsf{H}}\mathbf{v}$. □

8.2.3.2 Orthonormalisation and Computational Cost

Here, we assume[7] that a tensor $\mathbf{v} = \rho_{\mathrm{frame}}(\hat{\mathbf{a}}, (\hat{B}_j)_{j=1}^d)$ is given with a proper frame or non-orthonormal basis $\hat{B}_j$. Lemma 8.12 proposes two methods for generating orthonormal bases B_j. Another possibility is the computation of the HOSVD bases (cf. §8.3 and §8.3.3). These computations are more expensive, on the other hand they allow to determine orthonormal bases of the minimal subspaces $U_j^{\min}(\mathbf{v})$, whereas the following methods yield bases of possibly larger subspaces $U_j := \mathrm{range}(\hat{B}_j)$. We start with the QR decomposition. Given frames or bases $(\hat{B}_j)_{j=1}^d$, procedure $\mathbf{RQR}(n_j, \hat{r}_j, r_j, B_j, Q_j, R_j)$ from (2.29) yields the decomposition

$$\hat{B}_j = Q_j R_j \qquad (\hat{B}_j \in \mathbb{K}^{n_j \times \hat{r}_j}, Q_j \in \mathbb{K}^{n_j \times r_j}, R_j \in \mathbb{K}^{r_j \times \hat{r}_j})$$

with orthogonal matrices Q_j, where r_j is the rank of $\hat{B}_j$, Q_j, and R_j. Defining the Kronecker matrices

$$\hat{\mathbf{B}} := \bigotimes_{j=1}^d \hat{B}_j, \qquad \mathbf{Q} := \bigotimes_{j=1}^d Q_j, \qquad \text{and} \quad \mathbf{R} := \bigotimes_{j=1}^d R_j\,,$$

we get $\mathbf{v} = \hat{\mathbf{B}}\hat{\mathbf{a}} = \mathbf{Q}\mathbf{R}\hat{\mathbf{a}}$ (cf. (8.6c)). Besides the exact operation count, we give a bound in terms of

$$\bar{\hat{r}} := \max_j \hat{r}_j \qquad \text{and} \qquad n := \max_j n_j. \tag{8.15}$$

[7] Also tensors represented in the r-term format are related to subspaces U_j, for which orthonormal basis can be determined (cf. Remark 6.1). We can convert such tensors into the format ρ_{TS} according to §8.5.2.2 without arithmetical cost and apply the present algorithms.

Remark 8.18. Use the notations from above. The computational cost of all QR decompositions $\hat{B}_j = Q_j R_j$ and the cost of the product $\mathbf{a} := \mathbf{R}\hat{\mathbf{a}}$ add to

$$\sum_{j=1}^{d} \left[N_{\mathrm{QR}}(n_j, \hat{r}_j) + \prod_{k=1}^{j} r_k \cdot \prod_{k=j}^{d} \hat{r}_k \right] \leq 2dn\bar{\hat{r}}^2 + d\bar{\hat{r}}^{d+1}. \tag{8.16}$$

Proof. The second term describes the cost of $\mathbf{R}\hat{\mathbf{a}}$ considered in (13.27a). Because of the triangular structure, a factor two can be saved. □

The second approach from Lemma 8.12 is based on the Cholesky decomposition, provided that $(\hat{B}_j)_{j=1}^d$ represents bases. Because of the latter assumption, $\hat{r}_j = r_j$ holds. Note that, in particular for the case $r \ll n$, the resulting cost in (8.17) is almost the same as in (8.16).

Remark 8.19. With the notations from above, the computational cost of the Cholesky approach in Lemma 8.12b is

$$\sum_{j=1}^{d} \left[2n_j r_j^2 + \frac{1}{3} r_j^3 + r_j \prod_{k=1}^{d} r_k \right] \leq d \left(2n + \frac{\bar{r}}{3} \right) \bar{r}^2 + d\bar{r}^{d+1}. \tag{8.17}$$

Proof. The product $\hat{B}_j^{\mathsf{H}} \hat{B}_j$ takes $\frac{1}{2}(2n_j - 1) r_j (r_j + 1) \approx n_j r_j^2$ operations. The Cholesky decomposition into $L_j L_j^{\mathsf{H}}$ requires $\frac{1}{3} r_j^3$ operations (cf. Remark 2.18). Further $n_j r_j^2$ operations are needed to build the new basis $B_j := \hat{B}_j L_j^{-\mathsf{H}}$. The transformation $\mathbf{a} = \mathbf{L}^{\mathsf{H}} \hat{\mathbf{a}}$ costs $\left(\sum_{j=1}^{d} r_j \right) \cdot \prod_{j=1}^{d} r_j$ operations (cf. Remark 2.18). □

For larger d, the major part of the computational cost in Remark 8.18 is $d\bar{r}^{d+1}$, which is caused by the fact that $\hat{\mathbf{a}}$ is organised as full tensor in $\mathbb{K}^{\mathbf{J}}$. Instead, the hybrid format discussed in §8.2.4 uses the r-term format for $\hat{\mathbf{a}}$. The resulting cost of $\mathbf{R}\hat{\mathbf{a}}$ ($\mathbf{R}$, $\hat{\mathbf{a}}$ as in Remark 8.18) described in (13.28b) is given in the following corollary.

Corollary 8.20. Let the coefficient tensor $\hat{\mathbf{a}} \in \mathcal{R}_r$ be given in r-term format. The following transformations yield the new coefficient tensor $\mathbf{a}$ in the same format.

(a) Using the QR decompositions from Remark 8.18, the cost of $\mathbf{a} := \mathbf{R}\hat{\mathbf{a}}$ is

$$r \sum_{j=1}^{d} r_j \left(2\hat{r}_j - r_j \right) \lesssim dr\bar{\hat{r}}^2,$$

while the QR cost $\sum_{j=1}^{d} N_{\mathrm{QR}}(n_j, \hat{r}_j)$ does not change. The total cost is bounded by $d(2n + r)\bar{\hat{r}}^2$.

(b) In the Cholesky approach from Remark 8.19 the coefficient tensor $\mathbf{a} := \mathbf{L}^{\mathsf{H}} \hat{\mathbf{a}}$ can be obtained by $r \sum_{j=1}^{d} r_j^2$ operations, yielding the total bound $d(2n + \frac{\bar{r}}{3} + r)\bar{r}^2$.

8.2.3.3 Generalisation

The [orthonormal] tensor subspace representation (8.13) [or (8.14a)] can be used to represent several tensors *simultaneously* in the same tensor subspace:

$$\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(m)} \in \mathbf{U} = \bigotimes\nolimits_{j=1}^{d} U_j.$$

In this case, the data $(B_j)_{1\leq j\leq d}$ need to be stored only once. Each tensor $\mathbf{v}^{(\mu)}$ requires a coefficient tensor $\mathbf{a}^{(\mu)}$ $(1 \leq \mu \leq m)$. The required data size is $\bar{r}dn+m\bar{r}^d$, where $n := \max_j \#I_j$ and $\bar{r} := \max_j r_j$.

8.2.4 Hybrid Format

Let $\mathbf{v} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a_i} \bigotimes_{j=1}^{d} b_{i_j}^{(j)}$ be the standard tensor subspace representation ρ_{TS} or ρ_{orth}. An essential drawback of this format is the fact that the coefficient tensor $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$ is still represented in full format. Although $\mathbf{J} = J_1 \times \ldots \times J_d$ might be of much smaller size than $\mathbf{I} = I_1 \times \ldots \times I_d$, the exponential increase of $\#\mathbf{J}$ with respect to d proves disadvantageous. An obvious idea is to represent $\mathbf{a}$ itself by one of the tensor formats described so far. Using again a tensor subspace representation for $\mathbf{a}$ does not yield a new format as seen in Remark 8.21 below.

An interesting approach is the choice of an r-term representation of the coefficient tensor $\mathbf{a}$. Often, such an approach goes together with an approximation, but here we consider an exact representation of $\mathbf{a}$ by

$$\mathbf{a} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} a_\nu^{(j)} \in \mathbb{K}^{\mathbf{J}} \qquad \text{with } a_\nu^{(j)} \in \mathbb{K}^{J_j}. \tag{8.18}$$

The tensor subspace format $\mathbf{v} = \rho_{\mathrm{TS}}(\mathbf{a}, (B_j)_{j=1}^d)$ combined with the r-term representation $\mathbf{a} = \rho_{\text{r-term}}\big(r, (a_\nu^{(j)})_{1\leq j\leq d, 1\leq\nu\leq r}\big)$ from (8.18) yields the *hybrid format*, which may be interpreted in two different ways.

The *first interpretation* views $\mathbf{v}$ as a particular tensor from $\mathcal{T}_{\mathbf{r}}$ (with $r_j = \#J_j$) described by the iterated representation

$$\begin{aligned}\rho_{\mathrm{hybr}}\Big(r, (a_\nu^{(j)})_{\substack{1\leq j\leq d\\ 1\leq\nu\leq r}}, (B_j)_{j=1}^d\Big) &:= \rho_{\mathrm{TS}}\Big(\rho_{\text{r-term}}\Big(r, (a_\nu^{(j)})_{\substack{1\leq j\leq d\\ 1\leq\nu\leq r}}\Big), (B_j)_{j=1}^d\Big) \\ &= \sum_{\mathbf{i}\in\mathbf{J}} \Big(\sum_{\nu=1}^{r} \prod_{j=1}^{d} a_\nu^{(j)}[i_j]\Big) \bigotimes_{j=1}^{d} b_{i_j}^{(j)} \end{aligned} \tag{8.19}$$

with ρ_{TS} from (8.6c) and $\mathbf{a} = \rho_{\text{r-term}}(\ldots)$ from (7.7a). Similarly, we may define

$$\rho_{\mathrm{orth}}^{\mathrm{hybr}}\Big(r, (a_\nu^{(j)}), (B_j)_{j=1}^d\Big) := \rho_{\mathrm{orth}}\Big(\rho_{\text{r-term}}\big(r, (a_\nu^{(j)})\big), (B_j)_{j=1}^d\Big), \tag{8.20}$$

provided that B_j describes orthonormal bases.

The *second interpretation* views $\mathbf{v}$ as a particular tensor from $\mathcal{R}_r$:

$$\mathbf{v} = \sum_{\nu=1}^{r} \sum_{\mathbf{i}\in\mathbf{J}} \bigotimes_{j=1}^{d} a_\nu^{(j)}[i_j]\, b_{i_j}^{(j)} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} \Bigl(\sum_{i\in J_j} a_\nu^{(j)}[i]\, b_i^{(j)} \Bigr).$$

The right-hand side may be seen as $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$, where, according to modification (7.13), $v_\nu^{(j)}$ is described by means of the basis $\{b_i^{(j)} : i \in J_j\}$, which yields the matrix B_j. The format is abbreviated by

$$\rho_{\text{r-term}}^{\text{hybr}}\Bigl(r, \mathbf{J}, (a_\nu^{(j)})_{\substack{1\le j\le d,\\ 1\le\nu\le r}}, (B_j)_{j=1}^{d}\Bigr) = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} \Biggl(\sum_{i\in J_j} a_\nu^{(j)}[i]\, b_i^{(j)} \Biggr). \tag{8.21}$$

Note that the formats (8.19) and (8.21) are equivalent in the sense that they use the same data representing the same tensor.

Another characterisation of a tensor $\mathbf{v}$ in the hybrid format is

$$\mathbf{v} \in \mathcal{R}_r \cap \mathcal{T}_{\mathbf{r}}$$

with r, $\mathbf{r} = (r_1, \ldots, r_d)$, $r_j = \# J_j$ from (8.19) and (8.21).

The hybrid format is intensively used in Espig [52, Satz 2.2.4] (cf. §9.5.1) and in Khoromskij-Khoromskaja [119].

Finally, we discuss the situation of a coefficient tensor $\mathbf{a}$ given again in tensor subspace format.

Remark 8.21. (a) Consider the following nested tensor subspace formats:

$$\mathbf{v} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a_i} \bigotimes_{j=1}^{d} b_{i_j}^{(j)}, \qquad \mathbf{a} = \sum_{\mathbf{k}\in\mathbf{K}} \mathbf{c_k} \bigotimes_{j=1}^{d} \beta_{k_j}^{(j)}, \tag{8.22a}$$

where $\mathbf{v} \in \mathbf{V} = \mathbb{K}^{\mathbf{I}}$ with $\mathbf{I} = I_1 \times \ldots \times I_d$, $b_{i_j}^{(j)} \in V_j = \mathbb{K}^{I_j}$, $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$, $\beta_{k_j}^{(j)} \in \mathbb{K}^{J_j}$, $\mathbf{c} \in \mathbb{K}^{\mathbf{K}}$ with $\mathbf{K} = K_1 \times \ldots \times K_d$. Then, $\mathbf{v}$ has the standard tensor subspace representation

$$\mathbf{v} = \sum_{\mathbf{k}\in\mathbf{K}} \mathbf{c_k} \bigotimes_{j=1}^{d} \hat{b}_k^{(j)} \quad \begin{cases} \text{with } \hat{b}_k^{(j)} \in \mathbb{K}^{I_j} \text{ defined by} \\ \hat{b}_k^{(j)} := \sum_{i\in J_j} \beta_k^{(j)}[i]\, b_i^{(j)} \quad (k \in K_j). \end{cases} \tag{8.22b}$$

(b) Using $B_j := [b_1^{(j)} \cdots b_{r_j}^{(j)}]$ $(r_j := \# J_j)$, $\mathbf{B} := \bigotimes_{j=1}^{d} B$, $\boldsymbol{\beta}_j := [\beta_1^{(j)} \cdots \beta_{s_j}^{(j)}]$ $(s_j := \# K_j)$, $\boldsymbol{\beta} := \bigotimes_{j=1}^{d} \boldsymbol{\beta}_j$, and $\hat{B}_j := [\hat{b}_1^{(j)} \cdots \hat{b}_{s_j}^{(j)}]$, we rewrite (8.22a,b) as

$$\mathbf{v} = \mathbf{B}\mathbf{a}, \quad \mathbf{a} = \boldsymbol{\beta}\mathbf{c}, \quad \mathbf{v} = \hat{\mathbf{B}}\mathbf{c} \quad \text{with } \hat{\mathbf{B}} := \mathbf{B}\boldsymbol{\beta}. \tag{8.22c}$$

The equations in (8.22c) can be interpreted as transformation: set $\mathbf{a}_{\text{new}} = \mathbf{c}$, $\mathbf{S} = \boldsymbol{\beta}$, and $\mathbf{B}_{\text{new}} = \hat{\mathbf{B}}$ in Corollary 8.10. The computation of $\hat{\mathbf{B}}$, i.e., of all products $\hat{B}_j = B_j \beta_j$ requires $2\sum_{j=1}^{d} n_j r_j s_j$ operations, where $n_j := \# I_j$.

(c) Orthonormal tensor subspace representations for $\mathbf{v}$ and $\mathbf{a}$ in (8.22a) yield again an orthonormal tensor subspace representation in (8.22b).

8.3 Higher-Order Singular Value Decomposition (HOSVD)

In the following, matrices denoted by U, V or even U_j, V_j appear in the singular value decomposition. These matrices are to be distinguished from the (sub)spaces $U_j, U_j^{\min}$ and V_j with similar or even equal names.

As stated in Remark 3.46a, there is no true generalisation of the singular value decomposition (SVD) for $d \geq 3$. However, it is possible to extend parts of the SVD structure to higher dimensions as sketched below. Considering a (reduced) singular value decomposition of a *matrix*, we observe the following properties:

(a_1) $M = U\Sigma V^{\mathsf{T}} = \sum_{i=1}^{r} \sigma_i u_i v_i^{\mathsf{T}}$ can be exploited, e.g., for truncations.
(a_2) In fact, $M_s := \sum_{i=1}^{s} \sigma_i u_i v_i^{\mathsf{T}}$ is the best approximation of rank s.
(b_1) We may use u_i and v_i as new basis vectors.
(b_2) The basis transformation from (b_1) maps M into diagonal form.

HOSVD will also be helpful for truncation (as in (a_1)), and, in fact, this property will be a very important feature in practice. However, the result of truncation is not necessarily optimal, i.e., (a_2) does not extend to $d \geq 3$. As in (b_1), HOSVD will provide new bases, but the tensor expressed with respect to these bases is by no means diagonal, not even sparse, i.e., (b_2) has no tensor counterpart.

8.3.1 Definitions

We start with the tensor space $\mathbf{V} = \bigotimes_{j=1}^{d} \mathbb{K}^{I_j}$. Given $\mathbf{v} \in \mathbf{V}$, we consider the matricisation $M := \mathcal{M}_j(\mathbf{v})$ which is a matrix of size $I_j \times I_{[j]}$ with $I_{[j]} = \times_{k \neq j} I_k$. Its reduced singular value decomposition is

$$M = U\Sigma V^{\mathsf{T}} = \sum\nolimits_{i=1}^{r_j} \sigma_i u_i v_i^{\mathsf{T}} \in \mathbb{K}^{I_j \times I_{[j]}}, \tag{8.23}$$

where u_i and v_i are the columns of the respective orthogonal matrices $U \in \mathbb{K}^{I_j \times r_j}$ and $V \in \mathbb{K}^{I_{[j]} \times r_j}$, $\sigma_1 \geq \sigma_2 \geq \ldots > 0$ are the singular values, and $r_j = \operatorname{rank}(M) = \operatorname{rank}_j(\mathbf{v})$ (cf. (5.6b)). While U may be of reasonable size, $V \in \mathbb{K}^{I_{[j]} \times r_j}$ is expected to have a huge number of rows, which one does not like to compute. Moreover, it turns out that the matrix V is not needed.

We recall the 'left-sided singular value decomposition': as mentioned in Remark 2.24b, we may ask only for U and Σ in the singular value decomposition $M = U\Sigma V^{\mathsf{T}}$, and the computation of U and Σ may be based on $MM^{\mathsf{H}} = U\Sigma^2 U^{\mathsf{H}}$. The diagonal matrix Σ controls the truncation procedure (see item (a_1) from above), while U defines an orthonormal basis (item (b_1)). We remark that $\operatorname{range}(U) = \operatorname{range}(M) = U_j^{\min}(\mathbf{v})$ (cf. Remark 8.4).

Different from the case $d = 2$, we have d different matricisations $\mathcal{M}_j(\mathbf{v})$ leading to a tuple of d different decompositions (8.23), called '*higher-order singular value decomposition* (HOSVD)' by De Lathauwer et al. [41]. To distinguish the matricisations, we ornament the quantities of (8.23) with the index j referring to $\mathcal{M}_j(\mathbf{v}) \in V_j \otimes \mathbf{V}_{[j]}$.

In the next definition, $\mathbf{V}$ is a general Hilbert tensor space. This space as well as all $\mathbf{V}_{[j]} = \bigotimes_{k \neq j} V_k$ are equipped with the corresponding induced scalar product. All scalar products in $\mathbf{V}$, V_j, and $\mathbf{V}_{[j]}$ are denoted by $\langle \cdot, \cdot \rangle$.

Definition 8.22 (HOSVD basis). Let $\mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v}) \subset {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$. An orthonormal basis $B_j = (b_i^{(j)}, \ldots, b_{r_j}^{(j)}\}$ of $U_j^{\min}(\mathbf{v})$ is called j-th *HOSVD basis* for $\mathbf{v}$, if the following (singular value) decomposition is valid:

$$\begin{aligned} &\mathcal{M}_j(\mathbf{v}) = \sum_{i=1}^{r_j} \sigma_i^{(j)} b_i^{(j)} \otimes m_i^{(j)} && \text{with} \\ &\sigma_1^{(j)} \geq \sigma_2^{(j)} \geq \ldots > 0 && \text{and} \\ &\text{orthonormal } \{m_i^{(j)} : 1 \leq i \leq r_j\} \subset \mathbf{V}_{[j]} := {}_{\|\cdot\|}\bigotimes\nolimits_{k \neq j} V_k \ . \end{aligned} \tag{8.24}$$

$\sigma_i^{(j)}$ are called the singular values of the j-th matricisation. For infinite dimensional Hilbert spaces V_j and topological tensors, $r_j = \infty$ may occur.

Similarly, for a subset $\emptyset \neq \alpha \subsetneqq \{1, \ldots, d\}$, an orthonormal basis $(b_i^{(\alpha)})_{i=1}^{r_\alpha}$ of $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ is called an α-*HOSVD basis* for $\mathbf{v}$, if

$$\begin{aligned} &\mathcal{M}_\alpha(\mathbf{v}) = \sum_{i=1}^{r_\alpha} \sigma_i^{(\alpha)} b_i^{(\alpha)} \otimes m_i^{(\alpha)} && \text{with} \\ &\sigma_1^{(\alpha)} \geq \sigma_2^{(\alpha)} \geq \ldots > 0 && \text{and} \\ &\text{orthonormal } \{m_i^{(\alpha)} : 1 \leq i \leq r_j\} \subset \mathbf{V}_{\alpha^c}. \end{aligned} \tag{8.25}$$

Definition 8.23 (HOSVD representation). A tensor subspace representation $\mathbf{v} = \rho_{\mathrm{orth}}(\mathbf{a}, (B_j)_{1 \leq j \leq d})$ is a *higher-order singular value decomposition (HOSVD)* (or *'HOSVD tensor subspace representation'* or shortly *'HOSVD representation'*) of $\mathbf{v}$, if all bases B_j $(1 \leq j \leq d)$ are HOSVD bases for $\mathbf{v}$.[8] For B_j satisfying these conditions, we write

$$\mathbf{v} = \rho_{\mathrm{HOSVD}}\big(\mathbf{a}, (B_j)_{1 \leq j \leq d}\big). \tag{8.26}$$

The storage requirements of HOSVD are the same as for the general case which is described in Lemma 8.16b.

The next statement follows from Lemma 5.6.

Lemma 8.24. *(a) A tensor* $\mathbf{v} = \mathbf{B}\mathbf{a}$ *with* $\mathbf{B} = \bigotimes_{j=1}^{d} B_j$ *yields* $\mathcal{M}_j(\mathbf{v}) = B_j \mathcal{M}_j(\mathbf{a}) B_{[j]}^{\mathsf{T}}$ *with* $B_{[j]} = \bigotimes_{k \neq j} B_k$. *If, at least for* $k \neq j$, *the bases* B_k *are orthonormal, the matricisations of* $\mathbf{v}$ *and* $\mathbf{a}$ *are related by*

$$\mathcal{M}_j(\mathbf{v})\, \mathcal{M}_j(\mathbf{v})^{\mathsf{H}} = B_j \left[\mathcal{M}_j(\mathbf{a})\, \mathcal{M}_j(\mathbf{a})^{\mathsf{H}}\right] B_j^{\mathsf{H}}. \tag{8.27a}$$

(b) If also B_j *contains an orthonormal basis, a diagonalisation* $\mathcal{M}_j(\mathbf{a})\, \mathcal{M}_j(\mathbf{a})^{\mathsf{H}} = \hat{U}_j \Sigma_j^2 \hat{U}_j^{\mathsf{H}}$ *yields the left-sided singular value decomposition*

$$\mathcal{M}_j(\mathbf{v})\, \mathcal{M}_j(\mathbf{v})^{\mathsf{H}} = U_j \Sigma_j^2 U_j^{\mathsf{H}} \qquad \textit{with } U_j := B_j \hat{U}_j. \tag{8.27b}$$

[8] Because of the orthogonality property (8.24) for all $1 \leq j \leq d$, such a tensor representation is called *all-orthogonal* by De Lathauwer et al. [40], [106].

As a consequence, HOSVD representations of $\mathbf{v}$ and $\mathbf{a}$ are closely connected.

Corollary 8.25. Let $\mathbf{v} \in \bigotimes_{j=1}^{d} \mathbb{K}^{I_j}$ be given by an orthonormal tensor subspace representation (8.14a): $\mathbf{v} = \rho_{\mathrm{orth}}(\mathbf{a}, (B_j)_{1\le j\le d})$ with $B_j^{\mathsf{H}} B_j = I$. Then, $(B_j)_{1\le j\le d}$ describes the j-th HOSVD basis of $\mathbf{v}$ if and only if

$$\begin{aligned} &\mathcal{M}_j(\mathbf{a})\,\mathcal{M}_j(\mathbf{a})^{\mathsf{H}} = \Sigma_j^2 && \text{with} \\ &\Sigma_j = \operatorname{diag}\{\sigma_1^{(j)}, \sigma_2^{(j)}, \ldots\} && \text{and} \quad \sigma_1^{(j)} \ge \sigma_2^{(j)} \ge \ldots > 0. \end{aligned}$$

8.3.2 Examples

We give two examples of the HOSVD for the simple case $d = 3$ and the symmetric situation $r_1 = r_2 = r_3 = 2$, $V_1 = V_2 = V_3 =: V$.

Example 8.26. Let $x, y \in V$ be two orthonormal vectors and set[9]

$$\mathbf{v} := x \otimes x \otimes x + \sigma y \otimes y \otimes y \in \mathbf{V} := \otimes^3 V. \tag{8.28}$$

(8.28) is already the HOSVD representation of $\mathbf{v}$. For all $1 \le j \le 3$, (8.24) holds with

$$r_j = 2,\ \sigma_1^{(j)} = 1,\ \sigma_2^{(j)} = \sigma,\ b_1^{(j)} = x,\ b_2^{(j)} = y,\ m_1^{(j)} = x \otimes x,\ m_2^{(j)} = y \otimes y.$$

While $\mathbf{v}$ from (8.28) has tensor rank 2, the next tensor has rank 3.

Example 8.27. Let $x, y \in V$ be two orthonormal vectors and set

$$\mathbf{v} = \alpha x \otimes x \otimes x + \beta x \otimes x \otimes y + \beta x \otimes y \otimes x + \beta y \otimes x \otimes x \in \mathbf{V} := \otimes^3 V. \tag{8.29a}$$

For the choice

$$\alpha := \sqrt{1 - \tfrac{3}{2}\sqrt{2}\,\sigma + \sigma^2} \quad \text{and} \quad \beta := \sqrt{\sigma/\sqrt{2}}, \tag{8.29b}$$

the singular values are again $\sigma_1^{(j)} = 1,\ \sigma_2^{(j)} = \sigma$. The HOSVD basis is given by

$$b_1^{(j)} = \frac{\sqrt{1 - \frac{\sigma}{\sqrt{2}}}\,x + \sqrt{\sigma\left(\frac{1}{\sqrt{2}} - \sigma\right)}\,y}{\sqrt{(1+\sigma)(1-\sigma)}},\ b_2^{(j)} = \frac{\sqrt{\sigma\left(\frac{1}{\sqrt{2}} - \sigma\right)}\,x - \sqrt{1 - \frac{\sigma}{\sqrt{2}}}\,y}{\sqrt{(1+\sigma)(1-\sigma)}}. \tag{8.29c}$$

In principle, the HOSVD can also be performed in Hilbert tensor spaces $\mathbf{V} := {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ with induced scalar product. In the general case, the HOSVD bases are infinite (cf. Theorem 4.114). If $\mathbf{v} := {}_a\bigotimes_{j=1}^{d} V_j$ is an algebraic tensor, finite bases are ensured as in the next example referring to the polynomial from Example 8.5b. Note that here V_j is the function space $L^2([0,1])$.

[9] The coefficient tensor $\mathbf{a}$ has the entries $\mathbf{a}[1,1,1] = 1$, $\mathbf{a}[2,2,2] = \sigma$, and zero, otherwise.

Example 8.28. The HOSVD bases and the corresponding singular values[10] of the polynomial $f(x,y,z) = xz + x^2y \in \mathbf{V} := {}_a\bigotimes_{j=1}^d V_j\,,\ V_j = L^2([0,1])$, are

$$\begin{aligned}
b_1^{(1)} &= 0.99953x + 0.96327x^2, & \sigma_1^{(1)} &= \sqrt{\tfrac{109}{720} + \tfrac{1}{45}\sqrt{46}} &&\approx 0.54964,\\
b_2^{(1)} &= 6.8557x - 8.8922x^2, & \sigma_2^{(1)} &= \sqrt{\tfrac{109}{720} - \tfrac{1}{45}\sqrt{46}} &&\approx 0.025893,\\
b_1^{(2)} &= 0.58909 + 0.77158y, & \sigma_1^{(2)} &= \sqrt{\tfrac{109}{720} + \tfrac{1}{360}\sqrt{2899}} &&\approx 0.54859,\\
b_2^{(2)} &= 1.9113 - 3.3771y, & \sigma_2^{(2)} &= \sqrt{\tfrac{109}{720} - \tfrac{1}{360}\sqrt{2899}} &&\approx 0.042741,\\
b_1^{(3)} &= 0.44547 + 1.0203z, & \sigma_1^{(3)} &= \sigma_1^{(2)},\\
b_2^{(3)} &= 1.9498 - 3.3104z, & \sigma_2^{(3)} &= \sigma_2^{(2)}.
\end{aligned}$$

Proof. The matricisations $\mathcal{M}_j(f)$ define integral operators $\mathcal{K}_j := \mathcal{M}_j(f)\mathcal{M}_j^*(f) \in \mathcal{L}(L^2([0,1]), L^2([0,1]))$ of the form $(\mathcal{K}_j(g))(\xi) = \int_0^1 k_j(\xi,\xi')g(\xi')\mathrm{d}\xi'$ (cf. Example 5.16). The involved kernel functions are

$$\begin{aligned}
k_1(x,x') &= \int_0^1\int_0^1 f(x,y,z)f(x',y,z)\mathrm{d}y\mathrm{d}z = \frac{1}{3}xx' + \frac{1}{4}x^2x' + \frac{1}{4}xx'^2 + \frac{1}{3}x^2x'^2,\\
k_2(y,y') &= \frac{1}{9} + \frac{1}{8}y + \frac{1}{8}y' + \frac{1}{5}yy', \quad k_3(z,z') = \frac{1}{15} + \frac{1}{8}z + \frac{1}{8}z' + \frac{1}{3}zz'.
\end{aligned}$$

The eigenfunctions of $\mathcal{K}_1$ are $x - \frac{1}{6}(\sqrt{46}+1)x^2$ and $x + \frac{1}{6}(\sqrt{46}-1)x^2$ with the eigenvalues $\lambda_{1,2} = \frac{109}{720} \pm \frac{1}{45}\sqrt{46}$. Normalising the eigenfunctions and extracting the square root of $\lambda_{1,2}$, we obtain the orthonormal basis functions $b_i^{(1)}$ and $\sigma_i^{(1)}$ $(i = 1,2)$ from above.

Similarly, the eigenfunctions $1 + \frac{-\sqrt{2899}\pm 8}{35}y$ of $\mathcal{K}_2$ and $1 + \frac{8\pm\sqrt{2899}}{27}z$ of $\mathcal{K}_3$ yield the indicated results. □

8.3.3 Computation and Computational Cost

Let $V_j = \mathbb{K}^{I_j}$ with $n_j := \#I$ and $\mathbf{V} = \bigotimes_{j=1}^d V_j$. For subsets $\emptyset \neq \alpha \subsetneqq \{1,\ldots,d\}$, we use the notations $\alpha^c := \{1,\ldots,d\}\backslash\alpha$ and $\mathbf{V}_\alpha = {}_a\bigotimes_{k\in\alpha} V_k$. The usual choice is $\alpha = \{j\}$. We introduce the mapping

$$(B_\alpha, \varSigma_\alpha) := \mathrm{HOSVD}_\alpha(\mathbf{v}) \tag{8.30a}$$

characterised by the left-side singular value decomposition

$$\mathcal{M}_\alpha(\mathbf{v})\mathcal{M}_\alpha(\mathbf{v})^{\mathsf{H}} = B_\alpha\varSigma_\alpha^2 B_\alpha^{\mathsf{H}},\ B_\alpha \in \mathbb{K}^{n_\alpha\times r_\alpha},\ 0 \le \varSigma_\alpha \in \mathbb{K}^{r_\alpha\times r_\alpha},\ B_\alpha^{\mathsf{H}}B_\alpha = I,$$

[10] Since $\sum_{i=1}^2 \left(\sigma_1^{(j)}\right)^2 = \frac{109}{720}$ for all $1 \le j \le 3$, the values pass the test by Remark 5.12b.

where $r_\alpha := \mathrm{rank}_\alpha(\mathbf{v}) = \dim(\mathbf{U}_\alpha^{\min}(\mathbf{v}))$. Since the singular value decomposition is not always unique (cf. Corollary 2.21b), the map HOSVD_α is not well-defined in all cases. If multiple solutions exist, one may pick a suitable one.

Performing $\mathrm{HOSVD}_j(\mathbf{v})$ for all $1 \le j \le d$, we obtain the complete higher order singular value decomposition

$$(B_1, \Sigma_1, B_2, \Sigma_2, \ldots, B_d, \Sigma_d) := \mathrm{HOSVD}(\mathbf{v}). \tag{8.30b}$$

The computational realisation of HOSVD_j depends on the various formats. Here, we discuss the following cases:

(A) Tensor $\mathbf{v}$ given in full format.

(B) $\mathbf{v}$ given in r-term format $\rho_{\text{r-term}}\big(r, (v_\nu^{(j)})_{1\le j\le d,\ 1\le\nu\le r}\big)$.

(C) $\mathbf{v}$ given in the orthonormal tensor subspace format $\rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$, where the coefficient tensor $\mathbf{a}$ may have various formats.

(D) $\mathbf{v}$ given in the general tensor subspace format $\rho_{\mathrm{frame}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$, which means that B_j is not necessarily orthogonal.

As review we list the cost (up to lower order terms) for the various cases:

format of $\mathbf{v}$	computational cost	details in
full	n^{d+1}	Remark 8.29
r-term	$d\big[2nr\min(n,r)+nr^2+2n\bar{r}^2+2r^2\bar{r}+3r\bar{r}^2+\frac{8}{3}\bar{r}^3\big]$	Remark 8.30
ρ_{orth}	$3d\overline{\hat{r}}^{d+1}+2d\overline{\hat{r}}^2(n+\frac{4}{3}\overline{\hat{r}})$	(8.35c)
$\rho_{\mathrm{orth}}^{\mathrm{hybr}}$	$2dnr\overline{\hat{r}}+(d+2)\,r^2\overline{\hat{r}}+2dr\overline{\hat{r}}\min(\overline{\hat{r}},r)+3r\overline{\hat{r}}^2+\frac{14}{3}d\overline{\hat{r}}^3$	(8.36)

8.3.3.1 Case A: Full Format

Set $n_j := \#I_j$, $I_{[j]} = \times_{k\ne j} I_k$, and $n := \max_j n_j$. The data $\mathrm{HOSVD}_j(\mathbf{v}) = (B_j, \Sigma_j)$ can be determined by procedure $\mathbf{LSVD}(\#I_j, \#I_{[j]}, r_j, \mathcal{M}_j(\mathbf{v}), B_j, \Sigma_j)$ from (2.32), where

$$r_j = \dim(U_j^{\min}(\mathbf{v}))$$

describes the size of $B_j \in \mathbb{K}^{I_j \times r_j}$. The cost $N_{\mathrm{LSVD}}(n_j, \#I_{[j]})$ summed over all $1 \le j \le d$ yields

$$\sum_{j=1}^{d}\left[\left(2\#I_{[j]}-1\right)\frac{n_j}{2}(n_j+1)+\frac{8n_j^3}{3}\right] \approx \sum_{j=1}^{d} n_j\left[\frac{8n_j^2}{3}+\prod_{k=1}^{d} n_k\right] \le dn^{d+1}+\frac{8}{3}dn^3.$$

For $d \ge 3$, the dominant part of the cost is dn^{d+1} arising from the evaluation of the matrix entries of $M_j := \mathcal{M}_j(\mathbf{v})\mathcal{M}_j(\mathbf{v})^{\mathsf{H}}$:

$$M_j[\nu,\mu] = \sum_{\mathbf{i}\in I_{[j]}} \mathbf{v}[i_1,\cdots,i_{j-1},\nu,i_{j+1},\cdots,i_d]\,\overline{\mathbf{v}[i_1,\cdots,i_{j-1},\mu,i_{j+1},\cdots,i_d]}.$$

If the HOSVD tensor subspace format of $\mathbf{v}$ is desired, one has to determine the coefficient tensor $\mathbf{a}$. Lemma 8.17 implies that

$$\mathbf{a} = \mathbf{B}^{\mathsf{H}}\mathbf{v} \qquad \text{with } \mathbf{B} := \bigotimes_{j=1}^{d} B_j. \tag{8.31}$$

i.e., $\mathbf{a}[k_1, \ldots, k_d] = \sum_{i_1 \in I_1} \cdots \sum_{i_d \in I_d} B_1[k_1, i_1] \cdots B_d[k_d, i_d]\, \mathbf{v}[i_1 i_2 \cdots i_d]$. The cost for evaluating $\mathbf{a}$ is

$$\sum_{j=1}^{d} (2n_j - 1) \cdot \prod_{k=1}^{j} r_k \cdot \prod_{k=j+1}^{d} n_k \lesssim 2r_1 n^d.$$

In this estimate we assume that $r_j \ll n_j$, so that the terms for $j > 1$ containing $2r_1 r_2 n^{d-1}$, $2r_1 r_2 r_3 n^{d-2}, \ldots$ are much smaller than the first term. Obviously, the summation should be started with $j^* = \operatorname{argmin}\{r_j : 1 \le j \le d\}$.

Above, we have first determined the HOSVD bases and afterwards performed the projection (8.31). In fact, it is advantageous to apply the projection by $B_j B_j^{\mathsf{H}}$ immediately after the computation of B_j, since the projection reduces the size of the tensor:

```
start:  | $\mathbf{v}_0 := \mathbf{v}$
loop:   | for $j := 1$ to $d$ do
        | begin $(B_j, \Sigma_j) :=$ HOSVD$_j(\mathbf{v}_{j-1})$;
        |         $\mathbf{v}_j := \left(id \otimes \ldots \otimes id \otimes B_j^{\mathsf{H}} \otimes id \otimes \ldots \otimes id\right) \mathbf{v}_{j-1}$        (8.32)
        | end;
return: | $\mathbf{a} := \mathbf{v}_d$
```

Set $\mathbf{B}^{(1,d)} := B_1^{\mathsf{H}} \otimes \bigotimes_{j=2}^{d} id$ and $\mathbf{B}^{(1,d-1)} := B_1^{\mathsf{H}} \otimes \bigotimes_{j=3}^{d} id$. Lemma 5.6 implies the identity $\mathcal{M}_2(\mathbf{v}_1) = \mathcal{M}_2(\mathbf{B}^{(1,d)}\mathbf{v}) = \mathcal{M}_2(\mathbf{v})\mathbf{B}^{(1,d-1)\mathsf{T}}$. Since $\mathbf{B}^{(1,d-1)\mathsf{H}}\mathbf{B}^{(1,d-1)}$ is the projection onto the subspace $U_1^{\min}(\mathbf{v}) \otimes \bigotimes_{j=2}^{d} V_j$ which contains $\mathbf{v}$, one has $\mathbf{B}^{(1,d-1)\mathsf{H}}\mathbf{B}^{(1,d-1)}\mathbf{v} = \mathbf{v}$. This proves

$$\begin{aligned} \mathcal{M}_2(\mathbf{v}_1)\mathcal{M}_2(\mathbf{v}_1)^{\mathsf{H}} &= \mathcal{M}_2(\mathbf{v}) \left(\mathcal{M}_2(\mathbf{v})\, \mathbf{B}^{(1,d-1)\mathsf{T}}\, \overline{\mathbf{B}^{(1,d-1)}}\right)^{\mathsf{H}} \\ &= \mathcal{M}_2(\mathbf{v}) \left(\mathcal{M}_2(\mathbf{B}^{(1,d-1)\mathsf{H}}\, \mathbf{B}^{(1,d-1)}\mathbf{v})\right)^{\mathsf{H}} = \mathcal{M}_2(\mathbf{v})\mathcal{M}_2(\mathbf{v})^{\mathsf{H}}. \end{aligned}$$

Similarly, one proves the identity $\mathcal{M}_j(\mathbf{v}_{j-1})\mathcal{M}_j(\mathbf{v}_{j-1})^{\mathsf{H}} = \mathcal{M}_j(\mathbf{v})\mathcal{M}_j(\mathbf{v})^{\mathsf{H}}$ implying $(B_j, \Sigma_j) = \mathrm{HOSVD}_j(\mathbf{v}_{j-1}) = \mathrm{HOSVD}_j(\mathbf{v})$.

Remark 8.29. The cost of algorithm (8.32) is

$$\sum_{j=1}^{d} \left[(n_j + 2r_j) \cdot \prod_{k=1}^{j-1} r_k \cdot \prod_{k=j}^{d} n_k + \frac{8}{3} n_j^3 \right]. \tag{8.33}$$

Under the assumptions $r_j \ll n_j$ and $d \ge 3$, the dominant part is $n_1 \prod_{k=1}^{d} n_k$.

8.3.3.2 Case B: r-Term Format

In §8.5 we shall discuss conversions between format. The present case is already such a conversion from r-term format into HOSVD tensor subspace representation (other variants will be discussed in §8.5.2).

Let $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)} \in \mathcal{R}_r$ be given. First, all scalar products $\langle v_\nu^{(j)}, v_\mu^{(j)} \rangle$ $(1 \le j \le d,\ 1 \le \nu, \mu \le r)$ are to be computed. We discuss in detail the computation of $(B_1, \Sigma_1) = \mathrm{HOSVD}_1(\mathbf{v})$. The first matricisation is given by

$$\mathcal{M}_1(\mathbf{v}) = \sum_{\nu=1}^{r} v_\nu^{(1)} \otimes v_\nu^{[1]} \qquad \text{with} \quad v_\nu^{[1]} = \bigotimes_{j=2}^{d} v_\nu^{(j)}.$$

By definition, B_1 and Σ_1 from $\mathrm{HOSVD}_1(\mathbf{v})$ results from the diagonalisation of $M_1 := \mathcal{M}_1(\mathbf{v})\mathcal{M}_1(\mathbf{v})^{\mathsf{H}} = B_1 \Sigma_1^2 B_1^{\mathsf{H}}$. We exploit the special structure of $\mathcal{M}_1(\mathbf{v})$:

$$M_1 = \sum_{\nu=1}^{r}\sum_{\mu=1}^{r} \left\langle v_\nu^{[1]}, v_\mu^{[1]} \right\rangle v_\nu^{(1)} (v_\mu^{(1)})^{\mathsf{H}} = \sum_{\nu=1}^{r}\sum_{\mu=1}^{r} \left(\prod_{j=2}^{d} \left\langle v_\nu^{(j)}, v_\mu^{(j)} \right\rangle \right) v_\nu^{(1)} (v_\mu^{(1)})^{\mathsf{H}}.$$

M_1 has the form

$$M_1 = A_1 C_1 A_1^{\mathsf{H}} \quad \text{with} \begin{cases} A_1 := [v_1^{(1)}\ v_2^{(1)}\ \cdots v_r^{(1)}] \text{ and} \\ C_1 := \bigodot_{j=2}^{d} G_j \text{ with } G_j := \left(\langle v_\nu^{(j)}, v_\mu^{(j)} \rangle \right)_{\nu,\mu=1}^{r} \end{cases}$$

(here, $\bigodot_{j=2}^{d}$ denotes the multiple Hadamard product; cf. §4.6.4). As explained in Remark 2.36, the diagonalisation $M_1 = B_1 \Sigma_1^2 B_1^{\mathsf{H}}$ is not performed directly. Instead, one uses $M_1 = A_1 C_1 A_1^{\mathsf{H}} = Q_1 R_1 C_1 R_1^{\mathsf{H}} Q_1^{\mathsf{H}}$ and diagonalises $R_1 C_1 R_1^{\mathsf{H}}$. In the following algorithm, the index j varies from 1 to d:

form Gram matrices $G_j := \left(\langle v_\nu^{(j)}, v_\mu^{(j)} \rangle \right)_{\nu,\mu=1}^{r}$;	$G_j \in \mathbb{K}^{r \times r}$	1
compute Hadamard products $C_j := \bigodot_{k \ne j} G_k$;	$C_j \in \mathbb{K}^{r \times r}$	2
$[v_1^{(j)} \cdots v_r^{(j)}] = Q_j R_j$ $(Q_j \in \mathbb{K}^{n_j \times r_j}, R_j \in \mathbb{K}^{r_j \times r})$;	$r_j = \mathrm{rank}(Q_j R_j)$	3
form products $A_j := R_j C_j R_j^{\mathsf{H}}$;	$A_j \in \mathbb{K}^{r_j \times r_j}$	4
diagonalise $A_j = U_j \Lambda_j U_j^{\mathsf{H}}$;	$U_j, \Lambda_j \in \mathbb{K}^{r_j \times r_j}$	5
return $\Sigma_j := \Lambda_j^{1/2}$ and $B_j := Q_j U_j$;	$B_j \in \mathbb{K}^{n_j \times r_j}$	6

(8.34a)

In Line 3, rank $r_j = \dim(U_j^{\min}(\mathbf{v}))$ is determined. Therefore, $\mathbf{r} = (r_1, \ldots, r_d)$ is the Tucker rank which should be distinguished from the representation rank r of $\mathbf{v} \in \mathcal{R}_r$. Line 6 delivers the values $(B_j, \Sigma_j) = \mathrm{HOSVD}_j(\mathbf{v})$.

It remains to determine the coefficient tensor $\mathbf{a} \in \bigotimes_{j=1}^{d} \mathbb{K}^{r_j}$ of $\mathbf{v}$. As known from Theorem 8.36, also $\mathbf{a}$ possesses an r-term representation. Indeed,

$$\mathbf{a} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} u_\nu^{(j)} \qquad \text{with } u_\nu^{(j)}[i] = \langle v_\nu^{(j)}, b_i^{(j)} \rangle = \left(R_j^{\mathsf{H}} U_j \right)[\nu, i]. \tag{8.34b}$$

Hence, one obtains the hybrid format (8.19) of $\mathbf{v}$. If wanted, one may convert $\mathbf{a}$ into full representation. This would yield the standard tensor subspace format of $\mathbf{v}$.

Remark 8.30. The computational cost of (8.34a,b), up to lower order terms, is

$$r\sum_{j=1}^{d}\left[n_j\left(r+2\min(n_j,r)\right)+r_j\left(2r+3r_j\right)\right]+\sum_{j=1}^{d}\left(\tfrac{8}{3}r_j+2n_j\right)r_j^2$$
$$\leq d\left(2nr\min(n,r)+nr^2+2n\bar{r}^2+2r^2\bar{r}+3r\bar{r}^2+\tfrac{8}{3}\bar{r}^3\right)$$

with $n:=\max_j n_j$ and $\bar{r}:=\max_j r_j$.

Proof. The cost of each line in (8.34a) is $\frac{r(r+1)}{2}\sum_{j=1}^{d}(2n_j-1)$ (line 1), $(d-1)\,r(r+1)$ (line 2), $\sum_{j=1}^{d}N_{\mathrm{QR}}(n_j,r) = 2r\sum_{j=1}^{d}n_j\min(n_j,r)$ (line 3), $(2r-1)\sum_{j=1}^{d}r_j(r+\frac{r_j+1}{2})$ (line 4), $\frac{8}{3}\sum_{j=1}^{d}r_j^3$ (line 5), and $\sum_{j=1}^{d}r_j\left(1+n_j(2r_j-1)\right)$ (line 6), while (8.34b) requires $r\sum_{j=1}^{d}r_j(2r_j-1)$ operations. □

This approach is in particular favourable, if $r, r_j \ll n_j$, since n_j appears only linearly, whereas squares of r and third powers of r_j are present.

For later purpose we mention an approximative variant, which can be used for large r. The next remark is formulated for the first step $j=1$ in (8.34a). The other steps are analogous.

Remark 8.31. Assume that the norm[11] of $\mathbf{v}=\sum_{\nu=1}^{r}\bigotimes_{j=1}^{d}v_\nu^{(j)}\in\mathcal{R}_r$ is known, say, $\|\mathbf{v}\|\approx 1$. Normalise the vectors by $\|v_\nu^{(j)}\|=1$ for $2\leq j\leq d$. Instead of the exact QR decomposition $[v_1^{(1)}\cdots v_r^{(1)}]=Q_1R_1$ apply the algorithm from Corollary 2.39 and Remark 2.40. According to Corollary 2.39, we may omit sufficiently small terms and reduce r to r_0 (called m_0 in Corollary 2.39). Alternatively or additionally, one may use the approximate scalar product $\langle\cdot,\cdot\rangle_p$ from Remark 2.40 for the cheap approximate computation of G_j.

8.3.3.3 Case C: Orthonormal Tensor Subspace Format

Let $\mathbf{v}\in\mathbf{V}$ be represented by $\mathbf{v}=\rho_{\mathrm{orth}}(\hat{\mathbf{a}},(\hat{B}_j)_{1\leq j\leq d})$, i.e., $\mathbf{v}=\hat{\mathbf{B}}\,\hat{\mathbf{a}}$. According to Lemma 8.24b, the HOSVD bases of $\mathbf{v}$ can be derived from the HOSVD bases of the coefficient tensor $\hat{\mathbf{a}}$. Having determined $(\hat{U}_j,\Sigma_j):=\mathrm{HOSVD}_j(\hat{\mathbf{a}})$, we obtain $(B_j,\Sigma_j):=\mathrm{HOSVD}_j(\mathbf{v})$ by means of $B_j:=\hat{B}_j\hat{U}_j$.

Case C1: Assume that $\hat{\mathbf{a}}\in\mathbb{K}^{\hat{\mathbf{J}}}$ is given in full format with $\hat{\mathbf{J}}=\times_{j=1}^{d}\hat{J}_j$, $\hat{J}_j=\{1,\ldots,\hat{r}_j\}$. The computation of $(\hat{U}_j,\Sigma_j):=\mathrm{HOSVD}_j(\hat{\mathbf{a}})$ together with the evaluation of the coefficient tensor $\mathbf{a}\in\mathbb{K}^{\mathbf{J}}$, $\mathbf{J}=\times_{j=1}^{d}J_j$, $J_j=\{1,\ldots,r_j\}$, with the property $\hat{\mathbf{a}}=\hat{\mathbf{U}}\,\mathbf{a}$, $\hat{\mathbf{U}}:=\bigotimes_{j=1}^{d}\hat{U}_j$, requires

[11] Because of the instability discussed later, the norm of $\mathbf{v}$ may be much smaller than the sum of the norms of all terms (cf. Definition 9.15).

$$\sum_{j=1}^{d}\Big[(\hat{r}_j+2r_j)\cdot\prod_{k=1}^{j-1}r_k\cdot\prod_{k=j}^{d}\hat{r}_k+\frac{8}{3}\hat{r}_j^3\Big] \tag{8.35a}$$

operations (cf. (8.33)), where $r_j=\dim(U_j^{\min}(\hat{\mathbf{a}}))=\dim(U_j^{\min}(\mathbf{v}))$. Because of $r_j\le\hat{r}_j$, we get the estimate by $\sum_{j=1}^{d}\big[3\hat{r}_j\cdot\prod_{k=1}^{d}\hat{r}_k+\frac{8}{3}\hat{r}_j^3\big]\le 3d\overline{\hat{r}}^{d+1}+\frac{8}{3}d\overline{\hat{r}}^3$ with $\overline{\hat{r}}:=\max_j\hat{r}_j$.

The cost of $B_j:=\hat{B}_j\hat{U}_j$ for all $1\le j\le d$ is

$$\sum_{j=1}^{d}(2\hat{r}_j-1)\,n_jr_j. \tag{8.35b}$$

In total, the computational work is estimated by

$$3d\overline{\hat{r}}^{d+1}+2d\overline{\hat{r}}^2\Big(n+\frac{4}{3}\overline{\hat{r}}\Big)\quad\text{with } n,\overline{\hat{r}} \text{ from (8.15).} \tag{8.35c}$$

Case C2: Assume that $\hat{\mathbf{a}}\in\mathbb{K}^{\hat{\mathbf{J}}}$ is given in r-term format $\hat{\mathbf{a}}=\sum_{\nu=1}^{r}a_\nu\bigotimes_{j=1}^{d}v_\nu^{(j)}$. By Remark 8.30 (with n_j replaced by $\hat{r}_j$), the cost of $(\hat{U}_j,\Sigma_j):=\mathrm{HOSVD}_j(\hat{\mathbf{a}})$ including the computation of $\mathbf{a}$ with $\hat{\mathbf{a}}=\hat{\mathbf{U}}\mathbf{a}$ amounts to

$$r\sum_{j=1}^{d}\big[\hat{r}_j\,(r+2\min(\hat{r}_j,r))+r_j\,(2r+3r_j)\big]+\sum_{j=1}^{d}\Big(\frac{8}{3}r_j+2\hat{r}_j\Big)r_j^2$$
$$\le(d+2)\,r^2\overline{\hat{r}}+2dr\overline{\hat{r}}\min(\overline{\hat{r}},r)+3r\overline{\hat{r}}^2+\frac{14}{3}d\overline{\hat{r}}^3.$$

Adding the cost (8.35b) of $B_j:=\hat{B}_j\hat{U}_j$, we obtain the following operation count.

Remark 8.32. If the coefficient tensor $\hat{\mathbf{a}}$ in $\mathbf{v}=\rho_{\mathrm{orth}}(\hat{\mathbf{a}},(\hat{B}_j)_{1\le j\le d})$ is represented as $\hat{\mathbf{a}}=\rho_{\text{r-term}}(r,(v_\nu^{(j)})_{1\le j\le d,1\le\nu\le r})$, the computation of the HOSVD bases B_j and of the coefficient tensor $\mathbf{a}\in\mathcal{R}_r$ in $\mathbf{v}=\rho_{\mathrm{HOSVD}}(\mathbf{a},(B_j)_{1\le j\le d})$ requires

$$\sum_{j=1}^{d}\Big[r\hat{r}_j\,(r+2\min(\hat{r}_j,r))+rr_j\,(2r+3r_j)+\Big(\frac{8}{3}r_j+2\hat{r}_j\Big)r_j^2+2n_j\hat{r}_jr_j\Big]$$
$$\le 2dnr\overline{\hat{r}}+(d+2)\,r^2\overline{\hat{r}}+2dr\overline{\hat{r}}\min(\overline{\hat{r}},r)+3r\overline{\hat{r}}^2+\frac{14}{3}d\overline{\hat{r}}^3 \tag{8.36}$$

operations, where $\overline{\hat{r}}:=\max_j\hat{r}_j$ and $n:=\max_j n_j$ as in (8.15).

8.3.3.4 Case D: General and Hybrid Tensor Subspace Format

In the case of $\mathbf{v}=\rho_{\mathrm{frame}}(\hat{\mathbf{a}},(\hat{B}_j)_{j=1}^{d})$ with non-orthonormal bases, the simplest approach combines the following steps:

Step 1: convert the representation into orthonormal tensor subspace format $\mathbf{v} = \rho_{\mathrm{orth}}(\mathbf{a}', (B_j')_{j=1}^d)$ by one of the methods described in §8.2.3.2.

Step 2: apply the methods from §8.3.3.3 to obtain $\mathbf{v} = \rho_{\mathrm{HOSVD}}(\mathbf{a}, (B_j)_{j=1}^d)$.

Alternatively, one may determine $\rho_{\mathrm{HOSVD}}(\mathbf{a}, (B_j)_{j=1}^d)$ directly from the full tensor $\hat{\mathbf{a}}$ and the bases $(\hat{B}_j)_{j=1}^d$ in $\mathbf{v} = \rho_{\mathrm{frame}}(\hat{\mathbf{a}}, (\hat{B}_j)_{j=1}^d)$. However, this approach turns out to be more costly.[12]

The situation differs, at least for $r < n_j$, for the hybrid format, when the coefficient tensor $\hat{\mathbf{a}}$ is given in r-term format: $\hat{\mathbf{a}} = \sum_{\nu=1}^r a_\nu \bigotimes_{j=1}^d v_\nu^{(j)} \in \mathcal{R}_r$. First, the Gram matrices

$$G^{(j)} := \Big(\big\langle \hat{b}_\nu^{(j)}, \hat{b}_\mu^{(j)} \big\rangle\Big)_{\nu,\mu=1}^{\hat{r}_j} \in \mathbb{K}^{\hat{r}_j \times \hat{r}_j} \qquad (1 \le j \le d) \tag{8.37}$$

are to be generated (cost: $\sum_j n_j \hat{r}_j^2$). A different kind of Gram matrices are $G_j \in \mathbb{K}^{r\times r}$ with $G_j[\nu,\mu] := \langle G^{(j)} v_\nu^{(j)}, v_\mu^{(j)} \rangle$, whose computation costs $\sum_{j=1}^d (2r\hat{r}_j^2 + r^2\hat{r}_j)$. If $\hat{B}_j$ represents a basis, a modification is possible: compute the Cholesky decompositions $G^{(j)} = L^{(j)} L^{(j)\mathsf{H}}$ and use $G_j[\nu,\mu] = \langle L^{(j)\mathsf{H}} v_\nu^{(j)}, L^{(j)\mathsf{H}} v_\mu^{(j)} \rangle$. Then, the operation count $\sum_{j=1}^d (\frac{1}{3}\hat{r}_j^3 + \hat{r}_j^2 r + r^2 \hat{r}_j)$ is reduced because of the triangular shape of $L^{(j)\mathsf{H}}$. The matrices G_j correspond to the equally named matrices in the first line of (8.34a). Since the further steps are identical to those in (8.34a), one has to compare the costs $d(n\overline{\hat{r}}^2 + 2r\overline{\hat{r}}^2 + r^2\overline{\hat{r}})$ or $d(n\overline{\hat{r}}^2 + \frac{1}{3}\overline{\hat{r}}^3 + \overline{\hat{r}}^2 r + r^2\overline{\hat{r}})$ from above with the sum of $d\overline{\hat{r}}^2(2n + r)$ from Corollary 8.20 plus $dr^2 n$ for (8.34a$_1$) ($n, \overline{\hat{r}}$ from (8.15)). Unless $r \gg n$, the direct approach is cheaper. We summarise the results below. For the sake of simplicity, we compare the upper bounds.

Remark 8.33. Let $\mathbf{v} = \rho_{\mathrm{frame}}(\hat{\mathbf{a}}, (\hat{B}_j)_{j=1}^d)$ with $\hat{B}_j \in \mathbb{K}^{n_j \times \hat{r}_j}$ and assume that $\hat{\mathbf{a}}$ has an r-term representation. Then the direct method is advantageous if $n > r$. Its cost is by $d\big[\,(n-r)\,\overline{\hat{r}}^2 + \big(n - \overline{\hat{r}}\big)\,r^2\big]$ cheaper than the combination of Step 1 and 2 from above. The Cholesky modification mentioned above is even cheaper by $d\big[\,\big(n - \overline{\hat{r}}/3\big)\,\overline{\hat{r}}^2 + \big(n - \overline{\hat{r}}\big)\,r^2\big]$.

8.4 Sensitivity

We have two types of parameters: the coefficient tensor $\mathbf{a}$ and the basis vectors $b_i^{(j)}$. Perturbations in $\mathbf{a}$ are very easy to describe, since they appear linearly.

A perturbation

$$\tilde{\mathbf{a}} := \mathbf{a} + \delta\mathbf{a}$$

leads to a perturbation $\delta\mathbf{v} = \sum_{\mathbf{i}} \delta\mathbf{a}_{\mathbf{i}} \bigotimes_{j=1}^d b_{i_j}^{(j)}$ of $\mathbf{v}$. In the case of a general

[12] Starting from (8.37) and Kronecker products $\mathbf{G}^{[j]} := \bigotimes_{k\neq j} G^{(k)}$, one has to determine matrices M_j with entries $\big\langle \mathbf{G}^{[j]}\hat{\mathbf{a}}, \hat{\mathbf{a}} \big\rangle_{[j]}$ using the partial scalar product in $\bigotimes_{k\neq j} \mathbb{K}^{\hat{r}_j}$ (cf. §4.5.4).

crossnorm and a general basis, we have

$$\|\delta\mathbf{v}\| \leq \sum_{\mathbf{i}} |\delta\mathbf{a}_{\mathbf{i}}| \prod_{j=1}^{d} \|b_{i_j}^{(j)}\|.$$

For an orthonormal basis and Hilbert norm, we can use that the products $\bigotimes_{j=1}^{d} b_{i_j}^{(j)}$ are pairwise orthonormal and get

$$\|\delta\mathbf{v}\| \leq \sqrt{\sum\nolimits_{\mathbf{i}} |\delta\mathbf{a}_{\mathbf{i}}|^2} =: \|\delta\mathbf{a}\|\,,$$

where the norm on the right-hand side is the Euclidean norm.

For perturbations of the basis vectors we give only a differential analysis, i.e., we consider only a small perturbation in one component. Furthermore, we assume that the $b_i^{(j)}$ form orthonormal bases. Without loss of generality we may suppose that $b_1^{(1)}$ is perturbed into $b_1^{(1)} + \delta_1^{(1)}$. Then

$$\tilde{\mathbf{v}} = \mathbf{v} + \sum_{i_2,\ldots,i_d} \mathbf{a}[1, i_2, \ldots, i_d]\; \delta_1^{(1)} \otimes b_{i_2}^{(2)} \otimes b_{i_3}^{(3)} \otimes \cdots \otimes b_{i_d}^{(d)},$$

i.e., $\delta\mathbf{v} = \sum_{i_2\cdots i_d} \mathbf{a}[1, i_2, \ldots, i_d]\; \delta_1^{(1)} \otimes \bigotimes_{j=2}^{d} b_{i_j}^{(j)}$. Terms with different $(i_2, \ldots, i_d)$ are orthogonal. Therefore

$$\|\delta\mathbf{v}\| = \|\delta_1^{(1)}\| \sqrt{\sum_{i_2,\ldots,i_d} |\mathbf{a}[1, i_2, \ldots, i_d]|^2}.$$

As a consequence, for small perturbations $\delta_i^{(j)}$ of all $b_i^{(j)}$, the first order approximation is

$$\begin{aligned} \|\delta\mathbf{v}\| &\approx \sum\nolimits_{j=1}^{d} \sum_{\ell} \|\delta_\ell^{(j)}\| \sqrt{\sum_{i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_d} |\mathbf{a}[i_1, \ldots, i_{j-1}, \ell, i_{j+1}, \ldots, i_d]|^2} \\ &\leq \|\mathbf{v}\| \sum\nolimits_{j=1}^{d} \sqrt{\sum\nolimits_{\ell} \|\delta_\ell^{(j)}\|^2} \leq \|\mathbf{v}\| \sqrt{d}\, \sqrt{\sum\nolimits_{j,\ell} \|\delta_\ell^{(j)}\|^2}. \end{aligned} \tag{8.38}$$

Here, we have used the Schwarz inequality

$$\begin{aligned} &\sum_{\ell} \|\delta_\ell^{(j)}\| \sqrt{\sum_{i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_d} |\mathbf{a}[i_1, \ldots, i_{j-1}, \ell, i_{j+1}, \ldots, i_d]|^2} \\ &\leq \sqrt{\sum\nolimits_{\ell} \|\delta_\ell^{(j)}\|^2}\, \sqrt{\sum\nolimits_{\mathbf{i}} |\mathbf{a}_{\mathbf{i}}|^2} \end{aligned}$$

together with $\sum_{\mathbf{i}} |\mathbf{a}_{\mathbf{i}}|^2 = \|\mathbf{v}\|^2$ (cf. Exercise 8.15). The last inequality in (8.38) is again Schwarz' inequality.

8.5 Relations between the Different Formats

So far, three formats (full representation, $\mathcal{R}_r$, $\mathcal{T}_\mathbf{r}$) have been introduced; in addition, there is the hybrid format $\mathcal{R}_r \cap \mathcal{T}_\mathbf{r}$. A natural question is how to convert one format into another one. It will turn out that conversions between $\mathcal{R}_r$ and $\mathcal{T}_\mathbf{r}$ lead to the hybrid format introduced in §8.2.4. The conversions $\mathcal{R}_r \to \mathcal{T}_\mathbf{r}$ and $\mathcal{T}_\mathbf{r} \to \mathcal{R}_r$ are described in §8.5.2 and §8.5.3. The mapping from $\mathcal{R}_r$ into the HOSVD representation is already mentioned in §8.3.3.2. For completeness, the full format is considered in §8.5.1 (see also §7.6.1).

8.5.1 *Conversion from Full Representation into Tensor Subspace Format*

Assume that $\mathbf{v} \in \mathbf{V} := \bigotimes_{j=1}^{d} \mathbb{K}^{n_j}$ is given in full representation. The translation into tensor subspace format is $\mathbf{v} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{v}_\mathbf{i} \bigotimes_{j=1}^{d} b_{i_j}^{(j)}$ with the unit basis vectors $b_i^{(j)} := e^{(i)} \in \mathbb{K}^{n_j}$ from (2.2). Here, the tensor $\mathbf{v}$ and its coefficient tensor are identical. The memory cost of the tensor subspace format is even larger because of the additional basis vectors. In order to reduce the memory, one may determine the minimal subspaces $U_j^{\min}(\mathbf{v}) \subset \mathbb{K}^{n_j}$, e.g., by the HOSVD representation. If $r_j = \dim(U_j^{\min}(\mathbf{v})) < n_j$, the memory cost $\prod_{j=1}^{d} n_j$ is reduced to $\prod_{j=1}^{d} r_j$.

8.5.2 *Conversion from $\mathcal{R}_r$ to $\mathcal{T}_\mathbf{r}$*

The letter 'r' is the standard variable name for all kinds of ranks. Here, one has to distinguish the tensor rank or representation rank r in $\mathbf{v} \in \mathcal{R}_r$ from the vector-valued tensor subspace rank $\mathbf{r}$ with components r_j.

8.5.2.1 Theoretical Statements

First, we recall the special situation of $d = 2$ (matrix case). The matrix rank is equal to all ranks introduced for tensors: $\mathit{matrix\text{-}rank} = \mathit{tensor\text{-}rank} = \mathrm{rank}_1 = \mathrm{rank}_2$. Therefore, there is an r-term representation $\mathbf{v} = \sum_{i=1}^{r} v_i^{(1)} \otimes v_i^{(2)}$ with $r = \mathrm{rank}(\mathbf{v})$. Since $\{v_i^{(1)} : 1 \leq i \leq r\}$ and $\{v_i^{(2)} : 1 \leq i \leq r\}$ are sets of linearly independent vectors, they can be used as bases and yield a tensor subspace representation (8.6b) for $\mathbf{r} = (r, r)$ with the coefficients $\mathbf{a}_{ij} = \delta_{ij}$. The singular value decomposition yields another r-term representation $\mathbf{v} = \sum_{i=1}^{r} \sigma_i u_i \otimes v_i$, which is an orthonormal tensor subspace representation (8.14a) (with orthonormal bases $\{u_i\}$, $\{v_i\}$, and coefficients $\mathbf{a}_{ij} = \delta_{ij}\sigma_i$). The demonstrated identity $\mathcal{R}_r = \mathcal{T}_{(r,r)}$ of the formats does not extend to $d \geq 3$.

Given an r-term representation

$$\mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)}, \tag{8.39}$$

Remark 6.1 states that $\mathbf{v} \in {}_a\bigotimes_{j=1}^{d} U_j$ (cf. (6.1)) with $U_j := \mathrm{span}\{v_1^{(j)}, \ldots, v_d^{(j)}\}$. This proves $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ for $\mathbf{r} = (r_1, \ldots, r_d)$, $r_j := \dim(U_j)$. We recall the relations between the different ranks.

Theorem 8.34. *(a) The minimal tensor subspace rank* $\mathbf{r} = (r_1, \ldots, r_d)$ *of* $\mathbf{v} \in \mathbf{V}$ *is given by* $r_j = \mathrm{rank}_j(\mathbf{v})$ *(cf. (5.6b)). The tensor rank* $r = \mathrm{rank}(\mathbf{v})$ *(cf. Definition 3.32) satisfies* $r \geq r_j$*. The tensor rank may depend on the field:* $r_\mathbb{R} \geq r_\mathbb{C} \geq r_j$ *(cf. Proposition 3.40), while* $r_j = \mathrm{rank}_j(\mathbf{v})$ *is independent of the field.*

(b) The inequalities of Part (a) are also valid for the border rank from (9.11) instead of $r,\ r_\mathbb{R},\ r_\mathbb{C}$.

Proof. Part (a) follows from Remark 6.21. In the case of (b), $\mathbf{v}$ is the limit of a sequence of tensors $\mathbf{v}_n \in \mathcal{R}_r$. There are minimal subspaces $U_j^{\min}(\mathbf{v}_n)$ $(1 \leq j \leq d$, $n \in \mathbb{N})$ of dimension $r_{j,n}$ satisfying $r := \underline{\mathrm{rank}}(\mathbf{v}) \geq r_{j,n}$. By Theorem 6.24 the minimal subspace $U_j^{\min}(\mathbf{v})$ has dimension $r_j := \dim(U_j^{\min}(\mathbf{v})) \leq \liminf_{n\to\infty} r_{j,n} \leq r$. Hence, $r \geq r_j$ is proved. □

Using a basis B_j of $U_j = \mathrm{span}\{v_1^{(j)}, \ldots, v_d^{(j)}\}$, we shall construct a tensor subspace representation $\mathbf{v} = \rho_{\mathrm{TS}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$ in §8.5.2.3. In general, these subspaces U_j may be larger than necessary; however, under the assumptions of Proposition 7.8, $U_j = U_j^{\min}(\mathbf{v})$ are the minimal subspaces. This proves the following statement.

Remark 8.35. If $\mathbf{v}$ is given by the r-term representation (8.39) with $r = \mathrm{rank}(\mathbf{v})$, constructions based[13] on $U_j := \mathrm{span}\{v_1^{(j)}, \ldots, v_d^{(j)}\}$ yield $\mathbf{v} = \rho_{\mathrm{TS}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$ with $r_j = \mathrm{rank}_j(\mathbf{v})$.

Having converted $\mathbf{v} = \rho_{\text{r-term}}(\ldots)$ into $\mathbf{v} = \rho_{\mathrm{TS}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$, the next statement describes the r-term structure of the coefficient tensor $\mathbf{a}$. This helps, in particular, to obtain the hybrid format from §8.2.4.

Theorem 8.36. *Let* $\mathbf{v} = \rho_{\mathrm{TS}}(\mathbf{a}, (B_j)_{j=1}^d)$ *be any tensor subspace representation with the coefficient tensor* $\mathbf{a} \in \mathbb{K}^\mathbf{J}$ *and bases*[14] B_j*. Then the ranks of* $\mathbf{v}$ *and* $\mathbf{a}$ *coincide in two different meanings. First, the true* tensor ranks *satisfy*

$$\mathrm{rank}(\mathbf{a}) = \mathrm{rank}(\mathbf{v}).$$

[13] Representations with $r_j = \mathrm{rank}_j(\mathbf{v})$ can be obtained anyway by HOSVD. These decompositions, however, cannot be obtained from U_j alone.

[14] For general frames B_j the coefficient tensor is *not* uniquely defined and, in fact, different (equivalent) coefficient tensors may have different ranks. However, the minimum of $\mathrm{rank}(\mathbf{a})$ over all equivalent $\mathbf{a}$ coincides with $\mathrm{rank}(\mathbf{v})$.

Second, given (8.39) with representation rank r, *the r-term representation of* $\mathbf{a}$ *(with same number r) can be obtained constructively as detailed in* §*8.5.2.3. The resulting tensor subspace representation of* $\mathbf{v}$ *is the hybrid format (8.19).*

Proof. Consider $\mathbf{v}$ as an element of $\bigotimes_{j=1}^{d} U_j$ with $U_j = \mathrm{span}\{v_1^{(j)}, \ldots, v_d^{(j)}\} \cong \mathbb{K}^{J_j}$. By Lemma 3.36a, the rank is invariant. This proves $\mathrm{rank}(\mathbf{a}) = \mathrm{rank}(\mathbf{v})$. The second part of the statement follows from the constructions in §8.5.2.3. □

8.5.2.2 Conversion into General Tensor Subspace Format

A rather trivial translation of $\mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)} \in \mathcal{R}_r$ into $\mathbf{v} \in \mathcal{T}_{\mathbf{r}}$ with $\mathbf{r} = (r, \ldots, r)$ can be obtained without any arithmetical cost by choosing the frames

$$B_j = [v_1^{(j)}, \ldots, v_r^{(j)}] \tag{8.40a}$$

and the diagonal coefficient tensor

$$\mathbf{a}[i, \ldots, i] = 1 \text{ for } 1 \le i \le r, \qquad \text{and } \mathbf{a}[\mathbf{i}] = 0, \text{ otherwise.} \tag{8.40b}$$

Obviously, $\mathbf{v} = \rho_{\text{r-term}}(r, (v_i^{(j)})) = \rho_{\mathrm{TS}}(\mathbf{a}, (B_j)_{j=1}^{d})$ holds. Note that there is no guarantee that the frames are bases, i.e., the subspaces $U_j = \mathrm{range}(B_j)$ from (6.2b) may have a dimension less than r (cf. Remark 3.39).

The diagonal tensor $\mathbf{a}$ is a particular case of a sparse tensor (cf. (7.5)).

8.5.2.3 Conversion into Orthonormal Hybrid Tensor Subspace Format

Let $V_j = \mathbb{K}^{I_j}$. An orthonormal basis of the subspace U_j from above can be obtained by a QR decomposition of the matrix $A_j := [v_1^{(j)} \cdots v_r^{(j)}] \in \mathbb{K}^{I_j \times r}$. Procedure **RQR** from (2.29) yields $A_j = B_j R_j$ with an orthogonal matrix $B_j \in \mathbb{K}^{I_j \times r_j}$, where $r_j = \mathrm{rank}(A_j) = \dim(U_j)$. The second matrix R_j allows the representations $v_k^{(j)} = B_j R_j[\bullet, k] = \sum_{i=1}^{r_j} r_{i,k}^{(j)} b_i^{(j)}$. From this we derive

$$\mathbf{v} = \sum_{k=1}^{r} \bigotimes_{j=1}^{d} v_k^{(j)} = \sum_{k=1}^{r} \bigotimes_{j=1}^{d} \sum_{i_j=1}^{r_j} r_{i_j,k}^{(j)} b_{i_j}^{(j)} \tag{8.41a}$$

$$= \sum_{i_1=1}^{r_1} \cdots \sum_{i_d=1}^{r_d} \underbrace{\left[\sum_{k=1}^{r} \prod_{j=1}^{d} r_{i_j,k}^{(j)} \right]}_{\mathbf{a}[i_1 \cdots i_d]} \bigotimes_{j=1}^{d} b_{i_j}^{(j)},$$

proving $\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^{d}\big)$ with the coefficient tensor $\mathbf{a}$ described in the r-term format

$$\mathbf{a} = \sum_{k=1}^{r} \bigotimes_{j=1}^{d} r_k^{(j)}, \qquad r_k^{(j)} := \left(r_{i,k}^{(j)}\right)_{i=1}^{r_j} \in \mathbb{K}^{r_j}. \tag{8.41b}$$

This yields the orthonormal hybrid format $\mathbf{v} = \rho_{\mathrm{orth}}^{\mathrm{hybr}}\big(r, (r_k^{(j)}), (B_j)_{j=1}^d\big)$ (cf. (8.20)).

The conversion cost caused by the QR decomposition is $\sum_{j=1}^n N_{\mathrm{QR}}(n_j, r)$ with $n_j := \#I_j$. If $r \le n := \max_j n$, the cost is bounded by $2nr^2$.

The construction from above can equivalently be obtained by the following two steps: (i) apply the approach of §8.5.2.2 and (ii) perform an orthonormalisation of the frame as described in §8.2.2. Note that the transformation to new orthonormal bases destroys the sparsity of the coefficient tensor (8.40b).

8.5.2.4 Case of Large *r*

As seen from (8.41a,b), the representation rank r of $\mathbf{v} \in \mathcal{R}_r(\mathbb{K}^{\mathbf{I}})$ is inherited by the coefficient tensor $\mathbf{a} \in \mathcal{R}_r(\mathbb{K}^{\mathbf{J}})$. In §7.6.2, the conversion of $\mathbf{v} \in \mathcal{R}_r$ into full format is proposed, provided that $r > N := \big(\prod_{j=1}^d n_j\big)/\max_{1\le i\le d} n_i$. Now, the same consideration can be applied to $\mathbf{a} \in \mathcal{R}_r$, provided that

$$r > R := \Big(\prod\nolimits_{j=1}^{d} r_j\Big) / \max_{1\le i\le d} r_i, \tag{8.42}$$

where r_j is obtained in §8.5.2.3 as size of the basis $B_j \in \mathbb{K}^{I_j \times J_j}$, $J_j = \{1, \ldots, r_j\}$. Lemma 7.16 and Remark 7.17 show that a conversion of $\mathbf{a} \in \mathcal{R}_r(\mathbb{K}^{\mathbf{J}})$ into full format or R-term format $\mathbf{a} \in \mathcal{R}_R(\mathbb{K}^{\mathbf{J}})$ requires $2r\prod_{j=1}^d r_j$ operations. The cost to obtain $\mathbf{a} \in \mathcal{R}_r$ is $\sum_{j=1}^n N_{\mathrm{QR}}(n_j, r) \le 2r\sum_{j=1}^n n_j \min\{r, n_j\}$ (cf. §8.5.2.3). This yields the following result.

Lemma 8.37. *Assume* $\mathbf{v} = \rho_{\text{r-term}}\big(r, (v_\nu^{(j)})\big)$ *with* r *satisfying (8.42). Then,* $\mathbf{v}$ *can be converted into* $\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$ *or a hybrid format with* $\mathbf{a} \in \mathcal{R}_R(\mathbb{K}^{\mathbf{J}})$ *requiring* $2r\big(\prod_{j=1}^d r_j + \sum_{j=1}^n n_j \min\{r, n_j\}\big)$ *operations.*

8.5.3 Conversion from $\mathcal{T}_\mathbf{r}$ to $\mathcal{R}_r$

The tensor subspace representation

$$\mathbf{v} = \sum_{k_1=1}^{r_1} \sum_{k_2=1}^{r_2} \cdots \sum_{k_d=1}^{r_d} \mathbf{a}[k_1 k_2 \cdots k_d]\, b_{k_1}^{(1)} \otimes b_{k_2}^{(2)} \otimes \cdots \otimes b_{k_d}^{(d)}$$

from (8.6b) or (8.14a) is an r-term representation of $\mathbf{v}$ with $r := \prod_{j=1}^d r_j$ terms. To reach the format (7.7a): $\mathbf{v} = \sum_{\mathbf{k}} \bigotimes_{j=1}^d v_{\mathbf{k}}^{(j)}$, the vectors $v_{\mathbf{k}}^{(1)}$ for $j = 1$ could be defined by $\mathbf{a}[k_1 k_2 \cdots k_d] b_{k_1}^{(1)}$, while $v_{\mathbf{k}}^{(j)} := b_{k_j}^{(j)}$ for $j > 1$. Following the proof of Lemma 3.41, an improvement is possible. Choose the largest r_ℓ. Without loss of generality, assume $r_1 \ge r_j$ for all j. Rewrite $\mathbf{v}$ as

$$\mathbf{v} = \sum_{k_2=1}^{r_2} \cdots \sum_{k_d=1}^{r_d} \underbrace{\left(\sum_{k_1=1}^{r_1} \mathbf{a}[k_1 k_2 \cdots k_d] \, b_{k_1}^{(1)} \right)}_{=: \ \hat{b}^{(1)}[k_2 \cdots k_d]} \otimes b_{k_2}^{(2)} \otimes \cdots \otimes b_{k_d}^{(d)}. \tag{8.43}$$

This is an r-term representation of $\mathbf{v}$ with $r := \prod_{j=2}^{d} r_j$ terms.

Because of the presumably very large number r of terms, the storage requirement $r \sum_{j=1}^{d} size(U_j)$ of the r-term representation of $\mathbf{v}$ seems huge. An improvement can be based on the fact that the r factors $v_\nu^{(j)}$ $(1 \le \nu \le r)$ in $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$ for $j \ge 2$ are not r different vectors, but can be expressed by the basis $(b_i^{(j)})_{i=1}^{r_j}$, which is already stored (cf. (7.13)). This leads again to the hybrid format, now considered as a particular case of the r-term format (cf. (8.21)).

Next, we assume that $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ is given in the hybrid format from (8.19):

$$\mathbf{v} = \sum_{i_1=1}^{r_1} \cdots \sum_{i_d=1}^{r_d} \left(\sum_{\nu=1}^{r} \prod_{j=1}^{d} a_\nu^{(j)}[i_j] \right) \bigotimes_{j=1}^{d} b_{i_j}^{(j)}.$$

Using the reformulation (8.43), we have to compute the vectors

$$\hat{b}^{(1)}[i_2 \cdots i_d] := \sum_{i_1=1}^{r_1} \left(\sum_{\nu=1}^{r} \prod_{j=1}^{d} a_\nu^{(j)}[i_j] \right) b_{i_1}^{(1)} \in V_1 = \mathbb{K}^{n_1}.$$

As above, the direction $k = 1$ is chosen because of the assumption $r_1 \ge r_j$ for all j, so that the resulting representation rank $N = \prod_{j=2}^{d} n_j$ is minimal.

Remark 8.38. Let $\mathbf{v}$ be given in the hybrid format from (8.19) with r_j, r as above. Conversion into N-term format with $N = \prod_{j=2}^{d} n_j$ requires $N\big((d-1)r + 2n_1 r_1\big)$ operations.

8.5.4 Comparison of Both Representations

We summarise the results from above.

Remark 8.39. (a) If $\mathbf{v} \in \mathcal{R}_r$, then $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ with $\mathbf{r} = (r, \ldots, r)$.

(b) If $\mathbf{v} \in \mathcal{R}_r$, a hybrid format $\mathcal{R}_r \cap \mathcal{T}_\mathbf{r}$ can be constructed with $\mathbf{r} = (r_1, \ldots, r_d)$, $r_j = \mathrm{rank}_j(\mathbf{v})$.

(c) If $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ with $\mathbf{r} = (r_1, \ldots, r_d)$, then $\mathbf{v} \in \mathcal{R}_r$ with $r := \frac{\prod_{j=1}^{d} r_j}{\max_{1 \le j \le d} r_j}$.

For simplification, we assume in the following that $V_j = \mathbb{K}^n$ (i.e., dimension n independent of j).

The transfer between both representations is quite non-symmetric. According to Remark 8.39a, $\mathbf{v} \in \mathcal{R}_r$ yields $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ with $\mathbf{r} = (r, r, \ldots, r)$. Note that vectors of same size have to be stored:

$$N_{\mathrm{mem}}^{r\text{-term}} = r \cdot d \cdot n = N_{\mathrm{mem}}^{\mathrm{TSR}}\left((B_j)_{1\le j\le d}\right) \qquad \text{(cf. (7.8c) and (8.6d)).}$$

Additionally, the tensor subspace representation needs storage for the coefficient tensor $\mathbf{a}$:

$$N_{\mathrm{mem}}^{\mathrm{TSR}}(\mathbf{a}) = r^d$$

(cf. Remark 8.7b). This large additional memory cost makes the tensor subspace representation clearly less advantageous. The hybrid format from Remark 8.39b needs the storage

$$N_{\mathrm{mem}}^{\mathrm{hybr}} = (n+r)\sum_{j=1}^{d} r_j,$$

which may be smaller than $N_{\mathrm{mem}}^{r\text{-term}}$, if $r_j < r < n$.

On the other hand, if a tensor $\mathbf{v} \in \mathcal{T}_{\mathbf{r}}$ with $\mathbf{r} = (r, \cdots, r)$ is converted into the r^{d-1}-term representation from (8.43), the latter format requires storage of size

$$N_{\mathrm{mem}}^{N\text{-term}} = r^{d-1} \cdot d \cdot n.$$

Since $r \le n$ (and usually $r \ll n$), the inequality $N_{\mathrm{mem}}^{\mathrm{TSR}} = r \cdot d \cdot n + r^d \ll r^{d-1} \cdot n < r^{d-1} \cdot d \cdot n = N_{\mathrm{mem}}^{r\text{-term}}$ indicates that the tensor subspace representation is by far better.

The previous examples underline that none of the formats $\mathcal{R}_r$ or $\mathcal{T}_{\mathbf{r}}$ are, in general, better than the other. It depends on the nature of the tensor what format is to be preferred. Often, the hybrid format is the best compromise.

8.5.5 r-Term Format for Large $r > N$

In §7.6.3, we have discussed the case of $r > N$, where N is the bound of the maximal rank from (8.44b). In particular in the case of $d = 3$ one may use an intermediate tensor subspace representation (and, possibly, approximation tools for this format; see §10). Here, we assume that a tensor is represented in the r-term format,

$$\mathbf{v} = \sum_{i=1}^{r} v_i^{(1)} \otimes v_i^{(2)} \otimes v_i^{(3)} \in \mathbb{K}^{n_1\times n_2\times n_3}, \tag{8.44a}$$

with rather large r. The term 'rather large' may, e.g., mean

$$r > N := \min\{n_1 n_2, n_1 n_3, n_2 n_3\}. \tag{8.44b}$$

Instead of a large r, one may also assume N to be rather small.

In this case, the following procedure yields an *exact* N'-term representation with $N' \le N < r$.

Step 1. Convert $\mathbf{v}$ from r-term format into tensor subspace format (hybrid variant from §8.2.4).

Step 2. Convert $\mathbf{v}$ back into N'-term format with $N' \le N$ (N from (8.44b); see §8.5.3).

8.6 Joining two Tensor Subspace Representation Systems

8.6.1 Setting of the Problem

Let $\mathbf{v}' = \rho_{\mathrm{TS}}(\mathbf{a}', (B'_j)_{j=1}^d)$ and $\mathbf{v}'' = \rho_{\mathrm{TS}}(\mathbf{a}'', (B''_j)_{j=1}^d)$ be two tensors from $\mathbf{V} = \bigotimes_{j=1}^d V_j$ involving different subspaces U'_j and U''_j. Obviously, the sum $\mathbf{v}' + \mathbf{v}''$ requires the spaces U_j defined by

$$U_j := U'_j + U''_j \qquad \text{for } 1 \le j \le d. \tag{8.45}$$

A common systems B_j spanning U_j is to be constructed. A subtask is to transform the coefficient tensors $\mathbf{a}'$ of $\mathbf{v}'$ and $\mathbf{a}''$ of $\mathbf{v}''$ into the new coefficient tensors referring to the new bases B_j.

8.6.2 Trivial Joining of Frames

The least requirement is that

$$\begin{aligned} B'_j &= [b_1'^{(j)}, b_2'^{(j)}, \ldots, b_{r'_j}'^{(j)}] \in \left(U'_j\right)^{r'_j} \qquad \text{and} \\ B''_j &= [b_1''^{(j)}, b_2''^{(j)}, \ldots, b_{r'_j}''^{(j)}] \in \left(U''_j\right)^{r''_j} \end{aligned}$$

are frames spanning the respective subspaces U'_j and U''_j. The respective index sets are $\mathbf{J}' = \times_{j=1}^d J'_j$ and $\mathbf{J}'' = \times_{j=1}^d J''_j$ with $J'_j = \{1, \ldots, r'_j\}$ and $J''_j = \{1, \ldots, r''_j\}$. Since no linear independence is required, the simple definition

$$\begin{aligned} B_j &:= \left[B'_j \; B''_j\right] = \left[b_1'^{(j)}, b_2'^{(j)}, \ldots, b_{r'_j}'^{(j)}, b_1''^{(j)}, b_2''^{(j)}, \ldots, b_{r'_j}''^{(j)}\right], \\ J_j &:= \{1, \ldots, r_j\} \text{ with } r_j := r'_j + r''_j, \end{aligned}$$

yields a frame associated with the subspace $U_j := U'_j + U''_j$. The representation rank r_j is the sum of the previous ones, even if the subspaces U'_j and U''_j overlap.

An advantage is the easy construction of the coefficients. The columns of $B_j = [b_1^{(j)}, \ldots, b_{r_j}^{(j)}]$ are $b_i^{(j)} := b_i'^{(j)}$ for $1 \le i \le r'_j$ and $b_{i+r'_j}^{(j)} := b_i''^{(j)}$ for $1 \le i \le r''_j$. The coefficient $\mathbf{a}'$ of

$$\mathbf{v}' = \sum_{\mathbf{i}' \in \mathbf{J}'} \mathbf{a}'_{\mathbf{i}'} \bigotimes_{j=1}^d b_{i'_j}'^{(j)} \in \bigotimes_{j=1}^d U'_j \tag{8.46a}$$

becomes

$$\mathbf{v}' = \sum_{\mathbf{i} \in \mathbf{J}} \mathbf{a}_{\mathbf{i}} \bigotimes_{j=1}^d b_{i_j}^{(j)} \in \bigotimes_{j=1}^d U_j \tag{8.46b}$$

with $\mathbf{a}_{\mathbf{i}} := \mathbf{a}'_{\mathbf{i}}$ for $\mathbf{i} \in \mathbf{J}'$ and $\mathbf{a}_{\mathbf{i}} := 0$ for $\mathbf{i} \in \mathbf{J} \backslash \mathbf{J}'$. Analogously, a coefficient $\mathbf{a}''$ of $\mathbf{v}'' = \sum_{\mathbf{i}'' \in \mathbf{J}''} \mathbf{a}'_{\mathbf{i}'} \bigotimes_{j=1}^d b_{i''_j}''^{(j)}$ becomes $\mathbf{a}_{\mathbf{i}+\mathbf{r}'} := \mathbf{a}''_{\mathbf{i}}$ for $\mathbf{i} \in \mathbf{J}''$ with $\mathbf{r}' = (r'_1, \ldots, r'_d)$ and $\mathbf{a}_{\mathbf{i}} := 0$ otherwise.

Remark 8.40. The joining of frames requires only a rearrangement of data, but no arithmetical operations.

8.6.3 Common Bases

Now, we assume that B'_j and B''_j are bases of U'_j and U''_j. We want to construct a new, common basis B_j for the sum $U_j = U'_j + U''_j$. Applying the procedure

$$\mathbf{JoinBases}(B'_j, B''_j, r_j, B_j, T'_j, T''_j) \tag{8.47a}$$

from (2.35), we produce a common basis $B_j = [\, b_1^{(j)}, \ldots, b_{r_j}^{(j)}]$ of dimension r_j and transformation matrices T'_j and T''_j with the property

$$B'_j = B_j T'_j \quad \text{and} \quad B''_j = B_j T''_j \tag{8.47b}$$

(cf. (2.34)).

Lemma 8.41. *Let* $\mathbf{a}' \in \mathbb{K}^{\mathbf{J}'}$ *be the coefficient tensor of* $\mathbf{v}'$ *from (8.46a) with respect to the bases* B'_j. *The coefficient tensor* $\mathbf{a}'_{\mathsf{new}} \in \mathbb{K}^{\mathbf{J}}$ *of* $\mathbf{v}'$ *with respect to the bases* B_j *satisfying (8.47b) is given by*

$$\mathbf{a}'_{\mathrm{new}} = \left(\bigotimes_{j=1}^{d} T'_j \right) \mathbf{a}'$$

(cf. (8.46b)). Similarly, the coefficient tensor $\mathbf{a}'' \in \mathbb{K}^{\mathbf{J}''}$ *representing* $\mathbf{v}'' \in \bigotimes_{j=1}^{d} U''_j$ *transforms into* $\mathbf{a}''_{\mathsf{new}} = (\bigotimes_{j=1}^{d} T''_j)\mathbf{a}'' \in \mathbb{K}^{\mathbf{J}}$ *with respect to the bases* B_j. *A possible option in procedure* **JoinBases** *is to take* B'_j *as the first part of* B_j, *which leads to* $T'_j = \left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right]$.

Remark 8.42. Assume $V_j = \mathbb{K}^{n_j}$. The cost of (8.47a) is $N_{\mathrm{QR}}(n_j, r'_j + r''_j)$. The coefficient tensor $\mathbf{a}'_{\mathsf{new}}$ is without arithmetical cost under the option mentioned in Lemma 8.41, while $\mathbf{a}''_{\mathsf{new}}$ requires $2\sum_{j=1}^{d} \left[(\prod_{\ell=1}^{j} r''_\ell)(\prod_{\ell=j}^{d} r_\ell) \right]$ operations. If $n_j \leq n$ and $r_j \leq r$, the total cost can be estimated by $8dnr^2 + 2dr^{d+1}$.

The cost for the basis transformation is much less, if $\mathbf{v}'$ and $\mathbf{v}''$ are given in hybrid format. Here, we use that $\mathbf{a}'' = \sum_{\nu=1}^{r''} \bigotimes_{j=1}^{d} a_\nu^{\prime\prime(j)}$ and that $\mathbf{a}''_{\mathsf{new}} = (\bigotimes_{j=1}^{d} T''_j)\mathbf{a}'' = \sum_{\nu=1}^{r''} \bigotimes_{j=1}^{d} a_{\nu,\mathsf{new}}^{\prime\prime(j)}$ with $a_{\nu,\mathsf{new}}^{\prime\prime(j)} = T''_j a_\nu^{\prime\prime(j)}$ costs $2r'' \sum_{j=1}^{d} r_j r''_j$ operations.

Remark 8.43. In the case of hybrid tensors $\mathbf{v}'$ and $\mathbf{v}''$, the computation of common bases and the transformation of the coefficient tensors cost $N_{\mathrm{QR}}(n_j, r'_j + r''_j) + 2r'' \sum_{j=1}^{d} r_j r''_j$. If all ranks are bounded by r and $n_j \leq n$, the total cost can be estimated by $8dnr^2 + 2dr^3$.

In the case of a tensor subspace representation with *orthonormal* bases, procedure **JoinBases** is to be replaced by **JoinONB**.

Chapter 9
r-Term Approximation

Abstract In general, one tries to approximate a tensor $\mathbf{v}$ by another tensor $\mathbf{u}$ requiring less data. The reason is twofold: the memory size should decrease and, hopefully, operations involving $\mathbf{u}$ should require less computational work. In fact, $\mathbf{u} \in \mathcal{R}_r$ leads to decreasing cost for storage and operations as r decreases. However, the other side of the coin is an increasing approximation error. Correspondingly, in *Sect. 9.1* two approximation strategies are presented, where either the representation rank r of $\mathbf{u}$ or the accuracy is prescribed. Before we study the approximation problem in general, two particular situations are discussed. *Section 9.2* is devoted to $r = 1$, when $\mathbf{u} \in \mathcal{R}_1$ is an elementary tensor. The matrix case $d = 2$ is recalled in *Sect. 9.3*. The properties observed in the latter two sections contrast with the true tensor case studied in *Sect. 9.4*. Numerical algorithms solving the approximation problem will be discussed in *Sect. 9.5*. Modified approximation problems are addressed in *Sect. 9.6*.

9.1 Two Approximation Problems

In (7.8c), the storage requirement of an r-term representation $\mathbf{u} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)} \in \mathcal{R}_r$ under the assumption $n_j = n$ for all $1 \le j \le d$ is described by $N_{\mathrm{mem}}^{r\text{-term}}(p) = r \cdot d \cdot n$. On the other hand, the size of r is bounded by $r \le n^{d-1}$ (cf. Lemma 3.41). If we insert this inequality, we get an upper bound of $N_{\mathrm{mem}}^{r\text{-term}}(p) \le d \cdot n^d$ which is worse than the storage size for the full representation (cf. (7.4)).

The consequence is that the r-term representation makes sense only if the involved rank r is of moderate size. Since the true rank r may be large (or even infinite), an exact representation may be impossible, and instead one has to accept approximations of the tensor.

Let $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ be a Banach tensor space. The *approximation problem* (truncation problem) can be formulated in two different versions. In the *first version* we fix the representation rank r and look for approximations in $\mathcal{R}_r$:

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_9,

$$\begin{array}{l}\text{Given } \mathbf{v} \in \mathbf{V} \text{ and } r \in \mathbb{N}_0, \\ \text{determine } \mathbf{u} \in \mathcal{R}_r \text{ minimising } \|\mathbf{v} - \mathbf{u}\| .\end{array} \tag{9.1}$$

Here, $\|\cdot\|$ is an appropriate norm[1] on $\mathbf{V}$.

We shall see that, in general, a minimiser $\mathbf{u} \in \mathcal{R}_r$ of Problem (9.1) need not exist, but we can form the infimum

$$\varepsilon(\mathbf{v}, r) := \varepsilon(r) := \inf \{\|\mathbf{v} - \mathbf{u}\| : \mathbf{u} \in \mathcal{R}_r\} . \tag{9.2}$$

In §9.3 we shall discuss modified formulations of Problem (9.1).

In the next variant, the rôles of r and $\varepsilon(r)$ are reversed:

$$\begin{array}{l}\text{Given } \mathbf{v} \in \mathbf{V} \text{ and } \varepsilon > 0, \\ \text{determine } \mathbf{u} \in \mathcal{R}_r \text{ with } \|\mathbf{v} - \mathbf{u}\| \leq \varepsilon \text{ for minimal } r.\end{array} \tag{9.3}$$

There are two trivial cases which will not be discussed further on. One case is $r=0$, because $\mathcal{R}_0 = \{0\}$ leads to the solution $\mathbf{u}=0$. The second case is $d=1$, since then $\mathcal{R}_r = \mathbf{V}$ for all $r \geq 1$ and $\mathbf{u} := \mathbf{v}$ is the perfect minimiser of (9.1) and (9.3).

Remark 9.1. Problem (9.3) has always a solution.

Proof. Let $N_r := \{\|\mathbf{v} - \mathbf{u}\| : \mathbf{u} \in \mathcal{R}_r\} \subset [0, \infty)$ be the range of the norm. Given $\varepsilon > 0$, we have to ensure that there are some $r \in \mathbb{N}_0$ and $\varepsilon' \in N_r$ with $\varepsilon \geq \varepsilon'$. Then

$$N(\varepsilon) := \{r \in \mathbb{N}_0 : \text{there is some } \varepsilon' \in N_r \text{ with } \varepsilon' \leq \varepsilon\}$$

is a non-empty subset of $\mathbb{N}_0$ and a minimum $r := \min\{n \in N(\varepsilon)\}$ must exist.

First we consider the finite dimensional case. Then there is a finite $r_{\max}$ with $\mathbf{V} = \mathcal{R}_{r_{\max}}$. Hence, $\varepsilon' := 0 \in N_{r_{\max}}$ satisfies $\varepsilon' \leq \varepsilon$.

In the infinite dimensional case, ${}_a\bigotimes_{j=1}^d V_j$ is dense in $\mathbf{V}$. This implies that there is a $\mathbf{u}^\varepsilon \in {}_a\bigotimes_{j=1}^d V_j$ with $\varepsilon' := \|\mathbf{v} - \mathbf{u}^\varepsilon\| \leq \varepsilon/2$. By definition of the algebraic tensor space, $\mathbf{u}^\varepsilon$ has a representation of r elementary tensors for some $r \in \mathbb{N}_0$. This proves that $\varepsilon \geq \varepsilon' \in N_r$. □

Solutions $\mathbf{u} \in \mathcal{R}_r$ of Problem (9.3) satisfy $\|\mathbf{v} - \mathbf{u}\| \leq \varepsilon$ for a minimal r. Fixing this rank r, we may still ask for the best approximation among all $\mathbf{u} \in \mathcal{R}_r$. This leads again to Problem (9.1).

Lemma 9.2. *Let $(V_j, \langle\cdot,\cdot\rangle_j)$ be Hilbert spaces, while $(\mathbf{V}, \langle\cdot,\cdot\rangle)$ is endowed with the induced scalar product $\langle\cdot,\cdot\rangle$ and the corresponding norm $\|\cdot\|$. If a minimiser $\mathbf{u}^* \in \mathcal{R}_r$ of Problem (9.1) exists, then*

$$\mathbf{u}^* \in \mathbf{U}(\mathbf{v}) := \bigotimes_{j=1}^d U_j^{\min}(\mathbf{v}).$$

If, furthermore, $\mathbf{v}$ satisfies linear constraints (cf. §6.8), these are also fulfilled by $\mathbf{u}^*$. If $\mathbf{V} = \mathbf{V}^{(\mathbf{0})}$ is subspace of a Hilbert intersection space $\mathbf{V}^{(\mathbf{n})}$ and $\mathbf{v} \in \mathbf{V}^{(\mathbf{n})}$, then also $\mathbf{u}^* \in \mathbf{V}^{(\mathbf{n})}$ (cf. Uschmajew [188]). In the case of Problem (9.3), one of the solutions satisfies $\mathbf{u}^* \in \mathbf{U}(\mathbf{v})$ and the conclusions about possible constraints.

[1] Some results are stated for general norms, however, most of the practical algorithms will work for Hilbert tensor spaces with induced scalar product.

Proof. Let $\mathbf{P} : \mathbf{V} \to \mathbf{U}(\mathbf{v}) := \bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$ be the orthogonal projection onto $\mathbf{U}(\mathbf{v})$. Because of $\|\mathbf{v}-\mathbf{u}\|^2 \underset{\mathbf{Pv}=\mathbf{v}}{=} \|\mathbf{v}-\mathbf{Pu}\|^2 + \|(\mathbf{I}-\mathbf{P})\mathbf{u}\|^2$ and $\mathbf{Pu} \in \mathcal{R}_r$, the minimiser $\mathbf{u}$ must satisfy $(\mathbf{I}-\mathbf{P})\mathbf{u} = 0$, i.e., $\mathbf{u} \in \mathbf{U}(\mathbf{v})$. The further statements follow from $\mathbf{u} \in \mathbf{U}(\mathbf{v})$. □

9.2 Discussion for $r = 1$

Because of Exercise 8.2b, the following results can be derived from the later proven properties of $\mathcal{T}_{\mathbf{r}}$. Nevertheless, we discuss the case $\mathcal{R}_1$ as an exercise (cf. Zhang-Golub [201]) and as demonstration of the contrast to the case $r > 1$.

Let $\mathbf{v} \in \mathbf{V}$ be given. Problem (9.1) with $r = 1$ requires the minimisation of

$$\left\| \mathbf{v} - \bigotimes_{j=1}^{d} u^{(j)} \right\|. \tag{9.4}$$

In the following we assume that the vector spaces V_j are finite dimensional (corresponding results for infinitely dimensional spaces follow from Theorem 10.8 combined with Exercise 8.2b). Note that the choice of the norm is not restricted.

Lemma 9.3. *Let $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ be a finite dimensional normed tensor space. Then for any $\mathbf{v} \in \mathbf{V}$ there are tensors $\mathbf{u}_{\min} = \bigotimes_{j=1}^{d} u^{(j)} \in \mathcal{R}_1$ minimising (9.4):*

$$\|\mathbf{v} - \mathbf{u}_{\min}\| = \min_{u^{(1)} \in V_1, \ldots, u^{(d)} \in V_d} \left\| \mathbf{v} - \bigotimes_{j=1}^{d} u^{(j)} \right\|. \tag{9.5}$$

If $d \geq 2$ and $\dim(V_j) \geq 2$ for at least two indices j, the minimiser $\mathbf{u}_{\min}$ may be not unique.

Proof. 1) If $\mathbf{v} = 0$, $\mathbf{u} = 0 \in \mathcal{R}_1$ is the unique solution.

2) For the rest of the proof assume $\mathbf{v} \neq 0$. Furthermore, we may assume, without loss of generality, that there are norms $\|\cdot\|_j$ on V_j scaled in such a way that

$$\left\| \bigotimes_{j=1}^{d} v^{(j)} \right\| \geq \prod_{j=1}^{d} \|v^{(j)}\|_j \qquad \text{(cf. (4.27))}. \tag{9.6}$$

3) For the minimisation in $\min_{\mathbf{u} \in \mathcal{R}_1} \|\mathbf{v}-\mathbf{u}\|$, the set $\mathcal{R}_1$ may be reduced to the subset $C := \{\mathbf{u} \in \mathcal{R}_1 : \|\mathbf{u}\| \leq 2\|\mathbf{v}\|\}$, since otherwise

$$\|\mathbf{v}-\mathbf{u}\| \underset{(4.2)}{\geq} \|\mathbf{u}\| - \|\mathbf{v}\| > 2\|\mathbf{v}\| - \|\mathbf{v}\| = \|\mathbf{v}\| = \|\mathbf{v} - 0\|,$$

i.e., $0 \in \mathcal{R}_1$ is a better approximation than $\mathbf{u}$. Consider the subsets

$$C_j := \left\{ u^{(j)} \in V_j : \|v^{(j)}\|_j \leq (2\|\mathbf{v}\|)^{1/d} \right\} \subset V_j$$

and note that $C \subset C' := \left\{ \bigotimes_{j=1}^{d} u^{(j)} : u^{(j)} \in C_j \right\}$ because of (9.6). We conclude that

$$\inf_{\mathbf{u}\in\mathcal{R}_1}\|\mathbf{v}-\mathbf{u}\| = \inf_{\mathbf{u}\in C'}\|\mathbf{v}-\mathbf{u}\| = \inf_{u^{(j)}\in C_j}\left\|\mathbf{v}-\bigotimes\nolimits_{j=1}^{d}u^{(j)}\right\|.$$

Let $\mathbf{u}^\nu := \bigotimes\limits_{j=1}^{d} u^{j,\nu}$ $(u^{j,\nu}\in C_j)$ be a sequence with $\|\mathbf{v}-\mathbf{u}^\nu\| \to \inf\limits_{\mathbf{u}\in\mathcal{R}_1}\|\mathbf{v}-\mathbf{u}\|$. Since the sets C_j are bounded and closed, the finite dimension of V_j implies compactness. We find a subsequence so that $u^{j,\nu}\to u_*^{(j)}\in V_j$ and $\mathbf{u}_*:=\bigotimes_{j=1}^{d}u_*^{(j)}\in\mathcal{R}_1$ satisfies $\|\mathbf{v}-\mathbf{u}_*\| = \inf_{\mathbf{u}\in\mathcal{R}_1}\|\mathbf{v}-\mathbf{u}\|$.

4) For $d=1$, $\mathbf{v}\in\mathbf{V}$ belongs to $\mathcal{R}_1$ so that $\mathbf{u}_*:=\mathbf{v}$ is the only minimiser. Already for $d=2$, the matrix $\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]$ has the two different minimisers $\mathbf{u}_* = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\otimes\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] = \left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right]$ and $\mathbf{u}_{**} = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\otimes\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] = \left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right]$ with respect to the Frobenius norm. □

The practical computation of $\mathbf{u}_{\min}$ from (9.5) is rendered more difficult by the following fact.

Remark 9.4. The function $\Phi(u^{(1)},\ldots,u^{(d)}) = \|\mathbf{v}-\bigotimes_{j=1}^{d}u^{(j)}\|$ may have local minima larger than the global minimum.

For Problems (9.1) and (9.3) with $r\geq 2$ we have to distinguish the cases $d=2$ (see §9.3) and $d\geq 3$ (see §9.4).

9.3 Discussion in the Matrix Case $d = 2$

First we discuss the general case of finite dimensional vector spaces V_1 and V_2 and the tensor space $\mathbf{V}=V_1\otimes V_2$ with arbitrary norm $\|\cdot\|$. Introducing bases $\{b_i^{(j)}: i\in I_j\}$ in V_1 and V_2, we obtain isomorphisms $V_j\cong\mathbb{K}^{I_j}$. Similarly, the tensor space $\mathbf{V}$ is isomorphic to $\mathbb{K}^{I_1\times I_2}$. Given a norm $\|\cdot\|$ on $\mathbf{V}$, we define the equally named norm

$$\|M\| := \left\|\sum_{\nu\in I_1}\sum_{\mu\in I_2} M_{\nu\mu}\, b_\nu^{(1)}\otimes b_\mu^{(2)}\right\| \qquad \text{for } M\in\mathbb{K}^{I_1\times I_2}.$$

This makes the isomorphism $\mathbf{V}\cong\mathbb{K}^{I_1\times I_2}$ isometric. Therefore, Problem (9.1) is equivalent to

$$\begin{array}{l}\text{Given a matrix } M\in\mathbb{K}^{I_1\times I_2} \text{ and } r\in\mathbb{N}_0,\\ \text{determine } R\in\mathcal{R}_r \text{ minimising } \|M-R\|,\end{array} \tag{9.7}$$

with $\mathcal{R}_r = \left\{M\in\mathbb{K}^{I_1\times I_2} : \operatorname{rank}(M)\leq r\right\}$ from (2.6).

Proposition 9.5. *For $d=2$, Problems (9.1) and (9.7) have a solution, i.e., the minima $\min_{\mathbf{u}\in\mathcal{R}_r}\|\mathbf{v}-\mathbf{u}\|$ and $\min_{R\in\mathcal{R}_r}\|M-R\|$ are attained.*

Proof. Since the problems (9.1) and (9.7) are equivalent, we focus to Problem (9.7). As in the proof of Lemma 9.3 we find that the minimisation in (9.7) may be reduced to the bounded subset $\mathcal{R}_{r,M} := \mathcal{R}_r\cap\left\{R\in\mathbb{K}^{I\times J} : \|R\|\leq 2\|M\|\right\}$. We consider a sequence $R^{(k)}\in\mathcal{R}_{r,M}$ such that

$$\|M - R^{(k)}\| \to \inf_{R\in\mathcal{R}_r} \|M - R\| .$$

Since $\{R\in\mathbb{K}^{I\times J} : \|R\| \leq 2\,\|M\|\}$ is a compact set, there is a subsequence (denoted again by $R^{(k)}$) with $\lim R^{(k)} =: R^* \in \mathbb{K}^{I_1\times I_2}$. Continuity of $\|\cdot\|$ (cf. §4.1.1) implies that $\inf_{R\in\mathcal{R}_r} \|M - R\| = \|M - R^*\|$. It remains to show that $R^* \in \mathcal{R}_r$. Lemma 2.4 proves that indeed $\operatorname{rank}(R^*) = \operatorname{rank}(\lim R^{(k)}) \leq \liminf_{k\to\infty} \operatorname{rank}(R^{(k)}) \leq r$. □

Next, we consider the Frobenius norm[2] $\|\cdot\| = \|\cdot\|_\mathsf{F}$.

Proposition 9.6. *In the case of Problem (9.7) with the Frobenius norm $\|\cdot\| = \|\cdot\|_\mathsf{F}$, the characterisation of the solution is derived from the singular value decomposition $M = U\Sigma V^\mathsf{H} = \sum_{i=1}^{s} \sigma_i u_i v_i^\mathsf{H}$ (cf. (2.21)). Then*

$$R := \sum_{i=1}^{\min\{r,s\}} \sigma_i\, u_i\, v_i^\mathsf{H}$$

is a solution to Problem (9.7). It is unique if $r = s$ or $\sigma_{r+1} < \sigma_r$. The remaining error is

$$\|M - R\|_\mathsf{F} = \sqrt{\sum\nolimits_{i=r+1}^{s} \sigma_i^2}.$$

Proof. Use Lemma 2.30. □

For $\|\cdot\| = \|\cdot\|_\mathsf{F}$, also Problem (9.3) has an immediate solution. The result is deduced from (2.26b).

Remark 9.7. The problem

$$\begin{array}{l}\text{Given } M \in \mathbb{K}^{I\times J} \text{ and } \varepsilon > 0,\\ \text{determine } R \in \mathcal{R}_r \text{ with } \|M - R\|_\mathsf{F} \leq \varepsilon \text{ for minimal } r\end{array} \tag{9.8}$$

has the solution $R := \sum_{i=1}^{r_\varepsilon} \sigma_i u_i v_i^\mathsf{H}$, where $M = \sum_{i=1}^{s} \sigma_i u_i v_i^\mathsf{H}$ with $s := \operatorname{rank}(M)$ (cf. (2.21)) is the singular value decomposition and

$$r_\varepsilon = \min\left\{r \in \{0, \ldots, s\} : \sum\nolimits_{i=r+1}^{s} \sigma_i^2 \leq \varepsilon^2\right\}.$$

There is a connection between the case $r = 1$ from §9.2 and Problem (9.7) for $\|\cdot\| = \|\cdot\|_\mathsf{F}$. We may determine the solution $R \in \mathcal{R}_r$ of (9.8) sequentially by a *deflation technique*, where in each step we determine a best rank-1 matrix $R^{(i)} \in \mathcal{R}_1$:

1) let $R^{(1)} \in \mathcal{R}_1$ be the minimiser of $\min\limits_{s\in\mathcal{R}_1} \|M - S\|_\mathsf{F}$ and set $M^{(1)} := M - R^{(1)}$;

2) let $R^{(2)} \in \mathcal{R}_1$ be the minimiser of $\min\limits_{S\in\mathcal{R}_1} \|M^{(1)} - S\|_\mathsf{F}$ and set $M^{(2)} := M - R^{(2)}$;

⋮

r) let $R^{(r)} \in \mathcal{R}_1$ be the minimiser of $\min\limits_{S\in\mathcal{R}_1} \|M^{(r-1)} - S\|_\mathsf{F}$, set $R := \sum\limits_{i=1}^{r} R^{(i)} \in \mathcal{R}_r$.

[2] We may also choose the matrix norm $\|\cdot\|_2$ from (2.13) or any unitarily invariant matrix norm.

Remark 9.8. The solution of the previous deflation algorithm yields the approximation $R \in \mathcal{R}_r$ which is identical to the solution of Proposition 9.6. The error $\|M^{(r)}\|_\mathsf{F} = \|M - R\|_\mathsf{F}$ is as in Proposition 9.6.

The solutions discussed above satisfy an important *stability property*. We shall appreciate this property later when we find situations, where stability is lacking.

Lemma 9.9. *The solutions $R := \sum_{i=1}^r \sigma_i u_i v_i^\mathsf{H}$ to Problems (9.1) and (9.3) satisfy*[3]

$$\sum\nolimits_{i=1}^r \left\|\sigma_i u_i v_i^\mathsf{H}\right\|_\mathsf{F}^2 = \|R\|_\mathsf{F}^2\,, \tag{9.9}$$

i.e., the terms $\sigma_i u_i v_i^\mathsf{H}$ are pairwise orthogonal with respect to the Frobenius scalar product.

Proof. The Frobenius scalar product (2.10) yields the value $\left\langle \sigma_i u_i v_i^\mathsf{H}, \sigma_j u_j v_j^\mathsf{H} \right\rangle_\mathsf{F} = \sigma_i\sigma_j \left\langle u_i v_i^\mathsf{H}, u_j v_j^\mathsf{H} \right\rangle = \sigma_i\sigma_j \operatorname{trace}\left((u_j v_j^\mathsf{H})^\mathsf{H}(u_i v_i^\mathsf{H})\right) = \sigma_i\sigma_j \operatorname{trace}\left(v_j u_j^\mathsf{H} u_i v_i^\mathsf{H}\right)$. Since the singular vectors u_i are orthogonal, $u_j^\mathsf{H} u_i = 0$ holds for $i \neq j$ proving the orthogonality $\left\langle \sigma_i u_i v_i^\mathsf{H}, \sigma_j u_j v_j^\mathsf{H} \right\rangle_\mathsf{F} = 0$. □

9.4 Discussion in the Tensor Case $d \geq 3$

9.4.1 Non-Closedness of $\mathcal{R}_r$

In the following we assume that the tensor space of order $d \geq 3$ is non-degenerate (cf. Definition 3.24).

A serious difficulty for the treatment of tensors of order $d \geq 3$ is based on the fact that Proposition 9.5 does not extend to $d \geq 3$. The following result stems from De Silva-Lim [44] and is further discussed in Stegeman [175], [176]. However, an example of such a type can already be found in Bini-Lotti-Romani [19].

Proposition 9.10. *Let $\mathbf{V}$ be a non-degenerate tensor space of order $d \geq 3$. Then, independently of the choice of the norm, there are tensors $\mathbf{v} \in \mathbf{V}$ for which Problem (9.1) possesses no solution.*

Proof. Consider the tensor space $\mathbf{V} = V_1 \otimes V_2 \otimes V_3$ with $\dim(V_j) \geq 2$ and choose two linearly independent vectors $v_j, w_j \in V_j$. The tensor

$$\mathbf{v} := v^{(1)} \otimes v^{(2)} \otimes w^{(3)} + v^{(1)} \otimes w^{(2)} \otimes v^{(3)} + w^{(1)} \otimes v^{(2)} \otimes v^{(3)}$$

has tensor rank 3 as proved in Lemma 3.42. Next, we define

$$\begin{aligned} \mathbf{v}_n := {} & \left(w^{(1)} + nv^{(1)}\right) \otimes \left(v^{(2)} + \tfrac{1}{n} w^{(2)}\right) \otimes v^{(3)} \\ & + \qquad v^{(1)} \otimes \qquad v^{(2)} \qquad \otimes \left(w^{(3)} - nv^{(3)}\right) \quad \text{for } n \in \mathbb{N}. \end{aligned} \tag{9.10}$$

[3] Eq. (9.9) has a similar flavour as the estimate $\sum_i \|u_i \otimes v_i\|_\wedge \leq (1+\varepsilon)\|\mathbf{v}\|_\wedge$ for a suitable representation $\mathbf{v} = \sum_i u_i \otimes v_i$ with respect to the projective norm $\|\cdot\|_\wedge$ from §4.2.4.

Exercise 3.43 shows that $\mathrm{rank}(\mathbf{v}_n) = 2$. The identity $\mathbf{v} - \mathbf{v}_n = -\frac{1}{n} w^{(1)} \otimes w^{(2)} \otimes v^{(3)}$ is easy to verify; hence, independently of the choice of norm, one obtains

$$\lim_{n \to \infty} \mathbf{v}_n = \mathbf{v}.$$

This shows

$$3 = \mathrm{rank}(\mathbf{v}) = \mathrm{rank}(\lim \mathbf{v}_n) > \mathrm{rank}(\mathbf{v}_n) = 2$$

in contrary to (2.7).

The tensor space of order 3 from above can be embedded into higher order tensor spaces, so that the statement extends to non-degenerate tensor spaces with $d \geq 3$. □

The proof reveals that the set $\mathcal{R}_2$ is not closed.

Lemma 9.11. *Let* $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ *be a non-degenerate tensor space of order* $d \geq 3$. *Then* $\mathcal{R}_1 \subset \mathbf{V}$ *is closed, but* $\mathcal{R}_r \subset \mathbf{V}$ *for* $2 \leq r \leq \min_{1 \leq j \leq d} \dim(V_j)$ *is not closed.*[4]

Proof. 1) Consider a sequence $\mathbf{v}_n := \bigotimes_{j=1}^{d} u^{j,n} \in \mathcal{R}_1$ with $\mathbf{v} := \lim_{n \to \infty} \mathbf{v}_n \in \mathbf{V}$. Hence, $\inf_{\mathbf{u} \in \mathcal{R}_1} \|\mathbf{v} - \mathbf{u}\| \leq \inf_n \|\mathbf{v} - \mathbf{v}_n\| = 0$. On the other hand, Lemma 9.3 states that the minimum is attained: $\min_{\mathbf{u} \in \mathcal{R}_1} \|\mathbf{v} - \mathbf{u}\| = \|\mathbf{v} - \mathbf{u}_{\min}\|$ for some $\mathbf{u}_{\min} \in \mathcal{R}_1$. Together, we obtain from $0 = \inf_n \|\mathbf{v} - \mathbf{v}_n\| = \|\mathbf{v} - \mathbf{u}_{\min}\|$ that $\mathbf{v} = \mathbf{u}_{\min} \in \mathcal{R}_1$, i.e., $\mathcal{R}_1$ is closed.

2) The fact that $\mathcal{R}_2$ is not closed, is already proved for $d = 3$. The extension to $d \geq 3$ is mentioned in the proof of Proposition 3.40c.

3) For the discussion of $r > 2$ we refer to [44, Theorem 4.10]. □

De Silva and Lim [44] have shown that tensors $\mathbf{v}$ without a minimiser $\mathbf{u}^* \in \mathcal{R}_r$ of $\inf_{\mathbf{u} \in \mathcal{R}_r} \|\mathbf{v} - \mathbf{u}\|$ are not of measure zero, i.e., there is a positive expectation that random tensors $\mathbf{v}$ are of this type.

9.4.2 Border Rank

The observed properties lead to a modification of the tensor rank (cf. Bini et al. [19]), where $\mathcal{R}_r$ is replaced by its closure.

Definition 9.12. The *tensor border rank* is defined by

$$\underline{\mathrm{rank}}(\mathbf{v}) := \min \left\{ r : \mathbf{v} \in \overline{\mathcal{R}_r} \right\} \in \mathbb{N}_0. \tag{9.11}$$

Concerning estimates between $\underline{\mathrm{rank}}(\mathbf{v})$ and the tensor subspace ranks r_j see Theorem 8.34b. A practical application to Kronecker products is given in the following remark.

[4] The limitation $r \leq \min_j \{\dim(V_j)\}$ is used for the proof in [44, Theorem 4.10]. It is not claimed that $\mathcal{R}_r$ is closed for larger r.

Remark 9.13. For $A^{(j)}, B^{(j)} \in \mathcal{L}(V_j, V_j)$, the Kronecker product

$$\mathbf{A} := A^{(1)} \otimes B^{(2)} \otimes \ldots \otimes B^{(d)} + B^{(1)} \otimes A^{(2)} \otimes \ldots \otimes B^{(d)} + \ldots + B^{(1)} \otimes B^{(2)} \otimes \ldots \otimes A^{(d)}$$

has border rank $\underline{\mathrm{rank}}(\mathbf{A}) \leq 2$.

Proof. $\mathbf{A}$ is the derivative $\frac{d}{dt}\mathbf{C}(t)$ of $\mathbf{C}(t) := \bigotimes_{j=1}^{d} \left(B^{(j)} + tA^{(j)}\right)$ at $t = 0$. Since $\mathrm{rank}(\mathbf{C}(t)) = 1$ for all t and $\frac{1}{h}(\mathbf{C}(h) - \mathbf{C}(0)) \to \mathbf{A}$, the result follows. □

9.4.3 Stable and Unstable Sequences

For practical use, it would be sufficient to replace the (non-existing) minimiser $\mathbf{u} \in \mathcal{R}_r$ of $\|\mathbf{v} - \mathbf{u}\|$ by some $\mathbf{u}_\varepsilon \in \mathcal{R}_r$ with $\|\mathbf{v} - \mathbf{u}_\varepsilon\| \leq \inf_{\mathbf{u} \in \mathcal{R}_r} \|\mathbf{v} - \mathbf{u}\| + \varepsilon$ for an ε small enough. However, those $\mathbf{u}_\varepsilon$ with $\|\mathbf{v} - \mathbf{u}_\varepsilon\|$ close to $\inf_{\mathbf{u} \in \mathcal{R}_r} \|\mathbf{v} - \mathbf{u}\|$ suffer from the following instability (below, ε is replaced by $1/n$).

Remark 9.14. $\mathbf{v}_n$ from (9.10) is the sum $\mathbf{v}_n = \mathbf{v}_{n,1} + \mathbf{v}_{n,2}$ of two elementary tensors. While $\|\mathbf{v}_{n,1} + \mathbf{v}_{n,2}\| \leq C$ stays bounded, the norms $\|\mathbf{v}_{n,1}\|$ and $\|\mathbf{v}_{n,2}\|$ grow like n: $\|\mathbf{v}_{n,1}\|, \|\mathbf{v}_{n,2}\| \geq C'n$. Hence, the cancellation of both terms is the stronger the smaller $\|\mathbf{v} - \mathbf{v}_n\|$ is.

Cancellation is an unpleasant numerical effect leading to a severe error amplification. A typical example is the computation of $\exp(-20)$ by $\sum_{\nu=0}^{n} \frac{(-20)^\nu}{\nu!}$ with suitable n. Independently of n, the calculation with standard machine precision $eps = 10^{-16}$ yields a completely wrong result. The reason is that rounding errors produce an absolute error of size $\sum_{\nu=0}^{n} \left| (-20)^\nu / \nu! \right| \cdot eps \approx \exp(+20) \cdot eps$. Hence, the relative error is about $\exp(40) \cdot eps \approx 2.4_{10}17 \cdot eps$.

In general, the 'condition' of a sum $\sum_\nu a_\nu$ of reals can be described by the quotient

$$\sum\nolimits_\nu |a_\nu| \;\Big/\; \left|\sum\nolimits_\nu a_\nu\right|.$$

A similar approach leads us to the following definition of a stable representation (see also (7.12)).

Definition 9.15. Let $\mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$ be a normed tensor space.
(a) For any representation $0 \neq \mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)}$ we define[5]

$$\varkappa\left(\left(v_i^{(j)}\right)_{1 \leq i \leq r}^{1 \leq j \leq d}\right) := \left(\sum_{i=1}^{r} \left\| \bigotimes_{j=1}^{d} v_i^{(j)} \right\|\right) \Big/ \left\| \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)} \right\|. \tag{9.12a}$$

(b) For $\mathbf{v} \in \mathcal{R}_r$ we set

[5] The first sum on the right-hand side of (9.12a) resembles the projective norm (cf. §4.2.4). The difference is that here only r terms are allowed.

$$\varkappa(\mathbf{v}, r) := \inf\left\{\varkappa\Big((v_i^{(j)})_{1\leq i\leq r}^{1\leq j\leq d}\Big) : \mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)}\right\}. \tag{9.12b}$$

(c) A sequence $\mathbf{v}_n \in \mathcal{R}_r$ $(n \in \mathbb{N})$ is called *stable* in $\mathcal{R}_r$, if

$$\varkappa((\mathbf{v}_n)_{n\in\mathbb{N}}, r) := \sup_{n\in\mathbb{N}} \varkappa(\mathbf{v}_n, r) < \infty; \tag{9.12c}$$

otherwise, the sequence is unstable.

The instability observed in Remark 9.14 does not happen accidentally, but is a necessary consequence of the non-closedness of $\mathcal{R}_2$.

Proposition 9.16. *Suppose* $\dim(V_j) < \infty$. *If a sequence* $\mathbf{v}_n \in \mathcal{R}_r \subset {}_a\bigotimes_{j=1}^d V_j$ *is stable and convergent, then* $\lim_{n\to\infty} \mathbf{v}_n \in \mathcal{R}_r$.

Proof. Set $C := 2\varkappa((\mathbf{v}_n), r)$ and $\mathbf{v} := \lim_{n\to\infty} \mathbf{v}_n$. After choosing a subsequence, $\mathbf{v}_n \to \mathbf{v}$ holds with representations $\sum\limits_{i=1}^{r} \bigotimes\limits_{j=1}^{d} v_{n,i}^{(j)}$ such that $\sum\limits_{i=1}^{r} \Big\| \bigotimes\limits_{j=1}^{d} v_{n,i}^{(j)} \Big\| \leq C\,\|\mathbf{v}\|$ holds. The vectors $v_{n,i}^{(j)}$ can be scaled equally so that all $\{v_{n,i}^{(j)} \in V_j : n \in \mathbb{N}\}$ are uniformly bounded. Choosing again a subsequence, limits $\hat{v}_i^{(j)} := \lim_{n\to\infty} v_{n,i}^{(j)}$ exist and $\mathbf{v} = \lim \mathbf{v}_n = \lim \sum_{i=1}^r \bigotimes_{j=1}^d v_{n,i}^{(j)} = \sum_{i=1}^r \bigotimes_{j=1}^d \lim v_{n,i}^{(j)} = \sum_{i=1}^r \bigotimes_{j=1}^d \hat{v}_i^{(j)} \in \mathcal{R}_r$ proves the assertion. □

A generalisation of the last proposition to the infinite dimensional case follows.

Theorem 9.17. *Let* $\mathbf{V}$ *be a reflexive Banach space with a norm not weaker than* $\|\cdot\|_\vee$ *(cf. (6.18)). For any bounded and stable sequence* $\mathbf{v}_n \in \mathbf{V}$, *there is a weakly convergent subsequence* $\mathbf{v}_{n_\nu} \rightharpoonup \mathbf{v} \in \mathcal{R}_r$. *Moreover, if* $\mathbf{v}_n = \sum_{i=1}^r \bigotimes_{j=1}^d v_{n,i}^{(j)} \in \mathcal{R}_r$ *holds with balanced*[6] *factors* $v_{n,i}^{(j)}$, *i.e.,*

$$\sup_n \left\{ \max_j \|v_{n,i}^{(j)}\|_{V_j} / \min_j \|v_{n,i}^{(j)}\|_{V_j} \right\} < \infty \qquad \text{for all } 1 \leq i \leq r,$$

then there are $v_i^{(j)} \in V_j$ *and a subsequence such that*

$$v_{n_\nu,i}^{(j)} \rightharpoonup v_i^{(j)} \quad \text{and} \quad \mathbf{v}_{n_\nu} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_{n_\nu,i}^{(j)} \rightharpoonup \mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)}. \tag{9.13}$$

Proof. 1) By definition of stability, there are $v_{i,n}^{(j)} \in V_j$ such that $\mathbf{v}_{i,n} := \bigotimes_{j=1}^d v_{i,n}^{(j)}$ satisfies $\mathbf{v}_n = \sum_{i=1}^r \mathbf{v}_{i,n}$ and $\|\mathbf{v}_{i,n}\| \leq C\,\|\mathbf{v}_n\|$, e.g., for $C := \varkappa((\mathbf{v}_n), r) + 1$. Corollary 4.26 states that $\mathbf{v}_{i,n} \rightharpoonup \mathbf{v}_i$ and $\mathbf{v}_n \rightharpoonup \mathbf{v} = \sum_{i=1}^r \mathbf{v}_i$ are valid after restricting n to a certain subsequence. Note that $\mathbf{v}_{i,n} \in \mathcal{R}_1 = \mathcal{T}_{(1,\ldots,1)}$ and that $\mathcal{T}_{(1,\ldots,1)}$ is weakly closed (cf. Lemma 8.6). Hence, $\mathbf{v}_i \in \mathcal{R}_1$ implies that $\mathbf{v} = \sum_{i=1}^r \mathbf{v}_i \in \mathcal{R}_r$.

[6] This can be ensured by scaling the factors such that $\|v_{n,i}^{(j)}\|_{V_j} = \|v_{n,i}^{(k)}\|$ for all $1 \leq j, k \leq d$.

2) Taking the subsequence from Part 1), $\mathbf{v}_{i,n} = \bigotimes_{j=1}^{d} v_{i,n}^{(j)} \rightharpoonup \mathbf{v}_i \in \mathcal{R}_1$ holds, i.e., $\mathbf{v}_i = \bigotimes_{j=1}^{d} \hat{v}_i^{(j)}$. Since $\mathbf{v}_i = 0$ is a trivial case, assume $\mathbf{v}_i \neq 0$. Choose functionals $\varphi^{(j)} \in V_j^*$ with $\varphi^{(j)}(\hat{v}_i^{(j)}) = 1$ and define $\boldsymbol{\Phi}^{[k]} := \bigotimes_{j\neq k} \varphi^{(j)} : \mathbf{V} \to V_k$. The weak convergence $\mathbf{v}_{i,n} \rightharpoonup \mathbf{v}_i$ implies $\boldsymbol{\Phi}^{[k]}(\mathbf{v}_{i,n}) \rightharpoonup \boldsymbol{\Phi}^{[k]}(\mathbf{v}_i) = \hat{v}_i^{(k)}$ (cf. Lemma 6.23). By construction, $\boldsymbol{\Phi}^{[k]}(\mathbf{v}_{i,n}) = \alpha_n^{[k]} v_{i,n}^{(k)}$ holds with $\alpha_n^{[k]} := \prod_{j\neq k} \varphi^{(j)}(v_{i,n}^{(j)})$. Since $\|v_{n,i}^{(j)}\|_{V_j} \in [a,b]$ for some $0 < a \leq b < \infty$, also the sequence $\{\alpha_n^{[k]}\}$ is bounded. For $k = 1$, we extract a convergent subsequence such that $\alpha_n^{[1]} \to \alpha^{[1]}$ and hence $\alpha^{[1]} v_{i,n}^{(1)} \rightharpoonup \hat{v}_i^{(1)}$. Since $\hat{v}_i^{(1)} \neq 0$, $\alpha^{[1]}$ cannot vanish and allows to define $v_i^{(1)} := (1/\alpha^{[1]})\hat{v}_i^{(1)}$ as weak limit of $v_{i,n}^{(1)}$. Restricting the considerations to this subsequence, we proceed with the sequence $\{\alpha_n^{[2]}\}$ and derive the weak convergence $v_{i,n}^{(1)} \rightharpoonup v_i^{(2)} := (1/\alpha^{[2]})\hat{v}_i^{(2)}$, etc. The now defined $v_i^{(j)}$ satisfy statement (9.13). Note that $v_i^{(j)}$ and $\hat{v}_i^{(j)}$ differ only by an uninteresting scaling, since $\prod_{j=1}^{d} \alpha^{[j]} = 1$. □

9.4.4 A Greedy Algorithm

Finally, we hint to another important difference between the matrix case $d = 2$ and the true tensor case $d \geq 3$. Consider again Problem (9.1), where we want to find an approximation $\mathbf{u} \in \mathcal{R}_r$ of $\mathbf{v} \in \mathbf{V}$.

In principle, one can try to repeat the deflation method from §9.3:

1) determine the best approximation $\mathbf{u}_1 \in \mathcal{R}_1$ to $\mathbf{v} \in \mathbf{V}$ according to Lemma 9.3, set $\mathbf{v}_1 := \mathbf{v} - \mathbf{u}_1$,

2) determine the best approximation $\mathbf{u}_2 \in \mathcal{R}_1$ to $\mathbf{v}_1 \in \mathbf{V}$ according to Lemma 9.3, set $\mathbf{v}_2 := \mathbf{v}_1 - \mathbf{u}_2$,

$\vdots$

r) determine the best approximation $\mathbf{u}_r \in \mathcal{R}_1$ to $\mathbf{v}_{r-1} \in \mathbf{V}$.

Then $\hat{\mathbf{u}} := \mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_r \in \mathcal{R}_r$ can be considered as an approximation of $\mathbf{v} \in \mathbf{V}$. The described algorithm belongs to the class of '*greedy algorithms*', since in each single step one tries to reduce the error as good as possible.

Remark 9.18. (a) In the matrix case $d = 2$ with $\|\cdot\| = \|\cdot\|_{\mathsf{F}}$ (cf. §9.3), the algorithm from above yields the best approximation $\hat{\mathbf{u}} \in \mathcal{R}_r$ of $\mathbf{v} \in \mathbf{V}$, i.e., $\hat{\mathbf{u}}$ solves (9.1).
(b) In the true tensor case $d \geq 3$, the resulting $\hat{\mathbf{u}} \in \mathcal{R}_r$ is, in general, a rather poor approximation, i.e., $\|\mathbf{v} - \hat{\mathbf{u}}\|$ is much larger than $\inf_{\mathbf{u}\in\mathcal{R}_r} \|\mathbf{v} - \mathbf{u}\|$.

Proof. 1) The matrix case is discussed in Remark 9.8.

2) If $\inf_{\mathbf{u}\in\mathcal{R}_r} \|\mathbf{v} - \mathbf{u}\|$ has no minimiser, $\hat{\mathbf{u}}$ cannot be a solution of Problem (9.1). But even if $\mathbf{v} \in \mathcal{R}_r$ so that $\mathbf{u} := \mathbf{v}$ is the unique minimiser, practical examples (see below) show that the greedy algorithm yields a poor approximation $\hat{\mathbf{u}}$. □

As an example, we choose the tensor space $\mathbf{V} = \mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$ and the tensor

$$\mathbf{v} = \begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\2\end{bmatrix} \otimes \begin{bmatrix}3\\2\end{bmatrix} \otimes \begin{bmatrix}3\\1\end{bmatrix} \in \mathcal{R}_2 \tag{9.14}$$

with Euclidean norm $\|\mathbf{v}\| = \sqrt{2078} \approx 45.585$.

The best approximation $\mathbf{u}_1 \in \mathcal{R}_1$ of $\mathbf{v} \in \mathbf{V}$ is

$$\mathbf{u}_1 = 27.14270606 \cdot \begin{bmatrix}1\\0.7613363832\end{bmatrix} \otimes \begin{bmatrix}1\\0.7613363836\end{bmatrix} \otimes \begin{bmatrix}1\\0.3959752430\end{bmatrix}.$$

The approximation error is $\|\mathbf{v} - \mathbf{u}_1\| = 2.334461003$.

In the second step, the best approximation $\mathbf{u}_2 \in \mathcal{R}_1$ of $\mathbf{v} - \mathbf{u}_1$ turns out to be

$$\mathbf{u}_2 = 0.03403966791 \cdot \begin{bmatrix}1\\-4.86171875\end{bmatrix} \otimes \begin{bmatrix}1\\-3.91015625\end{bmatrix} \otimes \begin{bmatrix}1\\2.469921875\end{bmatrix}.$$

It yields the approximation error $\|\mathbf{v} - (\mathbf{u}_1 + \mathbf{u}_2)\| = 1.465604638$, whereas the best approximation in $\mathcal{R}_2$ is $\mathbf{u} = \mathbf{v}$ with vanishing error.

The reason why the algorithm fails to find the best approximation, becomes obvious from the first correction step. The correction $\mathbf{u}_1$ is close to the second term in (9.14), but not equal. Therefore $\mathbf{v} - \mathbf{u}_1$ belongs to $\mathcal{R}_3$ and it is impossible to reach the best approximation in the second step.

Nevertheless, the algorithm can be used as an iteration (cf. §17.1). Concerning convergence compare [47] and [58].

9.5 Numerical Approaches for the r-Term Approximation

If $\inf_{\mathbf{u}\in\mathcal{R}_r} \|\mathbf{v} - \mathbf{u}\|$ has no minimum, any numerical method is in trouble. First, no computed sequence can converge, and second, the instability will spoil the computation. On the other hand, if $\min_{\mathbf{u}\in\mathcal{R}_r} \|\mathbf{v} - \mathbf{u}\|$ exists and moreover essential uniqueness holds (see Remark 7.4b), there is hope for a successful numerical treatment.

9.5.1 Use of the Hybrid Format

When $\mathbf{v} \in \mathbf{V}$ is to be approximated by $\mathbf{u} \in \mathcal{R}_r$, the computational cost will depend on the representation of $\mathbf{v}$. The cost is huge, if $\mathbf{v}$ is represented in full format. Fortunately, in most of the applications, $\mathbf{v}$ is already given in R-term format with some $R > r$ or, possibly, in a tensor subspace format. We start with the latter case.

We consider the case of $V_j = \mathbb{K}^{I_j}$ and $\mathbf{V} = \mathbb{K}^{\mathbf{I}}$, $\mathbf{I} = I_1 \times \ldots \times I_d$, equipped with the Euclidean norm $\|\cdot\|$. Assume that $\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$ with orthogonal matrices $B_j \in \mathbb{K}^{I_j \times J_j}$ and $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$, $\mathbf{J} = J_1 \times \ldots \times J_d$. Also the Euclidean norm in $\mathbb{K}^{\mathbf{J}}$ is denoted by $\|\cdot\|$. Set $\mathbf{B} := \bigotimes_{j=1}^d B_j$ and note that $\mathbf{v} = \mathbf{B}\mathbf{a}$ (cf. (8.6b)).

Lemma 9.19. *Let* $\mathbf{v}$, $\mathbf{a}$, $\mathbf{B}$ *be as above. Any* $\mathbf{c} \in \mathbb{K}^{\mathbf{J}}$ *together with* $\mathbf{u} := \mathbf{Bc} \in \mathbb{K}^{\mathbf{I}}$ *satisfies* $\|\mathbf{a} - \mathbf{c}\| = \|\mathbf{v} - \mathbf{u}\|$. *Furthermore,* $\mathbf{c}$ *and* $\mathbf{u}$ *have equal tensor rank. Minimisation of* $\|\mathbf{a} - \mathbf{c}\|$ *over all* $\mathbf{c} \in \mathcal{R}_r(\mathbb{K}^{\mathbf{J}})$ *is equivalent to minimisation of* $\|\mathbf{v} - \mathbf{u}\|$ *over all* $\mathbf{u} \in \mathcal{R}_r(\mathbb{K}^{\mathbf{I}})$.

Proof. The coincidence of the norm holds because orthonormal bases are used: $\mathbf{B}^{\mathsf{H}}\mathbf{B} = \mathbf{I}$. Theorem 8.36a states that $\operatorname{rank}(\mathbf{c}) = \operatorname{rank}(\mathbf{u})$. □

Therefore, the strategy consists of three steps:

(i) Given $\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big) \in \mathbb{K}^{\mathbf{I}}$, focus to $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$,
(ii) approximate $\mathbf{a}$ by some $\mathbf{c} \in \mathcal{R}_r(\mathbb{K}^{\mathbf{J}})$,
(iii) define $\mathbf{u} := \mathbf{Bc} \in \mathcal{R}_r(\mathbb{K}^{\mathbf{I}})$ as approximant of $\mathbf{v}$.

This approach is of practical relevance, since $\#\mathbf{J} \le \#\mathbf{I}$ holds and often $\#\mathbf{J} \ll \#\mathbf{I}$ is expected. The resulting tensor $\mathbf{u} = \mathbf{Bc}$ has hybrid format $\mathbf{v} = \rho_{\mathrm{orth}}^{\mathrm{hybr}}(\ldots)$ (cf. (8.20)). This approach is, e.g., recommended in Espig [52].

The previous step (ii) depends again on the format of $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$. In the general case of $\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$, the coefficient tensor $\mathbf{a}$ is given in full format. A more favourable case is the hybrid format $\mathbf{v} = \rho_{\mathrm{orth}}^{\mathrm{hybr}}(\ldots)$, where $\mathbf{a}$ is given in R-term format.

Next, we assume that $\mathbf{v} \in \mathbf{V}$ is given in R-term format with a possibly large representation rank R, which is to be reduced to $r \le R$ (either r fixed, or indirectly determined by a prescribed accuracy). §8.5.2.3 describes the conversion of the tensor $\mathbf{v} = \rho_{\text{r-term}}(R, \ldots)$ into the hybrid format $\mathbf{v} = \rho_{\mathrm{orth}}^{\mathrm{hybr}}(\ldots) = \mathbf{Ba}$, i.e., with $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$ given again in the R-term format $\mathbf{a} = \rho_{\text{r-term}}(R, \ldots)$. According to Lemma 9.19, the approximation is applied to the coefficient tensor $\mathbf{a}$.

We summarise the reduction of the approximation problems for the various formats:

format of original tensor $\mathbf{v}$	format of coefficient tensor $\mathbf{a}$
$\rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$	full format
$\rho_{\mathrm{orth}}^{\mathrm{hybr}}\big(R, (a_\nu^{(j)}), (B_j)_{j=1}^d\big)$	R-term format
R-term format	R-term format

(9.15)

The equivalence of minimising $\mathbf{c}$ in $\|\mathbf{a} - \mathbf{c}\|$ and $\mathbf{u}$ in $\|\mathbf{v} - \mathbf{u}\|$ leads again to the statement that the minimiser $\mathbf{u}^*$ of $\min_{\mathbf{u}} \|\mathbf{v} - \mathbf{u}\|$ belongs to $\bigotimes_{j=1}^d U_j^{\min}(\mathbf{v})$ (cf. Lemma 9.2).

The hybrid format is also involved in the approach proposed by Khoromskij-Khoromskaja [120]. It applies in the case of a large representation rank r in $\mathbf{v} \in \mathcal{R}_r$ and $d = 3$, and consists of two steps:

Step 1: convert the tensor $\mathbf{v} \in \mathcal{R}_r$ approximately into an HOSVD representation $\mathbf{v}' = \rho_{\mathrm{HOSVD}}\big(\mathbf{a}, (B_j)_{1 \le j \le d}\big)$;

Step 2: exploit the sparsity pattern of $\mathbf{a}$ to reconvert to $\mathbf{v}'' \in \mathcal{R}_{r'}$ with hopefully much smaller $r' < r$.

For Step 1 one might use methods like described in Remark 8.31. Because of the HOSVD structure, the entries of the coefficient tensor $\mathbf{a}$ are not of equal size. In practical applications one observes that a large part of the entries can be dropped yielding a sparse tensor (cf. §7.6.5), although a theoretical guarantee cannot be given. A positive result about the sparsity of $\mathbf{a}$ can be stated for sparse grid bases instead of HOSVD bases (cf. §7.6.5).

9.5.2 Alternating Least-Squares Method

9.5.2.1 Alternating Methods in General

Assume that Φ is a real-valued function of variables $\mathbf{x} := (x_\omega)_{\omega\in\Omega}$ with an ordered index set Ω. We want to find a minimiser $\mathbf{x}^*$ of $\Phi(\mathbf{x}) = \Phi(x_{\omega_1}, x_{\omega_2}, \ldots)$. A standard iterative approach is the successive minimisation with respect to the single variables x_ω. The iteration starts with some $\mathbf{x}^{(0)}$. Each step of the iteration maps $\mathbf{x}^{(m-1)}$ into $\mathbf{x}^{(m)}$ and has the following form:

Start	choose $x_\omega^{(0)}$ for $\omega \in \Omega$.
Iteration $m = 1, 2, \ldots$	for $i := 1, \ldots, \#\Omega$ do $x_{\omega_i}^{(m)} :=$ minimiser of $\Phi(\ldots, x_{\omega_{i-1}}^{(m)}, \xi, x_{\omega_{i+1}}^{(m-1)}, \ldots)$ w.r.t. ξ (9.16)

Note that in the last line the variables $x_{\omega_\ell}^{(m-1)}$ for $\ell > i$ are taken from the last iterate $\mathbf{x}^{(m-1)}$, while for $\ell < i$ the new values are inserted.

The underlying assumption is that minimisation with respect to a single variable is much easier and cheaper than minimisation with respect to all variables simultaneously. The form of the iteration is well-known from the Gauss-Seidel method (cf. [81, §4.2.2]). Obviously, the value $\Phi(\mathbf{x}^{(m)})$ is weakly decreasing during the computation. Whether the iterates converge depends on properties of Φ and on the initial value. In case $\mathbf{x}^{(m)}$ converges, the limit may be a local minimum.

Next, we mention some variations of the general method.

(α) Minimisation may be replaced by maximisation.

(β) Using $i := 1, 2, \ldots, \#\Omega - 1, \#\Omega, \#\Omega - 1, \ldots, 2$ as i loop in (9.16), we ensure a certain symmetry. Prototype is the symmetric Gauss-Seidel iteration (cf. [81, §4.8.3]).

(γ) Instead of single variables, one may use groups of variables, e.g., minimise first with respect to (x_1, x_2), then with respect to (x_3, x_4), etc. After rewriting $(X_1 := (x_1, x_2), \ldots)$ we get the same setting as in (9.16). Since we have not fixed the format of x_j, each variable x_j may be vector-valued. The corresponding variant of the Gauss-Seidel iteration is called block-Gauss-Seidel (cf. [81, §4.5.2]).

(δ) The groups of variables may overlap, e.g., minimise first with respect to (x_1, x_2), then with respect to (x_2, x_3), etc.

(ε) Usually, we do not determine the *exact* minimiser, as required in (9.16). Since the method is iterative anyway, there is no need for an exact minimisation. The weak decrease of $\Phi(\mathbf{x}^{(m)})$ can still be ensured.

(ζ) Previous iterates can be used to form some nonlinear analogues of the cg, Krylov, or GMRES methods.

For any $1 \le k \le p$, let $\Phi(x_1, \ldots, x_p)$ with fixed x_j ($j \neq k$) be a quadratic function[7] in x_k. Then the minimisation in (9.16) is a least-squares problem, and algorithm (9.16) is called *alternating least-squares method*. This situation will typically arise, when Φ is a squared multilinear function.

9.5.2.2 ALS Algorithm for the r-Term Approximation

Let $\mathbf{V} \in \bigotimes_{j=1}^{d} V_j$ and $V_j = \mathbb{K}^{I_j}$ be equipped with the Euclidean norm[8] and set $\mathbf{I} := I_1 \times \ldots \times I_d$. For $\mathbf{v} \in \mathbf{V}$ and a representation rank $r \in \mathbb{N}_0$ we want to minimise[9]

$$\|\mathbf{v} - \mathbf{u}\|^2 = \left\| \mathbf{v} - \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} u_\nu^{(j)} \right\|^2 = \sum_{\mathbf{i} \in \mathbf{I}} \left| \mathbf{v}[\mathbf{i}] - \sum_{\nu=1}^{r} \prod_{j=1}^{d} u_\nu^{(j)}[i_j] \right|^2 \tag{9.17a}$$

with respect to all entries $u_\nu^{(j)}[i]$. To construct the *alternating least-squares method* (abbreviation: *ALS*), we identify the entries $u_\nu^{(k)}[i]$ with the variables x_ω from above. The indices are $\omega = (k, \nu, i) \in \Omega := \{1, \ldots, d\} \times \{1, \ldots, r\} \times I_k$. For this purpose, we introduce the notations $\mathbf{I}_{[k]} := \times_{j \neq k} I_j$ and $\mathbf{u}_\nu^{[k]} := \bigotimes_{j \neq k} u_\nu^{(j)}$ (cf. (3.21d)), so that $\mathbf{u} = \sum_{\nu=1}^{r} u_\nu^{(k)} \otimes \mathbf{u}_\nu^{[k]}$ and

$$\|\mathbf{v} - \mathbf{u}\|^2 = \sum_{i \in I_k} \sum_{\boldsymbol{\ell} \in \mathbf{I}_{[k]}} \left| \mathbf{v}[\ell_1, \ldots, \ell_{k-1}, i, \ell_{k+1}, \ldots, \ell_d] - \sum_{\nu=1}^{r} u_\nu^{(k)}[i] \cdot \mathbf{u}_\nu^{[k]}[\boldsymbol{\ell}] \right|^2.$$

For fixed $\omega = (k, \nu, i) \in \Omega$, this equation has the form

$$\|\mathbf{v} - \mathbf{u}\|^2 = \alpha_\omega |x_\omega|^2 - 2\,\mathfrak{Re}\,(\beta_\omega x_\omega) + \gamma_\omega \qquad \text{with} \tag{9.17b}$$
$$x_\omega = u_\nu^{(k)}[i], \;\; \alpha_\omega = \sum_{\boldsymbol{\ell} \in \mathbf{I}_{[k]}} \left| \mathbf{u}_\nu^{[k]}[\boldsymbol{\ell}] \right|^2, \;\; \beta_\omega = \sum_{\boldsymbol{\ell} \in \mathbf{I}_{[k]}} \mathbf{u}_\nu^{[k]}[\boldsymbol{\ell}]\, \overline{\mathbf{v}[\ldots, \ell_{k-1}, i, \ell_{k+1}, \ldots]}$$

and is minimised by $x_\omega = \beta_\omega / \alpha_\omega$. We add some comments:

[7] In the case of a Hilbert space V over $\mathbb{K} = \mathbb{C}$, 'quadratic function in $x_k \in V$' means $\Phi(x_k) = \langle Ax_k, x_k \rangle + \langle b, x_k \rangle + \langle x_k, b \rangle + c$, where $\langle \cdot, \cdot \rangle$ is the scalar product in V, $A = A^{\mathsf{H}}$, $b \in V$, and $c \in \mathbb{R}$.

[8] More generally, the norm $\|\cdot\|_j$ of V_j may be generated by any scalar product $\langle \cdot, \cdot \rangle_j$, provided that $\mathbf{V}$ is equipped with the induced scalar product.

[9] In the case of incomplete tensor data, the sum $\sum_{\mathbf{i} \in \mathbf{I}}$ has to be reduced to $\sum_{\mathbf{i} \in \mathring{\mathbf{I}}}$, where $\mathring{\mathbf{I}} \subset \mathbf{I}$ is the index set of the given entries.

1) For fixed k, ν, the minimisation with respect to $u_\nu^{(k)}[i]$, $i \in I_k$, is independent and can be performed in parallel. The coefficient α_ω from (9.17b) is independent of $i \in I_k$. These facts will be used in (9.18).

2) As a consequence of 1), the block version with $\Omega' := \{1, \ldots, d\} \times \{1, \ldots, r\}$ and $x_{\omega'} = u_\nu^{(k)} \in V_k$ ($\omega' = (k, \nu) \in \Omega'$) delivers identical results.

3) The starting value $\mathbf{u}^0$ should be chosen carefully. Obviously, $\mathbf{u}^0 = 0$ is not a good choice as $\alpha_\omega \neq 0$ is needed in (9.17b). Different starting values may lead to different minima of $\|\mathbf{v} - \mathbf{u}\|$ (unfortunately, also local minima different from $\min\|\mathbf{v} - \mathbf{u}\|$ must be expected).

4) Since $x_\omega = \beta_\omega / \alpha_\omega$, the computational cost is dominated by the evaluation of α_ω and β_ω for all $\omega \in \Omega$.

Rewriting the general scheme (9.16) for Problem (9.17a,b), we obtain the following algorithm. The iterate $x_\omega^{(m)}$ takes the form $u_\nu^{(j,m)} \in V_j$, where $\omega = (j, \nu) \in \Omega := \{1, \ldots, d\} \times \{1, \ldots, r\}$, since the components $u_\nu^{(j,m)}[i]$, $i \in I_k$, are determined in parallel.

Start	choose some starting values $u_\nu^{(j)} \in V_j$ for $1 \le j \le d$; for $j := 1, \ldots, d$ do for $\nu := 1, \ldots, r$ do $\tau_\nu^{(j)} := \Vert u_\nu^{(j)} \Vert_j^2$;

Iteration
$m = 1, 2, \ldots$
	for $k := 1, \ldots, d$ do for $\nu := 1, \ldots, r$ do
	begin $\alpha_\nu := \prod_{j \neq k}^d \tau_\nu^{(j)}$; $\quad u_\nu^{(k)} := \frac{1}{\alpha_\nu} \left\langle \bigotimes_{j \neq k} u_\nu^{(j)}, \mathbf{v} \right\rangle_{[k]}$;
	$\tau_\nu^{(k)} := \Vert u_\nu^{(k)} \Vert_k^2$
	end; $\{u_\nu^{(j)} = u_\nu^{(j,m)}$ are the coefficients of the m-th iterate$\}$

(9.18)

9.5.2.3 Computational Cost

Remark 9.20. According to §9.5.1, we should replace the tensors $\mathbf{v}, \mathbf{u} \in \mathbb{K}^{\mathbf{I}}$ ($\mathbf{u}$ approximant of $\mathbf{v}$) by their coefficient tensors $\mathbf{a}, \mathbf{c} \in \mathbb{K}^{\mathbf{J}}$. This does not change the algorithm; only when we discuss the computational cost, we have to replace $\#I_j$ by $\#J_j$.

The bilinear mapping $\langle \cdot, \cdot \rangle_{[k]} : (\bigotimes_{j \neq k} V_j) \times \mathbf{V} \to V_k$ from above is defined by

$$\langle \mathbf{w}, \mathbf{v} \rangle_{[k]} [i] = \sum_{\boldsymbol{\ell} \in \mathbf{I}_{[k]}} \mathbf{w}[\boldsymbol{\ell}] \, \overline{\mathbf{v}[\ell_1, \ldots, \ell_{k-1}, i, \ell_{k+1}, \ldots, \ell_d]} \quad \text{for all } i \in I_k. \tag{9.19}$$

The variables $\tau_\nu^{(j)}$ are introduced in (9.18) to show that only one norm $\|u_\nu^{(k)}\|_k$ is to be evaluated per (k, ν)-iteration.

So far, we have not specified how the input tensor $\mathbf{v}$ is represented. By Remark 9.20 and (9.15), the interesting formats are the full and R-term formats.

If the tensor $\mathbf{v}$ $[\mathbf{a}]$ is represented in full format, the partial scalar product $\left\langle \bigotimes_{j \neq k} u_\nu^{(j)}, \mathbf{v} \right\rangle_{[k]}$ takes $2 \prod_{j=1}^d \#I_j$ $\left[2 \prod_{j=1}^d \#J_j \right]$ operations (the quantities in brackets refer to the interpretation by Remark 9.20). All other operations are of lower order.

Remark 9.21. If $\mathbf{v} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$, ALS can be applied to the coefficient tensor $\mathbf{a}$ requiring $2\prod_{j=1}^d r_j$ operations per iteration ($r_j = \#J_j$).

The standard situation is that $\mathbf{v}$ is given in R-term representation, but with a large representation rank R which should be reduced to $r < R$. Let

$$\mathbf{v} = \sum_{\mu=1}^{R} \bigotimes_{j=1}^{d} v_\mu^{(j)}.$$

Then, $\langle \mathbf{w}, \mathbf{v}\rangle_{[k]} = \sum_{\mu=1}^R v_\mu^{(k)} \prod_{j\neq k} \sigma_{\nu\mu}^{(j)}$ holds with $\sigma_{\nu\mu}^{(j)} := \langle u_\nu^{(j)}, v_\mu^{(j)}\rangle_j$. The scalar products $\sigma_{\nu\mu}^{(j)}$ have to be evaluated for the starting values and as soon as new $u_\nu^{(j)}$ are computed. In total, $dr(R+1)$ scalar products or norm evaluations are involved per iteration. Assuming that the scalar product in $V_j = \mathbb{R}^{I_j}$ costs $2n_j - 1$ operations ($n_j := \#I_j$, cf. Remark 7.12), we conclude that the leading part of the cost is $2rR\sum_{j=1}^d n_j$. This cost has still to be multiplied by the number of iterations in (9.16). The cost described in the next remark uses the interpretation from Remark 9.20.

Remark 9.22. If the coefficient tensor of $\mathbf{v}$ is given in R-term format, the cost of one ALS iteration is

$$2rR\sum_{j=1}^{d} r_j \leq 2dr\bar{r}R \leq 2drR^2, \qquad \text{where } \bar{r} := \max_j r_j,\ r_j = \#J_j.$$

The use of Remark 9.20 requires as preprocessing the conversion of $\mathbf{v}$ from R-term into hybrid format (cost: $\sum_{j=1}^d N_{\mathrm{QR}}(n_j, R)$) and as postprocessing the multiplication $\mathbf{Bc}$ (cost: $2r\sum_{j=1}^d n_j r_j$).

The following table summarises the computational cost per iteration and possibly for pre- and postprocessing.

format of $\mathbf{v}$	cost per iteration	pre-, postprocessing
full	$2\prod_{j=1}^d n_j \quad (n_j := \#I_j)$	0
tensor subspace	$2\prod_{j=1}^d r_j \quad (r_j := \#J_j)$	$2r\sum_{j=1}^d n_j r_j$
R-term	$2rR\sum_{j=1}^d r_j$	$\sum_{j=1}^d N_{\mathrm{QR}}(n_j, R) + 2r\sum_{j=1}^d n_j r_j$
hybrid	$2rR\sum_{j=1}^d r_j$	$2r\sum_{j=1}^d n_j r_j$

9.5.2.4 Properties of the Iteration

Given an approximation $\mathbf{u}_m$, the loop over $\omega \in \Omega$ produces the next iterate $\mathbf{u}_{m+1}$. The obvious questions are whether the sequence $\{\mathbf{u}_m\}_{m=0,1,\ldots}$ converges, and in the positive case, whether it converges to $\mathbf{u}^*$ with $\|\mathbf{v}-\mathbf{u}^*\| = \min_{\mathbf{u}\in\mathcal{R}_r}\|\mathbf{v}-\mathbf{u}\|$.

Statements can be found in Mohlenkamp [149, §4.3]:

a) The sequence $\{\mathbf{u}_m\}$ is bounded,

b) $\|\mathbf{u}_m - \mathbf{u}_{m+1}\| \to 0$,

c) $\sum_{m=0}^{\infty} \|\mathbf{u}_m - \mathbf{u}_{m+1}\|^2 < \infty$,

d) the set of accumulation points of $\{\mathbf{u}_m\}$ is connected and compact.

The negative statements are: the properties from above do not imply convergence, and in the case of convergence, the limit $\mathbf{u}^*$ may be a local minimum with $\|\mathbf{v}-\mathbf{u}^*\| > \min_{\mathbf{u}\in\mathcal{R}_r} \|\mathbf{v}-\mathbf{u}\|$. A simple example for the latter fact is given in [149, §4.3.5].

Under suitable assumptions, which need not hold in general, local convergence is proved by Uschmajew [187].

9.5.3 Stabilised Approximation Problem

As seen in §9.4, the minimisation problem $\min_{\mathbf{u}\in\mathcal{R}_r} \|\mathbf{v}-\mathbf{u}\|$ is unsolvable, if and only if infimum sequences are unstable. An obvious remedy is to enforce stability by adding a penalty term:

$$\Phi_\lambda\left((u_i^{(j)})_{1\le i\le r}^{1\le j\le d}\right) := \min_{u_i^{(j)}\in V_j} \sqrt{\left\|\mathbf{v} - \sum_{i=1}^{r} \bigotimes_{j=1}^{d} u_i^{(j)}\right\|^2 + \lambda^2 \sum_{i=1}^{r} \left\|\bigotimes_{j=1}^{d} u_i^{(j)}\right\|^2}, \quad (9.20a)$$

where $\lambda > 0$ and $\|\bigotimes_{j=1}^{d} u_i^{(j)}\|^2 = \prod_{j=1}^{d} \|u_i^{(j)}\|^2$. Alternatively, stability may be requested as a side condition ($C > 0$):

$$\Phi_C\left((u_i^{(j)})_{1\le i\le r}^{1\le j\le d}\right) := \min_{\substack{u_i^{(j)}\in V_j \text{ subject to} \\ \sum_{i=1}^{r} \|\bigotimes_{j=1}^{d} u_i^{(j)}\|^2 \le C^2 \|\mathbf{v}\|^2}} \left\|\mathbf{v} - \sum_{i=1}^{r} \bigotimes_{j=1}^{d} u_i^{(j)}\right\|. \quad (9.20b)$$

If $\mathbf{u}_n = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} u_{i,n}^{(j)}$ is a sequence with $\|\mathbf{v}-\mathbf{u}_n\| \searrow \inf_{\mathbf{u}} \|\mathbf{v}-\mathbf{u}\|$ subject to the side condition from (9.20b), it is a stable sequence: $\varkappa((\mathbf{v}_n), r) \le C$. Hence, we infer from Theorem 9.17 that this subsequence converges to some $\mathbf{u}^* \in \mathcal{R}_r$.

In the penalty case of (9.20a), we may assume $\Phi_\lambda \le \|\mathbf{v}\|$, since already the trivial approximation $\mathbf{u} = 0$ ensures this estimate. Then $\varkappa((\mathbf{v}_n), r) \le \lambda$ follows and allows the same conclusion as above. Even for a general minimising sequence $\mathbf{u}_n = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} u_{i,n}^{(j)}$ with $c := \lim_n \Phi_\lambda\left((u_{i,n}^{(j)})_{1\le i\le r}^{1\le j\le d}\right)$, we conclude that $\varkappa((\mathbf{v}_n), r) \le c$ holds asymptotically.

If $\min_{\mathbf{u}\in\mathcal{R}_r} \|\mathbf{v}-\mathbf{u}\|$ possesses a stable minimising sequence with $\varkappa((\mathbf{u}_n), r) \le C$ and limit $\mathbf{u}^*$, the minimisation of Φ_C from (9.20b) yields the same result. Lemma 9.2 also holds for the solution $\mathbf{u}^*$ of the regularised solution.

9.5.4 Newton's Approach

The alternative to the successive minimisation is the *simultaneous* minimisation of $\varPhi(\mathbf{x})$ in all variables $\mathbf{x} = (u_{i,n}^{(j)})_{1\le i\le r}^{1\le j\le d}$. For this purpose, iterative methods can be applied like the *gradient method* or the *Newton method*. Both are of the form

$$\mathbf{x}^{(m+1)} := \mathbf{x}^{(m)} - \alpha_m\, s_m \qquad (s_m\text{: search direction, } \alpha_m \in \mathbb{K}).$$

The gradient method is characterised by $s_m = \nabla\varPhi(\mathbf{x}^{(m)})$, while Newton's method uses $s_m = H(\mathbf{x}^{(m)})^{-1}\,\nabla\varPhi(\mathbf{x}^{(m)})$ and $\alpha_m = 1$. Here, H is the matrix of the second partial derivatives: $H_{\omega\omega'} = \partial^2\varPhi/\partial\mathbf{x}_\omega\partial\mathbf{x}_{\omega'}$. However, there are a plenty of variations between both methods. The damped Newton method has a reduced parameter $0<\alpha_m<1$. The true Hessian H may be replaced by approximations $\tilde{H}$ which are easier to invert. For $\tilde{H}=I$, we regain the gradient method. Below we use a block diagonal part of H.

In Espig [52] and Espig-Hackbusch [54], a method is described which computes the minimiser of [10] $\varPhi_\lambda$ from (9.20a). It is a modified Newton method with an approximate Hessian matrix $\tilde{H}$ allowing for a continuous transition from the Newton to a gradient-type method. Although the Hessian H is a rather involved expression, its particular structure can be exploited when the system $H(\mathbf{x}^{(m)})s_m = \nabla\varPhi(\mathbf{x}^{(m)})$ is to be solved. This defines a procedure $\mathbf{RNM}(\mathbf{v},\mathbf{u})$ which determines the best approximation $\mathbf{u}\in\mathcal{R}_r$ of $\mathbf{v}\in\mathcal{R}_R$ by the stabilised Newton method (cf. [54, Alg. 1]). For details and numerical examples, we refer to [54]. The cost per iteration is

$$O\left(r(r+R)d^2 + dr^3 + r(r+R+d)\sum_{j=1}^{d} r_j\right)$$

with $r_j := \#J_j$ and J_j from Lemma 9.19.

In the following, we use the symbols $\mathbf{v},\mathbf{u}$ for the tensors involved in the optimisation problem. For the computation one should replace the tensors from $\mathbf{V}$ by the coefficient tensors in $\mathbb{K}^{\mathbf{J}}$ as detailed in §9.5.1. Newton's method is well-known for its fast convergence as soon as $\mathbf{x}^{(m)}$ is sufficiently close to a zero of $\nabla\varPhi(\mathbf{x}) = 0$. Usually, the main difficulty is the choice of suitable starting values. If a fixed rank is given (cf. Problem (9.1)), a rough initial guess can be constructed by the method described in Corollary 15.6.

A certain kind of nested iteration (cf. [83, §5], [81, §12.5]) can be exploited for solving Problem (9.3). The framework of the algorithm is as follows:

```
given data | $\mathbf{v}\in\mathcal{R}_R$, initial guess $\mathbf{u}\in\mathcal{R}_r$ with $r < R$, $\varepsilon > 0$
loop       | $\mathbf{RNM}(\mathbf{v},\mathbf{u})$;  $\boldsymbol{\rho} := \mathbf{v} - \mathbf{u}$; if $\|\boldsymbol{\rho}\| \le \varepsilon$ then return;
           | if $r = R$ then begin $\mathbf{u} := \mathbf{v}$; return end;
           | find a minimiser $\mathbf{w} \in \mathcal{R}_1$ of $\min_{\omega\in\mathcal{R}_1}\|\boldsymbol{\rho} - \boldsymbol{\omega}\|$ ;
           | $\mathbf{u} := \mathbf{u} + \mathbf{w} \in \mathcal{R}_{r+1}$;  $r := r + 1$; repeat the loop
```
(9.21)

[10] In fact, a further penalty term is added to enforce $\|u_i^{(j)}\| = \|u_i^{(k)}\|$ for $1 \le j,k \le d$.

Line 1: The initial guess $\mathbf{u} \in \mathcal{R}_r$ also defines the starting rank r.
Line 2: The best approximation $\mathbf{u} \in \mathcal{R}_r$ is accepted, if $\|\mathbf{v} - \mathbf{u}\| \leq \varepsilon$.
Line 3: If no approximation in $\mathcal{R}_r$ with $r < R$ is sufficiently accurate, $\mathbf{u} = \mathbf{v} \in \mathcal{R}_r$ must be returned.
Line 4: The best approximation problem in $\mathcal{R}_1$ can be solved by $\mathbf{RNM}$ or ALS. Here, no regularisation is needed (cf. §9.2).
Line 5: $\mathbf{u} + \mathbf{w}$ is the initial guess in $\mathcal{R}_{r+1}$.

Obviously, the $\mathcal{R}_1$ optimisation in Line 4 is of low cost compared with the other parts. This fact can be exploited to improve the initial guesses. Before calling $\mathbf{RNM}(\mathbf{v}, \mathbf{u})$ in Line 2, the following procedure can be applied. Here, $App_1(\mathbf{v}, \mathbf{w})$ is a rough $\mathcal{R}_1$ approximation of $\mathbf{v}$ using $\mathbf{w} \in \mathcal{R}_1$ as starting value (a very cheap method makes use of Remark 15.7):

data	$\mathbf{v} \in \mathcal{R}_R,\ \mathbf{u} = \sum_{i=1}^{r} \mathbf{u}_i \in \mathcal{R}_r,\ \mathbf{u}_i \in \mathcal{R}_1.$
loop	for $\nu = 1$ to r do begin $\mathbf{d} := \mathbf{u} - \sum_{i \neq \nu} \mathbf{u}_i;\ \mathbf{u}_\nu := App_1(\mathbf{d}, \mathbf{u}_\nu)$ end;

This improvement of the approximation $\mathbf{u}$ can be applied in Line 2 of (9.21) before calling $\mathbf{RNM}(\mathbf{v}, \mathbf{u})$. Details are given in [54].

9.6 Generalisations

Here we refer to §7.7, where subsets $A_j \subset V_j$ and $\mathcal{R}_r\big((A_j)_{j=1}^d\big) \subset \mathcal{R}_r$ have been introduced. The corresponding approximation problem is:

$$\begin{aligned}&\text{Given } \mathbf{v} \in \mathbf{V} \text{ and } r \in \mathbb{N}_0, \\ &\text{determine } \mathbf{u} \in \mathcal{R}_r\big((A_j)_{j=1}^d\big) \text{ minimising } \|\mathbf{v} - \mathbf{u}\|.\end{aligned} \tag{9.22}$$

Though the practical computation of the minimiser may be rather involved, the theoretical aspects can be simpler than in the standard case.

Lemma 9.23. *Assume that $\mathbf{V}$ is either finite dimensional or a reflexive Banach space. Let A_j be weakly closed subsets of V_j $(1 \leq j \leq d)$. If there is a stable subsequence $\mathbf{u}_n \in \mathcal{R}_r\big((A_j)_{j=1}^d\big)$ with $\lim_{n\to\infty} \|\mathbf{v} - \mathbf{u}_n\| = \inf_{\mathbf{u} \in \mathcal{R}_r((A_j)_{j=1}^d)} \|\mathbf{v} - \mathbf{u}\|$, then Problem (9.22) is solvable.*

Proof. By Theorem 9.17, there is a subsequence such that $\mathbf{u}_n = \sum_{i=1}^r \bigotimes_{j=1}^d u_{i,n}^{(j)} \rightharpoonup \mathbf{u} \in \mathcal{R}_r$ with $\mathbf{u} = \sum_{i=1}^r \bigotimes_{j=1}^d u_i^{(j)}$ and $u_{i,n}^{(j)} \rightharpoonup u_i^{(j)}$ satisfying $\|\mathbf{v} - \mathbf{u}\| = \inf_{\mathbf{w} \in \mathcal{R}_r((A_j)_{j=1}^d)} \|\mathbf{v} - \mathbf{w}\|$. Since $u_{i,n}^{(j)} \in A_j$ and A_j is weakly closed, $u_i^{(j)} \in A_j$ follows, proving $\mathbf{u} \in \mathcal{R}_r\big((A_j)_{j=1}^d\big)$. □

In §7.7, the first two examples of A_j are the subset $\{v \in V_j : v \geq 0\}$ of non-negative vectors ($V_j = \mathbb{R}^{n_j}$) or functions ($V_j = L^p$). Standard norms like the ℓ^p norm (cf. (4.3)) have the property

$$\|v + w\|_{V_j} \geq \|v\|_{V_j} \qquad \text{for all } v, w \in A_j \ (1 \leq j \leq d). \tag{9.23a}$$

Furthermore, these examples satisfy $A_j + A_j \subset A_j$, i.e.,

$$v, w \in A_j \Rightarrow v + w \in A_j \qquad (1 \leq j \leq d)\,. \tag{9.23b}$$

Remark 9.24. Conditions (9.23a,b) imply the stability estimate $\varkappa(\mathbf{v}, r) \leq r$, provided that in definition (9.12b) the vectors $v_i^{(j)}$ are restricted to A_j. Hence, any sequence $\mathbf{v}_n \in \mathcal{R}_r\big((A_j)_{j=1}^d\big)$ is stable and Lemma 9.23 can be applied.

For matrix spaces $V_j = \mathbb{C}^{n_j \times n_j}$ equipped with the spectral or Frobenius norm, $A_j = \{M \in V_j : M \text{ positive semidefinite}\}$ also satisfies conditions (9.23a,b). The set $A_j = \{M \in V_j : M \text{ Hermitean}\}$ is a negative example for (9.23a,b). Indeed, (9.10) with $v^{(j)}, w^{(j)} \in A_j$ is an example for an unstable sequence.

The subset $A_j = \{M \in V_j : M \text{ positive definite}\}$ is not closed, hence the minimiser of Problem (9.22) is expected in $\mathcal{R}_r\big((\overline{A_j})_{j=1}^d\big)$ instead of $\mathcal{R}_r\big((A_j)_{j=1}^d\big)$. Nevertheless, the following problem has a minimiser in $\mathcal{R}_1\big((A_j)_{j=1}^d\big)$.

Exercise 9.25. For $V_j = \mathbb{C}^{n_j \times n_j}$ equipped with the spectral or Frobenius norm, $\mathbf{M} \in \bigotimes_{j=1}^d V_j$ positive definite and $A_j = \{M \in V_j : M \text{ positive definite}\}$, $\inf\left\{\left\|\mathbf{M} - \bigotimes\limits_{j=1}^{d} M^{(j)}\right\| : M^{(j)} \in A_j\right\}$ is attained by some $\bigotimes\limits_{j=1}^{d} M^{(j)}$ with $M^{(j)} \in A_j$.

9.7 Analytical Approaches for the r-Term Approximation

The previous approximation methods are black box-like techniques which are applicable for any tensor. On the other side, for very particular tensors (e.g., multivariate functions) there are special analytical tools which yield an r-term approximation. Differently from the approaches above, the approximation error can be described in dependence on the parameter r. Often, the error is estimated with respect to the supremum norm $\|\cdot\|_\infty$, whereas the standard norm[11] considered above is ℓ^2 or L^2.

Analytical approaches will also be considered for the approximation in tensor subspace format. Since $\mathcal{R}_r = \mathcal{T}_{(r,r)}$ for dimension $d = 2$, these approaches from §10.4 can also be interesting for the r-term format.

Note that analytically derived approximations can serve two different purposes:

1. *Constructive approximation.* Most of the following techniques are suited for practical use. Such applications are described, e.g., in §9.7.2.5 and §9.7.2.6.
2. *Theoretical complexity estimates.* A fundamental question concerning the use of the formats $\mathcal{R}_r$ or $\mathcal{T}_\mathbf{r}$ is, how the best approximation error $\varepsilon(\mathbf{v}, r)$ from (9.2) depends on r. Any explicit error estimate of a particular (analytical) approximation yields an upper bound of $\varepsilon(\mathbf{v}, r)$. Under the conditions of this section, we shall obtain exponential convergence, i.e., $\varepsilon(\mathbf{v}, r) \leq O(\exp(-cr^\alpha)$ with $c, \alpha > 0$ is valid for the considered tensors $\mathbf{v}$.

[11] The optimisation problems from §9.5.2 and §9.5.4 can also be formulated for the ℓ^p norm with large, even p, which, however, would not make the task easier.

Objects of approximation are not only tensors of vector type, but also matrices described by Kronecker products. Early papers of such kind are [92], [88, 89].

9.7.1 Quadrature

Let V_j be Banach spaces of functions defined on $I_j \subset \mathbb{R}$, and $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ the space of multivariate functions on $I := \times_{j=1}^d I_j$. Assume that $\mathbf{f} \in \mathbf{V}$ has an integral representation

$$\mathbf{f}(x_1, \ldots, x_d) = \int_\Omega g(\omega) \prod_{j=1}^d f_j(x_j, \omega)\, \mathrm{d}\omega \qquad \text{for } x_j \in I_j, \tag{9.24}$$

where Ω is some parameter domain, such that the functions f_j are defined on $I_j \times \Omega$. For fixed ω, the integrand is an elementary tensor $\bigotimes_{j=1}^d f_j(\cdot, \omega) \in \mathbf{V}$. Since the integral $\int_\Omega$ is a limit of Riemann sums $\sum_{i=1}^r \ldots \in \mathcal{R}_r$, $\mathbf{f}$ is a topological tensor. A particular example of the right-hand side in (9.24) is the Fourier integral transform of $g(\omega) = g(\omega_1, \ldots, \omega_d)$:

$$\int_{\mathbb{R}^d} g(\omega) \exp\left(\mathrm{i} \sum_{j=1}^d x_j \omega_j\right) \mathrm{d}\omega.$$

A quadrature method for $\int_\Omega G(\omega)\, \mathrm{d}\omega$ is characterised by a sum $\sum_{i=1}^r \gamma_i G(\omega_i)$ with quadrature weights $(\gamma_i)_{i=1}^r$ and quadrature points $(\omega_i)_{i=1}^r$. Applying such a quadrature method to (9.24), we get the r-term approximation

$$\mathbf{f}_r \in \mathcal{R}_r \quad \text{with } \mathbf{f}_r(x_1, \ldots, x_d) := \sum_{i=1}^r \gamma_i\, g(\omega_i) \prod_{j=1}^d f_j(x_j, \omega_i). \tag{9.25}$$

Usually, there is a family of quadrature rules for all $r \in \mathbb{N}$, which leads to a sequence $(\mathbf{f}_r)_{r \in \mathbb{N}}$ of approximations. Under suitable smoothness conditions on the integrand of (9.24), one may try to derive error estimates of $\|\mathbf{f} - \mathbf{f}_r\|$. An interesting question concerns the (asymptotic) convergence speed $\|\mathbf{f} - \mathbf{f}_r\| \to 0$.

There is a connection to §9.6 and the subset A_j of non-negative functions. Assume that the integrand in (9.24) is non-negative. Many quadrature method (like the Gauss quadrature) have positive weights: $\gamma_i > 0$. Under this condition, also the terms in (9.25) are non-negative, i.e., $\mathbf{f}_r \in \mathcal{R}_r\big((A_j)_{j=1}^d\big)$.

So far, only the general setting is described. The concrete example of the sinc quadrature will follow in §9.7.2.2.

The described technique is not restricted to standard functions. Many of the tensors of finite dimensional tensor spaces can be considered as grid functions, i.e., as functions with arguments $x_1, \ldots, x_d$ restricted to a grid $\times_{j=1}^d G_j$, $\#G_j < \infty$. This fact does not influence the approach. If the error $\|\mathbf{f} - \mathbf{f}_r\|$ is the supremum norm of the associated function space, the restriction of the function to a grid is bounded by the same quantity.

9.7.2 Approximation by Exponential Sums

Below we shall focus to the (best) approximation with respect to the supremum norm $\|\cdot\|_\infty$. Optimisation with respect to the ℓ^2 norm is, e.g., considered by Golub-Pereyra [68]. However, in Proposition 9.31 $\|\cdot\|_\infty$ estimates will be needed, while ℓ^2 norm estimates are insufficient.

9.7.2.1 General Setting

For scalar-valued functions defined on a set D, we denote the supremum norm by

$$\|f\|_{D,\infty} := \sup\{|f(x)| : x \in D\}. \tag{9.26}$$

If the reference to D is obvious from the context, we also write $\|\cdot\|_\infty$ instead.

Exponential sums are of the form

$$E_r(t) = \sum_{\nu=1}^{r} a_\nu \exp(-\alpha_\nu t) \qquad (t \in \mathbb{R}) \tag{9.27a}$$

with $2r$ (real or complex) parameters a_ν and α_ν. Exponential sums are a tool to approximate certain univariate functions (details about their computation in §9.7.2 and §9.7.2.3).

Assume that a univariate function f in an interval $I \subset \mathbb{R}$ is approximated by some exponential sum E_r with respect to the supremum norm in I:

$$\|f - E_r\|_{I,\infty} \leq \varepsilon \tag{9.27b}$$

(we expect an exponential decay of $\varepsilon = \varepsilon_r$ with respect to r; cf. Theorem 9.29). Then the multivariate function

$$F(\mathbf{x}) = F(x_1, \ldots, x_d) := f\left(\sum_{j=1}^{d} \phi_j(x_j)\right) \tag{9.27c}$$

obtained by the substitution $t = \sum_{j=1}^{d} \phi_j(x_j)$, is approximated equally well by $F_r(\mathbf{x}) := E_r\left(\sum_{j=1}^{d} \phi_j(x_j)\right)$:

$$\|F - F_r\|_{\mathbf{I},\infty} \leq \varepsilon \qquad \text{for } \mathbf{I} := \mathop{\times}_{i=1}^{d} I_j, \tag{9.27d}$$

provided that

$$\left\{\sum_{j=1}^{d} \phi_j(x_j) : x_j \in I_j\right\} \subset I \qquad \text{with } I \text{ from (9.27b)}. \tag{9.27e}$$

For instance, condition (9.27e) holds for $\phi_j(x_j) = x_j$ and $I_j = I = [0, \infty)$.

By the property of the exponential function, we have

$$F_r(\mathbf{x}) := E_r\left(\sum_{j=1}^{d}\phi_j(x_j)\right) = \sum_{\nu=1}^{r} a_\nu \exp\left(-\alpha_\nu \sum_{j=1}^{d}\phi_j(x_j)\right) \tag{9.27f}$$
$$= \sum_{\nu=1}^{r} a_\nu \prod_{j=1}^{d} \exp\left(-\alpha_\nu \phi_j(x_j)\right).$$

Expressing the multivariate function E_r as a tensor product of univariate functions, we arrive at

$$F_r = \sum_{\nu=1}^{r} a_\nu \bigotimes_{j=1}^{d} E_\nu^{(j)} \in \mathcal{R}_r \qquad \text{with } E_\nu^{(j)}(x_j) := \exp\left(-\alpha_\nu \phi_j(x_j)\right), \tag{9.27g}$$

i.e., (9.27g) is an r-term representation of the tensor $F_r \in C(\mathbf{I}) = {}_\infty\bigotimes_{j=1}^{d} C(I_j)$, where the left suffix ∞ indicates the completion with respect to the supremum norm in $\mathbf{I} \subset \mathbb{R}^d$. A simple, but important observation is the following conclusion, which shows that the analysis of the univariate function f and its approximation by E_r is sufficient.

Conclusion 9.26. *The multivariate function $F_r(\mathbf{x})$ has tensor rank r independently of the dimension d. Also the approximation error (9.27d) is independent of the dimension d, provided that (9.27b) and (9.27e) are valid.*

Approximations by sums of Gaussians, $G_r(\xi) = \sum_{\nu=1}^{r} a_\nu \mathrm{e}^{-\alpha_\nu \xi^2}$, are equivalent to the previous exponential sums via $E_r(t) := G_r(\sqrt{t}) = \sum_{\nu=1}^{r} a_\nu \mathrm{e}^{-\alpha_\nu t}$.

A particular, but important substitution of the form considered in (9.27c) is $t = \|\mathbf{x}\|^2$ leading to

$$F_r(\mathbf{x}) := E_r\left(\sqrt{\sum\nolimits_{j=1}^{d} x_j^2}\right) = \sum_{\nu=1}^{r} a_\nu \exp\left(-\alpha_\nu \sum_{j=1}^{d} x_j^2\right) = \sum_{\nu=1}^{r} a_\nu \prod_{j=1}^{d} \mathrm{e}^{-\alpha_\nu x_j^2},$$
$$\text{i.e.,} \quad F_r = \sum_{\nu=1}^{r} a_\nu \bigotimes_{j=1}^{d} G_\nu^{(j)} \qquad \text{with } G_\nu^{(j)}(x_j) := \exp\left(-\alpha_\nu x_j^2\right). \tag{9.28}$$

Inequality (9.27b) implies

$$\|F - F_r\|_{D,\infty} \le \varepsilon \qquad \text{with } D := \left\{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \in I\right\}.$$

Remark 9.27. In the applications from above we make use of the fact that estimates with respect to the supremum norm are invariant under substitutions. When we consider an L^p norm $(1 \le p < \infty)$ instead of the supremum norm, the relation between the one-dimensional error bound (9.27b) and the multi-dimensional one in (9.27d) is more involved and depends on d.

9.7.2.2 Quadrature Based Exponential Sum Approximations

Approximations by exponential sums may be based on quadrature methods[12]. Assume that a function f with domain $I \subset \mathbb{R}$ is defined by the Laplace transform

$$f(x) = \int_0^\infty \mathrm{e}^{-tx} g(t) \mathrm{d}t \qquad \text{for } x \in I.$$

Any quadrature method $Q(F) := \sum_{\nu=1}^{r} \omega_\nu F(t_\nu)$ for a suitable integrand F defined on $[0, \infty)$ yields an exponential sum of the form (9.27a):

$$f(x) \approx Q(\mathrm{e}^{-\bullet x} g) := \sum_{\nu=1}^{r} \underbrace{\omega_\nu g(t_\nu)}_{=: a_\nu} \mathrm{e}^{-t_\nu x} \in \mathcal{R}_r.$$

Note that the quadrature error $f(x) - Q(\mathrm{e}^{-\bullet x} g)$ is to be controlled for all parameter values $x \in I$.

A possible choice of Q is the *sinc quadrature*. For this purpose one chooses a suitable substitution $t = \varphi(\tau)$ with $\varphi : \mathbb{R} \to [0, \infty)$ to obtain

$$f(x) = \int_{-\infty}^{\infty} \mathrm{e}^{-\varphi(\tau)x} g(\varphi(\tau)) \, \varphi'(\tau) \, \mathrm{d}\tau.$$

The sinc quadrature can be applied to analytic functions defined on $\mathbb{R}$:

$$\int_{-\infty}^{\infty} F(x) dx \approx T(F, h) := h \sum_{k=-\infty}^{\infty} F(kh) \approx T_N(F, h) := h \sum_{k=-N}^{N} F(kh).$$

$T(F, h)$ can be interpreted as the infinite trapezoidal rule with step size h, while $T_N(F, h)$ is a truncated finite sum. In fact, $T(F, h)$ and $T_N(F, h)$ are interpolatory quadratures, i.e., they are exact integrals $\int_{\mathbb{R}} C(f, h)(t) \mathrm{d}t$ and $\int_{\mathbb{R}} C_N(f, h)(t) \mathrm{d}t$ involving the sinc interpolations $C(f, h)$ and $C_N(f, h)$ defined in (10.37a,b).

The error analysis of $T(F, h)$ depends on the behaviour of the holomorphic function $F(z)$ in the complex strip $\mathfrak{D}_\delta$ defined in (10.38) and the norm (10.39). A typical error bound is of the form $C_1 \exp(-\sqrt{2\pi\delta\alpha N})$ with $C_1 = C_1(\|F\|_{\mathfrak{D}_\delta})$ and δ from (10.38), while α describes the decay of F: $|F(x)| \le O(\exp(-\alpha\,|x|)$. For a precise analysis see Stenger [177] and Hackbusch [86, §D.4]. Sinc quadrature applied to $F(t) = F(t; x) := \mathrm{e}^{-\varphi(t)x} g(\varphi(t)) \, \varphi'(t)$ yields

$$T_N(F, h) := h \sum_{k=-N}^{N} \mathrm{e}^{-\varphi(kh)x} g(\varphi(kh)) \, \varphi'(kh).$$

The right-hand side is an exponential sum (9.27a) with $r := 2N + 1$ and coefficients $a_\nu := h\, g\big(\varphi((\nu - 1 - N)\, h)\big)\, \varphi'\big(\,(\nu - 1 - N)\, h\big)$, $\alpha_\nu := \varphi\big(\,(\nu - 1 - N)\, h\big)$. Since

[12] Quadrature based approximation are very common in computational quantum chemistry. For a discussion from the mathematical side compare Beylkin-Monzón [18].

the integrand $F(\bullet; x)$ depends on the parameter $x \in I$, the error analysis must be performed *uniformly* in $x \in I$ to prove an estimate (9.27b): $\|f - E_r\|_{I,\infty} \le \varepsilon$.

Even if the obtainable error bounds possess an almost optimal asymptotic behaviour, they are inferior to the best approximations discussed next.

9.7.2.3 Approximation of $1/x$ and $1/\sqrt{x}$

Negative powers $x^{-\lambda}$ belong to the class of those functions which can be well approximated by exponential sums in $(0, \infty)$. Because of their importance, we shall consider the particular functions $1/x$ and $1/\sqrt{x}$. For the general theory of approximation by exponentials we refer to Braess [25]. The first statement concerns the existence of a best approximation and stability of the approximation expressed by positivity of its terms.

Theorem 9.28 ([25, p. 194]). *Given the function $f(x) = x^{-\lambda}$ with $\lambda > 0$ in an interval $I = [a, b]$ (including $b = \infty$) with $a > 0$, and $r \in \mathbb{N}$, there is a unique best approximation $E_{r,I}(x) = \sum_{\nu=1}^{r} a_{\nu,I} \exp(-\alpha_{\nu,I} x)$ such that*

$$\varepsilon(f, I, r) := \|f - E_{r,I}\|_{I,\infty} = \inf\left\{\left\| f - \sum_{\nu=1}^{r} b_\nu \mathrm{e}^{-\beta_\nu x}\right\|_{I,\infty} : b_\nu, \beta_\nu \in \mathbb{R}\right\}. \quad (9.29)$$

Moreover, this $E_{r,I}$ has positive coefficients: $a_\nu, \alpha_\nu > 0$ for $1 \le \nu \le r$.

In the case of $f(x) = 1/x$, substitution $x = at$ $(1 \le t \le b/a)$ shows that the best approximation for $I = [a, b]$ can be derived from the best approximation in $[1, b/a]$ via the transform

$$a_{\nu,[a,b]} := \frac{a_{\nu,[1,b/a]}}{a}, \quad \alpha_{\nu,[a,b]} := \frac{\alpha_{\nu,[1,b/a]}}{a}, \quad \varepsilon(f, [a, b], r) = \frac{\varepsilon(f, [1, b/a], r)}{a}. \quad (9.30a)$$

In the case of $f(x) = 1/\sqrt{x}$, the relations are

$$a_{\nu,[a,b]} = \frac{a_{\nu,[1,b/a]}}{\sqrt{a}}, \quad \alpha_{\nu,[a,b]} = \frac{\alpha_{\nu,[1,b/a]}}{a}, \quad \varepsilon(f, [a, b], r) = \frac{\varepsilon(f, [1, b/a], r)}{\sqrt{a}}. \quad (9.30b)$$

Therefore, it suffices to study the best approximation on standardised intervals $[1, R]$ for $R \in (1, \infty)$. The reference [84] points to a web page containing the coefficients $\{a_\nu, \alpha_\nu : 1 \le \nu \le r\}$ for various values of R and r.

Concerning convergence, we first consider a fixed interval $[1, R] = [1, 10]$. The error $\|1/x - E_{r,[1,10]}\|_{[1,10],\infty}$ is shown below:

$r = 1$	2	3	4	5	6	7
8.556_{10}-2	8.752_{10}-3	7.145_{10}-4	5.577_{10}-5	4.243_{10}-6	3.173_{10}-7	2.344_{10}-8

One observes an exponential decay like $O(\exp(-cr))$ with $c > 0$.

If R varies from 1 to ∞, there is a certain finite value $R^* = R_r^*$ depending on r, such that $\varepsilon(f,[1,R],r)$ as a function of R strictly increases in $[1,R^*]$, whereas the approximant $E_{r,[1,R]}$ as well as the error $\varepsilon(f,[1,R],r)$ is constant in $[R^*,\infty)$. This implies that the approximation $E_{r,[1,R^*]}$ is already the best approximation in the semi-infinite interval $[1,\infty)$. The next table shows R_r^* and $\varepsilon(1/x,[1,R_r^*],r) = \varepsilon(1/x,[1,\infty),r)$:

r	9	16	25	36	49
R_r^*	28387	2.027_{10}+6	1.513_{10}+8	1.162_{10}+10	9.074_{10}+11
$\varepsilon(\frac{1}{x},[1,\infty),r)$	2.611_{10}-5	3.659_{10}-7	4.898_{10}-9	6.382_{10}-11	8.172_{10}-13
$25\exp\left(-\pi\sqrt{2r}\right)$	4.068_{10}-5	4.785_{10}-7	5.628_{10}-9	6.619_{10}-11	7.786_{10}-13

Here, the accuracy behaves like the function $25\exp\left(-\pi\sqrt{2r}\right)$ given for comparison. The behaviour of $f(x) = 1/\sqrt{x}$ is quite similar:

r	9	16	25	36	49
R_r^*	7.994_{10}+6	4.129_{10}+9	2.17_{10}+12	1.15_{10}+15	6.10_{10}+17
$\varepsilon(1/\sqrt{x},[1,\infty),r)$	3.072_{10}-4	1.352_{10}-5	5.898_{10}-7	2.564_{10}-8	1.116_{10}-9
$4\exp\left(-\pi\sqrt{r}\right)$	3.228_{10}-4	1.395_{10}-5	6.028_{10}-7	2.605_{10}-8	1.126_{10}-9

The observed asymptotic decay from the last row of the table is better than the upper bound in the next theorem.

Theorem 9.29. *Let $f(x) = x^{-\lambda}$ with $\lambda > 0$. The asymptotic behaviour of the error $\varepsilon(f,I,r)$ is*

$$\varepsilon(f,I,r) \leq \begin{cases} C\exp(-cr) & \text{for a finite positive interval } I = [a,b] \subset (0,\infty)\,, \\ C\exp(-c\sqrt{r}\,) & \text{for a semi-infinite interval } I = [a,\infty),\ a > 0, \end{cases}$$

where the constants $C, c > 0$ depend on I. For instance, for $\lambda = 1/2$ and $a = 1$, upper bounds are

$$\varepsilon(1/\sqrt{x},[1,R],r) \leq 8\sqrt{2}\exp\left(-\pi^2 r/\sqrt{\log(8R)}\right),$$
$$\varepsilon(1/\sqrt{x},[1,\infty),r) \leq 8\sqrt{2}\exp\left(-\pi\sqrt{r/2}\right).$$

For general $a > 0$ use (9.30a,b).

Proof. Details about the constants can be found in Braess-Hackbusch [27], [28]. The latter estimates can be found in [28, Eqs. (33), (34)]. □

Best approximations with respect to the supremum norm can be performed by the Remez algorithm (cf. Remez [162]). For details of the implementation in the case of exponential sums see [84, §7] and [28, §7].

9.7.2.4 Other Exponential Sums

Another well-known type of exponential sums are trigonometric series. A periodic function in $[0, 2\pi]$ has the representation $f(x) = \sum_{\nu\in\mathbb{Z}} a_\nu \mathrm{e}^{\mathrm{i}\nu x}$. The coefficients a_ν decay the faster the smoother the function is. In that case, $f_n(x) = \sum_{|\nu|\leq n} a_\nu \mathrm{e}^{\mathrm{i}\nu x}$ yields a good approximation. f_n is of the form (9.27a) with imaginary coefficients $\alpha_\nu := \mathrm{i}\nu$.

Besides real coefficients α_ν like in Theorem 9.28 and imaginary ones as above, also complex coefficients with positive real part appear in applications. An important example is the Bessel function J_0, which is approximated by exponential sums in Beylkin-Monzón [17].

9.7.2.5 Application to Multivariate Functions

9.7.2.5.1 Multivariate Functions Derived from $1/x$

We start with an application for $f(x) = 1/x$. Let $f_j \in C(D_j)$ $(1 \leq j \leq d)$ be functions with values in $I_j \subset (0, \infty)$. Set

$$I := \sum_{j=1}^{d} I_j = \left\{ \sum_{j=1}^{d} y_j : y_j \in I_j \right\} = [a, b],$$

possibly with $b = \infty$. Choose an optimal exponential sum E_r for $\frac{1}{x}$ on[13] I with error bound $\varepsilon(\frac{1}{x}, I, r)$. As in the construction (9.27c), we obtain a best approximation of $F(\mathbf{x}) = F(x_1, \ldots, x_d) := 1/\sum_{j=1}^{d} f_j(x_j)$ by

$$\left\| \frac{1}{\sum_{j=1}^{d} f_j(x_j)} - \sum_{\nu=1}^{r} a_{\nu,I} \prod_{j=1}^{d} \exp\left(-\alpha_{\nu,I}\, f_j(x_j)\right) \right\|_{I,\infty} \leq \varepsilon(\tfrac{1}{x}, I, r),$$

i.e., $\|F - F_r\|_{I,\infty} \leq \varepsilon(\frac{1}{x}, I, r)$ with $F_r := \sum_{\nu=1}^{r} a_{\nu,I} \bigotimes_{j=1}^{d} E_\nu^{(j)} \in \mathcal{R}_r$, where $E_\nu^{(j)} = \exp(-\alpha_{\nu,I}\, f_j(\cdot))$.

Since $a_{\nu,I} > 0$ (cf. Theorem 9.28), the functions $E_\nu^{(j)}$ belong to the class A_j of positive functions. In the notation of §7.7, $F_r \in \mathcal{R}_r\big((A_j)_{j=1}^d\big)$ is valid (cf. §9.6).

In quantum chemistry, a so-called MP2 energy denominator $\frac{1}{\varepsilon_a+\varepsilon_b-\varepsilon_i-\varepsilon_j}$ appears, where $\varepsilon_a, \varepsilon_b > 0$ and $\varepsilon_i, \varepsilon_j < 0$ (more than four energies $\varepsilon_\bullet$ are possible). The denominator is contained in $[A, B]$ with $A := 2\,(\varepsilon_{\mathrm{LUMO}} - \varepsilon_{\mathrm{HOMO}}) > 0$ being related to the HOMO-LUMO gap, while $B := 2\,(\varepsilon_{\max} - \varepsilon_{\min})$ involves the maximal and minimal orbital energies (cf. [181]). Further computations are significantly accelerated, if the dependencies of $\varepsilon_a, \varepsilon_b, \varepsilon_i, \varepsilon_j$ can be separated. For this purpose, the optimal exponential sum E_r for $\frac{1}{x}$ on $[A, B]$ can be used:

$$\frac{1}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \approx \sum_{\nu=1}^{r} a_{\nu,I\nu}\, \mathrm{e}^{-\alpha_\nu \varepsilon_a} \cdot \mathrm{e}^{-\alpha_\nu \varepsilon_b} \cdot \mathrm{e}^{\alpha_\nu \varepsilon_i} \cdot \mathrm{e}^{\alpha_\nu \varepsilon_j} \in \mathcal{R}_r, \tag{9.31}$$

[13] For a larger interval I', E_r yields a (non-optimal) error bound with $\varepsilon(\frac{1}{x}, I', r)$.

where the error can be uniformly estimated by $\varepsilon(\frac{1}{x}, [A, B], r)$.

In quantum chemistry, the usual derivation of the exponential sum approximation starts from the Laplace transform $\frac{1}{x} = \int_0^\infty \exp(-tx)\mathrm{d}t$ and applies certain quadrature methods as described in §9.7.2.2 (cf. Almlöf [2]). However, in this setting it is hard to describe how the best quadrature rule should be chosen. Note that the integrand $\exp(-tx)$ is parameter dependent.

9.7.2.5.2 Multivariate Functions Derived from $1/\sqrt{x}$

The function

$$\mathbf{P}(\mathbf{x}) := \frac{1}{\|\mathbf{x}\|} = \frac{1}{\left\|\sum_{j=1}^3 x_j^2\right\|} \qquad \text{for } \mathbf{x} \in \mathbb{R}^3$$

is called *Newton potential*, if gravity is described, and *Coulomb potential* in connection with an electrical field. Mathematically, $4\pi\mathbf{P}$ is the singularity function of the Laplace operator

$$\varDelta = \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} \tag{9.32}$$

for $d = 3$ (cf. [82, §2.2]). Usually, it appears in a convolution integral $\mathbf{P} \star \mathbf{f}$. If $\mathbf{f}$ is the mass [charge] density,

$$4\pi \int_{\mathbb{R}^3} \frac{\mathbf{f}(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|}\mathrm{d}\mathbf{y} = 4\pi\left(\mathbf{P} \star \mathbf{f}\right)(\mathbf{x})$$

describes the gravitational [electrical] field.

Obviously, it is impossible to approximate $\mathbf{P}$ uniformly on the whole $\mathbb{R}^3$ by exponential sums. Instead, we choose some $\eta > 0$ which will be fixed in Lemma 9.30. Take an optimal approximation E_r of $1/\sqrt{t}$ on $I := [\eta^2, \infty)$. Following the strategy from (9.28), we substitute $t = \|\mathbf{x}\|^2 = \sum_{j=1}^3 x_j^2$ and obtain

$$E_r(\|\mathbf{x}\|^2) = \sum_{\nu=1}^{r} a_{\nu,I} \prod_{j=1}^{3} \exp(-\alpha_{\nu,I}\, x_j^2),$$

i.e., $E_{r,[\eta^2,\infty)}(\|\cdot\|^2) = \sum_{\nu=1}^r a_{\nu,I} \bigotimes_{j=1}^3 E_\nu^{(j)} \in \mathcal{R}_r$ with $E_\nu^{(j)}(\xi) = \mathrm{e}^{-\alpha_{\nu,I}\,\xi^2}$. The uniform estimate

$$\left|\mathbf{P}(\mathbf{x}) - E_{r,[\eta^2,\infty)}(\|\mathbf{x}\|^2)\right| \le \varepsilon(\tfrac{1}{\sqrt{\cdot}}, [\eta^2, \infty), r) = \frac{\varepsilon}{\eta} \qquad \begin{cases} \text{for } \eta \le \|\mathbf{x}\| < \infty \text{ and} \\ \varepsilon := \varepsilon(\frac{1}{\sqrt{\cdot}}, [1, \infty), r) \end{cases}$$

excludes the neighbourhood $U_\eta := \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| \le \eta\}$ of the singularity. Here, we use

$$\left|\mathbf{P}(\mathbf{x}) - E_{r,[\eta^2,\infty)}(\|\mathbf{x}\|^2)\right| \le \mathbf{P}(\mathbf{x}) \text{ for } \mathbf{x} \in U_\eta \quad \text{and} \quad \int_{U_\eta} \mathbf{P}(\mathbf{x})\mathrm{d}\mathbf{x} = 2\pi\eta^2.$$

Lemma 9.30. *Assume $\|\mathbf{f}\|_{L^1(\mathbb{R}^3)} \le C_1$ and $\|\mathbf{f}\|_{L^\infty(\mathbb{R}^3)} \le C_\infty$ set. Then*

$$\left| \int_{\mathbb{R}^3} \frac{\mathbf{f}(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|} \mathrm{d}\mathbf{y} - \int_{\mathbb{R}^3} E_r(\|\mathbf{x}-\mathbf{y}\|^2)\mathbf{f}(\mathbf{y})\mathrm{d}\mathbf{y} \right| \le 2\pi\eta^2 C_\infty + \frac{\varepsilon}{\eta} C_1$$

holds with $\varepsilon := \varepsilon(\frac{1}{\sqrt{\cdot}}, [1,\infty), r)$ for all $\mathbf{x} \in \mathbb{R}^3$. The error bound is minimised for $\eta = \sqrt[3]{\frac{C_1 \varepsilon}{4\pi C_\infty}}$:

$$\left\| \int_{\mathbb{R}^3} \frac{\mathbf{f}(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|} \mathrm{d}\mathbf{y} - \int_{\mathbb{R}^3} E_r(\|\mathbf{x}-\mathbf{y}\|^2)\mathbf{f}(\mathbf{y})\mathrm{d}\mathbf{y} \right\|_{\mathbb{R}^3,\infty} \le \underbrace{\tfrac{3}{2} 2^{\frac{2}{3}} \sqrt[3]{\pi}}_{=3.4873} \sqrt[3]{C_1^2 C_\infty}\, \varepsilon^{\frac{2}{3}}.$$

Inserting the asymptotic behaviour $\varepsilon = 8\sqrt{2}\exp(-\pi\sqrt{r/2})$ from Theorem 9.29, we obtain a bound of the same form $C\exp(-c\sqrt{r})$ with $c = \sqrt{2}\pi/3$. The observed behaviour is better: $O(\exp(-\frac{2\pi}{3}\sqrt{r})$. We conclude from Lemma 9.30 that the convolution $\mathbf{P} \star \mathbf{f}$ may be replaced by the convolution $E_r(\|\cdot\|^2) \star \mathbf{f}$, while the accuracy is still exponentially improving.

In the following, we assume for simplicity that $\mathbf{f}$ is an elementary tensor:

$$\mathbf{f}(\mathbf{y}) = f_1(y_1) \cdot f_2(y_2) \cdot f_3(y_3).$$

As seen in (4.75c), the convolution with $E_r(\|\mathbf{x}-\mathbf{y}\|^2)$ can be reduced to three one-dimensional convolutions:

$$\begin{aligned} \int_{\mathbb{R}^3} \frac{\mathbf{f}(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|} \mathrm{d}\mathbf{y} &\approx \int_{\mathbb{R}^3} E_r(\|\mathbf{x}-\mathbf{y}\|^2)\mathbf{f}(\mathbf{y})\mathrm{d}\mathbf{y} \\ &= \sum_{\nu=1}^{r} a_{\nu,I} \prod_{j=1}^{3} \int_{\mathbb{R}} \exp(-\alpha_{\nu,I}(x_j - y_j)^2) f_j(y_j)\mathrm{d}y_j. \end{aligned}$$

Numerical examples related to integral operators involving the Newton potential can be found in [90].

9.7.2.6 Application to Operators

Functions of matrices and operators are discussed in §4.6.6. Now we consider the situation of two functions f and $\tilde{f}$ applied to a matrix of the form UDU^{H}, where $\tilde{f}$ is considered as approximation of f.

Proposition 9.31. *Let $M = UDU^{\mathsf{H}}$ (U unitary, D diagonal) and assume that f and $\tilde{f}$ are defined on the spectrum $\sigma(M)$. Then the approximation error with respect to the spectral norm $\|\cdot\|_2$ is bounded by*

$$\|f(M) - \tilde{f}(M)\|_2 \le \|f - \tilde{f}\|_{\sigma(M),\infty}. \tag{9.33}$$

The estimate extends to selfadjoint operators. For diagonalisable matrices $M = TDT^{-1}$, the right-hand side becomes $\|T\|_2 \|T^{-1}\|_2 \|f - \tilde{f}\|_{\sigma(M),\infty}$.

Proof. Since $f(M)-\tilde{f}(M)=Uf(D)U^{\mathsf{H}}-U\tilde{f}(D)U^{\mathsf{H}}=U[f(D)-\tilde{f}(D)]U^{\mathsf{H}}$ and unitary transformations do not change the spectral norm, $\|f(M)-\tilde{f}(M)\|_2=\|f(D)-\tilde{f}(D)\|_2=\max\{|f(\lambda)-\tilde{f}(\lambda)|:\lambda\in\sigma(M)\}=\|f-\tilde{f}\|_{\sigma(M),\infty}$ follows. □

The supremum norm on the right-hand side in (9.33) cannot be relaxed to an L^p norm with $p<\infty$. This fact makes the construction of best approximations with respect to the supremum norm so important.

Under stronger conditions on f and $\tilde{f}$, general operators $M\in\mathcal{L}(V,V)$ can be admitted (cf. [86, Satz 13.2.4]).

Proposition 9.32. *Let f and $\tilde{f}$ be holomorphic in a complex domain Ω containing $\sigma(M)$ for some operator $M\in\mathcal{L}(V,V)$. Then*

$$\|f(M)-\tilde{f}(M)\|_2\leq\frac{1}{2\pi}\oint_{\partial\Omega}|f(\zeta)-\tilde{f}(\zeta)|\,\|(\zeta I-M)^{-1}\|_2\,\mathrm{d}\zeta.$$

Proof. Use the representation (4.77a). □

Quite another question is, how $f(M)$ behaves under perturbations of M. Here, the following result for Hölder continuous f is of interest.

Theorem 9.33 ([1]). *Let $f\in C^{\alpha}(\mathbb{R})$ with $\alpha\in(0,1)$, i.e., $|f(x)-f(y)|\leq C\,|x-y|^{\alpha}$ for $x,y\in\mathbb{R}$. Then symmetric matrices (or general selfadjoint operators) M' and M'' satisfy the analogous inequality $\|f(M')-f(M'')\|\leq C'\|M'-M''\|^{\alpha}$.*

The corresponding statement for Lipschitz continuous f (i.e., for $\alpha=1$) is wrong, but generalisations to functions of the Hölder-Zygmund class are possible (cf. [1]).

The inverse of M can be considered as the application of the function $f(x)=1/x$ to M, i.e., $f(M)=M^{-1}$. Assume that M is Hermitean (selfadjoint) and has a positive spectrum $\sigma(M)\subset[a,b]\subset(0,\infty]$. As approximation $\tilde{f}$ we choose the best exponential sum $E_{r,I}(x)=\sum_{\nu=1}^{r}a_{\nu,I}\exp(-\alpha_{\nu,I}x)$ on I, where $I\supset[a,b]$. Then

$$E_{r,I}(M)=\sum\nolimits_{\nu=1}^{r}a_{\nu,I}\exp(-\alpha_{\nu,I}M)\tag{9.34}$$

approximates M^{-1} exponentially well:

$$\|f(M)-\tilde{f}(M)\|_2\leq\varepsilon(\tfrac{1}{x},I,r).\tag{9.35}$$

The approximation of M^{-1} seems to be rather impractical, since matrix exponentials $\exp(-t_{\nu}M)$ have to be evaluated. The interesting applications, however, are matrices which are sums of certain Kronecker products. We recall Lemma 4.139b:

$$\mathbf{M}=\sum_{j=1}^{d}I\otimes\cdots\otimes M^{(j)}\otimes\cdots\otimes I\in\mathcal{R}_d,\quad M^{(j)}\in\mathbb{K}^{I_j\times I_j}\tag{9.36}$$

(factor $M^{(j)}$ at j-th position) has the exponential

$$\exp(\mathbf{M})=\bigotimes_{j=1}^{d}\exp(M^{(j)}).\tag{9.37}$$

Let $M^{(j)}$ be positive definite with extreme eigenvalues $0 < \lambda_{\min}^{(j)} \leq \lambda_{\max}^{(j)}$ for $1 \leq j \leq d$. Since the spectrum of $\mathbf{M}$ is the sum $\sum_{j=1}^{d} \lambda^{(j)}$ of all $\lambda^{(j)} \in \sigma(M^{(j)})$, the interval $[a,b]$ containing the spectrum $\sigma(\mathbf{M})$ is given by $a := \sum_{j=1}^{d} \lambda_{\min}^{(j)} > 0$ and $b := \sum_{j=1}^{d} \lambda_{\max}^{(j)}$. In the case of an unbounded selfadjoint operator, $b = \infty$ holds. These preparations lead us to the following statement, which is often used for the case $M^{(j)} = I$.

Proposition 9.34. *Let $M^{(j)}, A^{(j)} \in \mathbb{K}^{I_j \times I_j}$ be positive definite matrices with $\lambda_{\min}^{(j)}$ and $\lambda_{\max}^{(j)}$ being the extreme eigenvalues of the generalised eigenvalue problem $A^{(j)}x = \lambda M^{(j)}x$ and set*

$$\begin{aligned}\mathbf{A} =& A^{(1)} \otimes M^{(2)} \otimes \ldots \otimes M^{(d)} + M^{(1)} \otimes A^{(2)} \otimes \ldots \otimes M^{(d)} + \ldots \\ &+ M^{(1)} \otimes \ldots \otimes M^{(d-1)} \otimes A^{(d)}.\end{aligned} \tag{9.38a}$$

Then $\mathbf{A}^{-1}$ can be approximated by

$$\mathbf{B} := \left[\sum_{\nu=1}^{r} a_{\nu,I} \bigotimes_{j=1}^{d} \exp\left(-\alpha_{\nu,I}\left(M^{(j)}\right)^{-1} A^{(j)}\right)\right] \cdot \left[\bigotimes_{j=1}^{d} \left(M^{(j)}\right)^{-1}\right]. \tag{9.38b}$$

The error is given by

$$\left\|\mathbf{A}^{-1} - \mathbf{B}\right\|_2 \leq \varepsilon(\tfrac{1}{x}, [a,b], r) \left\|\mathbf{M}^{-1}\right\|_2 \tag{9.38c}$$

with $\mathbf{M} = \bigotimes_{j=1}^{d} M^{(j)}$, $a := \sum_{j=1}^{d} \lambda_{\min}^{(j)}$, and $b := \sum_{j=1}^{d} \lambda_{\max}^{(j)}$.

Proof. Write $\mathbf{A} = \mathbf{M}^{1/2} \cdot \hat{\mathbf{A}} \cdot \mathbf{M}^{1/2}$ with $\hat{\mathbf{A}} = \hat{A}^{(1)} \otimes I \ldots \otimes I + \ldots$, where $\hat{A}^{(j)} := (M^{(j)})^{-1/2} A^{(j)} (M^{(j)})^{-1/2}$. Note that $\lambda_{\min}^{(j)}$ and $\lambda_{\max}^{(j)}$ are the extreme eigenvalues of $\hat{A}^{(j)}$. Apply (9.35) to $\hat{\mathbf{A}}$ instead of M. For $\exp(-\alpha_{\nu,I}\hat{\mathbf{A}})$ appearing in $\hat{\mathbf{B}} := E_{r,I}(\hat{\mathbf{A}})$ (cf. (9.34)) use the representation (9.37) with the error estimate $\|\hat{\mathbf{A}}^{-1} - \hat{\mathbf{B}}\|_2 \leq \varepsilon(\frac{1}{x}, [a,b], r)$. Note that $\mathbf{B} = \mathbf{M}^{-1/2} \cdot E_{r,I}(\hat{\mathbf{A}}) \cdot \mathbf{M}^{-1/2}$. Hence, $\|\mathbf{A}^{-1} - \mathbf{B}\|_2 = \|\mathbf{M}^{-1/2} \cdot [\hat{\mathbf{A}}^{-1} - E_{r,I}(\hat{\mathbf{A}})] \cdot \mathbf{M}^{-1/2}\|_2 \leq \|\hat{\mathbf{A}}^{-1} - \hat{\mathbf{B}}\|_2 \|\mathbf{M}^{-1/2}\|_2^2$. The identity $\|\mathbf{M}^{-1/2}\|_2^2 = \|\mathbf{M}_2^{-1}\|$ completes the proof. □

It remains to compute the exponentials of $-\alpha_{\nu,I}\left(M^{(j)}\right)^{-1} A^{(j)}$. As described in [86, §13.3.1] and [66], the hierarchical matrix technique allows us to approximate $\exp\left(-\alpha_{\nu,I}(M^{(j)})^{-1} A^{(j)}\right)$ with a cost almost linear in $\#I_j$. The total number of arithmetical operations is $O\left(r \sum_{j=1}^{d} \#I_j \log^* \#I_j\right)$. For $\#I_j = n$ $(1 \leq j \leq d)$, this expression is $O(rdn \log^* n)$ and depends only linearly on d. For identical $A^{(j)} = A^{(k)}$, $M^{(j)} = M^{(k)}$ $(1 \leq j, k \leq d)$, the cost $O(rdn \log^* n)$ reduces to $O(rn \log^* n)$.

Proposition 9.34 can in particular be applied to the Laplace operator and its discretisations as detailed below.

Remark 9.35. (a) The negative Laplace operator (9.32) in[14] $H_0^1([0,1]^d)$ has the d-term format (9.38a) with $M^{(j)} = id$, $A^{(j)} = -\partial^2/\partial x_j^2$ and $\lambda_{\min}^{(j)} = \pi^2$, $\lambda_{\max}^{(j)} = \infty$.

[14] The reference to $H_0^1([0,1]^d)$ means that zero Dirichlet values are prescribed on the boundary.

(b) If we discretise by a finite difference scheme in an equidistant grid of step size $1/n$, $A^{(j)}$ is the tridiagonal matrix[15] $n^{-2} \cdot tridiag\{-1, 2, -1\}$, while $M^{(j)} = I$. The extreme eigenvalues are $\lambda_{\min}^{(j)} = 4n^2 \sin^2(\frac{\pi}{2n}) \approx \pi^2$, $\lambda_{\max}^{(j)} = 4n^2 \cos^2(\frac{\pi}{2n}) \approx 4n^2$.
(c) A finite element discretisation with piecewise linear elements in the same grid leads to the same[16] matrix $A^{(j)}$, but $M^{(j)}$ is the mass matrix $tridiag\{1/6, 2/3, 1/6\}$.

This approach to the inverse allows to treat cases with large n and d. Grasedyck [72] presents examples with $n = 1024$ and $d \approx 1000$. Note that in this case the matrix is of size $\mathbf{A}^{-1} \in \mathbb{R}^{M \times M}$ with $M \approx 10^{3000}$.

The approximation method can be extended to separable differential operators in tensor domains $D = \times_{j=1}^{d} D_j$ with appropriate spectra.

Definition 9.36. A differential operator L is called *separable* if $L = \sum_{j=1}^{d} L_j$ and L_j contains only derivatives with respect to x_j and has coefficients which only depend on x_j.

So far, we have applied the exponential sum $E_r \approx 1/x$. Analogous statements can be made about the application of $E_r \approx 1/\sqrt{x}$. Then r-term approximations of $\mathbf{A}^{-1/2}$ can be computed.

9.7.3 Sparse Grids

The mixed Sobolev space $H_{\text{mix}}^{2,p}([0,1]^d)$ for $2 \le p \le \infty$ is the completion of ${}_a\bigotimes^d H^{2,p}([0,1])$ with respect to the norm $\|f\|_{2,p,\text{mix}} = \big(\sum_{\|\nu\|_\infty \le 2} \int |D^\nu f(x)|^p \big)^{1/p}$ for $p < \infty$ and the obvious modification for $p = \infty$.

The approximation properties of sparse grids can be used to estimate $\varepsilon(\mathbf{v}, r)$ from (9.2) with respect to the L^p norm of $\mathbf{V} = {}_{\|\cdot\|_p}\bigotimes^d L^p([0,1])$.

Remark 9.37. For $\mathbf{v} \in H_{\text{mix}}^{2,p}([0,1]^d)$, the quantity $\varepsilon(\mathbf{v}, r)$ equals

$$\varepsilon(\mathbf{v}, r) = \inf\left\{ \|\mathbf{v} - \mathbf{u}\|_p : \mathbf{u} \in \mathcal{R}_r(\mathbf{V}) \right\} \le O\big(r^{-2} \log^{3(d-1)}(\log r)\big).$$

Proof. $\mathbf{V}_{\text{sg},\ell}$ is defined in (7.18). Note that the completion of $\bigcup_{\ell \in \mathbb{N}} \mathbf{V}_{\text{sg},\ell}$ yields $\mathbf{V}$. Consider $r = \dim(\mathbf{V}_{\text{sg},\ell}) \approx 2^\ell \log^{d-1}(\ell)$ (cf. [29, (3.63)]). The interpolant $\mathbf{u} \in \mathbf{V}_{\text{sg},\ell}$ of $\mathbf{v}$ satisfies $\|\mathbf{v} - \mathbf{u}\|_p \le O(2^{-2\ell} \log^{d-1}(\ell))$ (cf. [29, Theorem 3.8]). The inequality

$$2^{-2\ell} \log^{d-1}(\ell) \le r^{-2} \log^{3(d-1)}(\ell) \le O(r^{-2} \log^{3(d-1)}(\log r))$$

proves the assertion. □

[15] In this case, a cheap, exact evaluation of $\exp(A^{(j)})$ can be obtained by diagonalisation of $A^{(j)}$.

[16] In fact, both matrices are to be scaled by a factor $1/n$.

Chapter 10
Tensor Subspace Approximation

Abstract The exact representation of $\mathbf{v} \in \mathbf{V} = \bigotimes_{j=1}^{d} V_j$ by a tensor subspace representation (8.6b) may be too expensive because of the high dimensions of the involved subspaces or even impossible since $\mathbf{v}$ is a topological tensor admitting no finite representation. In such cases we must be satisfied with an approximation $\mathbf{u} \approx \mathbf{v}$ which is easier to handle. We require that $\mathbf{u} \in \mathcal{T}_\mathbf{r}$, i.e., there are bases $\{b_1^{(j)}, \ldots, b_{r_j}^{(j)}\} \subset V_j$ such that

$$\mathbf{u} = \sum\nolimits_{i_1=1}^{r_1} \cdots \sum\nolimits_{i_d=1}^{r_d} \mathbf{a}[i_1 \cdots i_d] \bigotimes\nolimits_{j=1}^{d} b_{i_j}^{(j)}. \tag{10.1}$$

The basic task of this chapter is the following problem:

$$\text{Given } \mathbf{v} \in \mathbf{V}, \text{ find a suitable approximation } \mathbf{u} \in \mathcal{T}_\mathbf{r} \subset \mathbf{V}, \tag{10.2}$$

where $\mathbf{r} = (r_1, \ldots, r_d) \in \mathbb{N}^d$. Finding $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ means finding coefficients $\mathbf{a}[i_1 \cdots i_d]$ as well as basis vectors $b_i^{(j)} \in V_j$ in (10.1). Problem (10.2) is formulated rather vaguely. If an accuracy $\varepsilon > 0$ is prescribed, $\mathbf{r} \in \mathbb{N}^d$ as well as $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ are to be determined. The strict minimisation of $\|\mathbf{v} - \mathbf{u}\|$ is often replaced by an appropriate approximation $\mathbf{u}$ requiring low computational cost. Instead of $\varepsilon > 0$, we may prescribe the rank vector $\mathbf{r} \in \mathbb{N}^d$ in (10.2).

Optimal approximations (so-called 'best approximations') will be studied in *Sect. 10.2*. While best approximations require an iterative computation, quasi-optimal approximations can be determined explicitly using the HOSVD basis introduced in Sect. 8.3. The latter approach is explained in *Sect. 10.1*.

10.1 Truncation to $\mathcal{T}_\mathbf{r}$

The term 'truncation' (to $\mathcal{T}_\mathbf{r}$) is used here for a (nonlinear) map $\tau = \tau_\mathbf{r} : \mathbf{V} \to \mathcal{T}_\mathbf{r}$ with quasi-optimality properties. Truncation should be seen as a cheaper alternative to the best approximation, which will be discussed in §10.2. Below we describe such truncations based on the higher order singular value decomposition (HOSVD) and study the introduced truncation error.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_10,

One of the advantages of the tensor subspace format is the constructive existence of the higher order singular value decomposition (cf. §8.3). The related truncation is described in §10.1.1, while §10.1.2 is devoted to the successive HOSVD projection. Examples of HOSVD projections can be found in §10.1.3. A truncation starting from an r-term representation is mentioned in §10.1.4.

Throughout this section, $\mathbf{V}$ is a Hilbert tensor space with induced scalar product. Often, we assume $V_j = \mathbb{K}^{I_j}$, where $n_j := \#I_j$ denotes the dimension.

10.1.1 HOSVD Projection

The tensor to be approximated will be called $\mathbf{v} \in \mathbf{V}$, while the approximant is denoted by $\mathbf{u}$ (possibly with further subscripts). The standard assumption is that $\mathbf{v}$ is represented in tensor subspace format, i.e., $\mathbf{v} \in \mathcal{T}_\mathbf{s}$ for some $\mathbf{s} \in \mathbb{N}_0^d$, whereas the approximant $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ is sought for some[1] $\mathbf{r} \lneqq \mathbf{s}$. If $\mathbf{v} \in \mathbf{V} = \bigotimes_{j=1}^d \mathbb{K}^{I_j}$ is given in full representation (cf. §7.2), it can be interpreted as $\mathbf{v} \in \mathcal{T}_\mathbf{s}$ with $\mathbf{s} := \mathbf{n} = (n_1, \ldots, n_d)$.

Optimal approximants $\mathbf{u}_\text{best}$ are a favourite subject in theory, but in applications[2] one is often satisfied with *quasi-optimal* approximations. We say that $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ is quasi-optimal, if there is a constant C such that

$$\|\mathbf{v} - \mathbf{u}\| \le C \, \|\mathbf{v} - \mathbf{u}_\text{best}\| \, . \tag{10.3}$$

The following approach is based on the higher-order singular value decomposition from §8.3. Therefore, the truncation is practically feasible if and only if the higher-order singular value decomposition is available. In particular, the cost of the HOSVD projection is identical to the cost of the HOSVD calculation discussed in §8.3.3. Moreover, it requires a Hilbert space structure of $\mathbf{V}$ as mentioned above.

We recall that, given $\mathbf{v} \in \mathbf{V}$, the j-th HOSVD basis $B_j = [b_1^{(j)} \cdots b_{s_j}^{(j)}]$ (cf. Definition 8.22) is a particular orthonormal basis of $U_j^{\min}(\mathbf{v})$, where each $b_i^{(j)}$ is associated with a singular value $\sigma_i^{(j)}$. The basis vectors are ordered according to $\sigma_1^{(j)} \ge \sigma_2^{(j)} \ge \ldots$ We use $s_j = \dim(U_j^{\min}(\mathbf{v}))$ instead of r_j in Definition 8.22. The following projections P_j correspond to the SVD projections from Remark 2.31.

Lemma 10.1 (HOSVD projection). *Let $\mathbf{v} \in \mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$, where $\mathbf{V}$ is a Hilbert tensor space. Let $\{b_i^{(j)} : 1 \le i \le s_j\}$ be the HOSVD basis of $U_j^{\min}(\mathbf{v})$. The j-th* HOSVD projection $P_j^\text{HOSVD} = P_{j,\text{HOSVD}}^{(r_j)}$ *corresponding to $r_j \le s_j$ is the orthogonal projection onto $U_{j,\text{HOSVD}}^{(r_j)} := \operatorname{span}\{b_i^{(j)} : 1 \le i \le r_j\}$. Its explicit description is*[3]

[1] The notation $\mathbf{r} \lneqq \mathbf{s}$ means that $r_j \le s_j$ for all $1 \le j \le d$, but $r_j < s_j$ for at least one index j.

[2] A prominent example is the Galerkin approximation technique, where the Lemma of Céa proves that the Galerkin solution in a certain subspace is quasi-optimal compared with the best approximation in that subspace (cf. [82, Theorem 8.2.1]).

[3] Note that P_j^HOSVD and $\mathbf{P}_\mathbf{r}^\text{HOSVD}$ depend on the tensor $\mathbf{v} \in \mathbf{V}$ whose singular vectors $b_i^{(j)}$ enter their definition. However, we avoid the notation $P_j^\text{HOSVD}(\mathbf{v})$ since this looks like the application of the projection onto $\mathbf{v}$.

$$P_j^{\mathrm{HOSVD}} = P_{j,\mathrm{HOSVD}}^{(r_j)} = \sum_{i=1}^{r_j} b_i^{(j)} b_i^{(j)*} = B_j^{(r_j)} \left(B_j^{(r_j)}\right)^* \in \mathcal{L}(V_j, V_j),$$

with $B_j^{(r_j)} := [b_1^{(j)} \cdots b_{r_j}^{(j)}] \in (V_j)^{r_j}$*. The overall HOSVD projection* $\mathbf{P}_\mathbf{r}^{\mathrm{HOSVD}}$ *is the orthogonal projection onto the tensor subspace* $\bigotimes_{j=1}^d U_j^{\mathrm{HOSVD}}$*, described by*

$$\mathbf{P}_\mathbf{r}^{\mathrm{HOSVD}} := \bigotimes_{j=1}^d P_{j,\mathrm{HOSVD}}^{(r_j)} \in \mathcal{L}(\mathbf{V}, \mathbf{V}) \qquad \textit{with } \mathbf{r} := (r_1, \ldots, r_d)\,.$$

We repeat the HOSVD representation in the case of $V_j = \mathbb{K}^{I_j}$ with $n_j := \#I_j$. The coefficient tensor $\mathbf{a} \in \mathbb{K}^{\hat{\mathbf{J}}}$ of $\mathbf{v}$ uses the index sets $\hat{J}_j = \{1, \ldots, s_j\}$ with s_j from above, whereas the index set $J_j = \{1, \ldots, r_j\}$ refers to $r_j \le s_j$.

Corollary 10.2. Let $\mathbf{V} = \bigotimes_{j=1}^d \mathbb{K}^{I_j}$ be endowed with the Euclidean scalar product. The HOSVD representation of $\mathbf{v} \in \mathbf{V}$ by $\mathbf{v} = \rho_{\mathrm{HOSVD}}\big(\mathbf{a}, (B_j)_{1\le j\le d}\big) = \mathbf{B}\mathbf{a}$ (cf. (8.26)) is characterised by

$$\mathbf{B} = \bigotimes_{j=1}^d B_j, \quad B_j \in \mathbb{K}^{I_j \times \hat{J}_j}, \quad s_j = \#\hat{J}_j := \mathrm{rank}(\mathcal{M}_j(\mathbf{a})), \quad B_j^{\mathsf{H}} B_j = I,$$

$$\mathcal{M}_j(\mathbf{a})\mathcal{M}_j(\mathbf{a})^{\mathsf{H}} = \mathrm{diag}\{\sigma_1^{(j)}, \sigma_2^{(j)}, \ldots, \sigma_{s_j}^{(j)}\}, \quad \sigma_1^{(j)} \ge \sigma_2^{(j)} \ge \ldots \ge \sigma_{s_j}^{(j)} > 0$$

(cf. Corollary 8.25 and (8.24) with s_j instead of r_j). For given rank vector $\mathbf{r} \in \mathbb{N}^d$ with $r_j \le s_j$ let $B_j^{(r_j)}$ be the restriction of the matrix B_j to the first r_j columns. Then $\mathbf{P}_\mathbf{r}^{\mathrm{HOSVD}} := \mathbf{B}^{(\mathbf{r})}\mathbf{B}^{(\mathbf{r})\mathsf{H}}$ with $\mathbf{B}^{(\mathbf{r})} = \bigotimes_{j=1}^d B_j^{(r_j)}$ is the orthogonal projection onto $\mathbf{U}_{\mathrm{HOSVD}}^{(\mathbf{r})} = \bigotimes_{j=1}^d U_{j,\mathrm{HOSVD}}^{(r_j)}$ with $U_{j,\mathrm{HOSVD}}^{(r_j)} = \mathrm{range}\{B_j^{(r_j)}\}$.

The first inequality in (10.4b) below is described by De Lathauwer et al. [41, Property 10]. While this first inequality yields a concrete error estimate, the second one in (10.4b) states quasi-optimality. The constant $C = \sqrt{d}$ shows independence of the dimensions $\mathbf{r} \le \mathbf{s} \le \mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$.

Theorem 10.3. *Let* $\mathbf{V} = \bigotimes_{j=1}^d \mathbb{K}^{I_j}$ *be endowed with the Euclidean scalar product. Define the orthogonal projection* $\mathbf{P}_\mathbf{r}^{\mathrm{HOSVD}}$ *and the singular values* $\sigma_i^{(j)}$ *as in Corollary 10.2. Then the* HOSVD truncation *is defined by*

$$\mathbf{u}_{\mathrm{HOSVD}} := \mathbf{P}_\mathbf{r}^{\mathrm{HOSVD}}\mathbf{v} \in \mathcal{T}_\mathbf{r}. \tag{10.4a}$$

This approximation is quasi-optimal:

$$\|\mathbf{v} - \mathbf{u}_{\mathrm{HOSVD}}\| \le \sqrt{\sum_{j=1}^d \sum_{i=r_j+1}^{s_j} \left(\sigma_i^{(j)}\right)^2} \le \sqrt{d}\,\|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\|\,, \tag{10.4b}$$

where $\mathbf{u}_{\mathrm{best}} \in \mathcal{T}_\mathbf{r}$ *yields the minimal error (i.e.,* $\|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\| = \min_{\mathbf{u}\in\mathcal{T}_\mathbf{r}} \|\mathbf{v} - \mathbf{u}\|$*).*

Proof (cf. [73]). We introduce the shorter notations $P_j = P^{(r_j)}_{j,\mathrm{HOSVD}}$, $U_j = U^{(r_j)}_{j,\mathrm{HOSVD}}$, $B_j = B_j^{(r_j)}$, and $\mathbf{B} = \mathbf{B}^{(\mathbf{r})}$.

$$P_j := I \otimes \ldots \otimes I \otimes B_j B_j^{\mathsf{H}} \otimes I \otimes \ldots \otimes I \tag{10.5}$$

is the projection onto

$$\mathbf{V}^{(j)} := \mathbb{K}^{I_1} \otimes \ldots \otimes \mathbb{K}^{I_{j-1}} \otimes U_j \otimes \mathbb{K}^{I_{j+1}} \otimes \ldots \otimes \mathbb{K}^{I_d}.$$

Then the projection $\mathbf{P_r} = \mathbf{BB}^{\mathsf{H}} = \bigotimes_{j=1}^d B_j B_j^{\mathsf{H}}$ is the product $\prod_{j=1}^d P_j$ and yields $\|\mathbf{v} - \mathbf{u}_{\mathrm{HOSVD}}\| = \|(I - \prod_{j=1}^d P_j)\mathbf{v}\|$. Lemma 4.123b proves the estimate

$$\|\mathbf{v} - \mathbf{u}_{\mathrm{HOSVD}}\|^2 \leq \sum_{j=1}^d \|(I - P_j)\,\mathbf{v}\|^2 .$$

The singular value decomposition of $\mathcal{M}_j(\mathbf{v})$ used in HOSVD implies that $(I - P_j)\mathbf{v}$ is the best approximation of $\mathbf{v}$ in $\mathbf{V}^{(j)}$ under the condition $\dim(U_j) = r_j$. Error estimate (2.19c) implies $\|(I - P_j)\mathbf{v}\|^2 = \sum_{i=r_j+1}^{s_j} (\sigma_i^{(j)})^2$. Thereby, the first inequality in (10.4b) is shown. The best approximation $\mathbf{u}_{\mathrm{best}}$ belongs to

$$\mathbb{K}^{I_1} \otimes \ldots \otimes \mathbb{K}^{I_{j-1}} \otimes \tilde{U}_j \otimes \mathbb{K}^{I_{j+1}} \otimes \ldots \otimes \mathbb{K}^{I_d}$$

with some subspace $\tilde{U}_j$ of dimension r_j. Since $(I - P_j)\mathbf{v}$ is the best approximation in this respect,

$$\|(I - P_j)\,\mathbf{v}\|^2 \leq \|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\|^2 \tag{10.6}$$

holds and proves the second inequality in (10.4b). □

Corollary 10.4. If $r_j = s_j$ (i.e., no reduction in the j-th direction), the sum $\sum_{i=r_j+1}^{s_j}$ in (10.4b) vanishes and the bound $\sqrt{d}$ can be improved by $\sqrt{\#\{j : r_j < s_j\}}$.

Since $\mathbf{u}_{\mathrm{HOSVD}} \in \bigotimes_{j=1}^d U_j^{\min}(\mathbf{v})$, the statements from the second part of Lemma 10.7 are still valid.

10.1.2 Successive HOSVD Projection

The algorithm behind Theorem 10.3 reads as follows: For all $1 \leq j \leq d$ compute the left-sided singular value decomposition of $\mathcal{M}_j(\mathbf{v})$ in order to obtain B_j and $\sigma_i^{(j)}$ $(1 \leq i \leq s_j)$. After all data are computed, the projection $\mathbf{P}_{\mathbf{r}}^{\mathrm{HOSVD}}$ is applied.

Instead, the projections can be applied sequentially, so that the result of the previous projections is already taken into account. The projection $\tilde{P}_j$ from (10.7) is again $P^{(r_j)}_{j,\mathrm{HOSVD}}$, but referring to the singular value decomposition of the actual tensor $\mathbf{v}_{j-1}$ (instead of $\mathbf{v}_0 = \mathbf{v}$):

Start	$\mathbf{v}_0 := \mathbf{v}$
Loop $j = 1$ to d	Perform the left-sided SVD of $\mathcal{M}_j(\mathbf{v}_{j-1})$ yielding the basis B_j and the singular values $\tilde{\sigma}_i^{(j)}$. Let $\tilde{B}_j$ be the restriction of B_j to the first r_j columns and set $\tilde{P}_j := I \otimes \ldots \otimes I \otimes \tilde{B}_j \tilde{B}_j^{\mathsf{H}} \otimes I \otimes \ldots \otimes I$. Define $\mathbf{v}_j := \tilde{P}_j \mathbf{v}_{j-1}$.
Return	$\tilde{\mathbf{u}}_{\mathrm{HOSVD}} := \mathbf{v}_d$.

(10.7)

The projection $\tilde{P}_j$ maps V_j onto some subspace U_j of dimension r_j. Hence, $\mathbf{v}_j$ belongs to $\mathbf{V}^{(j)} := U_1 \otimes \ldots \otimes U_j \otimes V_{j+1} \otimes \ldots \otimes V_d$. One advantage is that the computation of the left-sided singular value decomposition of $\mathcal{M}_j(\mathbf{v}_{j-1})$ is *cheaper* than the computation for $\mathcal{M}_j(\mathbf{v})$, since $\dim(\mathbf{V}^{(j)}) \le \dim(\mathbf{V})$. There is also an argument, why this approach may yield better results. Let $\mathbf{v}_1 := P_1^{\mathrm{HOSVD}}\mathbf{v} = \tilde{P}_1\mathbf{v}$ (note that $P_1^{\mathrm{HOSVD}} = \tilde{P}_1$, where P_j^{HOSVD} from Lemma 10.1 belongs to $\mathbf{v}$) be the result of the first step $j = 1$ of the loop. Projection P_1^{HOSVD} splits $\mathbf{v}$ into $\mathbf{v}_1 + \mathbf{v}_1^{\perp}$. If we use the projection P_2^{HOSVD} from Theorem 10.3, the singular values $\sigma_i^{(2)}$ select the basis $\hat{B}_j$. The singular value $\sigma_i^{(2)}$ corresponds to the norm of[4] $b_i^{(2)}b_i^{(2)\mathsf{H}}\mathbf{v}$, but what really matters is the size of $b_i^{(2)}b_i^{(2)\mathsf{H}}\mathbf{v}_1$, which is the singular value $\tilde{\sigma}_i^{(2)}$ computed in (10.7) from $\mathbf{v}_1$. This proves that the projection $\tilde{P}_2$ yields a better result than P_2^{HOSVD} from Theorem 10.3, i.e., $\|\mathbf{v} - \mathbf{v}_2\| = \|\mathbf{v} - \tilde{P}_2\tilde{P}_1\mathbf{v}\| \le \|\mathbf{v} - P_2^{\mathrm{HOSVD}}P_1^{\mathrm{HOSVD}}\mathbf{v}\|$.

One can prove an estimate corresponding to the first inequality in (10.4b), but now (10.4b) becomes an equality. Although $\tilde{\sigma}_i^{(j)} \le \sigma_i^{(j)}$ holds, this does not imply that $\|\mathbf{v} - \tilde{\mathbf{u}}_{\mathrm{HOSVD}}\| \le \|\mathbf{v} - \mathbf{u}_{\mathrm{HOSVD}}\|$. Nevertheless, in general one should expect the sequential version to be better.

Theorem 10.5. *The error of $\tilde{\mathbf{u}}_{\mathrm{HOSVD}}$ from (10.7) is equal to*

$$\|\mathbf{v} - \tilde{\mathbf{u}}_{\mathrm{HOSVD}}\| = \sqrt{\sum_{j=1}^{d}\sum_{i=r_j+1}^{s_j}\left(\tilde{\sigma}_i^{(j)}\right)^2} \le \sqrt{d}\,\|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\|\,. \tag{10.8}$$

The arising singular values satisfy $\tilde{\sigma}_i^{(j)} \le \sigma_i^{(j)}$, where the values $\sigma_i^{(j)}$ belong to the algorithm from Theorem 10.3.

Proof. 1) We split the difference into

$$\begin{aligned}\mathbf{v} - \tilde{\mathbf{u}}_{\mathrm{HOSVD}} &= (I - \tilde{P}_d\tilde{P}_{d-1}\cdots\tilde{P}_1)\mathbf{v} \\ &= (I - \tilde{P}_1)\mathbf{v} + (I - \tilde{P}_2)\tilde{P}_1\mathbf{v} + \ldots + (I - \tilde{P}_d)\tilde{P}_{d-1}\cdots\tilde{P}_1\mathbf{v}.\end{aligned}$$

Since the projections commute ($\tilde{P}_j\tilde{P}_k = \tilde{P}_k\tilde{P}_j$), all terms on the right-hand side are orthogonal. Setting $\mathbf{v}_j = \tilde{P}_j\cdots\tilde{P}_1\mathbf{v}$, we obtain

$$\|\mathbf{v} - \tilde{\mathbf{u}}_{\mathrm{HOSVD}}\|^2 = \sum_{j=1}^{d}\left\|\left(I - \tilde{P}_j\right)\mathbf{v}_{j-1}\right\|^2.$$

[4] $b_i^{(2)}b_i^{(2)\mathsf{H}}$ applies to the 2nd component: $(b_i^{(2)}b_i^{(2)\mathsf{H}})\bigotimes_{j=1}^{d} v^{(j)} = v^{(1)} \otimes \langle v^{(j)}, b_i^{(2)}\rangle b_i^{(2)} \otimes v^{(3)} \otimes \ldots$

Now, $\|(I-\tilde{P}_j)\mathbf{v}_j\|^2 = \sum_{i=r_j+1}^{s_j}(\tilde{\sigma}_i^{(j)})^2$ finishes the proof of the first part.

2) For $j=1$, the same HOSVD basis is used so that $\tilde{\sigma}_i^{(1)} = \sigma_i^{(1)}$ ($\sigma_i^{(j)}$ are the singular values from Theorem 10.3). For $j \geq 2$ the sequential algorithm uses the singular value decomposition of $\mathcal{M}_j(\mathbf{v}_{j-1}) = \mathcal{M}_j(\tilde{P}_{j-1}\cdots\tilde{P}_1\mathbf{v})$. The product $\tilde{P}_{j-1}\cdots\tilde{P}_1$ is better written as Kronecker product $\tilde{\mathbf{P}}\otimes\mathbf{I}$, where $\tilde{\mathbf{P}} = \bigotimes_{k=1}^{j-1}\tilde{P}_k$ and $\mathbf{I} = \bigotimes_{k=j}^{d} I$. According to (5.5),

$$\mathcal{M}_j(\mathbf{v}_{j-1})\mathcal{M}_j(\mathbf{v}_{j-1})^{\mathsf{H}} = \mathcal{M}_j(\mathbf{v})\tilde{\mathbf{P}}^{\mathsf{T}}\overline{\tilde{\mathbf{P}}}\mathcal{M}_j(\mathbf{v})^{\mathsf{H}} \leq \mathcal{M}_j(\mathbf{v})\mathcal{M}_j(\mathbf{v})^{\mathsf{H}}$$

holds because of $\tilde{\mathbf{P}}^{\mathsf{T}}\overline{\tilde{\mathbf{P}}} = \tilde{\mathbf{P}} \leq I$ (cf. Remark 4.122d). By Lemma 2.27b, the singular values satisfy $\tilde{\sigma}_i^{(j)} \leq \sigma_i^{(j)}$ ($\tilde{\sigma}_i^{(j)}$: singular values of $\mathcal{M}_j(\mathbf{v}_{j-1})$, $\sigma_i^{(j)}$: those of $\mathcal{M}_j(\mathbf{v})$). Therefore, the last inequality in (10.8) follows from (10.4b). □

10.1.3 Examples

Examples 8.26 and 8.27 describe two tensors from $\mathcal{T}_{(2,2,2)} \subset V\otimes V\otimes V$. Here, we discuss their truncation to $\mathcal{T}_{(1,1,1)} = \mathcal{R}_1$.

Tensor $\mathbf{v} = x\otimes x\otimes x + \sigma y\otimes y\otimes y$ from (8.28) is already given in HOSVD representation. Assuming $1>\sigma>0$, the HOSVD projection $P_j^{(1)} := P_{j,\mathrm{HOSVD}}^{(1)}$ is the projection onto $\mathrm{span}\{x\}$, i.e.,

$$\mathbf{u}_{\mathrm{HOSVD}} := \mathbf{P}_{(1,1,1)}^{\mathrm{HOSVD}}\mathbf{v} = x\otimes x\otimes x \in \mathcal{T}_{(1,1,1)} \tag{10.9}$$

is the HOSVD projection from Theorem 10.3. Obviously, the error is

$$\|\mathbf{v}-\mathbf{u}_{\mathrm{HOSVD}}\| = \|\sigma y\otimes y\otimes y\| = \sigma = \sigma_2^{(1)}$$

(cf. Example 8.26), and therefore smaller than the upper bound in (10.4b). The reason becomes obvious, when we apply the factors in $\mathbf{P}_{(1,1,1)}^{\mathrm{HOSVD}} = P_1^{(1)}\otimes P_2^{(1)}\otimes P_3^{(1)}$ sequentially. Already the first projection maps $\mathbf{v}$ into the final value $P_1^{(1)}\mathbf{v} = x\otimes x\otimes x$; the following projections cause no further approximation errors. Accordingly, if we apply the successive HOSVD projection from §10.1.2, the first step of algorithm (10.7) yields $P_1^{(1)}\mathbf{v} = x\otimes x\otimes x \in \mathcal{T}_{(1,1,1)}$, and no further projections are needed (i.e., $P_2^{(1)} = P_3^{(1)} = id$). Therefore, $\tilde{\mathbf{u}}_{\mathrm{HOSVD}} := P_1^{(1)}\mathbf{v}$ holds, and only the singular value $\sigma_2^{(1)}$ for $j=1$ appears in the error estimate (10.8).

Example 8.27 uses the tensor $\mathbf{v} = \alpha x\otimes x\otimes x + \beta x\otimes x\otimes y + \beta x\otimes y\otimes x + \beta y\otimes x\otimes x$, where α,β are chosen such that again $1 = \sigma_1^{(j)} > \sigma_2^{(j)} = \sigma \in [0,1)$ are the singular values for all j (cf. (8.29b)). The HOSVD bases $\{b_1^{(j)}, b_2^{(j)}\}$ are given in (8.29c). Since $b_i^{(j)} = b_i$ is independent of j, we omit the superscript j. The HOSVD projection yields

$$\mathbf{u}_{\mathrm{HOSVD}} = \gamma\, b_1\otimes b_1\otimes b_1 \qquad \text{with} \quad \gamma := \alpha\varkappa^3 + 3\beta\varkappa^2\lambda \text{ and} \tag{10.10}$$
$$x = \varkappa b_1 + \lambda b_2,\ y = \lambda b_1 - \varkappa b_2, \quad b_1 = \varkappa x + \lambda y,\ b_2 = \lambda x - \varkappa y,$$

where the coefficients $\varkappa = \sqrt{\frac{1-\sigma/\sqrt{2}}{(1+\sigma)(1-\sigma)}}$ and $\lambda = \sqrt{\frac{\sigma\left(1/\sqrt{2}-\sigma\right)}{(1+\sigma)(1-\sigma)}}$ $\left(\varkappa^2 + \lambda^2 = 1\right)$ are functions of the singular value $\sigma = \sigma_2^{(j)}$. The error is given by

$$\|\mathbf{v} - \mathbf{u}_{\mathrm{HOSVD}}\| = \sqrt{3/2}\,\sigma + O(\sigma^2).$$

For the special choice $\sigma = 1/10$, the approximation $\mathbf{u}_{\mathrm{HOSVD}}$ and its error are

$$\mathbf{u}_{\mathrm{HOSVD}} = \otimes^3(0.968135\,x + 0.247453\,y)\,,\ \|\mathbf{v}-\mathbf{u}_{\mathrm{HOSVD}}\| = 0.120158. \quad (10.11)$$

Next, we consider the successive HOSVD projection from §10.1.2. The first projection yields

$$\begin{aligned}\mathbf{u}^{(1)} &= b_1^{(1)} \otimes [(\alpha\varkappa + \beta\lambda)\ x \otimes x + \beta\varkappa\, x \otimes y + \beta\varkappa\, y \otimes x]\\ &= b_1^{(1)} \otimes [0.93126\, x \otimes x + 0.25763\, x \otimes y + 0.25763\, y \otimes x] \quad \text{for } \sigma = 1/10\end{aligned}$$

and $\|\mathbf{v} - \mathbf{u}^{(1)}\| = \sigma = \sigma_2^{(j)}$. The second projection needs the left-sided singular value decomposition of

$$\mathcal{M}_2(u^{(1)}) = x \otimes \left[b_1^{(1)} \otimes ((\alpha\varkappa + \beta\lambda)\, x + \beta\varkappa y)\right] + y \otimes \left[b_1^{(1)} \otimes \beta\varkappa x\right].$$

The singular values and left singular vectors for $\sigma = 1/10$ are

$$\begin{aligned}\sigma_1^{(2)} &= 0.99778, & b_1^{(2)} &= 0.96824\, x + 0.25\, y,\\ \sigma_2^{(2)} &= 0.066521, & b_2^{(2)} &= 0.25\, x - 0.96824\, y.\end{aligned}$$

$b_1^{(2)}$ is quite close to $b_1^{(1)} = 0.96885\, x + 0.24764\, y$. The second projection yields

$$\begin{aligned}\mathbf{u}^{(2)} &= b_1^{(1)} \otimes b_1^{(2)} \otimes [0.96609\, x + 0.24945\, y]\\ &= [0.96885\, x + 0.24764\, y] \otimes [0.96824\, x + 0.25\, y] \otimes [0.96609\, x + 0.24945\, y]\end{aligned}$$

with the error $\|\mathbf{u}^{(1)} - \mathbf{u}^{(2)}\| = \sigma_2^{(2)}$. Since $\mathrm{rank}_3(\mathbf{u}^{(2)}) = 1$, a third projection is not needed, i.e., $\tilde{\mathbf{u}}_{\mathrm{HOSVD}} := \mathbf{u}^{(2)}$. The total error is

$$\|\mathbf{v} - \tilde{\mathbf{u}}_{\mathrm{HOSVD}}\| = \sqrt{\left(\sigma_2^{(1)}\right)^2 + \left(\sigma_2^{(2)}\right)^2} = 0.12010.$$

One observes that $\tilde{\mathbf{u}}_{\mathrm{HOSVD}}$ is a bit better than $\mathbf{u}_{\mathrm{HOSVD}}$. However, as a consequence of the successive computations, the resulting tensor $\tilde{\mathbf{u}}_{\mathrm{HOSVD}}$ is not symmetric.

10.1.4 Other Truncations

Starting with an r-term representation $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_\nu^{(j)}$, the procedure from §8.3.3.2 allows an HOSVD representation in the hybrid format from §8.2.4, i.e., the coefficient tensor $\mathbf{a}$ of $\mathbf{v} \in \mathcal{T}_{\mathbf{r}}$ is represented in the r-term format $\mathcal{R}_r$.

To avoid the calculations from §8.3.3.2 for large r, there are proposals to simplify the truncation. In[5] [120], reduced singular value decompositions of the matrices $[v_1^{(j)},\ldots,v_r^{(j)}]$ are used to project $v_i^{(j)}$ onto a smaller subspace. For the correct scaling of $v_\nu^{(j)}$ define

$$\omega_\nu^{(j)} := \prod\nolimits_{k\neq j} \|v_\nu^{(k)}\|, \qquad A_j := \left[\omega_1^{(j)}v_1^{(j)},\cdots,\omega_r^{(j)}v_r^{(j)}\right] \in \mathbb{K}^{I_j\times r}.$$

The reduced left-sided singular value decomposition of $A_j = \sum_{i=1}^{s_j}\sigma_i^{(j)}u_i^{(j)}w_i^{(j)\mathsf{T}}$ ($s_j = \operatorname{rank}(A_j)$) yields $\sigma_i^{(j)}$ and $u_i^{(j)}$. Note that $s_j \ll r$ if $\#I_j \ll r$. Define the orthogonal projection $P_j^{(r_j)} = \sum_{i=1}^{r_j} u_i^{(j)}u_i^{(j)\mathsf{H}}$ from $\mathbb{K}^{I_j}$ onto $\operatorname{span}\{u_i^{(j)} : 1\leq i\leq r_j\}$ for some $r_j \leq s_j$. Application of $\mathbf{P_r} := \bigotimes_{j=1}^d P_j^{(r_j)}$ to $\mathbf{v}$ yields the truncated tensor

$$\tilde{\mathbf{v}} := \mathbf{P_r v} = \sum_{\nu=1}^{r}\bigotimes_{j=1}^{d}\left(P_j^{(r_j)}v_\nu^{(j)}\right) = \sum_{\nu=1}^{r}\bigotimes_{j=1}^{d}\sum_{i=1}^{r_j}\left\langle v_\nu^{(j)},u_i^{(j)}\right\rangle u_i^{(j)}.$$

The right-hand side is given in hybrid format (8.21). The error $\tilde{\mathbf{v}}-\mathbf{v}$ is caused by

$$d_\nu^{(j)} := \left(P_j^{(r_j)}-I\right)v_\nu^{(j)} = \sum_{i=r_j+1}^{s_j}\left\langle v_\nu^{(j)},u_i^{(j)}\right\rangle u_i^{(j)} = \sum_{i=r_j+1}^{s_j}\frac{\sigma_i^{(j)}w_{\nu,i}^{(j)}}{\omega_\nu^{(j)}}u_i^{(j)}.$$

The latter equality uses the singular value decomposition of A_j. The relative error introduced in (7.10c) is $\delta_\nu^{(j)} = \|d_\nu^{(j)}\|/\|v_\nu^{(j)}\|$. Note that $\omega_\nu^{(j)}\|v_\nu^{(j)}\| = \|\mathbf{v}_\nu\|$ with $\mathbf{v}_\nu = \bigotimes_{j=1}^d v_\nu^{(j)}$. Since $\{u_i^{(j)} : 1\leq i\leq s_j\}$ are orthonormal,

$$(\delta_\nu^{(j)})^2 = \|\mathbf{v}_\nu\|^{-2}\sum_{i=r_j+1}^{s_j}\left(\sigma_i^{(j)}\right)^2\left(w_{\nu,i}^{(j)}\right)^2$$

follows. Orthonormality of $w_i^{(j)}$ proves

$$\sum_{\nu=1}^{r}\|\mathbf{v}_\nu\|^2\sum_{j=1}^{d}(\delta_\nu^{(j)})^2 = \sum_{\nu=1}^{r}\sum_{j=1}^{d}\sum_{i=r_j+1}^{s_j}\left(\sigma_i^{(j)}\right)^2\left(w_{\nu,i}^{(j)}\right)^2 = \sum_{j=1}^{d}\sum_{i=r_j+1}^{s_j}\left(\sigma_i^{(j)}\right)^2.$$

Remark 10.6. Given a tolerance $\varepsilon>0$, choose the minimal j-rank $r_j\leq s_j$ such that $\sum_{j=1}^d\sum_{i=r_j+1}^{s_j}(\sigma_i^{(j)})^2\leq\varepsilon^2$. Then, the total error is bounded by

$$\|\tilde{\mathbf{v}}-\mathbf{v}\| \leq \sqrt{r}\,\varepsilon.$$

Proof. Apply Remark 7.11. □

Differently from Theorem 10.3, no comparison with the best approximation can be given.[6] Therefore, starting from a given error bound $\sqrt{r}\varepsilon$, the obtained reduced

[5] In [120, Theorem 2.5d], this approach is called ‘reduced HOSVD approximation’, although there is no similarity to HOSVD as defined in §8.3.

[6] As counterexample consider a tensor $\mathbf{v} = \mathbf{v}' + \varepsilon\mathbf{v}''$, where $\mathbf{v}'' = \mathbf{v}_n \in \mathcal{R}_2$ is taken from (9.10) with $n \gg 1/\varepsilon$, while $\mathbf{v}' \in \mathcal{R}_{r-2}$ has a stable representation. Together, $\mathbf{v}$ has an r-term

ranks r_j may be much larger than those obtained from HOSVD. In this case, the truncation $\mathbf{v} \mapsto \mathbf{P}\mathbf{v}$ can be followed by the ALS iteration from §10.3.

A favourable difference to the HOSVD projection is the fact that the projections P_j are determined independently.

10.2 Best Approximation in the Tensor Subspace Format

10.2.1 General Setting

As in §9.1, two approximation problems can be formulated. Let $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ be a Banach tensor space with norm $\|\cdot\|$. In the *first version* we fix the format $\mathcal{T}_\mathbf{r}$:

$$\begin{aligned} &\text{Given } \mathbf{v} \in \mathbf{V} \text{ and } \mathbf{r} = (r_1, \ldots, r_d) \in \mathbb{N}^d, \\ &\text{determine } \mathbf{u} \in \mathcal{T}_\mathbf{r} \text{ minimising } \|\mathbf{v} - \mathbf{u}\|. \end{aligned} \tag{10.12}$$

Again, we may form the infimum

$$\varepsilon(\mathbf{v}, \mathbf{r}) := \varepsilon(\mathbf{r}) := \inf \{\|\mathbf{v} - \mathbf{u}\| : \mathbf{u} \in \mathcal{T}_\mathbf{r}\}. \tag{10.13}$$

The variation over all $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ includes the variation over all subspaces $U_j \subset V_j$ of dimension r_j:

$$\varepsilon(\mathbf{v}, \mathbf{r}) = \inf_{\substack{U_1 \subset V_1 \text{ with} \\ \dim(U_1) = r_1}} \inf_{\substack{U_2 \subset V_2 \text{ with} \\ \dim(U_2) = r_2}} \cdots \inf_{\substack{U_d \subset V_d \text{ with} \\ \dim(U_d) = r_d}} \left\{ \inf_{\mathbf{u} \in \bigotimes_{j=1}^{d} U_j} \|\mathbf{v} - \mathbf{u}\| \right\}.$$

The existence of a best approximation $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ with $\|\mathbf{v} - \mathbf{u}\| = \varepsilon(\mathbf{v}, \mathbf{r})$ will be discussed in §10.2.2.2. Practical computations are usually restricted to the choice of the Euclidean norm (see §§10.2.2.3-4).

In the following *second variant* the rôles of $\mathbf{r}$ and $\varepsilon(\mathbf{r})$ are reversed:[7]

$$\begin{aligned} &\text{Given } \mathbf{v} \in \mathbf{V} \text{ and } \varepsilon > 0, \\ &\text{determine } \mathbf{u} \in \mathcal{T}_\mathbf{r} \text{ with } \|\mathbf{v} - \mathbf{u}\| \le \varepsilon \text{ and minimal storage size.} \end{aligned} \tag{10.14}$$

The following lemma is the analogue of Lemma 9.2 and can be proved similarly.

Lemma 10.7. *Assume that* $\mathbf{V}$ *is a Hilbert tensor space with induced scalar product. The best approximation from Problem (10.12) and at least one of the solutions of Problem (10.14) belong to the subspace* $\mathbf{U}(\mathbf{v}) := {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$ *(cf. (6.21)). Consequently, the statements from Lemma 9.2 are valid again.*

representation, where the two terms related to $\mathbf{v}_n$ are dominant and lead to the largest singular values $\sigma_i^{(j)}$. The projection described above omits parts of $\mathbf{v}'$, while $\varepsilon\mathbf{v}''$ is hardly changed. The ratio $\sigma_i^{(j)}/\sigma_1^{(j)} \cong \sigma_i^{(j)} n\varepsilon$ is not related to the relative error.

[7] In principle, we would like to ask for $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ with $\|\mathbf{v} - \mathbf{u}\| \le \varepsilon$ and $\mathbf{r}$ as small as possible, but this question may not lead to a unique $\mathbf{r}_{\min}$. The storage size of $\mathbf{u}$ is a scalar value depending of $\mathbf{r}$ and attains a minimum.

As an illustration, we discuss the Examples 8.26 and 8.27. The HOSVD projection (10.9) from Example 8.26 is already the best approximation. For Example 8.27 (with $\sigma = 1/10$) we make the symmetric ansatz $\mathbf{u}(\xi,\eta) := \otimes^3 (\xi\, x + \eta\, y)$. Minimisation of $\|\mathbf{v} - \mathbf{u}(\xi,\eta)\|$ over $\xi,\eta \in \mathbb{R}$ yields the optimum

$$\mathbf{u}_{\text{best}} := \otimes^3 (0.96756588\, x + 0.24968136\, y)\,, \qquad \|\mathbf{v} - \mathbf{u}_{\text{best}}\| = 0.120083,$$

which is only insignificantly better than $\|\mathbf{v} - \mathbf{u}_{\text{HOSVD}}\| = 0.120158$ from (10.11).

In §10.2.2 we shall analyse Problem (10.12), where the rank vector $\mathbf{r}$ is fixed. The second Problem (10.14) will be addressed in §10.3.3.

10.2.2 Approximation with Fixed Format

10.2.2.1 Matrix Case $d = 2$

The solution of Problem (10.12) is already discussed in Conclusion 2.32 for the Euclidean (Frobenius) norm. Let $r = r_1 = r_2 < \min\{n_1, n_2\}$. Determine the singular value decomposition $\sum_{i=1}^{\min\{n_1,n_2\}} \sigma_i u_i \otimes v_i$ of the tensor $\mathbf{v} \in \mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2}$. Then $B_1 = [u_1, \ldots, u_r]$ and $B_2 = [v_1, \ldots, v_r]$ contain the optimal orthonormal bases. The solution of Problem (10.12) is $\mathbf{u} = \sum_{i=1}^{r} \sigma_i u_i \otimes v_i$. The coefficient tensor is $\mathbf{a} = \mathbf{B}^{\mathsf{H}}\mathbf{v} = \operatorname{diag}\{\sigma_1, \ldots, \sigma_r\}$. The error $\|\mathbf{v} - \mathbf{u}\|$ equals $\sqrt{\sum_{i=r+1}^{\min\{n_1,n_2\}} \sigma_i^2}$ (cf. (2.26b)), while the maximised value $\|\mathbf{B}^{\mathsf{H}}\mathbf{v}\|$ is $\sqrt{\sum_{i=1}^{r} \sigma_i^2}$. Non-uniqueness occurs if $\sigma_r = \sigma_{r+1}$ (cf. Conclusion 2.32).

10.2.2.2 Existence of a Minimiser

The following assumptions hold in particular in the finite dimensional case.

Theorem 10.8. *Let $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j$ be a reflexive Banach tensor space with a norm not weaker than $\|\cdot\|_{\vee}$ (cf. (6.18)). Then the subset $\mathcal{T}_{\mathbf{r}} \subset \mathbf{V}$ is weakly closed. For any $\mathbf{v} \in \mathbf{V}$, Problem (10.12) has a solution, i.e., for given finite representation ranks $r_j \leq \dim(V_j)$ there are subspaces $U_j \subset V_j$ with $\dim(U_j) = r_j$ and a tensor $\mathbf{u}_{\min} \in \mathbf{U} = \bigotimes_{j=1}^{d} U_j$ such that*

$$\|\mathbf{v} - \mathbf{u}_{\min}\| = \inf_{\mathbf{u} \in \mathcal{T}_{\mathbf{r}}} \|\mathbf{v} - \mathbf{u}\|\,.$$

Proof. By Lemma 8.6, $\mathcal{T}_{\mathbf{r}}$ is weakly closed. Thus, Theorem 4.28 proves the existence of a minimiser. □

For (infinite dimensional) Hilbert spaces V_j, the statement of Theorem 10.8 is differently proved by Uschmajew [186, Corollary 23].

Concerning non-uniqueness of the best approximation, the observations for the matrix case $d = 2$ mentioned in §10.2.2.1 are still valid for larger d.

Remark 10.9. If $d \geq 2$ and $\dim(V_j) > r_j > 0$ for at least one $j \in \{0, \ldots, d\}$, uniqueness[8] of the minimiser $\mathbf{u}_{\min}$ cannot be guaranteed.

10.2.2.3 Optimisation with Respect to the Euclidean Norm

The Hilbert structure enables further characterisations. Concerning orthogonal projections we refer to §4.4.3.

Lemma 10.10. *(a) Given a fixed subspace* $\mathbf{U} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j$, *the minimiser of* $\|\mathbf{v} - \mathbf{u}\|$ *over all* $\mathbf{u} \in \mathbf{U}$ *is explicitly described by*

$$\mathbf{u} = P_{\mathbf{U}}\mathbf{v}, \tag{10.15a}$$

where $P_{\mathbf{U}}$ *is the orthogonal projection onto* $\mathbf{U}$. *Pythagoras' equality yields*

$$\|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v} - \mathbf{u}\|^2. \tag{10.15b}$$

(b) $P_{\mathbf{U}}$ *may be written as Kronecker product*

$$P_{\mathbf{U}} = \bigotimes_{j=1}^{d} P_{\overline{U_j}} \qquad (P_{\overline{U_j}} \text{ orthogonal projection onto } \overline{U_j}). \tag{10.15c}$$

Proof. 1) By definition of $\mathbf{U} := {}_{\|\cdot\|}\bigotimes_{j=1}^{d} U_j$, this subspace is closed and (10.15a) follows from Remark 4.122c. By Remark 4.122e, $I - P_{\mathbf{U}}$ is the orthogonal projection onto $\mathbf{U}^{\perp}$. Since $P_{\mathbf{U}}\mathbf{v} \in \mathbf{U}$ and $(I - P_{\mathbf{U}})\,\mathbf{v} \in \mathbf{U}^{\perp}$ are orthogonal,

$$\|\mathbf{v}\|^2 = \|P_{\mathbf{U}}\mathbf{v} + (I - P_{\mathbf{U}})\,\mathbf{v}\|^2 = \|P_{\mathbf{U}}\mathbf{v}\|^2 + \|\,(I - P_{\mathbf{U}})\,\mathbf{v}\|^2$$

follows. Now, $P_{\mathbf{U}}\mathbf{v} = \mathbf{u}$ and $(I - P_{\mathbf{U}})\,\mathbf{v} = \mathbf{v} - \mathbf{u}$ yield (10.15b).

2) (10.15c) is trivial. Note that $P_{\overline{U_j}}$ uses the closed subspace, since closeness of U_j is not yet assumed. □

Next, we consider the special case of finite dimensional $V_j = \mathbb{K}^{I_j}$, $n_j = \#I_j$, endowed with the Euclidean norm (and therefore also the Euclidean scalar product). Let $\mathbf{r} = (r_1, \ldots, r_d)$ be the prescribed dimensions and set $J_j := \{1, \ldots, r_j\}$. With each subspace U_j of dimension r_j we associate an orthonormal basis $B_j = [b_1^{(j)} \cdots b_{r_j}^{(j)}] \in \mathbb{K}^{I_j \times J_j}$. Then, $P_{U_j} = B_j B_j^{\mathsf{H}} \in \mathbb{K}^{I_j \times I_j}$ is the orthogonal projection onto[9] U_j (cf. Remark 4.122f). Using (10.15c), we obtain the representation

$$P_{\mathbf{U}} = \bigotimes_{j=1}^{d} B_j B_j^{\mathsf{H}} = \mathbf{B}\mathbf{B}^{\mathsf{H}} \qquad \text{with } \mathbf{B} := \bigotimes_{j=1}^{d} B_j \in \mathbb{K}^{\mathbf{I} \times \mathbf{J}}, \tag{10.16}$$

where $\mathbf{I} := I_1 \times \ldots \times I_d$ and $\mathbf{J} := J_1 \times \ldots \times J_d$.

[8] Since there are often misunderstandings we emphasise that uniqueness of $\mathbf{u}_{\min}$ is meant, not uniqueness of its representation by $\mathbf{a}[i_1 \cdots i_d]$ and $b_i^{(j)}$.

[9] Note that $U_j = \overline{U_j}$ because of the finite dimension.

Remark 10.11. Under the assumptions from above, the following minimisation problems are equivalent:

$$\min_{\mathbf{u}\in\mathcal{T}_{\mathbf{r}}} \|\mathbf{v}-\mathbf{u}\| = \min_{\mathbf{B}\in\mathbb{K}^{I_j\times J_j}} \left\{ \|\mathbf{v}-\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v}\| : \mathbf{B}=\bigotimes_{j=1}^{d} B_j,\ B_j^{\mathsf{H}}B_j=I \right\}. \tag{10.17}$$

Proof. Any $\mathbf{u}\in\mathcal{T}_{\mathbf{r}}$ belongs to some subspace $\mathbf{U}=\bigotimes_{j=1}^{d}U_j$ with $\dim(U_j)=r_j$; hence, $\mathbf{u}=\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v}$ holds for a suitable $\mathbf{B}$ proving $\min_{\mathbf{B}}\|\mathbf{v}-\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v}\|\leq\|\mathbf{v}-\mathbf{u}\|$. On the other hand, $\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v}$ belongs to $\mathcal{T}_{\mathbf{r}}$ so that $\min_{\mathbf{u}\in\mathcal{T}_{\mathbf{r}}}\|\mathbf{v}-\mathbf{u}\|\leq\|\mathbf{v}-\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v}\|$.□

Lemma 10.12. *The minimisation problem* $\mathbf{u}^*:=\arg\min_{\mathbf{u}\in\mathcal{T}_{\mathbf{r}}}\|\mathbf{v}-\mathbf{u}\|$ *is equivalent to the following maximisation problem:*[10]

$$\begin{aligned}&\textit{Find } \mathbf{B} \textit{ with } \mathbf{B}=\textstyle\bigotimes_{j=1}^{d}B_j,\ B_j\in\mathbb{K}^{I_j\times J_j},\ B_j^{\mathsf{H}}B_j=I,\\ &\textit{such that } \|\mathbf{B}^{\mathsf{H}}\mathbf{v}\| \textit{ is maximal.}\end{aligned} \tag{10.18}$$

If $\hat{\mathbf{B}}:=\arg\max_{\mathbf{B}}\|\mathbf{B}^{\mathsf{H}}\mathbf{v}\|$, *then* $\mathbf{u}^*=\hat{\mathbf{B}}\mathbf{a}$ *with* $\mathbf{a}:=\hat{\mathbf{B}}^{\mathsf{H}}\mathbf{v}\in\mathbb{K}^{\mathbf{J}}$. *If* $\mathbf{B}$ *is a solution of (10.18), also* $\mathbf{BQ}$ *with* $\mathbf{Q}=\bigotimes_{j=1}^{d}Q_j$ *and unitary* $Q_j\in\mathbb{K}^{J_j\times J_j}$, *is a solution.*

Proof. As a consequence of (10.15b), minimisation of $\|\mathbf{v}-\mathbf{u}\|$ is equivalent to the maximisation of $\|\mathbf{u}\|$. By Remark 10.11, $\mathbf{u}=\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v}$ holds for some orthogonal matrix $\mathbf{B}$ so that

$$\|\mathbf{u}\|^2=\langle\mathbf{u},\mathbf{u}\rangle=\left\langle\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v},\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v}\right\rangle=\left\langle\mathbf{B}^{\mathsf{H}}\mathbf{v},\mathbf{B}^{\mathsf{H}}\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{v}\right\rangle=\left\langle\mathbf{B}^{\mathsf{H}}\mathbf{v},\mathbf{B}^{\mathsf{H}}\mathbf{v}\right\rangle=\|\mathbf{B}^{\mathsf{H}}\mathbf{v}\|^2$$

(cf. Exercise 8.15). The last assertion follows from $\|\mathbf{B}^{\mathsf{H}}\mathbf{v}\|=\|\mathbf{Q}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{v}\|$. □

The reformulation (10.18), which is due to De Lathauwer-De Moor-Vandewalle [43, Theorem 4.2]), is the basis of the *ALS method* described in the next section.

10.3 Alternating Least-Squares Method (ALS)

10.3.1 Algorithm

Problem (10.18) is an optimisation problem, where the d parameters are orthogonal matrices $B_j\in\mathbb{K}^{I_j\times J_j}$. The function

$$\Phi(B_1,\ldots,B_d):=\|\mathbf{B}^{\mathsf{H}}\mathbf{v}\|^2$$

is a quadratic function of the B_j entries. As discussed in §9.5.2.1, a standard method for optimising multivariate functions is the iterative optimisation with respect to

[10] The orthogonal matrices B_j from (10.18) form the so-called Stiefel manifold.

a single parameter. In this case, we consider B_j as one parameter and obtain the following iteration (cf. De Lathauwer-De Moor-Vandevalle [43, Alg. 4.2], where it is called HOOI: higher-order orthogonal iteration. We use the term 'alternating least-squares method', although it is an alternating *largest*-squares method with the side conditions $B_j^{\mathsf{H}} B_j = I$.).

Start	Choose $B_j^{(0)} \in \mathbb{K}^{I_j \times J_j}$ $(1 \leq j \leq d)$ (cf. Remark 10.16c), set $m := 1$.
Loop	For $j = 1$ to d do compute $B_j^{(m)}$ as maximiser of $B_j^{(m)} := \underset{B_j \text{ with } B_j^{\mathsf{H}} B_j = I}{\operatorname{argmax}} \Phi(B_1^{(m)}, \ldots, B_{j-1}^{(m)}, B_j, B_{j+1}^{(m-1)}, \ldots, B_d^{(m-1)})$ (10.19) Set $m := m + 1$ and repeat the iteration.

The concrete realisation will be discussed in §10.3.2. Here, we give some general statements. Define $\mathbf{v}_{j,m} \in \mathbb{K}^{J_1 \times \ldots \times J_{j-1} \times I_j \times J_{j+1} \times \ldots \times J_d}$ by

$$\mathbf{v}_{j,m} := \left(B_1^{(m)} \otimes \ldots \otimes B_{j-1}^{(m)} \otimes id \otimes B_{j+1}^{(m-1)} \otimes \ldots \otimes B_d^{(m-1)} \right)^{\mathsf{H}} \mathbf{v} \qquad (10.20)$$

During the iteration (10.19) one is looking for an orthogonal matrix $B_j \in \mathbb{K}^{I_j \times J_j}$ so that $B_j^{\mathsf{H}} \mathbf{v}_{j,m}$ has maximal norm. Here and in the sequel, the short notation B_j, when applied to a tensor, means $id \otimes \ldots \otimes B_j \otimes \ldots \otimes id$.

Lemma 10.13. *The maximiser* $B_j = [b_1^{(j)} \cdots b_{r_j}^{(j)}] \in \mathbb{K}^{I_j \times J_j}$ *is given by the first* r_j *columns (singular vectors) of* U *in the reduced left-sided singular value decomposition* $\mathcal{M}_j(\mathbf{v}_{j,m}) = U \Sigma V^{\mathsf{T}}$. *Moreover,* $B_j B_j^{\mathsf{H}} = P_{j,\mathrm{HOSVD}}^{(r_j)}$ *is the HOSVD projection.*

Proof. The statements are easily derived by $\|B_j^{\mathsf{H}} \mathbf{v}_{j,m}\| \underset{\text{Remark 5.8}}{=} \|\mathcal{M}_j(B_j^{\mathsf{H}} \mathbf{v}_{j,m})\| = \underset{\text{Lemma 5.6}}{=} \|B_j^{\mathsf{H}} \mathcal{M}_j(\mathbf{v}_{j,m})\| = \|B_j^{\mathsf{H}} U \Sigma V^{\mathsf{T}}\| \underset{V \text{ unitary}}{=} \|B_j^{\mathsf{H}} U \Sigma\|$. □

Remark 10.14. (a) The construction of B_j requires $\operatorname{rank}(\mathcal{M}_j(\mathbf{v}_{j,m})) \geq r_j$, since, otherwise, U has not enough columns. In the latter case, one either adds arbitrarily chosen orthonormal vectors from $\operatorname{range}(U)^{\perp}$ or one continues with decreased j-th representation rank r_j.
(b) If $\operatorname{rank}(\mathcal{M}_j(\mathbf{v}_{j,m})) = r_j$, any orthonormal basis B_j of $\operatorname{range}(\mathcal{M}_j(\mathbf{v}_{j,m}))$ is the solution of (10.19).
(c) Note that initial values $B_j^{(0)}$ with $\operatorname{range}(B_j^{(0)}) \perp \operatorname{range}(\mathcal{M}_j(\mathbf{v}))$ for at least one $j \geq 2$ lead to $\mathbf{v}_{1,1} = 0$.

In the sequel, we assume that such failures of (10.19) do not appear. We introduce the index sets

$$I_j = \{1, \ldots, n_j\}, \; J_j = \{1, \ldots, r_j\}, \; \mathbf{I}_{[j]} = \bigtimes_{k \in \{1,\ldots,d\} \setminus j} I_k, \; \mathbf{J}_{[j]} = \bigtimes_{k \in \{1,\ldots,d\} \setminus \{j\}} J_k \quad (10.21)$$

and the tensor $\mathbf{B}_{[j]} := B_1^{(m)} \otimes \ldots \otimes B_{j-1}^{(m)} \otimes B_{j+1}^{(m-1)} \otimes \ldots \otimes B_d^{(m-1)} \in \mathbb{K}^{I_j \times \mathbf{J}_{[j]}}$. Remark 5.8 shows that

$$\mathcal{M}_j(\mathbf{v}_{j,m}) = \mathcal{M}_j(\mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{v}) = \mathcal{M}_j(\mathbf{v})\overline{\mathbf{B}_{[j]}}.$$

This proves the first part of the following remark.

Remark 10.15. Assume $\operatorname{rank}(\mathcal{M}_j(\mathbf{v}_{j,m})) \geq r_j$. Matrix U from Lemma 10.13 is obtained by diagonalising

$$\mathcal{M}_j(\mathbf{v}_{j,m})\mathcal{M}_j(\mathbf{v}_{j,m})^{\mathsf{H}} = \mathcal{M}_j(\mathbf{v})\overline{\mathbf{B}_{[j]}}\mathbf{B}_{[j]}^{\mathsf{T}}\mathcal{M}_j(\mathbf{v})^{\mathsf{H}} = U\Sigma^2U^{\mathsf{H}}.$$

All maximisers $B_j^{(m)}$ from (10.19) satisfy $\operatorname{range}(B_j^{(m)}) \subset U_j^{\min}(\mathbf{v})$.

Proof. Use $\operatorname{range}(B_j^{(m)}) \subset \operatorname{range}(\mathcal{M}_j(\mathbf{v})\overline{\mathbf{B}_{[j]}}) \subset \operatorname{range}(\mathcal{M}_j(\mathbf{v})) = U_j^{\min}(\mathbf{v})$. □

Remark 10.16. (a) The function values $\Phi(B_1^{(m)}, \ldots, B_j^{(m)}, B_{j+1}^{(m-1)}, \ldots, B_d^{(m-1)})$ increase weakly monotonously to a maximum of Φ. The sequence $B_j^{(m)}$ has a convergent subsequence.
(b) The determined maximum of Φ may be a local one.
(c) The better the starting values $B_j^{(0)}$ are, the better are the chances to obtain the global maximum of Φ. A good choice of $B_j^{(0)}$ can be obtained from the HOSVD projection $\mathbf{P}_{\mathbf{r}}^{\mathrm{HOSVD}} = \bigotimes_{j=1}^{d} B_j^{(0)} B_j^{(0)\mathsf{H}}$, denoted by $\mathbf{B}^{(\mathbf{r})}\mathbf{B}^{(\mathbf{r})\mathsf{H}}$ in Corollary 10.2.

For a detailed discussion of this and related methods we refer to De Lathauwer-De Moor-Vandevalle [43]. In particular, it turns out that the chance to obtain fast convergence to the global maximum is the greater the larger the gaps $\sigma_{r_j}^{(j)} - \sigma_{r_j+1}^{(j)}$ are.

10.3.2 ALS for Different Formats

The realisation described above involves $\mathcal{M}_j(\mathbf{v})\overline{\mathbf{B}_{[j]}} \in \mathbb{K}^{I_j \times \mathbf{J}_{[j]}}$ and its left-sided singular value decomposition. The corresponding computations depend on the format of $\mathbf{v}$. Note that the choice of the format is independent of the fact that the optimal solution $\mathbf{u}$ is sought in tensor subspace format $\mathcal{T}_{\mathbf{r}}$. We start with the case of the full tensor representation.

10.3.2.1 Full Format

The tensors $\mathbf{v}_{j,m}$ or equivalently their matricisations $\mathcal{M}_j(\mathbf{v}_{j,m})$ have to be determined.[11] The direct computation of the iterate $\mathbf{v}_{j,m} = \mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{v}$ from the tensor $\mathbf{v}$ costs $2\sum_{k=1}^{j-1}\prod_{\ell=1}^{k} r_\ell \prod_{\ell=k}^{d} n_\ell + 2\frac{n_j}{r_j}\sum_{k=j+1}^{d}\prod_{\ell=1}^{k} r_\ell \prod_{\ell=k}^{d} n_\ell$ operations. This number

[11] A precomputation of the Gram matrix $C := \overline{\mathbf{B}_{[j]}}\mathbf{B}_{[j]}^{\mathsf{T}}$ followed by the evaluation of the product $\mathcal{M}_j(\mathbf{v})C\mathcal{M}_j(\mathbf{v})^{\mathsf{H}} \in \mathbb{K}^{I_1 \times I_1}$ is more expensive.

is bounded by $2\sum_{k=1}^{j-1}\bar{r}^k n^{d-k+1} + 2\sum_{k=j+1}^{d}\bar{r}^{k-1}n^{d-k+2}$, where $n := \max n_j$ and $\bar{r} := \max r_j$. If $\bar{r} \ll n$, the leading term is $2r_1\prod_{\ell=1}^{d} n_\ell \leq 2\bar{r}n^d$.

Instead, one can determine $B_d^{(m-1)\mathsf{H}}\mathbf{v}$, $(B_{d-1}^{(m-1)}\otimes B_d^{(m-1)})^{\mathsf{H}}\mathbf{v}, \ldots$ at the expense of more memory. Note that the sizes of these tensors are decreasing. Having computed a new $B_1^{(m)}$, one can obtain $\mathbf{v}_{1,m}$ from $B_1^{(m)\mathsf{H}}(B_3^{(m-1)} \otimes \ldots \otimes B_d^{(m-1)})^{\mathsf{H}}\mathbf{v}$ etc. Using these data, we need

$$2\sum_{j=2}^{d}\left[\prod_{\ell=1}^{j} n_\ell\right]\left[\prod_{\ell=j}^{d} r_\ell\right] + 2\sum_{j=2}^{d}\sum_{k=1}^{j-1}\left[\prod_{\ell=1}^{k} r_\ell\right]\left[\prod_{\ell=k}^{j} n_\ell\right]\left[\prod_{\ell=j+1}^{d} r_\ell\right]$$

operations to determine *all* d tensors $\mathbf{v}_{1,m}, \mathbf{v}_{2,m}, \ldots, \mathbf{v}_{d,m}$.

As soon as $\mathbf{v}_{j,m}$ is determined, the computation of $\mathcal{M}_j(\mathbf{v}_{j,m})\mathcal{M}_j(\mathbf{v}_{j,m})^{\mathsf{H}}$ and its diagonalisation requires $n_j^2\prod_{\ell\neq j} r_\ell + \frac{8}{3}n_j^3$ operations.

We summarise the total cost per iteration $m \mapsto m+1$ for different ratios r/n.

(a) If $\bar{r} \ll n$, the leading cost is $4\bar{r}n^d$.

(b) If $\bar{r} < n$, the asymptotic cost is $n^d\bar{r}\big[4 + 6\frac{\bar{r}}{n} + (\frac{\bar{r}}{n})^{d-2} + O\big((\frac{\bar{r}}{n})^2\big)\big]$.

(c) If $\bar{r} \approx n$, so that $\bar{r} \leq n$ can be used, the leading bound is $(d^2 + 2d - 2)n^{d+1}$.

10.3.2.2 r-Term Format

Now $\mathbf{v} = \sum_{i=1}^{r}\bigotimes_{j=1}^{d} v_i^{(j)}$ is assumed. The projections $\mathbf{v} \mapsto B_j^{\mathsf{H}}\mathbf{v}$ can be performed independently for all j:

$$w_i^{(j)} := B_j^{\mathsf{H}} v_i^{(j)} \in \mathbb{K}^{J_j}$$

(cost: $2r\sum_{j=1}^{d} r_j n_j$). The projected iterate $\mathbf{v}_{j,m} = \mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{v}$ (cf. (10.20)) takes the form

$$\mathbf{v}_{j,m} = \sum_{i=1}^{r}\left(\bigotimes_{k=1}^{j-1} w_i^{(k)}\right)\otimes v_i^{(j)} \otimes \left(\bigotimes_{k=j+1}^{d} w_i^{(k)}\right).$$

Therefore, the computation of $\mathcal{M}_j(\mathbf{v}_{j,m})\mathcal{M}_j(\mathbf{v}_{j,m})^{\mathsf{H}} \in \mathbb{K}^{I_j\times I_j}$ requires similar Gram matrices G_k as in (8.34a), but now the entries are scalar products $\langle w_\nu^{(k)}, w_\mu^{(k)}\rangle$ in $\mathbb{K}^{J_j}$ instead of $\mathbb{K}^{I_j}$. Furthermore, the QR decomposition $[v_1^{(j)} \cdots v_r^{(j)}] = Q_j R_j$ with $\tilde{r}_j = \operatorname{rank}(Q_j R_j) \leq \min\{n_j, r\}$ (line 3 in (8.34a)) can be performed once for all. The matrix U_j from the diagonalisation $\mathcal{M}_j(\mathbf{v}_{j,m})\mathcal{M}_j(\mathbf{v}_{j,m})^{\mathsf{H}} = U_j\Sigma_j U_j^{\mathsf{H}}$ is obtained as in (8.34a). Repeating the proof of Remark 8.30, we obtain a total computational cost of

$$\sum_{j=1}^{d}\left(r^2 r_j + 2r^2\tilde{r}_j + r\tilde{r}_j^2 + \frac{8}{3}\tilde{r}_j^3 + 2r_j\tilde{r}_j n_j + 2r r_j n_j\right) \tag{10.22}$$

per iteration $m \mapsto m+1$. For $r \ll n := \max_j n_j$, $\bar{r} := \max_j r_j \ll n$, the dominating term is $4dr\bar{r}n$.

10.3.2.3 Tensor Subspace Format

We recall that the index sets J_j, $\mathbf{J}$, and $\mathbf{J}_{[j]}$ together with the representation ranks r_j are fixed in (10.21) and used for the format of the optimal solution $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ of Problem (10.12). For $\mathbf{v} \in \mathbf{V} = \mathbb{K}^\mathbf{I}$ we introduce representation ranks s_j and

$$\hat{J}_j = \{1, \ldots, s_j\}, \quad \hat{\mathbf{J}} = \hat{J}_1 \times \ldots \times \hat{J}_d, \quad \hat{\mathbf{J}}_{[j]} = \bigtimes_{k\in\{1,\ldots,d\}\backslash\{j\}} \hat{J}_k, \tag{10.23a}$$

and assume

$$\mathbf{v} = \rho_{\mathrm{orth}}\Big(\hat{\mathbf{a}}, (\hat{B}_j)_{j=1}^d\Big) = \hat{\mathbf{B}}\hat{\mathbf{a}} \quad \text{with} \begin{cases} \hat{\mathbf{a}} \in \mathbb{K}^{\hat{\mathbf{J}}},\ \hat{B}_j \in \mathbb{K}^{I_j \times \hat{J}_j}, \\ \hat{\mathbf{B}} = \bigotimes_{j=1}^d \hat{B}_j \in \mathbb{K}^{\mathbf{I}\times\hat{\mathbf{J}}} \end{cases} \tag{10.23b}$$

and I_j and $\mathbf{I}$ from (10.21). If $\mathbf{v}$ is given in the general format ρ_{frame}, one has first to orthonormalise the bases (cf. §8.2.3.2).

The bases from (10.23b) determine the spaces $U_j := \mathrm{range}(\hat{B}_j)$. According to Remark 10.16c, one should compute the HOSVD representation (and the corresponding truncation to $\mathcal{T}_\mathbf{r}$). If B_j is the HOSVD basis, even $U_j = U_j^{\min}(\mathbf{v})$ holds; otherwise, $U_j^{\min}(\mathbf{v}) \subset U_j$. We recall that the best approximation $\mathbf{u} = \mathbf{u}_{\mathrm{best}} \in \mathcal{T}_\mathbf{r}$ of Problem (10.12) belongs to $U_j^{\min}(\mathbf{v})$ (cf. Lemma 10.7):

$$\mathbf{u}_{\mathrm{best}} \in \mathbf{U}(\mathbf{v}) := \bigotimes_{j=1}^d U_j^{\min}(\mathbf{v}) \quad \text{and} \quad U_j^{\min}(\mathbf{v}) \subset U_j = \mathrm{range}(\hat{B}_j). \tag{10.23c}$$

The usual advantage of the tensor subspace format is that the major part of the computations can be performed using the smaller coefficient tensor $\hat{\mathbf{a}}$. This statement is also true for the best approximation in $\mathcal{T}_\mathbf{r}$. Because of (10.23c), there is a coefficient tensor $\hat{\mathbf{c}}_{\mathrm{best}}$ such that

$$\mathbf{u}_{\mathrm{best}} = \hat{\mathbf{B}}\,\hat{\mathbf{c}}_{\mathrm{best}}.$$

Orthonormality of the bases $\hat{B}_j$ ensures that

$$\|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\| = \|\hat{\mathbf{a}} - \hat{\mathbf{c}}_{\mathrm{best}}\|$$

holds for the Euclidean norms in $\mathbb{K}^\mathbf{I}$ and $\mathbb{K}^{\hat{\mathbf{J}}}$, respectively.

Proposition 10.17. *(a) Minimisation of* $\|\mathbf{v}-\mathbf{u}\|$ *over* $\mathcal{T}_\mathbf{r}(\mathbb{K}^\mathbf{I})$ *is equivalent to minimisation of* $\|\hat{\mathbf{a}} - \hat{\mathbf{c}}\|$ *over* $\mathcal{T}_\mathbf{r}(\mathbb{K}^{\hat{\mathbf{J}}})$. *If* $\hat{\mathbf{c}}_{\mathrm{best}}$ *is found,* $\mathbf{u}_{\mathrm{best}} := \hat{\mathbf{B}}\,\hat{\mathbf{c}}_{\mathrm{best}}$ *is the desired solution.*
(b) Let $\hat{\mathbf{c}}_{\mathrm{best}} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (\beta_j)_{j=1}^d\big)$ *with* $\mathbf{a} \in \mathbb{K}^\mathbf{J}$ *and* $\beta_j = [\beta_1^{(j)} \cdots \beta_{r_j}^{(j)}] \in \mathbb{K}^{\hat{J}_j \times J_j}$ *be the representation in* $\mathcal{T}_\mathbf{r}(\mathbb{K}^{\hat{\mathbf{J}}})$ *and set* $\boldsymbol{\beta} := \bigotimes_{j=1}^d \beta_j$. *Then*

$$\mathbf{u}_{\mathrm{best}} = \rho_{\mathrm{orth}}\big(\mathbf{a}, (B_j)_{j=1}^d\big) = \mathbf{B}\mathbf{a} \quad \textit{with} \begin{cases} \mathbf{a} \in \mathbb{K}^\mathbf{J},\ B_j := \hat{B}_j\beta_j \in \mathbb{K}^{I_j \times \hat{J}_j}, \\ \mathbf{B} := \hat{\mathbf{B}}\boldsymbol{\beta} \in \mathbb{K}^{\mathbf{I}\times\mathbf{J}} \end{cases}$$

is the orthonormal tensor subspace representation of $\mathbf{u}_{\mathrm{best}}$.

Proof. Part (b) follows by Remark 8.21. □

Application of the ALS iteration to $\hat{\mathbf{a}} \in \mathbb{K}^{\hat{\mathbf{J}}}$ requires the cost stated in §10.3.2.1 with n_j replaced by $s_j := \#\hat{J}_j$.

10.3.2.4 Hybrid Format

Proposition 10.17 holds again, but now $\hat{\mathbf{a}}$ is given in r-term format. The cost of one ALS iteration applied to $\hat{\mathbf{a}} \in \mathbb{K}^{\hat{\mathbf{J}}}$ is given by (10.22) with n_j replaced by $s_j := \#\hat{J}_j$.

10.3.2.5 Special Case $\mathbf{r} = (1, \ldots, 1)$

An interesting special case is given by

$$r_j = 1 \quad \text{for all } 1 \le j \le d, \qquad \text{i.e., } \mathbf{r} = (1, \ldots, 1),$$

since then

$$\max_{\mathbf{B}} \|\mathbf{B}^{\mathsf{H}}\mathbf{v}\| = \|\mathbf{v}\|_{\vee}$$

describes the injective crossnorm from §4.3.4 and §4.5.2. De Lathauwer-De Moor-Vandewalle [43] propose an iteration, which they call the *higher order power method*, since, for $d = 2$, it corresponds to the power method.

The basis $B_j \in \mathbb{K}^{I_j \times J_j}$ from (10.17) reduces to one vector $b^{(j)} := b_1^{(j)}$. The mapping

$$\mathbf{B}^{[j]} = b^{(1)\mathsf{H}} \otimes \ldots \otimes b^{(j-1)\mathsf{H}} \otimes id \otimes b^{(j-1)\mathsf{H}} \otimes \ldots \otimes b^{(d)\mathsf{H}} \in \mathcal{L}(\mathbf{V}, V_j)$$

acts on elementary vectors as

$$\mathbf{B}^{[j]} \left(\bigotimes_{k=1}^{d} v^{(k)} \right) = \left[\prod_{k \neq j} \left\langle v^{(k)}, b^{(k)} \right\rangle \right] v^{(j)}.$$

The higher order power method applied to $\mathbf{v} \in \mathbf{V}$ can be formulated as follows:

start	choose $b^{(j)}$, $1 \le j \le d$, with $\|b^{(j)}\| = 1$

iteration
$m = 1, 2, \ldots$
	for $j := 1$ to d do
	begin $b^{(j)} := \mathbf{B}^{[j]}(\mathbf{v})$; $\lambda := \|b^{(j)}\|$; $b^{(j)} := b^{(j)}/\lambda$ end;

return	$\mathbf{u} := \lambda \bigotimes_{j=1}^{d} b^{(j)} \in \mathcal{T}_{(1,\ldots,1)}$.

For further comments on this method see the article De Lathauwer-De Moor-Vandevalle [43, §3].

10.3.3 Approximation with Fixed Accuracy

Consider the tensor space $\mathbf{V} = \bigotimes_{j=1}^d V_j$ with $V_j = \mathbb{K}^{I_j}$, $n_j = \#I_j$, equipped with the Euclidean norm. A tensor from $\mathbf{V}$ given in the format $\mathbf{v} \in \mathcal{T}_\mathbf{s}$ with tensor subspace rank $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}_0^d$ requires storage of size

$$N_\mathbf{s} := \sum_{j=1}^d s_j n_j + \prod_{j=1}^d s_j \qquad \text{(cf. Remark 8.7a,b).}$$

An approximation $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ with $\mathbf{r} \lneqq \mathbf{s}$ leads to a reduced storage $N_\mathbf{r}$ (cf. Footnote 1). Given some $\varepsilon > 0$, there is a subset $R_\varepsilon \subset \mathbb{N}_0^d$ of smallest vectors $\mathbf{r} \in \mathbb{N}_0^d$ with the property $\min_{\mathbf{u} \in \mathcal{T}_\mathbf{r}} \|\mathbf{v} - \mathbf{u}\| \le \varepsilon$, i.e.,[12]

$$R_\varepsilon := \left\{ \mathbf{r} \in \mathbb{N}_0^d : \begin{array}{l} \mathbf{0} \le \mathbf{r} \le \mathbf{s},\ \min_{\mathbf{u} \in \mathcal{T}_\mathbf{r}} \|\mathbf{v} - \mathbf{u}\| \le \varepsilon, \text{ and} \\ \mathbf{r} = \mathbf{0} \text{ or } \min_{\mathbf{u} \in \mathcal{T}_\mathbf{s}} \|\mathbf{v} - \mathbf{u}\| > \varepsilon \text{ for all } \mathbf{0} \le \mathbf{s} \lneqq \mathbf{r} \end{array} \right\}.$$

Let $\mathbf{r}^*$ be the minimiser of $\min\{N_\mathbf{r} : \mathbf{r} \in R_\varepsilon\}$ and choose a minimiser[13] $\mathbf{u}^* \in \mathcal{T}_{\mathbf{r}^*}$ of $\min\{\|\mathbf{v} - \mathbf{u}\| : \mathbf{u} \in \mathcal{T}_{\mathbf{r}^*}\}$. Then, $\mathbf{u}^*$ is the solution of Problem (10.14). Since neither $\mathbf{r}^* \in \mathbb{N}_0^d$ nor $\mathbf{u}^* \in \mathcal{T}_{\mathbf{r}^*}$ are necessarily unique minimisers, and in particular because of the comment in Footnote 13, Problem (10.14) admits, in general, many solutions.

To obtain R_ε we need to know the minimal errors $\varepsilon_\mathbf{r} := \min_{\mathbf{u} \in \mathcal{T}_\mathbf{r}} \|\mathbf{v} - \mathbf{u}\|$. As seen in §10.3, the computation of $\varepsilon_\mathbf{r}$ is not a trivial task. Instead, we shall use a heuristic strategy which is again based on the higher order singular value decomposition.

First, one has to compute the HOSVD tensor subspace representation of $\mathbf{v}$. This includes the determination of the singular values $\sigma_i^{(j)}$ $(1 \le j \le d,\ 1 \le i \le s_j)$ and the corresponding basis vectors $b_i^{(j)}$. The reduction in memory size is the larger the more basis vectors $b_i^{(j)}$ can be omitted. More precisely, the saved storage

$$\Delta N_j(\mathbf{s}) := N_\mathbf{s} - N_{\mathbf{s}^{(j)}} = n_j + \prod_{k \ne j} s_k \text{ with } \mathbf{s}^{(j)} := (\ldots, s_{j-1}, s_j - 1, s_{j+1}, \ldots),$$

depends on the size of n_j and s. Consider the maximum over all $\Delta N_j(\mathbf{s})/(\sigma_{s_j}^{(j)})^2$, which is attained for some j^*. Dropping the component corresponding to $b_{s_{j^*}}^{(j^*)}$ yields the best decrease of storage cost. Hence, $\mathbf{v}$ is replaced by $\mathbf{u} := P_{j^*,\text{HOSVD}}^{(s_{j^*}-1)} \mathbf{v}$ (projection defined in Lemma 10.1). Note that $\mathbf{u} \in \mathcal{T}_\mathbf{r}$ for $\mathbf{r} := \mathbf{s}^{(j^*)}$. Since the values $\Delta N_j(\mathbf{r})$ for $j \ne j^*$ are smaller than $\Delta N_j(\mathbf{s})$, the singular values are now weighted by $\Delta N_j(\mathbf{r})/(\sigma_{r_j}^{(j)})^2$. Their maximiser j^* is used for the next projection

[12] The exceptional case $\mathbf{r} = \mathbf{0}$ occurs if $\|\mathbf{v}\| \le \varepsilon$. Then $\mathbf{u} = \mathbf{0} \in \mathcal{T}_\mathbf{0}$ is a sufficient approximation.

[13] To solve Problem (10.14), it suffices to take any $\mathbf{u}^* \in \mathcal{T}_{\mathbf{r}^*}$ with $\|\mathbf{v} - \mathbf{u}^*\| \le \varepsilon$. Among all possible $\mathbf{u}^*$ with this property, the minimiser is the most appreciated solution.

$\mathbf{u} := P^{(r_{j^*}-1)}_{j^*,\mathrm{HOSVD}}\mathbf{u}$. These reductions are repeated until the sum of the omitted squared singular values $\sigma^{(j^*)}_{r_{j^*}}$ does not exceed ε^2. The corresponding algorithm reads as follows:

```
start | $\mathbf{u} := \mathbf{v}$; $\mathbf{r} := \mathbf{s}$; $\tau := \varepsilon^2$, compute the HOSVD of $\mathbf{v}$                      | 1
loop  | $J := \{j \in \{1,\dots,d\} : (\sigma^{(j)}_{r_j})^2 < \tau\}$;                                      | 2
      | if $J = \emptyset$ then halt;                                                                        | 3
      | determine $\Delta N_j(\mathbf{r})$ for $j \in J$;                                                    | 4
      | $j^* := \mathrm{argmax}\{\Delta N_j(\mathbf{r})/(\sigma^{(j)}_{r_j})^2 : j \in J\}$;                 | 5
      | $\mathbf{u} := P^{(r_{j^*}-1)}_{j^*,\mathrm{HOSVD}}\mathbf{u}$; $\tau := \tau - (\sigma^{(j^*)}_{r_{j^*}})^2$; | 6
      | if $r_{j^*} > 1$ then set $\mathbf{r} := \mathbf{r}^{(j^*)}$ and repeat the loop                    | 7
```
(10.24)

In line 2, directions are selected for which a projection $P^{(r_j-1)}_{j,\mathrm{HOSVD}}$ yields an approximation $\mathbf{u}$ satisfying the requirement $\|\mathbf{v}-\mathbf{u}\| \le \varepsilon$. In line 6, the previous approximation $\mathbf{u} \in \mathcal{T}_{\mathbf{r}}$ is projected into $\mathbf{u} \in \mathcal{T}_{\mathbf{r}^{(j^*)}}$, where the j^*-th rank is reduced from r_{j^*} to $r_{j^*}-1$. If $r_{j^*}=1$ occurs in line 7, $\mathbf{u} = \mathbf{0}$ holds. The algorithm terminates with values of $\mathbf{r} \in \mathbb{N}_0^d$, $\mathbf{u} \in \mathcal{T}_{\mathbf{r}}$, and $\tau \ge 0$.

Thanks to estimate (10.4b), the inequality $\|\mathbf{v}-\mathbf{u}\|^2 \le \varepsilon^2 - \tau \le \varepsilon^2$ holds with τ determined by (10.24). However, as discussed in §10.1.3, the estimate (10.4b) may be too pessimistic. Since $\mathbf{v}-\mathbf{u} \perp \mathbf{u}$, the true error can be computed from

$$\|\mathbf{v}-\mathbf{u}\|^2 = \|\mathbf{v}\|^2 - \|\mathbf{u}\|^2$$

(note that $\|\mathbf{v}\|^2 = \sum_{i=1}^{s_j}(\sigma_i^{(j)})^2$ for any j). If a further reduction is wanted, algorithm (10.24) can be repeated with $\mathbf{u}$, $\mathbf{r}$ and $\varepsilon^2 - \|\mathbf{v}-\mathbf{u}\|^2$ instead of $\mathbf{v}, \mathbf{s}$ and ε^2.

The proposed algorithm requires only one HOSVD computation in (10.24). In principle, after getting a new approximation $\mathbf{u}$ in line 6, one may compute a new HOSVD. In that case, the accumulated squared error $\varepsilon^2 - \tau$ is the true error $\|\mathbf{v}-\mathbf{u}\|^2$ (cf. Theorem 10.5).

10.4 Analytical Approaches for the Tensor Subspace Approximation

In the following, we consider multivariate function spaces $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ and use interpolation of univariate functions from V_j. We may also replace functions $f \in V_j$ by grid functions $\hat{f} \in \mathbb{K}^{I_j}$ with the interpretation $\hat{f}_i = f(\xi_i)$ (ξ_i, $i \in I_j$: grid nodes). Then, all interpolation points appearing below must belong to the grid $\{\xi_i : i \in I_j\}$. Interpolation will map the functions into a fixed tensor subspace

$$\mathbf{U} = \bigotimes_{j=1}^d U_j \subset \mathbf{V}, \tag{10.25}$$

which is equipped with the norm of $\mathbf{V}$. Note that $\mathbf{U} \subset \mathcal{T}_{\mathbf{r}}$ with $\mathbf{r} = (r_1, \ldots, r_d)$.

As remarked at the beginning of §9.7, the following techniques can be used for practical constructions as well as for theoretical estimates of the best approximation error $\varepsilon(\mathbf{v}, \mathbf{r})$ from (10.13).

10.4.1 Linear Interpolation Techniques

10.4.1.1 Linear Interpolation Problem

Here, we omit the index j of the direction and rename r_j, U_j, V_j by r, U, V.

For $r \in \mathbb{N}_0$ fix a subspace $U \subset V$ of dimension r and define linear functionals $\Lambda_i \in V^*$ $(1 \le i \le r)$. Then the linear interpolation problem in V reads as follows:

$$\begin{aligned} &\text{Given } \lambda_i \in \mathbb{K}\ (1 \le i \le r), \quad \text{find } f \in U \text{ with} \\ &\quad \Lambda_i(f) = \lambda_i \qquad \text{for } 1 \le i \le r. \end{aligned} \tag{10.26a}$$

In most of the cases, Λ_i are Dirac functionals at certain interpolation points ξ_i, i.e., $\Lambda_i(f) = f(\xi_i)$. Then, the interpolation conditions for $f \in U$ become

$$f(\xi_i) = \lambda_i \qquad \text{for } 1 \le i \le r. \tag{10.26b}$$

The Dirac functionals are continuous in $C(I)$ or Hilbert spaces with sufficient smoothness (Sobolev embedding). In spaces like $L^2(I)$, other functionals Λ_i must be chosen for (10.26a).

Remark 10.18. (a) Problem (10.26a) is uniquely solvable for all $\lambda_i \in \mathbb{K}$, if and only if the functionals $\{\Lambda_i : 1 \le i \le r\}$ are linearly independent on U.
(b) In the positive case, there are so-called *Lagrange functions* L_i defined by $L_i \in U$ $(1 \le i \le r)$ and

$$\Lambda_\nu(L_\mu) = \delta_{\nu\mu} \qquad (1 \le \nu, \mu \le r). \tag{10.27a}$$

Then, problem (10.26a) has the solution

$$f = \sum_{i=1}^{r} \lambda_i L_i. \tag{10.27b}$$

In the following, we assume that the interpolation problem is solvable. The Lagrange functions define the interpolation operator $\mathcal{I} \in \mathcal{L}(V, U)$ by

$$\mathcal{I}(f) = \sum_{i=1}^{r} \Lambda_i(f) L_i \in U_j. \tag{10.28}$$

Remark 10.19. $\mathcal{I} \in \mathcal{L}(V, U)$ is a projection onto U. The norm $C_{\text{stab}} := \|\mathcal{I}_j\|_{V \leftarrow V}$ is called the *stability constant* of the interpolation $\mathcal{I}$.

The estimation of the interpolation error $f - \mathcal{I}(f)$ requires a Banach subspace $W \subset V$ with a stronger norm of $f \in W$ (e.g., $W = C^{p+1}(I) \subsetneqq V = C(I)$ in (10.35)).

10.4.1.2 Linear Product Interpolation

Let the function $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, \ldots, x_d)$ be defined on a product domain $\mathbf{I} := \times_{j=1}^{d} I_j$. For each direction j, we assume an interpolation operator

$$\mathcal{I}_j(f) = \sum_{i=1}^{r_j} \Lambda_i^{(j)}(f) L_i^{(j)}$$

involving functionals $\Lambda_i^{(j)}$ and Lagrange functions $L_i^{(j)}$. Interpolations points are denoted by $\xi_i^{(j)}$.

In the sequel, we assume that $\|\cdot\|$ is a uniform crossnorm. Then, the multivariate interpolation operator (product interpolation operator)

$$\mathcal{I} := \bigotimes_{j=1}^{d} \mathcal{I}_j : \mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^{d} V_j \to \mathbf{U} = \bigotimes_{j=1}^{d} U_j \subset \mathbf{V} \tag{10.29}$$

is bounded by $C_{\mathrm{stab}} := \prod_{j=1}^{d} C_{\mathrm{stab},j}$. Application of $\mathcal{I}$ to $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, \ldots, x_d)$ can be performed recursively. The following description refers to the Dirac functionals in (10.26b). Application of $\mathcal{I}_1$ yields $\mathbf{f}_{(1)}(\mathbf{x}) = \sum_{i_1=1}^{r_1} \mathbf{f}(\xi_{i_1}^{(1)}, x_2, \ldots, x_d) L_{i_1}^{(1)}(x_1)$. $\mathcal{I}_2$ maps into

$$\mathbf{f}_{(2)}(\mathbf{x}) = \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \mathbf{f}(\xi_{i_1}^{(1)}, \xi_{i_2}^{(2)}, x_2, \ldots, x_d) L_{i_1}^{(1)}(x_1) L_{i_2}^{(2)}(x_2).$$

After d steps the final result is reached:

$$\mathbf{f}_{(d)}(\mathbf{x}) = \mathcal{I}(\mathbf{f})(\mathbf{x}) = \sum_{i_1=1}^{r_1} \cdots \sum_{i_d=1}^{r_d} \mathbf{f}(\xi_{i_1}^{(1)}, \ldots, \xi_{i_d}^{(d)}) \prod_{j=1}^{d} L_{i_j}^{(j)}(x_j) \in \mathbf{U}.$$

Remark 10.20. Even the result $\mathbf{f}_{(d-1)}$ is or interest. The function

$$\mathbf{f}_{(d-1)}(\mathbf{x}) = \sum_{i_1=1}^{r_1} \cdots \sum_{i_{d-1}=1}^{r_d} \mathbf{f}(\xi_{i_1}^{(1)}, \ldots, \xi_{i_{d-1}}^{(d-1)}, x_d) \prod_{j=1}^{d-1} L_{i_j}^{(j)}(x_j)$$

belongs to $\left(\bigotimes_{j=1}^{d-1} U_j\right) \otimes V_d$. For fixed values $\{\xi_i^{(j)} : 1 \leq i \leq r_j, 1 \leq j \leq d-1\}$ the function $\mathbf{f}(\xi_{i_1}^{(1)}, \ldots, \xi_{i_{d-1}}^{(d-1)}, \bullet)$ is already univariate.

Interpolation is a special form of approximation. Error estimates for interpolation can be derived from best approximation errors.

Lemma 10.21. *Let $C_{\mathrm{stab},j}$ be the stability constant of $\mathcal{I}_j$ (cf. Remark 10.19). Then the interpolation error can be estimated by*

$$\|f-\mathcal{I}_j(f)\|_{V_j} \le (1+C_{\mathrm{stab},j})\inf\{\|f-g\|_{V_j} : g\in U_j\}.$$

Proof. Split the error into $[f-\mathcal{I}_j(g)]+[\mathcal{I}_j(g)-\mathcal{I}_j(f)]$ for $g\in U_j$ and note that $\mathcal{I}_j(g)=g$ because of the projection property. Taking the infimum over $g\in U_j$ in

$$\|f-\mathcal{I}_j(f)\|_{V_j} \le \|f-g\|_{V_j}+\|\mathcal{I}_j(g-f)\|_{V_j} \le (1+C_{\mathrm{stab},j})\,\|f-g\|_{V_j}\,,$$

we obtain the assertion. □

The multivariate interpolation error can be obtained from univariate errors as follows. Let

$$\varepsilon_j(\mathbf{f}) := \inf\left\{\|\mathbf{f}-\mathbf{g}\| : \mathbf{g}\in\left[\bigotimes_{k=1}^{j-1}V_j\right]\otimes U_j\otimes\left[\bigotimes_{k=j+1}^{d}V_j\right]\right\} \tag{10.30}$$

for $1\le j\le d$ be the best approximation error in j-th direction.

Proposition 10.22. *Let the norm of* $\mathbf{V}$ *be a uniform crossnorm. With* $\varepsilon_j(\mathbf{f})$ *and* $C_{\mathrm{stab},j}$ *from above, the interpolation error of* $\mathcal{I}$ *from (10.29) can be estimated by*

$$\|\mathbf{f}-\mathcal{I}(\mathbf{f})\| \le \sum_{j=1}^{d}\left[\prod_{k=1}^{j-1}C_{\mathrm{stab},k}\right](1+C_{\mathrm{stab},j})\;\varepsilon_j(\mathbf{f}).$$

Proof. Consider the construction of $\mathbf{f}_{(j)}$ from above with $\mathbf{f}_{(0)}:=\mathbf{f}$ and $\mathbf{f}_{(d)}=\mathcal{I}(\mathbf{f})$. The difference $\mathbf{f}_{(j-1)}-\mathbf{f}_{(j)}$ in $\mathbf{f}-\mathcal{I}(\mathbf{f})=\sum_{j=1}^{d}\left(\mathbf{f}_{(j-1)}-\mathbf{f}_{(j)}\right)$ can be rewritten as

$$\begin{aligned}\mathbf{f}_{(j-1)}-\mathbf{f}_{(j)} &= [\mathcal{I}_1\otimes\mathcal{I}_2\otimes\ldots\otimes\mathcal{I}_{j-1}\otimes(I-\mathcal{I}_j)\otimes id\otimes\ldots\otimes id]\,(\mathbf{f})\\ &= \left[\left(\bigotimes_{k=1}^{j-1}\mathcal{I}_k\right)\otimes\left(\bigotimes_{k=j}^{d}id\right)\right][id\otimes\ldots\otimes id\otimes(I-\mathcal{I}_j)\otimes id\otimes\ldots\otimes id]\,(\mathbf{f}).\end{aligned}$$

The norm of $[\ldots\otimes id\otimes(I-\mathcal{I}_j)\otimes id\otimes\ldots]\,(\mathbf{f})$ is bounded by $(1+C_{\mathrm{stab},j})\,\varepsilon_j(\mathbf{f})$ (cf. Lemma 10.21). The operator norm of the first factor is $\prod_{k=1}^{j-1}C_{\mathrm{stab},k}$ because of the uniform crossnorm property. □

10.4.1.3 Use of Transformations

We return to the univariate case of a function f defined on an interval I. Let $\phi: J\to I$ be a mapping from a possibly different interval J onto I and define

$$F := f\circ\phi. \tag{10.31a}$$

The purpose of the transformation ϕ is an improvement of the smoothness properties. For instance, ϕ may remove a singularity.[14] Applying an interpolation $\mathcal{I}_J$ with interpolation points $\xi_i^J\in J$ to F, we get

[14] $f(x)=\sqrt{\sin(x)}$ in $I=[0,1]$ and $x=\phi(y):=y^2$ yield $F(y)=\sqrt{\sin(y^2)}\in C^\infty$.

$$F(y) \approx (\mathcal{I}_J(F))\,(y) = \sum_{i=1}^{r} F(\xi_i^J) L_i^J(y). \tag{10.31b}$$

The error estimate of $F - \mathcal{I}_J(F)$ may exploit the improved smoothness of F. We can reinterpret this quadrature rule on J as a new quadrature rule on I:

$$f(x) \approx (\mathcal{I}_J(F))\,(\phi^{-1}(x)) = \sum_{i=1}^{r} F(\xi_i^J) L_i^J(\phi^{-1}(x)) = \mathcal{I}_I(f)(x) \tag{10.31c}$$

with $\mathcal{I}_I(f) := \sum_{i=1}^{r} \Lambda_i^I(f) L_i^I$, $\Lambda_i^I(f) := f(\zeta_i^I)$, $\zeta_i^I := \phi(\xi_i^J)$, $L_i^I := L_i^J \circ \phi^{-1}$.

Remark 10.23. (a) Since $(\mathcal{I}_J(f \circ \phi))\,(\phi^{-1}(\cdot)) = \mathcal{I}_I(f)$, the supremum norms of the errors coincide: $\|f - \mathcal{I}_I(f)\|_{I,\infty} = \|F - \mathcal{I}_J(F)\|_{J,\infty}$.
(b) While L_i^J may be standard functions like, e.g., polynomials, $L_i^I = L_i^J \circ \phi^{-1}$ are non-standard.

10.4.2 Polynomial Approximation

10.4.2.1 Notations

Let $\mathbf{I} = \times_{j=1}^{d} I_j$. Choose the Banach tensor space $C(\mathbf{I}) = {}_\infty\bigotimes_{j=1}^{d} C(I_j)$, i.e., $V_j = C(I_j)$. The subspaces $U_j \subset V_j$ are polynomial spaces $\mathcal{P}_{p_j}$, where

$$\mathcal{P}_p := \left\{ \sum\nolimits_{\nu=0}^{p} a_\nu x^\nu : a_\nu \in \mathbb{K} \right\}.$$

The tensor subspace $\mathbf{U}$ from (10.25) is built from $U_j = \mathcal{P}_{p_j}$:

$$\mathcal{P}_{\mathbf{p}} := \bigotimes\nolimits_{j=1}^{d} \mathcal{P}_{p_j} \subset C(\mathbf{I}) \qquad \text{with } \mathbf{p} = (p_1, \ldots, p_d).$$

Note that $\mathbf{U} = \mathcal{P}_{\mathbf{p}} \subset \mathcal{T}_{\mathbf{r}}$ requires $p_j \le r_j - 1$.

10.4.2.2 Approximation Error

The approximation error is the smaller the smoother the function is. Optimal smoothness conditions hold for analytic functions. For this purpose, we assume that a univariate function is analytic (holomorphic) in a certain ellipse.

$$E_{a,b} := \left\{ z \in \mathbb{C} : z = x + iy,\ \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1 \right\}$$

is the ellipse with half-axes a and b. In particular,

$$\mathcal{E}_\rho := E_{\frac{1}{2}(\rho+1/\rho),\frac{1}{2}(\rho-1/\rho)} \qquad \text{for } \rho > 1$$

is the unique ellipse with foci ±1 and ρ being the sum of the half-axes. The interior of $\mathcal{E}_\rho$ is denoted by $\mathring{\mathcal{E}}_\rho$. Note that the interval $[-1,1]$ is contained in $\mathring{\mathcal{E}}_\rho$ because of $\rho > 1$. $\mathcal{E}_\rho$ will be called *regularity ellipse*, since the functions to be approximated are assumed to be holomorphic in $\mathring{\mathcal{E}}_\rho$.

The main result is Bernstein's theorem [14] (proof in [46, Sect. 8, Chap. 7]).

Theorem 10.24 (Bernstein). *Let f be holomorphic and uniformly bounded in $\mathring{\mathcal{E}}_\rho$ with $\rho > 1$. Then, for any $p \in \mathbb{N}_0$ there is a polynomial P_p of degree $\leq p$ such that*

$$\|f - P_p\|_{[-1,1],\infty} \leq \frac{2\rho^{-p}}{\rho - 1} \|f\|_{\mathring{\mathcal{E}}_\rho,\infty}. \tag{10.32}$$

A general real interval $[x_1, x_2]$ with $x_1 < x_2$ is mapped by $\Phi(z) := \frac{2(z-x_1)}{x_2-x_1} - 1$ onto $[-1,1]$. We set

$$\begin{aligned}&\mathcal{E}_\rho([x_1,x_2]) := \Phi^{-1}(\mathcal{E}_\rho)\\ &= \left\{ z \in \mathbb{C} :\ z = x + iy,\ \frac{\left(x - \frac{x_1+x_2}{2}\right)^2}{(\rho + 1/\rho)^2} + \frac{y^2}{(\rho - 1/\rho)^2} \leq \left(\frac{x_2 - x_1}{4}\right)^2 \right\}.\end{aligned}$$

Corollary 10.25. Assume that a function f defined on $I = [x_1, x_2]$ can be extended holomorphically onto $\mathring{\mathcal{E}}_\rho(I)$ with $M := \sup\{|f(z)| : z \in \mathring{\mathcal{E}}_\rho([x_1,x_2])\}$. Then, for any $p \in \mathbb{N}_0$ there is a polynomial P_p of degree $\leq p$ such that

$$\|f - P_p\|_{I,\infty} \leq \frac{2\rho^{-p}}{\rho - 1} M. \tag{10.33}$$

The next statement exploits only properties of f on a real interval (proof in Melenk-Börm-Löhndorf [146]).

Lemma 10.26. *Let f be an analytical function defined on the interval $I \subset \mathbb{R}$ of length L. Assume that there are constants $C, \gamma \geq 0$ such that*

$$\left\| \frac{\mathrm{d}^n}{\mathrm{d}x^n} f \right\|_{I,\infty} \leq C\, n!\, \gamma^n \qquad \textit{for all } n \in \mathbb{N}_0. \tag{10.34a}$$

Then, for any $p \in \mathbb{N}_0$ there is a polynomial P_p of degree $\leq p$ such that

$$\|f - P_p\|_{I,\infty} \leq 4e\, C\, (1 + \gamma L)\, (p+1) \left(1 + \frac{2}{\gamma L}\right)^{-(p+1)}. \tag{10.34b}$$

Corollary 10.27. With the notations from §10.4.2.1 assume that $\mathbf{f} \in C(\mathbf{I})$ is analytic in all arguments. Then the best approximation error $\varepsilon_j(\mathbf{f})$ from (10.30) can be estimated by

$$\varepsilon_j(\mathbf{f}) \leq \frac{2M_j}{\rho_j - 1} \rho_j^{-p_j} \qquad (\rho_j > 1,\ 1 \leq j \leq d),$$

if for all $x_k \in I_k$ $(k \neq j)$, the univariate function $\mathbf{f}(x_1, \ldots, x_{j-1}, \bullet, x_{j+1}, \ldots, x_d) \in C(I_j)$ satisfies the conditions of Corollary 10.25. The estimate

$$\varepsilon_j(\mathbf{f}) \leq C' \, (p+1) \, \rho_j^{-p_j} \qquad \text{with } \rho_j := 1 + \frac{2}{\gamma L}$$

holds, if $\mathbf{f}(x_1, \ldots, x_{j-1}, \bullet, x_{j+1}, \ldots, x_d)$ fulfils the inequalities (10.34a). In both cases, the bound of $\varepsilon_j(\mathbf{f})$ decays exponentially like $O(\rho_j^{-p_j})$ as $p_j \to \infty$.

10.4.3 Polynomial Interpolation

10.4.3.1 Univariate Interpolation

The univariate interpolation by polynomials is characterised by the interval $I = [a, b]$, the degree p, and the quadrature points $(\xi_i)_{i=0}^{p} \subset I$. The interval can be standardised to $[-1, 1]$. An interpolation operator $\mathcal{I}_{[-1,1]}$ on $[-1, 1]$ with quadrature points $(\xi_i)_{i=0}^{p} \subset [-1, 1]$ can be transferred to an interpolation operator $\mathcal{I}_{[a,b]}$ on $[a, b]$ with quadrature points $(\Phi(\xi_i))_{i=0}^{p}$, where $\Phi : [-1, 1] \to [a, b]$ is the affine mapping $\Phi(x) = a + \frac{1}{2}(b-a)(x+1)$. The interpolating polynomials satisfy $\mathcal{I}_{[a,b]}(f) = \mathcal{I}_{[-1,1]}(f \circ \Phi)$.

The Lagrange functions L_ν from (10.27a) are the *Lagrange polynomials*

$$L_\nu(x) = \prod_{i \in \{0,\ldots,p\} \setminus \{\nu\}} \frac{x - \xi_i}{\xi_\nu - \xi_i}.$$

They satisfy $L_\nu^{[a,b]} = L_\nu^{[-1,1]} \circ \Phi^{-1}$.

A well-known interpolation error estimate holds for functions $f \in C^{p+1}(I)$:

$$\|f - \mathcal{I}(f)\|_\infty \leq \frac{C_\omega(\mathcal{I})}{(p+1)!} \|f^{(p+1)}\|_\infty \text{ with } C_\omega(\mathcal{I}) := \left\| \prod_{i=0}^{p} (x - \xi_i) \right\|_{I,\infty}. \tag{10.35}$$

The natural Banach space is $V = \big(C(I), \|\cdot\|_{I,\infty}\big)$.

Remark 10.28. (a) If $\mathcal{I}_{[a,b]}(f) = \mathcal{I}_{[-1,1]}(f \circ \Phi)$ are polynomial interpolations of degree p, then $C_\omega(\mathcal{I}_{[a,b]}) = C_\omega(\mathcal{I}_{[-1,1]})(\frac{b-a}{2})^{p+1}$.
(b) The stability constant $C_{\text{stab}} = \|\mathcal{I}_{[a,b]}\|_{V \leftarrow V}$ does not depend on the interval $[a, b]$.

The smallest constant $C_\omega(\mathcal{I}_{[-1,1]})$ is obtained for the so-called Chebyshev interpolation using the Chebyshev quadrature points

$$\xi_i = \cos\left(\frac{i + 1/2}{p+1}\pi\right) \in [-1, 1] \qquad (i = 0, \ldots, p),$$

which are the zeros of the $(p+1)$-th Chebyshev polynomial T_{p+1}.

Remark 10.29 ([164]). The Chebyshev interpolation of polynomial degree p leads to the constants

$$C_\omega(\mathcal{I}_{[-1,1]}) = 2^{-p-1} \quad \text{and} \quad C_{\text{stab}} \leq 1 + \frac{2}{\pi} \log(p+1).$$

10.4.3.2 Product Interpolation

Given polynomial interpolation operators $\mathcal{I}_j$ of degree p_j on intervals I_j, the product interpolation is the tensor product

$$\mathcal{I} := \bigotimes_{j=1}^{d} \mathcal{I}_j : C(\mathbf{I}) \to \mathcal{P}_{\mathbf{p}} .$$

Under the conditions of Corollary 10.27, the approximation error $\varepsilon_j(\mathbf{f})$ from (10.30) decays exponentially: $\varepsilon_j(\mathbf{f}) \leq O(\rho_j^{-p_j})$. Hence, Proposition 10.22 yields the result

$$\|\mathbf{f} - \mathcal{I}(\mathbf{f})\| \leq \sum_{j=1}^{d} \left[\prod_{k=1}^{j-1} C_{\mathrm{stab},k}\right] (1 + C_{\mathrm{stab},j})\ \varepsilon_j(\mathbf{f}) \leq O(\max_j \rho_j^{-p_j}).$$

For Chebyshev interpolation, the stability constants $C_{\mathrm{stab},j} \leq 1 + \frac{2}{\pi}\log(p_j + 1)$ depend only very weakly on the polynomial degree p_j.

10.4.4 Sinc Approximations

The following facts are mainly taken from the monograph of Stenger [177].

10.4.4.1 Sinc Functions, Sinc Interpolation

The sinc function $\operatorname{sinc}(x) := \frac{\sin(\pi x)}{\pi x}$ is holomorphic in $\mathbb{C}$. Fixing a step size $h > 0$, we obtain a family of scaled and shifted functions

$$S(k,h)(x) := \operatorname{sinc}\Big(\frac{x}{h} - k\Big) = \frac{\sin\left[\pi(x - kh)/h\right]}{\pi\left(x - kh\right)/h} \quad (h > 0,\ k \in \mathbb{Z}). \tag{10.36}$$

Note that $S(k,h)$ is a function in x with two parameters k, h.

Remark 10.30. The entire function $S(k,h)$ satisfies $S(k,h)(\ell h) = \delta_{k,\ell}$ for all $\ell \in \mathbb{Z}$.

Because of Remark 10.30, $S(k,h)$ can be viewed as Lagrange function L_k corresponding to infinite many interpolation points $\{kh : k \in \mathbb{Z}\}$. This leads to the following definition.

Definition 10.31 (sinc interpolation). Let $f \in C(\mathbb{R})$ and $N \in \mathbb{N}_0$. The sinc interpolation in $2N+1$ points $\{kh : k \in \mathbb{Z},\ |k| \leq N\}$ is denoted by[15]

[15] Only for the sake of convenience we consider the sum $\sum_{k=-N}^{N} \cdots$ One may use instead $\sum_{k=-N_1}^{N_2}$, where N_1 and N_2 are adapted to the behaviour at $-\infty$ and $+\infty$, separately.

$$C_N(f,h) := \sum_{k=-N}^{N} f(kh)S(k,h). \tag{10.37a}$$

If the limit exists for $N \to \infty$, we write

$$C(f,h) := \sum_{k=-\infty}^{\infty} f(kh)S(k,h). \tag{10.37b}$$

The corresponding interpolation errors are

$$E_N(f,h) := f - C_N(f,h), \qquad E(f,h) := f - C(f,h). \tag{10.37c}$$

Lemma 10.32. *The stability constant in* $\|C_N(f,h)\|_\infty \le C_{\mathrm{stab}}(N)\,\|f\|_\infty$ *for all* $f \in C(\mathbb{R})$ *satisfies*

$$C_{\mathrm{stab}}(N) \le \frac{2}{\pi}\,(3 + \log(N)) \qquad \textit{(cf. [177, p. 142]).}$$

Under strong conditions on f, it coincides with $C(f,h)$ (cf. [177, (1.10.3)]). Usually, there is an error $E(f,h)$, which will be estimated in (10.40). The speed, by which $f(x)$ tends to zero as $\mathbb{R} \ni x \to \pm\infty$, determines the error estimate of $C(f,h) - C_N(f,h) = E_N(f,h) - E(f,h)$ (cf. Lemma 10.34), so that, finally, $E_N(f,h)$ can be estimated.

The error estimates are based on the fact that f can be extended analytically from $\mathbb{R}$ to a complex stripe $\mathfrak{D}_\delta$ satisfying $\mathbb{R} \subset \mathfrak{D}_\delta \subset \mathbb{C}$:

$$\mathfrak{D}_\delta := \{z \in \mathbb{C} : |\Im m\, z| < \delta\} \qquad (\delta > 0)\,. \tag{10.38}$$

Definition 10.33. For $\delta > 0$ and f holomorphic in $\mathfrak{D}_\delta$, define

$$\|f\|_{\mathfrak{D}_\delta} = \int_{\partial\mathfrak{D}_\delta} |f(z)|\,|\mathrm{d}z| = \int_{\mathbb{R}} (|f(x+i\delta)| + |f(x-i\delta)|)\,\mathrm{d}x \tag{10.39}$$

(set $\|f\|_{\mathfrak{D}_\delta} = \infty$, if the integral does not exist). Then, a Banach space is given by

$$\mathbf{H}(\mathfrak{D}_\delta) := \{f \text{ is holomorphic in } \mathfrak{D}_\delta \text{ and } \|f\|_{\mathfrak{D}_\delta} < \infty\}.$$

The residual theorem allows to represent the interpolation error exactly:

$$E(f,h)(z) = \frac{\sin(\pi z/h)}{2\pi i} \int_{\partial\mathfrak{D}_\delta} \frac{f(\zeta)}{(\zeta - z)\sin(\pi\zeta/h)}\mathrm{d}\zeta \qquad \text{for all } z \in \mathfrak{D}_\delta$$

(cf. [177, Thm 3.1.2]). Estimates with respect to the supremum norm $\|\cdot\|_{\mathbb{R},\infty}$ or $L^2(\mathbb{R})$ norm have the form[16]

[16] For a proof and further estimates in $L^2(\mathbb{R})$ compare [177, (3.1.12)] or [86, §D.2.3].

$$\|E(f,h)\|_\infty \le \frac{\|f\|_{\mathfrak{D}_\delta}}{2\pi\delta\,\sinh(\frac{\pi\delta}{h})} \tag{10.40}$$

Note that $\frac{1}{\sinh(\pi\delta/h)} \le 2\exp(\frac{-\pi\delta}{h})$ decays exponentially as $h \to 0$.

While $E(f,h)$ depends on $\|f\|_{\mathfrak{D}_\delta}$ and therefore on the behaviour of f in the complex plane, the difference $E(f,h) - E_N(f,h)$ hinges on decay properties of f on the real axis alone.

Lemma 10.34. *Assume that for $f \in \mathbf{H}(\mathfrak{D}_\delta)$ there are some $c \ge 0$ and $\alpha > 0$ such that*

$$|f(x)| \le c \cdot \mathrm{e}^{-\alpha|x|} \qquad \textit{for all } x \in \mathbb{R}. \tag{10.41}$$

Then, the difference $E(f,h) - E_N(f,h) = \sum_{|k|>N} f(kh)S(k,h)$ can be bounded by

$$\|E(f,h) - E_N(f,h)\|_\infty \le \frac{2c}{\alpha h}\mathrm{e}^{-\alpha N h}. \tag{10.42}$$

Proof. Since $E(f,h) - E_N(f,h) = \sum_{|k|>N} f(kh)S(k,h)$ and $\|S(k,h)\|_\infty \le 1$, the sum $\sum_{|k|>N} |f(kh)|$ can be estimated using (10.41). □

To bound $\|E_N(f,h)\|_\infty \le \|E(f,h)\|_\infty + \|E(f,h) - E_N(f,h)\|_\infty$ optimally, the step width h has to be chosen such that both terms are balanced.

Theorem 10.35. *Let $f \in \mathbf{H}(\mathfrak{D}_\delta)$ satisfy (10.41). Choose h by*

$$h := h_N := \sqrt{\frac{\pi\delta}{\alpha N}}. \tag{10.43}$$

Then the interpolation error is bounded by

$$\|E_N(f,h)\|_\infty \le \sqrt{\tfrac{N}{\delta}}\,\mathrm{e}^{-\sqrt{\pi\alpha\delta N}} \left[\frac{\|f\|_{\mathfrak{D}_\delta}}{\pi\left[1-\mathrm{e}^{-\pi\alpha\delta N}\right]\sqrt{N\delta}} + \frac{2c}{\sqrt{\pi\alpha}} \right]. \tag{10.44}$$

The right-hand side in (10.44) behaves like $\|E_N(f,h)\|_\infty \le O(\exp\{-C\sqrt{N}\})$.

Proof. Combine (10.40) and (10.42) with h from (10.43). □

Inequality (10.40) implies that, given an $\varepsilon > 0$, accuracy $\|E_N(f,h_N)\|_\infty \le \varepsilon$ holds for $N \ge C^{-2}\log^2(\frac{1}{\varepsilon}) + \mathcal{O}(\log\frac{1}{\varepsilon})$.

Corollary 10.36. A stronger decay than in (10.41) holds, if

$$|f(x)| \le c \cdot \mathrm{e}^{-\alpha\,|x|^\gamma} \qquad \text{for all } x \in \mathbb{R} \text{ and some } \gamma > 1. \tag{10.45}$$

Instead of (10.42), the latter condition implies that

$$\|E(f,h) - E_N(f,h)\|_\infty \le \frac{2c}{\alpha N^{\gamma-1}h^\gamma}\exp(-\alpha\,(Nh)^\gamma). \tag{10.46}$$

The optimal step size $h := h_N := \left(\frac{\pi\delta}{\alpha}\right)^{1/(\gamma+1)} N^{-\gamma/(\gamma+1)}$ leads to

$$\|E_N(f,h_N)\|_\infty \le O\Big(\mathrm{e}^{-C\,N^{\frac{\gamma}{\gamma+1}}}\Big) \qquad \text{for all } 0 < C < \alpha^{\frac{1}{\gamma+1}}\,(\pi\delta)^{\frac{\gamma}{\gamma+1}}\,. \tag{10.47}$$

To reach accuracy ε, the number N must be chosen $\ge \left(C^{-1}\log(1/\varepsilon)\right)^{(\gamma+1)/\gamma}$.

Proof. See [86, Satz D.2.11]. □

For increasing γ, the right-hand side in (10.47) approaches $O(\mathrm{e}^{-CN})$ as attained in Theorem 9.29 for a compact interval. A bound quite close to $O(\mathrm{e}^{-CN})$ can be obtained when f decays doubly exponentially.

Corollary 10.37. Assume that for $f \in \mathbf{H}(\mathfrak{D}_\delta)$ there are constants $c_1, c_2, c_3 > 0$ such that

$$|f(x)| \le c_1 \cdot \exp\{-c_2\mathrm{e}^{c_3|x|}\} \qquad \text{for all } x \in \mathbb{R}. \tag{10.48}$$

Then

$$\|E(f,h) - E_N(f,h)\|_\infty \le \frac{2c_1}{c_2c_3}\exp\big(-c_2\mathrm{e}^{c_3Nh}\big)\frac{\mathrm{e}^{-c_3Nh}}{h}. \tag{10.49}$$

Choosing $h := h_N := \frac{\log N}{c_3N}$, we obtain

$$\|E_N(f,h)\|_\infty \le C\exp\Big(\frac{-\pi\delta c_3N}{\log N}\Big) \quad \text{with } C \to \frac{\|f\|_{\mathfrak{D}_\delta}}{2\pi\delta} \text{ for } N \to \infty. \tag{10.50}$$

Accuracy $\varepsilon > 0$ follows from $N \ge C_\varepsilon\left(\log\frac{1}{\varepsilon}\right)\cdot\log\left(\log\frac{1}{\varepsilon}\right)$ with $C_\varepsilon = \frac{1+o(1)}{\pi\delta\,c_3}$ as $\varepsilon \to 0$.

Proof. See [86, Satz D.2.13]. □

10.4.4.2 Transformations and Weight Functions

As mentioned in §10.4.1.3, a transformation $\phi : J \to I$ may improve the smoothness of the function. Here, the transformation has another reason. Since the sinc interpolation is performed on $\mathbb{R}$, a function f defined on I must be transformed into $F := f \circ \phi$ for some $\phi : \mathbb{R} \to I$. Even if $I = \mathbb{R}$, a further substitution by $\phi : \mathbb{R} \to \mathbb{R}$ may lead to a faster decay of $|f(x)|$ as $|x| \to \infty$. We give some examples:

	I	transformations $x = \phi(\zeta)$
(a)	$(0,1]$	$\phi(\zeta) = \frac{1}{\cosh(\zeta)}$ or $\frac{1}{\cosh(\sinh(\zeta))}$ (cf. [112])
(b)	$[1,\infty)$	$\phi(\zeta) = \cosh(\zeta)$ or $\cosh(\sinh(\zeta))$
(c)	$(0,\infty)$	$\phi(\zeta) = \exp(\zeta)$
(d)	$(-\infty,\infty)$	$\phi(\zeta) = \sinh(\zeta)$

One has to check carefully, whether $F := f \circ \phi$ belongs to $\mathbf{H}(\mathfrak{D}_\delta)$ for a positive δ. The stronger the decay on the real axis is, the faster is the increase in the imaginary axis. The second transformations in the lines (a) and (b) are attempts to reach the doubly exponential decay from Corollary 10.37. The transformation from line (d) may improve the decay.

Let f be defined on I. To study the behaviour of $F(\zeta) = f(\phi(\zeta))$ for $\zeta \to \pm\infty$, it is necessary to have a look at the values of f at the end points of I. Here, different cases are to be distinguished.

Case A1: Assume that f is defined on $(0,1]$ with $f(x) \to 0$ as $x \to 0$. Then $F(\zeta) = f(1/\cosh(\zeta))$ decays even faster to zero as $\zeta \to \pm\infty$. Note that the boundary value $f(1)$ is arbitrary. In particular, $f(1) = 0$ is not needed. Furthermore, since $F(\zeta)$ is an even function, the interpolation $C_N(F,h) = \sum_{k=-N}^{N} F(kh)S(k,h)$ from (10.37a) can be simplified to $C_N(F,h) = F(0)S(0,h) + 2\sum_{k=1}^{N} F(kh)S(k,h)$, which, formulated with $f = F \circ \phi^{-1}$, yields the new interpolation scheme

$$\hat{C}_N(f,h)(x) = f(1)L_0(x) + 2\sum_{k=1}^{N} f(\tfrac{1}{\cosh(kh)})L_k(x)$$

with $L_k(x) := S(k,h)(\mathrm{Arcosh}(\frac{1}{x}))$. Area hyperbolic cosine Arcosh is the inverse of $\cosh$. Note that $\hat{C}_N$ involves only $N+1$ interpolation points $\xi_k = 1/\cosh(kh)$.

Case A2: Take f as above, but assume that $f(x) \to c \neq 0$ as $x \to 0$. Then $F(\zeta) \to c$ as $\zeta \to \pm\infty$ shows that F fails to fulfil the desired decay to zero.

Case B: Let f be defined on $(0,\infty)$ and set $F(\zeta) := f(\exp(\zeta))$. Here, we have to require $f(x) \to 0$ for $x \to 0$ as well as for $x \to \infty$. Otherwise, F fails as in Case A2.

In the following, we take Case A2 as model problem and choose a weight function $\omega(x)$ with $\omega(x) > 0$ for $x > 0$ and $\omega(x) \to 0$ as $x \to 0$. Then, obviously, $f_\omega := \omega \cdot f$ has the correct zero limit at $x = 0$. Applying the interpolation to $F_\omega(\zeta) := (\omega \cdot f)(\phi(\zeta))$ yields

$$F_\omega(\zeta) \approx C_N(F_\omega, h)(\zeta) = \sum_{k=-N}^{N} F_\omega(kh) \cdot S(k,h)(\zeta).$$

Backsubstitution yields $\omega(x)f(x) \approx \sum_{k=-N}^{N} (\omega \cdot f)(\phi(kh)) \cdot S(k,h)(\phi^{-1}(x))$ and

$$f(x) \approx \hat{C}_N(f,h)(x) := \sum_{k=-N}^{N} (\omega \cdot f)\,(\phi(kh)) \frac{S(k,h)(\phi^{-1}(x))}{\omega(x)} \quad \text{for } x \in (0,1].$$

The convergence of $|f(x) - \hat{C}_N(f,h)(x)|$ as $N \to \infty$ is no longer uniform. Instead, the weighted error

$$\|f - \hat{C}_N(f,h)\|_\omega := \|\omega[f - \hat{C}_N(f,h)]\|_\infty$$

satisfies the previous estimates. In many cases, this is still sufficient.

Example 10.38. (a) The function $f(x) = x^x$ is analytic in $(0,1]$, but singular at $x = 0$ with $\lim_{x\to 0} f(x) = 1$. Choose $\omega(x) := x^\lambda$ for some $\lambda > 0$ and transform by

$\phi(\zeta) = 1/\cosh(\zeta)$. Then $F_\omega(\zeta) = (\cosh(\zeta))^{-\lambda-1/\cosh(\zeta)}$ behaves for $\zeta\to\pm\infty$ like $2^\lambda \exp(-\lambda\,|\zeta|)$. It belongs to $\mathbf{H}(\mathcal{D}_\delta)$ for $\delta < \pi/2$ (note the singularity at $\zeta = \pm\pi\mathrm{i}/2$). Therefore, Lemma 10.34 can be applied.

(b) Even if f is unbounded like $f(x) = 1/\sqrt{x}$, the weight $\omega(x) := x^{1/2+\lambda}$ leads to the same convergence of F_ω as in Part (a).

10.4.4.3 Separation by Interpolation, Tensor Subspace Representation

We apply Remark 10.20 for $d = 2$, where the interpolation is the sinc interpolation with respect to the first variable. Consider a function $f(x, y)$ with $x \in X$ and $y \in Y$. If necessary, we apply a transformation $x = \phi(\zeta)$ with $\phi : \mathbb{R} \to X$, so that the first argument varies in $\mathbb{R}$ instead of X. Therefore, we may assume that $f(x, y)$ is given with $x \in X = \mathbb{R}$. Suppose that $f(x, y) \to 0$ as $|x| \to \pm\infty$. Sinc interpolation in x yields

$$f(x, y) \approx C_N(f(\cdot, y), h)(x) = \sum_{k=-N}^{N} f(kh, y)\cdot S(k, h)(x).$$

The previous convergence results require uniform conditions with respect to y.

Proposition 10.39. *Assume that there is some $\delta > 0$ so that $f(\cdot, y) \in \mathbf{H}(\mathcal{D}_\delta)$ for all $y \in Y$ and $|||\, f\,||| := \sup\{\|f(\cdot, y)\|_{\mathfrak{D}_\delta} : y \in Y\} < \infty$. Furthermore, suppose that there are $c \geq 0$ and $\alpha > 0$ such that*

$$|f(x, y)| \leq c\cdot \mathrm{e}^{-\alpha|x|} \qquad \textit{for all } x \in \mathbb{R},\ y \in Y.$$

Then the choice $h := \sqrt{\frac{\pi\delta}{\alpha N}}$ yields the uniform error estimate

$$|E_N(f(\cdot, y), h)| \leq \sqrt{\frac{N}{\delta}}\,\mathrm{e}^{-\sqrt{\pi\alpha\delta N}}\left[\frac{|||\, f\,|||}{\pi\left[1-\mathrm{e}^{-\pi\alpha\delta N}\right]\sqrt{N\delta}} + \frac{2c}{\sqrt{\pi\alpha}}\right] \qquad (y \in Y).$$

Proof. For any $y \in Y$, the univariate function $f(\cdot, y)$ satisfies the conditions of Theorem 10.35. Inequality $\|f(\cdot, y)\|_{\mathfrak{D}_\delta} \leq |||\, f\,|||$ proves the desired estimate. □

If $f(x, y) \to 0$ as $|x| \to \pm\infty$ is not satisfied, the weighting by $\omega(x)$ has to be applied additionally. We give an example of this type.

Example 10.40. Consider the function $f(x, y) := \frac{1}{x+y}$ for $x, y \in (0, \infty)$. Choose the weight function $\omega(x) := x^\alpha$ $(0 < \alpha < 1)$. Transformation $x = \phi(\zeta) := \exp(\zeta)$ yields

$$F_\omega(\zeta, y) := [\omega(x) f(x, y)]\Big|_{x=\phi(\zeta)} = \frac{\exp(\alpha\zeta)}{y + \exp(\zeta)}.$$

One checks that $F_\omega \in \mathbf{H}(\mathcal{D}_\delta)$ for all $\delta < \pi$. The norm $\|f(\cdot, y)\|_{\mathfrak{D}_\delta}$ is not uniquely bounded for $y \in (0, \infty)$, but behaves like $O(y^{\alpha-1})$:

$$\left|\omega(x)\left[f(x,y)-\hat{C}_N(f(\cdot,y),h)(x)\right]\right| \le O(y^{\alpha-1}\mathrm{e}^{-\sqrt{\pi\alpha\delta N}}) \qquad \text{for}$$

$$\hat{C}_N(f(\cdot,y),h)(x) := \frac{C_N(F_\omega(\cdot,y),h)(\log(x))}{\omega(x)} = \sum_{k=-N}^{N} \frac{\mathrm{e}^{\alpha kh}}{y+\mathrm{e}^{kh}} \frac{S(k,h)(\log(x))}{\omega(x)}.$$

The interpolation yields a sum of the form $F_N(x,y) := \sum_{k=-N}^{N} f_k(y)L_k(x)$ with L_k being the sinc function $S(k,h)$ possibly with a further transformation and additional weighting, while $f_k(y)$ are evaluations of $f(x,y)$ at certain x_k. Note that $F_N \in \mathcal{R}_{2N+1}$.

Next, we assume that $f(x_1,x_2,\ldots,x_d)$ is a d-variate function. For simplicity suppose that all x_j vary in $\mathbb{R}$ with $f \to 0$ as $|x_j| \to \pm\infty$ to avoid transformations. Since interpolation in x_1 yields

$$f_{(1)}(x_1,x_2,\ldots,x_d) = \sum_{k_1=-N_1}^{N_1} f_{(1),k_1}(x_2,\ldots,x_d)\cdot S(k_1,h)(x_1).$$

Application of sinc interpolation to $f_{(1),k}(x_2,\ldots,x_d)$ with respect to x_2 separates the x_2 dependence. Insertion into the previous sum yields

$$f_{(2)}(x_1,\ldots,x_d) = \sum_{k_1=-N_1}^{N_1} \sum_{k_2=-N_2}^{N_2} f_{(2),k_1,k_2}(x_3,\ldots,x_d)\cdot S(k_1,h)(x_1)\cdot S(k_2,h)(x_2).$$

After $d-1$ steps we reach at

$$f_{(d-1)}(x_1,\ldots,x_d) = \sum_{k_1=-N_1}^{N_1} \cdots \sum_{k_{d-1}=-N_{d-1}}^{N_{d-1}} f_{(d-1),k_1,\ldots,k_{d-1}}(x_d) \prod_{j=1}^{d-1} S(k_j,h)(x_j),$$

which belongs to $\bigotimes_{j=1}^{d} U_j$, where $\dim(U_j) = 2N_j + 1$, while U_d spanned by all $f_{(d-1),k_1,\ldots,k_{d-1}}$ for $-N_j \le k_j \le N_j$ is high-dimensional. The last step yields

$$f_{(d)}(x_1,x_2,\ldots,x_d) = \sum_{k_1=-N_1}^{N_1} \cdots \sum_{k_d=-N_d}^{N_d} f_{(d),k_1,\ldots,k_{d-1}} \prod_{j=1}^{d} S(k_j,h)(x_j) \in \mathcal{T}_{\mathbf{r}}\,,$$

where $\mathbf{r} = (2N_1+1,\ldots,2N_d+1)$ is the tensor subspace rank.

10.5 Simultaneous Approximation

Let $\mathbf{v}_1,\ldots,\mathbf{v}_m \in \mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$ be m tensors. As in Problem (10.12) we want to approximate all $\mathbf{v}_i$ by $\mathbf{u}_i \in \mathcal{T}_{\mathbf{r}}$, however, the involved subspace $\mathbf{U} = \bigotimes_{j=1}^{d} U_j$ with $\dim(U_j) = r_j$ should be the same for all $\mathbf{u}_i$. More precisely, we are looking for the minimisers $\mathbf{u}_i$ of the following minimisation problem:

$$\inf_{\substack{U_1\subset V_1 \text{ with}\\ \dim(U_1)=r_1}} \inf_{\substack{U_2\subset V_2 \text{ with}\\ \dim(U_2)=r_2}} \cdots \inf_{\substack{U_d\subset V_d \text{ with}\\ \dim(U_d)=r_d}} \left\{ \inf_{\mathbf{u}_i\in\bigotimes_{j=1}^d U_j} \sum_{i=1}^m \omega_i^2 \left\|\mathbf{v}_i-\mathbf{u}_i\right\|^2 \right\}, \tag{10.51}$$

where we have associated suitable weights $\omega_i^2 > 0$.

Such a problem arises for matrices (i.e., $d = 2$), e.g., in the construction of $\mathcal{H}^2$-matrices (cf. [86, §8]).

We refer to Lemma 3.26: a tuple $(\mathbf{v}_1, \ldots, \mathbf{v}_m) \in \mathbf{V}^m$ may be considered as a tensor of $\mathbf{W} = {}_a\bigotimes_{j=1}^{d+1} V_j$ with the $(d+1)$-th vector space $V_{d+1} := \mathbb{K}^m$. The Hilbert structure is discussed in the next remark.

Remark 10.41. Let $\mathbf{V} = {}_a\bigotimes_{j=1}^{d} V_j$ be a Hilbert space with scalar product $\langle\cdot,\cdot\rangle_{\mathbf{V}}$, while $\mathbb{K}^m$ is endowed with the scalar product $\langle x,y\rangle_{d+1} := \sum_{i=1}^m \omega_i^2 x_i \overline{y_i}$. Then $\mathbf{V}^m$ is isomorphic and isometric to $\mathbf{W} = {}_a\bigotimes_{j=1}^{d+1} V_j$ with $V_{d+1} := \mathbb{K}^m$. Tuples $(\mathbf{v}_1, \ldots, \mathbf{v}_m) \in \mathbf{V}^m$ are written as tensors $\mathbf{w} := \sum_{i=1}^m \mathbf{v}_i \otimes e^{(i)} \in \mathbf{W}$ ($e^{(i)}$ unit vectors, cf. (2.2)) with the following induced scalar product and norm:

$$\langle \mathbf{w}, \mathbf{w}'\rangle = \sum\nolimits_{i=1}^m \omega_i^2 \left\langle \mathbf{v}_i, \mathbf{v}_i'\right\rangle_{\mathbf{V}}, \qquad \|\mathbf{w}\| = \sqrt{\sum\nolimits_{i=1}^m \omega_i^2 \left\|\mathbf{v}_i\right\|_{\mathbf{V}}^2}.$$

Hence, Problem (10.51) is equivalent to

$$\inf_{\substack{U_1\subset V_1 \text{ with}\\ \dim(U_1)=r_1}} \inf_{\substack{U_2\subset V_2 \text{ with}\\ \dim(U_2)=r_2}} \cdots \inf_{\substack{U_d\subset V_d \text{ with}\\ \dim(U_d)=r_d}} \left\{ \inf_{\mathbf{u}\in\bigotimes_{j=1}^{d+1} U_j} \left\|\mathbf{w}-\mathbf{u}\right\|^2 \right\} \tag{10.52}$$

where the last subspace $U_{d+1} = V_{d+1} = \mathbb{K}^m$ has full dimension. This shows the equivalence to the basic Problem (10.12) (but with d replaced by $d+1$, $V_{d+1} = \mathbb{K}^m$ and $r_{d+1} = m$). The statements about the existence of a minimiser of (10.12) can be transferred to statements about Problem (10.51).

As an important application we consider the simultaneous approximation problem for matrices $\mathbf{v}_i = M_i \in \mathbb{K}^{I\times J}$ $(1 \le i \le m)$, where $\sum_{i=1}^m \omega_i^2 \left\|M_i - R_i\right\|_\mathsf{F}^2$ is to be minimised with respect to $R_i \in U_1 \otimes U_2$ with the side conditions $\dim(U_1) = r_1$ and $\dim(U_2) = r_2$. Here, $\mathbf{W} = \mathbb{K}^I \otimes \mathbb{K}^J \otimes \mathbb{K}^m$ is the underlying tensor space. The HOSVD bases of $\mathbb{K}^I$ and $\mathbb{K}^J$ result from the matrices $U^{(1)}$ and $U^{(2)}$ of the left-sided singular value decompositions $\mathbf{LSVD}(I, m\#J, r_1, \mathcal{M}_1(\mathbf{w}), U^{(1)}, \Sigma^{(1)})$ and $\mathbf{LSVD}(J, m\#I, r_2, \mathcal{M}_2(\mathbf{w}), U^{(2)}, \Sigma^{(2)})$:

$$\mathcal{M}_1(\mathbf{w}) = [\omega_1 M_1,\ \omega_2 M_2, \ldots,\ \omega_m M_m] = U^{(1)}\Sigma^{(1)}V^{(1)\mathsf{T}} \in \mathbb{K}^{I\times(J\times m)},$$
$$\mathcal{M}_2(\mathbf{w}) = \left[\omega_1 M_1^\mathsf{T},\ \omega_2 M_2^\mathsf{T}, \ldots,\ \omega_m M_m^\mathsf{T}\right] = U^{(2)}\Sigma^{(2)}V^{(2)\mathsf{T}} \in \mathbb{K}^{J\times(I\times m)}.$$

Equivalently, one has to perform the diagonalisations

$$\mathcal{M}_1(\mathbf{w})\mathcal{M}_1(\mathbf{w})^\mathsf{T} = \sum\nolimits_{i=1}^m \omega_i^2 M_i M_i^\mathsf{T} = U^{(1)}(\Sigma^{(1)})^2 U^{(1)\mathsf{T}} \in \mathbb{K}^{I\times I},$$
$$\mathcal{M}_2(\mathbf{w})\mathcal{M}_2(\mathbf{w})^\mathsf{T} = \sum\nolimits_{i=1}^m \omega_i^2 M_i^\mathsf{T} M_i = U^{(2)}(\Sigma^{(2)})^2 U^{(2)\mathsf{T}} \in \mathbb{K}^{J\times J}.$$

The HOSVD basis of $\mathbb{K}^m$ is not needed since we do not want to introduce a strictly smaller subspace. The HOSVD projection from Lemma 10.1 takes the special form $R_k := P_1^{(r_1)} M_k P_2^{(r_2)}$, where $P_i^{(r_i)} = \sum_{\nu=1}^{r_i} u_\nu^{(i)} u_\nu^{(i)\mathsf{H}}$ $(i = 1, 2)$ uses the ν-th column $u_\nu^{(i)}$ of $U^{(i)}$. The error estimate (10.4b) becomes

$$\sum\nolimits_{i=1}^{m} \omega_i^2 \left\| M_i - R_i \right\|_{\mathsf{F}}^2 = \left\| \mathbf{w} - \mathbf{u}_{\mathrm{HOSVD}} \right\|^2 \leq \sum_{j=1}^{2} \sum_{i=r_j+1}^{n_j} \left(\sigma_i^{(j)} \right)^2$$

$$\leq 2 \left\| \mathbf{v} - \mathbf{u}_{\mathrm{best}} \right\|^2 = 2 \sum_{i=1}^{m} \omega_i^2 \left\| M_i - R_i^{\mathrm{best}} \right\|_{\mathsf{F}}^2$$

with $n_1 := \#I$ and $n_2 := \#J$. Here, we have made use of Corollary 10.4 because of $r_3 = n_3 := m$.

10.6 Résumé

The discussion of the two traditional formats (r-term and tensor subspace format) has shown that the analysis as well as the numerical treatment is by far more complicated for $d \geq 3$ than for the matrix case $d = 2$. The main disadvantages are:

⊖ Truncation of $\mathbf{v} \in \mathcal{R}_R$ to $\mathbf{u} \in \mathcal{R}_r$ with smaller rank $r < R$ is not easy to perform. This nonlinear optimisation problem, which usually needs regularisation, does not only lead to an involved numerical algorithm, but also the result is not reliable, since a local minimum may be obtained and different starting value can lead to different optima.

⊖ There is no decomposition of $\mathbf{v} \in \mathcal{R}_r$ into highly and less important terms, which could help for the truncation. In contrast, large terms may be negligible since they add up to a small tensor.

⊖ The disadvantage of the tensor subspace format is the data size $\prod_{j=1}^{d} r_j$ of the coefficient tensor. For larger d and $r_j \geq r$, the exponential increase of r^d leads to severe memory problems.

⊖ While the ranks r_j of the tensor subspace format are bounded by $n_j = \dim(V_j)$, the upper bound of the tensor rank has exponential increase with respect to d (cf. Lemma 3.41). Therefore, the size of r in $\mathbf{v} \in \mathcal{R}_r$ may become problematic.

On the other hand, both formats have their characteristic advantages:

⊕ If the rank r of $\mathbf{v} \in \mathcal{R}_r$ is moderate, the representation of $\mathbf{v}$ by the r-term format requires a rather small storage size, which is proportional to r, d, and the dimension of the involved vector spaces V_j.

⊕ The tensor subspace format together with the higher order singular value decompositions (HOSVD) allows a simple truncation to smaller ranks. For the important case of $d = 3$, the data size $\prod_{j=1}^{d} r_j$ is still tolerable.

Chapter 11
Hierarchical Tensor Representation

Abstract The hierarchical tensor representation (notation: $\mathcal{H}_{\mathfrak{r}}$) allows to keep the advantages of the subspace structure of the tensor subspace format $\mathcal{T}_{\mathbf{r}}$, but has only linear cost with respect to the order d concerning storage and operations. The hierarchy mentioned in the name is given by a 'dimension partition tree'. The fact that the tree is binary, allows a simple application of the singular value decomposition and enables an easy truncation procedure.

After an introduction in *Sect.* *11.1*, the algebraic structure of the hierarchical tensor representation is described in *Sect.* *11.2*. While the algebraic representation uses subspaces, the concrete representation in *Sect.* *11.3* introduces frames or bases and the associated coefficient matrices in the hierarchy. Again, higher order singular value decompositions (HOSVD) can be applied and the corresponding singular vectors can be used as basis. In *Sect.* *11.4*, the approximation in the $\mathcal{H}_{\mathfrak{r}}$ format is studied with respect to two aspects. First, the best approximation within $\mathcal{H}_{\mathfrak{r}}$ can be solved. Second, the HOSVD bases allow a quasi-optimal truncation. *Section* *11.5* discusses the joining of two representations. This important feature is needed if two tensors described by two different hierarchical tensor representations require a common representation. Finally, *Sect.* *11.6* shows how the sparse grid format can be mapped into the hierarchical tensor format.

11.1 Introduction

11.1.1 Hierarchical Structure

In the following, we want to keep the positive properties of the tensor subspace representation but avoiding exponential increase of the coefficient tensor. The dimension of the subspaces $U_j \subset V_j$ is bounded by r_j, but their d-fold tensor product is again high-dimensional. In the approach of the hierarchical tensor format, we repeat the concept of tensor subspaces on higher levels: we do not form tensor products of *all* U_j, but choose again subspaces of pairs of subspaces so that

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_11,

the dimension is reduced. The recursive use of the subspace idea leads to a certain tree structure describing a hierarchy of subspaces. In particular, we shall use binary trees, since this fact will allow us to apply standard singular value decompositions to obtain HOSVD bases and to apply HOSVD truncations.

The previous r-term and tensor subspace representations are invariant with respect to the ordering of the spaces V_j. This is different for the hierarchical format.[1] Before giving a strict definition, we illustrate the idea by examples.

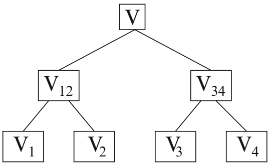


Fig. 11.1 Hierarchy

We start with the case of $d = 4$, where $\mathbf{V} = \bigotimes_{j=1}^{4} V_j$. In §3.2.4, a tensor space of order 4 has been introduced as $((V_1 \otimes V_2) \otimes V_3) \otimes V_4$ using the definition of binary tensor products. However, the order of binary tensor products can be varied. We may first introduce the spaces $\mathbf{V}_{12} := V_1 \otimes V_2$ and $\mathbf{V}_{34} := V_3 \otimes V_4$ and then $\mathbf{V}_{12} \otimes \mathbf{V}_{34} = (V_1 \otimes V_2) \otimes (V_3 \otimes V_4) \cong V_1 \otimes V_2 \otimes V_3 \otimes V_4$. This approach is visualised in Fig. 11.1. Following the tensor subspace idea, we introduce subspaces for all spaces in Fig. 11.1. This leads to the following construction:

$$\begin{array}{ccccccc}
 & & & \mathbf{U}_{\{1,2,3,4\}} \subset \mathbf{U}_{\{1,2\}} \otimes \mathbf{U}_{\{3,4\}} \subset \mathbf{V} & & & \\
 & & / & & \backslash & & \\
 & \mathbf{U}_{\{1,2\}} \subset U_1 \otimes U_2 & & & & \mathbf{U}_{\{3,4\}} \subset U_3 \otimes U_4 & \\
 / & & \backslash & & / & & \backslash \\
U_1 \subset V_1 & & U_2 \subset V_2 & & U_3 \subset V_3 & & U_4 \subset V_4
\end{array} \tag{11.1}$$

The tensor to be represented must be contained in the upper subspace $\mathbf{U}_{\{1,2,3,4\}}$. Assume that there are suitable subspaces U_j $(1 \leq j \leq 4)$ of dimension r. Then the tensor products $U_1 \otimes U_2$ and $U_3 \otimes U_4$ have the increased dimension r^2. The hope is to find again subspaces $\mathbf{U}_{\{1,2\}}$ and $\mathbf{U}_{\{3,4\}}$ of smaller dimension, say, r. In this way, the exponential increase of the dimension could be avoided.

The construction by (11.1b) is still invariant with respect to permutations $1 \leftrightarrow 2$, $3 \leftrightarrow 4$, $\{1,2\} \leftrightarrow \{3,4\}$; however, the permutation $2 \leftrightarrow 3$ yields another tree.

A perfectly balanced tree like in Fig. 11.1 requires that $d = 2^L$, where L is the depth of the tree. The next example $d = 7$, i.e., $\mathbf{V} = \bigotimes_{j=1}^{7} V_j$, shows a possible construction of $\mathbf{V}$ by binary tensor products:

$$\begin{array}{lcccccccc}
\text{level 0} & & & & \mathbf{V} & & & & \\
 & & & / & & \backslash & & & \\
\text{level 1} & & \mathbf{V}_{123} & & & & \mathbf{V}_{4567} & & \\
 & / & \backslash & & & / & & \backslash & \\
\text{level 2} & V_1 & & \mathbf{V}_{23} & & \mathbf{V}_{45} & & \mathbf{V}_{67} & \\
 & & / & \backslash & / & \backslash & / & \backslash & \\
\text{level 3} & & V_2 & V_3 & V_4 & V_5 & V_6 & & V_7
\end{array} \tag{11.2}$$

Again, the hierarchical format replaces the full spaces by subspaces.

[1] Instead of 'hierarchical tensor format' also the term '$\mathcal{H}$-Tucker format' is used (cf. [74], [131]), since the subspace idea of the Tucker format is repeated recursively.

The position of $\mathbf{V}$ in the tree is called the *root* of tree. The factors of a binary tensor product appear in the tree as two *sons*, e.g., V_1 and $\mathbf{V}_{23}$ are the sons of $\mathbf{V}_{123}$ in the last example. Equivalently, $\mathbf{V}_{123}$ is called the *father* of V_1 and $\mathbf{V}_{23}$. The spaces V_j, which cannot be decomposed further, are the *leaves* of the tree. The root is associated with the level 0. The further levels are defined recursively: sons of a father at level ℓ have the level number $\ell + 1$.

The fact that we choose a binary tree (i.e., the number of sons is either 2 or 0) is essential, since we want to apply matrix techniques. If we decompose tensor products of order d recursively in $(d/2) + (d/2)$ factors for even d and in $((d-1)/2) + (d/2)$ factors for odd d, it follows easily that[2]

$$L := \lceil \log_2 d \rceil \tag{11.3}$$

is the largest level number (depth of the tree).

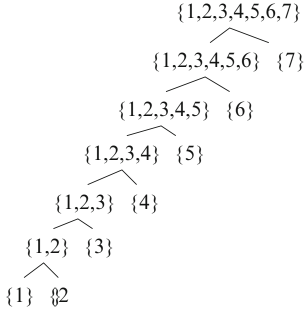


Fig. 11.2 Linear tree T_D^{TT}

Quite another division strategy is the splitting into $1+(d-1)$ factors. The latter example $\bigotimes_{j=1}^{7} V_j$ leads to the partition tree T_D^{TT} depicted in Fig. 11.2. In this case, the depth of the tree is maximal:

$$L := d - 1. \tag{11.4}$$

Another derivation of the hierarchical representation can be connected with the minimal subspaces from §6. The fact that a tensor $\mathbf{v}$ might have small j-th ranks, gives rise to the standard tensor subspace format with $U_j = U_j^{\min}(\mathbf{v})$. However, minimal subspaces $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ of dimension r_α are also existing for subsets $\emptyset \subsetneqq \alpha \subsetneqq D := \{1, \ldots, d\}$ (cf. §6.4). For instance, $\mathbf{U}_{\{1,2\}}$ and $\mathbf{U}_{\{3,4\}}$ in (11.1) can be chosen as $\mathbf{U}_{\{1,2\}}^{\min}(\mathbf{v})$ and $\mathbf{U}_{\{3,4\}}^{\min}(\mathbf{v})$. As shown in (6.13), they are nested is the sense of $\mathbf{U}_{\{1,2\}} \subset U_1 \otimes U_2$, etc. Note that the last minimal subspace is one-dimensional: $\mathbf{U}_{\{1,2,3,4\}} = \mathbf{U}_{\{1,2,3,4\}}^{\min}(\mathbf{v}) = \mathrm{span}\{\mathbf{v}\}$.

11.1.2 Properties

We give a preview of the favourable features of the actual approach.

1. Representations in the formats $\mathcal{R}_r$ (cf. §11.2.4.2 and §11.3.5) or $\mathcal{T}_\mathbf{r}$ (cf. §11.2.4.1) can be converted into hierarchical format with similar storage cost. Later this will be shown for further formats (sparse-grid representation, TT representation). As a consequence, the approximability by the hierarchical format is at least as good as by the aforementioned formats.
2. The cost is strictly linear in the order d of the tensor space $\bigotimes_{j=1}^{d} V_j$. This is in contrast to $\mathcal{T}_\mathbf{r}$, where the coefficient tensor causes problems for higher d. These

[2] $\lceil x \rceil$ is the integer with $x \le \lceil x \rceil < x + 1$.

statements hold under the assumption that the rank parameters stay bounded when d varies. If, however, the rank parameters increase with d, all formats have problems.

3. The binary tree structure allows us to compute all approximations by means of standard linear algebra tools. This is essential for the truncation procedure.
4. The actual representation may be much cheaper than a representation within $\mathcal{R}_r$ or $\mathcal{T}_{\mathbf{r}}$ (see next Example 11.1).

Example 11.1. Consider the Banach tensor space $L^2([0,1]^d) = {}_{L^2}\bigotimes_{j=1}^{d} L^2([0,1])$ for $d = 4$ and the particular function

$$f(x_1, x_2, x_3, x_4) = P_1(x_1, x_2) \cdot P_2(x_3, x_4),$$

where P_1 and P_2 are polynomials of degree p.

(a) *Realisation in hierarchical format*. The example is such that the dimension partition tree T_D from Fig. 11.3 is optimal. Writing $P_1(x_1, x_2)$ as

$$\sum_{i=1}^{p+1} P_{1,i}(x_1) x_2^{i-1} = \sum_{i=1}^{p+1} P_{1,i}(x_1) \otimes x_2^{i-1},$$

we see that $\dim(U_1) = \dim(U_2) = p+1$ is sufficient to have $P_1(x_1, x_2) \in U_1 \otimes U_2$. Similarly, $\dim(U_3) = \dim(U_4) = p+1$ is enough to ensure $P_2(x_3, x_4) \in U_3 \otimes U_4$. The subspaces $\mathbf{U}_{12}$ and $\mathbf{U}_{34}$ may be one-dimensional: $\mathbf{U}_{12} = \operatorname{span}\{P_1(x_1, x_2)\}$, $\mathbf{U}_{34} = \operatorname{span}\{P_2(x_3, x_4)\}$. The last subspace $\mathbf{U}_{14} = \operatorname{span}\{f\}$ is one-dimensional anyway. Hence, the highest dimension is $r_j = p+1$ for $1 \le j \le 4$.

(b) *Realisation in $\mathcal{T}_{\mathbf{r}}$*. The subspaces U_j coincide with U_j from above, so that $r_j = \dim(U_j) = p+1$. Hence, $\mathbf{U} = \bigotimes_{j=1}^{4} U_j$ has dimension $(p+1)^d = (p+1)^4$.

(c) *Realisation in $\mathcal{R}_r$*. $P_1(x_1, x_2) = \sum_{i=1}^{p+1} P_{1,i}(x_1) x_2^{i-1}$ as well as $P_2(x_3, x_4) = \sum_{i=1}^{p+1} P_{2,i}(x_3) x_4^{i-1}$ have $p+1$ terms; hence, their product has $r = (p+1)^2$ terms.

The background of this example is that the formats $\mathcal{T}_{\mathbf{r}}$ and $\mathcal{R}_r$ are symmetric in the treatment of the factors V_j in $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$. The simple structure $f = P_1 \otimes P_2$ (while P_1 and P_2 are not elementary tensors) cannot be mapped into the $\mathcal{T}_{\mathbf{r}}$ or $\mathcal{R}_r$ structure. One may extend the example to higher d by choosing $f = \bigotimes_{j=1}^{d/2} P_j$, where $P_j = P_j(x_{2j-1}, x_{2j})$. Then the cost of the $\mathcal{T}_{\mathbf{r}}$ and $\mathcal{R}_r$ formats is exponential in d (more precisely, $(p+1)^d$ for $\mathcal{T}_{\mathbf{r}}$ and $r = (p+1)^{d/2}$ for $\mathcal{R}_r$), while for the hierarchical format the dimensions are $r_j = p+1$ $(1 \le j \le d)$ and $r_\alpha = 1$ for all other subsets α appearing in the tree.

11.1.3 Historical Comments

The hierarchical idea has also appeared as 'sequential unfolding' of a tensor. For instance, $\mathbf{v} \in \mathbb{K}^{n \times n \times n \times n}$ can be split by a singular value decomposition into $\mathbf{v} = \sum_\nu \mathbf{e}_\nu \otimes \mathbf{f}_\nu$ with $\mathbf{e}_\nu, \mathbf{f}_\nu \in \mathbb{K}^{n \times n}$. Again, each $\mathbf{e}_\nu, \mathbf{f}_\nu$ has a decomposition $\mathbf{e}_\nu = \sum_\mu a_{\nu,\mu} \otimes b_{\nu,\mu}$ and $\mathbf{f}_\nu = \sum_\lambda c_{\nu,\lambda} \otimes d_{\nu,\lambda}$ with vectors $a_{\nu,\mu}, \ldots, d_{\nu,\lambda} \in \mathbb{K}^n$.

Together, we obtain $\mathbf{v} = \sum_{\nu,\mu,\lambda} a_{\nu,\mu} \otimes b_{\nu,\mu} \otimes c_{\nu,\lambda} \otimes d_{\nu,\lambda}$. The required data size is $4r^2n$, where $r \leq n$ is the maximal size of the index sets for ν, μ, λ. In the general case of $\mathbf{v} \in \bigotimes_{j=1}^{d} \mathbb{K}^n$ with $d = 2^L$, the required storage is $dr^L n$. Such an approach is considered in Khoromskij [115, §2.2] and repeated in Salmi-Richter-Koivunen [165]. It has two drawbacks. The required storage is still exponentially increasing (but only with exponent $L = \log_2 d$), but the major problem is that singular value decompositions are to be computed for extremely large matrices.

The remedy is to require that all $\mathbf{e}_\nu$ from above belong *simultaneously* to one r-dimensional subspace ($\mathbf{U}_{12}$ in (11.1)). The author has borrowed this idea from a similar approach leading to the $\mathcal{H}^2$ technique of hierarchical matrices (cf. Hackbusch [86, §8], Börm [21]). The hierarchical tensor format is described 2009 in Hackbusch-Kühn [94]. A further analysis is given by Grasedyck [73]. However, the method has been mentioned much earlier by Vidal [191] in the quantum computing community. In quantum chemistry, MCTDH abbreviates 'multi-configuration time-dependent Hartree' (cf. Meyer et al. [147]). In Wang-Thoss [194], a multilayer formulation of the MCTDH theory is presented which might contain the idea of a hierarchical format. At least Lubich [144, p. 45] has translated the very specific quantum chemistry language of that paper into a mathematical formulation using the key construction (11.24) of the hierarchical tensor representation.

A closely related method is the matrix product system which is the subject of the next chapter (§12).

11.2 Basic Definitions

11.2.1 Dimension Partition Tree

We consider the tensor space[3]

$$\mathbf{V} = {}_a\bigotimes_{j\in D} V_j \tag{11.5}$$

with a finite index set D. To avoid trivial cases, we assume $\#D \geq 2$.

Definition 11.2 (dimension partition tree). The tree T_D is called a *dimension partition tree* (of D) if

1) all vertices[4] $\alpha \in T_D$ are non-empty subsets of D,
2) D is the root of T_D,
3) every vertex $\alpha \in T_D$ with $\#\alpha \geq 2$ has two sons $\alpha_1, \alpha_2 \in T_D$ such that

$$\alpha = \alpha_1 \cup \alpha_2, \qquad \alpha_1 \cap \alpha_2 = \emptyset.$$

The set of sons of α is denoted by $S(\alpha)$. If $S(\alpha) = \emptyset$, α is called a leaf. The set of leaves is denoted by $\mathcal{L}(T_D)$.

[3] The tensors represented next belong to the *algebraic* tensor space $\mathbf{V}_{\text{alg}} = {}_a\bigotimes_{j\in D} V_j$. Since $\mathbf{V}_{\text{alg}} \subset \mathbf{V}_{\text{top}} = {}_{\|\cdot\|}\bigotimes_{j\in D} V_j$, they may also be seen as elements of $\mathbf{V}_{\text{top}}$.

[4] Elements of a tree are called '*vertices*'.

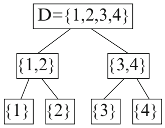


Fig. 11.3 Dimension partition tree

The tree T_D corresponding to (11.1) is illustrated in Fig. 11.3. The numbers $1, \ldots, d$ are chosen according to Remark 11.4.

As mentioned in §11.1, the level number of vertices are defined recursively by

$$level(D) = 0, \quad \sigma \in S(\alpha) \Rightarrow level(\sigma) = level(\alpha) + 1. \tag{11.6}$$

The *depth of the tree* defined below is often abbreviated by L:

$$L := depth(T_D) := \max \{level(\alpha) : \alpha \in T_D\}. \tag{11.7}$$

Occasionally, the following level-wise decomposition of the tree T_D is of interest:

$$T_D^{(\ell)} := \{\alpha \in T_D : level(\alpha) = \ell\} \qquad (0 \leq \ell \leq L). \tag{11.8}$$

Remark 11.3. Easy consequences of Definition 11.2 are:
(a) T_D is a binary tree,
(b) The set of leaves, $\mathcal{L}(T_D)$, consists of all singletons of D:

$$\mathcal{L}(T_D) = \{\{j\} : j \in D\}. \tag{11.9}$$

(c) The number of vertices in T_D is $2\#D - 1$.

Considering D as a set, no ordering is prescribed. A total ordering not only of D, but also of all vertices of the tree T_D can be defined as follows.

Remark 11.4 (ordering of T_D). (a) Choose any ordering of the two elements in $S(\alpha)$ for any $\alpha \in T_D \backslash \mathcal{L}(T_D)$, i.e., there is a *first* son α_1 and a *second* son α_2 of α. The ordering is denoted by $\alpha_1 < \alpha_2$. Then for two different $\beta, \gamma \in T_D$ there are the following cases:

(i) If β, γ are disjoint, there is some $\alpha \in T_D$ with sons $\alpha_1 < \alpha_2$ such that $\beta \subset \alpha_1$ and $\gamma \subset \alpha_2$ [or $\gamma \subset \alpha_1$ and $\beta \subset \alpha_2$]. Then define $\beta < \gamma$ [or $\gamma < \beta$, respectively];

(ii) If β, γ are not disjoint, either $\beta \subset \gamma$ or $\gamma \subset \beta$ must hold. Then define $\beta < \gamma$ or $\gamma < \beta$, respectively.

(b) Let T_D be ordered and denote the elements of D by $1, \ldots, d$ according to their ordering. The vertices α are of the form $\alpha = \{i \in D : i_{\min}^{\alpha} \leq i \leq i_{\max}^{\alpha}\}$ with bounds $i_{\min}^{\alpha}, i_{\max}^{\alpha}$. For the sons $\alpha_1 < \alpha_2$ of α we have $i_{\min}^{\alpha} = i_{\min}^{\alpha_1} \leq i_{\max}^{\alpha_1} = i_{\min}^{\alpha_2} - 1 < i_{\max}^{\alpha_2} = i_{\max}^{\alpha}$. Interchanging the ordering of α_1 and α_2 yields a permutation of D, but the hierarchical representation will be completely isomorphic (cf. Remark 11.20).

The notations $\alpha_1, \alpha_2 \in S(\alpha)$ or $S(\alpha) = \{\alpha_1, \alpha_2\}$ tacitly imply that $\alpha_1 < \alpha_2$.

Taking the example (11.1) corresponding to T_D from Fig. 11.3, we may define that the left son precedes the right one. Then the total ordering

$$\{1\} < \{2\} < \{1,2\} < \{3\} < \{4\} < \{3,4\} < \{1,2,3,4\}$$

of T_D results. In particular, one obtains the ordering $1 < 2 < 3 < 4$ of D, where we identify j with $\{j\}$.

Remark 11.5. (a) The minimal depth of T_D is $L = \lceil \log_2 d \rceil$ (cf. (11.3)), which is obtained under the additional condition

$$|\#\alpha_1 - \#\alpha_2| \leq 1 \qquad \text{for all } \alpha_1, \alpha_2 \in S(\alpha) \backslash \mathcal{L}(T_D).$$

(b) The maximal depth of T_D is $L = d - 1$ (cf. (11.4)).

Different trees T_D will lead to different formats and in the case of approximations to different approximation errors. Example 11.1 shows that for a given tensor there may be more and less appropriate trees.

We recall the notation $\mathbf{V}_\alpha := \bigotimes_{j\in\alpha} V_j$ for $\alpha \subset D$ (cf. (5.4)). For leaves $\alpha = \{j\}$ with $j \in D$, the latter definition reads $\mathbf{V}_{\{j\}} = V_j$. The matricisation $\mathcal{M}_\alpha$ denotes the isomorphism $\mathbf{V} \cong \mathbf{V}_\alpha \otimes \mathbf{V}_{\alpha^c}$, where $\alpha^c := D\backslash\alpha$ is the complement.

Definition 11.6 **(T_α).** For any $\alpha \in T_D$ the subtree $T_\alpha := \{\beta \in T_D : \beta \subset \alpha\}$ is defined by the root α and the same set $S(\beta)$ of sons as in T_D.

11.2.2 Algebraic Characterisation, Hierarchical Subspace Family

In the case of $\mathcal{T}_\mathbf{r}$, the definition in (8.3) is an algebraic one using subspaces of certain dimensions. The concrete tensor subspace representation $\mathbf{v} = \rho_{\mathrm{TS}}\big(\mathbf{a}, (B_j)_{j=1}^d\big)$ uses bases and coefficients. Similarly, we start here with a definition based on subspace properties and later in §11.3.1 introduce bases and coefficient matrices.

Let a dimension partition tree T_D together with a tensor $\mathbf{v} \in \mathbf{V} = {}_a\bigotimes_{j\in D} V_j$ be given. The hierarchical representation of $\mathbf{v}$ is characterised by finite dimensional subspaces[5]

$$\mathbf{U}_\alpha \subset \mathbf{V}_\alpha := {}_a\bigotimes_{j\in\alpha} V_j \qquad \text{for all } \alpha \in T_D\,. \tag{11.10}$$

The basis assumptions on $\mathbf{U}_\alpha$ depend on the nature of the vertex $\alpha \in T_D$. Here, we distinguish (a) the root $\alpha = D$, (b) leaves $\alpha \in \mathcal{L}(T_D)$, (c) non-leaf vertices $\alpha \in T_D\backslash\mathcal{L}(T_D)$.

- The aim of the construction is to obtain a subspace $\mathbf{U}_D$ at the root $D \in T_D$ such that
$$\mathbf{v} \in \mathbf{U}_D\,. \tag{11.11a}$$
Since $D \in T_D$ is not a leaf, also condition (11.11c) must hold.
- At a leaf $\alpha = \{j\} \in \mathcal{L}(T_D)$, $\mathbf{U}_\alpha = U_j$ is a subspace[6] of V_j (same situation as in (8.3) for $\mathcal{T}_\mathbf{r}$):
$$U_j \subset V_j \qquad \text{for all } j \in D, \text{ i.e., } \alpha = \{j\} \in \mathcal{L}(T_D). \tag{11.11b}$$

[5] We use the bold-face notation $\mathbf{U}_\alpha$ for tensor spaces, although for $\alpha \in \mathcal{L}(T_D)$, $\mathbf{U}_{\{j\}} = U_j$ is a subspace of the standard vector space V_j.

[6] We identify the notations V_j $(j \in D)$ and $\mathbf{V}_\alpha = \mathbf{V}_{\{j\}}$ for $\alpha = \{j\} \in \mathcal{L}(T_D)$. Similar for $U_j = \mathbf{U}_{\{j\}}$. Concerning the relation of $\mathcal{L}(T_D)$ and D see Remark 11.3b.

- For any vertex $\alpha \in T_D \backslash \mathcal{L}(T_D)$ with sons $\alpha_1, \alpha_2 \in S(\alpha)$, the subspace $\mathbf{U}_\alpha$ (cf. (11.10)) must be related to the subspaces $\mathbf{U}_{\alpha_1}$ and $\mathbf{U}_{\alpha_1}$ by the crucial *nestedness property*

$$\mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2} \qquad \text{for all } \alpha \in T_D \backslash \mathcal{L}(T_D),\ \alpha_1, \alpha_2 \in S(\alpha). \tag{11.11c}$$

Diagram (11.1) depicts these conditions for the tree T_D from Fig. 11.3. Note that $\mathbf{U}_{\{1,2\}} \subset \mathbf{U}_{\{1\}} \otimes \mathbf{U}_{\{2\}}$ is a *subspace,* but not a *tensor subspace* (cf. abstract in §8).

Remark 11.7. (a) Since the subspace $\mathbf{U}_D \subset \mathbf{V}$ has to satisfy only that $\mathbf{v} \in \mathbf{U}_D$ for a given tensor $\mathbf{v} \in \mathbf{V}$, it is sufficient to define $\mathbf{U}_D$ by

$$\mathbf{U}_D = \operatorname{span}\{\mathbf{v}\}, \qquad \text{i.e., } \dim(\mathbf{U}_D) = 1. \tag{11.12}$$

(b) Another situation arises if a family $F \subset \mathbf{U} = \bigotimes_{j \in D} U_j$ of tensors is to be represented (cf. §6.2.3). Then requirement (11.11a) is replaced by $F \subset \mathbf{U}_D$. The minimal choice is

$$\mathbf{U}_D = \operatorname{span}(F).$$

Definition 11.8 (hierarchical subspace family). (a) We call $\{\mathbf{U}_\alpha\}_{\alpha \in T_D}$ a *hierarchical subspace family* (associated with $\mathbf{V} = {}_a\bigotimes_{j \in D} V_j$), if T_D is a dimension partition tree of D and the subspaces $\mathbf{U}_\alpha$ satisfy (11.11b,c).
(b) We say that a tensor $\mathbf{v}$ is represented by the hierarchical subspace family $\{\mathbf{U}_\alpha\}_{\alpha \in T_D}$, if $\mathbf{v} \in \mathbf{U}_D$ (cf. (11.11a)).

Definition 11.9. A *successor* of $\alpha \in T_D$ is any $\sigma \in T_D$ with $\sigma \subset \alpha$. A set $\{\sigma_i\}$ of disjoint successors of $\alpha \in T_D$ is called *complete* if $\alpha = \dot{\bigcup}_i \sigma_i$.

For instance, in the case of the tree from Fig. 11.3, the set $\{\{1\}, \{2\}, \{3, 4\}\}$ is a complete set of successors of $\{1, 2, 3, 4\}$.

Lemma 11.10. *Let $\{\mathbf{U}_\alpha\}_{\alpha \in T_D}$ be a hierarchical subspace family. For any $\alpha \in T_D$ and any complete set Σ of successors of α,*

$$\mathbf{U}_\alpha \subset \bigotimes_{\sigma \in \Sigma} \mathbf{U}_\sigma \subset \mathbf{V}_\alpha \tag{11.13}$$

holds with $\mathbf{V}_\alpha$ from (11.10).

Proof. The fundamental structure (11.11c) implies $\mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$. The chain of inclusions can be repeated inductively to prove the first inclusion in (11.13). Since $\mathbf{U}_\sigma \subset \mathbf{V}_\sigma$ (cf. (11.10)), $\bigotimes_{\sigma \in \Sigma} \mathbf{U}_\sigma \subset \bigotimes_{\sigma \in \Sigma} \mathbf{V}_\sigma = \mathbf{V}_\alpha$ proves the last inclusion. □

The analogue of $\mathcal{T}_\mathbf{r}$ is the set $\mathcal{H}_\mathfrak{r}$ which is defined next. We start from given (bounds of) dimensions

$$\mathfrak{r} := (r_\alpha)_{\alpha \in T_D} \in \mathbb{N}_0^{T_D}, \tag{11.14}$$

and consider subspaces $\mathbf{U}_\alpha \subset \mathbf{V}_\alpha$ with $\dim(\mathbf{U}_\alpha) \leq r_\alpha$ for all $\alpha \in T_D$.

Definition 11.11 **($\mathcal{H}_\mathfrak{r}$).** Fix some $\mathfrak{r}$ from (11.14) and let $\mathbf{V} = {}_a\bigotimes_{k\in D} V_k$. Then $\mathcal{H}_\mathfrak{r} = \mathcal{H}_\mathfrak{r}(\mathbf{V}) \subset \mathbf{V}$ is the set[7]

$$\mathcal{H}_\mathfrak{r} := \left\{ \mathbf{v} \in \mathbf{V} : \begin{array}{l} \text{there is a hierarchical subspace family } \{\mathbf{U}_\alpha\}_{\alpha\in T_D} \\ \text{with } \mathbf{v} \in \mathbf{U}_D \text{ and } \dim(\mathbf{U}_\alpha) \le r_\alpha \text{ for all } \alpha \in T_D\,. \end{array} \right\} \tag{11.15}$$

11.2.3 Minimal Subspaces

We recall the definition of the minimal subspace associated to $\alpha \subset D$ with complement $\alpha^c = D\backslash\alpha$:

$$\mathbf{U}_\alpha^{\min}(\mathbf{v}) = \left\{ \varphi_{\alpha^c}(\mathbf{v}) : \varphi_{\alpha^c} \in \left(\bigotimes_{j\in\alpha^c} V_j \right)' \right\} \text{ for all } \alpha \in T_D\backslash\{D\} \tag{11.16a}$$

(cf. (6.14)). A possible computation uses the matricisation $\mathcal{M}_\alpha(\mathbf{v})$ from Definition 5.3. The (left-sided) singular value decomposition of $\mathcal{M}_\alpha(\mathbf{v})$ yields the data $\sigma_i^{(\alpha)}$ and $\mathbf{u}_i^{(\alpha)}$ of $\mathbf{v} = \sum_{i=1}^r \sigma_i^{(\alpha)} \mathbf{u}_i^{(\alpha)} \otimes \mathbf{v}_i^{(\alpha^c)}$ with $\sigma_1^{(\alpha)} \ge \ldots \ge \sigma_r^{(\alpha)} > 0$. Then

$$\mathbf{U}_\alpha^{\min}(\mathbf{v}) = \operatorname{span}\{\mathbf{u}_i^{(\alpha)} : 1 \le i \le r\}. \tag{11.16b}$$

For $\alpha = \{j\} \in \mathcal{L}(T_D)$, these subspaces coincide with the subspaces $U_j^{\min}(\mathbf{v})$ of the tensor subspace representation (cf. (6.10a)). For $\alpha \in T_D\backslash\mathcal{L}(T_D)$, the subspaces fulfil the *nestedness property*

$$\mathbf{U}_\alpha^{\min}(\mathbf{v}) \subset \mathbf{U}_{\alpha_1}^{\min}(\mathbf{v}) \otimes \mathbf{U}_{\alpha_2}^{\min}(\mathbf{v}) \qquad (\alpha_1, \alpha_2 \text{ sons of } \alpha) \tag{11.16c}$$

as stated in (6.13).

Next, we give a simple characterisation of the property $\mathbf{v} \in \mathcal{H}_\mathfrak{r}$. Moreover, the subspace family $\{\mathbf{U}_\alpha\}_{\alpha\in T_D}$ can be described explicitly.

Theorem 11.12. *For $\mathfrak{r} := (r_\alpha)_{\alpha\in T_D} \in \mathbb{N}_0^{T_D}$, a tensor $\mathbf{v}$ belongs to $\mathcal{H}_\mathfrak{r}$ if and only if $\operatorname{rank}_\alpha(\mathbf{v}) \le r_\alpha$ holds for all $\alpha \in T_D$. A possible hierarchical subspace family is given by $\{\mathbf{U}_\alpha^{\min}(\mathbf{v})\}_{\alpha\in T_D}$.*

Proof. 1) The α-rank definition $\operatorname{rank}_\alpha(\mathbf{v}) = \dim(\mathbf{U}_\alpha^{\min}(\mathbf{v}))$ (cf. (6.15)) and $\operatorname{rank}_\alpha(\mathbf{v}) \le r_\alpha$ imply the condition $\dim(\mathbf{U}_\alpha) \le r_\alpha$ for $\mathbf{U}_\alpha := \mathbf{U}_\alpha^{\min}(\mathbf{v})$ in (11.15). Property (11.16c) proves that $\{\mathbf{U}_\alpha\}_{\alpha\in T_D}$ is a hierarchical subspace family.

2) Assume $\mathbf{v} \in \mathcal{H}_\mathfrak{r}$ with some hierarchical subspace family $\{\mathbf{U}_\alpha\}_{\alpha\in T_D}$. Fix some $\alpha \in T_D$ and choose a complete set Σ of successors such that $\alpha \in \Sigma$. Statement (11.13) with D instead of α shows that $\mathbf{v} \in \mathbf{U}_D \subset \bigotimes_{\sigma\in\Sigma} \mathbf{U}_\sigma$. The definition of a minimal subspace implies $\mathbf{U}_\sigma^{\min}(\mathbf{v}) \subset \mathbf{U}_\sigma$, in particular, $\mathbf{U}_\alpha^{\min}(\mathbf{v}) \subset \mathbf{U}_\alpha$ for the fixed but arbitrary $\alpha \in T_D$. Hence, $\operatorname{rank}_\alpha(\mathbf{v}) = \dim(\mathbf{U}_\alpha^{\min}(\mathbf{v})) \le \dim(\mathbf{U}_\alpha) \le r_\alpha$ proves the reverse direction of the theorem. □

The proof reveals the following statement.

[7] By analogy with (8.3) one might prefer $\dim(\mathbf{U}_\alpha) = r_\alpha$ instead of $\dim(\mathbf{U}_\alpha) \le r_\alpha$. The reason for the latter choice is the fact that, otherwise, without the conditions (11.17) $\mathcal{H}_\mathfrak{r} = \emptyset$ may occur.

Corollary 11.13. $\mathfrak{r}=(r_\alpha)_{\alpha\in T_D}$ with $r_\alpha:=\dim(\mathbf{U}_\alpha^{\min}(\mathbf{v}))$ are the smallest integers so that $\mathbf{v}\in\mathcal{H}_\mathfrak{r}$. This tuple $\mathfrak{r}$ is called the *hierarchical rank* of $\mathbf{v}$.

Remark 11.14. Let $\alpha\in T_D\backslash\{D\}$ be a vertex with sons $\alpha_1,\alpha_2\in S(\alpha)$ so that $\mathbf{U}_\alpha^{\min}(\mathbf{v})\subset\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})\otimes\mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})$. Then $\mathbf{U}_{\alpha_i}^{\min}(\mathbf{v})$ can be interpreted differently:

$$\mathbf{U}_{\alpha_i}^{\min}(\mathbf{v})=\mathbf{U}_{\alpha_i}^{\min}(F)\quad\text{for } F:=\mathbf{U}_\alpha^{\min}(\mathbf{v}),\ i=1,2\qquad\text{(cf. (6.10c))}.$$

Proof. $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ is the span of all $\mathbf{u}_i^{(\alpha)}$ appearing in the reduced singular value decomposition $\mathbf{v}=\sum_{i=1}^{r_\alpha}\sigma_i^{(\alpha)}\,\mathbf{u}_i^{(\alpha)}\otimes\mathbf{v}_i^{(\alpha^c)}$. Hence, $\mathbf{U}_{\alpha_1}^{\min}(F)=\sum_{i=1}^{r_\alpha}\mathbf{U}_{\alpha_1}^{\min}(\mathbf{u}_i^{(\alpha)})$ holds. The singular value decomposition of each $\mathbf{u}_i^{(\alpha)}$ by

$$\mathbf{u}_i^{(\alpha)}=\sum_{j=1}^{r}\tau_j^{(i)}\,\mathbf{a}_j^{(i)}\otimes\mathbf{b}_j^{(i)}\quad\text{with } \mathbf{a}_j^{(i)}\in\mathbf{V}_{\alpha_1},\ \mathbf{b}_j^{(i)}\in\mathbf{V}_{\alpha_2},\ \tau_1^{(i)}\geq\ldots\geq\tau_r^{(i)}>0$$

yields $\mathbf{U}_{\alpha_1}^{\min}(\mathbf{u}_i^{(\alpha)})=\operatorname{span}\{\mathbf{a}_j^{(i)}:1\leq j\leq r\}$ so that

$$\mathbf{U}_{\alpha_1}^{\min}(F)=\operatorname{span}\left\{\mathbf{a}_j^{(i)}:1\leq j\leq r,\ 1\leq i\leq r_\alpha\right\}.$$

There are functionals $\beta_j^{(i)}\in\mathbf{V}'_{\alpha_2}$ with $\beta_j^{(i)}(\mathbf{b}_k^{(i)})=\delta_{jk}$ (cf. (2.1)). As stated in (11.16a), $\mathbf{u}_i^{(\alpha)}=\varphi_{\alpha^c}(\mathbf{v})$ holds for some $\varphi_{\alpha^c}\in\mathbf{V}'_{\alpha^c}$. The functional $\varphi:=\beta_j^{(i)}\otimes\varphi_{\alpha^c}$ belongs to $\mathbf{V}'_{\alpha_1^c}$ (note that $\alpha_1^c=\alpha_2\cup\alpha^c$). It follows from (11.16a) that

$$\varphi(\mathbf{v})=\beta_j^{(i)}(\mathbf{u}_i^{(\alpha)})=\sum_{k=1}^{r}\tau_k^{(i)}\,\beta_j^{(i)}(\mathbf{b}_k^{(i)})\,\mathbf{a}_k^{(i)}=\tau_j^{(i)}\,\mathbf{a}_j^{(i)}\in\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v}).$$

Hence, $\mathbf{a}_j^{(i)}\in\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})$ for all i,j, i.e., $\mathbf{U}_{\alpha_1}^{\min}(F)\subset\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})$. On the other hand, $\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})\subset\mathbf{U}_{\alpha_1}^{\min}(F)$ follows from $\mathbf{v}=\sum_{i=1}^{r_\alpha}\sum_{j=1}^{r}\tau_j^{(i)}\sigma_{\alpha,i}\,\mathbf{a}_j^{(i)}\otimes\big(\mathbf{b}_j^{(i)}\otimes\mathbf{v}_i^{(\alpha^c)}\big)$. □

Finally, we consider the reverse setting: we specify dimensions r_α and construct a tensor $\mathbf{v}$ such that $r_\alpha=\operatorname{rank}_\alpha(\mathbf{v})=\dim(\mathbf{U}_\alpha^{\min}(\mathbf{v}))$. The following restrictions are necessary:

$$\begin{array}{ll} r_\alpha\leq r_{\alpha_1}r_{\alpha_2},\ r_{\alpha_1}\leq r_{\alpha_2}r_\alpha,\ r_{\alpha_2}\leq r_{\alpha_1}r_\alpha, & \text{for } \alpha\in T_D\backslash\mathcal{L}(T_D),\\ r_\alpha\leq\dim(V_j) & \text{for } \alpha=\{j\}\in\mathcal{L}(T_D),\\ r_D=1 & \text{for } \alpha=D, \end{array}\tag{11.17}$$

where α_1,α_2 are the sons of α. The first line follows from (6.17b) and Corollary 6.20a (note that $\operatorname{rank}_\alpha(\mathbf{v})=\operatorname{rank}_{\alpha^c}(\mathbf{v})$).

Lemma 11.15. *Let* $\mathfrak{r}:=(r_\alpha)_{\alpha\in T_D}\in\mathbb{N}^{T_D}$ *satisfy (11.17). Then there is a tensor* $\mathbf{v}\in\mathcal{H}_\mathfrak{r}(\mathbf{V})$ *with* $\operatorname{rank}_\alpha(\mathbf{v})=\dim(\mathbf{U}_\alpha^{\min}(\mathbf{v}))=r_\alpha$.

Proof. 1) We construct the subspaces $\mathbf{U}_\alpha\subset\mathbf{V}_\alpha$ from the leaves to the root. For $\alpha=\{j\}\in T_D\backslash\mathcal{L}(T_D)$ choose any $\mathbf{U}_{\{j\}}$ of dimension $r_{\{j\}}$ (here, the second line in (11.17) is needed).

2) Assume that for a vertex $\alpha \in T_D \backslash \mathcal{L}(T_D)$ the subspaces $\mathbf{U}_{\alpha_1}$ and $\mathbf{U}_{\alpha_2}$ for the leaves with dimensions r_{α_1} and r_{α_2} are already constructed. Choose any bases $\{\mathbf{b}_\ell^{(\alpha_i)} : 1 \leq \ell \leq r_{\alpha_i}\}$ of $\mathbf{U}_{\alpha_i}$ $(i=1,2)$. Without loss of generality assume $r_{\alpha_1} \geq r_{\alpha_2}$. For the most critical case $r_{\alpha_1} = r_{\alpha_2} r_\alpha$ set $\mathbf{U}_\alpha := \operatorname{span}\{\mathbf{b}_\ell^{(\alpha)} : 1 \leq \ell \leq r_\alpha\}$ with

$$\mathbf{b}_\ell^{(\alpha)} := \sum_{i,j=1}^{r_{\alpha_2}} \mathbf{b}_{i+(\ell-1)r_{\alpha_2}}^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}.$$

One observes that $\mathbf{b}_\ell^{(\alpha)} \in \mathbf{V}_\alpha$ satisfies

$$\mathbf{U}_{\alpha_1}^{\min}(\mathbf{b}_\ell^{(\alpha)}) = \operatorname{span}\{\mathbf{b}_{i+(\ell-1)r_{\alpha_2}}^{(\alpha_1)} : 1 \leq i \leq r_{\alpha_2}\}, \quad \text{while } \mathbf{U}_{\alpha_2}^{\min}(\mathbf{b}_\ell^{(\alpha)}) = \mathbf{U}_{\alpha_2}.$$

From Exercise 6.14 we conclude that $\mathbf{U}_{\alpha_1}^{\min}(\{\mathbf{b}_\ell^{(\alpha)} : 1 \leq \ell \leq r_\alpha\}) = \mathbf{U}_{\alpha_1}^{\min}(\mathbf{U}_\alpha) = \operatorname{span}\{\mathbf{b}_i^{(\alpha_1)} : 1 \leq i \leq r_{\alpha_2} r_\alpha\} \underset{r_{\alpha_1} = r_{\alpha_2} r_\alpha}{=} \mathbf{U}_{\alpha_1}$. For $r_{\alpha_1} < r_{\alpha_2} r_\alpha$ (but $r_\alpha \leq r_{\alpha_1} r_{\alpha_2}$) it is easy to produce linearly independent $\{\mathbf{b}_\ell^{(\alpha)} : 1 \leq \ell \leq r_\alpha\}$ with $\mathbf{U}_{\alpha_i}^{\min}(\mathbf{U}_\alpha) = \mathbf{U}_{\alpha_i}$.

3) For $\alpha = D$, the first and third lines of (11.17) imply that $r_{\alpha_1} = r_{\alpha_2}$. Set $\mathbf{v} := \sum_{i=1}^{r_{\alpha_1}} \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_i^{(\alpha_2)}$ and $\mathbf{U}_D = \operatorname{span}\{\mathbf{v}\}$. Obviously, $\mathbf{U}_{\alpha_i}^{\min}(\mathbf{v}) = \mathbf{U}_{\alpha_i}$ holds for $i = 1, 2$, proving the assertion $r_{\alpha_i} = \operatorname{rank}_{\alpha_i}(\mathbf{v})$. For the further vertices use Remark 11.14. Induction from the root to the leaves shows that $\mathbf{U}_\alpha = \mathbf{U}_\alpha^{\min}(\mathbf{v})$ implies $\mathbf{U}_{\alpha_i}^{\min}(\mathbf{U}_\alpha) = \mathbf{U}_{\alpha_i}^{\min}(\mathbf{v})$. Because of the identities $\mathbf{U}_{\alpha_i}^{\min}(\mathbf{U}_\alpha) = \mathbf{U}_{\alpha_i}$ and $\dim(\mathbf{U}_{\alpha_i}) = \operatorname{rank}_{\alpha_i}(\mathbf{v}) = r_{\alpha_i}$, the lemma is proved. □

Remark 11.16. With probability one a random tensor from $\bigotimes_{j\in D} \mathbb{R}^{n_j}$ possesses the maximal hierarchical rank $\mathfrak{r}$ with

$$r_\alpha = \min\left\{\prod_{j\in\alpha} n_j, \prod_{j\in D\backslash\alpha} n_j\right\} \qquad (\alpha \in T_D).$$

Proof. Apply Remark 2.5 to the matrix $\mathcal{M}_\alpha(\mathbf{v})$ and note that $r_\alpha = \operatorname{rank}(\mathcal{M}_\alpha(\mathbf{v}))$. □

11.2.4 Conversions

Tensors from $\mathcal{T}_\mathbf{r}$ or $\mathcal{R}_r$ can be represented exactly in $\mathcal{H}_\mathfrak{r}$ with at least similar storage cost.

11.2.4.1 Conversion from $\mathcal{T}_\mathbf{r}$ to $\mathcal{H}_\mathfrak{r}$, Maximal Subspaces

Assume that a tensor is given in the tensor subspace representation:

$$\mathbf{v} \in \mathbf{U} = \bigotimes_{j\in D} U_j \subset \mathcal{T}_\mathbf{r} \subset \mathbf{V} = \bigotimes_{j\in D} V_j$$

with $\dim(U_j) = r_j$. The *maximal* choice of the subspaces is

$$\mathbf{U}_\alpha := \begin{cases} U_j & \text{for } \alpha = \{j\}, \\ \bigotimes_{j\in\alpha} U_j & \text{for } \alpha \in T_D \backslash \mathcal{L}(T_D). \end{cases} \tag{11.18}$$

From $\mathbf{v} \in \mathcal{T}_\mathbf{r}$ we derive that $\dim(U_j) = r_j$ and, in general, $\dim(\mathbf{U}_\alpha) = \prod_{j\in\alpha} r_j$ for $\alpha \neq D$. The large dimension $\dim(\mathbf{U}_\alpha) = \prod_{j\in\alpha} r_j$ corresponds exactly to the large data size of the coefficient tensor $\mathbf{a} \in \bigotimes_{j\in D} \mathbb{K}^{r_j}$ from (8.6). On the positive side, this approach allows us to represent *any* $\mathbf{v} \in \mathbf{U}$ in the hierarchical format $\{\mathbf{U}_\alpha\}_{\alpha\in T_D}$.

11.2.4.2 Conversion from $\mathcal{R}_r$

Assume an r-term representation of $\mathbf{v} \in \mathcal{R}_r \subset \mathbf{V} = \bigotimes_{j\in D} V_j$ by

$$\mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j\in D} v_i^{(j)} \qquad \text{with } v_i^{(j)} \in V_j.$$

Set

$$\mathbf{U}_\alpha := \begin{cases} \operatorname{span}\left\{ \bigotimes_{j\in\alpha} v_i^{(j)} : 1 \leq i \leq r \right\} & \text{for } \alpha \neq D, \\ \operatorname{span}\{\mathbf{v}\} & \text{for } \alpha = D. \end{cases} \tag{11.19}$$

Obviously, conditions (11.11a-c) are fulfilled. This proves the next statement.

Theorem 11.17. *Let T_D be any dimension partition tree. Then $\mathbf{v} \in \mathcal{R}_r \subset \mathbf{V} = \bigotimes_{j\in D} V_j$ belongs to $\mathcal{H}_\mathfrak{r}$ with $r_\alpha = r$ for $\alpha \neq D$ and $r_\alpha = 1$ for $\alpha = D$.*

Conversion from $\mathcal{R}_r$ to hierarchical format will be further discussed in §11.3.5.

11.3 Construction of Bases

As for the tensor subspace format, the involved subspaces have to be characterised by frames or bases. The particular problem in the case of the hierarchical format is the fact that a basis $[\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}]$ of $\mathbf{U}_\alpha$ consists of tensors $\mathbf{b}_i^{(\alpha)} \in \mathbf{V}_\alpha$ of order $\#\alpha$. A representation of $\mathbf{b}_i^{(\alpha)}$ by its entries would require a huge storage. It is essential to describe the vectors $\mathbf{b}_i^{(\alpha)}$ indirectly by means of the frames associated to the sons $\alpha_1, \alpha_2 \in S(\alpha)$. The general concept of the hierarchical representation is explained in §11.3.1. Of particular interest is the performance of basis transformations. Usually, one prefers an orthonormal basis representation which is discussed in §11.3.2. A special orthonormal basis is defined by the higher order singular value decomposition (HOSVD). Its definition and construction are given in §11.3.3. In §11.3.4, a sensitivity analysis is given, which describes how perturbations of the data influence the tensor. Finally, in §11.3.5, we mention that the conversion of r-term tensors into the hierarchical format yields very particular coefficient matrices.

11.3.1 Hierarchical Bases Representation

The term ‘bases’ in the heading may be replaced more generally by ‘frames’.

11.3.1.1 Basic Structure

In the most general case, the subspace $\mathbf{U}_\alpha$ ($\alpha \in T_D$) is generated by a frame:

$$\mathbf{B}_\alpha = \left[\mathbf{b}_1^{(\alpha)}, \mathbf{b}_2^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}\right] \in (\mathbf{U}_\alpha)^{r_\alpha}, \tag{11.20a}$$

$$\mathbf{U}_\alpha = \mathrm{span}\{\mathbf{b}_i^{(\alpha)} : 1 \le i \le r_\alpha\} \quad \text{for all } \alpha \in T_D. \tag{11.20b}$$

Except for $\alpha \in \mathcal{L}(T_D)$, the tensors $\mathbf{b}_i^{(\alpha)} \in \mathbf{U}_\alpha$ are not represented as full tensor. Therefore, the frame vectors $\mathbf{b}_i^{(\alpha)}$ serve for *theoretical purpose* only, while other data will be used in the later representation of a tensor $\mathbf{v} \in \mathbf{V}$. The integer[8] r_α is defined by (11.20b) denoting the size of the frame.

Concerning the choice of $\mathbf{B}_\alpha$ in (11.20a,b), the following possibilities exist:

1. **Frame.** A frame $\mathbf{B}_\alpha \in (\mathbf{U}_\alpha)^{r_\alpha}$ cannot be avoided as an intermediate representation (cf. §11.5.2), but usually one of the following choices is preferred.
2. **Basis.** If $\mathbf{B}_\alpha$ is a basis, the number r_α coincides with the dimension:
$$r_\alpha := \dim(\mathbf{U}_\alpha) \qquad \text{for all } \alpha \in T_D\,. \tag{11.21}$$
3. **Orthonormal basis.** Assuming a scalar product in $\mathbf{U}_\alpha$, we can construct an orthonormal basis $\mathbf{B}_\alpha$ (cf. §11.3.2).
4. **HOSVD.** The higher order singular value decomposition from §8.3 can be applied again and leads to a particular orthonormal basis $\mathbf{B}_\alpha$ (cf. §11.3.3).

Concerning the practical realisation, we have to distinguish leaves $\alpha \in \mathcal{L}(T_D)$ from non-leaf vertices.

Leaf vertices. Leaves $\alpha \in \mathcal{L}(T_D)$ are characterised by $\alpha = \{j\}$ for some $j \in D$. The subspace $U_j \subset V_j$ refers to V_j from $\mathbf{V} = {}_a\bigotimes_{j\in D} V_j$ and is characterised by a frame or basis $B_j = \left[b_1^{(j)}, \ldots, b_{r_j}^{(j)}\right] \in (U_j)^{r_j}$ from above. The vectors $b_i^{(j)}$ are stored directly.

Remark 11.18. (a) If $V_j = \mathbb{K}^{I_j}$, the memory cost for $B_j \in \mathbb{K}^{I_j \times r_j}$ is $r_j \# I_j$.
(b) Depending on the nature of the vector space V_j, one may use other data-sparse representations of $b_i^{(j)}$ (cf. §7.5 and §14.1.4.3).

Non-leaf vertices $\alpha \in T_K \backslash \mathcal{L}(T_K)$**.** The sons of α are denoted by $\alpha_1, \alpha_2 \in S(\alpha)$. Let $\mathbf{b}_i^{(\alpha_1)}$ and $\mathbf{b}_j^{(\alpha_2)}$ be the columns of the respective frames [bases] $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$. Then the tensor space $\mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ has the *canonical frame [basis]* consisting of the tensor products of the frame [basis] vectors of $\mathbf{U}_{\alpha_1}$ and $\mathbf{U}_{\alpha_2}$ as detailed below.

[8] See Footnote 6 for the notation $r_j = r_{\{j\}}$. Similarly, $b_i^{(j)} = \mathbf{b}_i^{(\{j\})}$ etc.

Remark 11.19. Let $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$ be generating systems of $\mathbf{U}_{\alpha_1}$ and $\mathbf{U}_{\alpha_2}$. Define the tuple $\mathcal{B}$ and the tensors $\mathbf{b}_{ij}^{(\alpha)} \in \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ by

$$\mathcal{B} := (\mathbf{b}_{ij}^{(\alpha)} := \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} : 1 \le i \le r_{\alpha_1}, 1 \le j \le r_{\alpha_2}). \tag{11.22}$$

(a) If $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$ are frames, $\mathcal{B}$ is a frame of $\mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$.
(b) If $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$ are bases, $\mathcal{B}$ is a basis of $\mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ (cf. Lemma 3.11a).
(c) If $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$ are orthonormal bases, $\mathcal{B}$ is an orthonormal basis of $\mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ (cf. Remark 4.125).

As a consequence, any tensor $\mathbf{w} \in \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ and, in particular, $\mathbf{w} \in \mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$, can be written in the form[9]

$$\mathbf{w} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha)} \mathbf{b}_{ij}^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha)}\, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}. \tag{11.23a}$$

Since the frame vectors $\mathbf{b}_{ij}^{(\alpha)}$ carry two indices, the coefficient vector $\left(c_{ij}^{(\alpha)}\right)$ has the special form of a *coefficient matrix*

$$C^{(\alpha)} = \left(c_{ij}^{(\alpha)}\right)_{\substack{i=1,\ldots,r_{\alpha_1} \\ j=1,\ldots,r_{\alpha_2}}} \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2}}. \tag{11.23b}$$

If $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$ are bases, the one-to-one correspondence between a tensor $\mathbf{w} \in \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ and its coefficient matrix $C^{(\alpha)}$ defines an isomorphism which we denote by

$$\Theta_\alpha : \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2} \to \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2}} \quad \text{for } \alpha \in T_D \backslash \mathcal{L}(T_D). \tag{11.23c}$$

The fact that $\mathbf{w} \in \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ can be coded by $r_{\alpha_1} \cdot r_{\alpha_2}$ numbers, is independent of how the frame vectors $\mathbf{b}_i^{(\alpha_1)}, \mathbf{b}_j^{(\alpha_2)}$ are represented. They may be given directly (as for the leaves $\alpha_\ell \in \mathcal{L}(T_D)$) or indirectly (as for non-leaves $\alpha_\ell \in T_D \backslash \mathcal{L}(T_D)$).

Now, we apply the representation (11.23a,b) to the frame vectors $\mathbf{b}_\ell^{(\alpha)} \in \mathbf{U}_\alpha$ from $\mathbf{B}_\alpha = \left[\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}\right]$ and denote the coefficient matrix by $C^{(\alpha,\ell)} \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2}}$:

$$\boxed{\begin{array}{ll} \mathbf{b}_\ell^{(\alpha)} = \sum\limits_{i=1}^{r_{\alpha_1}} \sum\limits_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)}\, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} & \{\alpha_1, \alpha_2\} = S(\alpha) \\ \text{with } C^{(\alpha,\ell)} = \left(c_{ij}^{(\alpha,\ell)}\right)_{\substack{1 \le i \le r_{\alpha_1} \\ 1 \le j \le r_{\alpha_2}}} \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2}} & \text{for } 1 \le \ell \le r_\alpha. \end{array}} \tag{11.24}$$

This is the key relation of the hierarchical format.

Remark 11.20. The formulation of the coefficient matrix $C^{(\alpha,\ell)}$ depends of the ordering of the sons α_1, α_2. If the sons are interchanged, $C^{(\alpha,\ell)}$ changes into the transposed matrix $C^{(\alpha,\ell)\mathsf{T}}$.

[9] In the case of a frame, the coefficients $c_{ij}^{(\alpha)}$ are not uniquely determined.

We summarise: Only for leaves $\alpha \in \mathcal{L}(T_D)$, the basis vectors $b_i^{(j)}$ are explicitly represented. For all other vertices, the vectors $\mathbf{b}_\ell^{(\alpha)} \in \mathbf{U}_\alpha$ are defined recursively by means of the coefficient matrices[10] $C^{(\alpha,\ell)}$. The practical representation of a tensor $\mathbf{v} \in \mathbf{V}$ uses the data $C^{(\alpha,\ell)}$ for $\alpha \in T_D \backslash \mathcal{L}(T_D)$ and $b_i^{(j)}$ for $\{j\} \in \mathcal{L}(T_D)$ only, while the theoretical discussion may still refer to $\mathbf{b}_\ell^{(\alpha)}$ and their properties.

One obtains, in particular, a frame [basis] $\mathbf{B}_D$ for the root $D \in T_D$. We can represent all tensors of $\mathbf{v} \in \mathbf{U}_D$ by a coefficient vector $c^{(D)} \in \mathbb{K}^{r_D}$:

$$\mathbf{v} = \sum_{i=1}^{r_D} c_i^{(D)} \mathbf{b}_i^{(D)}. \tag{11.25}$$

Remark 11.21. Since, usually, the basis $\mathbf{B}_D$ of $\mathbf{U}_D$ consists of one basis vector only (cf. Remark 11.7a), one might avoid the coefficient $c_1^{(D)}$ by choosing $\mathbf{b}_1^{(D)} = \mathbf{v}$ and $c_1^{(D)} = 1$. However, for systematic reasons (orthonormal basis, basis transforms), it is advantageous to separate the choice of the basis vector $\mathbf{b}_1^{(D)}$ from the value of $\mathbf{v}$.

11.3.1.2 Explicit Description

The definition of the basis vectors is recursive. Correspondingly, all operations will be performed recursively. In the following, we give an explicit description of the tensor $\mathbf{v}$ represented in the hierarchical format. However, this description will not be used for practical purposes.

Renaming ℓ by $\ell[\alpha]$, i by $\ell[\alpha_1]$, j by $\ell[\alpha_2]$, we rewrite (11.24) by

$$\mathbf{b}_{\ell[\alpha]}^{(\alpha)} = \sum_{\ell[\alpha_1]=1}^{r_{\alpha_1}} \sum_{\ell[\alpha_2]=1}^{r_{\alpha_2}} c_{\ell[\alpha_1],\ell[\alpha_2]}^{(\alpha,\ell[\alpha])} \mathbf{b}_{\ell[\alpha_1]}^{(\alpha_1)} \otimes \mathbf{b}_{\ell[\alpha_2]}^{(\alpha_2)}.$$

Insertion of the definitions of $\mathbf{b}_{\ell[\alpha_1]}^{(\alpha_1)}$ and $\mathbf{b}_{\ell[\alpha_2]}^{(\alpha_2)}$ yields

$$\mathbf{b}_{\ell[\alpha]}^{(\alpha)} = \sum_{\substack{\ell[\beta]=1 \\ \text{for } \beta \in T_\alpha \backslash \{\alpha\}}}^{r_\beta} \prod_{\beta \in T_\alpha \backslash \mathcal{L}(T_\alpha)} c_{\ell[\beta_1],\ell[\beta_2]}^{(\beta,\ell[\beta])} \bigotimes_{j=1}^{d} b_{\ell[\{j\}]}^{(j)} \quad (\beta_1, \beta_2 \text{ sons of } \beta).$$

The multiple summation involves all variable $\ell[\beta] \in \{1, \ldots, r_\beta\}$ with $\beta \in T_\alpha \backslash \{\alpha\}$, where T_α is the subtree from Definition 11.6.

The tensor $\mathbf{v} = \sum_{\ell=1}^{r_D} c_\ell^{(D)} \mathbf{b}_\ell^{(D)}$ (cf. (11.25)) has the representation

$$\mathbf{v} = \sum_{\substack{\ell[\alpha]=1 \\ \text{for } \alpha \in T_D}}^{r_\alpha} c_{\ell[D]}^{(D)} \prod_{\beta \in T_D \backslash \mathcal{L}(T_D)} c_{\ell[\beta_1],\ell[\beta_2]}^{(\beta,\ell[\beta])} \bigotimes_{j=1}^{d} b_{\ell[\{j\}]}^{(j)}. \tag{11.26}$$

[10] Note that also in wavelet representations, basis vectors do not appear explicitly. Instead the filter coefficients are used for the transfer of the basis vectors.

To interpret (11.26) correctly, note that β_1 or β_2 in $c^{(\beta,\ell[\beta])}_{\ell[\beta_1],\ell[\beta_2]}$ may belong to $\mathcal{L}(T_D)$, e.g., $\beta_1 = \{j\}$ for some $j \in D$. Then $\ell[\beta_1]$ coincides with the index $\ell[\{j\}]$ of $b^{(j)}_{\ell[\{j\}]}$.

For the case of $D = \{1,2,3\}$ and $S(D) = \{\{1,2\},\{3\}\}$, Eq. (11.26) becomes

$$\mathbf{v} = \sum_{\ell_1=1}^{r_1}\sum_{\ell_2=1}^{r_2}\sum_{\ell_3=1}^{r_3} \underbrace{c^{(D)}_1 \cdot \sum_{\nu=1}^{r_{\{1,2\}}} c^{(\{1,2\},\nu)}_{\ell_1,\ell_2} \cdot c^{(D,1)}_{\nu,\ell_3}}_{=:\,\mathbf{a}[\ell_1,\ell_2,\ell_3]} \bigotimes_{j=1}^{d} b^{(j)}_{\ell_j},$$

where we have assumed the standard case $r_D = 1$. The summation variables are $\nu = \ell[\{1,2\}]$, $\ell_j = \ell[\{j\}]$. When using minimal ranks, we obtain $r_3 = r_{\{1,2\}}$ (cf. (6.17a)). Hence, $\mathbf{C}_{\{1,2\}}$ can be considered as a tensor from $\mathbb{K}^{r_1\times r_2\times r_3}$ which has the same size as the coefficient tensor $\mathbf{a}$ of the tensor subspace representation.

We conclude from the last example that for $d = 3$ the tensor subspace format and the hierarchical format require almost the same storage (the data $c^{(D)}$ and $c^{(D,1)}$ are negligible compared with $\mathbf{C}_{\{1,2\}}$).

11.3.1.3 Hierarchical Representation

Equation (11.26) shows that $\mathbf{v}$ is completely determined by the data $C^{(\alpha,\ell)}$ (cf. (11.24)), $c^{(D)}$ (cf. (11.25)), and the bases B_j for the leaves $j \in D$. Also the frames $\mathbf{B}_\alpha$ for $\alpha \in T_D\backslash\mathcal{L}(T_D)$ are implicitly given by these data. The coefficient matrices $C^{(\alpha,\ell)}$ at vertex $\alpha \in T_D\backslash\mathcal{L}(T_D)$ are gathered in the tuple[11]

$$\mathbf{C}_\alpha := \left(C^{(\alpha,\ell)}\right)_{1\le\ell\le r_\alpha} \in \left(\mathbb{K}^{r_{\alpha_1}\times r_{\alpha_2}}\right)^{r_\alpha} \quad \text{for all } \alpha \in T_D\backslash\mathcal{L}(T_D). \tag{11.27}$$

Hence, the formal description of the hierarchical tensor representation is[12]

$$\mathbf{v} = \rho_{\mathrm{HTR}}\Big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\Big). \tag{11.28}$$

Remark 11.22. (a) The data size of T_D, $(\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}$, $c^{(D)}$, and $(B_j)_{j\in D}$ is

$$\begin{aligned} N^{\mathrm{HTR}}_{\mathrm{mem}}(T_D) &= 2\#D-1 \text{ vertices}, \\ N^{\mathrm{HTR}}_{\mathrm{mem}}((\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}) &= \sum_{\alpha\in T_D\backslash\mathcal{L}(T_D)} r_\alpha r_{\alpha_1} r_{\alpha_2} && (\alpha_1,\alpha_2 \text{ sons of } \alpha) \\ N^{\mathrm{HTR}}_{\mathrm{mem}}(c^{(D)}) &= r_D && (\text{cf. Remark 11.21}), \\ N^{\mathrm{HTR}}_{\mathrm{mem}}((B_j)_{j\in D}) &= \sum_{j=1}^{d} r_j \cdot size(U_j) && (\text{cf. Remark 8.7a}). \end{aligned} \tag{11.29a}$$

(b) Suppose $r_j = r$ for all $j \in D$. Then the data size $N^{\mathrm{HTR}}_{\mathrm{mem}}((B_j)_{j\in D})$ is the same as $N^{\mathrm{r\text{-}term}}_{\mathrm{mem}}$ for the r-term representation or $N^{\mathrm{TSR}}_{\mathrm{mem}}((B_j)_{j\in D})$ for the tensor

[11] $\mathbf{C}_\alpha$ may be viewed as $\Theta_\alpha(\mathbf{B}_\alpha)$, where the isomorphism Θ_α from (11.23c) is extended from $\mathbf{U}_\alpha$ to $(\mathbf{U}_\alpha)^{r_\alpha}$. In [73], $\mathbf{C}_\alpha$ is called 'transfer tensor'; cf. §11.3.1.8.

[12] HTR abbreviates 'hierarchical tensor representation'.

subspace representation. The terms $N_{\mathrm{mem}}^{\mathrm{HTR}}(T_D) + N_{\mathrm{mem}}^{\mathrm{HTR}}(c^{(D)})$ may be neglected. The dominant parts are

$$N_{\mathrm{mem}}^{\mathrm{HTR}}\left((B_j)_{j\in D}\right) \quad \text{and} \quad N_{\mathrm{mem}}^{\mathrm{HTR}}\left((\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}\right).$$

If $\mathbf{V} = \mathbb{K}^{\mathbf{I}}$ with $\mathbf{I} = I_1 \times \ldots \times I_d$ and $\#I_j = n$, full representation of the basis vectors leads to

$$N_{\mathrm{mem}}^{\mathrm{HTR}}\left((B_j)_{j\in D}\right) = d \cdot r \cdot n, \tag{11.29b}$$

while

$$N_{\mathrm{mem}}^{\mathrm{HTR}}\left((\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}\right) = (d-1)\, r^3. \tag{11.29c}$$

Proof. For T_D compare Remark 11.3. The coefficient matrix $C^{(\alpha,\ell)} \in \mathbb{K}^{r_{\alpha_1}\times r_{\alpha_2}}$ contains $r_{\alpha_1}r_{\alpha_2}$ entries. Since $1 \le \ell \le r_\alpha$, the size of $\mathbf{C}_\alpha$ is $r_{\alpha_1}r_{\alpha_2}r_\alpha$ for each $\alpha \in T_D\backslash\mathcal{L}(T_D)$. Eq. (11.29c) follows from $\#\left(T_D\backslash\mathcal{L}(T_D)\right) = \#T_D - \#\mathcal{L}(T_D) = (2d-1) - d = d-1$ (cf. Remark 11.3c). □

11.3.1.4 Transformations

There will be various reasons to change the frame from (11.20a) into another one. In general, even the generated subspaces may vary. For a vertex $\alpha \in T_D$ we consider the following 'old' and 'new' frames and subspaces:

$$\begin{aligned} \mathbf{B}_\alpha^{\mathrm{new}} &= \left[\mathbf{b}_{1,\mathrm{new}}^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha^{\mathrm{new}},\mathrm{new}}^{(\alpha)}\right], &\qquad \mathbf{U}_\alpha^{\mathrm{new}} &= \mathrm{range}\{\mathbf{B}_\alpha^{\mathrm{new}}\},\\ \mathbf{B}_\alpha^{\mathrm{old}} &= \left[\mathbf{b}_{1,\mathrm{old}}^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha^{\mathrm{old}},\mathrm{old}}^{(\alpha)}\right], &\qquad \mathbf{U}_\alpha^{\mathrm{old}} &= \mathrm{range}\{\mathbf{B}_\alpha^{\mathrm{old}}\},\end{aligned}$$

The replacement $\mathbf{B}_\alpha^{\mathrm{old}} \mapsto \mathbf{B}_\alpha^{\mathrm{new}}$ creates new coefficient matrices $C_{\mathrm{new}}^{(\alpha,\ell)}$ (cf. Lemma 11.23). Moreover, if $\alpha \neq D$, the coefficient matrices $C_{\mathrm{old}}^{(\beta,\ell)}$ associated to the father $\beta \in T_D$ of α must be renewed into $C_{\mathrm{new}}^{(\beta,\ell)}$, since these coefficients refer to $\mathbf{B}_\alpha^{\mathrm{new}}$ (cf. Lemma 11.24). If $\alpha = D$, the coefficient vector $c^{(D)}$ must be transformed instead (cf. Lemma 11.26).

We distinguish three different situations:

Case A. $\mathbf{B}_\alpha^{\mathrm{old}}$ and $\mathbf{B}_\alpha^{\mathrm{new}}$ generate the *same* subspace $\mathbf{U}_\alpha = \mathbf{U}_\alpha^{\mathrm{new}} = \mathbf{U}_\alpha^{\mathrm{old}}$. Then there are *transformation matrices* $T^{(\alpha)} \in \mathbb{K}^{r_\alpha^{\mathrm{new}}\times r_\alpha^{\mathrm{old}}}$, $S^{(\alpha)} \in \mathbb{K}^{r_\alpha^{\mathrm{old}}\times r_\alpha^{\mathrm{new}}}$ such that

$$\mathbf{B}_\alpha^{\mathrm{old}} = \mathbf{B}_\alpha^{\mathrm{new}} T^{(\alpha)}, \quad \text{i.e., } \mathbf{b}_{j,\mathrm{old}}^{(\alpha)} = \sum_{k=1}^{r_\alpha^{\mathrm{new}}} T_{kj}^{(\alpha)}\, \mathbf{b}_{k,\mathrm{new}}^{(\alpha)} \quad (1 \le j \le r_\alpha^{\mathrm{old}}), \tag{11.30a}$$

$$\mathbf{B}_\alpha^{\mathrm{new}} = \mathbf{B}_\alpha^{\mathrm{old}} S^{(\alpha)}, \quad \text{i.e., } \mathbf{b}_{k,\mathrm{new}}^{(\alpha)} = \sum_{j=1}^{r_\alpha^{\mathrm{old}}} S_{jk}^{(\alpha)}\, \mathbf{b}_{j,\mathrm{old}}^{(\alpha)} \quad (1 \le k \le r_\alpha^{\mathrm{new}}). \tag{11.30b}$$

In the standard case, $\mathbf{B}_\alpha^{\mathrm{old}}$ and $\mathbf{B}_\alpha^{\mathrm{new}}$ are bases. Then $r_\alpha^{\mathrm{old}} = r_\alpha^{\mathrm{new}} = \dim(\mathbf{U}_\alpha)$ holds, and $T^{(\alpha)}$ and $S^{(\alpha)}$ are uniquely defined satisfying

$$S^{(\alpha)} := (T^{(\alpha)})^{-1}. \tag{11.30c}$$

In the case of frames, the representation ranks $r_\alpha^{\mathrm{old}}, r_\alpha^{\mathrm{new}} \geq \dim(\mathbf{U}_\alpha)$ may be different so that $T^{(\alpha)}$ and $S^{(\alpha)}$ are rectangular matrices. If $r_\alpha^{\mathrm{new}} > \dim(\mathbf{U}_\alpha)$ $[r_\alpha^{\mathrm{old}} > \dim(\mathbf{U}_\alpha)]$, the matrix $T^{(\alpha)}$ $[S^{(\alpha)}]$ satisfying (11.30a [b]) is not unique.

There may be reasons to change the subspace. In Case B we consider $\mathbf{U}_\alpha^{\mathrm{new}} \subset \mathbf{U}_\alpha^{\mathrm{old}}$, and in Case C the opposite inclusion $\mathbf{U}_\alpha^{\mathrm{new}} \supset \mathbf{U}_\alpha^{\mathrm{old}}$.

Case B. Assume $\mathbf{U}_\alpha^{\mathrm{new}} \subsetneqq \mathbf{U}_\alpha^{\mathrm{old}}$. This is a typical step, when we truncate the tensor representation. Note that a transformation matrix $S^{(\alpha)}$ satisfying (11.30b) exists, whereas there is no $T^{(\alpha)}$ satisfying (11.30a).

Case C. Assume $\mathbf{U}_\alpha^{\mathrm{new}} \supsetneqq \mathbf{U}_\alpha^{\mathrm{old}}$. This happens, when we enrich $\mathbf{U}_\alpha^{\mathrm{old}}$ by further vectors. Then a transformation matrix $T^{(\alpha)}$ satisfying (11.30a) exists, but no $S^{(\alpha)}$ with (11.30b).

In Cases A and B the transformation matrix $T^{(\alpha)}$ exists. Then (11.30b) proves the following result.

Lemma 11.23. *If (11.30b) holds for $\alpha \in T_D \backslash \mathcal{L}(T_D)$, the new basis vectors $\mathbf{b}_{k,\mathrm{new}}^{(\alpha)}$ have coefficient matrices $C_{\mathrm{new}}^{(\alpha,k)}$ defined by*

$$\mathbf{C}_\alpha^{\mathrm{new}} = \mathbf{C}_\alpha^{\mathrm{old}} S^{(\alpha)}, \quad \textit{i.e., } C_{\mathrm{new}}^{(\alpha,k)} = \sum_{j=1}^{r_\alpha^{\mathrm{old}}} S_{jk}^{(\alpha)} C_{\mathrm{old}}^{(\alpha,j)} \quad \left(1 \leq k \leq r_\alpha^{\mathrm{new}}\right). \tag{11.31}$$

The arithmetical cost of (11.31) is $2 r_\alpha^{\mathrm{new}} r_\alpha^{\mathrm{old}} r_{\alpha_1} r_{\alpha_2}$ (α_1, α_2 sons of α).

Next, we consider the influence of a transformation upon the coefficient matrices of the father. Here, we rename the father vertex by $\alpha \in T_D \backslash \mathcal{L}(T_D)$ and assume that the bases $\mathbf{B}_{\alpha_1}^{\mathrm{old}}$ and $\mathbf{B}_{\alpha_2}^{\mathrm{old}}$ for at least one of the sons α_1, α_2 of α are changed into $\mathbf{B}_{\alpha_1}^{\mathrm{new}}$ and $\mathbf{B}_{\alpha_2}^{\mathrm{new}}$. If only one basis is changed, set $S^{(\alpha_i)} = T^{(\alpha_i)} = I$ for the other son. Since the transformation matrix $T^{(\alpha_i)}$ is used, the following lemma applies to Cases A and C.

Lemma 11.24. *Let α_1, α_2 be the sons of $\alpha \in T_D \backslash \mathcal{L}(T_D)$. Basis transformations (11.30a) at the son vertices α_1, α_2, i.e., $\mathbf{B}_{\alpha_i}^{\mathrm{new}} T^{(\alpha_i)} = \mathbf{B}_{\alpha_i}^{\mathrm{old}}$ ($i = 1, 2$), lead to a transformation of the coefficients at vertex α by*

$$C_{\mathrm{old}}^{(\alpha,\ell)} \mapsto C_{\mathrm{new}}^{(\alpha,\ell)} = T^{(\alpha_1)}\, C_{\mathrm{old}}^{(\alpha,\ell)}\, (T^{(\alpha_2)})^{\mathsf{T}} \qquad \textit{for } 1 \leq \ell \leq r_\alpha. \tag{11.32}$$

The arithmetical cost for (11.32) is $2 r_\alpha r_{\alpha_1}^{\mathrm{old}} r_{\alpha_2}^{\mathrm{old}} \left(r_{\alpha_1}^{\mathrm{new}} + r_{\alpha_2}^{\mathrm{new}}\right)$. If the basis is changed only at α_1 (i.e., $T^{(\alpha_2)} = I$), the cost reduces to $2 r_\alpha r_{\alpha_1}^{\mathrm{new}} r_{\alpha_1}^{\mathrm{old}} r_{\alpha_2}^{\mathrm{old}}$.

Proof. The basis vector $\mathbf{b}_\ell^{(\alpha)}$ at vertex α has the representation

$$\mathbf{b}_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}^{\mathrm{old}}} \sum_{j=1}^{r_{\alpha_2}^{\mathrm{old}}} c_{ij,\mathrm{old}}^{(\alpha,\ell)}\, \mathbf{b}_{i,\mathrm{old}}^{(\alpha_1)} \otimes \mathbf{b}_{j,\mathrm{old}}^{(\alpha_2)} \qquad \text{for } 1 \leq \ell \leq r_\alpha$$

with respect to $\mathbf{B}_{\alpha_1}^{\mathrm{old}}$ and $\mathbf{B}_{\alpha_2}^{\mathrm{old}}$. Using $\mathbf{B}_{\alpha_1}^{\mathrm{old}} = \mathbf{B}_{\alpha_1}^{\mathrm{new}} T^{(\alpha_1)}$ and $\mathbf{B}_{\alpha_2}^{\mathrm{old}} = \mathbf{B}_{\alpha_2}^{\mathrm{new}} T^{(\alpha_2)}$ (cf. (11.30a)), we obtain

$$\begin{aligned}
\mathbf{b}_\ell^{(\alpha)} &= \sum_{i=1}^{r_{\alpha_1}^{\mathrm{old}}} \sum_{j=1}^{r_{\alpha_2}^{\mathrm{old}}} c_{ij,\mathrm{old}}^{(\alpha,\ell)} \left(\sum_{k=1}^{r_{\alpha_1}^{\mathrm{new}}} T_{ki}^{(\alpha_1)} \mathbf{b}_{k,\mathrm{new}}^{(\alpha_1)} \right) \otimes \left(\sum_{m=1}^{r_{\alpha_2}^{\mathrm{new}}} T_{mj}^{(\alpha_2)} \mathbf{b}_{m,\mathrm{new}}^{(\alpha_2)} \right) \\
&= \sum_{k=1}^{r_{\alpha_1}^{\mathrm{new}}} \sum_{m=1}^{r_{\alpha_2}^{\mathrm{new}}} \left(\sum_{i=1}^{r_{\alpha_1}^{\mathrm{old}}} \sum_{j=1}^{r_{\alpha_2}^{\mathrm{old}}} T_{ki}^{(\alpha_1)}\, c_{ij,\mathrm{old}}^{(\alpha,\ell)}\, T_{mj}^{(\alpha_2)} \right) \mathbf{b}_{k,\mathrm{new}}^{(\alpha_1)} \otimes \mathbf{b}_{m,\mathrm{new}}^{(\alpha_2)} \\
&= \sum_{k=1}^{r_{\alpha_1}^{\mathrm{new}}} \sum_{m=1}^{r_{\alpha_2}^{\mathrm{new}}} c_{km,\mathrm{new}}^{(\alpha,\ell)}\, \mathbf{b}_{k,\mathrm{new}}^{(\alpha_1)} \otimes \mathbf{b}_{m,\mathrm{new}}^{(\alpha_2)}
\end{aligned}$$

with $c_{km,\mathrm{new}}^{(\alpha,\ell)} := \sum_{i=1}^{r_{\alpha_1}^{\mathrm{old}}} \sum_{j=1}^{r_{\alpha_2}^{\mathrm{old}}} T_{ki}^{(\alpha_1)} c_{ij,\mathrm{old}}^{(\alpha,\ell)} T_{mj}^{(\alpha_2)}$. This corresponds to the matrix formulation (11.32). □

The next lemma uses the transformation matrix $S^{(\alpha_i)}$ from Cases A and B.

Lemma 11.25. *Let $\alpha \in T_D \backslash \mathcal{L}(T_D)$ be a vertex with sons $\{\alpha_1, \alpha_2\} = S(\alpha)$. Assume that the coefficient matrices $C_{\mathrm{old}}^{(\alpha,\ell)}$ admit a decomposition*

$$C_{\mathrm{old}}^{(\alpha,\ell)} = S^{(\alpha_1)}\, C_{\mathrm{new}}^{(\alpha,\ell)}\, (S^{(\alpha_2)})^{\mathsf{T}} \qquad \textit{for } 1 \le \ell \le r_\alpha. \tag{11.33a}$$

Then $C_{\mathrm{new}}^{(\alpha,\ell)}$ are the coefficient matrices with respect to the new bases

$$\mathbf{B}_{\alpha_i}^{\mathrm{new}} := \mathbf{B}_{\alpha_i}^{\mathrm{old}}\, S^{(\alpha_i)} \qquad (i = 1, 2) \tag{11.33b}$$

at the son vertices (cf. (11.30b)). Since the frame $\mathbf{B}_{\alpha_i}^{\mathrm{new}}$ is not used in computations, no arithmetical operations accrue.

Proof. We start with (11.24) and insert $C_{\mathrm{old}}^{(\alpha,\ell)} = S^{(\alpha_1)}\, C_{\mathrm{new}}^{(\alpha,\ell)}\, (S^{(\alpha_2)})^{\mathsf{T}}$. Then,

$$\begin{aligned}
\mathbf{b}_\ell^{(\alpha)} &= \sum_{i=1}^{r_{\alpha_1}^{,\mathrm{old}}} \sum_{j=1}^{r_{\alpha_2}^{,\mathrm{old}}} c_{ij,\mathrm{old}}^{(\alpha,\ell)}\, \mathbf{b}_{i,\mathrm{old}}^{(\alpha_1)} \otimes \mathbf{b}_{j,\mathrm{old}}^{(\alpha_2)} \\
&= \sum_{i=1}^{r_{\alpha_1}^{,\mathrm{old}}} \sum_{j=1}^{r_{\alpha_2}^{,\mathrm{old}}} \left(S^{(\alpha_1)}\, C_{\mathrm{new}}^{(\alpha,\ell)}\, (S^{(\alpha_2)})^{\mathsf{T}} \right)_{i,j} \mathbf{b}_{i,\mathrm{old}}^{(\alpha_1)} \otimes \mathbf{b}_{j,\mathrm{old}}^{(\alpha_2)} \\
&= \sum_{i=1}^{r_{\alpha_1}^{,\mathrm{old}}} \sum_{j=1}^{r_{\alpha_2}^{,\mathrm{old}}} \sum_{k=1}^{r_{\alpha_1}^{\mathrm{new}}} \sum_{m=1}^{r_{\alpha_2}^{\mathrm{new}}} S_{ik}^{(\alpha_1)}\, c_{km,\mathrm{new}}^{(\alpha,\ell)}\, S_{jm}^{(\alpha_2)}\, \mathbf{b}_{i,\mathrm{old}}^{(\alpha_1)} \otimes \mathbf{b}_{j,\mathrm{old}}^{(\alpha_2)} \\
&= \sum_{k=1}^{r_{\alpha_1}^{\mathrm{new}}} \sum_{m=1}^{r_{\alpha_2}^{\mathrm{new}}} c_{km,\mathrm{new}}^{(\alpha,\ell)} \left(\sum_{i=1}^{r_{\alpha_1}^{\mathrm{old}}} S_{ik}^{(\alpha_1)}\, \mathbf{b}_{i,\mathrm{old}}^{(\alpha_1)} \right) \otimes \left(\sum_{j=1}^{r_{\alpha_2}^{\mathrm{old}}} S_{jm}^{(\alpha_2)}\, \mathbf{b}_{j,\mathrm{old}}^{(\alpha_2)} \right)
\end{aligned}$$

$$= \sum_{k=1}^{r_{\alpha_1}^{\mathrm{new}}} \sum_{m=1}^{r_{\alpha_2}^{\mathrm{new}}} c_{km,\mathrm{new}}^{(\alpha,\ell)} \, \mathbf{b}_{k,\mathrm{new}}^{(\alpha_1)} \otimes \mathbf{b}_{m,\mathrm{new}}^{(\alpha_2)}$$

has the coefficients $C_{\mathrm{new}}^{(\alpha,\ell)}$ with respect to the new basis $\mathbf{B}_{\alpha_i}^{\mathrm{new}} := \mathbf{B}_{\alpha_i}^{\mathrm{old}} S^{(\alpha_i)}$. □

At the root $\alpha = D$, the tensor $\mathbf{v}$ is expressed by $\mathbf{v} = \sum_{i=1}^{r_D} c_i^{(D)} \mathbf{b}_i^{(D)} = \mathbf{B}_D c^{(D)}$ (cf. (11.25)). A change of the basis $\mathbf{B}_D$ is considered next.

Lemma 11.26. *Assume a transformation by $\mathbf{B}_D^{\mathrm{new}} T^{(D)} = \mathbf{B}_D^{\mathrm{old}}$ (cf. (11.30a)). Then the coefficient vector $c_{\mathrm{old}}^{(D)}$ must be transformed into*

$$c_{\mathrm{new}}^{(D)} := T^{(D)} \, c_{\mathrm{old}}^{(D)}. \tag{11.34}$$

The arithmetical cost is $2 r_D^{\mathrm{old}} r_D^{\mathrm{new}}$.

Proof. $\mathbf{v} = \mathbf{B}_D^{\mathrm{old}} c_{\mathrm{old}}^{(D)} = \mathbf{B}_D^{\mathrm{new}} \, T^{(D)} c_{\mathrm{old}}^{(D)} = \mathbf{B}_D^{\mathrm{new}} c_{\mathrm{new}}^{(D)}$. □

11.3.1.5 Multiplication by Kronecker Products

In §11.3.1.4, the bases and consequently also some coefficient matrices have been changed, but the tensor $\mathbf{v}$ is fixed. Now, we map $\mathbf{v}$ into another tensor $\mathbf{w} = \mathbf{A}\mathbf{v}$, where $\mathbf{A} = \bigotimes_{j \in D} A_j$, but the coefficient matrices stay constant.

Proposition 11.27. *Let the tensor $\mathbf{v} = \rho_{\mathrm{HTR}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big)$ and the elementary Kronecker product $\mathbf{A} = \bigotimes_{j \in D} A^{(j)}$ be given. Then $\mathbf{w} := \mathbf{A}\mathbf{v}$ has the representation $\mathbf{w} = \rho_{\mathrm{HTR}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j^w)_{j \in D}\big)$, where only the frames (bases) $B_j = [b_1^{(j)}, \ldots, b_{r_j}^{(j)}]$ are replaced by the new frames $B_j^w = [A^{(j)} b_1^{(j)}, \ldots, A^{(j)} b_{r_j}^{(j)}]$.*

Proof. Consider $\alpha \in T_D \backslash \mathcal{L}(T_D)$ with sons α_1, α_2. Application of $\mathbf{A}^{(\alpha)} = \mathbf{A}^{(\alpha_1)} \otimes \mathbf{A}^{(\alpha_2)}$ to $\mathbf{b}_\ell^{(\alpha)}$ from (11.24) yields

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

Although the quantities $\mathbf{A}^{(\alpha)} \mathbf{b}_\ell^{(\alpha)}$, $\mathbf{A}^{(\alpha_1)} \mathbf{b}_i^{(\alpha_1)}$, $\mathbf{A}^{(\alpha_2)} \mathbf{b}_i^{(\alpha_2)}$ are new, the coefficient matrix $C^{(\alpha,\ell)}$ is unchanged. □

11.3.1.6 Gram Matrices of Bases

The Gram matrix $G(\mathbf{B}_\alpha) = \mathbf{B}_\alpha^{\mathsf{H}} \mathbf{B}_\alpha \in \mathbb{K}^{r_\alpha \times r_\alpha}$ will frequently appear later on. Its entries are

$$G(\mathbf{B}_\alpha) = (g_{ij}^{(\alpha)}) \quad \text{with } g_{ij}^{(\alpha)} = \left\langle \mathbf{b}_j^{(\alpha)}, \mathbf{b}_i^{(\alpha)} \right\rangle. \tag{11.35}$$

The recursive structure (11.24) allows a recursive definition of the Gram matrices.

Lemma 11.28. *For $\alpha \in T_D \backslash \mathcal{L}(T_D)$ let $C^{(\alpha,\bullet)}$ be the coefficient matrices from (11.24). Then $G(\mathbf{B}_\alpha)$ can be derived from $G(\mathbf{B}_{\alpha_1}), G(\mathbf{B}_{\alpha_2})$ (α_1, α_2 sons of α) by*

$$\begin{aligned} g_{\ell k}^{(\alpha)} &= \operatorname{trace}\left(C^{(\alpha,k)} G(\mathbf{B}_{\alpha_2})^\mathsf{T} (C^{(\alpha,\ell)})^\mathsf{H} G(\mathbf{B}_{\alpha_1})\right) \qquad (1 \le \ell, k \le r_\alpha) \\ &= \left\langle C^{(\alpha,k)} G(\mathbf{B}_{\alpha_2})^\mathsf{T}, G(\mathbf{B}_{\alpha_1}) C^{(\alpha,\ell)} \right\rangle_\mathsf{F} \\ &= \left\langle G(\mathbf{B}_{\alpha_1})^{\frac{1}{2}} C^{(\alpha,k)} G(\mathbf{B}_{\alpha_2})^{\frac{1}{2}\mathsf{T}}, G(\mathbf{B}_{\alpha_1})^{\frac{1}{2}} C^{(\alpha,\ell)} G(\mathbf{B}_{\alpha_2})^{\frac{1}{2}\mathsf{T}} \right\rangle_\mathsf{F}. \end{aligned}$$

Proof. $g_{\ell k}^{(\alpha)} = \langle \mathbf{b}_k^{(\alpha)}, \mathbf{b}_\ell^{(\alpha)} \rangle = \langle \sum_{ij} c_{ij}^{(\alpha,k)} \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}, \sum_{pq} c_{pq}^{(\alpha,\ell)} \mathbf{b}_p^{(\alpha_1)} \otimes \mathbf{b}_q^{(\alpha_2)} \rangle = \sum_{ijpq} c_{ij}^{(\alpha,k)} \langle \mathbf{b}_j^{(\alpha_2)}, \mathbf{b}_q^{(\alpha_2)} \rangle \overline{c_{pq}^{(\alpha,\ell)}} \langle \mathbf{b}_i^{(\alpha_1)}, \mathbf{b}_p^{(\alpha_1)} \rangle$. Use (2.10). □

11.3.1.7 Ordering of the Directions

The construction of the tree T_D groups the directions $1, 2, \ldots, d$ in a certain way. Different trees T_D lead to different nodes $\alpha \subset D$ and therefore also to different dimensions r_α. Theoretically, one would prefer a tree T_D such that

$$N_{\mathrm{mem}}^{\mathrm{HTR}}((\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}) = \sum_{\alpha \in T_D \backslash \mathcal{L}(T_D)} r_\alpha r_{\alpha_1} r_{\alpha_2} \qquad (\alpha_1, \alpha_2 \text{ sons of } \alpha)$$

is minimal. However, this minimisation is hard to perform, since usually the ranks $r_\alpha = \operatorname{rank}_\alpha(\mathbf{v})$ are not known in advance.

On the other hand, given a tree T_D, we can identify all permutations π of $D = \{1, 2, \ldots, d\}$ such that the tensor $\mathbf{v}_{i_{\pi(1)} \cdots i_{\pi(d)}}$ with interchanged directions is organised by almost the same tree.

Lemma 11.29. *Any node $\alpha \in T_D \backslash \mathcal{L}(T_D)$ gives rise to a permutation $\pi_\alpha : D \to D$ by interchanging the positions of the sons α_1 and α_2. Set*

$$P := \left\{ \prod_{\alpha \in A} \pi_\alpha : A \subset T_D \backslash \mathcal{L}(T_D) \right\}.$$

Any permutation $\pi \in P$ lets the tree T_D invariant, only the ordering of the sons α_1 and α_2 may be reversed. According to Remark 11.20, the coefficient matrix $C^{(\alpha,\ell)}$ becomes $C^{(\alpha,\ell)\mathsf{T}}$, if the factor π_α appears in $\pi \in P$. Hence, all tensors $\mathbf{v}_{i_{\pi(1)} \cdots i_{\pi(d)}}$ for $\pi \in P$ have almost the same representation. Since $\#\left(T_D \backslash \mathcal{L}(T_D)\right) = d - 1$, there are 2^{d-1} permutations in P.

Remark 11.30. A particular permutation is the reversion

$$\pi : (1, 2, \ldots, d) \mapsto (d, d-1, \ldots, 1).$$

Since $\pi = \prod_{\alpha \in T_D \backslash \mathcal{L}(T_D)} \pi_\alpha$, this permutation is contained in the set P from above. Hence, the tensor representation of $\mathbf{w} \in V_d \otimes \ldots \otimes V_1$ defined by $\mathbf{w} = \pi(\mathbf{v})$ (cf. (3.44)) is obtained from the representation (11.28) of $\mathbf{v}$ by transposing all coefficient matrices $C^{(\alpha,\ell)}$.

11.3.1.8 Interpretation of $\mathbf{C}_\alpha$ as Tensor of Order Three

Above, $\mathbf{C}_\alpha$ from (11.27) is seen as a tuple of matrices. According to Lemma 3.26, such a tuple is equivalent to a tensor of order $d = 3$. In this case, the indices are rearranged:

$$\mathbf{C}_\alpha \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2} \times r_\alpha} \quad \text{with entries } C^{(\alpha)}_{ijk} := C^{(\alpha,k)}_{ij}.$$

The transformation of the coefficients looks different in the new notation:

$$\begin{aligned}
&\text{Eq. (11.31):} && \mathbf{C}^{\text{new}}_\alpha = \left(I \otimes I \otimes (S^{(\alpha)})^{\mathsf{T}}\right) \mathbf{C}^{\text{old}}_\alpha, \\
&\text{Eq. (11.32):} && \mathbf{C}^{\text{old}}_\alpha \mapsto \mathbf{C}^{\text{new}}_\alpha = \left(T^{(\alpha_1)} \otimes T^{(\alpha_2)} \otimes I\right) \mathbf{C}^{\text{old}}_\alpha, \\
&\text{Eq. (11.33a):} && \mathbf{C}^{\text{old}}_\alpha = \left(S^{(\alpha_1)} \otimes S^{(\alpha_2)} \otimes I\right) \mathbf{C}^{\text{new}}_\alpha.
\end{aligned}$$

11.3.2 Orthonormal Bases

The choice of orthonormal bases has many advantages, numerical stability is one reason, the later truncation procedure from §11.4.2.1 another one.

11.3.2.1 Bases at Leaf Vertices

Assume that the spaces U_j involved in $\mathbf{U} = \bigotimes_{j \in D} U_j$ possess scalar products, which are denoted by $\langle \cdot, \cdot \rangle$ (the reference to the index j is omitted). Also the induced scalar product on the tensor spaces $\mathbf{U}_\alpha = \bigotimes_{j \in \alpha} U_j$ for $\alpha \in T_D$ is written as $\langle \cdot, \cdot \rangle$.

Even if $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j \in D} V_j$ is a Banach tensor space with a non-Hilbert norm, one can introduce another (equivalent) norm on the finite dimensional subspace $\mathbf{U}$, in particular, one may define a scalar product (see also Exercise 2.16c). A Hilbert structure in $\mathbf{V}$ outside of $\mathbf{U} = \bigotimes_{k \in D} U_k$ is not needed.

If the given bases $B_j = [b^{(j)}_1, \ldots, b^{(j)}_{r_j}]$ of U_j are not orthonormal, one can apply the techniques discussed in §8.2.3.2 (see also §13.4.4). For instance, one may use the Gram matrix $G(B_j)$ from (11.35) (cf. Lemma 8.12b). After obtaining the new orthonormal basis B^{new}_j, we replace the old one by B^{new}_j and rename it B_j. If $\{j\} \in T_D$ has a father $\alpha \in T_D$ for which already a basis of $\mathbf{U}_\alpha$ is defined, the corresponding coefficients have to be transformed according to Lemma 11.24.

11.3.2.2 Bases at Non-Leaf Vertices

Now, we are considering a vertex $\alpha \in T_D \backslash \mathcal{L}(T_D)$ and assume that orthonormal bases $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$ at the son vertices α_1, α_2 are already determined. According to Remark 11.19c, the tensor space $\mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ at vertex α has the canonical *orthonormal* basis

$$\{\mathbf{b}^{(\alpha)}_{\nu\mu} := \mathbf{b}^{(\alpha_1)}_\nu \otimes \mathbf{b}^{(\alpha_2)}_\mu : \ 1 \le i \le r_{\alpha_1}, \ 1 \le j \le r_{\alpha_2}\}. \tag{11.36}$$

Assume that a subspace $\mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ is defined as the span of some basis $\{\mathbf{b}_i^{(\alpha)} : 1 \le i \le r_\alpha\}$, which gives rise to $\mathbf{B}_\alpha := \left[\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}\right] \in (\mathbf{U}_\alpha)^{r_\alpha}$. In case the basis $\mathbf{B}_\alpha$ is not already orthonormal, we may follow Lemma 8.12b and determine the Gram matrix $G(\mathbf{B}_\alpha)$ (cf. Lemma 11.28).

Lemma 11.31. *Let α_1 and α_2 be the sons of $\alpha \in T_D$. Suppose that $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$ represent orthonormal*[13] *bases. For any vectors $\mathbf{v}, \mathbf{w} \in \mathbf{U}_\alpha$ with representations*

$$\mathbf{v} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{\mathbf{v}}\, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} \quad \textit{and} \quad \mathbf{w} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{\mathbf{w}}\, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}$$

involving coefficient matrices $C_\mathbf{v} := \left(c_{ij}^{\mathbf{v}}\right), C_\mathbf{w} := \left(c_{ij}^{\mathbf{w}}\right) \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2}}$, the scalar product of $\mathbf{v}$ and $\mathbf{w}$ equals the Frobenius scalar product (2.10) of $C_\mathbf{v}$ and $C_\mathbf{w}$:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle C_\mathbf{v}, C_\mathbf{w} \rangle_\mathsf{F} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{\mathbf{v}}\, \overline{c_{ij}^{\mathbf{w}}}\,. \tag{11.37a}$$

Hence, the isomorphism Θ_α from (11.23c) is unitary. In particular, the coefficient matrices of $\mathbf{b}_\nu^{(\alpha)}$ and $\mathbf{b}_\mu^{(\alpha)}$ yield

$$g_{\nu\mu}^{(\alpha)} = \left\langle \mathbf{b}_\mu^{(\alpha)}, \mathbf{b}_\nu^{(\alpha)} \right\rangle = \left\langle C^{(\alpha,\mu)}, C^{(\alpha,\nu)} \right\rangle_\mathsf{F} \quad (1 \le \nu, \mu \le r_\alpha). \tag{11.37b}$$

The basis $\{\mathbf{b}_\nu^{(\alpha)}\}$ is orthonormal if and only if $\{C^{(\alpha,\nu)}\}$ is orthonormal with respect to the Frobenius scalar product. The computational cost for all entries $g_{\nu\mu}^{(\alpha)}$ $(1 \le \nu, \mu \le r_\alpha)$ is $2r_\alpha^2 r_{\alpha_1} r_{\alpha_2}$.

Proof. By

$$\begin{aligned}\langle \mathbf{v}, \mathbf{w} \rangle &= \left\langle \sum_{i,j} c_{ij}^{\mathbf{v}}\, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}, \sum_{k,\ell} c_{k\ell}^{\mathbf{w}}\, \mathbf{b}_k^{(\alpha_1)} \otimes \mathbf{b}_\ell^{(\alpha_2)} \right\rangle \\ &= \sum_{k,\ell} \sum_{i,j} c_{ij}^{\mathbf{v}}\, \overline{c_{k\ell}^{\mathbf{w}}}\, \underbrace{\left\langle \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}, \mathbf{b}_k^{(\alpha_1)} \otimes \mathbf{b}_\ell^{(\alpha_2)} \right\rangle}_{=\delta_{i,k}\delta_{j,\ell}} = \sum_{i,j} c_{ij}^{\mathbf{v}}\, \overline{c_{k\ell}^{\mathbf{w}}}\,,\end{aligned}$$

we arrive at the Frobenius scalar product. (11.37b) follows also from Lemma 11.28 because of $G(\mathbf{B}_{\alpha_i}) = I$. □

We denote the hierarchical format with orthonormal bases by

$$\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{orth}}\left(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \setminus \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\right), \tag{11.38}$$

which requires $B_j^\mathsf{H} B_j = I$ for all $j \in D$ and $\langle C^{(\alpha,\mu)}, C^{(\alpha,\nu)} \rangle_\mathsf{F} = \delta_{\nu\mu}$. By Lemma 11.31, these conditions imply orthonormality: $\mathbf{B}_\alpha^\mathsf{H} \mathbf{B}_\alpha = I$.

Adding the cost of the orthonormalisation of a basis and of the transformations involved, we get the following result.

Remark 11.32. Given a hierarchical representation with general bases or frames, the orthonormalisation costs asymptotically $2dnr^2 + 4r^4(d-1)$ operations ($r := \max_\alpha r_\alpha$, $n := \max_j n_j$, details in (13.16b)).

[13] If the bases are not orthonormal, compare Lemma 11.44 and its proof.

11.3.2.3 Transformation between Orthonormal Bases

Lemmata 11.24 and 11.25 remain valid for orthonormal bases. In order not to lose orthonormality, the transformation matrices must be unitary or (in the case of rectangular matrices) orthogonal. Note that orthogonal $n \times m$ matrices require $n \geq m$. The situation $r_{\alpha_i}^{\text{new}} \leq r_{\alpha_i}^{\text{old}}$ is covered by Part (a) of the next corollary, while Part (b) requires $r_{\alpha_i}^{\text{new}} \geq r_{\alpha_i}^{\text{old}}$.

Corollary 11.33. Let α_1, α_2 be the sons of $\alpha \in T_D \backslash \mathcal{L}(T_D)$. (a) If the transformations

$$\begin{array}{l} C_{\text{old}}^{(\alpha,\ell)} = S^{(\alpha_1)}\, C_{\text{new}}^{(\alpha,\ell)}\, (S^{(\alpha_2)})^{\mathsf{T}} \quad \text{for } 1 \leq \ell \leq r_\alpha, \\ \mathbf{B}_{\alpha_1}^{\text{new}} = \mathbf{B}_{\alpha_1}^{\text{old}}\, S^{(\alpha_1)}, \quad \mathbf{B}_{\alpha_2}^{\text{new}} = \mathbf{B}_{\alpha_2}^{\text{old}}\, S^{(\alpha_2)} \end{array} \tag{11.39a}$$

hold with orthogonal matrices $S^{(\alpha_i)}$ $(i = 1, 2)$, the bases $\mathbf{B}_{\alpha_i}^{\text{new}}$ inherit orthonormality from $\mathbf{B}_{\alpha_i}^{\text{old}}$, while the Frobenius scalar product of the coefficient matrices is invariant:

$$\left\langle C_{\text{new}}^{(\alpha,\ell)}, C_{\text{new}}^{(\alpha,k)} \right\rangle_{\mathsf{F}} = \left\langle C_{\text{old}}^{(\alpha,\ell)}, C_{\text{old}}^{(\alpha,k)} \right\rangle_{\mathsf{F}} \qquad \text{for all } 1 \leq \ell, k \leq r_\alpha. \tag{11.39b}$$

(b) If $B_{\alpha_i}^{\text{new}} T^{(\alpha_i)} = B_{\alpha_i}^{\text{old}}$ holds for $i=1,2$ with orthogonal matrices $T^{(\alpha_i)}$, the new coefficients defined by $C_{\text{new}}^{(\alpha,\ell)} = T^{(\alpha_1)} C_{\text{old}}^{(\alpha,\ell)} (T^{(\alpha_2)})^{\mathsf{T}}$ satisfy again (11.39b).

Proof. $\mathbf{B}_{\alpha_1}^{\text{new}\mathsf{H}} \mathbf{B}_{\alpha_1}^{\text{new}} = S^{(\alpha_1)\mathsf{H}} \mathbf{B}_{\alpha_1}^{\text{old}\mathsf{H}} \mathbf{B}_{\alpha_1}^{\text{old}}\, S^{(\alpha_1)} = S^{(\alpha_1)\mathsf{H}}\, S^{(\alpha_1)} = I$ proves orthonormality of the basis $B_{\alpha_1}^{\text{new}}$.

The identity $\langle C_{\text{old}}^{(\alpha,\ell)}, C_{\text{old}}^{(\alpha,k)} \rangle_{\mathsf{F}} = \langle S^{(\alpha_1)} C_{\text{new}}^{(\alpha,\ell)}\, S^{(\alpha_2)\mathsf{T}}, S^{(\alpha_1)} C_{\text{new}}^{(\alpha,k)} S^{(\alpha_2)\mathsf{T}} \rangle_{\mathsf{F}} = \langle C_{\text{new}}^{(\alpha,\ell)}, C_{\text{new}}^{(\alpha,k)} \rangle_{\mathsf{F}}$ follows from Exercise 2.11b. □

As in §11.3.2.1, a transformation of $\mathbf{B}_\alpha$ should be followed by an update of $\mathbf{C}_\beta$ for the father β of α (cf. Lemma 11.24). If D is the father of α, the coefficient c^D must be updated (cf. Lemma 11.26).

11.3.2.4 Unitary Mappings

Here, we consider the analogue of the mappings from §11.3.1.5 under orthonormality preserving conditions. For $j \in D$, let $B_j = [b_1^{(j)}, \ldots, b_{r_j}^{(j)}]$ be an orthonormal basis. $U_j := \operatorname{range}(B_j)$ is a subspace of V_j. Let $A_j : U_j \to \hat{U}_j \subset V_j$ be a mapping such that $\hat{B}_j = [\hat{b}_1^{(j)}, \ldots, \hat{b}_{r_j}^{(j)}]$ with $\hat{b}_i^{(j)} := A_j b_i^{(j)}$ is again an orthonormal basis. Proposition 11.27 applied to $\mathbf{A} = \bigotimes_{j \in D} A_j$ takes the following form.

Proposition 11.34. *Let the tensor* $\mathbf{v} = \rho_{\text{HTR}}^{\text{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big)$ *and the elementary Kronecker product* $\mathbf{A} = \bigotimes_{j \in D} A_j$ *with unitary mappings* $A_j : U_j \to \hat{U}_j \subset V_j$ *be given. Then* $\mathbf{w} := \mathbf{A}\mathbf{v}$ *has the representation* $\mathbf{w} = \rho_{\text{HTR}}^{\text{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (\hat{B}_j)_{j \in D}\big)$*, where only the orthonormal bases* B_j *are replaced by the orthonormal bases* $\hat{B}_j = [A_j b_1^{(j)}, \ldots, A_j b_{r_j}^{(j)}]$.

Let $\mathfrak{A} \subset T_D$ be a complete set of successors of D (cf. Definition 11.9). Consider Kronecker products $\mathbf{A} = \bigotimes_{\alpha\in\mathfrak{A}} A_\alpha$ and assume that

$$A_\alpha : \mathbf{U}_\alpha \to \hat{\mathbf{U}}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2} \text{ is unitary for all } \alpha \in \mathfrak{A}.$$

Hence, the orthonormal basis $\mathbf{B}_\alpha = [\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}]$ is mapped into a new orthonormal basis $\hat{\mathbf{B}}_\alpha = [\hat{\mathbf{b}}_1^{(\alpha)}, \ldots, \hat{\mathbf{b}}_{r_\alpha}^{(\alpha)}]$ with $\hat{\mathbf{b}}_\ell^{(\alpha)} := A_\alpha \mathbf{b}_\ell^{(\alpha)}$. To represent $\hat{\mathbf{b}}_\ell^{(\alpha)}$, new coefficient matrices $\hat{C}^{(\alpha,\ell)}$ are to be defined with the property

$$A_\alpha \mathbf{b}_\ell^{(\alpha)} = \sum\nolimits_{ij} \hat{c}_{ij}^{(\alpha,\ell)} \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}.$$

The result $\mathbf{Av}$ has the representation $\rho_{\mathrm{HTR}}^{\mathrm{orth}}\big(T_D, (\hat{\mathbf{C}}_\beta)_{\beta\in T_D\setminus\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\big)$, where $\hat{\mathbf{C}}_\beta = \mathbf{C}_\beta$ for all $\beta \notin \mathfrak{A}$. Only for $\beta \in \mathfrak{A}$, new coefficient matrices appear as defined above.

11.3.3 HOSVD Bases

11.3.3.1 Definitions, Computation of $\mathcal{M}_\alpha(\mathbf{v})\mathcal{M}_\alpha(\mathbf{v})^{\mathsf{H}}$

A by-product of the representation (11.16b) in Theorem 11.12 is stated below.

Remark 11.35. The left singular vectors $\mathbf{u}_i^{(\alpha)}$ of $\mathcal{M}_\alpha(\mathbf{v})$ (cf. (11.16b)) may be chosen as orthonormal basis: $\mathbf{B}_\alpha = [\mathbf{u}_1^{(\alpha)} \cdots \mathbf{u}_{r_\alpha}^{(\alpha)}]$. They form the HOSVD basis corresponding to the tensor $\mathbf{v} \in \mathbf{V}$ and to the vertex $\alpha \in T_D$ (cf. Definition 8.22).

Definition 11.36 (hierarchical HOSVD representation). The hierarchical HOSVD representation denoted by

$$\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{HOSVD}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\setminus\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\big)$$

indicates that these data correspond to HOSVD bases $\mathbf{B}_\alpha$ for all $\alpha \in T_D$.

Remark 11.35 states the existence of a basis $\mathbf{B}_\alpha^{\mathrm{HOSVD}} = \big[\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}\big]$, but for the practical implementation one needs the corresponding coefficient matrix family $\mathbf{C}_\alpha^{\mathrm{HOSVD}} = (C_{\mathrm{HOSVD}}^{(\alpha,\ell)})_{1\le\ell\le r_\alpha}$. In the following, we describe a simple realisation of its computation.

The left singular value decomposition of $\mathcal{M}_\alpha(\mathbf{v})$ is equivalent to the diagonalisation of $\mathcal{M}_\alpha(\mathbf{v})\mathcal{M}_\alpha(\mathbf{v})^{\mathsf{H}}$. We recall that $\mathcal{M}_\alpha(\mathbf{v})\mathcal{M}_\alpha(\mathbf{v})^{\mathsf{H}}$ for a certain vertex $\alpha \in T_D$ is the matrix version of the partial scalar product $\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha^c} \in \mathbf{V}_\alpha \otimes \mathbf{V}_\alpha$ (cf. §5.2.3). In the case of the tensor subspace format, its computation has to refer to the complete coefficient tensor. Similarly, for the r-term format, the computation of M_α involves all coefficients. For the orthonormal hierarchical format, the situation is simpler. Only the coefficients $\mathbf{C}_\beta$ for all predecessors $\beta \supset \alpha$ are involved.

Theorem 11.37. *For* $\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\big)$ *define the matrices* $E_\alpha = \big(e_{ij}^{(\alpha)}\big) \in \mathbb{K}^{r_\alpha\times r_\alpha}$ *by*

$$\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha^c} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)}\, \mathbf{b}_i^{(\alpha)} \otimes \overline{\mathbf{b}_j^{(\alpha)}} \in \mathbf{V}_\alpha \otimes \mathbf{V}_\alpha. \tag{11.40a}$$

For $\alpha = D$*, the matrix* E_D *(usually of size* 1×1*) equals*

$$E_D := c^{(D)}(c^{(D)})^{\mathsf{H}} \in \mathbb{K}^{r_D\times r_D}. \tag{11.40b}$$

Let α_1, α_2 *be the sons of* $\alpha \in T_D\backslash\mathcal{L}(T_D)$*. Given* $E_\alpha \in \mathbb{K}^{r_\alpha\times r_\alpha}$*, one determines* E_{α_1} *and* E_{α_2} *from*

$$E_{\alpha_1} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)} C^{(\alpha,i)}(C^{(\alpha,j)})^{\mathsf{H}}, \quad E_{\alpha_2} = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\alpha)} (C^{(\alpha,i)})^{\mathsf{T}}\, \overline{C^{(\alpha,j)}}. \tag{11.40c}$$

Proof. Use Theorem 5.14 and note that the Gram matrix is the identity, since the bases are orthonormal. □

Even for non-orthonormal bases, Theorem 5.14 provides a recursion for E_α.

Theorem 11.37 allows us to determine E_α by a recursion from the root to the leaves. However, it helps also for the computation of the HOSVD bases. A HOSVD basis at vertex α is characterised by

$$E_\alpha = \operatorname{diag}\big\{\big(\sigma_1^{(\alpha)}\big)^2, \ldots, \big(\sigma_{r_\alpha}^{(\alpha)}\big)^2\big\}, \tag{11.41}$$

corresponding to the diagonalisation $\langle \mathcal{M}_\alpha(\mathbf{v}), \mathcal{M}_\alpha(\mathbf{v})\rangle_{\alpha^c} = \sum_{i=1}^{r_\alpha} \big(\sigma_i^{(\alpha)}\big)^2 \mathbf{b}_i^{(\alpha)} \otimes \overline{\mathbf{b}_i^{(\alpha)}}$.

Theorem 11.38. *Given the data* $\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\big)$*, assume that (11.41) holds at the vertex* $\alpha \in T_D\backslash\mathcal{L}(T_D)$*. Then*

$$E_{\alpha_1} = \sum_{i=1}^{r_\alpha} \big(\sigma_i^{(\alpha)}\big)^2 C^{(\alpha,i)}(C^{(\alpha,i)})^{\mathsf{H}}, \; E_{\alpha_2} = \sum_{i=1}^{r_\alpha} \big(\sigma_1^{(\alpha)}\big)^2 (C^{(\alpha,i)})^{\mathsf{T}}\, \overline{C^{(\alpha,i)}} \tag{11.42}$$

holds at the son vertices α_1 *and* α_2*. The diagonalisations* $E_{\alpha_1} = U_{\alpha_1}\Sigma_{\alpha_1}^2 U_{\alpha_1}^{\mathsf{H}}$ *and* $E_{\alpha_2} = U_{\alpha_2}\Sigma_{\alpha_2}^2 U_{\alpha_2}^{\mathsf{H}}$ *yield the HOSVD bases* $\mathbf{B}_{\alpha_k}^{\mathrm{HOSVD}} = \mathbf{B}_{\alpha_k} U_{\alpha_k}$ *for* $k = 1, 2$*, where* $U_{\alpha_k} = [u_1^{(\alpha_k)}, \ldots, u_{r_{\alpha_k}^{HOSVD}}^{(\alpha_k)}]$.

Proof. $E_{\alpha_1} = U_{\alpha_1}\Sigma_{\alpha_1}^2 U_{\alpha_1}^{\mathsf{H}}$ can be rewritten as $E_{\alpha_1} = \sum_{\nu=1}^{r_{\alpha_1}^{\mathrm{HOSVD}}} (\sigma_\nu^{(\alpha_1)})^2 u_\nu^{(\alpha_1)} u_\nu^{(\alpha_1)\mathsf{H}}$ and $e_{ij}^{(\alpha_1)} = \sum_{\nu=1}^{r_{\alpha_1}^{\mathrm{HOSVD}}} (\sigma_\nu^{(\alpha_1)})^2 u_{\nu,i}^{(\alpha_1)} u_{\nu,j}^{(\alpha_1)\mathsf{H}}$. Hence,

$$\langle \mathcal{M}_{\alpha_1}(\mathbf{v}), \mathcal{M}_{\alpha_1}(\mathbf{v})\rangle_{\alpha_1^c} = \sum_{i,j=1}^{r_{\alpha_1}} \sum_{\nu=1}^{r_{\alpha_1}^{\mathrm{HOSVD}}} (\sigma_\nu^{(\alpha_1)})^2 u_{\nu,i}^{(\alpha_1)} u_{\nu,j}^{(\alpha_1)\mathsf{H}}\, \mathbf{b}_i^{(\alpha_1)} \otimes \overline{\mathbf{b}_j^{(\alpha_1)}} =$$

$$= \sum_{\nu=1}^{r_{\alpha_1}^{\text{HOSVD}}} (\sigma_\nu^{(\alpha_1)})^2 \, \mathbf{b}_{\nu,\text{HOSVD}}^{(\alpha_1)} \otimes \overline{\mathbf{b}_{\nu,\text{HOSVD}}^{(\alpha_1)}}$$

with $\mathbf{b}_{\nu,\text{HOSVD}}^{(\alpha_1)} = \sum_{i=1}^{r_\alpha} u_\nu^{(\alpha_1)}[i] \, \mathbf{b}_i^{(\alpha_1)}$, i.e., $\mathbf{B}_{\alpha_1}^{\text{HOSVD}} = \mathbf{B}_{\alpha_1} U_{\alpha_1}$. Similarly for α_2. □

In the following, we shall determine the HOSVD bases $\mathbf{B}_\alpha^{\text{HOSVD}}$ together with the singular values $\sigma_i^{(\alpha)} > 0$ which can be interpreted as weights of $b_{i,\text{HOSVD}}^{(\alpha)}$. We add $\Sigma_\alpha = \operatorname{diag}\{\sigma_1^{(\alpha)}, \ldots, \sigma_{r_\alpha^{\text{HOSVD}}}^{(\alpha)}\}$ to the representation data.

In the following, we start from orthonormal bases $\mathbf{B}_\alpha$ and their coefficient matrices $\mathbf{C}_\alpha$ and construct the new HOSVD bases $\mathbf{B}_\alpha^{\text{HOSVD}}$, weights Σ_α, and matrices $\mathbf{C}_\alpha^{\text{HOSVD}}$. The coefficient matrices $C^{(\alpha,\ell)}$ will change twice: a transform $\mathbf{B}_\alpha \mapsto \mathbf{B}_\alpha^{\text{HOSVD}}$ creates new basis vectors and therefore also new coefficient matrices $\hat{C}^{(\alpha,\ell)}$. Since the coefficients refer to the basis vectors of the sons, a basis change in these vertices leads to the second transform $\hat{C}^{(\alpha,\ell)} \mapsto C_{\text{HOSVD}}^{(\alpha,\ell)}$ into their final state. Also the number of basis vectors in $\mathbf{B}_{\alpha_i}$ for the sons $\alpha_1, \alpha_2 \in S(\alpha)$ may change from r_{α_i} to $r_{\alpha_i}^{\text{HOSVD}}$. For simplicity, we shall overwrite the old values by the new ones without changing the symbol.

Because of $r_D = 1$, the treatment of the root (cf. §11.3.3.2) differs from the treatment of the inner vertices of T_D (cf. §11.3.3.3). The inductive steps can be combined in different ways to get (a) the complete HOSVD representation $\rho_{\text{HTR}}^{\text{HOSVD}}$, (b) the HOSVD at one vertex, and (c) the coefficients at one level $T_D^{(\ell)}$ of the tree (cf. §11.3.3.4).

11.3.3.2 Treatment of the Root

In order to apply Theorem 11.38, we assume in the following that all bases $\mathbf{B}_\alpha$ $(\alpha \in T_D)$ are orthonormal: $\mathbf{v} = \rho_{\text{HTR}}^{\text{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \setminus \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big)$.

The root $\alpha = D$ plays a special rôle, since we assume $\mathbf{U}_D = \operatorname{span}\{\mathbf{v}\}$ to be one-dimensional (otherwise replace $\mathbf{U}_D$ by $\operatorname{span}\{\mathbf{v}\}$; note that the HOSVD is associated to one tensor $\mathbf{v}$ only). Hence, there is only one basis vector $\mathbf{b}_1^{(D)}$. We may assume that $\mathbf{b}_1^{(D)} = \mathbf{v} / \|\mathbf{v}\|$. The definition of the weight

$$\sigma_1^{(D)} := \|\mathbf{v}\| \tag{11.43a}$$

coincides with (11.40b): $\Sigma_D^2 := E_D := c^{(D)}(c^{(D)})^{\mathsf{H}} \in \mathbb{K}^{1\times 1}$.

Let $C^{(D,1)} \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2}}$ be the coefficient matrix of the vector $\mathbf{b}_1^{(D)}$, where α_1, α_2 are the sons of D (cf. (11.24)). Determine the reduced (both-sided) singular value decomposition of $\sigma_1^{(D)} C^{(D,1)}$:

$$\sigma_1^{(D)} C^{(D,1)} = U \Sigma V^{\mathsf{T}} \qquad \left\{ \begin{array}{l} U \in \mathbb{K}^{r_{\alpha_1} \times r}, \ V \in \mathbb{K}^{r_{\alpha_2} \times r} \text{ orthogonal}, \\ \sigma_1 \geq \ldots \geq \sigma_r > 0, \ r := \operatorname{rank}(C^{(D,1)}), \\ \Sigma = \operatorname{diag}\{\sigma_1, \ldots, \sigma_r\} \in \mathbb{K}^{r \times r}. \end{array} \right\} \tag{11.43b}$$

The rank $r_{\alpha_k}^{\text{HOSVD}} := r$ may be smaller than r_{α_k} $(k = 1, 2)$. The bases at the son vertices α_1, α_2 are changed via[14]

$$\mathbf{B}_{\alpha_1} \mapsto \mathbf{B}_{\alpha_1}^{\text{HOSVD}} := \mathbf{B}_{\alpha_1} U \quad \text{and} \quad \mathbf{B}_{\alpha_2} \mapsto \mathbf{B}_{\alpha_2}^{\text{HOSVD}} := \mathbf{B}_{\alpha_2} V, \tag{11.43c}$$

i.e., Lemma 11.25 applies with $S^{(\alpha_1)} := U$ and $S^{(\alpha_2)} := V$, and shows that

$$C_{\text{HOSVD}}^{(D,1)} = \Sigma = \operatorname{diag}\left\{\frac{\sigma_1}{\sigma_1^{(\text{D})}}, \ldots, \frac{\sigma_r}{\sigma_1^{(\text{D})}}\right\}. \tag{11.43d}$$

The size of the bases $\mathbf{B}_{\alpha_i}^{\text{HOSVD}}$ is $r_{\alpha_i}^{\text{HOSVD}} := r = \operatorname{rank}(C^{(D,1)})$. According to Lemma 11.23, the coefficient matrices of the new basis vectors $\mathbf{b}_{\ell,\text{HOSVD}}^{(\alpha_i)}$ are

$$\mathbf{C}_{\alpha_1}^{\text{HOSVD}} := \mathbf{C}_{\alpha_1} U \qquad \text{and} \qquad \mathbf{C}_{\alpha_2}^{\text{HOSVD}} := \mathbf{C}_{\alpha_2} V, \tag{11.43e}$$

i.e., $C_{\text{HOSVD}}^{(\alpha_1,\ell)} := \sum_{k=1}^{r_{\alpha_1}} U_{k\ell}\, C^{(\alpha_1,k)}$ and $C_{\text{HOSVD}}^{(\alpha_2,\ell)} := \sum_{k=1}^{r_{\alpha_2}} V_{k\ell}\, C^{(\alpha_2,k)}$ for $1 \le \ell \le r$. To simplify the notation, we omit the suffix 'HOSVD' and write r_{α_i}, $\mathbf{B}_{\alpha_i}$, $\mathbf{C}_{\alpha_i}$ for the new quantities at α_i. The newly introduced weights are defined by the singular values of (11.43b):

$$\Sigma_{\alpha_1} := \Sigma_{\alpha_2} := \operatorname{diag}\{\sigma_1, \ldots, \sigma_r\} \text{ with } \sigma_i = \sigma_i^{(\alpha_1)} = \sigma_i^{(\alpha_2)} \text{ from (11.43b).} \tag{11.43f}$$

As mentioned above, the old dimensions r_{α_i} of the subspaces are changed into

$$r_{\alpha_1} := r_{\alpha_2} := r. \tag{11.43g}$$

Remark 11.39. The computational work of (11.43) consists of

a) $N_{\text{SVD}}(r_{\alpha_1}, r_{\alpha_2})$ for U, σ_i, V,

b) $2(r_{\alpha_1} r_{\alpha_1}^{\text{HOSVD}} r_{\alpha_{11}} r_{\alpha_{12}} + r_{\alpha_2} r_{\alpha_2}^{\text{HOSVD}} r_{\alpha_{21}} r_{\alpha_{22}})$ for (11.43e), where[15] $\alpha_{k1}, \alpha_{k2} \in S(\alpha_k)$. Bounding all r_γ $(\gamma \in T_D)$ by r, the total asymptotic work is $4r^4$.

11.3.3.3 Inner Vertices

Assume that at vertex $\alpha \in T_D \backslash (\{D\} \cup \mathcal{L}(T_D))$ a new basis $\mathbf{B}_\alpha = [\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}]$ with weight tuples $\Sigma_\alpha = \operatorname{diag}\{\sigma_1^{(\alpha)}, \ldots, \sigma_{r_\alpha}^{(\alpha)}\}$ is already determined. The corresponding coefficient matrices are gathered in $\mathbf{C}^{(\alpha)} = (C^{(\alpha,\ell)})_{1 \le \ell \le r_\alpha}$ (note that in the previous step these matrices have been changed). Form the matrices[16]

$$\mathbf{Z}_{\alpha_1} := \left[\sigma_1^{(\alpha)} C^{(\alpha,1)}, \sigma_2^{(\alpha)} C^{(\alpha,2)}, \ldots, \sigma_{r_\alpha}^{(\alpha)} C^{(\alpha,r_\alpha)}\right] \in \mathbb{K}^{r_{\alpha_1} \times \left(r_\alpha r_{\alpha_2}\right)}, \tag{11.44a}$$

$$\mathbf{Z}_{\alpha_2} := \begin{bmatrix} \sigma_1^{(\alpha)} C^{(\alpha,1)} \\ \vdots \\ \sigma_{r_\alpha}^{(\alpha)} C^{(\alpha,r_\alpha)} \end{bmatrix} \in \mathbb{K}^{\left(r_\alpha r_{\alpha_1}\right) \times r_{\alpha_2}}, \tag{11.44b}$$

[14] Note that E_{α_1} from (11.42) equals $(\sigma_1^{(D)} C^{(D,1)})(\sigma_1^{(D)} C^{(D,1)})^{\mathsf{H}} = U_{\alpha_1} \Sigma_{\alpha_1}^2 U_{\alpha_1}^{\mathsf{H}}$ with $U_{\alpha_1} = U$ (U from (11.43b)). Analogously, $U_{\alpha_2} = V$ holds.

[15] If α_1 (or α_2) is a leaf, the numbers change as detailed in Remark 11.40.

[16] $\mathbf{Z}_{\alpha_1}$ may be interpreted as $\Theta_\alpha(\mathbf{B}_\alpha \Sigma_\alpha)$ from (11.23c).

where α_1 and α_2 are the sons of $\alpha \in T_D$. Compute the left-sided reduced singular value decomposition of $\mathbf{Z}_{\alpha_1}$ and the right-sided one of $\mathbf{Z}_{\alpha_2}$:

$$\mathbf{Z}_{\alpha_1} = U\Sigma_{\alpha_1}\hat{V}^\mathsf{T} \qquad \text{and} \qquad \mathbf{Z}_{\alpha_2} = \hat{U}\Sigma_{\alpha_2}V^\mathsf{T}. \tag{11.44c}$$

The matrices $\hat{V}$ and $\hat{U}$ are not needed. Only the matrices

$$\begin{aligned} &U \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_1}^{\text{HOSVD}}}, \quad &&\Sigma_{\alpha_1} = \operatorname{diag}\{\sigma_1^{(\alpha_1)}, \ldots, \sigma_{r_{\alpha_1}^{\text{HOSVD}}}^{(\alpha_1)}\} \in \mathbb{K}^{r_{\alpha_1}^{\text{HOSVD}} \times r_{\alpha_1}^{\text{HOSVD}}}, \\ &V \in \mathbb{K}^{r_{\alpha_2} \times r_{\alpha_2}^{\text{HOSVD}}}, \quad &&\Sigma_{\alpha_2} = \operatorname{diag}\{\sigma_1^{(\alpha_2)}, \ldots, \sigma_{r_{\alpha_2}^{\text{HOSVD}}}^{(\alpha_2)}\} \in \mathbb{K}^{r_{\alpha_2}^{\text{HOSVD}} \times r_{\alpha_2}^{\text{HOSVD}}} \end{aligned} \tag{11.44d}$$

are of interest, where $r_{\alpha_i}^{\text{HOSVD}} := \operatorname{rank}(\mathbf{Z}_{\alpha_i}) < r_{\alpha_i}$ may occur. The data (11.44d) are characterised by the diagonalisations of the matrices E_{α_1} and E_{α_2} from (11.42):

$$\begin{aligned} E_{\alpha_1} &= \mathbf{Z}_{\alpha_1}\mathbf{Z}_{\alpha_1}^\mathsf{H} = \textstyle\sum_{\ell=1}^{r_\alpha}(\sigma_\ell^{(\alpha)})^2 C^{(\alpha,\ell)}C^{(\alpha,\ell)\mathsf{H}} = U\Sigma_{\alpha_1}^2U^\mathsf{H}, \\ E_{\alpha_2} &= \mathbf{Z}_{\alpha_2}^\mathsf{T}\overline{\mathbf{Z}_{\alpha_2}} = \textstyle\sum_{\ell=1}^{r_\alpha}(\sigma_\ell^{(\alpha)})^2 C^{(\alpha,\ell)\mathsf{T}}\overline{C^{(\alpha,\ell)}} = V\Sigma_{\alpha_2}^2V^\mathsf{H}. \end{aligned}$$

The inclusions $\operatorname{range}(C^{(\alpha,\ell)}) \subset \operatorname{range}(U)$ and $\operatorname{range}(C^{(\alpha,\ell)\mathsf{T}}) = \operatorname{range}(V)$ are valid by construction; hence, $C^{(\alpha,\ell)}$ allows a representation $C^{(\alpha,\ell)} = UC_{\text{HOSVD}}^{(\alpha,\ell)}V^\mathsf{T}$. Since U and V are orthogonal matrices, the coefficient matrices at vertex α are transformed by

$$C^{(\alpha,\ell)} \mapsto C_{\text{HOSVD}}^{(\alpha,\ell)} := U^\mathsf{H}C^{(\alpha,\ell)}\overline{V} \in \mathbb{K}^{r_{\alpha_1}^{\text{HOSVD}} \times r_{\alpha_2}^{\text{HOSVD}}} \qquad (1 \le \ell \le r_\alpha). \tag{11.44e}$$

According to Lemmata 11.25 and 11.23, the bases and coefficient matrices at the son vertices α_1, α_2 transform as follows:

$$\begin{aligned} \mathbf{B}_{\alpha_1} \mapsto \mathbf{B}_{\alpha_1}^{\text{HOSVD}} &:= \mathbf{B}_{\alpha_1}U \;\text{ and }\; \mathbf{B}_{\alpha_2} \mapsto \mathbf{B}_{\alpha_2}^{\text{HOSVD}} := \mathbf{B}_{\alpha_2}V, \\ \mathbf{C}_{\alpha_1}^{\text{HOSVD}} &:= \mathbf{C}_{\alpha_1}U \;\text{ and }\; \mathbf{C}_{\alpha_2}^{\text{HOSVD}} := \mathbf{C}_{\alpha_2}V. \end{aligned} \tag{11.44f}$$

Again, we write $\mathbf{B}_{\alpha_i}$ and $\mathbf{C}_{\alpha_i}$ instead of $\mathbf{B}_{\alpha_i}^{\text{HOSVD}}$ and $\mathbf{C}_{\alpha_i}^{\text{HOSVD}}$ and redefine r_{α_i} by

$$r_{\alpha_i} := r_{\alpha_i}^{\text{HOSVD}}.$$

The related weights are the diagonal matrices of Σ_{α_1} and Σ_{α_2} from (11.44d).

Remark 11.40. The computational work of the steps (11.44a-f) consists of a) $(r_{\alpha_1} + r_{\alpha_2})\,r_\alpha r_{\alpha_1} r_{\alpha_2}$ for forming $\mathbf{Z}_{\alpha_1}\mathbf{Z}_{\alpha_1}^\mathsf{H}$ and $\mathbf{Z}_{\alpha_2}^\mathsf{T}\overline{\mathbf{Z}_{\alpha_2}}$ and $\frac{8}{3}(r_{\alpha_1}^3 + r_{\alpha_2}^3)$ for the diagonalisation producing $U, \Sigma_{\alpha_1}, V, \Sigma_{\alpha_2}$, b) $2r_\alpha r_{\alpha_1}^{\text{HOSVD}} r_{\alpha_2}(r_{\alpha_1} + r_{\alpha_2}^{\text{HOSVD}})$ for (11.44e), c) $2(r_{\alpha_1} r_{\alpha_1}^{\text{HOSVD}} r_{\alpha_{11}} r_{\alpha_{12}} + r_{\alpha_2} r_{\alpha_2}^{\text{HOSVD}} r_{\alpha_{21}} r_{\alpha_{22}})$ for (11.44f), where $\alpha_{k1}, \alpha_{k2} \in S(\alpha_k)$, provided that $\alpha_k \in T_D \backslash \mathcal{L}(T_D)$. Otherwise, if $\alpha_1 = \{j\}$, (11.44f) costs $2(r_j r_j^{\text{HOSVD}} n_j)$, where $n_j = \dim(V_j)$. Bounding all r_γ by r and n_j by n, the total asymptotic work is $10r^4$ (if $\alpha_1, \alpha_2 \notin \mathcal{L}(T_D)$), $8r^4 + 2r^2n$ (if one leaf in $\{\alpha_1, \alpha_2\}$), and $6r^4 + 4r^2n$ (if $\alpha_1, \alpha_2 \in \mathcal{L}(T_D)$).

Exercise 11.41. Let $\{\alpha_1, \alpha_2\} = S(\alpha)$. $\mathbf{Z}_{\alpha_1}$ and $\mathbf{Z}_{\alpha_2}$ from (11.44a,b) formulated with the transformed matrices $C^{(\alpha,\ell)} = C^{(\alpha,i)}_{\mathrm{HOSVD}}$ satisfy

$$\begin{aligned} \mathbf{Z}_{\alpha_1}\mathbf{Z}_{\alpha_1}^{\mathsf{H}} &= \textstyle\sum_{\ell=1}^{r_\alpha}(\sigma_\ell^{(\alpha)})^2\, C^{(\alpha,\ell)}C^{(\alpha,\ell)\mathsf{H}} = \Sigma_{\alpha_1}^2,\\ \mathbf{Z}_{\alpha_2}^{\mathsf{T}}\,\overline{\mathbf{Z}_{\alpha_2}} &= \textstyle\sum_{\ell=1}^{r_\alpha}(\sigma_\ell^{(\alpha)})^2\, C^{(\alpha,\ell)\mathsf{T}}\overline{C^{(\alpha,\ell)}} = \Sigma_{\alpha_2}^2. \end{aligned} \tag{11.45}$$

Note that the case of the root $\alpha = D$ is not really different from inner vertices. Because of $r_D = 1$, the identity $\mathbf{Z}_{\alpha_1} = \mathbf{Z}_{\alpha_2}$ holds.

11.3.3.4 HOSVD Computation

First, we want to compute the HOSVD bases at *all* vertices. The algorithm starts at the root and proceeds to the leaves. The underlying computational step at vertex $\alpha \in T_D \backslash \mathcal{L}(T_D)$ is abbreviated as follows:

$$\boxed{\begin{array}{l} \textbf{procedure HOSVD}(\alpha); \qquad (\text{for } \alpha \in T_D\backslash\mathcal{L}(T_D) \text{ with sons } \alpha_1, \alpha_2)\\ \text{transform } C^{(\alpha,\ell)}\ (1 \le \ell \le r_\alpha) \text{ according to } \begin{cases} (11.43\text{d}) \text{ if } \alpha = D,\\ (11.44\text{e}) \text{ if } \alpha \ne D; \end{cases}\\ \text{transform } \left\{\begin{array}{ll} C^{(\alpha_1,\ell)} & (1 \le \ell \le r_{\alpha_1})\\ C^{(\alpha_2,\ell)} & (1 \le \ell \le r_{\alpha_2}) \end{array}\right\} \text{ according to } \begin{cases} (11.43\text{e}) \text{ if } \alpha = D,\\ (11.44\text{f}) \text{ if } \alpha \ne D; \end{cases}\\ \text{define the weights } \left\{\begin{array}{l} \Sigma_{\alpha_1} := \operatorname{diag}\{\sigma_1^{(\alpha_1)}, \ldots, \sigma_{r_{\alpha_1}}^{(\alpha_1)}\}\\ \Sigma_{\alpha_2} := \operatorname{diag}\{\sigma_1^{(\alpha_2)}, \ldots, \sigma_{r_{\alpha_2}}^{(\alpha_2)}\} \end{array}\right\}\\ \text{according to } \left\{\begin{array}{l} (11.43\text{f}) \text{ if } \alpha = D\\ (11.44\text{c}) \text{ if } \alpha \ne D \end{array}\right\} \text{ with possibly new } r_{\alpha_1}, r_{\alpha_2}; \end{array}} \tag{11.46a}$$

The complete computation of HOSVD bases at all vertices of T_D is performed by the call $\mathbf{HOSVD}^*(D)$ of the recursive procedure $\mathbf{HOSVD}^*(\alpha)$ defined by

$$\boxed{\begin{array}{l} \textbf{procedure HOSVD}^*(\alpha);\\ \text{if } \alpha \notin \mathcal{L}(T_D) \text{ then}\\ \text{begin } \mathbf{HOSVD}(\alpha);\ \text{for all sons } \sigma \in S(\alpha) \text{ do } \mathbf{HOSVD}^*(\sigma) \text{ end;} \end{array}} \tag{11.46b}$$

The derivation of the algorithm yields the following result.

Theorem 11.42. *Assume* $\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\big)$. *The result of* $\mathbf{HOSVD}^*(D)$ *is* $\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{HOSVD}}\big(T_D, (\mathbf{C}_\alpha^{\mathrm{HOSVD}})_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c_{\mathrm{HOSVD}}^{(D)}, (B_j^{\mathrm{HOSVD}})_{j\in D}\big)$. *The implicitly defined bases* $\mathbf{B}_\alpha^{\mathrm{HOSVD}} = [\mathbf{b}_{1,\mathrm{HOSVD}}^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha,\mathrm{HOSVD}}^{(\alpha)}]$ *for* $\alpha \in T_D$ *are the HOSVD bases. The computed tuples* Σ_α *contain the singular values.*

The computational cost for $\alpha = D$ and $\alpha \in T_D \backslash \{D \cup \mathcal{L}(T_D)\}$ is discussed in the Remarks 11.39 and 11.40. The total cost of $\mathbf{HOSVD}^*(D)$ sums to

$$N_{\mathrm{SVD}}(r_{\sigma_1}, r_{\sigma_2}) + 2\sum\nolimits_{j=1}^{d} r_j r_j^{\mathrm{HOSVD}} n_j \tag{11.46c}$$
$$+\,2 \sum_{\alpha\in T_D\backslash\{D\cup\mathcal{L}(T_D)\}} r_\alpha r_{\alpha_2}\bigl(r_\alpha^{\mathrm{HOSVD}} r_{\alpha_1} + r_{\alpha_1}^{\mathrm{HOSVD}}(r_{\alpha_1} + r_{\alpha_2}^{\mathrm{HOSVD}})\bigr),$$

where $\{\sigma_1, \sigma_2\} = D$ and $\{\alpha_1, \alpha_2\} = S(\alpha)$. If $r_\alpha \leq r$ and $n_j \leq n$, the asymptotic cost is $3(d-2)r^4 + 2dr^2 n$.

Remark 11.43. Algorithm (11.46b) uses a recursion over the tree T_D. Computations at the sons of a vertex are completely independent. This allows an easy *parallelisation*. This reduces the computational time by a factor $d/\log_2 d$.

We can use a similar recursion to obtain the HOSVD basis of a *single* vertex $\alpha \in T$. The algorithm shows that only the predecessors of α are involved.

```
procedure HOSVD$^{**}(\alpha)$;                                                    (11.46d)
begin if $\alpha \neq D$ then
      begin $\beta :=$ father$(\alpha)$; if HOSVD not yet installed at $\beta$ then HOSVD$^{**}(\beta)$
      end;
      HOSVD$(\alpha)$
end;
```

We recall that the tree T_D is decomposed in $T_D^{(\ell)}$ for the levels $0 \leq \ell \leq L$ (cf. (11.8)). The quite general recursion in (11.46b) can be performed levelwise:

```
procedure HOSVD-lw$(\ell)$;                                       (11.47a)
for all $\alpha \in T_D^{(\ell)}\backslash\mathcal{L}(T_D)$ do HOSVD$(\alpha)$;
```

To determine the HOSVD bases on level ℓ, we may call **HOSVD-lw**(ℓ), provided that HOSVD bases are already installed on level $\ell - 1$ (or if $\ell = 0$). Otherwise, one has to call

```
procedure HOSVD$^*$-lw$(\ell)$;                                   (11.47b)
for $\lambda = 0$ to $\ell$ do HOSVD-lw$(\lambda)$;
```

11.3.4 Sensitivity

The data of the hierarchical format, which may be subject to perturbations, consist mainly of the coefficients $c_{ij}^{(\alpha,\ell)}$ and the bases B_j at the leaves. Since basis vectors $\mathbf{b}_\ell^{(\alpha)}$ appear only implicitly, perturbations of $\mathbf{b}_\ell^{(\alpha)}$ are caused by perturbations of $c_{ij}^{(\alpha,\ell)}$ (usually, some coefficients $c_{ij}^{(\alpha,\ell)}$ are replaced by zero).

An important tool for the analysis are Gram matrices, which are considered in §11.3.4.1. The error analysis for the general (sub)orthonormal case is given in §11.3.4.2. HOSVD bases are considered in §11.3.4.3.

11.3.4.1 Gram Matrices and Suborthonormal Bases

We recall the definition of a Gram matrix (cf. (2.16)). For any basis $\mathbf{B}_\alpha$ ($\alpha \in T_D$) or B_j ($j \in D$) we set

$$G(\mathbf{B}_\alpha) := (g_{\nu\mu}^{(\alpha)})_{1\le\nu,\mu\le r_\alpha} \quad \text{with} \quad g_{\nu\mu}^{(\alpha)} := \left\langle \mathbf{b}_\mu^{(\alpha)}, \mathbf{b}_\nu^{(\alpha)} \right\rangle. \tag{11.48a}$$

Similarly, the tuple $\mathbf{C}_\alpha := \left(C^{(\alpha,\ell)}\right)_{1\le\ell\le r_\alpha}$ of coefficient matrices is associated with

$$G(\mathbf{C}_\alpha) := (g_{\nu\mu})_{1\le\nu,\mu\le r_\alpha} \quad \text{with} \quad g_{\nu\mu} := \left\langle C^{(\alpha,\mu)}, C^{(\alpha,\nu)} \right\rangle_{\mathsf{F}}. \tag{11.48b}$$

First we consider the situation of a vertex $\alpha \in T_D \backslash \mathcal{L}(T_D)$ with sons α_1, α_2. In the following lemma the data $\mathbf{B}_{\alpha_1}, \mathbf{B}_{\alpha_2}$ are general variables; they may be bases or their perturbations. The crucial question is whether the mapping $(\mathbf{B}_{\alpha_1}, \mathbf{B}_{\alpha_2}) \mapsto \mathbf{B}_\alpha$ defined in (11.24) is stable.

Lemma 11.44. *Let $\alpha_1, \alpha_2 \in S(\alpha)$. $\mathbf{B}_{\alpha_1} \in (\mathbf{V}_{\alpha_1})^{r_{\alpha_1}}$ and $\mathbf{B}_{\alpha_2} \in (\mathbf{V}_{\alpha_2})^{r_{\alpha_2}}$ are mapped by $\mathbf{b}_\ell^{(\alpha)} = \sum_{i,j} c_{ij}^{(\alpha,\ell)} \, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}$ into $\mathbf{B}_\alpha \in (\mathbf{V}_\alpha)^{r_\alpha}$. The related Gram matrices satisfy*

$$\|G(\mathbf{B}_\alpha)\|_2 \le \|G(\mathbf{C}_\alpha)\|_2 \, \|G(\mathbf{B}_{\alpha_1})\|_2 \, \|G(\mathbf{B}_{\alpha_2})\|_2 \, . \tag{11.49}$$

Proof. According to Lemma 2.17, there are coefficients $\xi_i \in \mathbb{K}$ with $\sum_{\ell\in1}^{r_\alpha} |\xi_\ell|^2 = 1$ and $\|G(\mathbf{B}_\alpha)\|_2 = \left\| \sum_{\ell=1}^{r_\alpha} \xi_\ell \mathbf{b}_\ell^{(\alpha)} \right\|_2^2$. Summation over ℓ yields $c_{ij}^{(\alpha)} := \sum_{\ell=1}^{r_\alpha} \xi_\ell \, c_{ij}^{(\alpha,\ell)}$ and the matrix $C_\alpha = (c_{ij}^{(\alpha)})$. With this notation and the abbreviations $G_i := G(\mathbf{B}_{\alpha_i})$ for $i = 1, 2$, we continue:

$$\begin{aligned}
\left\| \sum_{\ell=1}^{r_\alpha} \xi_\ell \, \mathbf{b}_\ell^{(\alpha)} \right\|_2^2 &= \left\| \sum_{i,j} c_{ij}^{(\alpha)} \, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} \right\|_2^2 \\
&= \left\langle \sum_{i,j} c_{ij}^{(\alpha)} \, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}, \sum_{i',j'} c_{i'j'}^{(\alpha)} \, \mathbf{b}_{i'}^{(\alpha_1)} \otimes \mathbf{b}_{j'}^{(\alpha_2)} \right\rangle \\
&= \sum_{i,j} \sum_{i',j'} c_{ij}^{(\alpha)} \, g_{j'j}^{(\alpha_2)} \, \overline{c_{i'j'}^{(\alpha)}} \, g_{i'i}^{(\alpha_1)} \\
&= \operatorname{trace}(C_\alpha G_2^{\mathsf{T}} C_\alpha^{\mathsf{H}} G_1).
\end{aligned}$$

Set $\hat{C} := G_1^{1/2} C_\alpha (G_2^{1/2})^{\mathsf{T}}$. Exercise 2.7a allows us to rewrite the trace as follows:

$$\operatorname{trace}(C_\alpha G_2^{\mathsf{T}} C_\alpha^{\mathsf{H}} G_1) = \operatorname{trace}(\hat{C}\hat{C}^{\mathsf{H}}) = \langle \hat{C}, \hat{C} \rangle_{\mathsf{F}} = \|\hat{C}\|_{\mathsf{F}}^2.$$

Thanks to Lemma 2.10, we can estimate by[17]

[17] Here we use that positive semi-definite matrices satisfy $\|G^{1/2}\|^2 = (\rho(G^{1/2}))^2 = \rho(G) = \|G\|$ (ρ: spectral radius from (4.76)).

$$\|\hat{C}\|_\mathsf{F}^2 = \|G_1^{\frac{1}{2}} C_\alpha (G_2^{\frac{1}{2}})^\mathsf{T}\|_\mathsf{F}^2 \le \left[\|G_1^{\frac{1}{2}}\|_2 \|C_\alpha\|_\mathsf{F} \|G_2^{\frac{1}{2}}\|_2\right]^2 = \|G_1\|_2 \|C_\alpha\|_\mathsf{F}^2 \|G_2\|_2$$

Now we use $C_\alpha = \sum_{\ell=1}^{r_\alpha} \xi_\ell \, C^{(\alpha,\ell)}$ and apply Lemma 2.17 (with the Euclidean scalar product replaced by the Frobenius scalar product):

$$\|C_\alpha\|_\mathsf{F}^2 = \left\|\sum_{\ell\in 1}^{r_\alpha} \xi_\ell C^{(\alpha,\ell)}\right\|_\mathsf{F}^2 \le \max_{\sum_\ell |\eta_\ell|^2=1} \left\|\sum_{\ell\in 1}^{r_\alpha} \eta_\ell C^{(\alpha,\ell)}\right\|_\mathsf{F}^2 = \|G(\mathbf{C}_\alpha)\|_2 \,.$$

Putting all estimates together, we obtain the desired estimate. □

This result deserves some comments.

1) Orthonormality of the basis $\mathbf{B}_{\alpha_i}$ is equivalent to $G(\mathbf{B}_{\alpha_i}) = I$. According to Lemma 11.31, orthonormal matrices $C^{(\alpha,\ell)}$ produce an orthonormal basis $\mathbf{B}_\alpha$, i.e., $G(\mathbf{B}_\alpha) = I$. Under these assumptions, inequality (11.49) takes the form $1 \le 1 \cdot 1 \cdot 1$.

2) The quantity $\|G(\mathbf{B}_\alpha)\|_2$ is a reasonable one, since it is an estimate for all expressions $\| \sum_{\ell=1}^{r_\alpha} \xi_\ell \mathbf{b}_\ell^{(\alpha)} \|_2^2$ with $\sum_{\ell=1}^{r_\alpha} |\xi_\ell|^2 = 1$ (cf. Lemma 2.17).

3) Starting from the orthonormal setting (i.e., $\|G(\ldots)\|_2 = 1$), we shall see that truncations lead to $\|G(\ldots)\|_2 \le 1$. Therefore, errors will not be amplified.

A typical truncation step at vertex α_1 omits a vector of the basis, say, $\mathbf{b}_{r_{\alpha_1}}^{(\alpha_1)}$ keeping $\mathbf{B}_{\alpha_1}^{\text{new}} = [\mathbf{b}_1^{(\alpha_1)} \cdots \mathbf{b}_{r_{\alpha_1}-1}^{(\alpha_1)}]$. Although $\mathbf{B}_{\alpha_1}^{\text{new}}$ and $\mathbf{B}_{\alpha_2}$ represent orthonormal bases, the resulting basis $\mathbf{B}_\alpha^{\text{new}}$ at the father vertex α is no longer orthonormal. $G(\mathbf{B}_\alpha^{\text{new}})$ corresponds to $G(\mathbf{C}_\alpha^{\text{new}})$, where $\mathbf{C}_\alpha^{\text{new}} = (C_{\text{new}}^{(\alpha,\ell)})_{1\le\ell\le r_\alpha}$ is obtained by omitting the r_{α_1}-th row in $C^{(\alpha,\ell)}$. However, still the inequality $G(\mathbf{B}_\alpha^{\text{new}}) \le I$ can be shown (cf. Exercise 11.48).

Exercise 11.45. Prove $\|\mathbf{b}_\ell^{(\alpha)}\|_2^2 \le \|G(\mathbf{B}_{\alpha_1})\|_2 \|G(\mathbf{B}_{\alpha_2})\|_2 \|C^{(\alpha,\ell)}\|_\mathsf{F}^2$ for $1 \le \ell \le r_\alpha$.

Exercise 11.46. (a) Prove that $\sqrt{\|G(\mathbf{B}+\mathbf{C})\|_2} \le \sqrt{\|G(\mathbf{B})\|_2} + \sqrt{\|G(\mathbf{C})\|_2}$.
(b) Let $\mathbf{B}, \mathbf{C} \in \mathbb{K}^{n\times m}$ be pairwise orthogonal, i.e., $\mathbf{B}^\mathsf{H}\mathbf{C} = 0$. Prove that

$$G(\mathbf{B}+\mathbf{C}) = G(\mathbf{B}) + G(\mathbf{C}),$$
$$\|G(\mathbf{B}+\mathbf{C})\|_2 \le \|G(\mathbf{B})\|_2 + \|G(\mathbf{C})\|_2.$$

Definition 11.47. An n-tuple of linearly independent vectors $\mathfrak{x} = (x_1, \ldots, x_n)$ is called *suborthonormal*, if the corresponding Gram matrix satisfies

$$0 < G(\mathfrak{x}) \le I \qquad \text{(cf. (2.14))}\,.$$

Exercise 11.48. Show for any $\mathbf{B} \in \mathbb{K}^{n\times m}$ and any orthogonal projection $P \in \mathbb{K}^{n\times n}$ that

$$G(P\mathbf{B}) \le G(\mathbf{B}).$$

Hint: Use Remark 2.14b with $P = P^\mathsf{H} P \le I$.

11.3.4.2 Orthonormal Bases

Here we suppose that all bases $\mathbf{B}_\alpha \in (\mathbf{V}_\alpha)^{r_\alpha}$ ($\alpha \in T_D$) are orthonormal or, more generally, suborthonormal. This fact implies $G(\mathbf{B}_\alpha) \leq I$, $G(\mathbf{C}_\beta) \leq I$ and thus $\|G(\mathbf{B}_\alpha)\|_2, \|G(\mathbf{C}_\beta)\|_2 \leq 1$. Perturbations may be caused as follows.

1. At a leaf $j \in D$, the basis B_j may be altered. We may even reduce the dimension by omitting one of the basis vectors (which changes some $b_i^{(j)}$ into 0). Whenever the new basis B_j^{new} is not orthonormal or the generated subspace has a smaller dimension, the implicitly defined bases $\mathbf{B}_\alpha^{\mathrm{new}}$ with $j \in \alpha$ also lose orthonormality.
2. Let $\alpha \in T_D \backslash \mathcal{L}(T_D)$ have sons α_1, α_2. Changing $\mathbf{C}_\alpha$, we can rotate the basis $\mathbf{B}_\alpha$ into a new orthonormal basis $\mathbf{B}_\alpha^{\mathrm{new}}$ satisfying the nestedness property $\mathrm{range}(\mathbf{B}_\alpha) \subset \mathrm{range}(\mathbf{B}_{\alpha_1}) \otimes \mathrm{range}(\mathbf{B}_{\alpha_2})$. This does not change $G(\mathbf{B}_\beta^{\mathrm{new}}) = I$ and $G(\mathbf{C}_\beta) = I$ for all predecessors $\beta \in T_D$ (i.e., $\beta \supset \alpha$).
3. We may omit, say, $\mathbf{b}_{r_\alpha}^{(\alpha)}$ from $\mathbf{B}_\alpha \in (\mathbf{V}_\alpha)^{r_\alpha}$ by setting $C_{\mathrm{new}}^{(\alpha, r_\alpha)} := 0$. If $\mathbf{B}_\alpha$ is (sub)orthonormal, $\mathbf{B}_\alpha^{\mathrm{new}}$ is so too. For $\beta \supset \alpha$, $G(\mathbf{B}_\beta^{\mathrm{new}}) \leq G(\mathbf{B}_\beta)$ follows, i.e., an orthonormal basis $\mathbf{B}_\beta$ becomes a suborthonormal basis $\mathbf{B}_\alpha^{\mathrm{new}}$. The inequality $G(\mathbf{C}_\alpha^{\mathrm{new}}) \leq G(\mathbf{C}_\alpha)$ holds.

We consider a general perturbation $\delta\mathbf{B}_\alpha \in (\mathbf{V}_\alpha)^{r_\alpha}$, i.e., the exact $\mathbf{B}_\alpha$ is changed into

$$\mathbf{B}_\alpha^{\mathrm{new}} := \mathbf{B}_\alpha - \delta\mathbf{B}_\alpha$$

at *one* vertex $\alpha \in T_D$. Let $\beta \in T_D$ be the father of α such that $\beta_1, \beta_2 \in S(\beta)$ are the sons and, e.g., $\alpha = \beta_1$. A perturbation $\delta\mathbf{B}_\alpha$ causes a change of $\mathbf{B}_\beta$ into $\mathbf{B}_\beta^{\mathrm{new}} = \mathbf{B}_\beta - \delta\mathbf{B}_\beta$. Because of the linear dependence, $\delta\mathbf{b}_\ell^{(\beta)} = \sum_{i,j} c_{ij}^{(\beta,\ell)}\, \delta\mathbf{b}_i^{(\alpha)} \otimes \mathbf{b}_j^{(\beta_2)}$ holds for the columns of $\delta\mathbf{B}_\beta$, $\delta\mathbf{B}_\alpha$, and inequality (11.49) implies

$$\|G(\delta\mathbf{B}_\beta)\|_2 \leq \|G(\mathbf{C}_\beta)\|_2 \, \|G(\delta\mathbf{B}_\alpha)\|_2 \, \|G(\mathbf{B}_{\beta_2})\|_2 \leq \|G(\delta\mathbf{B}_\alpha)\|_2 \, .$$

The inequality $\|G(\delta\mathbf{B}_\beta)\|_2 \leq \|G(\delta\mathbf{B}_\alpha)\|_2$ can be continued up to the root D yielding $\|G(\delta\mathbf{B}_D)\|_2 \leq \|G(\delta\mathbf{B}_\alpha)\|_2$. However, since $r_D = 1$, $G(\delta\mathbf{B}_D) = \|\delta\mathbf{b}_1^{(D)}\|_2^2$ is an 1×1 matrix. This proves the following result.

Theorem 11.49. *Given* $\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big)$ *as defined in (11.38), consider a perturbation* $\delta\mathbf{B}_\alpha$ *at some* $\alpha \in T_D$ *(in particular,* δB_j *for* $\alpha = \{j\}$ *is of interest). Then* $\mathbf{v}$ *is changed into* $\mathbf{v}^{\mathrm{new}} = \mathbf{v} - \delta\mathbf{v}$ *with*

$$\|\delta\mathbf{v}\|_2 \leq |c_1^{(D)}| \sqrt{\|G(\delta\mathbf{B}_\alpha)\|_2}.$$

If β is the father of $\alpha \in T_D$, the fact that the perturbed basis $\mathbf{B}_\alpha^{\mathrm{new}} = \mathbf{B}_\alpha - \delta\mathbf{B}_\alpha$ is, in general, no longer orthonormal, implies that also $\mathbf{B}_\beta^{\mathrm{new}} = \mathbf{B}_\beta - \delta\mathbf{B}_\beta$ loses orthonormality, although the coefficients $\mathbf{C}_\beta$ still satisfy $G(\mathbf{C}_\beta) = I$.

Corollary 11.50. Let $A \subset T_D$ be a subset of size $\#A$ (hence, $\#A \leq 2d - 1$). Perturbations $\delta\mathbf{B}_\alpha$ at all $\alpha \in A$ yield the error

$$\|\delta\mathbf{v}\|_2 \le |c_1^{(D)}|\sqrt{\#A}\sqrt{\sum\nolimits_{\alpha\in A}\|G(\delta\mathbf{B}_\alpha)\|_2} + \text{higher order terms.}$$

A first perturbation changes some $\mathbf{B}_\beta$ into $\mathbf{B}_\beta^{\mathrm{new}}$. If $\|G(\mathbf{B}_\beta)\|_2 \le 1$, i.e., if $\mathbf{B}_\beta^{\mathrm{new}}$ is still suborthonormal, the second perturbation is not amplified. However, $\|G(\mathbf{B}_\beta^{\mathrm{new}})\|_2$ may become larger than one. Exercise 11.46a allows us to bound the norm by

$$\begin{aligned}\|G(\mathbf{B}_\beta^{\mathrm{new}})\|_2 &\le \left[\sqrt{\|G(\mathbf{B}_\beta)\|_2} + \sqrt{\|G(\delta\mathbf{B}_\beta)\|_2}\right]^2 \\ &= \left[1 + \sqrt{\|G(\delta\mathbf{B}_\beta)\|_2}\right]^2 = 1 + O\left(\sqrt{\|G(\delta\mathbf{B}_\beta)\|_2}\right).\end{aligned}$$

If this factor is not compensated by other factors smaller than 1, higher order terms appear in the estimate.

So far, we have considered general perturbations $\delta\mathbf{B}_\alpha$ leading to perturbations $\delta\mathbf{v}_\alpha$ of $\mathbf{v}$. If the perturbations $\delta\mathbf{v}_\alpha$ are orthogonal, we can derive a better estimate of $\|G(\cdot)\|_2$ (cf. Exercise 11.46b).

In connection with the HOSVD truncations discussed in §11.4.2, we shall continue the error analysis.

11.3.4.3 HOSVD Bases

Since HOSVD bases are particular orthonormal bases, the results of §11.3.4.2 are still valid. However, now the weights Σ_α from (11.43f) and (11.44d) enable a weighted norm of the error, which turns out to be optimal for our purpose.

First, we assume that HOSVD bases are installed at all vertices $\alpha \in T_D$. The coefficient matrices $C^{(\alpha,\ell)}$ together with the weights $\Sigma_{\alpha_i} = \mathrm{diag}\{\sigma_1^{(\alpha_i)}, \ldots, \sigma_{r_{\alpha_i}}^{(\alpha_i)}\}$ $(i = 1, 2)$ at the son vertices $\{\alpha_1, \alpha_2\} = S(\alpha)$ satisfy (11.45) (cf. Exercise 11.41). A perturbation $\delta\mathbf{B}_\tau$ of the basis $\mathbf{B}_\tau$ is described by $\delta\mathbf{B}_\tau = [\delta\mathbf{b}_1^{(\tau)}, \ldots, \delta\mathbf{b}_{r_\tau}^{(\tau)}]$. The weights $\sigma_i^{(\tau)}$ from Σ_τ are used for the formulation of the error:

$$\varepsilon_\tau := \sqrt{\sum_{i=1}^{r_\tau}\left(\sigma_i^{(\tau)}\|\delta\mathbf{b}_i^{(\tau)}\|\right)^2} \qquad (\tau \in T_D). \tag{11.50}$$

The next proposition describes the error transport from the son α_1 to the father α.

Proposition 11.51. *Suppose that the basis $\mathbf{B}_\alpha$ satisfies the first HOSVD property (11.45) for $\alpha_1 \in S(\alpha)$. Let $\delta\mathbf{B}_{\alpha_1}$ be a perturbation*[18] *of $\mathbf{B}_{\alpha_1} = [\mathbf{b}_1^{(\alpha_1)}, \ldots, \mathbf{b}_{r_{\alpha_1}}^{(\alpha_1)}]$ into $\mathbf{B}_{\alpha_1} - \delta\mathbf{B}_{\alpha_1}$. The perturbation is measured by ε_{α_1} from (11.50). The (exact) basis $\mathbf{B}_\alpha = [\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}]$ with $\mathbf{b}_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}}\sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)}\,\mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}$ is perturbed into $\mathbf{B}_\alpha - \delta\mathbf{B}_\alpha$ with*

[18] This and following statements are formulation for the *first* son α_1. The corresponding result for α_2 is completely analogous.

$$\delta\mathbf{B}_\alpha = \left[\delta\mathbf{b}_1^{(\alpha)},\ldots,\delta\mathbf{b}_{r_\alpha}^{(\alpha)}\right] \quad \textit{and} \quad \delta\mathbf{b}_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}}\sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)}\,\delta\mathbf{b}_i^{(\alpha_1)}\otimes\mathbf{b}_j^{(\alpha_2)}.$$

Then the perturbations $\delta\mathbf{b}_\ell^{(\alpha)}$ *lead to an error of equal size:*

$$\varepsilon_\alpha := \sqrt{\sum_{\ell=1}^{r_\alpha}\left(\sigma_\ell^{(\alpha)}\|\delta\mathbf{b}_\ell^{(\alpha)}\|\right)^2} = \varepsilon_{\alpha_1}.$$

Proof. By orthonormality of the basis $\mathbf{B}_{\alpha_2}$ we have

$$\|\delta\mathbf{b}_\ell^{(\alpha)}\|^2 = \left\|\sum_{i=1}^{r_{\alpha_1}}\sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)}\,\delta\mathbf{b}_i^{(\alpha_1)}\otimes\mathbf{b}_j^{(\alpha_2)}\right\|^2 = \sum_{j=1}^{r_{\alpha_2}}\left\|\sum_i c_{ij}^{(\alpha,\ell)}\,\delta\mathbf{b}_i^{(\alpha_1)}\right\|^2$$

$$= \sum_{j=1}^{r_{\alpha_2}}\left\langle \sum_i c_{ij}^{(\alpha,\ell)}\,\delta\mathbf{b}_i^{(\alpha_1)}, \sum_{i'} c_{i'j}^{(\alpha,\ell)}\,\delta\mathbf{b}_{i'}^{(\alpha_1)}\right\rangle = \sum_{i,i',j} c_{ij}^{(\alpha,\ell)}\overline{c_{i'j}^{(\alpha,\ell)}}\left\langle\delta\mathbf{b}_i^{(\alpha_1)},\delta\mathbf{b}_{i'}^{(\alpha_1)}\right\rangle.$$

The Gram matrix $G = G(\delta\mathbf{B}_{\alpha_1}) \in \mathbb{K}^{r_{\alpha_1}\times r_{\alpha_1}}$ has the entries $G_{i'i} := \left\langle\delta\mathbf{b}_i^{(\alpha_1)},\delta\mathbf{b}_{i'}^{(\alpha_1)}\right\rangle$. The sum over i, i', j from above equals

$$\operatorname{trace}\left((C^{(\alpha,\ell)})^{\mathsf{H}}\,G\,C^{(\alpha,\ell)}\right) = \operatorname{trace}\left(G^{1/2}C^{(\alpha,\ell)}(C^{(\alpha,\ell)})^{\mathsf{H}}G^{1/2}\right)$$

(cf. Exercise 2.7a). From (11.45) we derive

$$\sum_{\ell=1}^{r_\alpha}\left(\sigma_\ell^{(\alpha)}\|\delta\mathbf{b}_\ell^{(\alpha)}\|\right)^2 = \sum_{\ell=1}^{r_\alpha}\left(\sigma_\ell^{(\alpha)}\right)^2\operatorname{trace}(G^{1/2}C^{(\alpha,\ell)}(C^{(\alpha,\ell)})^{\mathsf{H}}G^{1/2})$$

$$= \operatorname{trace}\left\{G^{\frac12}\left[\textstyle\sum_{\ell=1}^{r_\alpha}\left(\sigma_\ell^{(\alpha)}\right)^2 C^{(\alpha,\ell)}(C^{(\alpha,\ell)})^{\mathsf{H}}\right]G^{\frac12}\right\} = \operatorname{trace}\left(G^{\frac12}\varSigma_{\alpha_1}^2G^{\frac12}\right)$$

$$= \operatorname{trace}\left(\varSigma_{\alpha_1}\,G\,\varSigma_{\alpha_1}\right) = \sum_{i=1}^{r_\alpha}\left(\sigma_i^{(\alpha_1)}\|\delta b_i^{(\alpha_1)}\|\right)^2 = \varepsilon_{\alpha_1}^2$$

concluding the proof. □

The HOSVD condition (11.45) can be weakened:

$$\sum_{\ell=1}^{r_\alpha}(\sigma_\ell^{(\alpha)})^2\,C^{(\alpha,\ell)}(C^{(\alpha,\ell)})^{\mathsf{H}} \le \varSigma_{\alpha_1}^2,\quad \sum_{\ell=1}^{r_\alpha}(\sigma_\ell^{(\alpha)})^2\,(C^{(\alpha,\ell)})^{\mathsf{H}}C^{(\alpha,\ell)} \le \varSigma_{\alpha_2}^2 \tag{11.51}$$

for $\alpha_1, \alpha_2 \in S(\alpha)$. Furthermore, the basis $\mathbf{B}_\tau = [\mathbf{b}_1^{(\tau)},\ldots,\mathbf{b}_{r_\tau}^{(\tau)}]$ ($\tau \in T_D$) may be suborthonormal, i.e., the Gram matrix $G(\mathbf{B}_\tau) := \mathbf{B}_\tau^{\mathsf{H}}\mathbf{B}_\tau$ satisfies

$$G(\mathbf{B}_\tau) \le I \qquad (\tau \in T_D). \tag{11.52}$$

The statement of Proposition 11.51 can be generalised for the weak HOSVD condition.

Corollary 11.52. Suppose that the basis $\mathbf{B}_\alpha$ satisfies the first equality in (11.51) for $\alpha_1 \in S(\alpha)$. Let $\delta\mathbf{B}_{\alpha_1}$ be a perturbation of $\mathbf{B}_{\alpha_1}$ measured by ε_{α_1} from (11.50). The basis $\mathbf{B}_{\alpha_2}$ at the other son $\alpha_2 \in S(\alpha)$ may be suborthonormal (cf. (11.52)). Then the perturbations $\delta\mathbf{b}_\ell^{(\alpha)}$ are estimated by

$$\varepsilon_\alpha := \sqrt{\sum_{\ell=1}^{r_\alpha}\left(\sigma_\ell^{(\alpha)}\|\delta\mathbf{b}_\ell^{(\alpha)}\|\right)^2} \le \varepsilon_{\alpha_1}.$$

Proof. 1) We have $\|\delta\mathbf{b}_\ell^{(\alpha)}\|^2 = \left\|\sum_{j=1}^{r_{\alpha_2}}\mathbf{c}_j\otimes\mathbf{b}_j^{(\alpha_2)}\right\|^2$ for $\mathbf{c}_j := \sum_{i=1}^{r_{\alpha_1}} c_{ij}^{(\alpha,\ell)}\,\delta\mathbf{b}_i^{(\alpha_1)}$. Using the Gram matrices $C := G(\mathbf{c})$ and $G := G(\mathbf{B}_{\alpha_2})$, the identity

$$\begin{aligned}\left\|\sum\nolimits_j \mathbf{c}_j\otimes\mathbf{b}_j^{(\alpha_2)}\right\|^2 &= \sum\nolimits_{j,k}\langle\mathbf{c}_j,\mathbf{c}_k\rangle\left\langle\mathbf{b}_j^{(\alpha_2)},\mathbf{b}_k^{(\alpha_2)}\right\rangle \underset{C=C^{\mathsf{H}}}{=} \operatorname{trace}(CG)\\ &= \operatorname{trace}(C^{1/2}GC^{1/2})\end{aligned}$$

holds. Applying (11.52) and Remark 2.14d, we can use the inequality

$$\begin{aligned}\operatorname{trace}(C^{1/2}GC^{1/2}) \le \operatorname{trace}(C^{1/2}C^{1/2}) &= \operatorname{trace}(C) = \sum\nolimits_j\langle\mathbf{c}_j,\mathbf{c}_j\rangle\\ &= \sum_{j=1}^{r_{\alpha_2}}\left\|\sum_{i=1}^{r_{\alpha_1}} c_{ij}^{(\alpha,\ell)}\,\delta\mathbf{b}_i^{(\alpha_1)}\right\|^2\end{aligned}$$

to continue with the estimates in the proof of Proposition 11.51. In the last lines of the proof $\operatorname{trace}\{\ldots\} = \operatorname{trace}\left(\delta^{1/2}\Sigma_{\alpha_1}^2\delta^{1/2}\right)$ has to be replaced by the inequality $\operatorname{trace}\{\ldots\} \le \operatorname{trace}\left(\delta^{1/2}\Sigma_{\alpha_1}^2\delta^{1/2}\right)$. □

Theorem 11.53. *Suppose $\mathbf{v}\in\mathcal{H}_\mathfrak{r}$. Assume that all $\mathbf{B}_\tau$, $\tau\in T_D$, are weak HOSVD bases is the sense of (11.52) and that all coefficient matrices $C^{(\alpha,\ell)}$ together with the weights Σ_α satisfy (11.51). Then a perturbation of the basis at vertex $\alpha\in T_D$ by*

$$\varepsilon_\alpha := \sqrt{\sum_{\ell=1}^{r_\alpha}\left(\sigma_\ell^{(\alpha)}\|\delta\mathbf{b}_\ell^{(\alpha)}\|\right)^2}$$

leads to an absolute error of $\mathbf{v}$ by

$$\|\delta\mathbf{v}\| \le \varepsilon_\alpha.$$

Proof. Corollary 11.52 shows that the same kind of error at the father vertex does not increase. By induction, one obtains $\varepsilon_D = \sqrt{\sum_{\ell=1}^{r_D}(\sigma_\ell^{(D)}\|\delta\mathbf{b}_\ell^{(D)}\|)^2} \le \varepsilon_\alpha$ at the root $D\in T_D$. Since $r_D = 1$ and $\sigma_1^{(D)} = \|\mathbf{v}\|$, it follows that $\varepsilon_D = \|\mathbf{v}\|\|\delta\mathbf{b}_1^{(D)}\|$. On the other hand, $\mathbf{v} = c_1^{(D)}\mathbf{b}_1^{(D)}$ with $\|\mathbf{v}\| = |c_1^{(D)}|$ proves that the perturbation is $\|\delta\mathbf{v}\| = \|c_1^{(D)}\delta\mathbf{b}_1^{(D)}\| = \|\mathbf{v}\|\|\delta\mathbf{b}_1^{(D)}\|$. □

11.3.5 Conversion from $\mathcal{R}_r$ to $\mathcal{H}_\mathfrak{r}$ Revisited

In §11.2.4.2, the conversion from r-term representation $\mathcal{R}_r$ into hierarchical format $\mathcal{H}_\mathfrak{r}$ has been discussed on the level of subspaces. Now we consider the choice of bases. The input tensor is

$$\mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j\in D} v_i^{(j)} \qquad \text{with } v_i^{(j)} \in V_j \text{ for } j \in D. \tag{11.53a}$$

Let T_D be a suitable dimension partition tree for D. The easiest approach is to choose

$$\mathbf{b}_i^{(\alpha)} := \bigotimes_{j\in\alpha} v_i^{(j)} \qquad (1 \le i \le r,\ \alpha \in T_D\backslash\{D\}) \tag{11.53b}$$

as a frame (cf. (11.19)). Note that $r_\alpha = r$ for all $\alpha \in T_D\backslash\{D\}$. Because of

$$\mathbf{b}_i^{(\alpha)} = \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_i^{(\alpha_2)} \quad (1 \le i \le r,\ \alpha_1, \alpha_2 \text{ sons of } \alpha), \tag{11.53c}$$

the coefficient matrices are of extremely sparse form:

$$c_{ij}^{(\alpha,\ell)} = \left\{ \begin{matrix} 1 & \text{if } \ell = i = j \\ 0 & \text{otherwise.} \end{matrix} \right\} \quad 1 \le i, j, \ell \le r,\ \alpha \in T_D\backslash\{D\}. \tag{11.53d}$$

Only for $\alpha = D$, the definition $\mathbf{b}_1^{(D)} = \sum_{i=1}^r \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_i^{(\alpha_2)}$ is used to ensure $r_D = 1$, i.e., $C^{(D,1)} = I$. In particular, the matrices $C^{(\alpha,\ell)}$ are sparse, diagonal, and of rank one for $\alpha \neq D$.

Now we denote the basis vectors $\mathbf{b}_i^{(\alpha)}$ from (11.53b,c) by $\mathbf{b}_{i,\text{old}}^{(\alpha)}$. Below we construct an orthonormal hierarchical representation.

Step 1 (orthogonalisation at the leaves): By QR or Gram matrix techniques one can obtain an orthonormal bases $\{b_i^{(j)} : 1 \le i \le r_j\}$ of $\operatorname{span}\{v_i^{(j)} : 1 \le i \le r\}$ with $r_j \le r$ (cf. Lemma 8.12). Set $B_j = [b_1^{(j)}, \ldots, b_{r_j}^{(j)}]$. Then there is a transformation matrix $T^{(j)}$ with

$$B_j^{\text{old}} = [v_1^{(j)}, \ldots, v_r^{(j)}] = B_j T^{(j)}, \quad \text{i.e., } v_\ell^{(j)} = \sum_{i=1}^{r_k} t_{i\ell}^{(j)} b_i^{(j)} \ (1 \le \ell \le r)\,.$$

Step 2 (non-leaf vertices): The sons of $\alpha \in T_D\backslash\mathcal{L}(T_D)$ are denoted by α_1 and α_2. By induction, we have $\mathbf{b}_{\ell,\text{old}}^{(\alpha_\nu)} = \sum_i t_{i\ell}^{(\alpha_\nu)} \mathbf{b}_i^{(\alpha_\nu)}$ $(\nu = 1, 2)$. From (11.53c) we conclude that

$$\mathbf{b}_{\ell,\text{old}}^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} t_{i\ell}^{(\alpha_1)} t_{j\ell}^{(\alpha_2)}\ \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_i^{(\alpha_2)} \qquad (1 \le \ell \le r) \tag{11.54a}$$

with the coefficient matrix

$$C_{\mathrm{old}}^{(\alpha,\ell)} = \left(c_{ij}^{(\alpha,\ell)}\right), \qquad c_{ij}^{(\alpha,\ell)} := t_{i\ell}^{(\alpha_1)} t_{j\ell}^{(\alpha_2)} \tag{11.54b}$$

$(1 \le \ell \le r,\ 1 \le i \le r_{\alpha_1},\ 1 \le j \le r_{\alpha_2})$. The identity $C_{\mathrm{old}}^{(\alpha,\ell)} = ab^{\mathsf{T}}$ for the vectors $a = T_{\bullet,\ell}^{(\alpha_1)}$, $b = T_{\bullet,\ell}^{(\alpha_2)}$ proves the first statement of the next remark.

Remark 11.54. (a) The matrices $C_{\mathrm{old}}^{(\alpha,\ell)}$ from (11.54b) are of rank 1.
(b) The Gram matrix $G_\alpha = G(\mathbf{C}_\alpha)$ has coefficients $g_{m\ell}^{(\alpha)} = \left\langle C_{\mathrm{old}}^{(\alpha,\ell)}, C_{\mathrm{old}}^{(\alpha,m)} \right\rangle_{\mathsf{F}}$ (cf. (11.37b)) of the form

$$g_{m\ell}^{(\alpha)} = \left(\sum_{i=1}^{r_{\alpha_1}} s_{i\ell}^{(\alpha_1)} s_{im}^{(\alpha_1)}\right)\left(\sum_{j=1}^{r_{\alpha_2}} s_{j\ell}^{(\alpha_2)} s_{jm}^{(\alpha_2)}\right).$$

Therefore, the computational cost for orthonormalisation is $2r_\alpha^2\,(r_{\alpha_1} + r_{\alpha_2})$ (instead of $2r_\alpha^2 r_{\alpha_1} r_{\alpha_2}$ as stated in Lemma 11.31 for the general case).

Using G_α, we can construct an orthonormal basis $\{\mathbf{b}_\ell^{(\alpha)} : 1 \le \ell \le r_\alpha\}$ leading to $\mathbf{B}_\alpha^{\mathrm{old}} = \mathbf{B}_\alpha T^{(\alpha)}$ and an updated $C^{(\alpha,\ell)}$.

Step 3 (root): For $\alpha = D$ choose $r_D = 1$ with obvious modifications of Step 2.

11.4 Approximations in $\mathcal{H}_\mathfrak{r}$

11.4.1 Best Approximation in $\mathcal{H}_\mathfrak{r}$

11.4.1.1 Existence

Lemma 8.6 has the following counterpart for $\mathcal{H}_\mathfrak{r}$.

Lemma 11.55. *Let $\mathbf{V} = {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j$ be a Banach tensor space with a norm not weaker than $\|\cdot\|_\vee$ (cf. (6.18)). Then the subset $\mathcal{H}_\mathfrak{r} \subset \mathbf{V}$ is weakly closed.*

Proof. Let $\mathbf{v}^{(\nu)} \in \mathcal{H}_\mathfrak{r}$ be a weakly convergent: $\mathbf{v}^{(\nu)} \rightharpoonup \mathbf{v} \in \mathbf{V}$. Because of $\mathbf{v}^{(\nu)} \in \mathcal{H}_\mathfrak{r}$ we know that $\dim(\mathbf{U}_\alpha^{\min}(\mathbf{v}^{(\nu)})) \le r_j$. By Theorem 6.24, $\dim(\mathbf{U}_\alpha^{\min}(\mathbf{v})) \le r_j$ follows. Also the nestedness property $\mathbf{U}_\alpha^{\min}(\mathbf{v}) \subset \mathbf{U}_{\alpha_1}^{\min}(\mathbf{v}) \otimes \mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})$ follows from (6.13). Hence $\mathbf{v} \in \mathcal{H}_\mathfrak{r}$ is proved. □

The minimisation problem for $\mathcal{H}_\mathfrak{r}$ reads as follows:

$$\boxed{\begin{array}{l}\text{Given } \mathbf{v} \in {}_{\|\cdot\|}\bigotimes_{j=1}^d V_j \text{ and } \mathfrak{r} = (r_\alpha)_{\alpha\in T_D} \in \mathbb{N}^{T_D},\\ \text{determine } \mathbf{u} \in \mathcal{H}_\mathfrak{r} \text{ minimising } \|\mathbf{v} - \mathbf{u}\|.\end{array}} \tag{11.55}$$

The supposition of the next theorem is, in particular, satisfied if $\dim(V_j) < \infty$.

Theorem 11.56. *Suppose that the Banach space* $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ *is reflexive with a norm not weaker than* $\|\cdot\|_{\vee}$. *Then Problem (11.55) has a solution for all* $\mathbf{v} \in \mathbf{V}$, *i.e., for given representation ranks* $\mathfrak{r} = (r_\alpha)_{\alpha \in T_D}$, *there is a tensor* $\mathbf{u}_{\text{best}} \in \mathcal{H}_{\mathfrak{r}} \subset \mathbf{V}$ *which solves*

$$\|\mathbf{v} - \mathbf{u}_{\text{best}}\| = \inf_{\mathbf{u} \in \mathcal{H}_{\mathfrak{r}}} \|\mathbf{v} - \mathbf{u}\|.$$

Proof. By Lemma 11.55, $\mathcal{H}_{\mathfrak{r}}$ is weakly closed. Thus, Theorem 4.28 proves the existence of a minimiser. □

In the rest of this chapter we assume that $\mathbf{V}$ is a Hilbert tensor space with induced scalar product. Since $\mathbf{u}_{\text{best}} \in \bigotimes_{j=1}^{d} U_j^{\min}(\mathbf{v})$, the statements from the second part of Lemma 10.7 are still valid. These are also true for the truncation results from §11.4.2.

11.4.1.2 ALS Method

As in the case of the tensor subspace format, one can improve the approximation iteratively. Each iteration contains a loop over all vertices of $T_D \backslash \{D\}$. The action at $\alpha \in T_D \backslash \{D\}$ is in principle as follows. Let $\mathbf{v} \in \mathbf{V}$ be the given tensor and $\mathbf{u} = \rho_{\text{HTR}}^{\text{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big)$ the present representation in $\mathcal{H}_{\mathfrak{r}}$.

If $\alpha = \{k\} \in \mathcal{L}(T_D)$, $\mathbf{u}$ is replaced by $\mathbf{u}_{\text{new}}$ which is the minimiser of

$$\left\{ \left\| \mathbf{v} - \rho_{\text{HTR}}^{\text{orth}}\big(T_D, (\mathbf{C}_\alpha), c^{(D)}, (B_j)_{j \in D}\big) \right\| : B_k \in \mathbb{K}^{n_k \times r_k} \text{ with } B_k^{\mathsf{H}} B_k = I \right\},$$

i.e., a new r_k-dimensional subspace $U_k^{\text{new}} \subset V_k$ is optimally chosen and represented by an orthonormal basis B_k^{new}. Replacing the previous basis B_k from $\mathbf{u} = \rho_{\text{HTR}}^{\text{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big)$ by B_k^{new}, we obtain the representation of $\mathbf{u}_{\text{new}}$. Note that the change from B_k to B_k^{new} corresponds to a unitary mapping A_k. By Proposition 11.34, $\mathbf{u}_{\text{new}} = (A_k \otimes A_{[k]})\mathbf{u}$ holds with $A_{[k]} = I$.

If $\alpha \in T_D \backslash \mathcal{L}(T_D)$, a new r_α-dimensional subspace $\mathbf{U}_\alpha^{\text{new}} \subset \mathbf{V}_\alpha$ is to be determined. Since $\mathbf{U}_\alpha = \text{range}(\mathbf{B}_\alpha)$, one has in principle to minimise over all $\mathbf{B}_\alpha \in (\mathbf{V}_\alpha)^{r_\alpha}$ with $\mathbf{B}_\alpha^{\mathsf{H}} \mathbf{B}_\alpha = \mathbf{I}$. Since the basis $\mathbf{B}_\alpha$ does not appear explicitly in the representation $\rho_{\text{HTR}}^{\text{orth}}(\ldots)$, one has to use $\mathbf{C}_\alpha$ instead. Note that $\mathbf{B}_\alpha^{\mathsf{H}} \mathbf{B}_\alpha = \mathbf{I}$ holds if and only if $G(\mathbf{C}_\alpha) = I$ holds for the Gram matrix $G(\mathbf{C}_\alpha)$ (cf. (11.48b)). Hence, the previous approximation $\mathbf{u} = \rho_{\text{HTR}}^{\text{orth}}\big(T_D, (\mathbf{C}_\beta)_{\beta \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big)$ is replaced by $\mathbf{u}_{\text{new}}$ which is the minimiser of

$$\left\{ \left\| \mathbf{v} - \rho_{\text{HTR}}^{\text{orth}}\big(T_D, (\mathbf{C}_\beta)_{\beta \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big) \right\| : \mathbf{C}_\alpha \text{ with } G(\mathbf{C}_\alpha) = I \right\}.$$

$\mathbf{C}_\alpha^{\text{new}} = \big(C_{\text{new}}^{(\alpha,\ell)}\big)_{\ell=1}^{r_\alpha}$ defines the basis $\mathbf{B}_\alpha^{\text{new}} = [\mathbf{b}_{1,\text{new}}^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha,\text{new}}^{(\alpha)}]$ by (11.24). If $A_\alpha : \text{range}(\mathbf{B}_\alpha) \to \text{range}(\mathbf{B}_\alpha^{\text{new}})$ with $A_\alpha \mathbf{b}_i^{(\alpha)} = \mathbf{b}_{i,\text{new}}^{(\alpha)}$ is a unitary mapping, the transition from $\mathbf{C}_\alpha$ to $\mathbf{C}_\alpha^{\text{new}}$ produces $\mathbf{u}_{\text{new}} = (A_\alpha \otimes A_{\alpha^c})\mathbf{u}$ with $A_{\alpha^c} = I$ (see second part of §11.3.2.4).

11.4.2 HOSVD Truncation to $\mathcal{H}_\mathfrak{r}$

As for the tensor subspace format in §10.1, the higher order singular value decomposition can be used to project a tensor into $\mathcal{H}_\mathfrak{r}$. In the case of the tensor subspace format $\mathcal{T}_\mathbf{r}$ we have discussed two versions: (i) independent HOSVD projections in all d directions (§10.1.1) and (ii) a successive version with renewed HOSVD after each partial step in §10.1.2.

In the case of the hierarchical format, a uniform version (i) will be discussed in §11.4.2.1. However, now the successive variant (ii) splits into two versions (iia) and (iib). The reason is that not only different directions exists but also different levels of the vertices. Variant (iia) in §11.4.2.2 uses the direction root-to-leaves, while variant (iib) in §11.4.2.3 proceeds from the leaves to the root.

11.4.2.1 Basic Form

We assume that $\mathbf{v}$ is represented in $\mathcal{H}_\mathfrak{s}$ and should be truncated into a representation in $\mathcal{H}_\mathfrak{r}$ for a fixed rank vector $\mathfrak{r} \leq \mathfrak{s}$. More precisely, the following compatibility conditions are assumed:

$$\begin{array}{ll} r_\alpha \leq s_\alpha & \text{for all } \alpha \in T_D, \\ r_D = 1, & r_{\sigma_1} = r_{\sigma_2} \text{ for the sons } \sigma_1, \sigma_2 \text{ of } D, \\ r_\alpha \leq r_{\alpha_1} r_{\alpha_2} & \text{for all } \alpha \in T_D \backslash \mathcal{L}(T_D) \text{ and } \{\alpha_1, \alpha_2\} = S(\alpha). \end{array} \tag{11.56}$$

The last inequality follows from the nestedness property (11.11c). The equation $r_{\sigma_1} = r_{\sigma_2}$ for the sons σ_1, σ_2 of D is due to the fact that for $\mathbf{V} = \mathbf{V}_{\sigma_1} \otimes \mathbf{V}_{\sigma_2}$ we have the matrix case: The minimal subspaces $\mathbf{U}_{\sigma_1}^{\min}(\mathbf{u})$ and $\mathbf{U}_{\sigma_2}^{\min}(\mathbf{u})$ of some approximation $\mathbf{u} \in \mathbf{V}$ have identical dimensions (cf. Corollary 6.6).

First we describe a truncation to $\mathbf{u}_{\text{HOSVD}} \in \mathcal{H}_\mathfrak{r}$ which is completely analogous to the truncation to $\mathcal{T}_\mathbf{r}$ discussed in Theorem 10.3 for the tensor subspace representation. Nevertheless, there is a slight difference. In the case of $\mathcal{T}_\mathbf{r}$, there are single projections P_j $(1 \leq j \leq d)$ from (10.5) and $\mathbf{u}_{\text{HOSVD}} = \mathbf{P}_\mathbf{r}\mathbf{v}$ holds for the product $\mathbf{P}_\mathbf{r}$ of all P_j. Since the P_j commute, the ordering of the P_j in the product does not matter. In the hierarchical case, one has to take into consideration that not all projections commute.

In the case of $\mathcal{H}_\mathfrak{s}$, let $\mathbf{B}_\alpha^{\text{HOSVD}} = [\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{s_\alpha}^{(\alpha)}]$ be the HOSVD basis discussed in §11.3.3. Denote the reduction to the first r_α bases vectors (assuming $r_\alpha \leq s_\alpha$) by $\mathbf{B}_\alpha^{\text{red}} := [\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}]$. The projection P_α from $\mathbf{V}_\alpha$ onto $\text{range}(\mathbf{B}_\alpha^{\text{red}})$ is determined by $P_\alpha = \mathbf{B}_\alpha^{\text{red}}(\mathbf{B}_\alpha^{\text{red}})^\mathsf{H}$. The same symbol P_α denotes its extension to $\mathbf{V}$ via

$$P_\alpha := P_\alpha \otimes id_{\alpha^c}, \text{ i.e., } P_\alpha\Big(\bigotimes\nolimits_{j\in D} u_j\Big) = \Big[P_\alpha\Big(\bigotimes\nolimits_{j\in\alpha} u_j\Big)\Big] \otimes \Big[\bigotimes\nolimits_{j\in\alpha^c} u_j\Big]$$

for any $u_j \in V_j$ (cf. (3.39a,b)).

Given $\mathbf{v} = \rho_{\text{HTR}}^{\text{HOSVD}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\big)$ (cf. Definition 11.36), the projection P_α must be expressed by means of the coefficient matrices $\mathbf{C}_\alpha$, as detailed in the following procedure with $r = r_\alpha$:

```
procedure REDUCE$(\alpha, r)$;
1) delete $C^{(\alpha,\ell)}$ for $\ell > r$ ;
2) if $level(\alpha) \geq 1$ then for the father $\beta$ do
  2a) if $\alpha = \beta_1$ is the first son of $\beta$ then
      reduce $C^{(\beta,\ell)} \in \mathbb{K}^{s_1 \times s_2}$ to $\bigl(c_{i,j}^{(\beta,\ell)}\bigr)_{1\leq i\leq r,\, 1\leq j\leq s_2} \in \mathbb{K}^{r \times s_2}$ else
  2b) if $\alpha = \beta_2$ is the second son of $\beta$ then
      reduce $C^{(\beta,\ell)} \in \mathbb{K}^{s_1 \times s_2}$ to $\bigl(c_{i,j}^{(\beta,\ell)}\bigr)_{1\leq i\leq s_1,\, 1\leq j\leq r} \in \mathbb{K}^{s_1 \times r}$.
```

Concerning the ordering of the sons see Remark 11.4. $s_1 \times s_2$ denotes the actual size of $C^{(\beta,\ell)}$ before reduction. Note that the remaining coefficients $c_{i,j}^{(\alpha,\ell)}$, $c_{i,j}^{(\beta,\ell)}$ are unchanged, only those referring to $\{\mathbf{b}_\ell^{(\alpha)} : r+1 \leq \ell \leq s_\alpha\}$ are deleted by reducing the size of $\mathbf{C}_\alpha$, $\mathbf{C}_\beta$.

The recursive version of this procedure is

```
procedure REDUCE*$(\alpha, \mathfrak{r})$;
begin REDUCE$(\alpha, r_\alpha)$;
      if $\alpha \notin \mathcal{L}(T_D)$ then for all $\sigma \in S(\alpha)$ do REDUCE*$(\sigma, \mathfrak{r})$
end;
```

$\mathbf{REDUCE}^*(D, \mathfrak{r})$ maps $\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{HOSVD}}(\cdot) \in \mathcal{H}_{\mathfrak{s}}$ into $\mathbf{u}_{\mathrm{HOSVD}} = \mathbf{P}_{\mathfrak{r}}\mathbf{v} \in \mathcal{H}_{\mathfrak{r}}$. Since no arithmetical operations are performed, the computational cost of the procedures $\mathbf{REDUCE}$ and $\mathbf{REDUCE}^*$ is zero.

Remark 11.57. (a) If $\alpha, \beta \in T_D$ are disjoint, the projections commute: $P_\alpha P_\beta = P_\beta P_\alpha$. Otherwise, $P_\alpha P_\beta$ and $P_\beta P_\alpha$ may differ.
(b) If P_α, P_β are orthogonal projections onto subspaces U_α, U_β with $U_\alpha \subset U_\beta$, the identity $P_\alpha = P_\alpha P_\beta = P_\beta P_\alpha$ is valid.
(c) Let σ_1 and σ_2 be the sons of $D \in T_D$. The projections P_{σ_1} and P_{σ_2} (as mappings on $\mathbf{V}$) satisfy $P_{\sigma_1}\mathbf{v} = P_{\sigma_2}\mathbf{v}$.

Proof. For Part (c) note that $\mathbf{v} = \sum_\nu \sigma_\nu \, \mathbf{b}_\nu^{(\sigma_1)} \otimes \mathbf{b}_\nu^{(\sigma_2)}$. □

To study the consequence of Remark 11.57a, we consider $\alpha \in T_D \backslash \mathcal{L}(T_D)$ and its sons α_1 and α_2. While P_{α_1} and P_{α_2} commute, they do not commute with P_α, which can be seen as follow. First we describe the action of P_α. A tensor $\mathbf{v}_\alpha \in \mathbf{U}_\alpha$ has the representation

$$\mathbf{v}_\alpha = \sum_{\ell=1}^{s_\alpha} c_\ell^{(\alpha)} \mathbf{b}_\ell^{(\alpha)}$$

with $\mathbf{b}_\ell^{(\alpha)}$ from the HOSVD basis $\mathbf{B}_\alpha^{\mathrm{HOSVD}} = [\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{s_\alpha}^{(\alpha)}]$. Hence, the projection yields

$$P_\alpha \mathbf{v}_\alpha = \sum_{\ell=1}^{r_\alpha} c_\ell^{(\alpha)} \mathbf{b}_\ell^{(\alpha)} = \sum_{\ell=1}^{r_\alpha} c_\ell^{(\alpha)} \sum_{i=1}^{s_{\alpha_1}} \sum_{j=1}^{s_{\alpha_2}} c_{i,j}^{(\alpha,\ell)} \, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)},$$

where we use the representation (11.24) of $\mathbf{b}_\ell^{(\alpha)}$ (index bound s_α replaced by r_α!). The projections P_{α_1} and P_{α_2} produce

$$P_{\alpha_2} P_{\alpha_1} P_\alpha \mathbf{v}_\alpha = \sum_{\ell=1}^{r_\alpha} c_\ell^{(\alpha)} \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{i,j}^{(\alpha,\ell)} \, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} \tag{11.57}$$

(summation up to r_{α_i} instead of s_{α_i}). Now, $P_{\alpha_2} P_{\alpha_1} P_\alpha \mathbf{v}_\alpha$ belongs to the subspace $\tilde{\mathbf{U}}_\alpha := \mathrm{span}\{\tilde{\mathbf{b}}_\ell^{(\alpha)} : 1 \leq \ell \leq r_\alpha\}$ with the modified vectors

$$\tilde{\mathbf{b}}_\ell^{(\alpha)} := \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{i,j}^{(\alpha,\ell)} \, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}. \tag{11.58}$$

$\tilde{\mathbf{U}}_\alpha$ is a subspace of $\mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ for $\mathbf{U}_{\alpha_i} := \mathrm{span}\{\mathbf{b}_\ell^{(\alpha_i)} : 1 \leq \ell \leq r_{\alpha_i}\}$, i.e., the nestedness property (11.11c) holds.

On the other hand, if we first apply $P_{\alpha_2} P_{\alpha_1}$, we get

$$P_{\alpha_2} P_{\alpha_1} \mathbf{v}_\alpha = \sum_{\ell=1}^{s_\alpha} c_\ell^{(\alpha)} \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{i,j}^{(\alpha,\ell)} \, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} = \sum_{\ell=1}^{s_\alpha} c_\ell^{(\alpha)} \tilde{\mathbf{b}}_\ell^{(\alpha)}$$

with the modified vectors $\tilde{\mathbf{b}}_\ell^{(\alpha)}$ from (11.58). The next projection P_α yields some vector $P_\alpha P_{\alpha_2} P_{\alpha_1} \mathbf{v}_\alpha$ in $\mathbf{U}_\alpha := \mathrm{range}(\mathbf{B}_\alpha) = \mathrm{span}\{\mathbf{b}_\ell^{(\alpha)} : 1 \leq \ell \leq r_\alpha\} \neq \tilde{\mathbf{U}}_\alpha$. Therefore, in general,

$$P_{\alpha_2} P_{\alpha_1} P_\alpha \mathbf{v}_\alpha \neq P_\alpha P_{\alpha_2} P_{\alpha_1} \mathbf{v}_\alpha$$

holds proving noncommutativity. Further, $\mathbf{U}_\alpha$ is *not* a subspace of $\mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$, i.e., the construction does not satisfy the nestedness property (11.11c).

These considerations show that the HOSVD projections have to be applied from the *root to the leaves*. A possible description is as follows. For all level numbers $1 \leq \ell \leq L := \mathit{depth}(T_D)$ (cf. (11.7)) we set[19]

$$P^{(\ell)} := \prod\nolimits_{\alpha \in T_D,\, \mathit{level}(\alpha) = \ell} P_\alpha. \tag{11.59}$$

Since all P_α with $\mathit{level}(\alpha) = \ell$ commute (cf. Remark 11.57a), the ordering in the product does not matter and $P^{(\ell)}$ itself is a projection. Then we apply these projections in the order

$$\mathbf{u}_{\mathrm{HOSVD}} := P^{(L)} P^{(L-1)} \cdots P^{(2)} P^{(1)} \mathbf{v}. \tag{11.60}$$

The following theorem is the analogue of Theorem 10.3 for the tensor subspace representation. Again, we refer to the best approximation $\mathbf{u}_{\mathrm{best}} \in \mathcal{H}_{\mathfrak{r}}$, which exists as stated in Theorem 11.56.

Theorem 11.58. *Let $\mathbf{V} = {}_a\bigotimes_{j \in D} V_j$ and V_j be pre-Hilbert spaces with induced scalar product*[20]. *For $\mathbf{v} \in \mathcal{H}_{\mathfrak{s}}$ and $\mathfrak{r} \leq \mathfrak{s}$ satisfying (11.56) the approximation $\mathbf{u}_{\mathrm{HOSVD}} \in \mathcal{H}_{\mathfrak{r}}$ from (11.60) is quasi-optimal:*

$$\|\mathbf{v} - \mathbf{u}_{\mathrm{HOSVD}}\| \leq \sqrt{\sum_\alpha \sum_{i \geq r_\alpha + 1} (\sigma_i^{(\alpha)})^2} \leq \sqrt{2d-3}\, \|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\|. \tag{11.61}$$

$\sigma_i^{(\alpha)}$ are the singular values of $\mathcal{M}_\alpha(\mathbf{v})$. The sum $\sum_\alpha$ is taken over all $\alpha \in T_D \backslash \{D\}$ except that only one son σ_1 of D is involved.

[19] At level $\ell = 0$ no projection P_0 is needed because of (11.12), which holds for HOSVD bases.

[20] The induced scalar product is also used for $\mathbf{V}_\alpha$, $\alpha \in T_D \backslash L(T_D)$.

Proof. The vertex $\alpha = D$ is exceptional, since $P^{(1)} = \prod_{\alpha\in T_D,\, level(\alpha)=1} P_\alpha = P_{\sigma_2}P_{\sigma_1}$ (σ_1, σ_2 sons of D) can be replaced by P_{σ_1} alone (cf. Remark 11.57c). Since $\mathbf{u}_{\text{HOSVD}} := P^{(L)}\cdots P^{(2)}P_{\sigma_1}\mathbf{v}$, Lemma 4.123b yields

$$\left\|\mathbf{v} - \mathbf{u}_{\text{HOSVD}}\right\|^2 \le \sum\nolimits_\alpha \left\|(I - P_\alpha)\,\mathbf{v}\right\|^2.$$

The number of projections P_α involved is $2d-3$. The last estimate in

$$\left\|(I - P_\alpha)\,\mathbf{v}\right\|^2 = \sum\nolimits_{i\ge r_\alpha+1} (\sigma_i^{(\alpha)})^2 \le \left\|\mathbf{v} - \mathbf{u}_{\text{best}}\right\|^2 \tag{11.62}$$

follows as in the proof of Theorem 10.3. Summation of (11.62) over α proves (11.61). □

For all α with $r_\alpha = r_{\alpha_1}r_{\alpha_2}$ or $r_\alpha = s_\alpha$, the projections P_α may be omitted. This improves the error bound.

The practical performance of (11.60) is already illustrated by (11.57).

Proposition 11.59. *The practical performance of the HOSVD projection (11.60) is done in three steps:*

1) Install HOSVD bases at all vertices as described in §11.3.3. For an orthonormal basis this is achieved by $\mathbf{HOSVD}^*(D)$ *from (11.46b).*

2) Delete the basis vectors $\mathbf{b}_i^{(\alpha)}$ *with* $r_\alpha < i \le s_\alpha$*. Practically this means that the coefficient matrices* $C^{(\alpha,\ell)} \in \mathbb{K}^{s_{\sigma_1}\times s_{\sigma_2}}$ *for* $\ell > r_\alpha$ *are deleted, whereas those* $C^{(\alpha,\ell)}$ *with* $1 \le \ell \le r_\alpha$ *are reduced to matrices of the size* $\mathbb{K}^{r_{\alpha_1}\times r_{\alpha_2}}$*. This is performed by* $\mathbf{REDUCE}^*(D, \mathfrak{r})$*.*

3) Finally, $\mathbf{u}_{\text{HOSVD}}$ *is represented by* $\rho_{\text{HTR}}^{\text{HOSVD}}\big(T_D, (\tilde{\mathbf{C}}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (\tilde{B}_j)_{j\in D}\big)$ *referring to the bases* $\tilde{\mathbf{B}}_\alpha = [\tilde{\mathbf{b}}_1^{(\alpha)}, \ldots, \tilde{\mathbf{b}}_{r_\alpha}^{(\alpha)}]$ *generated recursively from* $\tilde{B}_j$ *and* $\tilde{\mathbf{C}}_\alpha$*:*

$$\begin{aligned}
\tilde{\mathbf{b}}_i^{(\alpha)} &:= \mathbf{b}_i^{(\alpha)} && \text{for } \alpha \in \mathcal{L}(T_D) \text{ and } 1 \le i \le r_\alpha,\\
\tilde{\mathbf{b}}_\ell^{(\alpha)} &:= \textstyle\sum_{i=1}^{r_{\alpha_1}}\sum_{j=1}^{r_{\alpha_2}} c_{i,j}^{(\alpha,\ell)}\tilde{\mathbf{b}}_i^{(\alpha_1)} \otimes \tilde{\mathbf{b}}_j^{(\alpha_2)} && \text{for } \alpha \in T_D\backslash\mathcal{L}(T_D) \text{ and } 1 \le \ell \le r_\alpha.
\end{aligned}$$

Note that these bases are suborthonormal, not orthonormal. *To reinstall orthonormality, the orthonormalisation procedure from §11.3.2 must be applied.*

According to Remark 11.40, the cost of Step 1) is about $(10r^4 + 2r^2n)d$, while Steps 2) and 3) are free. The cost for a possible re-orthonormalisation is discussed in Remark 11.32.

11.4.2.2 Sequential Truncation

In the case of the tensor subspace representation, a sequential truncation is formulated in (10.7). Similarly, the previous algorithm can be modified. In the algorithm from Proposition 11.59 the generation of the HOSVD bases in Step 1 is completed before the truncation starts in Step 2. Now, both parts are interweaved. Again, we want to truncate from $\mathcal{H}_\mathfrak{s}$ to $\mathcal{H}_\mathfrak{r}$, where $\mathfrak{s} \ge \mathfrak{r}$.

The following loop is performed from the root to the leaves:

1) **Start**: Tensor given in orthonormal hierarchical representation. Set $\alpha := D$.

2) **Loop**: a) If $\alpha \notin \mathcal{L}(T_D)$, compute the HOSVD bases $\tilde{\mathbf{B}}_{\alpha_1}$ and $\tilde{\mathbf{B}}_{\alpha_2}$ for the sons α_1 and α_2 of α by $\mathbf{HOSVD}(\alpha)$.

b) Restrict the bases at the vertices α_i to the first r_{α_i} vectors, i.e., call the procedures $\mathbf{REDUCE}(\alpha_1, r_{\alpha_1})$ and $\mathbf{REDUCE}(\alpha_2, r_{\alpha_2})$.

c) As long as the sons satisfy $\alpha_i \notin \mathcal{L}(T_D)$ repeat the loop for $\alpha := \alpha_i$.

Let the tensor $\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j \in D}\big) \in \mathcal{H}_\mathfrak{s}$ be given by an orthonormal hierarchical representation. The call

$$\mathbf{HOSVD\text{-}TrSeq}(D, \mathfrak{r})$$

yields the truncated tensor $\tilde{\mathbf{v}} = \rho_{\mathrm{HTR}}\big(T_D, (\tilde{\mathbf{C}}_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, c^{(D)}, (\tilde{B}_j)_{j \in D}\big) \in \mathcal{H}_\mathfrak{r}$:

```
procedure $\mathbf{HOSVD\text{-}TrSeq}(\alpha, \mathfrak{r})$;
if $\alpha \notin \mathcal{L}(T_D)$ then
begin $\mathbf{HOSVD}(\alpha)$; let $\alpha_1, \alpha_2 \in S(\alpha)$;
      $\mathbf{REDUCE}(\alpha_1, r_{\alpha_1})$; $\mathbf{REDUCE}(\alpha_2, r_{\alpha_2})$;
      $\mathbf{HOSVD\text{-}TrSeq}(\alpha_1, \mathfrak{r})$; $\mathbf{HOSVD\text{-}TrSeq}(\alpha_2, \mathfrak{r})$
end;
```

(11.63)

Remark 11.60. (a) Note that the order by which the vertices are visited is not completely fixed. The only restriction is the root-to-leaves direction. This fact enables a *parallel computation*. The result does not depend on the choice of the ordering. In particular, computations at the vertices of a fixed level can be performed in parallel. This reduces the factor d in the computational work to $\log_2 d$. Another saving of the computational work is caused by the fact that $\tilde{\mathbf{C}}_\alpha$ has smaller data size than $\mathbf{C}_\alpha$.
(b) When the HOSVD basis $\tilde{\mathbf{B}}_{\alpha_1}$ is created, this is by definition an orthonormal basis. If, however, the computation proceeds at $\alpha := \alpha_1$, the truncation of the basis at the son vertex of α_1 destroys orthonormality. As in the basic version, an orthonormal basis may be restored afterwards. However, even without re-orthonormalisation the sensitivity analysis from Theorem 11.53 guarantees stability for the resulting suborthonormal bases.

Below we use the sets $T_D^{(\ell)}$ defined in (11.8) and $L = depth(T_D)$ (cf. (11.7)).

Theorem 11.61. *The algorithm* $\mathbf{HOSVD\text{-}TrSeq}(D, \mathfrak{r})$ *yields a final approximation* $\mathbf{u}_\mathfrak{r} \in \mathcal{H}_\mathfrak{r}$ *with*

$$\begin{aligned} \|\mathbf{v} - \mathbf{u}_\mathfrak{r}\| &\le \sum_{\ell=1}^{L} \sqrt{\sum\nolimits_{\alpha \in T_D^{(\ell)}} \sum\nolimits_{i \ge r_\alpha + 1} \big(\tilde{\sigma}_i^{(\alpha)}\big)^2} \\ &\le \left[1 + \sum_{\ell=2}^{L} \sqrt{\# T_D^{(\ell)}}\right] \|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\|, \end{aligned} \tag{11.64}$$

where $\tilde{\sigma}_i^{(\alpha)}$ *are the singular values computed during the algorithm. The sum* $\sum_\alpha$ *is understood as in Theorem 11.58, i.e., at level* $\ell = 1$ *only one son of* D *is involved.*

Proof. 1) When bases at $\alpha_1, \alpha_2 \in S(\alpha)$ are computed, they are orthonormal, i.e., $G(\mathbf{B}_{\alpha_i}) = I$ holds for the Gram matrix. All later changes are applications of projections. Thanks to Exercise 11.48, $G(\mathbf{B}_{\alpha_i}) \le I$ holds for all modifications of the basis in the course of the algorithm. This is important for the later application of Theorem 11.53.

2) Algorithm $\mathbf{HOSVD\text{-}TrSeq}(D, \mathfrak{r})$ starts at level $\ell = 0$ and reduces recursively the bases at the son vertices at the levels $\ell = 1$ to $L = depth(T_D)$. $\mathbf{v}^0 := \mathbf{v}$ is the starting value. Let $\mathbf{v}^\ell$ denote the result after the computations at level ℓ. The final result is $\mathbf{u}_\mathfrak{r} := \mathbf{v}^L$. The standard triangle inequality yields

$$\|\mathbf{v} - \mathbf{u}_\mathfrak{r}\| \le \sum_{\ell=1}^{L} \|\mathbf{v}^\ell - \mathbf{v}^{\ell-1}\|.$$

As in (11.59), we define the product $\tilde{P}^{(\ell)} := \prod_{\alpha \in T_D^{(\ell)}} \tilde{P}_\alpha$, but now $\tilde{P}_\alpha$ describes the orthogonal projection onto $\mathrm{span}\{\tilde{\mathbf{b}}_i^{(\alpha)} : 1 \le i \le r_\alpha\}$, where $\tilde{\mathbf{b}}_i^{(\alpha)}$ are the (no more orthonormal) basis vectors computed by the present algorithm. We observe that $\mathbf{v}^\ell = \tilde{P}^{(\ell)} \mathbf{v}^{\ell-1}$. Lemma 4.123b allows us to estimate by

$$\|\mathbf{v}^\ell - \mathbf{v}^{\ell-1}\|^2 = \|\big(I - \tilde{P}^{(\ell)}\big)\mathbf{v}^{\ell-1}\|^2 \le \sum_{\alpha \in T_D^{(\ell)}} \|\big(I - \tilde{P}_\alpha\big)\mathbf{v}^{\ell-1}\|^2.$$

Theorem 11.53 states that

$$\|\big(I - \tilde{P}_\alpha\big)\mathbf{v}^{\ell-1}\|^2 \le \sum_{i \ge r_\alpha + 1} \big(\tilde{\sigma}_i^{(\alpha)}\big)^2, \tag{11.65}$$

since the perturbations are $\delta\mathbf{b}_i^{(\alpha)} = 0$ for $1 \le i \le r_\alpha$, but $\delta\mathbf{b}_i^{(\alpha)} = -\mathbf{b}_i^{(\alpha)}$ with $\|\mathbf{b}_i^{(\alpha)}\| = 1$ for $i \ge r_\alpha + 1$. This proves the inequality in (11.64). Concerning $\ell = 1$, one uses again that $\tilde{P}_{\alpha_1}\tilde{P}_{\alpha_2} = \tilde{P}_{\alpha_1}$ $(\alpha_1, \alpha_2 \in S(D))$.

3) Next we prove that the involved singular values $\tilde{\sigma}_i^{(\alpha)}$ are not larger than those $\sigma_i^{(\alpha)}$ from the basic algorithm in Theorem 11.58. During the sequential process we visit each vertex $\alpha \in T_D$ and create an orthogonal basis $\mathbf{b}_i^{(\alpha)}$ $(1 \le i \le s_\alpha)$ by calling $\mathbf{HOSVD}(\alpha)$. For theoretical purpose, we choose coefficient matrices $\mathbf{C}_\alpha$ corresponding to these bases. The truncation process $\mathbf{v} \mapsto \ldots \mapsto \mathbf{v}' \mapsto \mathbf{v}'' \mapsto \ldots \mapsto \mathbf{u}_\mathfrak{r}$ starts with the original tensor $\mathbf{v}$ and ends with $\mathbf{u}_\mathfrak{r}$. Fix a vertex $\beta \in T_D$ and let $\mathbf{v}', \mathbf{v}''$ be the tensors before and after the truncation at β. Via $\mathcal{M}_\alpha(\mathbf{v}')$ and $\mathcal{M}_\alpha(\mathbf{v}'')$ we obtain singular values, for which $\sigma_i''^{(\alpha)} \le \sigma_i'^{(\alpha)}$ is to be proved for all $\alpha \subset \beta$. For these cases we apply (5.12c): $E_{\beta_1} = \sum_{i,j=1}^{r_{\beta_1}} e_{ij}^{(\beta)} C^{(\beta,i)} C^{(\beta,j)\mathsf{H}}$ and Corollary 5.15: the squared singular values are the eigenvalues of the matrices E_α. More precisely, we have E_α' and E_α'' corresponding to $\mathbf{v}'$ and $\mathbf{v}''$.

3a) Case $\alpha = \beta$. $E_\alpha' = \mathrm{diag}\{\sigma_1'^{(\alpha)}, \ldots, \sigma_{s_\alpha}'^{(\alpha)}\}^2$ holds because of the particular choice of basis $\{\mathbf{b}_i^{(\alpha)}\}$, while truncation yields $E_\alpha'' = \mathrm{diag}\{\sigma_1''^{(\alpha)}, \ldots, \sigma_{s_\alpha}''^{(\alpha)}\}^2$ with $\sigma_i''^{(\alpha)} = \sigma_i'^{(\alpha)}$ for $1 \le i \le r_\alpha$ and $\sigma_i''^{(\alpha)} = 0$ for $r_\alpha < i \le s_\alpha$. Hence, $E_\alpha'' \le E_\alpha'$ holds.

3b) Case $\alpha \subsetneqq \beta$. We apply induction in the subtree T_β (cf. Definition 11.6). It suffices to explain the case of α being the first son of β. Equation (5.12c) states that $E_\alpha = \sum_{i,j=1}^{r_\alpha} e_{ij}^{(\beta)} C^{(\beta,i)} C^{(\beta,j)\mathsf{H}}$ (note that $G_\bullet = I$ because of orthonormality). Using this identity for $E'_\bullet$ and $E''_\bullet$ instead of $E_\bullet$, the inequality $E''_\beta \le E'_\beta$ together with Lemma 2.15 proves $E''_\alpha \le E'_\alpha$ and, by Lemma 2.27a, $\sigma_i^{\prime\prime(\alpha)} \le \sigma_i^{\prime(\alpha)}$. This sequence of inequalities proves

$$\tilde{\sigma}_i^{(\alpha)} \le \sigma_i^{(\alpha)} \qquad \text{for } 1 \le i \le s_\alpha,\ \alpha \in T_\beta, \tag{11.66}$$

4) Thanks to (11.66), inequality (11.65) can be continued by the comparison $\sum_{i \ge r_\alpha+1} (\sigma_i^{(\alpha)})^2 \le \|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\|$ with the best approximation (cf. (11.62)). □

The sum $\sum_\ell \sqrt{\cdots}$ appears in (11.64), since the perturbations from the different levels are not orthogonal. We may estimate the error by $\sqrt{\sum_\alpha \|(I - P_\alpha)\,\mathbf{v}\|^2}$ as in the proof of Theorem 11.58, but now the HOSVD basis at vertex α is not related to the singular value decomposition of $\mathcal{M}_\alpha(\mathbf{v})$, so that we cannot continue like in the mentioned proof.

Remark 11.62. The factor $C(T_k) := 1 + \sum_{\ell=2}^{L} \sqrt{\#T_D^{(\ell)}}$ in (11.64) depends on the structure of T_k. If $d = 2^L$, the perfectly balanced tree T_D leads to

$$C(T_k) = 1 + \sum_{\ell=0}^{L-2} 2^{(L-\ell)/2} = \left(2 + \sqrt{2}\right)\sqrt{d} - 1 - 2\sqrt{2} = 3.4142\sqrt{d} - 3.8284\,.$$

For general d, the tree T_k with minimal depth $\lceil \log_2 d \rceil$ (see Remark 11.5a) yields

$$C(T_k) = \sqrt{d - 2^{L-1}} + \sum_{\ell=1}^{L-2} 2^{(L-\ell)/2} + 1 < 4.1213 \cdot 2^{L/2} - 3.8284\,.$$

The worst factor appears for the tree of maximal depth $L = d - 1$ (see Remark 11.5b), where

$$C(T_k) = 1 + 2\,(d-2)\,.$$

Remark 11.63. In principle, Algorithm (11.63) can be modified such that after each reduction step (call of **REDUCE**) the higher order SVD is updated. Then all contributions in $\sqrt{\sum_\alpha \|(I - P_\alpha)\,\mathbf{v}\|^2}$ can be estimated by $\|\mathbf{v} - \mathbf{u}_{\mathrm{best}}\|$ and we regain the estimate in (11.61).

11.4.2.3 Leaves-to-Root Truncation

As pointed out in §11.4.2.1, the projections P_α should not proceed from the leaves to the root, since then the nestedness property is violated. Grasedyck [73] proposes a truncation from the leaves to the root modified in such a way that nestedness is ensured. Let a tensor $\mathbf{v} \in \mathcal{H}_\mathfrak{s}$ be given. A truncation at the sons $\alpha_1, \alpha_1 \in S(\alpha)$ reduces the size of the coefficient matrices $C^{(\alpha,\ell)}$ so that the computational work at vertex α is reduced. We assume that the target format $\mathcal{H}_\mathfrak{r}$ satisfies (11.56). We use the notation $T_D^{(\ell)}$ from (11.8) and the abbreviation $L := depth(T_D)$. The leaves-to-root direction is reflected by the loop $L, L-1, \ldots, 1$ in the following algorithm:

Start: The starting value is $\mathbf{u}^{L+1} := \mathbf{v}$.

Loop from $\ell := L$ to 1: For all $\alpha \in T_D^{(\ell)}$ determine the HOSVD from $\mathcal{M}_\alpha(\mathbf{u}^{\ell+1})$. Let $\mathbf{U}_\alpha$ be spanned by the first r_α left singular vectors of $\mathcal{M}_\alpha(\mathbf{u}^{\ell+1})$ and define the HOSVD projection P_α as orthogonal projection onto $\mathbf{U}_\alpha$. Set $\mathbf{u}^\ell := P^{(\ell)}\mathbf{u}^{\ell+1}$, where $P^{(\ell)} := \prod_{\alpha \in T_D^{(\ell)}} P_\alpha$.

The fact that in each step the matricisation $\mathcal{M}_\alpha(\mathbf{u}^{\ell+1})$ uses the last projected tensor $\mathbf{u}^{\ell+1}$ is essential, since it guarantees that the left-sided singular value decomposition of $\mathcal{M}_\alpha(\mathbf{u}^{\ell+1})$ for $\alpha \in T_D^{(\ell)}$ leads to basis vectors $\mathbf{b}_i^{(\alpha)} \in \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ (α_1, α_2 sons of α). This ensures nestedness.

Theorem 11.64. *The algorithm described above yields a final approximation $\mathbf{u}^1 \in \mathcal{H}_\mathfrak{r}$ with*

$$\left\|\mathbf{v} - \mathbf{u}^1\right\| \le \sqrt{\sum_\alpha \sum_{i \ge r_\alpha + 1} \sigma_i^{(\alpha)}} \le \sqrt{2d-3}\,\|\mathbf{v} - \mathbf{u}_{\text{best}}\|\,.$$

The sum $\sum_\alpha$ is understood as in Theorem 11.58. The singular values $\sigma_i^{(\alpha)}$ are those of $\mathcal{M}_\alpha(\mathbf{u}^{\ell+1})$ with $\ell = level(\alpha)$.

Proof. We remark that $\mathbf{V} = {}_a\bigotimes_{\alpha \in T_D^{(\ell)} \text{ or } \alpha \in \mathcal{L}(T_D),\, level(\alpha) < \ell} \mathbf{V}_\alpha$. Consider all

$$\mathbf{u} \in \Big[\, {}_a\bigotimes\nolimits_{\alpha \in T_D^{(\ell)}} \mathbf{U}_\alpha \Big] \otimes \Big[\, {}_a\bigotimes\nolimits_{\alpha \in \mathcal{L}(T_D),\, level(\alpha) < \ell} \mathbf{V}_\alpha \Big]$$

with subspaces $\mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ satisfying $\dim(\mathbf{U}_\alpha) = r_\alpha$, provided that $\alpha \in T_D^{(\ell)}$ (note that $\mathbf{U}_{\alpha_1}, \mathbf{U}_{\alpha_2}$ are already fixed). Let $\mathbf{u}_{\text{best}}^\ell$ be the best approximation in $\inf \|\mathbf{u}^{\ell+1} - \mathbf{u}\| = \|\mathbf{u}^{\ell+1} - \mathbf{u}_{\text{best}}^\ell\|$. This tensor satisfies

$$\mathbf{u}_{\text{best}}^\ell \in \Big[\, {}_a\bigotimes\nolimits_{\alpha \in T_D^{(\ell)}} \mathbf{U}_\alpha^* \Big] \otimes \Big[\, {}_a\bigotimes\nolimits_{\alpha \in \mathcal{L}(T_D),\, level(\alpha) < \ell} \mathbf{V}_\alpha \Big]$$

for certain subspaces $\mathbf{U}_\alpha^*$, $\alpha \in T_D^{(\ell)}$. Again, the proof of Theorem 10.3 yields

$$\left\|\mathbf{u}^{\ell+1} - \mathbf{u}^\ell\right\| \le \sqrt{\#T_D^{(\ell)}}\,\left\|\mathbf{u}^{\ell+1} - \mathbf{u}_{\text{best}}^\ell\right\|$$

for $\mathbf{u}^\ell := P^{(\ell)}\mathbf{u}^{\ell+1}$, since $P^{(\ell)}$ is the tensor product of $\#T_D^{(\ell)}$ projections P_α. We have $\mathbf{u}_{\text{best}}^\ell = P_\ell^* \mathbf{u}^{\ell+1} = P_\ell^* P^{(\ell+1)} P^{(\ell+2)} \cdots P^{(L)} \mathbf{v}$ with P_ℓ^* being the product of the orthogonal projections onto $\mathbf{U}_\alpha^*$. The nestedness property together with Remark 11.57b shows that

$$P_\ell^* P^{(\ell+1)} P^{(\ell+2)} \cdots P^{(L)} = P_\ell^* = P^{(\ell+1)} P^{(\ell+2)} \cdots P^{(L)} P_\ell^*.$$

This implies that

$$\begin{aligned}\left\|\mathbf{u}^{\ell+1}-\mathbf{u}^{\ell}_{\text{best}}\right\| &= \left\|P^{(\ell+1)}P^{(\ell+2)}\cdots P^{(L)}\mathbf{v}-P^{(\ell+1)}P^{(\ell+2)}\cdots P^{(L)}P^{*}_{\ell}\mathbf{v}\right\| \\ &= \left\|P^{(\ell+1)}P^{(\ell+2)}\cdots P^{(L)}\left(\mathbf{v}-P^{*}_{\ell}\mathbf{v}\right)\right\| \le \left\|\mathbf{v}-P^{*}_{\ell}\mathbf{v}\right\| \le \left\|\mathbf{v}-\mathbf{u}_{\text{best}}\right\|.\end{aligned}$$

Together, we see that

$$\left\|\mathbf{u}^{\ell+1}-\mathbf{u}^{\ell}\right\| \le \sqrt{\#T^{(\ell)}_D}\,\left\|\mathbf{v}-\mathbf{u}_{\text{best}}\right\|.$$

We claim that $\mathbf{u}^{\ell+1}-\mathbf{u}^{\ell}\perp\mathbf{u}^{m+1}-\mathbf{u}^{m}$ for $\ell\neq m$. Without loss of generality assume $\ell>m$. We have $\mathbf{u}^{m+1}-\mathbf{u}^{m}=(I-P^{(m)})P^{(m+1)}\cdots P^{(\ell)}\cdots P^{(L)}\mathbf{v}$. Again, Remark 11.57b and the nestedness property prove that all projections $P^{(i)}$ are pairwise commuting; hence, $\mathbf{u}^{m+1}-\mathbf{u}^{m}=P^{(\ell)}(\mathbf{u}^{m+1}-\mathbf{u}^{m})$ is orthogonal to $\mathbf{u}^{\ell+1}-\mathbf{u}^{\ell}=(I-P^{(\ell)})\mathbf{u}^{\ell+1}$. Therefore, we can estimate as follows:

$$\begin{aligned}\left\|\mathbf{v}-\mathbf{u}^{0}\right\| &= \left\|\left(\mathbf{v}-\mathbf{u}^{L}\right)+\left(\mathbf{u}^{L}-\mathbf{u}^{L-1}\right)+\ldots+\left(\mathbf{u}^{2}-\mathbf{u}^{1}\right)\right\| \\ &\le \sqrt{\left\|\mathbf{v}-\mathbf{u}^{L}\right\|^{2}+\left\|\mathbf{u}^{L}-\mathbf{u}^{L-1}\right\|^{2}+\ldots+\left\|\mathbf{u}^{2}-\mathbf{u}^{1}\right\|^{2}} \\ &\le \sqrt{\#T^{(L)}_D+\ldots+\#T^{(1)}_D}\,\left\|\mathbf{v}-\mathbf{u}_{\text{best}}\right\|.\end{aligned}$$

Again, we can argue that at level $\ell=1$ one projection is sufficient:

$$P^{(1)}\mathbf{u}^{2}=P_{\alpha_1}P_{\alpha_2}\mathbf{u}^{2}=P_{\alpha_2}\mathbf{u}^{2}.$$

This reduces $\#T^{(1)}_D=2$ to 1. In total, $\sqrt{\cdots}=\sqrt{2d-3}$ proves the assertion of the theorem. □

11.4.2.4 Error Controlled Truncation

So far, we have prescribed a fixed $\mathfrak{r}$ for the truncation. Instead, one can prescribe a tolerance $\varepsilon>0$. Given $\mathbf{v}\in\mathcal{H}_\mathfrak{s}$, we want to find an approximation $\tilde{\mathbf{v}}\in\mathcal{H}_\mathfrak{r}$ with $\mathfrak{r}\le\mathfrak{s}$ such that $\|\mathbf{v}-\tilde{\mathbf{v}}\|\le\varepsilon$. The following heuristic strategies will yield an $\mathfrak{r}\le\mathfrak{s}$ such that for any smaller component than $\mathfrak{r}_\alpha$ the error bound by ε cannot be ensured. The theoretically optimal choice of $\mathfrak{r}$ and $\tilde{\mathbf{v}}\in\mathcal{H}_\mathfrak{r}$ would be the minimiser $(\mathfrak{r},\tilde{\mathbf{v}})\in\mathbb{N}^{T_D}\times\mathcal{H}_\mathfrak{r}$ of

$$\min\left\{\text{memory cost of }\tilde{\mathbf{v}}\in\mathcal{H}_\mathfrak{r}\text{ with }\mathfrak{r}\le\mathfrak{s},\ \|\mathbf{v}-\tilde{\mathbf{v}}\|\le\varepsilon\right\}.$$

The truncation from §11.4.2.1 allows the easiest realisation of an error controlled truncation. Given $\mathbf{v}=\rho^{\text{HOSVD}}_{\text{HTR}}(\ldots)\in\mathcal{H}_\mathfrak{s}$, all singular values $\sigma^{(\alpha)}_i$ ($\alpha\in T_D\backslash\mathcal{L}(T_D)$, $1\le i\le s_\alpha$) are available. We may order them by size:

$$\sigma_1\ge\sigma_2\ge\ldots\ge\sigma_\nu\ge\ldots\ge\sigma_{\nu_{\max}},\qquad \nu_{\max}=\sum_{\alpha\in T_D\backslash\mathcal{L}(T_D)}s_\alpha. \tag{11.67}$$

For all indices ν, there are $\alpha[\nu] \in T_D \backslash \mathcal{L}(T_D)$ and $i[\nu] \in [1, s_{\alpha[\nu]}]$ such that $\sigma_\nu = \sigma_{i[\nu]}^{(\alpha[\nu])}$. It is easy to choose a minimal ν_ε such that

$$\sum_{\nu=\nu_\varepsilon}^{\nu_{\max}} \sigma_\nu^2 \leq \varepsilon^2. \tag{11.68}$$

Then define $r_\alpha = \min\{i[\nu] - 1 : \nu_\varepsilon \leq \nu \leq \nu_{\max} \text{ with } \alpha[\nu] = \alpha\}$ (and $r_\alpha = s_\alpha$ if the latter set is empty). This defines a format $\mathcal{H}_\mathfrak{r}$. Truncation to $\mathcal{H}_\mathfrak{r}$ as in §11.4.2.1 leads us to $\tilde{\mathbf{v}} \in \mathcal{H}_\mathfrak{r}$ with $\|\mathbf{v} - \tilde{\mathbf{v}}\| \leq \varepsilon$ because of (11.61).

For the sequential version from §11.4.2.3, one has to split ε^2 into $\varepsilon^2 = \sum_{\ell=0}^{L-1} \varepsilon_\ell^2$. Then, in each level ℓ, one can proceed as before but with ε replaced by ε_ℓ. This means that all $\sigma_i^{(\alpha)}$ with $\alpha \in T_D^{(\ell)} \backslash \mathcal{L}(T_D)$ are ordered as in (11.67).

The sequential version from §11.4.2.2 requires the splitting $\varepsilon = \sum_{\ell=0}^{L-1} \varepsilon_\ell$.

In the approaches from above, we have maximised

$$\sum_{\alpha \in T_D \backslash \mathcal{L}(T_D)} (s_\alpha - r_\alpha),$$

which is the number of omitted basis vectors. If the gained storage is to be maximised, the strategy has to be refined as follows (taking the example of the truncation from §11.4.2.1). Omitting one basis vector at vertex α saves a storage of size[21] $N_{\mathrm{mem}}(\alpha) = s_{\alpha_1} s_{\alpha_2} + s_\beta s_{\alpha'}$, where $\{\alpha_1, \alpha_2\} = S(\alpha)$ and $\{\alpha, \alpha'\} = S(\beta)$, i.e., β is the father of α. The term $s_{\alpha_1} s_{\alpha_2}$ corresponds to the deleted matrix $C^{(\alpha, s_\alpha)}$, while $s_\beta s_{\alpha'}$ is related to the omitted rows $C_{s_\alpha, \cdot}^{(\beta, \cdot)}$ (if α is the first son of β). Instead of (11.67), one can order the quantities

$$\hat{\sigma}_\nu = \sigma_{i[\nu]}^{(\alpha[\nu])} / \sqrt{N_{\mathrm{mem}}(\alpha[\nu])}.$$

The indices $\nu_\varepsilon \leq \nu \leq \nu_{\max}$ are selected with minimal ν_ε subject to (11.68) (with σ_ν, not $\hat{\sigma}_\nu$).

11.5 Joining two Hierarchical Tensor Representation Systems

11.5.1 Setting of the Problem

We consider the following situation: $\mathbf{v}' \in \mathcal{H}_{\mathfrak{r}'}$ and $\mathbf{v}'' \in \mathcal{H}_{\mathfrak{r}''}$ are tensors involving two hierarchical systems related to $\mathbf{V} = \bigotimes_{j=1}^d V_j$ with a common dimension partition tree T_D, but different basis systems

$$(\mathbf{B}'_\alpha)_{\alpha \in \mathcal{L}(T_D)}, \; (\mathbf{B}''_\alpha)_{\alpha \in \mathcal{L}(T_D)}, \; (\mathbf{C}'_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)}, \; (\mathbf{C}''_\alpha)_{\alpha \in T_D \backslash \mathcal{L}(T_D)},$$

[21] At the beginning, the ranks s_α are defined by $\mathfrak{s}$ from $\mathcal{H}_\mathfrak{s}$. After each truncation step, one of the s_α is replaced by $s_\alpha - 1$. Therefore also the memory save may decrease.

and different subspaces $\{\mathbf{U}'_\alpha\}_{\alpha\in T_D}$ and $\{\mathbf{U}''_\alpha\}_{\alpha\in T_D}$.

We want to construct a hierarchical system $\mathcal{H}_\mathfrak{r}$ by means of subspaces $\{\mathbf{U}_\alpha\}_{\alpha\in T_D}$ defined by

$$\mathbf{U}_\alpha := \mathbf{U}'_\alpha + \mathbf{U}''_\alpha \qquad \text{for } \alpha \in T_D.$$

Then, $\mathcal{H}_\mathfrak{r}$ represents all tensors of $\mathbf{U}_D = \mathbf{U}'_D + \mathbf{U}''_D$, i.e., the tensors $\mathbf{v}', \mathbf{v}''$ from above belong to the new, common basis systems.

Each tensor in $\mathcal{H}_{\mathfrak{r}'}$ and $\mathcal{H}_{\mathfrak{r}''}$ is characterised by the respective coefficients $c'^{(D)}$ and $c''^{(D)}$. A subtask is to transform these coefficients into new ones referring to the new basis of $\mathbf{U}_D$.

Before we discuss the solution of this problem under various requirements on the bases, we ensure that the problem makes sense.

Remark 11.65. The subspaces $\{\mathbf{U}_\alpha\}_{\alpha\in T_D}$ defined above satisfy the nestedness condition:

$$\mathbf{U}'_\alpha + \mathbf{U}''_\alpha \subset \left(\mathbf{U}'_{\alpha_1} + \mathbf{U}''_{\alpha_1}\right) \otimes \left(\mathbf{U}'_{\alpha_2} + \mathbf{U}''_{\alpha_2}\right)$$

for $\alpha \in T_D \backslash \mathcal{L}(T_D)$ and $\alpha_1, \alpha_2 \in S(\alpha)$.

11.5.2 Trivial Joining of Frames

The least requirement is that $\mathbf{B}'_\alpha = [\mathbf{b}'^{(\alpha)}_1, \ldots, \mathbf{b}'^{(\alpha)}_{r'_\alpha}] \in (\mathbf{U}'_\alpha)^{r'_\alpha}$ and $\mathbf{B}''_\alpha = [\mathbf{b}''^{(\alpha)}_1, \ldots, \mathbf{b}''^{(\alpha)}_{r''_\alpha}] \in (\mathbf{U}''_\alpha)^{r''_\alpha}$ are frames spanning the respective subspaces $\mathbf{U}'_\alpha$ and $\mathbf{U}''_\alpha$. Since no linear independence is required, the simple definition

$$\mathbf{B}_\alpha := \left[\mathbf{b}'^{(\alpha)}_1, \mathbf{b}'^{(\alpha)}_2, \ldots, \mathbf{b}'^{(\alpha)}_{r'_\alpha}, \mathbf{b}''^{(\alpha)}_1, \mathbf{b}''^{(\alpha)}_2, \ldots, \mathbf{b}''^{(\alpha)}_{r''_\alpha}\right]$$

is a frame generating $\mathbf{U}_\alpha := \mathbf{U}'_\alpha + \mathbf{U}''_\alpha$. The drawback is that the cardinality $r_\alpha := r'_\alpha + r''_\alpha$ of the new frame is fully additive, even if the subspaces $\mathbf{U}'_\alpha$, $\mathbf{U}''_\alpha$ overlap.

An advantage is the easy construction of the coefficients. Consider a vertex $\alpha \in T_D \backslash \mathcal{L}(T_D)$ with sons α_1, α_2. The coefficient matrix $C'^{(\alpha,\ell)}$ representing $\mathbf{b}'^{(\alpha)}_\ell$ and the matrix $C''^{(\alpha,\ell)}$ representing $\mathbf{b}''^{(\alpha)}_\ell$ lead to the block diagonal matrix

$$C^{(\alpha,\ell)} := \begin{bmatrix} C'^{(\alpha,\ell)} & 0 \\ 0 & C''^{(\alpha,\ell)} \end{bmatrix}$$

representing the new columns of $\mathbf{B}_\alpha$ by those from $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$.

If $\mathbf{v}' \in \mathbf{U}'_D$ is coded by the coefficient vector $c'^{(D)} \in \mathbb{K}^{r'_D}$, the new coefficient is $\left[\begin{smallmatrix} c'^{(D)} \\ 0 \end{smallmatrix}\right]$, while $\mathbf{v}'' \in \mathbf{U}''_D$ coded by $c''^{(D)} \in \mathbb{K}^{r''_D}$ is expressed by the coefficient vector $\left[\begin{smallmatrix} 0 \\ c''^{(D)} \end{smallmatrix}\right]$.

Remark 11.66. The joining of frames requires only a rearrangement of data, but no arithmetical operation.

11.5.3 Common Bases

Let $\mathbf{B}'_\alpha$ and $\mathbf{B}''_\alpha$ carry the bases of the respective spaces $\mathbf{U}'_\alpha$ and $\mathbf{U}''_\alpha$. For the joint space $\mathbf{U}_\alpha = \mathbf{U}'_\alpha + \mathbf{U}''_\alpha$ we want to construct a new common basis $\mathbf{B}_\alpha$.

11.5.3.1 General or Orthonormal Case

The computation proceeds over all $\alpha \in T_D$ from level $depth(T_D)$ to 0. In principle, the two bases $\mathbf{B}'_\alpha$ and $\mathbf{B}''_\alpha$ are joined into one basis by means of the procedure $\mathbf{JoinBases}$ from (2.35). Here, we have to distinguish the case of leaves $\alpha \in \mathcal{L}(T_D)$ from inner vertices, since the bases are explicitly known for leaves only.

The procedure takes the following form, where $T_D^{(\lambda)}$ is defined in (11.8).

```
for $\lambda := depth(T_D)$ to 0 do
begin for all $\alpha \in T_D^{(\lambda)}$ do
    if $\alpha \in \mathcal{L}(T_D)$ then apply (11.70a-c) else
    begin apply (11.71a);
        if $\alpha \neq D$ then apply (11.70b,c)
        else apply (11.71b)
end  end;
```
(11.69)

Case of leaves. If $\alpha = \{j\} \in \mathcal{L}(T_D)$, the matrices $B'_j = [b_1'^{(j)}, \ldots, b_{r'_j}'^{(j)}]$ and $B''_j = [b_1''^{(j)}, \ldots, b_{r''_j}''^{(j)}]$ contain the explicitly available basis vectors. The call of

$$\mathbf{JoinBases}(B'_j, B''_j, r_j, B_j, T', T'') \tag{11.70a}$$

yields a basis $B_j = [b_1^{(j)}, \ldots, b_{r_j}^{(j)}]$ together with its cardinality r_j and transformation matrices $T' \in \mathbb{K}^{r_j \times r'_j}$, $T'' \in \mathbb{K}^{r_j \times r''_j}$ with the property $B'_j = B_j T'$ and $B''_j = B_j T''$ (cf. (2.34)). The basis change at $\alpha = \{j\}$ influences $C'^{(\beta,\ell)}$ and $C''^{(\beta,\ell)}$ for the father vertex β of α. Assume that $\alpha = \beta_1$ is the first son of β. Then $C_{\mathrm{new}}'^{(\beta,\ell)} = T^{(\beta_1)} C_{\mathrm{old}}'^{(\beta,\ell)} (T^{(\beta_2)})^{\mathsf{T}}$ from (11.32) holds with $T^{(\beta_1)} = T'$ and $T^{(\beta_2)} = I$, i.e., the coefficient matrix from $\mathcal{H}_{\mathfrak{r}'}$ becomes $C_{\mathrm{new}}'^{(\beta,\ell)} := T' C_{\mathrm{old}}'^{(\beta,\ell)}$. If $\alpha = \beta_2$ is the second son of β, then $C_{\mathrm{new}}'^{(\beta,\ell)} := C_{\mathrm{old}}'^{(\beta,\ell)} T'^{\mathsf{T}}$ holds. The coefficient matrices from $\mathcal{H}_{\mathfrak{r}''}$ are treated analogously:

$$C_{\mathrm{new}}'^{(\beta,\ell)} := T' C_{\mathrm{old}}'^{(\beta,\ell)} \quad \text{or} \quad C_{\mathrm{new}}'^{(\beta,\ell)} := C_{\mathrm{old}}'^{(\beta,\ell)} T'^{\mathsf{T}}, \tag{11.70b}$$

$$C_{\mathrm{new}}''^{(\beta,\ell)} := T'' C_{\mathrm{old}}''^{(\beta,\ell)} \quad \text{or} \quad C_{\mathrm{new}}''^{(\beta,\ell)} := C_{\mathrm{old}}''^{(\beta,\ell)} T''^{\mathsf{T}} \tag{11.70c}$$

depending on whether α is the first [second] son of β (left [right] identities).

Case of inner vertices. Let $\alpha \in T_D \backslash \mathcal{L}(T_D)$ and assume that by procedure (11.69) the sons α_1, α_2 of α have already common bases $\mathbf{B}_{\alpha_1}$ and $\mathbf{B}_{\alpha_2}$. Therefore, the coefficient matrices $C'^{(\alpha,\ell)} \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2}}$ of $\mathcal{H}_{\mathfrak{r}'}$ and $C''^{(\alpha,\ell)} \in \mathbb{K}^{r_{\alpha_1} \times r_{\alpha_2}}$ of

$\mathcal{H}_{\mathfrak{r}''}$ refer to these common bases (see (11.70b,c)). The matrices are gathered in $\mathbf{C}'_\alpha = \left(C'^{(\alpha,\ell)}\right)_{\ell=1}^{r'_\alpha}$ and $\mathbf{C}''_\alpha = \left(C''^{(\alpha,\ell)}\right)_{\ell=1}^{r''_\alpha}$. The matrices in $\mathbf{C}'_\alpha$ are linearly independent if and only if the represented vectors $\mathbf{b}'^{(\alpha)}_\ell = \sum_{i,j} c'^{(\alpha,\ell)}_{ij} \mathbf{b}^{(\alpha_1)}_i \otimes \mathbf{b}^{(\alpha_2)}_j$ are linearly independent. The call of

$$\mathbf{JoinBases}(\mathbf{C}'_\alpha, \mathbf{C}''_\alpha, r_\alpha, \mathbf{C}_\alpha, T', T'') \tag{11.71a}$$

yields the collection $\mathbf{C}_\alpha = (C^{(\alpha,\ell)})_{\ell=1}^{r_\alpha}$ of r_α linearly independent matrices $C^{(\alpha,\ell)}$ representing the new basis vectors $\mathbf{b}^{(\alpha)}_\ell = \sum_{i,j} c^{(\alpha,\ell)}_{i,j} \mathbf{b}^{(\alpha_1)}_i \otimes \mathbf{b}^{(\alpha_2)}_j$ $(1 \le \ell \le r_\alpha)$. The matrices T', T'' describe the relations $\mathbf{C}'_\alpha = \mathbf{C}_\alpha T'$ and $\mathbf{C}''_\alpha = \mathbf{C}_\alpha T''$, i.e.,

$$C'^{(\alpha,\ell)} = \sum_{k=1}^{r_\alpha} T'_{k\ell} C^{(\alpha,k)} \ (1 \le \ell \le r'_\alpha), \quad C''^{(\alpha,\ell)} = \sum_{k=1}^{r_\alpha} T''_{k\ell} C^{(\alpha,k)} \ (1 \le \ell \le r''_\alpha)$$

(cf. (2.34)), which are equivalent to $\mathbf{B}'_\alpha = \mathbf{B}_\alpha T'$ and $\mathbf{B}''_\alpha = \mathbf{B}_\alpha T''$. Hence, again the transformations (11.70b,c) are to be applied.

If $\alpha = D$, the coefficients $c'^{(D)}$ $[c''^{(D)}]$ of the tensors $\mathbf{v}' \in \mathcal{H}_{\mathfrak{r}'}$ $[\mathbf{v}'' \in \mathcal{H}_{\mathfrak{r}''}]$ are to be updated:

$$c'^{(D)}_{\text{new}} := T' c'^{(D)} \text{ for } \mathbf{v}' \in \mathcal{H}_{\mathfrak{r}'}, \quad c''^{(D)}_{\text{new}} := T'' c''^{(D)} \text{ for } \mathbf{v}'' \in \mathcal{H}_{\mathfrak{r}''} \tag{11.71b}$$

(cf. (11.34)). Note that (11.71b) has to be performed for each tensor represented by these schemes (cf. Remark 11.7b).

Remark 11.67. (a) Provided that $\mathbf{B}'_\alpha$ is already a basis, one option of the procedure **JoinBases** is to produce a new basis $\mathbf{B}_\alpha$ whose first r'_α columns coincide with those of $\mathbf{B}'_\alpha$.[22] Then $T' = \left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right]$ holds, that means for instance that the definition $C'^{(\beta,\ell)}_{\text{new}} := T' C'^{(\beta,\ell)}_{\text{old}}$ in (11.70b) copies all entries $c'^{(\beta,\ell)}_{ij,\text{old}}$ for $1 \le i \le r'_{\beta_1}$, $1 \le j \le r'_{\beta_2}$ into $c'^{(\beta,\ell)}_{ij,\text{new}}$ and adds the entries $c'^{(\beta,\ell)}_{ij,\text{new}} := 0$ for $r'_{\beta_1} < i \le r_{\beta_1}$ if $\beta_1 = \alpha$ or $r'_{\beta_2} < j \le r_{\beta_2}$ if $\beta_2 = \alpha$.

(b) Assume that $V_k = \mathbb{K}^{n_k}$ for $k \in D$. Then the cost of **JoinBases** in (11.70a) is $N_{\text{QR}}(n_k, r'_k + r''_k)$. For $\alpha \in T_D \backslash \mathcal{L}(T_D)$, the cost of **JoinBases** in (11.71a) is $N_{\text{QR}}(r_{\alpha_1} \cdot r_{\alpha_2}, r'_\alpha + r''_\alpha)$ (cf. Lemma 2.19b).

(c) The computation of (11.70b,c) can be reduced to either (11.70b) or (11.70c) (see Part (a)). For $T'' \neq I$, the transformation (11.70c) costs $2 r_\alpha r''_\beta r''_{\beta_1} r''_{\beta_2}$ operations, where one of the sons $\{\beta_1, \beta_2\} = S(\beta)$ coincides with α.

(d) Assuming the bounds $r_\alpha, r'_\alpha, r''_\alpha \le r$, $n_k \le n$ and $2r \le n$, we can estimate the overall cost by $\le d r^2 \left(12 r^2 + 8n\right)$.

In the case of the hierarchical representation with *orthonormal* bases, procedure **JoinONB** is to be used instead of **JoinBases**. This does not change the arithmetical cost.

[22] The optimal choice is to retain the basis $\mathbf{B}'_\alpha$ or $\mathbf{B}''_\alpha$ of largest dimension.

11.5.3.2 HOSVD Case

Here we consider two tensors

$$\mathbf{v}' = \rho_{\mathrm{HTR}}^{\mathrm{HOSVD}}\big(T_D, (\mathbf{C}'_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c'^{(D)}, (B'_j)_{j\in D}\big),$$
$$\mathbf{v}'' = \rho_{\mathrm{HTR}}^{\mathrm{HOSVD}}\big(T_D, (\mathbf{C}''_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c''^{(D)}, (B''_j)_{j\in D}\big)$$

together with their weights Σ'_α and Σ''_α. Remark 5.17 states that the HOSVD basis $\mathbf{B}_\alpha$ of the family $[\mathbf{v}'\ \mathbf{v}'']$ at vertex α is obtained by diagonalisation of

$$\mathcal{M}_\alpha(\mathbf{v}')\mathcal{M}_\alpha(\mathbf{v}')^{\mathrm{H}} + \mathcal{M}_\alpha(\mathbf{v}'')\mathcal{M}_\alpha(\mathbf{v}'')^{\mathrm{H}} = \mathbf{B}'_\alpha\Sigma'^2_\alpha\mathbf{B}'^{\mathrm{H}}_\alpha + \mathbf{B}''_\alpha\Sigma''^2_\alpha\mathbf{B}''^{\mathrm{H}}_\alpha. \quad (11.72)$$

We start with the leaves $\alpha = \{j\}$. By $\mathbf{JoinONB}(B'_j, B''_j, r_j, \hat{B}_j, T', T'')$ we obtain an intermediate orthonormal basis $\hat{B}_j$ with the properties

$$B'_j = \hat{B}_j T' \quad \text{and} \quad B''_j = \hat{B}_j T''.$$

The right-hand side of Eq. (11.72) becomes $\hat{B}_j\big(T'\Sigma'^2_j T'^{\mathsf{H}} + T''\Sigma''^2_j T''^{\mathsf{H}}\big)\hat{B}_j^{\mathsf{H}}$. Diagonalisation of the $r_j \times r_j$ matrix

$$T'\Sigma'^2_j T'^{\mathsf{H}} + T''\Sigma''^2_j T''^{\mathsf{H}} = T\Sigma^2_j T^{\mathsf{H}} \quad (11.73a)$$

allows us to form the final basis

$$B_j := \hat{B}_j T \quad (11.73b)$$

Now, the right-hand side of (11.72) equals $B_j\Sigma_j^2 B_j^{\mathsf{H}}$, i.e., we have determined the HOSVD representation of $[\mathbf{v}'\ \mathbf{v}'']$ at vertex $\alpha = \{j\}$. The identities $B'_j = B_j T^{\mathsf{H}} T'$ and $B''_j = B_j T^{\mathsf{H}} T''$ follow from $T^{-1} = T^{\mathsf{H}}$. Therefore, the coefficients $C'^{(\beta,\ell)}$ and $C''^{(\beta,\ell)}$ with $\beta = \text{father}(\{j\})$ are transformed into $\hat{C}'^{(\beta,\ell)}$ and $\hat{C}''^{(\beta,\ell)}$ according to Lemma 11.24.

Assume that new bases are created at the son vertices of $\alpha \in T_D\backslash\mathcal{L}(T_D)$. The isomorphic formulation of (11.72) in terms of the coefficient matrices becomes

$$\sum_\ell \Big(\hat{C}'^{(\alpha,\ell)}\Sigma'^2_\alpha\hat{C}'^{(\alpha,\ell)\mathsf{H}} + \hat{C}''^{(\alpha,\ell)}\Sigma''^2_\alpha\hat{C}''^{(\alpha,\ell)\mathsf{H}}\Big).$$

By $\mathbf{JoinONB}(\hat{\mathbf{C}}'_\alpha, \hat{\mathbf{C}}''_\alpha, r_\alpha, \hat{\mathbf{C}}_\alpha, T', T'')$ we obtain a common orthonormal basis $\hat{\mathbf{C}}_\alpha$ (isomorphic to $\hat{\mathbf{B}}_\alpha$) and proceed as in (11.73a,b):

$$T'\Sigma'^2_\alpha T'^{\mathsf{H}} + T''\Sigma''^2_\alpha T''^{\mathsf{H}} = T\Sigma^2_\alpha T^{\mathsf{H}} \text{ and } \mathbf{C}_\alpha := \hat{\mathbf{C}}_\alpha T. \quad (11.73c)$$

Note that the basis change

$$\mathbf{B}'_\alpha, \mathbf{B}''_\alpha \mapsto \mathbf{B}_\alpha \quad \text{with } \mathbf{B}'_\alpha = \mathbf{B}_\alpha T^{\mathsf{H}} T' \text{ and } \mathbf{B}''_\alpha = \mathbf{B}_\alpha T^{\mathsf{H}} T''$$

involves again a transformation $C'^{(\beta,\ell)}, C''^{(\beta,\ell)} \mapsto \hat{C}'^{(\beta,\ell)}, \hat{C}''^{(\beta,\ell)}$ according to Lemma 11.24.

Finally, at the root $\alpha = D$ the coefficients $c'^{(D)}$ and $c''^{(D)}$ are updated by

$$c'^{(D)}_{\text{new}} := T^{\mathsf{H}} T' c'^{(D)}, \quad c''^{(D)}_{\text{new}} := T^{\mathsf{H}} T'' c''^{(D)}. \tag{11.73d}$$

Eventually, we obtain the common representations

$$\begin{aligned} \mathbf{v}' &= \rho^{\text{HOSVD}}_{\text{HTR}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\setminus\mathcal{L}(T_D)}, c'^{(D)}_{\text{new}}, (B_j)_{j\in D}\big), \\ \mathbf{v}'' &= \rho^{\text{HOSVD}}_{\text{HTR}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\setminus\mathcal{L}(T_D)}, c''^{(D)}_{\text{new}}, (B_j)_{j\in D}\big). \end{aligned} \tag{11.73e}$$

Remark 11.68. The generalisation from two tensors to a family $\{\mathbf{v}_i : 1 \le i \le p\}$ of tensors is obvious.

11.5.3.3 Truncation

One option is to determine the exact common representation (11.73e) first and then to truncate according to §11.4.2.

A second, cheaper approach performs truncation together with the computation of the common bases. After obtaining T from (11.73a), we define T_{tr} by the first $r_{j,\text{tr}}$ columns of T (either $r_{j,\text{tr}}$ is a prescribed dimension or determined implicitly from the singular values in the diagonal matrix Σ_j). Then, $B_{j,\text{tr}} := \hat{B}_j T_{\text{tr}}$ spans a subspace $U_{j,\text{tr}}$ of the reduced dimension $r_{j,\text{tr}}$. Let $P_j := B_{j,\text{tr}} B^{\mathsf{H}}_{j,\text{tr}}$ be the orthogonal projection onto $U_{j,\text{tr}}$. Consequently, $\mathbf{v}'$ and $\mathbf{v}''$ are replaced by $P_j\mathbf{v}'$ and $P_j\mathbf{v}''$. Furthermore, the bases B'_j and B''_j are to be replaced by $P_j B'_j$ and $P_j B''_j$, which are no longer orthonormal, but $B_{j,\text{tr}}$ represents an orthonormal basis. Instead of the previous relation $B'_j = B_j T^{\mathsf{H}} T'$, one now obtains

$$P_j B'_j = B_{j,\text{tr}} T^{\mathsf{H}}_{\text{tr}} T' \quad \text{and} \quad P_j B''_j = B_{j,\text{tr}} T^{\mathsf{H}}_{\text{tr}} T''.$$

Therefore the updates of $C'^{(\beta,\ell)}, C''^{(\beta,\ell)}$ for the father β of $\alpha = \{j\}$ involve $T^{\mathsf{H}}_{\text{tr}} T'$ and $T^{\mathsf{H}}_{\text{tr}} T''$. Since $T^{\mathsf{H}}_{\text{tr}} T' \in \mathbb{K}^{r_{j,\text{tr}}\times r'_j}$, the updated version $\hat{C}'^{(\beta,\ell)}_{\text{tr}}$ is of size[23] $r'_{\beta_1,\text{tr}} \times r'_{\beta_2,\text{tr}}$ instead of $r'_{\beta_1} \times r'_{\beta_2}$. Because of the reduced size, the following calculations are cheaper.

The further truncation at the inner vertices follows the same line.

Remark 11.69. The truncation controls the absolute error. If two tensors $\mathbf{v}'$ and $\mathbf{v}''$ are converted into a common representation in order to compute the difference $\mathbf{v}' - \mathbf{v}''$ with $\|\mathbf{v}' - \mathbf{v}''\| \ll \|\mathbf{v}'\|$, the usual cancellation effect may occur.

[23] This holds after truncation at both son vertices β_1 and β_2.

11.6 Conversion from Sparse-Grid

The sparse-grid space $\mathbf{V}_{\mathrm{sg}}$ and the spaces $V_{(\ell)}$ are introduced in (7.18) and (7.17). Given $\mathbf{v} \in \mathbf{V}_{\mathrm{sg}}$ and any dimension partition tree T_D, we define the characteristic subspaces $\mathbf{U}_\alpha$ ($\alpha \in T_D$) as follows:

$$\mathbf{U}_\alpha = \sum_{\sum_{j\in\alpha} \ell_j = \ell+d-1} \bigotimes_{j\in\alpha} V_{(\ell_j)} \qquad \text{for } \alpha \in T_D \backslash \{D\}, \tag{11.74}$$

and $\mathbf{U}_D = \operatorname{span}\{\mathbf{v}\}$.

Because of (7.17), we may replace the summation in (7.18) and (11.74) over $\sum \ell_j = L := \ell + d - 1$ by $\sum \ell_j \leq L$.

We have to prove the nestedness property (11.11c): $\mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$. It is sufficient to prove

$$\bigotimes_{j\in\alpha} V_{(\ell_j)} \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$$

for any tuple $(\ell_j)_{j\in\alpha}$ with $\sum_{j\in\alpha} \ell_j = L$. Obviously,

$$\bigotimes_{j\in\alpha} V_{(\ell_j)} = \left(\bigotimes_{j\in\alpha_1} V_{(\ell_j)} \right) \otimes \left(\bigotimes_{j\in\alpha_2} V_{(\ell_j)} \right).$$

Since $\sum_{j\in\alpha_1} \ell_j \leq L$, the inclusion $\bigotimes_{j\in\alpha_1} V_{(\ell_j)} \subset \mathbf{U}_{\alpha_1}$ holds and, analogously, $\bigotimes_{j\in\alpha_2} V_{(\ell_j)} \subset \mathbf{U}_{\alpha_2}$.

If we also define $\mathbf{U}_D$ by (11.74), $\mathbf{U}_D = \mathbf{V}_{\mathrm{sg}}$ follows. Hence, $\mathbf{v} \in \mathbf{V}_{\mathrm{sg}}$ belongs to $\mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}$ (α_1, α_2 sons of D) proving

$$\mathbf{U}_D = \operatorname{span}\{\mathbf{v}\} \subset \mathbf{U}_{\alpha_1} \otimes \mathbf{U}_{\alpha_2}.$$

All subspaces $\mathbf{U}_\alpha$ satisfy $\dim(\mathbf{U}_\alpha) \leq \dim(\mathbf{V}_{\mathrm{sg}}) = O(2^L L^{d-1})$. This proves that any sparse grid tensor from $\mathbf{V}_{\mathrm{sg}}$ can be exactly represented in hierarchical format with subspaces of dimensions not exceeding $\dim(\mathbf{V}_{\mathrm{sg}})$.

Chapter 12
Matrix Product Systems

Abstract The term ‘matrix-product state’ (MPS) is introduced in quantum physics (see, e.g., Verstraete-Cirac [190], [105, Eq. (2)]). The related tensor representation can be found already in Vidal [191] without a special naming of the representation. The method has been reinvented by Oseledets and Tyrtyshnikov ([152], [155], [159]) and called ‘TT decomposition’.[1]
We start in *Sect. 12.1* with the finite dimensional case. In *Sect. 12.2* we show that the TT representation is a special form of the hierarchical format. Finally three conversions are considered: conversion from r-term format to TT format (cf. §12.3.1), from TT format into hierarchical format with a general tree T_D (cf. §12.3.2), and vice versa, from general hierarchical format into TT format (cf. §12.3.3). A closely related variant of the TT format is the cyclic matrix product format. As we shall see in *Sect. 12.4*, the change from the tree structure to a proper graph structure may have negative consequences.
The algorithms for obtaining HOSVD bases and for truncations are mentioned only briefly. The reason is the equivalence to the hierarchical format, so that the algorithms defined there can be easily transferred. The interested reader finds such algorithms in [155].

12.1 Basic TT Representation

12.1.1 Finite Dimensional Case

Consider $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ with $V_j = \mathbb{K}^{I_j}$ and a tensor $\mathbf{v} \in \mathbf{V} = \bigotimes_{j=1}^{d} V_j$ written as

$$\mathbf{v}[i_1 i_2 \cdots i_d] = \sum_{k_1=1}^{\rho_1} \cdots \sum_{k_{d-1}=1}^{\rho_{d-1}} v^{(1)}_{i_1 k_1} \cdot v^{(2)}_{k_1 i_2 k_2} \cdot \ldots \cdot v^{(d-1)}_{k_{d-2} i_{d-1} k_{d-1}} \cdot v^{(d)}_{k_{d-1} i_d} \qquad (12.1a)$$

[1] While the first interpretation of ‘TT’ has been ‘Tree Tensor’, the later reading is ‘Tensor Train’. We avoid the term ‘decomposition’ (there is no uniqueness, cf. §7.1.3) and prefer the term ‘TT representation’ or ‘TT format’.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_12,

for all $(i_1, \ldots, i_d) \in \mathbf{I} := I_1 \times \ldots \times I_d$. The scalars $v^{(j)}_{k_{j-1} i_j k_j}$ can be considered as entries of a tensor of order three from $\mathbb{K}^{K_{j-1}} \otimes \mathbb{K}^{I_j} \otimes \mathbb{K}^{K_j}$, where

$$K_j = \{1, \ldots, \rho_j\} \qquad \text{for } 0 \leq j \leq d. \tag{12.1b}$$

In the cases of $j = 1$ and $j = d$, we set

$$\rho_0 = \rho_d = 1, \qquad v^{(1)}_{i_1 k_1} = v^{(1)}_{1, i_1 k_1}, \quad v^{(d)}_{k_{d-1} i_d} = v^{(d)}_{k_{d-1} i_d, 1}, \tag{12.1c}$$

so that $\mathbb{K}^{K_0} \otimes \mathbb{K}^{I_1} \otimes \mathbb{K}^{K_1} = \mathbb{K} \otimes \mathbb{K}^{I_1} \otimes \mathbb{K}^{K_1} \cong \mathbb{K}^{I_1} \otimes \mathbb{K}^{K_1}$ and $\mathbb{K}^{K_{d-1}} \otimes \mathbb{K}^{I_d} \otimes \mathbb{K}^{K_d} \cong \mathbb{K}^{K_{d-1}} \otimes \mathbb{K}^{I_d}$.

Rewriting $v^{(j)}_{k_{j-1} i_j k_j}$ as $v^{(j)}_{k_{j-1} k_j}[i_j]$, we reformulate (12.1a) as

$$\mathbf{v}[i_1 i_2 \cdots i_d] = \sum_{k_1=1}^{\rho_1} \cdots \sum_{k_{d-1}=1}^{\rho_{d-1}} v^{(1)}_{k_1}[i_1] \cdot v^{(2)}_{k_1 k_2}[i_2] \cdot \ldots \cdot v^{(d-1)}_{k_{d-2} k_{d-1}}[i_{d-1}] \cdot v^{(d)}_{k_{d-1}}[i_d]. \tag{12.1d}$$

Fixing the indices $i_1, \ldots, i_d$, we interpret $v^{(j)}_{k_{j-1} k_j}[i_j]$ as entries of the matrix

$$V^{(j)}[i_j] := \left(v^{(j)}_{k_{j-1} k_j}[i_j]\right)_{k_{j-1} \in K_{j-1},\, k_j \in K_j} \in \mathbb{K}^{K_{j-1} \times K_j} \qquad (i_j \in I_j) \tag{12.2a}$$

(using (12.1c)). Then the entries $\mathbf{v}[i_1 i_2 \cdots i_d]$ can be regarded as matrix products:

$$\mathbf{v}[i_1 i_2 \cdots i_d] = V^{(1)}[i_1] \cdot V^{(2)}[i_2] \cdot \ldots \cdot V^{(d-1)}[i_{d-1}] \cdot V^{(d)}[i_d] \in \mathbb{K}\,. \tag{12.2b}$$

This representation justifies the term ‘matrix-product representation’. Note that $V^{(1)}[i_1] \in \mathbb{K}^{1 \times \rho_1} \cong \mathbb{K}^{\rho_1}$ is a row vector, while $V^{(d)}[i_d] \in \mathbb{K}^{\rho_{d-1} \times 1} \cong \mathbb{K}^{\rho_{d-1}}$ is a column vector.

For fixed k_{j-1}, k_j, the entries $v^{(j)}_{k_{j-1} k_j}[i_j]$ define the vector $v^{(j)}_{k_{j-1} k_j} \in \mathbb{K}^{I_j} = V_j$ (for all $1 \leq j \leq d$, using (12.1c)). Then (12.1a) is equivalent to

$$\mathbf{v} = \sum_{k_1=1}^{\rho_1} \sum_{k_2=1}^{\rho_2} \cdots \sum_{k_{d-1}=1}^{\rho_{d-1}} v^{(1)}_{1,k_1} \otimes v^{(2)}_{k_1 k_2} \otimes v^{(3)}_{k_2 k_3} \otimes \ldots \otimes v^{(d-1)}_{k_{d-2} k_{d-1}} \otimes v^{(d)}_{k_{d-1},1}. \tag{12.3a}$$

Formulation (12.3a) can be used for general spaces V_j. Using K_j from (12.1b) together with (12.1c), we shorten the notation (12.3a) by

$$\mathbf{v} = \sum_{k_0 \in K_0} \cdots \sum_{k_d \in K_d} \bigotimes_{j=1}^{d} v^{(j)}_{k_{j-1} k_j} \qquad \text{with } v^{(j)}_{k_{j-1} k_j} \in V_j. \tag{12.3b}$$

Definition 12.1 ($\mathbb{T}_{\boldsymbol{\rho}}$)**.** Let $\mathbf{V} = {}_a\bigotimes_{j=1}^d V_j$ and fix a tuple $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{d-1}) \in \mathbb{N}^{d-1}$. The TT format is defined by

$$\mathbb{T}_{\boldsymbol{\rho}} = \mathbb{T}_{\boldsymbol{\rho}}(\mathbf{V}) := \left\{ \mathbf{v} \in \mathbf{V} : \begin{array}{l} \mathbf{v} = \sum\limits_{\substack{k_i \in K_i \\ (0 \le i \le d)}} \bigotimes\limits_{j=1}^{d} v_{k_{j-1}k_j}^{(j)} \text{ with } v_{k_{j-1}k_j}^{(j)} \in V_j \\ \text{and } \#K_j = \begin{cases} 1 & \text{for } j = 0 \text{ or } j = d \\ \rho_j & \text{for } 1 \le j \le d-1 \end{cases} \end{array} \right\}. \tag{12.4}$$

Theorem 12.2. *Let $\mathbf{v} \in \mathbb{T}_{\boldsymbol{\rho}}$ with $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{d-1})$. Then $\rho_j \ge \rho_j^*$ holds with*

$$\rho_j^* := \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v}) \qquad \textit{for } 1 \le j < d \tag{12.5}$$

(cf. (5.6a)). Furthermore, a representation $\mathbf{v} \in \mathbb{T}_{\boldsymbol{\rho}^*}$ for $\boldsymbol{\rho}^* = \left(\rho_1^*, \ldots, \rho_{d-1}^*\right)$ exists.

Proof. 1) Summation in (12.3) over all k_ν except for $\nu = j$ yields

$$\mathbf{v} = \sum_{k_j=1}^{\rho_j} \mathbf{v}_{k_j}^{\{1,\ldots,j\}} \otimes \mathbf{v}_{k_j}^{\{j+1,\ldots,d\}} \quad \text{with} \tag{12.6}$$

$$\mathbf{v}_{k_j}^{\{1,\ldots,j\}} := \sum_{\substack{k_i \in K_i \\ (0 \le i \le j-1)}} \bigotimes_{j=1}^{d} v_{k_{j-1}k_j}^{(j)} \quad \text{and} \quad \mathbf{v}_{k_j}^{\{j+1,\ldots,d\}} := \sum_{\substack{k_i \in K_i \\ (j+1 \le i \le d)}} \bigotimes_{j=1}^{d} v_{k_{j-1}k_j}^{(j)}.$$

This proves $\operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v}) \le \rho_j$ (cf. Lemma 6.5). The results of §12.2.4 will show that a representation (12.3) with $\rho_j^* = \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v})$ can be obtained, i.e., $\mathbf{v}$ belongs to $\mathbb{T}_{\boldsymbol{\rho}^*}$. □

The TT representation is denoted by

$$\rho_{\mathsf{TT}}\left(\boldsymbol{\rho}, \left(\left(v_{k_{j-1}k_j}^{(j)}\right)_{\substack{k_{j-1} \in K_{j-1} \\ k_j \in K_j}}\right)_{1 \le j \le d}\right) = \sum_{\substack{k_i \in K_i \\ (0 \le i \le d)}} \bigotimes_{j=1}^{d} v_{k_{j-1}k_j}^{(j)}, \tag{12.7}$$

where $\#K_j = \rho_j$ and $v_{k_{j-1}k_j}^{(j)} \in V_j$. Note that (12.1c), i.e., $\#K_0 = \#K_d = 1$, is always required.

12.1.2 Function Case

As already mentioned, the formulation (12.3a) holds for $v_{k_{j-1}k_j}^{(j)} \in V_j$, whatever the vector space is. In the case of a function space V_j we regain (12.1d) in the form

$$\mathbf{f}(x_1, \ldots, x_d) = \sum_{k_i \in K_j \ (1 \le j \le d-1)} v_{k_1}^{(1)}(x_1) \cdot v_{k_1k_2}^{(2)}(x_2) \cdot \ldots \cdot v_{k_{d-2}k_{d-1}}^{(d-1)}(x_{d-1}) \cdot v_{k_{d-1}}^{(d)}(x_d).$$

Here, $V^{(j)}(x_j)$ from (12.2a) can be regarded as matrix-valued function.

A further generalisation replaces the matrices by kernel functions. Then the representation rank ρ_j becomes infinite:

$$\mathbf{f}(x_1,\ldots,x_d)=\int\limits_{K_1\times\cdots\times K_{d-1}} v^{(1)}(\kappa_1,x_1)v^{(2)}(\kappa_1,x_2,\kappa_2)\cdots v^{(d)}(\kappa_{d-1},x_d)\mathrm{d}\kappa_1\ldots\mathrm{d}\kappa_{d-1}.$$

By quadrature approximation one can regain the foregoing form with finite rank.

12.2 TT Format as Hierarchical Format

12.2.1 Related Subspaces

Let $D:=\{1,\ldots,d\}$. The dimension partition tree T_D^{TT} is given in Fig. 11.2, i.e., T_D^{TT} consists of leaves $\{j\}$ $(j\in D)$ and interior nodes $\{1,\ldots j\}$ for $j\in D\backslash\{1\}$:

$$T_D^{\mathsf{TT}}=\{\{1,\ldots j\},\{j\}:1\leq j\leq d\} \tag{12.8a}$$

The first son of $\{1,\ldots,j\}$ is $\{1,\ldots,j-1\}$, the second one is $\{j\}$:

$$S(\{1,\ldots,j\})=\{\{1,\ldots,j-1\},\{j\}\}\quad\text{for }2\leq j\leq d. \tag{12.8b}$$

According to (11.10), given a tensor $\mathbf{v}\in\mathbb{T}_{\boldsymbol{\rho}}(\mathbf{V})$, we have to introduce subspaces $\mathbf{U}_\alpha\subset\mathbf{V}_\alpha$ for all $\alpha\in T_D^{\mathsf{TT}}$. For $j=1$ we choose

$$U_1=\mathbf{U}_{\{1\}}=\operatorname{span}\{v_{k_1}^{(1)}:k_1\in K_1\}. \tag{12.9a}$$

Here, the vectors $v_{k_1}^{(1)}$ (and later $v_{k_{j-1}k_j}^{(j)}$) are those from the representation (12.3a). For $j>1$ the trivial choice

$$U_j=V_j\qquad\text{for }j\in D\backslash\{1\} \tag{12.9b}$$

is made. The next interior node is $\{1,2\}\in T_D^{\mathsf{TT}}$. As in (12.6) we form

$$\mathbf{v}_{k_2}^{\{1,2\}}:=\sum_{k_1=1}^{\rho_1}v_{1,k_1}^{(1)}\otimes v_{k_1k_2}^{(2)}\quad\text{ and }\quad\mathbf{U}_{\{1,2\}}:=\operatorname{span}\{\mathbf{v}_{k_2}^{\{1,2\}}:k_2\in K_2\}.$$

In the general case, $\mathbf{v}_{k_j}^{\{1,\ldots,j\}}:=\sum_{\substack{k_i\in K_i\\(0\leq i\leq j-1)}}\bigotimes_{\ell=1}^{j}v_{k_{j-1}k_j}^{(\ell)}$ is obtained recursively by

$$\mathbf{v}_{k_j}^{\{1,\ldots,j\}}=\sum_{k_{j-1}=1}^{\rho_{j-1}}\mathbf{v}_{k_{j-1}}^{\{1,\ldots,j-1\}}\otimes v_{k_{j-1}k_j}^{(j)}\qquad(k_j\in K_j). \tag{12.9c}$$

These tensors define the subspace

$$\mathbf{U}_{\{1,\ldots,j\}} := \operatorname{span}\{\mathbf{v}_{k_j}^{\{1,\ldots,j\}} : k_j \in K_j\} \qquad \text{for } j \in D\backslash\{1\} \tag{12.9d}$$

(the case $j=1$ is already stated in (12.9a)).

Since $\mathbf{v}_{k_{j-1}}^{\{1,\ldots,j-1\}} \in \mathbf{U}_{\{1,\ldots,j-1\}}$ and $v_{k_{j-1}k_j}^{(j)} \in U_j = V_j$, we obtain the inclusion

$$\mathbf{U}_{\{1,\ldots,j\}} \subset \mathbf{U}_{\{1,\ldots,j-1\}} \otimes U_j \qquad \text{for } j \in D\backslash\{1\}, \tag{12.9e}$$

which is the nestedness condition (11.11c), since $\{1,\ldots,j-1\}$ and $\{j\}$ are the sons of $\{1,\ldots,j\}$. Because of $\#K_d=1$ (cf. (12.4)), there is only one tensor $\mathbf{v}_{k_d}^{\{1,\ldots,d\}}=\mathbf{v}$ which spans $\mathbf{U}_D$. This proves

$$\mathbf{v} \in \mathbf{U}_D \quad \text{and} \quad \dim(\mathbf{U}_D) = 1. \tag{12.9f}$$

Following Definition 11.8, the tensor $\mathbf{v} \in \mathbb{T}_{\boldsymbol{\rho}}$ is represented by the hierarchical subspace family $\{\mathbf{U}_\alpha\}_{\alpha \in T_D^{\mathrm{TT}}}$.

12.2.2 From Subspaces to TT Coefficients

Let the subspaces $\mathbf{U}_{\{1,\ldots,j\}} \subset \mathbf{V}_{\{1,\ldots,j\}}$ satisfy conditions (12.9e,f). Choose any basis (or frame) $\{b_k^{(1)} : k \in K_1\}$ of $\mathbf{U}_{\{1\}} = U_1$ and rename the basis vectors by $v_k^{(1)} = b_k^{(1)}$. For $j \in \{2,\ldots,d-1\}$ let $\{\mathbf{b}_k^{(j)} : k \in K_j\}$ be a basis (or frame) of $\mathbf{U}_{\{1,\ldots,j\}}$ and assume by induction that the tensors

$$\mathbf{b}_{k_{j-1}}^{(j-1)} = \sum_{k_1,\ldots,k_{j-2}} v_{k_1}^{(1)} \otimes v_{k_1k_2}^{(2)} \otimes \ldots \otimes v_{k_{j-2}k_{j-1}}^{(j-1)} \qquad (k_{j-1} \in K_{j-1}) \tag{12.10}$$

are already constructed. By inclusion (12.9e), the basis vector $\mathbf{b}_k^{(j)}$ has a representation

$$\mathbf{b}_{k_j}^{(j)} = \sum_{k_{j-1}\in K_{j-1}} \sum_{i_j \in I_j} c_{k_{j-1},i_j}^{(\alpha,k_j)} \mathbf{b}_{k_{j-1}}^{(j-1)} \otimes b_{i_j}^{(j)} \quad \text{with } \alpha = \{1,\ldots,j\}, \text{ cf. (11.24).}$$

Setting $v_{k_{j-1}k_j}^{(j)} = \sum_{i_j} c_{k_{j-1},i_j}^{(\alpha,k_j)} b_{i_j}^{(j)}$, (12.10) follows for j instead of $j-1$. For $j=d$, the tensor $\mathbf{v} \in \mathbf{U}_D \subset \mathbf{U}_{\{1,\ldots,d-1\}} \otimes U_d$ is written as $\mathbf{v} = \sum_{i_d} c_{k_{d-1},i_d}^{(D,1)} \mathbf{b}_{k_{d-1}}^{(d-1)} \otimes b_{i_d}^{(d)}$. Now, $v_{k_{d-1}}^{(d)} = \sum_{i_d} c_{k_{d-1},i_d}^{(D,1)} b_{i_d}^{(d)}$ defines the last coefficients in the representation (12.1a). Note that the cardinalities $\rho_j = \#K_j$ coincide with $\dim(\mathbf{U}_{\{1,\ldots,j\}})$, provided that bases (not frames) are used.

As a by-product, the construction shows how the data $v_{k_{j-1}k_j}^{(j)}$ are connected to the coefficients $c_{k_{j-1},i_j}^{(\{1,\ldots,j\},k_j)}$ of the hierarchical format.

12.2.3 From Hierarchical Format to TT Format

Now, we start from $\mathbf{v} \in \mathcal{H}_{\mathfrak{r}}$ with the underlying tree T_D^{TT} from (12.8a,b) and a rank tuple $\mathfrak{r} = (r_\alpha)_{\alpha \in T_D^{\mathsf{TT}}}$. We may construct a TT-representation based on the subspaces $(\mathbf{U}_\alpha)_{\alpha \in T_D^{\mathsf{TT}}}$ as in §12.2.2. Instead, we translate the data from

$$\mathbf{v} = \rho_{\mathrm{HTR}}\big(T_D^{\mathsf{TT}}, (\mathbf{C}_\alpha), c^{(D)}, (B_j)\big) \in \mathcal{H}_{\mathfrak{r}}$$

directly into the TT data of $\rho_{\mathsf{TT}}\big(\boldsymbol{\rho}, (v^{(j)}_{k_{j-1}k_j})\big)$ with $\rho_j = r_{\{1,\ldots,j\}}$. By (11.26) the explicit representation of $\mathbf{v}$ is

$$\mathbf{v} = \sum_{\substack{i[\alpha]=1 \\ \text{for } \alpha \in T_D^{\mathsf{TT}}}}^{r_\alpha} c^{(D)}_{i[D]} \left[\prod_{\beta \in T_D \backslash \mathcal{L}(T_D)} c^{(\beta, i[\beta])}_{i[\beta_1], i[\beta_2]} \right] \bigotimes_{j=1}^{d} b^{(j)}_{i[\{j\}]}.$$

We rename the indices as follows: for $\alpha = \{1, \ldots, j\} \in T_D^{\mathsf{TT}}$ we rewrite $i[\alpha]$ by k_j, and for leaves $\alpha = \{j\} \in T_D^{\mathsf{TT}}$, $j > 1$, we write i_j. This yields

$$\mathbf{v} = \sum_{\substack{i_\ell \in I_\ell \\ (2 \le \ell \le d)}} \sum_{\substack{k_\ell = 1 \\ (1 \le \ell \le d)}}^{r_{\{1,\ldots,\ell\}}} c^{(D)}_{k_d} \left[\prod_{j=2}^{d} c^{(\{1,\ldots,j\},k_j)}_{k_{j-1}, i_j} \right] b^{(1)}_{k_1} \otimes b^{(2)}_{i_2} \otimes \ldots \otimes b^{(d)}_{i_d}.$$

Because of the choice $U_j = V_j$ for $2 \le j \le d$, the basis $\{b^{(j)}_i : i \in I_j\}$ is the canonical one formed by the unit vectors of $V_j = \mathbb{K}^{I_j}$. This implies $b^{(j)}_i[\ell] = \delta_{i\ell}$. Therefore, the entries of $\mathbf{v}$ have the form

$$\begin{aligned}
\mathbf{v}[i_1 i_2 \cdots i_d] &= \sum_{\substack{k_\ell = 1 \\ (1 \le \ell \le d)}}^{\rho_\ell} c^{(D)}_{k_d} \left[\prod_{j=2}^{d} c^{(\{1,\ldots,j\},k_j)}_{k_{j-1}, i_j} \right] b^{(1)}_{k_1}[i_1] \\
&= \sum_{\substack{k_\ell = 1 \\ (1 \le \ell \le d-1)}}^{\rho_\ell} b^{(1)}_{k_1}[i_1] \cdot c^{(\{1,2\},k_2)}_{k_1, i_2} \cdot \ldots \cdot c^{(\{1,\ldots,d-1\},k_{d-1})}_{k_{d-2}, i_{d-1}} \cdot \sum_{k_d = 1}^{r_D} c^{(\{1,\ldots,d\},k_d)}_{k_{d-1}, i_d} c^{(D)}_{k_d}
\end{aligned}$$

with $\rho_\ell := r_{\{1,\ldots,\ell\}}$. Defining

$$\begin{aligned}
v^{(1)}_{k_1}[i_1] &:= b^{(1)}_{k_1}[i_1] && \text{for } j = 1, && i_1 \in I_1, \\
v^{(j)}_{k_{j-1},k_j}[i_j] &:= c^{(\{1,\ldots,j\},k_j)}_{k_{j-1}, i_j} && \text{for } 2 \le j \le d-1, && i_j \in I_j, \\
v^{(d)}_{k_{d-1}}[i_d] &:= \textstyle\sum_{k_d=1}^{\rho_d} c^{(\{1,\ldots,d\},k_d)}_{k_{d-1}, i_d} c^{(D)}_{k_d} && \text{for } j = d, && i_d \in I_d, \\
& \text{with } 1 \le k_j \le \rho_j \text{ for } 1 \le j \le d-1,
\end{aligned} \tag{12.11}$$

we get the matrix formulation

$$\mathbf{v}[i_1 i_2 \cdots i_d] = \sum_{\substack{k_\ell=1 \\ (1\le\ell\le d-1)}}^{\rho_\ell} v_{k_1}^{(1)}[i_1] \cdot v_{k_1,k_2}^{(2)}[i_2] \cdot \ldots \cdot v_{k_{d-2},k_{d-1}}^{(d-1)}[i_{d-1}] \cdot v_{k_{d-1}}^{(d)}[i_d],$$

i.e., $\mathbf{v} \in \mathbb{T}_{\boldsymbol{\rho}}$ with $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{d-1})$.

12.2.4 Construction with Minimal ρ_j

Given a tensor $\mathbf{v} \in \mathbf{V}$ and the dimension partition tree T_D^{TT}, the considerations of §11.2.3 show that a hierarchical representation exists involving the minimal subspaces $\mathbf{U}_\alpha = \mathbf{U}_\alpha^{\min}(\mathbf{v})$ for $\alpha = \{1, \ldots, j\}$. Hence $r_\alpha = \dim(\mathbf{U}_\alpha^{\min}(\mathbf{v})) = \operatorname{rank}_\alpha(\mathbf{v})$ (cf. (6.15)). As seen in §12.2.3, this hierarchical representation can be transferred into TT format with

$$\rho_j = \dim(\mathbf{U}_{\{1,\ldots,j\}}^{\min}(\mathbf{v})) = \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v}).$$

On the other side, the first part of Theorem 12.2 states that $\rho_j \geq \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v})$. This proves the second part of Theorem 12.2.

12.2.5 Extended TT Representation

Representation (12.7) requires to store $\rho_{j-1}\rho_j$ vectors $v_{k_{j-1},k_j}^{(j)}$ from V_j. In the optimal case, all $v_{k_{j-1},k_j}^{(j)}$ belong to $U_j^{\min}(\mathbf{v})$ whose dimension is r_j. It is not unlikely that $\rho_{j-1}\rho_j > r_j$ holds. Then it may be more advantageous to store a basis $\{b_i^{(j)} : 1 \leq i \leq r_j\}$ of $U_j^{\min}(\mathbf{v})$. The representations

$$v_{k_1}^{(1)} = \sum_{i=1}^{r_1} a_i^{(1,1,k_1)} b_i^{(1)}, \qquad v_{k_{d-1}}^{(d)} = \sum_{i=1}^{r_{d-1}} a_i^{(d,k_{d-1},1)} b_i^{(d)}, \tag{12.12}$$

$$v_{k_{j-1},k_j}^{(j)} = \sum_{i=1}^{r_j} a_i^{(j,k_{j-1},k_j)} b_i^{(j)} \qquad (2 \leq j \leq d-1),$$

lead to the overall storage cost

$$\sum_{j=1}^{d} r_j \rho_{j-1} \rho_j + \sum_{j=1}^{d} r_j \cdot size(V_j) \qquad \text{with } \rho_0 = \rho_d = 1.$$

In this case, the tensor $\mathbf{v}$ is given by[2]

[2] $a_{i_j}^{(j,k_{j-1},k_j)}$ is to be interpreted as $a_{i_1}^{(1,k_1)}$ for $j = 1$ and as $a_{i_d}^{(d,k_{d-1})}$ for $j = d$.

$$\mathbf{v} = \sum_{\substack{k_i \in K_i \\ (0 \le i \le d)}} \prod_{\substack{1 \le i_j \le r_j \\ (1 \le j \le d)}} a_i^{(j,k_{j-1},k_j)} \bigotimes_{j=1}^{d} b_{i_j}^{(j)}.$$

This format is called 'extended tensor-train decomposition' in [158, Eq. (11)]. Note that the optimal values of the decisive parameters (ranks) are

$$r_j = \operatorname{rank}_j(\mathbf{v}) \quad \text{and} \quad \rho_j = \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v}) \qquad \text{for } 1 \le j \le d. \tag{12.13}$$

This format is completely equivalent to the hierarchical format with the particular choice of the dimension partition tree T_D^{TT}.

12.2.6 Properties

Remark 12.3. The *storage cost* for $\mathbf{v} \in \mathbb{T}_{\boldsymbol{\rho}}$ is

$$\sum_{j=1}^{d} \rho_{j-1}\rho_j \cdot size(V_j) \quad \text{with } \rho_0 = \rho_d = 1,$$

where $size(V_j)$ denotes the storage size for a vector from V_j. Under the assumption

$$size(V_j) = n \qquad \text{and} \quad \rho_j = \rho \quad \text{for all } 1 \le j \le d-1,$$

the data need a storage of size

$$\big((d-2)\,\rho^2 + 2\rho\big)\, n.$$

The storage cost is less, if some factors $v_{k_{j-1},k_j}^{(j)}$ vanish (cf. Remark 12.4).

The storage cost improves for the extended TT format (cf. §12.2.5), since then it coincides with the storage cost of the hierarchical format.

Remark 11.4b states that any permutation of indices from D which correspond to the interchange of sons $\{\alpha_1, \alpha_2\} = S(\alpha)$, leads to an isomorphic situation. In the case of the TT representation, the only permutation keeping the linear tree structure and leading to the same ranks ρ_j is the reversion

$$(1, \ldots, d) \mapsto (d, d-1, \ldots, 1).$$

The reason is (6.17a): $\rho_j = \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v}) = \operatorname{rank}_{\{j+1,\ldots,d\}}(\mathbf{v})$.

Note that the underlying linear tree T_D^{TT} is not the optimal choice with respect to the following aspects. Its length $d-1$ is maximal, which might have negative effects, e.g., in Remark 11.62. Because of the linear structure, computations are sequential, while a balanced tree T_D supports parallel computations (cf. Remark 11.60).

12.2.7 HOSVD Bases and Truncation

In principle, the HOSVD computation is identical to the algorithm from §11.3.3.4. The do statement 'for all sons $\sigma \in S(\alpha)$' in (11.46b) can be rewritten. Since for the linear tree T_D^{TT}, the second son α_2 is a leaf, the recursion coincides with the loop from $\{1, \ldots, d\}$ to $\{1, 2\}$, i.e., from d to 2. The first matrix in this loop is $M_d := V^{(d)}[1] \in \mathbb{K}^{\rho_{d-1} \times n_d}$ (cf. (12.2b) and $\rho_d = 1$), to which a left-sided singular value decomposition is applied. Let H_d be the result (i.e., $M_d = H_d \Sigma_d G_d^{\mathsf{T}}$ with G_d, H_d orthogonal). The HOSVD basis of the d-th direction is given by $H_d^{\mathsf{H}} V^{(d)}[\bullet]$. The matrix-product representation $V^{(1)}[i_1] \cdot V^{(2)}[i_2] \cdot \cdots \cdot V^{(d-1)}[i_{d-1}] \cdot V^{(d)}[i_d]$ from (12.2b) is transformed into

$$V^{(1)}[i_1] \cdot V^{(2)}[i_2] \cdot \ldots \cdot V^{(d-1)}[i_{d-1}] H_d \cdot \underbrace{H_d^{\mathsf{H}} V^{(d)}[i_d]}_{V^{(d,\mathrm{HOSVD})}[i_d]}$$

with $V^{(d,\mathrm{HOSVD})}[\bullet] \in \mathbb{K}^{\rho_{d-1}^{\mathrm{HOSVD}} \times n_d}$, where the rank $\rho_{d-1}^{\mathrm{HOSVD}} = \operatorname{rank}_{\{1\cdots d-1\}}(\mathbf{v})$ is possibly smaller than ρ_{d-1}. For general $2 \le j \le d-1$, the matrix

$$M_j := [V^{(j)}[i_1] H_{j+1} \; V^{(j)}[i_2] H_{j+1} \; \cdots \; V^{(j)}[i_{n_j}] H_{j+1}] \in \mathbb{K}^{\rho_{j-1} \times \rho_j^{\mathrm{HOSVD}} n_j}$$

possesses a left-sided singular matrix $H_j \in \mathbb{K}^{\rho_{j-1}^{\mathrm{HOSVD}} \times \rho_{j-1}^{\mathrm{HOSVD}}}$ and the matrix-product is further transformed into

$$V^{(1)}[i_1] \cdots V^{(j-1)}[i_{j-1}] H_j \cdot \underbrace{H_j^{\mathsf{H}} V^{(j)}[i_j] H_{j+1}}_{V^{(j,\mathrm{HOSVD})}[i_j]} \cdots \underbrace{H_{d-1}^{\mathsf{H}} V^{(d-1)}[i_{d-1}] H_d}_{V^{(d-1,\mathrm{HOSVD})}[i_{d-1}]} \cdot \underbrace{H_d^{\mathsf{H}} V^{(d)}[i_d]}_{V^{(d,\mathrm{HOSVD})}[i_d]}.$$

The final HOSVD matrices are

$$V^{(1,\mathrm{HOSVD})}[\bullet] = V^{(1)}[\bullet] H_2, \qquad V^{(j,\mathrm{HOSVD})}[\bullet] = H_j^{\mathsf{H}} V^{(j)}[\bullet] H_{j+1}$$

for $2 \le j \le d-1$, and $V^{(d,\mathrm{HOSVD})}[\bullet] = H_d^{\mathsf{H}} V^{(d)}[\bullet]$. The computational cost can be estimated by

$$2 \sum_{j=2}^{d} \rho_{j-1}^2 \left(\rho_{j-2} n_{j-1} + 2 \rho_j n_j + \frac{8}{3} \rho_{j-1} \right). \tag{12.14}$$

In §11.4.2.1, the truncation based on the HOSVD bases leads to the estimate (11.61) with the factor $\sqrt{2d-3}$, since $2d-3$ projections are applied. This number reduces to $d-1$ for the TT format, since only $d-1$ projections are performed (reduction of H_j to the first ρ'_j columns). The hierarchical format requires further $d-2$ projections for the subspaces $U_j \subset V_j$ ($2 \le j \le d-1$) which are now fixed by $U_j = V_j$.

The truncation of §11.4.2.3 leads again to the factor $\sqrt{d-1}$ instead of $\sqrt{2d-3}$ (same reasons as above).

12.3 Conversions

12.3.1 Conversion from $\mathcal{R}_r$ to $\mathbb{T}_{\boldsymbol{\rho}}$

Remark 12.4. (a) Let $\mathbf{v} \in \mathcal{R}_r$, i.e., $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} u_\nu^{(j)}$. Then set $\rho_j := r$ for $1 \le j \le d-1$ and

$$v_{k_1}^{(1)} := u_{k_1}^{(1)}, \quad v_{k_{d-1}}^{(d)} = u_{k_{d-1}}^{(d)}, \quad v_{k_{j-1},k_j}^{(j)} := \begin{cases} u_\nu^{(j)} & \text{for } k_{j-1} = k_j = \nu \\ 0 & \text{otherwise} \end{cases}$$

for the factors in (12.3a). Because most of the $v_{k_{j-1},k_j}^{(j)}$ are vanishing, the storage cost from Remark 12.3 reduces to the storage needed for $\mathcal{R}_r$.
(b) Part (a) describes the implication $\mathbf{v} \in \mathcal{R}_r \Rightarrow \mathbf{v} \in \mathbb{T}_{\boldsymbol{\rho}}$ for $\boldsymbol{\rho} = (r, \ldots, r)$. If $r = \operatorname{rank}(\mathbf{v})$, $\boldsymbol{\rho} = (r, \ldots, r)$ is minimal in the case of $d \le 3$, while for $d \ge 4$, the ranks ρ_j^* $(j \notin \{1, d-1\})$ of $\boldsymbol{\rho}^*$ with $\mathbf{v} \in \mathbb{T}_{\boldsymbol{\rho}^*}$ may be smaller than r.

Proof. We consider Part (b) only and prove that $\rho_1 = \rho_{d-1} = r$. Lemma 3.38 defines $v_\nu^{[1]}$ $(1 \le \nu \le r)$ and states that these vectors are linearly independent. This implies that

$$\rho_1 = \operatorname{rank}_{\{1\}}(\mathbf{v}) = \operatorname{rank}_{\{2,\ldots,d\}}(\mathbf{v}) = \dim\{v_\nu^{[1]} : 1 \le \nu \le r\} = r.$$

For ρ_{d-1} use $\rho_{d-1} = \operatorname{rank}_{\{d\}}(\mathbf{v}) = \dim\{v_\nu^{[d]} : 1 \le \nu \le r\} = r$. Note that for $d \le 3$, all indices $1 \le j \le d-1$ belong to the exceptional set $\{1, d-1\}$. □

12.3.2 Conversion from $\mathbb{T}_{\boldsymbol{\rho}}$ to $\mathcal{H}_{\mathfrak{r}}$ with a General Tree

The format $\mathbb{T}_{\boldsymbol{\rho}}$ is connected with the tree T_D^{TT} and the ordering $1, \ldots, d$ of the vector spaces V_j. We assume that another dimension partition tree T_D is based on the same ordering.[3] The tensor $\mathbf{v} = \rho_{\mathsf{TT}}\big(\boldsymbol{\rho}, (v_{k_{j-1}k_j}^{(j)})\big) \in \mathbb{T}_{\boldsymbol{\rho}}$ is described by the data $v_{k_{j-1}k_j}^{(j)} \in V_j$.

By the assumption on the ordering of the indices, each $\alpha \in T_D$ has the form

$$\alpha = \{j'_\alpha, j'_\alpha + 1, \ldots, j''_\alpha\}$$

for suitable $j'_\alpha, j''_\alpha \in D$. We define

$$\mathbf{u}_{k_{j'_\alpha-1},k_{j''_\alpha}}^{(\alpha)} := \sum_{k_{j'_\alpha}} \cdots \sum_{k_{j''_\alpha-1}} v_{k_{j'_\alpha-1}k_{j'_\alpha}}^{(j)} \otimes v_{k_{j'_\alpha}k_{j'_\alpha}+1}^{(j)} \otimes \ldots \otimes v_{k_{j''_\alpha-1}k_{j''_\alpha}}^{(j)} \tag{12.15a}$$

for $k_{j'_\alpha-1} \in K_{j'_\alpha-1}$, $k_{j''_\alpha} \in K_{j''_\alpha}$ (with $K_0 = K_d = \{1\}$), and

[3] According to Remark 11.4, several orderings can be associated with T_D. One of them has to coincide with the ordering of T_D^{TT}.

$$\mathbf{U}_\alpha := \operatorname{span}\left\{u^{(\alpha)}_{k_{j'_\alpha-1},k_{j''_\alpha}} : k_{j'_\alpha-1}\in K_{j'_\alpha-1},\ k_{j''_\alpha}\in K_{j''_\alpha}\right\}. \tag{12.15b}$$

Note that $\rho_j := \#K_j$. Since the number of tensors on the right-hand side of (12.15b) is $\#K_{j'_\alpha-1}\#K_{j''_\alpha} = \rho_{j'_\alpha-1}\rho_{j''_\alpha}$, we obtain the estimate

$$\dim(\mathbf{U}_\alpha) \le \rho_{j'_\alpha-1}\rho_{j''_\alpha} \qquad \text{for } \alpha = \{j\in D : j'_\alpha \le j \le j''_\alpha\}. \tag{12.15c}$$

For $\alpha\in T_D\backslash\mathcal{L}(T_D)$ with sons α_1,α_2, we can rewrite (12.15a) as

$$\mathbf{u}^{(\alpha)}_{k_{j'_\alpha-1},k_{j''_\alpha}} = \sum_{k_{j''_{\alpha_2}}\in K_{j''_{\alpha_2}}} \mathbf{u}^{(\alpha_1)}_{k_{j'_{\alpha_1}-1},k_{j''_{\alpha_1}}} \otimes \mathbf{u}^{(\alpha_2)}_{k_{j''_{\alpha_1}},k_{j''_{\alpha_2}}}, \tag{12.16}$$

since $j'_{\alpha_1} = j'_\alpha$, $j''_{\alpha_1} = j'_{\alpha_2}-1$, and $j''_{\alpha_2} = j''_\alpha$. Equality (12.16) proves the nestedness property $\mathbf{U}_\alpha \subset \mathbf{U}_{\alpha_1}\otimes\mathbf{U}_{\alpha_2}$. Since for $\alpha = D$, $\mathbf{u}^{(D)}_{k_{j'_D-1},k_{j''_D}} = \mathbf{v}$ holds, also $\mathbf{v}\in\mathbf{U}_D$ is shown. Hence, $\{\mathbf{U}_\alpha\}_{\alpha\in T_D}$ is a hierarchical subspace family and (11.15) is satisfied: $\mathbf{v}\in\mathcal{H}_\mathfrak{r}$.

Proposition 12.5. *Let* $\mathbf{v}\in\mathbb{T}_{\boldsymbol{\rho}}$ *with* $\boldsymbol{\rho} = (\rho_1,\ldots,\rho_{d-1})$ *and consider a hierarchical format* $\mathcal{H}_\mathfrak{r}$ *involving any dimension partition tree* T_D *with the same ordering of* D.
(a) All $\mathbf{v}\in\mathbb{T}_{\boldsymbol{\rho}}$ *can be transformed into a representation* $\mathbf{v}\in\mathcal{H}_\mathfrak{r}$*, where the dimensions* $\mathfrak{r} = (r_\alpha : \alpha\in T_D)$ *are bounded by*

$$r_\alpha \le \rho_{j'_\alpha-1}\cdot\rho_{j''_\alpha} \qquad \textit{for } \alpha = \{j : j'_\alpha \le j \le j''_\alpha\}\in T_D. \tag{12.17}$$

ρ_j *are the numbers appearing in* $\boldsymbol{\rho} = (\rho_1,\ldots,\rho_{d-1})$*. The estimate remains true for* $\rho_j := \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v})$.

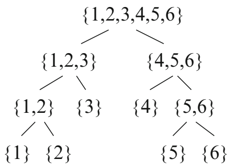
(b) If $T_D = T_D^{\mathsf{TT}}$ *(cf. (12.8a,b)),* $r_{\{1,\ldots,j\}} = \rho_j$ *holds.*

(c) If $d\le 6$*, a tree* T_D *with minimal depth can be chosen such that all* r_α *are bounded by some* r_j *or* ρ_j *from (12.13), i.e., no product like in (12.17) appears.*

Proof. 1) (12.17) corresponds to (12.15c). Using the definitions $r_\alpha = \operatorname{rank}_\alpha(\mathbf{v})$ for vertices $\alpha = \{j : j'_\alpha \le j \le j''_\alpha\}$ and (12.5), i.e., $\rho_{j'_\alpha-1} = \operatorname{rank}_\beta(\mathbf{v})$ for $\beta = \{1,\ldots,j'_\alpha-1\}$ and $\rho_{j''_\alpha} = \operatorname{rank}_\gamma(\mathbf{v}) = \operatorname{rank}_{\gamma^c}(\mathbf{v})$ for $\gamma^c = \{1,\ldots,j''_\alpha\}$, inequality (12.17) is a particular case of Lemma 6.19b.

2) Consider the case $d = 6$ in Part (c). Choose the tree depicted below. For a leaf $\alpha\in\mathcal{L}(T_D)$, the rank r_α is some r_j from (12.13). The vertices $\{1,2\}$ and $\{1,2,3\}$ lead to ρ_2 and ρ_3, while $r_{\{4,5,6\}} = r_{\{1,2,3\}} = \rho_3$ and $r_{\{5,6\}} = r_{\{1,2,3,4\}} = \rho_4$. □

A simplification of the previous proposition is as follows: if the TT representation uses the constant ranks $\boldsymbol{\rho} = (r,\ldots,r)$, the ranks of the hierarchical format are always bounded by r^2. This estimate is sharp for $d\ge 8$ as shown by an example in [75]. Up to $d = 6$, the better bound r can be achieved with optimally balanced trees.



12.3.3 Conversion from $\mathcal{H}_{\mathfrak{r}}$ to $\mathbb{T}_{\boldsymbol{\rho}}$

Given a hierarchical format $\mathcal{H}_{\mathfrak{r}}$ involving the tree T_D, we may consider $\mathbb{T}_{\boldsymbol{\rho}}$ with an optimal permutation of the dimension indices from $D = \{1, \ldots, d\}$. Rewriting this new ordering again by $1, \ldots, d$ means that the $\mathcal{H}_{\mathfrak{r}}$-vertices from of T_D are not necessarily of the form $\{j : j'_\alpha \leq j \leq j''_\alpha\}$. The $\mathbb{T}_{\boldsymbol{\rho}}$-ranks $\rho_j = \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v})$ can be estimated by products of the ranks $r_\alpha = \operatorname{rank}_\alpha(\mathbf{v})$ $(\alpha \in T_D)$ appearing in the hierarchical format as follows.

The set $\{1, \ldots, j\}$ can be represented (possibly in many ways) as a disjoint union of subsets of T_D:

$$\{1, \ldots, j\} = \bigcup_{\nu=1}^{\varkappa_j} \alpha_\nu \qquad (\text{disjoint } \alpha_\nu \in T_D). \tag{12.18}$$

The existence of such a representation is proved by the singletons (leaves of T_D): $\{1, \ldots, j\} = \bigcup_{\nu=1}^{j} \{\nu\}$, yielding the largest possible value $\varkappa_j = j$. In the best case, $\{1, \ldots, j\}$ is already contained in T_D and $\varkappa_j$ equals 1. Let $\varkappa'_{j,\min}$ be the smallest $\varkappa_j$ in (12.18) taken over all possible representations. Then Lemma 6.19b states that $\rho_j \leq \prod_{\nu=1}^{\varkappa'_{j,\min}} r_{\alpha_\nu^{\min}}$, where $\alpha_\nu^{\min}$ are the subsets with $\{1, \ldots, j\} = \bigcup_{\nu=1}^{\varkappa'_{j,\min}} \alpha_\nu^{\min}$. However, since $\rho_j = \operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v})$ coincides with $\operatorname{rank}_{\{j+1,\ldots,d\}}(\mathbf{v})$, one has also to consider partitions $\{j+1, \ldots, d\} = \bigcup_{\nu=1}^{\varkappa_j} \beta_\nu$ $(\beta_\nu \in T_D)$. Let $\varkappa''_{j,\min}$ and $\beta_\nu^{\min}$ be the optimal choice in the latter case. Then

$$\rho_j \leq \min\left\{ \prod_{\nu=1}^{\varkappa'_{j,\min}} r_{\alpha_\nu^{\min}}, \prod_{\nu=1}^{\varkappa''_{j,\min}} r_{\beta_\nu^{\min}} \right\}$$

follows. Assuming a hierarchical format $\mathcal{H}_{\mathfrak{r}}$ with $r_\alpha \leq r$ for all $\alpha \in T_D$, we obtain the estimate $\rho_j \leq r^{\min\{\varkappa'_{j,\min}, \varkappa''_{j,\min}\}}$. Introducing

$$\varkappa_{\max} := \max\left\{ \min\{\varkappa'_{j,\min}, \varkappa''_{j,\min}\} : 1 \leq j \leq d \right\},$$

we get

$$\max\{\rho_j : 1 \leq j \leq d\} \leq r^{\varkappa_{\max}}.$$

To understand how large $\varkappa_{\max}$ may become, we consider the regular case of a completely balanced tree for $d = 2^L$ with even L (cf. (11.1) for $L = 2$). We choose

$$j := \sum_{\nu=0}^{\frac{L}{2}-1} 4^\nu \tag{12.19}$$

and note that $d - j = 1 + \sum_{\nu=0}^{L/2-1} 2 \cdot 4^\nu$. Since all subsets from T_D have the size 2^μ for some $\mu \in \{0, \ldots, L\}$, one verifies that $\{1, \ldots, j\}$ needs exactly $\varkappa'_{j,\min} = L/2$ subsets $\alpha_{\nu,\min}$. Similarly, $\{j+1, \ldots, d\}$ is covered by at least $L/2+1$ subsets. This proves

$$\varkappa_{\max} = \frac{L}{2} = \frac{1}{2}\log_2 d.$$

Conclusion 12.6. *Let T_D be the balanced tree for $d = 2^L$. Assuming a hierarchical format $\mathcal{H}_\mathfrak{r}$ with $r_\alpha = r$ for all $\alpha \in T_D$, the rank ρ_j for $j = \sum_{\nu=0}^{\frac{L}{2}-1} 4^\nu$ is bounded by*

$$\rho_j \leq r^{\frac{1}{2}\log_2 d} = d^{\frac{1}{2}\log_2 r}. \tag{12.20}$$

Conjecture 12.7. One can construct tensors $\mathbf{v} \in \mathcal{H}_\mathfrak{r}$ such that inequality (12.20) becomes an equality.

The reasoning is as follows (see also [75]). By Lemma 6.19c, $\mathrm{rank}_{\alpha_\nu^{\min}}(\mathbf{v}) = \mathrm{rank}_{\beta_\nu^{\min}}(\mathbf{v}) = r$ implies $\rho_j = d^{\frac{1}{2}\log_2 r}$ for a suitable tensor $\mathbf{v} \in \mathbf{V}$. However, such a tensor may violate the conditions $\mathrm{rank}_\alpha(\mathbf{v}) \leq r$ for the other vertices $\alpha \in T_D$. Nevertheless, for the special j from (12.19) and the associated vertices $\alpha_\nu^{\min}$ and $\beta_\nu^{\min}$, equality (12.20) can be proved. The missing part is the argument that another partition of $\{1,\ldots,j\}$ consisting of $\varkappa > \varkappa_{\max}$ vertices may be such that the proof of $\rho_j = r^\varkappa$ does not apply. In this case, at least $\rho_j < r^{\varkappa-1}$ must hold. See also the example mentioned below.

In more detail, the following cases can be distinguished:

1) $1 \leq d \leq 5$: The ranks ρ_j of $\mathbb{T}_{\boldsymbol{\rho}}$ coincide with certain r_α from $\mathcal{H}_\mathfrak{r}$. For $d \leq 3$, any tree after reordering coincides with the linear tree. For $d = 4$ take the natural ordering of the balanced tree T_D. Then,

$$\rho_1 = r_{\{1\}}, \quad \rho_2 = r_{\{1,2\}}, \quad \rho_3 = \mathrm{rank}_{\{1,2,3\}}(\mathbf{v}) = \mathrm{rank}_{\{4\}}(\mathbf{v}) = r_{\{4\}}$$

holds. If $d=5$, four leaves (say 1-4) are contained in pair vertices (say $\{1,2\}$, $\{3,4\}$). Use the numbering $i_1=1$, $i_2=2$, $i_3=5$, $i_4=3$, $i_5=4$ for $\mathbb{T}_{\boldsymbol{\rho}}$. Again, we may use $\rho_3 = \mathrm{rank}_{\{1,2,5\}}(\mathbf{v}) = \mathrm{rank}_{\{3,4\}}(\mathbf{v}) = r_{\{3,4\}}$.

2) $6 \leq d \leq 16$: Assuming $r_\alpha \leq r$ for $\mathcal{H}_\mathfrak{r}$, the ranks ρ_j of $\mathbb{T}_{\boldsymbol{\rho}}$ are bounded by $\rho_j \leq r^2$ (equality sign is taken for suitable $\mathbf{v}$ and j). For the case $d = 6$ we choose the balanced tree[4] with three vertices $\{1,2\}$, $\{3,4\}$, $\{5,6\}$ of size 2. Then all triples $\{i_1, i_2, i_3\}$ as well as their complements are of size 3, while the vertices of T_D are of size 1, 2, 4, 6. Hence, the quantity $\min\{\varkappa'_{j,\min}, \varkappa''_{j,\min}\}$ from above is at least 2. An example in [75] proves that already for $d = 6$ the maximal TT rank is the square of the hierarchical ranks.

3) $d = 17$: Take the balanced tree of 16 leaves and add the vertices $\{17\}$ and $\{1,\ldots,17\}$. Then $\rho_j \leq r^3$ holds for some j. For a proof take any permutation $\{i_1,\ldots,i_{16}\}$ and consider the set $\{i_1,\ldots,i_7\}$. It requires $\varkappa'_{7,\min} \geq 3$ subsets. The complement $\{i_7,\ldots,i_{16}\}$ has 10 elements. To obtain $\varkappa''_{7,\min} = 2$ one must use $10 = 8 + 2$, i.e., $\{i_1,\ldots,i_7\}$ must be the union of two vertices $\alpha \in T_D$ with cardinality 8 and 2. This proves $17 \notin \{i_7,\ldots,i_{16}\}$ and $17 \in \{i_1,\ldots,i_7\}$. Now, 17 is contained in $\{i_1,\ldots,i_{10}\}$ and leads to $\varkappa'_{10,\min} \geq 3$, while $\{i_{11},\ldots,i_{17}\}$ with 7 elements proves $\varkappa''_{10,\min} \geq 3$.

[4] However, the tree containing the vertices $\{1,2\}$, $\{1,2,3\}$, $\{5,6\}$, $\{4,5,6\}$ has the same length and allows the estimate $\rho_j \leq r$.

12.4 Cyclic Matrix Products and Tensor Network States

The definition $\rho_0 = \rho_d = 1$ in (12.1c) or $\#K_j = 1$ for the index sets for $j = 0$ and $j = d$ has the purpose to avoid summations over these indices. Instead, one can identify the indices of $K_0 = K_d$ and allow $\rho_d > 1$:

$$\mathbf{v} = \sum_{k_1=1}^{\rho_1} \cdots \sum_{k_{d-1}=1}^{\rho_{d-1}} \sum_{k_d=1}^{\rho_d} v^{(1)}_{k_d,k_1} \otimes v^{(2)}_{k_1 k_2} \otimes \ldots \otimes v^{(d-1)}_{k_{d-2}k_{d-1}} \otimes v^{(d)}_{k_{d-1},k_d}. \tag{12.21}$$

This results into a cycle instead of a linear tree. In the following, we set $D = \mathbb{Z}_d$ which implies $0 = d$ and hence $\rho_0 = \rho_d$. Although this tensor representation looks quite similar to (12.3a), it has essentially different properties.

Proposition 12.8. *(a) If $\rho_j = 1$ for at least one $j \in D$, the tensor representation (12.21) coincides with (12.3a) with the ordering $\{j+1, j+1, \ldots, d, 1, \ldots, j\}$.*
(b) The minimal subspace $U_j^{\min}(\mathbf{v})$ is not related to a single parameter ρ_k in (12.21), i.e., ρ_k cannot be interpreted as a subspace dimension.
(c) Inequality $\operatorname{rank}_j(\mathbf{v}) = \dim(U_j^{\min}(\mathbf{v})) \le \rho_{j-1}\rho_j$ holds for (12.21).
(d) In general, a cyclic representation with $\operatorname{rank}_j(\mathbf{v}) = \rho_{j-1}\,\rho_j$ does not exist.

Proof. 1) Assume that $j = d$ in Part (a). Then $\rho_0 = \rho_d = 1$ yields (12.3a).

2) Fix $j = 1$. Then the representation $\mathbf{v} = \sum_{k_1=1}^{\rho_1} \sum_{k_d=1}^{\rho_d} v^{(1)}_{k_d,k_1} \otimes \mathbf{v}^{[1]}_{k_d,k_1}$ holds with $\mathbf{v}^{[1]}_{k_d,k_1} := \sum_{k_2,k_3,\ldots,k_{d-1}} \bigotimes_{\ell=2}^{d} v^{(\ell)}_{k_{\ell-1}k_\ell}$. Both indices k_d and k_1 enter the definition of $U_1^{\min}(\mathbf{v}) = \left\{\varphi(\mathbf{v}^{[1]}_{k_d,k_1}) : \varphi \in \mathbf{V}'_{[1]}\right\}$ in the same way proving Part (b). Obviously, the dimension is bounded by $\rho_d\rho_1 = \rho_0\rho_1$ as stated in Part (c).

3) If $\operatorname{rank}_j(\mathbf{v})$ is a prime number, $\operatorname{rank}_j(\mathbf{v}) = \rho_{j-1}\rho_j$ implies that $\rho_{j-1} = 1$ or $\rho_j = 1$. Hence, by Part (a), (12.21) cannot be a proper cyclic representation. □

According to Proposition 12.8a, we call (12.21) a proper representation, if all ρ_j are larger than 1. In the cyclic case, the ranks ρ_j are not related to the dimensions of $U_j^{\min}(\mathbf{v})$. Therefore, we cannot repeat the proof of Lemma 11.55 to prove closedness of the format (12.21). In fact, non-closedness is proved (cf. Theorem 12.9).

The cycle $\mathbb{Z}_d$ is only one example of tensor representations based on graphs (instead of trees). Examples[5] are given in Hübener-Nebendahl-Dür [105, p. 5], in particular, multi-dimensional grid-shaped graphs are considered instead of the one-dimensional chain $1 - 2 - \ldots - d$ used in (12.3a). Whenever the graph contains a cycle, statements as in Proposition 12.8b-d can be made and instability must be expected because of Theorem 12.9. If a connected graph contains no cycle, it describes a tree and corresponds to the hierarchical format (with the possible generalisation that the dimension partition tree T_D is not necessarily binary; cf. Definition 11.2).

The following result is proved by Landsberg-Qi-Ye [136]. It implies that a similar kind of instability may occur as for the r-term format (cf. Remark 9.14).

Theorem 12.9. *In general, a graph-based format containing a cycle is not closed.*

[5] The graph-based tensor format has several names: tensor network states, finitely correlated states (FCS), valance-bond solids (VBS), projected entangled pairs states (PEPS), etc. (cf. [136]).

Chapter 13
Tensor Operations

Abstract In §4.6 several tensor operations have been described. The numerical tensor calculus requires the practical realisation of these operations. In this chapter we describe the performance and arithmetical cost of the operations for the different formats. The discussed operations are the addition in *Sect. 13.1*, evaluation of tensor entries in *Sect. 13.2*, the scalar product and partial scalar product in *Sect. 13.3,* the change of bases in *Sect. 13.4*, general binary operations in *Sect. 13.5*, Hadamard product in *Sect. 13.6*, convolution of tensors in *Sect. 13.7*, matrix-matrix multiplication in *Sect. 13.8*, and matrix-vector multiplication in *Sect. 13.9*. *Section 13.10* is devoted to special functions applied to tensors. In the last *Sect. 13.11* we comment on the operations required for the treatment of Hartree-Fock and Kohn-Sham applications in quantum chemistry.

In connection with the tensorisation discussed in Chap. 14, further operations and their cost will be discussed.

We repeat the consideration from §7.1 concerning operations. Two mathematical entities $s_1, s_2 \in S$ are represented via parameters $p_1, p_2 \in P_S$, i.e., $s_1 = \rho_S(p_1)$ and $s_2 = \rho_S(p_2)$. A binary operation $\boxdot$ leads to $s := s_1 \boxdot s_2$. We have to find a parameter $p \in P_S$ such that $s = \rho_S(p)$. Therefore, on the side of the parameter representations, we have to perform (7.2):

$$p := p_1 \,\widehat{\boxdot}\, p_2 \qquad :\Longleftrightarrow \qquad \rho_S(p) = \rho_S(p_1) \boxdot \rho_S(p_2).$$

The related memory cost is denoted by $N_{\mathrm{mem}}(\cdot)$ and the number of arithmetical operations by $N_{\boxdot}$ with '$\boxdot$' replaced by the respective operation.

The following list yields an overview of the asymptotic cost of various operations for the different formats. Here, we assume the most general case (different bases etc.) and consider only upper bounds using $n = \min n_j$, $r = \max r_\alpha$ etc.[1] Furthermore, the cost of the j-th scalar product is assumed to be $2n_j - 1$.

[1] Note that the meaning of the bound r differs for the formats, since different kinds of ranks are involved.

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series
in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_13,

	full	r-term	tensor subspace	hierarchical
storage	n^d	dnr	$r^d + dnr$	$dr^3 + dnr$
basis change	$2dn^{d+1}$	$2dn^2r$	$2dr^{d+1}$	$2dr^3$
orthonormalisation			$2dr^{d+1} + 2dnr^2$	$2dnr^2 + 4dr^4$
$\mathbf{u}+\mathbf{v}$	n^d	0	$2dr^{d+1} + 2dnr^2$	$8dnr^2 + 8dr^4$
$\mathbf{v}_\mathbf{i}$ evaluation	0	dr	$2r^d$	$2dr^3$
$\langle\mathbf{u},\mathbf{v}\rangle$	$2n^d$	$dr^2 + 2dnr^2$	$2dr^{d+1} + 8dnr^2$	$2dnr^2 + 6dr^4$
$\langle\mathbf{u},\mathbf{v}\rangle_{\alpha^c}$	$2n^{d+\#\alpha}$	$r^2\#\alpha^c + 2\#\alpha^c nr^2$	$2dr^{d+\#\alpha} + 8\#\alpha^c nr^2$	$< 2dnr^2 + 6dr^4$
$\langle\mathbf{v},\mathbf{v}\rangle_{\{j\}^c}$	n^{d+1}	$\frac{1}{2}dr^2 + dnr^2$	$2dr^{d+1} + 8dnr^2$	$< 2dnr^2 + 6dr^4$
$\mathbf{u}\odot\mathbf{v}$	n^d	dnr^2	$r^{2d} + dnr^2$	$dnr^2 + (d-1)\,r^4$
$\mathbf{Av}$	$2dn^{d+1}$	$2dn^2r$	$2d\left(r^{d+1} + n^2r + nr^2\right)$	$2dn^2r$
truncation		$\sim dr^2R + d^2rR$	$3dr^{d+1} + 2dr^2n$	$2dr^2n + 3dr^4$

The cost of the truncation is cited from §9.5.4 for[2] $\mathcal{R}_r$, from §8.3.3 for $\mathcal{T}_\mathbf{r}$, and from (11.46c) for $\mathcal{H}_\mathfrak{r}$.

The terms involving n may be improved by applying the tensorisation technique from §14 as detailed in §14.1.4. In the best case, n may be replaced by $O(\log n)$.

13.1 Addition

Given $\mathbf{v}', \mathbf{v}'' \in \mathbf{V}$ in some representation, we have to represent $\mathbf{v} := \mathbf{v}' + \mathbf{v}''$.

13.1.1 Full Representation

Assume $\mathbf{V} = \bigotimes_{j=1}^{d} \mathbb{K}^{I_j}$ with $\mathbf{I} = \times_{j=1}^{d} I_j$ (cf. §7.2). Given $\mathbf{v}', \mathbf{v}'' \in \mathbf{V}$ in full representation, the sum is performed directly by

$$\mathbf{v}_\mathbf{i} := \mathbf{v}'_\mathbf{i} + \mathbf{v}''_\mathbf{i} \qquad \text{for all } \mathbf{i} \in \mathbf{I}.$$

The memory $N_{\text{mem}}(\mathbf{v}) = \#\mathbf{I}$ is the same as for each of $\mathbf{v}', \mathbf{v}''$. Also the number of arithmetical operations equals

$$N_+^{\text{full}} = N_{\text{mem}}(\mathbf{v}) = \#\mathbf{I} = \prod_{j=1}^{d} \#I_j.$$

As a variant, one may consider sparse tensors. Then, obviously, $\mathbf{v}$ is less sparse than $\mathbf{v}', \mathbf{v}''$, unless both terms possess the same sparsity pattern.

Another variant is the full functional representation by a function. Given two functions $function\ v1(\ldots)$ and $function\ v2(\ldots)$, the sum is represented by the $function\ v(i_1, \ldots, i_d)$ defined by $v(\ldots) := v1(\ldots) + v2(\ldots)$. Hence, the cost per call increases: $N_v = N_{v1} + N_{v2}$.

[2] Here, the cost of one iteration is given.

13.1.2 *r*-Term Representation

Given $\mathbf{v}' = \sum_{\nu=1}^{r} v_\nu^{(1)} \otimes \ldots \otimes v_\nu^{(d)} \in \mathcal{R}_r$ and $\mathbf{v}'' = \sum_{\nu=1}^{s} w_\nu^{(1)} \otimes \ldots \otimes w_\nu^{(d)} \in \mathcal{R}_s$, the sum $\mathbf{v} = \mathbf{v}' + \mathbf{v}''$ is performed by *concatenation*, i.e.,

$$\mathbf{v} = \sum_{\nu=1}^{r+s} v_\nu^{(1)} \otimes \ldots \otimes v_\nu^{(d)} \in \mathcal{R}_{r+s}, \text{ where } v_{r+\nu}^{(j)} := w_\nu^{(j)} \text{ for } 1 \le \nu \le s,\ 1 \le j \le d.$$

The memory is additive: $N_{\text{mem}}(\mathbf{v}) = N_{\text{mem}}(\mathbf{v}') + N_{\text{mem}}(\mathbf{v}'')$, while no arithmetical work is required: $N_+ = 0$.

Since the result lies in $\mathcal{R}_{r+s}$ with increased representation rank $r+s$, we usually need a truncation procedure to return to a lower representation rank.

Consider the *hybrid format* from (8.21). If $\mathbf{v}' = \rho_{\text{r-term}}^{\text{hybr}}\big(r', \mathbf{J}, (a_\nu'^{(j)}), (B_j)\big)$ and $\mathbf{v}'' = \rho_{\text{r-term}}^{\text{hybr}}\big(r'', \mathbf{J}, (a_\nu''^{(j)}), (B_j)\big)$ holds with identical bases, the procedure is as above. The coefficients are joined into $(a_\nu^{(j)})$ with $1 \le \nu \le r := r' + r''$. If different bases (B_j') and (B_j'') are involved, these have to be joined via $\mathbf{JoinBases}(B_j', B_j'', r_j, B_j, T'^{(j)}, T''^{(j)})$. Now, $\mathbf{v}''$ can be reformulated by coefficients $a^{(j)} = T''^{(j)} a''^{(j)}$ with respect to the basis B_j. $\mathbf{v}'$ may stay essentially unchanged, since B_j' can be taken as the first part of B_j. Afterwards, $\mathbf{v}'$ and $\mathbf{v}''$ have identical bases and can be treated as before. The total cost in the second case is

$$N_+ = \sum_{j=1}^{d} \left(N_{\text{QR}}(n_j, r_j' + r_j'') + r_j(2r_j'' - 1) \right).$$

13.1.3 Tensor Subspace Representation

Case I (two tensors from $\mathcal{T}_\mathbf{r}$ with same bases). First we consider the case that $\mathbf{v}', \mathbf{v}'' \in \mathcal{T}_\mathbf{r}$ belong to the *same* tensor subspace $\mathbf{U} := \bigotimes_{j=1}^{d} U_j$ with $r_j = \dim(U_j)$. The representation parameters are the coefficient tensors $\mathbf{a}', \mathbf{a}'' \in \mathbb{K}^\mathbf{J}$ with $\mathbf{J} = \times_{j=1}^{d} J_j$, where $i \in J_j = \{1, \ldots, r_j\}$ is associated to the basis vectors $b_i^{(j)}$ (cf. (8.6b)). The addition of $\mathbf{v}', \mathbf{v}'' \in \mathbf{U}$ reduces to the addition $\mathbf{a} := \mathbf{a}' + \mathbf{a}''$ of the coefficient tensors for which the full representation is used. Hence, §13.1.1 yields

$$N_{\text{mem}}(\mathbf{v}) = N_+ = \#\mathbf{J} = \prod_{j=1}^{d} r_j.$$

Case II (two tensors with different bases). Another situation arises, if different tensor subspaces are involved:

$$\mathbf{v}' = \sum_{\mathbf{i} \in \mathbf{J}'} \mathbf{a}_\mathbf{i}' \bigotimes_{j=1}^{d} b_{i_j}'^{(j)} \in \mathbf{U}' := \bigotimes_{j=1}^{d} U_j', \quad \mathbf{v}'' = \sum_{\mathbf{i} \in \mathbf{J}''} \mathbf{a}_\mathbf{i}'' \bigotimes_{j=1}^{d} b_{i_j}''^{(j)} \in \mathbf{U}'' := \bigotimes_{j=1}^{d} U_j''.$$

The sum $\mathbf{v} := \mathbf{v}' + \mathbf{v}''$ belongs to the larger space $\mathbf{U} := \mathbf{U}' + \mathbf{U}'' = \bigotimes_{j=1}^{d} U_j$ with

$$U_j := U_j' + U_j'' = \mathrm{range}(B_j') + \mathrm{range}(B_j''),$$

where $B_j' = [b_1'^{(j)}, \ldots, b_{r_j'}'^{(j)}]$ and $B_j'' = [b_1''^{(j)}, \ldots, b_{r_j''}''^{(j)}]$ are the given bases or frames of $U_j' \subset V_j$ and $U_j'' \subset V_j$, respectively. The further treatment depends on the requirement about $B_j := [b_1^{(j)}, \ldots, b_{r_j}^{(j)}]$.

1. If B_j may be any frame (cf. Remark 8.7d), one can set $r_i := r_j' + r_j''$ and

$$B_j := \left[b_1'^{(j)}, \ldots, b_{r_j'}'^{(j)}, b_1''^{(j)}, \ldots, b_{r_j''}''^{(j)}\right].$$

Then $\mathbf{a} \in \mathbb{K}^{\mathbf{J}}$ with $\mathbf{J} = \times_{j=1}^{d} J_j$ and $J_j = \{1, \ldots, r_j\}$ is obtained by concatenation: $\mathbf{a}_\mathbf{i} = \mathbf{a}_\mathbf{i}'$ for $\mathbf{i} \in \mathbf{J}' \subset \mathbf{J}$, $\mathbf{a}_{\mathbf{r}'+\mathbf{i}} = \mathbf{a}_\mathbf{i}''$ for $\mathbf{i} \in \mathbf{J}''$, where $\mathbf{r}' = (r_1', \ldots, r_d')$. All further entries are defined by zero. There is no arithmetical cost, i.e.,

$$N_+^{\mathcal{T}_\mathbf{r},\mathrm{frame}} = 0,$$

but the memory is largely increased: $N_{\mathrm{mem}}(\mathbf{v}) = \#\mathbf{J} = \prod_{j=1}^{d} r_j$ (note that, in general, $N_{\mathrm{mem}}(\mathbf{v}) \gg N_{\mathrm{mem}}(\mathbf{v}') + N_{\mathrm{mem}}(\mathbf{v}'')$).

2. If B_j should be a basis, we apply $\mathbf{JoinBases}(B_j', B_j'', r_j, B_j, T'^{(j)}, T''^{(j)})$ from (2.35), which determines a basis $B_j = \left[b_1^{(j)}, \ldots, b_{r_j}^{(j)}\right]$ of U_j together with transfer maps $b_k'^{(j)} = \sum_{i=1}^{r} T_{ik}'^{(j)} b_i^{(j)}$ and $b_k''^{(j)} = \sum_{i=1}^{r} T_{ik}''^{(j)} b_i^{(j)}$. It is advantageous to retain one part, say B_j', and to complement B_j' by the linearly independent contributions from B_j'', which leads to $T_{ik}'^{(j)} = \delta_{ik}$. The dimension $r_j = \dim(U_j)$ may take any value in $\max\{r_j', r_j''\} \le r_j \le r_j' + r_j''$. It defines the index sets $J_j = \{1, \ldots, r_j\}$ and $\mathbf{J} = \times_{j=1}^{d} J_j$. If $r_j = r_j' + r_j''$, the memory is as large as in Case 1. The work required by **JoinBases** depends on the representation of the vectors in V_j (cf. §7.5). Set[3] $\mathbf{T}' := \bigotimes_{j=1}^{d} T'^{(j)}$ and $\mathbf{T}'' := \bigotimes_{j=1}^{d} T''^{(j)}$. Lemma 8.9 states that

$$\begin{aligned} \mathbf{v}' &= \rho_{\mathrm{frame}}\left(\mathbf{a}', (B_j')_{1\le j\le d}\right) &= \rho_{\mathrm{frame}}\left(\mathbf{T}'\mathbf{a}', (B_j)_{1\le j\le d}\right), \\ \mathbf{v}'' &= \rho_{\mathrm{frame}}\left(\mathbf{a}'', (B_j'')_{1\le j\le d}\right) &= \rho_{\mathrm{frame}}\left(\mathbf{T}''\mathbf{a}'', (B_j)_{1\le j\le d}\right). \end{aligned}$$

Then $\mathbf{a} := \mathbf{T}'\mathbf{a}' + \mathbf{T}''\mathbf{a}''$ is the resulting coefficient tensor in $\mathbf{v} = \rho_{\mathrm{frame}}(\mathbf{a}, (B_j))$. The cost of **JoinBases** is $\sum_{j=1}^{d} N_{\mathrm{QR}}(n_j, r_j' + r_j'')$, if $V_j = \mathbb{K}^{n_j}$. The update $\mathbf{a} := \mathbf{T}'\mathbf{a}' + \mathbf{T}''\mathbf{a}''$ of the coefficient tensor leads to $2\#\mathbf{J} \sum_{j=1}^{d} r_j$ operations. If $n_j \le n$ and $r_j \le r$, the overall cost is

$$N_+^{\mathcal{T}_\mathbf{r}} \le 2dnr^2 + 2dr^{d+1}.$$

3. In the case of orthonormal bases B_j', B_j'', B_j, one applies **JoinONB** (cf. (2.36)). The coefficient tensors are treated as in Case 2. The cost is as in Item 2.

[3] $\mathbf{T}'$ can be chosen as trivial injection: $T_{\alpha\beta}'^{(j)} = \delta_{\alpha\beta}$.

13.1.4 Hierarchical Representation

Case I (two tensors with identical bases). Assume that both tensors $\mathbf{v}', \mathbf{v}'' \in \mathcal{H}_\mathfrak{r}$ are represented by the same data $\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, (B_j)_{j\in D}\big)$, only their coefficients $c'^{(D)}$ and $c''^{(D)}$ differ. Then the sum $\mathbf{v} := \mathbf{v}' + \mathbf{v}''$ is characterised by $\mathbf{v} = \rho_{\mathrm{HTR}}\left(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\right)$ with the coefficient vector $c^{(D)} := c'^{(D)} + c''^{(D)} \in \mathbb{K}^{r_D}$. The cost is marginal:

$$N_+^{\mathcal{H}_\mathfrak{r},\text{Case I}} = r_D.$$

Case II (two tensors with different bases). Here we assume that both terms $\mathbf{v}' \in \mathcal{H}_{\mathfrak{r}'}$ and $\mathbf{v}'' \in \mathcal{H}_{\mathfrak{r}''}$ use the same dimension partition tree T_D:

$$\mathbf{v}' = \rho_{\mathrm{HTR}}\big(T_D, (\mathbf{C}'_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c'^{(D)}, (B'_j)_{j\in D}\big),$$
$$\mathbf{v}'' = \rho_{\mathrm{HTR}}\big(T_D, (\mathbf{C}''_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c''^{(D)}, (B''_j)_{j\in D}\big).$$

First we consider the involved hierarchical subspace families from Definition 11.8a. Let $\{\mathbf{U}'_\alpha\}_{\alpha\in T_D}$ and $\{\mathbf{U}''_\alpha\}_{\alpha\in T_D}$ be the subspaces associated with $\mathbf{v}'$ and $\mathbf{v}''$, respectively. The sum $\mathbf{v} := \mathbf{v}' + \mathbf{v}''$ belongs to $\{\mathbf{U}_\alpha\}_{\alpha\in T_D}$ with $\mathbf{U}_\alpha := \mathbf{U}'_\alpha + \mathbf{U}''_\alpha$.

As in §13.1.3, we have to determine bases of the spaces $\mathbf{U}'_\alpha + \mathbf{U}''_\alpha$. This procedure has been described in §11.5. According to Remark 11.67d, the cost is bounded by $\leq 8dr^2(r^2+n)$, where $r := \max r_j$ and $n := \max n_j$. Having a common basis representation, we can apply Case I from above. Hence, the cost is

$$N_+^{\mathcal{H}_\mathfrak{r},\text{Case II}} \leq 8dnr^2 + 8dr^4.$$

An increase of storage is caused by the fact that, in the worst case, the subspaces $\mathbf{U}_\alpha = \mathbf{U}'_\alpha + \mathbf{U}''_\alpha$ have dimension $\dim(\mathbf{U}_\alpha) = \dim(\mathbf{U}'_\alpha) + \dim(\mathbf{U}''_\alpha)$. In particular, $\dim(\mathbf{U}_D) \geq 2$ can be reduced to 1 without loss of accuracy. Possibly, further subspaces $\mathbf{U}_\alpha$ can be reduced by the truncation procedure of §11.4.2.

13.2 Entry-wise Evaluation

For $V_j = \mathbb{K}^{n_j}$, the tensor $\mathbf{v} \in \mathbf{V} = \bigotimes_{j=1}^d V_j$ has entries $\mathbf{v}_\mathbf{i}$ with $\mathbf{i} = (i_1, \ldots, i_d) \in \mathbf{I}$ and the evaluation $\varLambda_\mathbf{i} : \mathbf{v} \mapsto \mathbf{v}_\mathbf{i} \in \mathbb{K}$ is of interest.

In connection with variants of the cross approximation (cf. §15), it is necessary to evaluate $\mathbf{v}_\mathbf{i}$ not only for one index $\mathbf{i}$, but for all $\mathbf{k}$ in the so-called fibre

$$\mathcal{F}(j,\mathbf{i}) := \big\{\mathbf{k} \in \mathbf{I} = I_1 \times \ldots \times I_d : \mathbf{k}_\ell = \mathbf{i}_\ell \text{ for } \ell \in \{1,\ldots,d\}\backslash\{j\}\big\}.$$

Note that the component k_j of $\mathbf{k} \in \mathcal{F}(j,\mathbf{i})$ takes all values from I_j, while all other components are fixed. The challenge is to perform the simultaneous evaluation cheaper than $\#I_j$ times the cost of a single evaluation.

The entry-wise evaluation may be viewed as scalar product by the unit vector $\mathbf{e}^{(\mathbf{i})} \in \mathbf{V}$ with $\mathbf{e}^{(\mathbf{i})}_{\mathbf{j}} = \delta_{\mathbf{ij}}$ ($\mathbf{i}, \mathbf{j} \in \mathbf{I}$), since $\mathbf{v}_{\mathbf{i}} = \langle \mathbf{v}, \mathbf{e}^{(\mathbf{i})} \rangle$. Therefore, the evaluation of the scalar product with an elementary tensor is closely related.

Full representation of $\mathbf{v}$ need not be discussed, since then $\mathbf{v}_{\mathbf{i}}$ is directly available, i.e., $N^{\text{r-term}}_{\text{eval}} = 0$.

13.2.1 r-Term Representation

If $\mathbf{v}$ is represented in r-term format $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v_{\nu}^{(j)} \in \mathcal{R}_r$, the entry $\mathbf{v}_{\mathbf{i}}$ equals $\sum_{\nu=1}^{r} \prod_{j=1}^{d} (v_{\nu}^{(j)})_{i_j}$. Its computation requires

$$N^{\text{r-term}}_{\text{eval}} = rd - 1$$

arithmetical operations.

The cost of the evaluation for all indices $\mathbf{k} \in \mathcal{F}(j, \mathbf{i})$ is

$$N^{\text{r-term}}_{\text{eval}}(\mathcal{F}(j, \mathbf{i})) = r\,(d + 2\#I_j - 2) - \#I_j.$$

Here, the products $\prod_{\ell \in \{1,\ldots,d\} \backslash \{j\}} (v_{\nu}^{(\ell)})_{i_\ell}$ $(1 \leq \nu \leq r)$ are computed first.

13.2.2 Tensor Subspace Representation

The tensor subspace representation $\mathbf{v} = \sum_{\mathbf{k} \in \mathbf{J}} \mathbf{a}_{\mathbf{k}} \bigotimes_{j=1}^{d} b_{k_j}^{(j)}$ yields

$$\mathbf{v}[i_1, \ldots, i_d] = \mathbf{v}_{\mathbf{i}} = \sum_{\mathbf{k} \in \mathbf{J}} \mathbf{a}_{\mathbf{k}} \prod_{j=1}^{d} (b_{k_j}^{(j)})_{i_j}$$

with $\mathbf{J} = J_1 \times \ldots \times J_d$ and $J_j := \{1 \leq i \leq r_j\}$. The evaluation starts with summation over k_1 yielding a reduced coefficient tensor $\mathbf{a}[k_2, \ldots, k_d]$ etc. Then the arithmetical operations amount to

$$N^{\mathcal{T}_{\mathbf{r}}}_{\text{eval}} \;=\; \sum_{\ell=1}^{d} (2r_\ell - 1) \prod_{j=\ell+1}^{d} r_j \;<\; 2\#\mathbf{J}.$$

Summation in the order $k_1, k_2, \ldots, k_d$ is optimal, if $r_1 \geq r_2 \geq \ldots \geq r_d$. Otherwise, the order of summation should be changed. If $r_j \leq r$, the cost is about $N^{\mathcal{T}_{\mathbf{r}}}_{\text{eval}} \leq 2r^d$.

For the simultaneous evaluation, the summations over all k_ℓ are to be performed in such an order that $\ell = j$ is the last one. For $j = d$, the cost is

$$N^{\mathcal{T}_{\mathbf{r}}}_{\text{eval}}(\mathcal{F}(j, \mathbf{i})) \;=\; \sum_{\ell=1}^{d-1} (2r_\ell - 1) \prod_{j=\ell+1}^{d} r_j + \#I_d\,(2r_d - 1)\,.$$

13.2.3 Hierarchical Representation

For $\alpha \subset D = \{1, \ldots, d\}$ the index $\mathbf{i}_\alpha$ belongs to $\mathbf{I}_\alpha = \times_{j \in \alpha} I_j$. The evaluation of the $\mathbf{i}_\alpha$ entry

$$\beta_\ell^{(\alpha)} := \mathbf{b}_\ell^{(\alpha)}[\mathbf{i}_\alpha] = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)} \, \mathbf{b}_i^{(\alpha_1)}[\mathbf{i}_{\alpha_1}] \mathbf{b}_j^{(\alpha_2)}[\mathbf{i}_{\alpha_2}] \tag{13.1}$$

of the basis vector $\mathbf{b}_\ell^{(\alpha)}$ is performed recursively from the leaves to the root:

```
procedure eval*($\alpha, \mathbf{i}$);
for $\ell := 1$ to $r_\alpha$ do
if $\alpha = \{j\}$ then $\beta_\ell^{(\alpha)} := b_\ell^{(j)}[i_j]$ else        {leaf}
begin eval*($\alpha_1, \mathbf{i}$); eval*($\alpha_2, \mathbf{i}$);          $\{\alpha_1, \alpha_2$ sons of $\alpha\}$
    $\beta_\ell^{(\alpha)} := \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)} \beta_i^{(\alpha_1)} \beta_j^{(\alpha_2)}$    {non-leaf vertex, cf. (13.1)}
end;
```

(cf. (11.26)). The evaluation of $\mathbf{v}[\mathbf{i}]$ is implemented by

```
function eval($\mathbf{v}, \mathbf{i}$);
begin eval*($D, \mathbf{i}$); $s := 0$;
    for $\ell := 1$ to $r_D$ do $s := s + c_\ell^{(D)} \cdot \beta_\ell^{(D)}$;
    eval := $s$
end;
```

The asymptotic computational cost is

$$N_{\text{eval}}^{\mathcal{H}_\mathfrak{r}} = 2 \sum_{\alpha \in T_D \setminus \mathcal{L}(T_D)} r_\alpha r_{\alpha_1} r_{\alpha_2} \qquad (\alpha_1, \alpha_2 \text{ sons of } \alpha)\,.$$

For $r_\alpha \leq r$, the cost is bounded by $N_{\text{eval}}^{\mathcal{H}_\mathfrak{r}} \leq 2dr^3$.

The cost of the simultaneous evaluation at $\mathcal{F}(j, \mathbf{i})$ amounts to

$$N_{\text{eval}}^{\mathcal{H}_\mathfrak{r}}(\mathcal{F}(j, \mathbf{i})) = N_{\text{eval}}^{\mathcal{H}_\mathfrak{r}} + 2\#I_j \sum_{\alpha \in T_D \setminus \mathcal{L}(T_D) \text{ with } j \in \alpha_1 \in S(\alpha)} r_\alpha r_{\alpha_1}.$$

The latter summation involves all non-leaf vertices α with a son α_1 containing j. The total cost is bounded by $2r^2 \left[dr + (depth(T_D) - 1)\, \#I_j\right]$ (cf. (11.7)). Note that the tree T_D can be constructed such that $depth(T_D) \approx \log_2 d$.

13.2.4 Matrix Product Representation

The TT format is introduced in (12.1a) by means of a representation of an entry $\mathbf{v}_\mathbf{i}$. Since the data $v_{k_{j-1} i_j k_j}^{(j)}$ are already separated with respect to i_j, only the matrices

$v^{(j)}_{k_{j-1}k_j}[i_j]$ from (12.2b) enter the computation. Correspondingly, the evaluation of the right-hand side requires less operations than in §13.2.3:

$$N^{\mathsf{TT}}_{\text{eval}} = 2\sum_{\ell=1}^{d-2} \rho_\ell\, \rho_{\ell+1}. \tag{13.2}$$

For $\rho_\ell \leq \rho$, this is $N^{\mathsf{TT}}_{\text{eval}} \approx 2\,(d-2)\,\rho^2$.

For the simultaneous evaluation in the case of $j \in \{2,\ldots,d-1\}$, perform the product of the matrices $V^{(\ell,i_\ell)}$ in (12.2b) such that $V_I \cdot V^{(j,k_j)} \cdot V_{II}$ holds with vectors $V_I \in \mathbb{K}^{\rho_{j-1}}$ and $V_{II} \in \mathbb{K}^{\rho_j}$. Its evaluation for all $k_j \in J_j$ yields

$$\begin{aligned} N^{\mathsf{TT}}_{\text{eval}}(\mathcal{F}(j,\mathbf{i})) &= \rho_{j-1}\rho_j\,(2\rho_j+1) + \sum_{\ell=1}^{j-2}(2\rho_\ell-1)\,\rho_{\ell+1} + \sum_{\ell=j}^{d-2}(2\rho_{\ell+1}-1)\,\rho_\ell \\ &\approx 2\Big(\rho_{j-1}\rho_j^2 + \sum\nolimits_{\ell=1}^{d-2}\rho_\ell\rho_{\ell+1}\Big). \end{aligned}$$

The cases $j=1$ and $j=d$ are left to the reader.

13.3 Scalar Product

Given pre-Hilbert spaces V_j with scalar product $\langle\cdot,\cdot\rangle_j$, the induced scalar product in $\mathbf{V} = {}_a\bigotimes_{j=1}^d V_j$ is defined in §4.5.1. The corresponding norms of V_j and $\mathbf{V}$ are denoted by $\|\cdot\|_j$ and $\|\cdot\|$. We suppose that the computation of $\langle u,v\rangle_j$ is feasible, at least for $u,v \in U_j \subset V_j$ from the relevant subspace U_j, and that its computational cost is

$$N_j \tag{13.3}$$

(cf. Remark 7.12). In the case of function spaces, $\langle u,v\rangle_j$ for $u,v \in U_j$ may be given analytically or approximated by a quadrature formula[4], provided that U_j is a subspace of sufficiently smooth functions.

The scalar product $\langle\mathbf{u},\mathbf{v}\rangle$ is considered in two situations. In the general case, both $\mathbf{u}$ and $\mathbf{v}$ are tensors represented in one of the formats. A particular, but also important case is the scalar product of $\mathbf{u}$—represented in one of the formats—and an elementary tensor

$$\mathbf{v} = \bigotimes_{j=1}^d v^{(j)} \tag{13.4}$$

represented by the vectors $v^{(j)} \in V_j$ (1-term format).

A related problem is the computation of the *partial scalar product* defined in §4.5.4. It is important since the left or right singular vectors of the singular value decomposition can be obtained via the partial scalar product (see Lemma 5.13).

[4] Whether approximations of the scalar product are meaningful or not, depends on the application. For instance, a quadrature formula with n quadrature points cannot be used to determine an (approximately) orthogonal system of more than n vectors.

13.3.1 Full Representation

For $\mathbf{V} = \mathbb{K}^{\mathbf{I}}$ with $\mathbf{I} = \times_{j=1}^{d} I_j$, the Euclidean scalar product $\langle \mathbf{u}, \mathbf{v} \rangle$ is to be computed by $\sum_{\mathbf{i}\in\mathbf{I}} \mathbf{u}_{\mathbf{i}}\overline{\mathbf{v}_{\mathbf{i}}}$ so that $N_{\langle\cdot,\cdot\rangle} = 2\#\mathbf{I}$. The computation may be much cheaper for sparse tensors. The full representation by a function is useful only in connection with a quadrature formula. The case of (13.4) is even a bit more expensive.

The partial scalar product in $\mathbb{K}^{\mathbf{I}}$ depends on the decomposition of $\{1, \ldots, d\}$ into disjoint and non-empty sets α and $\alpha^c := \{1, \ldots, d\}\backslash\alpha$. This induces a partition of $\mathbf{I}$ into $\mathbf{I} = \mathbf{I}_\alpha \times \mathbf{I}_{\alpha^c}$ with $\mathbf{I}_\alpha = \times_{j\in\alpha} I_j$ and $\mathbf{I}_{\alpha^c} = \times_{j\in\alpha^c} I_j$. Then $\mathbf{w} := \langle \mathbf{u}, \mathbf{v} \rangle_{\alpha^c} \in \mathbb{K}^{\mathbf{I}_\alpha} \otimes \mathbb{K}^{\mathbf{I}_\alpha}$ is defined by the entries[5]

$$\mathbf{w}_{\mathbf{i}',\mathbf{k}'} = \sum\nolimits_{\mathbf{i}''\in\mathbf{I}_{\alpha^c}} \mathbf{u}_{\mathbf{i}',\mathbf{i}''}\overline{\mathbf{v}_{\mathbf{k}',\mathbf{i}''}} \qquad \text{for } \mathbf{i}', \mathbf{k}' \in \mathbf{I}_\alpha.$$

The computational cost per entry is $2\#\mathbf{I}_{\alpha^c}$. Since $\mathbf{w}$ has $(\#\mathbf{I}_\alpha)^2$ entries, the overall cost is $N_{\langle\cdot,\cdot\rangle,\mathbf{I}_\alpha} = 2\,(\#\mathbf{I}_\alpha)^2\,\#\mathbf{I}_{\alpha^c}$.

13.3.2 r-Term Representation

For elementary tensors, the definition of $\langle\cdot,\cdot\rangle$ yields $\langle \mathbf{u}, \mathbf{v} \rangle = \prod_{j=1}^{d} \left\langle u^{(j)}, v^{(j)} \right\rangle_j$. Combining all terms from $\mathbf{u} \in \mathcal{R}_{r_\mathbf{u}}$ and $\mathbf{v} \in \mathcal{R}_{r_\mathbf{v}}$, we obtain the following result.

Remark 13.1. The scalar product of $\mathbf{u} \in \mathcal{R}_{r_\mathbf{u}}$ and $\mathbf{v} \in \mathcal{R}_{r_\mathbf{v}}$ costs

$$N_{\langle\cdot,\cdot\rangle}^{\text{r-term}} = r_\mathbf{u} r_\mathbf{v} \big(d + \sum\nolimits_{j=1}^{d} N_j\big)$$

operations with N_j from (13.3). The case of (13.4) is included by the choice $r_\mathbf{v} = 1$.

For *partial* scalar products we use the notations $\alpha, \alpha^c \subset \{1, \ldots, d\}$ and $\mathbf{w} := \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{I}_{\alpha^c}}$ as in §13.3.1. First, let $\mathbf{u} = \bigotimes_{j=1}^{d} u^{(j)}$ and $\mathbf{v} = \bigotimes_{j=1}^{d} v^{(j)}$ be elementary tensors. Then

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\alpha^c} = \Big(\prod_{j\in\alpha^c} \left\langle u^{(j)}, v^{(j)} \right\rangle_j \Big) \Big(\bigotimes_{j\in\alpha} u^{(j)} \Big) \otimes \overline{\Big(\bigotimes_{j\in\alpha} v^{(j)} \Big)} \in \mathbf{V}_\alpha \otimes \mathbf{V}_\alpha$$

with $\mathbf{V}_\alpha = \bigotimes_{j\in\alpha} V_j$ is again an elementary tensor. The same considerations as above lead to the next remark. Since $\langle \mathbf{v}, \mathbf{v} \rangle_{\alpha^c}$ (i.e., $\mathbf{u} = \mathbf{v}$) appears for the left-sided singular value decomposition of $\mathcal{M}_\alpha(\mathbf{v})$, this case is of special interest.

Remark 13.2. (a) The partial scalar product $\mathbf{w} := \langle \mathbf{u}, \mathbf{v} \rangle_{\alpha^c}$ of $\mathbf{u} \in \mathcal{R}_{r_\mathbf{u}}$ and $\mathbf{v} \in \mathcal{R}_{r_\mathbf{v}}$ costs $N_{\langle\cdot,\cdot\rangle} = r_\mathbf{u} r_\mathbf{v} (\#\alpha^c + \sum_{j\in\alpha^c} N_j)$ operations. The tensor $\mathbf{w} \in \mathbf{V}_\alpha \otimes \mathbf{V}_\alpha$ is given in the format $\mathcal{R}_r$ with $r := r_\mathbf{u} r_\mathbf{v}$.
(b) Because of symmetry, $N_{\langle\cdot,\cdot\rangle,\alpha^c}^{\mathcal{R}_r}$ reduces to $\frac{r_\mathbf{v}(r_\mathbf{v}+1)}{2} \left(\#\alpha^c + \sum_{j\in\alpha^c} N_j\right)$ for the computation of $\langle \mathbf{v}, \mathbf{v} \rangle_{\alpha^c}$.

[5] The notation $\mathbf{u}_{\mathbf{i}',\mathbf{i}''}$ assumes that $\alpha^c = \{j^*, \ldots, d\}$ for some $1 \leq j^* \leq d$. Otherwise, $\mathbf{u}_{\pi(\mathbf{i}',\mathbf{i}'')}$ with a suitable permutation π is needed. However, this does not effect the computational cost.

13.3.3 Tensor Subspace Representation

Case I (two tensors from $\mathcal{T}_\mathbf{r}$ with same bases). The easiest case is given by $\mathbf{u} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}^\mathbf{u}_\mathbf{i} \bigotimes_{j=1}^d b^{(j)}_{i_j} \in \mathcal{T}_\mathbf{r}$ and $\mathbf{v} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}^\mathbf{v}_\mathbf{i} \bigotimes_{j=1}^d b^{(j)}_{i_j} \in \mathcal{T}_\mathbf{r}$ belonging to the same subspace $\mathbf{U} := \bigotimes_{j=1}^d U_j$ with orthonormal bases $B_j = [b^{(j)}_1, \ldots, b^{(j)}_{r_j}]$ of U_j. Then

$$\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{a}^\mathbf{u}, \mathbf{a}^\mathbf{v}\rangle_\mathbf{J}$$

holds, where the latter is the Euclidean scalar product of the coefficient tensors in $\mathbb{K}^\mathbf{J}$. The cost is $N_{\langle\cdot,\cdot\rangle} = 2\#\mathbf{J} = 2\prod_{j=1}^d r_j$ (cf. §13.3.1).

Case II (tensor from $\mathcal{T}_\mathbf{r}$ and elementary tensor). If $\mathbf{u} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}_\mathbf{i} \bigotimes_{j=1}^d b^{(j)}_{i_j} \in \mathcal{T}_\mathbf{r}$, while $\mathbf{v}$ is an elementary tensor, $\langle \mathbf{u}, \mathbf{v}\rangle = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}_\mathbf{i} \bigotimes_{j=1}^d \langle b^{(j)}_{i_j}, v^{(j)}\rangle_j$ requires

$$\sum_{j=1}^d (2r_j - 1) \prod_{\ell=j+1}^d r_j + \sum_{j=1}^d r_j N_j$$

operations. The second sum corresponds to the scalar products $\beta^{(j)}_{i_j} := \langle b^{(j)}_{i_j}, v^{(j)}\rangle_j$, where N_j is the cost of a scalar product in V_j (cf. (13.3)). Performing first the summation of $\sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}_\mathbf{i} \prod_{j=1}^d \beta^{(j)}_{i_j}$ over $1 \le i_1 \le r_1$ for all combinations of $i_2, i_3, \ldots, i_d$, we obtain $\sum_{i_1=1}^{r_1} \mathbf{a}_\mathbf{i} \beta^{(1)}_{i_1}$ with the cost $(2r_1 - 1)\, r_2 \cdots r_d$. Proceeding with summation over $i_2, \ldots, i_d$, we obtain the cost given above.

Case III (tensors from $\mathcal{T}_{\mathbf{r}'}$ and $\mathcal{T}_{\mathbf{r}''}$). If $\mathbf{u}' = \sum_{\mathbf{i}\in\mathbf{J}'} \mathbf{a}'_\mathbf{i} \bigotimes_{j=1}^d b'^{(j)}_{i_j} \in \mathcal{T}_{\mathbf{r}'}$ and $\mathbf{u}'' = \sum_{\mathbf{i}\in\mathbf{J}''} \mathbf{a}''_\mathbf{i} \bigotimes_{j=1}^d b''^{(j)}_{i_j} \in \mathcal{T}_{\mathbf{r}''}$ use different bases, the computation of

$$\langle \mathbf{u}', \mathbf{u}''\rangle = \sum_{\mathbf{i}\in\mathbf{J}'} \sum_{\mathbf{k}\in\mathbf{J}''} \mathbf{a}'_\mathbf{i}\, \overline{\mathbf{a}''_\mathbf{k}} \prod_{j=1}^d \left\langle b'^{(j)}_{i_j}, b''^{(j)}_{k_j} \right\rangle_j \tag{13.5}$$

requires

$$N^{\mathcal{T}_\mathbf{r}}_{\langle\cdot,\cdot\rangle} = \sum_{j=1}^d \left(2r''_j + 1\right) r'_j \prod_{\ell=j+1}^d r'_\ell r''_\ell + \sum_{j=1}^d r'_j r''_j N_j$$

operations.

Remark 13.3. Assume $n_j \le n$ and $r_j, r'_\ell, r''_\ell \le r$. Then the asymptotic costs of Cases I-III can be estimated by

$$\text{I: } r^d, \qquad \text{II: } 2\left(r^d + dnr\right), \qquad \text{III: } 2\left(r^{2d} + dnr^2\right).$$

An alternative way for Case III is to transform $\mathbf{u}'$ and $\mathbf{u}''$ into a representation with a common orthonormal basis B_j as explained in §8.6.3. The expense is $8dnr^2 + 2dr^{d+1}$ (cf. Remark 8.42). Having a common basis of dimension $\le 2r$, we can apply Case I. Hence, the total cost is

$$\text{III': } N^{\mathcal{T}_\mathbf{r}}_{\langle\cdot,\cdot\rangle} = 8dnr^2 + 2dr^{d+1} + (2r)^d.$$

This leads to the following remark:

Remark 13.4. For Case III with $n < \frac{1}{3d}r^{2d-2} - \frac{1}{3}r^{d-1} - \frac{1}{6d}2^d r^{d-2}$, it is advantageous first to transform $\mathbf{u}'$ and $\mathbf{u}''$ into a representation with common orthonormal bases B_j. The cost is $O(dnr^2 + r^d \min\{r^d, 2^d + dr\})$.

In the case of a *partial scalar product* $\mathbf{w} := \langle \mathbf{u}, \mathbf{v} \rangle_{\alpha^c} \in \mathbb{K}^{\mathbf{I}_\alpha} \otimes \mathbb{K}^{\mathbf{I}_\alpha}$, we have similar cases as before.

Case I (two tensors from $\mathcal{T}_\mathbf{r}$ with same bases). Let $\mathbf{u} = \sum_{\mathbf{i}\in\mathbf{J}_\alpha} \mathbf{a}_\mathbf{i}^\mathbf{u} \bigotimes_{j=1}^d b_{i_j}^{(j)}$ and $\mathbf{v} = \sum_{\mathbf{i}\in\mathbf{J}_\alpha} \mathbf{a}_\mathbf{i}^\mathbf{v} \bigotimes_{j=1}^d b_{i_j}^{(j)}$. Again, under the assumption of orthonormal bases, the partial scalar product can be applied to the coefficient tensors:

$$\mathbf{w} = \sum_{(\mathbf{i},\mathbf{k})\in\mathbf{J}_\alpha\times\mathbf{J}_\alpha} \mathbf{c}_{\mathbf{i},\mathbf{k}} \bigotimes_{j\in\alpha} b_{i_j}^{(j)} \otimes \bigotimes_{j\in\alpha} b_{k_j}^{(j)} \qquad \text{with } \mathbf{c} := \langle \mathbf{a}^\mathbf{u}, \mathbf{a}^\mathbf{v} \rangle_{\alpha^c}.$$

Therefore, the cost is given by $N_{\langle\cdot,\cdot\rangle,\alpha^c}^{\mathcal{T}_\mathbf{r}} = 2\left(\#\mathbf{J}_\alpha\right)^2 \#\mathbf{J}_{\alpha^c}$ (cf. §13.3.1). Note that the resulting tensor $\mathbf{w} \in \mathbb{K}^{\mathbf{I}_\alpha} \otimes \mathbb{K}^{\mathbf{I}_\alpha}$ is again represented in tensor subspace format.

The most important case is $\mathbf{u} = \mathbf{v}$ and $\alpha = \{j\}$. Then,

$$N_{\langle\cdot,\cdot\rangle,\alpha^c}^{\mathcal{T}_\mathbf{r}} = (r_j + 1) \prod_{k=1}^d r_k \qquad \text{for } \mathbf{u} = \mathbf{v} \text{ and } \alpha = \{j\} \text{ with } r_j = \#J_j. \tag{13.6}$$

Case II (tensors from $\mathcal{T}_{\mathbf{r}'}$ and $\mathcal{T}_{\mathbf{r}''}$). Now we assume that $\mathbf{v}' = \sum_{\mathbf{i}\in\mathbf{J}'_\alpha} \mathbf{a}'_\mathbf{i} \bigotimes_{j=1}^d b_{i_j}^{\prime(j)}$ and $\mathbf{v}'' = \sum_{\mathbf{i}\in\mathbf{J}''_\alpha} \mathbf{a}''_\mathbf{i} \bigotimes_{j=1}^d b_{i_j}^{\prime\prime(j)}$ use not only different bases, but also different subspaces of possibly different dimensions. Basing the computation of the partial scalar product $\mathbf{w} := \langle \mathbf{v}', \mathbf{v}'' \rangle_{\mathbf{I}_{\alpha^c}}$ on the identity[6]

$$\mathbf{w} = \sum_{\mathbf{i}'\in\mathbf{J}'_\alpha} \sum_{\mathbf{k}'\in\mathbf{J}''_\alpha} \underbrace{\left[\sum_{\mathbf{i}''\in\mathbf{J}'_{\alpha^c}} \sum_{\mathbf{k}''\in\mathbf{J}''_{\alpha^c}} \mathbf{a}'_{\mathbf{i}',\mathbf{i}''} \mathbf{a}''_{\mathbf{k}',\mathbf{k}''} \prod_{j\in\alpha^c} \left\langle b_{i''_j}^{\prime(j)}, b_{k''_j}^{\prime\prime(j)} \right\rangle_j \right] \bigotimes_{j\in\alpha}}_{=:\mathbf{b}_{\mathbf{i}',\mathbf{k}'}} b_{i'_j}^{\prime(j)} \otimes \bigotimes_{j\in\alpha} b_{k'_j}^{\prime\prime(j)},$$

we need

$$N_{\langle\cdot,\cdot\rangle,\alpha^c}^{\mathcal{T}_\mathbf{r}} = 2\prod_{j=1}^d r'_j r''_j + \sum_{j\in\alpha^c} r'_j r''_j N_j + \text{lower order}$$

operations for the evaluation of the coefficient tensor $\mathbf{b} \in \mathbb{K}^{\mathbf{J}_\alpha} \otimes \mathbb{K}^{\mathbf{J}_\alpha}$. Here, we assume that all $r'_j = \#J'_j$ and $r''_j = \#J''_j$ are of comparable size.

The alternative approach is to determine common orthonormal bases for $j \in \alpha^c$ requiring $\#\alpha^c\left(8nr^2 + 2r^{d+1}\right)$ operations (assuming common bounds r and n for all directions). Then, Case I can be applied. The estimate of the total cost by

$$N_{\langle\cdot,\cdot\rangle,\alpha^c}^{\mathcal{T}_\mathbf{r}} \leq \#\alpha^c\left(8nr^2 + 2r^{d+1}\right) + 2^{\#\alpha^c} r^{d+\#\alpha}$$

shows that the second approach is cheaper under the assumptions $N_j = 2n_j - 1$ and $n \leq r^{2d-2}/\left(3\#\alpha^c\right)$ up to lower order terms.

[6] Note that the index $\mathbf{i} \in \mathbf{J}' = \times_{j=1}^d J'_j$ of $\mathbf{a}'_\mathbf{i} = \mathbf{a}'_{\mathbf{i}',\mathbf{i}''}$ is split into the pair $(\mathbf{i}', \mathbf{i}'')$, where $\mathbf{i}' \in \mathbf{J}_\alpha = \times_{j\in\alpha} J_j$ and $\mathbf{i}'' \in \mathbf{J}_{\alpha^c} = \times_{j\in\alpha^c} J_j$. Similarly for $\mathbf{k} \in \mathbf{J}''$.

13.3.4 Hybrid Format

The hybrid format $\mathbf{u} = \rho_{\mathrm{hybr}}(\cdot)$ from §8.2.4 implies that $\mathbf{u} = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}^{\mathbf{u}}_{\mathbf{i}} \bigotimes_{j=1}^{d} b^{(j)}_{i_j} \in \mathcal{T}_{\mathbf{r}}$, where $\mathbf{a}^{\mathbf{u}} \in \mathcal{R}_r(\mathbb{K}^{\mathbf{J}})$ is represented in r-term format. The cost of a scalar product in $\mathbb{K}^{J_j}$ is denoted by $\hat{N}_j$ (usually, this is $2\#J_j$).

Again, we distinguish the cases from above.

Case I (two hybrid tensors with same bases). Here, $\mathbf{u}, \mathbf{v} \in \mathcal{T}_{\mathbf{r}}$ are given with identical subspace $\mathbf{U} := \bigotimes_{j=1}^{d} U_j$ and orthonormal bases $B_j = [b^{(j)}_1, \ldots, b^{(j)}_{r_j}]$ of U_j. Again, the identity $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{a}^{\mathbf{u}}, \mathbf{a}^{\mathbf{v}} \rangle_{\mathbf{J}}$ holds. Since $\mathbf{a}^{\mathbf{u}}, \mathbf{a}^{\mathbf{v}} \in \mathcal{R}_{r_{\mathbf{u}}}(\mathbb{K}^{\mathbf{J}})$ and $\mathbf{a}^{\mathbf{v}} \in \mathcal{R}_{r_{\mathbf{v}}}(\mathbb{K}^{\mathbf{J}})$, the latter scalar product can be performed as discussed in §13.3.2. The cost is

$$r_{\mathbf{u}} r_{\mathbf{v}} \left(d + \sum_{j=1}^{d} \hat{N}_j \right).$$

Case II (hybrid tensor and elementary tensor). Let $\mathbf{u}$ be of hybrid format, while $\mathbf{v}$ is the elementary tensor (13.4). As in §13.3.3, the scalar products $\beta^{(j)}_{i_j} := \langle b^{(j)}_{i_j}, v^{(j)} \rangle_j$ are to be computed. Since $\mathbf{a} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} a^{(j)}_{\nu} \in \mathcal{R}_r$, we obtain $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{\mathbf{i}\in\mathbf{J}} \mathbf{a}_{\mathbf{i}} \prod_{j=1}^{d} \beta^{(j)}_{i_j} = \sum_{\nu} \prod_{j=1}^{d} \left\langle a^{(j)}_{\nu}, \beta^{(j)} \right\rangle$ involving the $\mathbb{K}^{\mathbf{J}}$-scalar product with $\beta^{(j)} = (\beta^{(j)}_i)_{i\in J_j}$. The total cost is

$$\sum_{j=1}^{d} r_j N_j + r \sum_{j=1}^{d} \hat{N}_j.$$

Case III (hybrid tensors with different bases). For hybrid tensors $\mathbf{u}', \mathbf{u}''$ with coefficient tensors $\mathbf{a}' = \sum_{\nu=1}^{r'} \bigotimes_{j=1}^{d} \mathbf{a}'^{(j)}_{\nu}$ and $\mathbf{a}'' = \sum_{\nu=1}^{r''} \bigotimes_{j=1}^{d} \mathbf{a}''^{(j)}_{\nu}$, the right-hand side in (13.5) can be written as

$$\langle \mathbf{u}', \mathbf{u}'' \rangle = \sum_{\nu,\mu} \prod_{j=1}^{d} \left[\sum_{i_j \in J'_j} \sum_{k_j \in J''_j} \mathbf{a}'^{(j)}_{\nu}[i_j] \, \overline{\mathbf{a}''^{(j)}_{\mu}[k_j]} \left\langle b'^{(j)}_{i_j}, b''^{(j)}_{k_j} \right\rangle_j \right] \tag{13.7}$$

and requires

$$N^{\mathcal{T}_{\mathbf{r}}}_{\langle\cdot,\cdot\rangle} = \sum_{j=1}^{d} r'_j r''_j \left(N_j + 3r'r'' \right)$$

operations, which are bounded by

$$2dnr^2 + 2r^4, \qquad \text{if } r'_j, r''_j, r', r'' \leq r \text{ and } N_j \leq 2n.$$

Alternatively, we introduce common bases. According to Remark 8.43, the cost including the transformation of the coefficient tensors $\mathbf{a}', \mathbf{a}''$ is $8dnr^2 + 2dr^3$. The addition of $\mathbf{a}' \in \mathcal{R}_{r'}$ and $\mathbf{a}'' \in \mathcal{R}_{r''}$ requires no arithmetical work, but increases the representation rank: $r = r' + r''$.

Remark 13.5. For Case III with $n \geq (r^2/d - r)/3$, the first variant based on (13.7) is more advantageous. The cost is bounded by $2r^2 \left(dn + r^2 \right)$.

13.3.5 Hierarchical Representation

We start with the scalar product $\langle \mathbf{u}, \mathbf{v} \rangle$ of $\mathbf{u} \in \mathcal{H}_{\mathfrak{r}}$ and an elementary tensor $\mathbf{v} = \bigotimes_{j=1}^{d} v^{(j)}$ from (13.4). Define $\mathbf{v}^{(\alpha)} := \bigotimes_{j\in\alpha} v^{(j)}$ and use the recursion

$$\left\langle \mathbf{b}_{\ell}^{(\alpha)}, \mathbf{v}^{(\alpha)} \right\rangle_{\alpha} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)} \left\langle \mathbf{b}_{i}^{(\alpha_1)}, \mathbf{v}^{(\alpha_1)} \right\rangle_{\alpha_1} \left\langle \mathbf{b}_{j}^{(\alpha_2)}, \mathbf{v}^{(\alpha_2)} \right\rangle_{\alpha_2} \tag{13.8}$$

(cf. (11.24)), where α_1, α_2 are the sons of α.

Remark 13.6. The computation of all $\left\langle \mathbf{b}_{\ell}^{(\alpha)}, \mathbf{v}^{(\alpha)} \right\rangle_{\alpha}$ for $\alpha \in T_D$, $1 \le \ell \le r_\alpha$, can be performed by

$$\sum_{\alpha\in T_D\backslash\mathcal{L}(T_D)} \left\{ r_\alpha \left(2r_{\alpha_1}+1\right) r_{\alpha_2} - 1 \right\} + \sum_{j=1}^{d} r_j N_j + 2r_D - 1 \qquad (\{\alpha_1, \alpha_2\} = S(\alpha))$$

arithmetical operations. Under the assumptions $r_\alpha \le r$ and $N_j \le 2n-1$, the asymptotic cost is

$$2\,(d-1)\,r^3 + 2rn.$$

Proof. Set $\beta_{\ell}^{(\alpha)} := \langle \mathbf{b}_{\ell}^{(\alpha)}, \mathbf{v}^{(\alpha)} \rangle_{\alpha}$ and $\beta^{(\alpha)} := (\beta_{\ell}^{(\alpha)})_{\ell=1}^{r_\alpha} \in \mathbb{K}^{r_\alpha}$. Given the vectors $\beta^{(\alpha_1)}$ and $\beta^{(\alpha_2)}$, (13.8) implies that $\beta_{\ell}^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)} \beta_i^{(\alpha_1)} \beta_j^{(\alpha_2)}$, i.e., $\beta_{\ell}^{(\alpha)} = (\beta^{(\alpha_1)})^{\mathsf{T}} C^{(\alpha,\ell)} \beta^{(\alpha_2)}$ for $1 \le \ell \le r_\alpha$. Therefore, the computation of $\beta^{(\alpha)}$ requires $r_\alpha\,(2r_{\alpha_1}+1)\,r_{\alpha_2} - 1$ operation. The recursion (13.8) terminates with the scalar products $\langle b_i^{(j)}, v^{(j)} \rangle_j$, which cost $\sum_{j=1}^{d} r_j N_j$ operations. Finally, the scalar product $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{\ell=1}^{r_D} c_{\ell}^{(D)} \beta_{\ell}^{(D)}$ takes $2r_D - 1$ operations. □

Next, we consider the scalar product $\langle \mathbf{u}, \mathbf{v} \rangle$ of general tensors $\mathbf{u}, \mathbf{v} \in \mathcal{H}_{\mathfrak{r}}$.

Case I (two tensors from $\mathcal{H}_{\mathfrak{r}}$ with identical bases). Two tensors

$$\mathbf{u} = \rho_{\mathrm{HTR}}^{\mathrm{orth}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c_u^{(D)}, (B_j)_{j\in D}\,\big),\ \mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{orth}}\big(\ldots, c_v^{(D)}, \ldots\big)$$

given in the *same* format $\mathcal{H}_{\mathfrak{r}}$ with *orthonormal* bases (cf. §11.3.2) satisfy

$$\langle \mathbf{u}, \mathbf{v} \rangle = \big\langle c_{\mathbf{u}}^{(D)}, c_{\mathbf{v}}^{(D)} \big\rangle, \tag{13.9}$$

where $\big\langle c_{\mathbf{u}}^{(D)}, c_{\mathbf{v}}^{(D)} \big\rangle$ is the Euclidean scalar product in $\mathbb{K}^{r_D}$. The cost is negligible: $N_{\langle\cdot,\cdot\rangle} = 2r_D - 1$. Note that Case I holds in particular for $\mathbf{u} = \mathbf{v}$.

Case II (two tensors from $\mathcal{H}_{\mathfrak{r}}$ with different bases). Next, we consider two tensors

$$\mathbf{u}' = \rho_{\mathrm{HTR}}\left(T_D, (\mathbf{C}'^{(\alpha)})_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c'^{(D)}, (\mathbf{B}'^{(\alpha)})_{\alpha\in\mathcal{L}(T_D)}\right) \quad \text{and} \tag{13.10a}$$

$$\mathbf{u}'' = \rho_{\mathrm{HTR}}\left(T_D, (\mathbf{C}''^{(\alpha)})_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c''^{(D)}, (\mathbf{B}''^{(\alpha)})_{\alpha\in\mathcal{L}(T_D)}\right), \tag{13.10b}$$

which are given with respect to *different* bases. Note that the bases need not be orthonormal. The next lemma uses the subtree T_α from Definition 11.6. $\langle\cdot,\cdot\rangle_\beta$ denotes the scalar product of the tensor space $\mathbf{V}_\beta := \bigotimes_{j\in\beta} V_j$.

Lemma 13.7. *Let* $\{\mathbf{b}_i'^{(\beta)} : 1 \le i \le r'_\beta\}$ *and* $\{\mathbf{b}_j''^{(\beta)} : 1 \le j \le r''_\beta\}$ *be the bases involved in (13.10a,b) for* $\beta \in T_\alpha$. *The computation of all scalar products*

$$\left\langle \mathbf{b}_i'^{(\beta)}, \mathbf{b}_j''^{(\beta)} \right\rangle_\beta \qquad \textit{for } 1 \le i \le r'_\beta,\ 1 \le j \le r''_\beta,\ \beta \in T_\alpha$$

costs

$$\sum_{j\in\alpha} r'_{\{j\}} r''_{\{j\}} N_j + 2 \sum_{\beta\in T_\alpha\backslash\mathcal{L}(T_\alpha)} r'_\beta \left(r'_{\beta_1} r''_{\beta_1} r''_{\beta_2} + r'_{\beta_1} r'_{\beta_2} r''_{\beta_1} + r''_\beta r'_{\beta_2} r''_{\beta_2} \right)$$

(β_1, β_2 *sons of* β*) arithmetical operations, if these quantities are computed as detailed in the proof.*

Proof. By property (11.24), we have the recursive equation

$$\left\langle \mathbf{b}_\ell'^{(\beta)}, \mathbf{b}_k''^{(\beta)} \right\rangle_\beta = \sum_{i=1}^{r'_{\beta_1}} \sum_{j=1}^{r'_{\beta_2}} \sum_{m=1}^{r''_{\beta_1}} \sum_{n=1}^{r''_{\beta_2}} c_{ij}'^{(\beta,\ell)} \overline{c_{mn}''^{(\beta,k)}} \left\langle \mathbf{b}_i'^{(\beta_1)}, \mathbf{b}_m''^{(\beta_1)} \right\rangle_{\beta_1} \left\langle \mathbf{b}_j'^{(\beta_2)}, \mathbf{b}_n''^{(\beta_2)} \right\rangle_{\beta_2}. \tag{13.11}$$

Let $S^{(\beta)} \in \mathbb{K}^{r'_\beta \times r''_\beta}$ be the matrix with the entries $S_{\ell k}^{(\beta)} = \langle \mathbf{b}_\ell'^{(\beta)}, \mathbf{b}_k''^{(\beta)} \rangle_\beta$. Eq. (13.11) involves the matrices $S^{(\beta_1)} \in \mathbb{K}^{r'_{\beta_1} \times r''_{\beta_1}}$ and $S^{(\beta_2)} \in \mathbb{K}^{r'_{\beta_2} \times r''_{\beta_2}}$ defined by the entries

$$S_{im}^{(\beta_1)} = \langle \mathbf{b}_i'^{(\beta_1)}, \mathbf{b}_m''^{(\beta_1)} \rangle_{\beta_1} \quad \text{and} \quad S_{jn}^{(\beta_2)} = \langle \mathbf{b}_j'^{(\beta_2)}, \mathbf{b}_n''^{(\beta_2)} \rangle_{\beta_2}.$$

Note that the fourfold sum in (13.11) can be expressed as the Frobenius scalar product $\left\langle S^{\beta_1\mathsf{T}} C'^{(\beta,\ell)} S^{\beta_2}, C''^{(\beta,k)} \right\rangle_\mathsf{F}$. The computation of

$$M_\ell := S^{\beta_1\mathsf{T}} C'^{(\beta,\ell)} S^{\beta_2} \quad \text{for all } 1 \le \ell \le r'_\beta$$

needs $2r'_\beta (r'_{\beta_1} r''_{\beta_1} r''_{\beta_2} + r'_{\beta_1} r'_{\beta_2} r''_{\beta_1})$ operations. The products $\left\langle M_\ell, C''^{(\alpha,k)} \right\rangle_\mathsf{F}$ for all $1 \le \ell \le r'_\beta$ and $1 \le k \le r''_\beta$ cost $2r'_\beta r''_\beta r'_{\beta_2} r''_{\beta_2}$.

The recursion (13.11) terminates for the scalar products $\langle\cdot,\cdot\rangle_\beta$ with respect to the leaves $\beta = \{j\}$ and $j \in \alpha$. In this case, $S^{(\beta)}$ requires the computation of $r'_{\{j\}} r''_{\{j\}}$ scalar products in V_j, each with the cost N_j. □

The scalar product of $\mathbf{u}' = \sum_{i=1}^{r_D} c_i'^{(D)} \mathbf{b}_i'^{(D)}$ and $\mathbf{u}'' = \sum_{j=1}^{r_D} c_j''^{(D)} \mathbf{b}_j''^{(D)}$ equals

$$\langle \mathbf{u}', \mathbf{u}'' \rangle = \sum_{\ell=1}^{r'_D} \sum_{k=1}^{r''_D} c_\ell'^{(D)} \overline{c_k''^{(D)}} \left\langle \mathbf{b}_\ell'^{(D)}, \mathbf{b}_k''^{(D)} \right\rangle.$$

The computation of $S_{\ell k}^{(D)} := \langle \mathbf{b}_\ell'^{(D)}, \mathbf{b}_k''^{(D)} \rangle$ is discussed in Lemma 13.7 for $\alpha := D$. The computation of $\langle \mathbf{u}', \mathbf{u}'' \rangle = (c'^{(D)})^\mathsf{T} S^{(D)} c''^{(D)}$ costs $2r'_D (r''_D + 1)$ operations. Altogether, we get the following result.

Remark 13.8. (a) The recursive computation of the scalar product $\langle \mathbf{u}', \mathbf{u}'' \rangle$ (see proof of Lemma 13.7) costs

$$N_{\langle\cdot,\cdot\rangle} = \sum_{j=1}^{d} r'_{\{j\}} r''_{\{j\}} N_j \tag{13.12a}$$
$$+ 2 \sum_{\alpha \in T_D \backslash \mathcal{L}(T_D)} r'_\alpha \left(r'_{\alpha_1} r''_{\alpha_1} r''_{\alpha_2} + r'_{\alpha_1} r'_{\alpha_2} r''_{\alpha_1} + r''_\alpha r'_{\alpha_2} r''_{\alpha_2} \right) + 2 r'_D \left(r''_D + 1 \right)$$

operations. Under the assumptions $r'_\alpha, r''_\alpha \leq r$ and $N_j \leq 2n$, the cost is bounded by

$$N_{\langle\cdot,\cdot\rangle} \leq 2dr^2 n + 6\,(d-1)\,r^4 + 2r^2. \tag{13.12b}$$

(b) Equation (13.9) does not hold for a non-orthonormal basis. In that case, the scalar product $\langle \mathbf{u}, \mathbf{v} \rangle$ has to be computed as in Case II. By Hermitean symmetry of $S^{(\alpha)}_{\ell k} = \langle \mathbf{b}^{(\alpha)}_\ell, \mathbf{b}^{(\alpha)}_k \rangle_\alpha$, the computational cost is only half of (13.12a).

An alternative approach is to join the bases of $\mathbf{u}'$ and $\mathbf{u}''$ as described in §11.5. By Remark 11.67, this procedure requires

$$N_{\langle\cdot,\cdot\rangle} \leq 8dr^2 \left(r^2 + n \right)$$

operations. Obviously, the latter cost is larger than (13.12b). However, for the special case considered in §14.1.3, this approach is advantageous.

Next we consider the *partial scalar product* $\langle \mathbf{u}', \mathbf{u}'' \rangle_{\alpha^c}$. Here we concentrate to the case of $\alpha \in T_D$. We recall that the partial scalar product $\langle \mathbf{v}, \mathbf{v} \rangle_{\alpha^c}$ is needed for the left-sided singular value decomposition of $\mathcal{M}_\alpha(\mathbf{v})$ (see Lemma 5.13). The result of $\langle \mathbf{u}', \mathbf{u}'' \rangle_{\alpha^c}$ is a tensor in the tensor space $V_\alpha \otimes V_\alpha$ for which a hierarchical format is still to be defined. Let α' be a copy of α disjoint to α and set $A(\alpha) := \alpha \dot{\cup} \alpha'$. The dimension partition tree $T_{A(\alpha)}$ is defined as follows: $A(\alpha)$ is the root with the sons α and α'. The subtree at vertex α is T_α (cf. Definition 11.6) and the subtree at vertex α' is the isomorphic copy $T_{\alpha'}$ of T_α. The bases $\mathbf{b}'^{(\beta)}_\ell$ $(\beta \in T_\alpha)$ $[\mathbf{b}''^{(\beta)}_\ell$ $(\beta \in T_{\alpha'})]$ of $\mathbf{u}'$ $[\mathbf{u}'']$ define the subspaces U_γ, $\gamma \in T_{A(\alpha)} \backslash A(\alpha)$, together with their bases, while the basis of the subspace $U_{A(\alpha)}$ is still to be determined.

The computation of $\langle \mathbf{u}', \mathbf{u}'' \rangle_{\alpha^c}$ follows the description in §4.5.4. First, we form $\mathbf{u}' \otimes \overline{\mathbf{u}''} \in \mathbf{V} \otimes \mathbf{V}$, which is represented in the hierarchical format with the tree $T_{A(D)}$. Let σ_1 and σ_2 be the sons of D. Since either $\alpha \subset \sigma_1$ or $\alpha \subset \sigma_2$, it follows that either $\alpha^c \supset \sigma_2$ or $\alpha^c \supset \sigma_1$. Without loss of generality, we assume $\alpha^c \supset \sigma_2$ and apply the contraction $\mathfrak{C}_{\sigma_2}$ from Definition 4.130:

$$\mathbf{u}' \otimes \overline{\mathbf{u}''} \mapsto \mathfrak{C}_{\sigma_2}(\mathbf{u}' \otimes \overline{\mathbf{u}''}) \in \mathbf{V}_{\sigma_1} \otimes \mathbf{V}_{\sigma_1}.$$

Let

$$\mathbf{u}' = \sum_{\ell=1}^{r_D} c'^{(D)}_\ell \mathbf{b}'^{(D)}_\ell = \sum_{\ell,i,j} c'^{(D)}_\ell c'^{(D,\ell)}_{ij} \mathbf{b}'^{(\sigma_1)}_i \otimes \mathbf{b}'^{(\sigma_2)}_j \quad \text{and}$$

$$\mathbf{u}'' = \sum_{k,m,n} c_k^{\prime\prime(D)} c_{mn}^{\prime\prime(D,k)} \mathbf{b}_m^{\prime\prime(\sigma_1)} \otimes \mathbf{b}_n^{\prime\prime(\sigma_2)}.$$

Then

$$\mathfrak{C}_{\sigma_2}(\mathbf{u}' \otimes \overline{\mathbf{u}''}) = \sum_{\ell,i,j} \sum_{k,m,n} c_\ell^{\prime(D)} c_{ij}^{\prime(D,\ell)} \overline{c_k^{\prime\prime(D)} c_{mn}^{\prime\prime(D,k)}} \overbrace{\left\langle \mathbf{b}_j^{\prime(\sigma_2)}, \mathbf{b}_n^{\prime\prime(\sigma_2)} \right\rangle_{\sigma_2}}^{=S_{jn}^{(\sigma_2)}} \mathbf{b}_i^{\prime(\sigma_1)} \otimes \overline{\mathbf{b}_m^{\prime\prime(\sigma_1)}}$$

holds. For each pair (i, m), the coefficient of $\mathbf{b}_i^{\prime(\sigma_1)} \otimes \overline{\mathbf{b}_m^{\prime\prime(\sigma_1)}}$ is the sum

$$\sum_{\ell,j,k,n} c_\ell^{\prime(D)} c_{ij}^{\prime(D,\ell)} \overline{c_k^{\prime\prime(D)}}\, \overline{c_{mn}^{\prime\prime(D,k)}}\, S_{jn}^{(\sigma_2)}.$$

Set

$$c_{ij}^{\prime(D)} := \sum_\ell c_\ell^{\prime(D)} c_{ij}^{\prime(D,\ell)} \quad \text{and} \quad c_{mn}^{\prime\prime(D)} := \sum_k c_k^{\prime\prime(D)} c_{mn}^{\prime\prime(D,k)}.$$

Then, the fourfold sum equals $C^{\prime(D)} S^{(\sigma_2)} (C^{\prime\prime(D)})^{\mathsf{H}} =: C^{(\sigma_1)}$ and yields the representation

$$\mathfrak{C}_{\sigma_2}(\mathbf{u}' \otimes \overline{\mathbf{u}''}) = \sum_{i,m} c_{im}^{(\sigma_1)} \mathbf{b}_i^{\prime(\sigma_1)} \otimes \overline{\mathbf{b}_m^{\prime\prime(\sigma_1)}} \in \mathbf{V}_{\sigma_1} \otimes \mathbf{V}_{\sigma_1}. \tag{13.13}$$

The computational cost (without the determination of $S_{jn}^{(\sigma_2)}$) is

$$2r'_D r'_{\sigma_1} r'_{\sigma_2} + 2r''_D r''_{\sigma_1} r''_{\sigma_2} + 2r'_{\sigma_1} r'_{\sigma_2} r''_{\sigma_2} + 2r'_{\sigma_1} r''_{\sigma_1} r''_{\sigma_2}.$$

Now we proceed recursively: if $\sigma_1 = \alpha$, we are ready. Otherwise, let σ_{11} and σ_{12} be the sons of σ_1, apply $\mathfrak{C}_{\sigma_{12}}$ [$\mathfrak{C}_{\sigma_{11}}$] if $\alpha^c \supset \sigma_{12}$ [if $\alpha^c \supset \sigma_{11}$] and repeat recursively.

The overall cost is given under the simplification $r'_\beta, r''_\beta \le r$. Then the computation of the coefficients in (13.13) requires $8r^3 level(\alpha)$. In addition, we need to compute the scalar products $\langle \mathbf{b}_j^{\prime(\beta)}, \mathbf{b}_n^{\prime\prime(\beta)} \rangle_\beta$ for all $\beta \in \{\gamma \in T_D : \gamma \cap \alpha = \emptyset\}$. The latter set contains $d - \#\alpha - level(\alpha)$ interior vertices and $d - \#\alpha$ leaf vertices (see (11.6) for the definition of the level). Hence, Lemma 13.7 yields a cost of $2\,(d - \#\alpha)\, r^2 n + 6\,(d - \#\alpha - level(\alpha))\, r^4$. The result is summarised in the following remark.

Remark 13.9. Assume $\alpha \in T_D$. The partial scalar product $\langle \mathbf{u}', \mathbf{u}'' \rangle_{\alpha^c}$ can be performed with the arithmetical cost

$$2\,(d - \#\alpha)\, r^2 n + 6\,(d - \#\alpha - level(\alpha))\, r^4 + 8r^3 level(\alpha).$$

The resulting tensor belongs to $\mathbf{V}_\alpha \otimes \mathbf{V}_\alpha$ and is given in the hierarchical format with the dimension partition tree $T_{A(\alpha)}$ explained above.

13.3.6 Orthonormalisation

One purpose of scalar product computations is the orthonormalisation of a basis (QR or Gram-Schmidt orthonormalisation). If the basis belongs to one of the directly represented vector spaces V_j, the standard procedures from §2.7 apply. This is different, when the basis vectors are tensors represented in one of the tensor formats. This happens, e.g., if the vectors from V_j are tensorised as proposed in §14.1.4. Assume that we start with s tensors

$$\mathbf{b}_j \in \mathbf{W} \qquad (1 \le j \le s)$$

given in some format with representation ranks r (i.e., $r = \max_j r_j$ in the case of $\mathbf{b}_j \in \mathcal{T}_\mathbf{r}$ and $r = \max_\alpha r_\alpha$ for $\mathbf{b}_j \in \mathcal{H}_\mathfrak{r}$). Furthermore, assume that $\dim(\mathbf{W})$ is sufficiently large. Here, one can choose between the following cases.

1. Perform the Gram-Schmidt orthonormalisation without truncation. Then an *exact* orthonormalisation can be achieved. In general, the representation ranks of the new basis elements $\mathbf{b}_j^{\mathrm{new}} \in \mathbf{W}$ equal jr for $1 \le j \le s$ leading to unfavourably large ranks.
2. The same procedure, but with truncation, produces basis elements $\mathbf{b}_j^{\mathrm{new}}$ which are *almost* orthonormal. This may be sufficient if an orthonormal basis is introduced for the purpose of stability.
3. Let $\mathbf{B} := [\mathbf{b}_1 \cdots \mathbf{b}_s] \in \mathbf{W}^s$ be the matrix corresponding to the basis and compute the Cholesky decomposition of the Gram matrix: $\mathbf{B}^\mathsf{H}\mathbf{B} = LL^\mathsf{H} \in \mathbb{K}^{s\times s}$. The (exactly) orthonormalised basis is $\mathbf{B}^{\mathrm{new}} = \mathbf{B}L^{-\mathsf{H}}$ (cf. Lemma 8.12b). In some applications is suffices to use the factorisation $\mathbf{B}L^{-\mathsf{H}}$ without performing the product.

13.4 Change of Bases

In general, the vector spaces V_j [or U_j] are addressed by means of a basis or frame $(b_i^{(j)})_{1\le i\le n_j}$ which gives rise to a matrix $B_j := [b_1^{(j)}\ b_2^{(j)}\ \ldots\ b_{n_j}^{(j)}] \in V_j^{n_j}$. Consider a new basis $(b_{i,\mathrm{new}}^{(j)})_{1\le i\le n_{j,\mathrm{new}}}$ and B_j^{new} together with the transformation

$$B_j = B_j^{\mathrm{new}} T^{(j)}, \qquad \text{i.e.,} \quad b_k^{(j)} = \sum_{i=1}^{n_{j,\mathrm{new}}} T_{ik}^{(j)} b_{i,\mathrm{new}}^{(j)} . \tag{13.14}$$

$n_j = n_{j,\mathrm{new}}$ holds for bases. If B_j is a frame, also $n_{j,\mathrm{new}} < n_j$ may occur.

We write r_j $[r_{j,\mathrm{new}}]$ instead of n_j $[n_{j,\mathrm{new}}]$, if the bases span only a subspace U_j.

In the case of the tensor subspace format and the hierarchical format, another change of bases may be of interest. If the subspaces are not described by orthonormal bases, an *orthonormalisation* can be performed. This includes the determination of some orthonormal basis and the corresponding transformation of the coefficients.

13.4.1 Full Representation

The full representation $\sum_{\mathbf{i}\in\mathbf{I}} \mathbf{a_i}\, b^{(1)}_{i_1} \otimes \ldots \otimes b^{(d)}_{i_d}$ (cf. (7.3)) is identical to the tensor subspace representation involving maximal subspaces $U_j = V_j$ with dimension $n_j = \#I_j$. The coefficient tensor $\mathbf{a} \in \mathbb{K}^{\mathbf{I}}$ is transformed into $\mathbf{a}_{\text{new}} := \mathbf{T}\,\mathbf{a} \in \mathbb{K}^{\mathbf{I}}$ with the Kronecker matrix $\mathbf{T} = \bigotimes_{j=1}^{d} T^{(j)}$ (cf. (8.7b)). The elementwise operations are $\mathbf{a}^{\text{new}}_{i_1 i_2 \cdots i_d} = \sum_{k_1=1}^{n_1} \cdots \sum_{k_d=1}^{n_d} T^{(1)}_{i_1 k_1} \cdots T^{(d)}_{i_d k_d} \mathbf{a}_{k_1 k_2 \cdots k_d}$ with $1 \le i_j, k_j \le n_j$. The arithmetical cost is

$$N^{\text{full}}_{\text{basis-change}} = 2\left(\sum_{j=1}^{d} n_j\right) \prod_{j=1}^{d} n_j \le 2dn^{d+1} \qquad \text{for } n := \max_j n_j.$$

13.4.2 Hybrid r-Term Representation

Let $\mathbf{v} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} v^{(j)}_\nu \in \mathcal{R}_r$. If $v^{(j)}_\nu \in \mathbb{K}^{I_j}$ represents the vector $\sum_{i=1}^{n_j} v^{(j)}_{\nu,i} b^{(j)}_i \in V_j$, the transformation with respect to the new bases B^{new}_j yields

$$\sum_{k=1}^{n_j} v^{(j)}_{\nu,k} b^{(j)}_k = \sum_{k=1}^{n_j} v^{(j)}_{\nu,k} \sum_{i=1}^{n_{j,\text{new}}} T^{(j)}_{ik} b^{(j)}_{i,\text{new}} = \sum_{i=1}^{n_{j,\text{new}}} \underbrace{\left(\sum_{k=1}^{n_j} T^{(j)}_{ik} v^{(j)}_{\nu,k}\right)}_{= \hat{v}^{(j)}_{\nu,i}} b^{(j)}_{i,\text{new}}$$

(cf. (13.14)). Hence, the transformed tensor is

$$\hat{\mathbf{v}} = \sum_{\nu=1}^{r} \bigotimes_{j=1}^{d} \hat{v}^{(j)}_\nu \in \mathcal{R}_r \qquad \text{with } \hat{v}^{(j)}_\nu = T^{(j)} v^{(j)}_\nu.$$

Multiplication by the $n_{j,\text{new}} \times n_j$ matrices $T^{(j)}$ leads to the total cost

$$N^{\mathcal{R}_r}_{\text{basis-change}} = r \sum_{j=1}^{d} n_{j,\text{new}}\,(2n_j - 1) \le 2drn^2 \qquad \text{for } n := \max_j n_j.$$

13.4.3 Tensor Subspace Representation

Here, it is assumed that only the bases representing the subspaces $U_j \subset V_j$ are changed. We assume that the basis vectors $b^{(j)}_{i,\text{new}}$, $1 \le i \le r_{j,\text{new}}$, are given together with the $r_j \times r_{j,\text{new}}$ matrices $T^{(j)}$. The cost for transforming the coefficient tensor is as in §13.4.1, but with n_j replaced by r_j:

$$N^{\mathcal{T}_{\mathbf{r}}}_{\text{basis-change}} = 2\sum_{j=1}^{d} \left(\prod_{k=1}^{j} r_{k,\text{new}} \prod_{k=j}^{d} r_k\right) \le 2dr^{d+1} \quad \text{for } r := \max_j\{r_j, r_{j,\text{new}}\}. \tag{13.15a}$$

Another type of basis transform is the orthonormalisation in the case that the format $\mathcal{T}_\mathbf{r}$ is described by general bases (or frames) $B_j^{\text{old}} := [b_{1,\text{old}}^{(j)}, \ldots, b_{r_{j,\text{old}},\text{old}}^{(j)}]$. By procedure $\mathbf{RQR}(n_j, r_{j,\text{old}}, r_j, B_j^{\text{old}}, Q, R)$ from (2.29) one obtains a new orthonormal basis $B_j^{\text{new}} = Q = [b_{1,\text{new}}^{(j)}, \ldots, b_{r_j,\text{new}}^{(j)}]$ together with the transformation matrix $T^{(j)} = R$, i.e., $B_j^{\text{new}} T^{(j)} = B_j^{\text{old}}$. Note that in the case of a frame B_j^{old}, the dimension r_j may be smaller than $r_{j,\text{old}}$. The cost for calling $\mathbf{RQR}$ is $N_{\text{QR}}(n_j, r_{j,\text{old}}) = 2n_j r_{j,\text{old}}^2$. The cost of an application of $T^{(j)}$ to the coefficient tensor is $N_{\text{basis-change}}^{\mathcal{T}_\mathbf{r}}$ from above. Altogether, the cost of orthonormalisation is

$$N_{\text{orthonormalisation}}^{\mathcal{T}_\mathbf{r}} \leq 2dr^{d+1} + 2dnr^2 \tag{13.15b}$$

with $r := \max_j r_{j,\text{old}}$ and $n := \max_j n_j$.

13.4.4 Hierarchical Representation

The transformation considered here, is a special case of §11.3.1.4. The basis transformations (13.14) influence the coefficient matrices $C^{(\alpha,\ell)}$ for vertices $\alpha \in T_D$ with at least one son $\{j\} \in \mathcal{L}(T_D)$. Let α_1 denote the first son and α_2 the second son of α. Then

$$C_{\text{new}}^{(\alpha,\ell)} := \left\{ \begin{array}{ll} T^{(j_1)}\, C_{\text{old}}^{(\alpha,\ell)}\, (T^{(j_2)})^\mathsf{T} & \text{if } \alpha_1 = \{j_1\} \text{ and } \alpha_2 = \{j_2\} \\ T^{(j_1)}\, C_{\text{old}}^{(\alpha,\ell)} & \text{if } \alpha_1 = \{j_1\} \text{ and } \alpha_2 \notin \mathcal{L}(T_D) \\ C_{\text{old}}^{(\alpha,\ell)}\, (T^{(j_2)})^\mathsf{T} & \text{if } \alpha_1 \notin \mathcal{L}(T_D) \text{ and } \alpha_2 = \{j_2\} \end{array} \right\}$$

for $1 \leq \ell \leq r_\alpha$. Otherwise, $C^{(\alpha,\ell)}$ is unchanged. The computational work consists of d matrix multiplications:

$$N_{\text{basis-change}}^{\mathcal{H}_\mathfrak{r}} = 2\sum_{j=1}^{d} r_j r_{j,\text{new}} r_{\text{brother}(j)} \leq 2dr^3. \tag{13.16a}$$

The brother of $\{j\}$ may be defined by $\{\text{brother}(j)\} := S(\text{father}(j)) \backslash \{j\}$.

Next, we assume that the bases (frames) $\mathbf{B}_\alpha := [\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}]$ ($\alpha \in T_D$) of $\mathcal{H}_\mathfrak{r}$ are to be orthonormalised. Generating orthonormal bases by $\mathbf{RQR}$ costs $N_{\text{QR}}(n_j, r_j) = 2n_j r_j^2$ for $1 \leq j \leq d$ and $N_{\text{QR}}(r_\alpha, r_\alpha) = 2r_\alpha^3$ for $\alpha \in T_D \backslash \mathcal{L}(T_D)$. Each transformation $T^{(\alpha)}$ ($\alpha \neq D$) leads to r_γ matrix multiplications (cf. (11.32)) with the cost $2r_\gamma r_\alpha r_\beta r_\alpha^{\text{new}}$, where $\gamma := \text{father}(\alpha)$ and $\beta := \text{brother}(\alpha)$. $T^{(D)}$ leads to $2r_D r_D^{\text{new}}$ operations (cf. (11.34)). Hence, orthonormalisation is realised by

$$\begin{aligned} & N_{\text{orthonormalisation}}^{\mathcal{H}_\mathfrak{r}} \qquad\qquad\qquad\qquad (\{\alpha_1, \alpha_2\} = S(\alpha)) \\ & = 2\sum_{j=1}^{d} n_j r_j^2 + 2\sum_{\alpha \in T_D \backslash \mathcal{L}(T_D)} \left(r_\alpha^3 + r_\alpha r_{\alpha_1} r_{\alpha_2} \left(r_{\alpha_1} + r_{\alpha_2} \right) \right) + 2r_D^2 \qquad (13.16b) \\ & \leq 2dnr^2 + 4\,(d-1)\, r^4 + 2\,(d-1)\, r^3 + 2r^2. \end{aligned}$$

operations, where $r := \max_\alpha r_\alpha$ and $n := \max_j n_j$.

13.5 General Binary Operation

Here we consider tensor spaces $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$, $\mathbf{W} = \bigotimes_{j=1}^{d} W_j$, $\mathbf{X} = \bigotimes_{j=1}^{d} X_j$, and any bilinear operation

$$\boxdot : \mathbf{V} \times \mathbf{W} \to \mathbf{X},$$

which satisfies[7]

$$\left(\bigotimes_{j=1}^{d} v^{(j)} \right) \boxdot \left(\bigotimes_{j=1}^{d} w^{(j)} \right) = \bigotimes_{j=1}^{d} \left(v^{(j)} \boxdot w^{(j)} \right), \quad v^{(j)} \boxdot w^{(j)} \in X_j, \quad (13.17)$$

for elementary tensors. We assume that the evaluation of $v^{(j)} \boxdot w^{(j)}$ for vectors $v^{(j)} \in V_j$ and $w^{(j)} \in W_j$ costs $N_j^{\boxdot}$ arithmetical operations.

13.5.1 *r*-Term Representation

Tensors $\mathbf{v} = \sum_{\nu=1}^{r_{\mathbf{v}}} \bigotimes_{j=1}^{d} v_\nu^{(j)} \in \mathcal{R}_{r_{\mathbf{v}}}(\mathbf{V})$ and $\mathbf{w} = \sum_{\mu=1}^{r_{\mathbf{w}}} \bigotimes_{j=1}^{d} w_\mu^{(j)} \in \mathcal{R}_{r_{\mathbf{w}}}(\mathbf{W})$ lead to $\mathbf{x} := \mathbf{v} \boxdot \mathbf{w} \in \mathcal{R}_{r_{\mathbf{v}} r_{\mathbf{w}}}(\mathbf{X})$ with

$$\mathbf{x} = \sum_{\nu=1}^{r_{\mathbf{v}}} \sum_{\mu=1}^{r_{\mathbf{w}}} \bigotimes_{j=1}^{d} \left(v_\nu^{(j)} \boxdot w_\mu^{(j)} \right) \in \mathcal{R}_{r_{\mathbf{x}}}, \qquad r_{\mathbf{x}} = r_{\mathbf{v}} \cdot r_{\mathbf{w}}.$$

Under the assumption about $N_j^{\boxdot}$, the total work is

$$N_{\boxdot}^{\mathcal{R}_r} = r_{\mathbf{v}} r_{\mathbf{w}} \sum_{j=1}^{d} N_j^{\boxdot}. \quad (13.18)$$

13.5.2 Tensor Subspace Representation

For $\mathbf{v} = \sum_{\mathbf{i} \in \mathbf{J}'} \mathbf{a}'_{\mathbf{i}} \bigotimes_{j=1}^{d} b_{i_j}^{\prime(j)} \in \mathcal{T}_{\mathbf{r}'}(\mathbf{V})$ and $\mathbf{w} = \sum_{\mathbf{k} \in \mathbf{J}''} \mathbf{a}''_{\mathbf{k}} \bigotimes_{j=1}^{d} b_{k_j}^{\prime\prime(j)} \in \mathcal{T}_{\mathbf{r}''}(\mathbf{W})$ we conclude from (13.17) that

$$\mathbf{w} := \mathbf{u} \boxdot \mathbf{v} = \sum_{\mathbf{i} \in \mathbf{J}'} \sum_{\mathbf{k} \in \mathbf{J}''} \mathbf{a}'_{\mathbf{i}} \, \mathbf{a}''_{\mathbf{k}} \bigotimes_{j=1}^{d} \left(b_{i_j}^{\prime(j)} \boxdot b_{k_j}^{\prime\prime(j)} \right).$$

We may define the frame $\mathfrak{b}^{(j)} := (b_i^{\prime(j)} \boxdot b_k^{\prime\prime(j)} : i \in J'_j,\, k \in J'_j)$ and the subspace $U_j = \operatorname{span}(\mathfrak{b}^{(j)}) \subset \mathbf{X}$. Then a possible representation is

[7] The map $\boxdot_j = \boxdot : V_j \times W_j \to X_j$ on the right-hand side is denoted by the same symbol $\boxdot$.

$$\mathbf{x} = \sum_{\mathbf{m}\in\mathbf{J}} \mathbf{a}_{\mathbf{m}} \bigotimes_{j=1}^{d} b_{m_j}^{(j)} \quad \text{with } \mathbf{J} := \mathop{\times}_{j=1}^{d} J_j, \quad J_j := J'_j \times J''_j, \tag{13.19}$$
$$b_{m_j}^{(j)} := b_{m'_j}^{\prime(j)} \boxdot b_{m''_j}^{\prime\prime(j)} \in \mathfrak{b}^{(j)} \qquad \text{for } m_j := \left(m'_j, m''_j\right) \in J_j,$$
$$\mathbf{a}_{\mathbf{m}} := \mathbf{a}'_{\mathbf{m}'} \mathbf{a}''_{\mathbf{m}''} \qquad \text{with } \mathbf{m} = \left(\left(m'_1, m''_1\right), \ldots, \left(m'_d, m''_d\right)\right)$$

and $\mathbf{m}' = \left(m'_1, \ldots, m'_d\right), \mathbf{m}'' = \left(m''_1, \ldots, m''_d\right)$.

The cost for computing all frame vectors $b_{m_j}^{(j)} \in \mathfrak{b}^{(j)}$ is $\#J_j N_j^{\boxdot}$. The coefficient tensor $\mathbf{a}_{\mathbf{m}}$ requires $\#J$ multiplications. The total work is

$$N_{\boxdot}^{\mathcal{T}_{\mathbf{r}}} = \sum_{j=1}^{d} \#J'_j \#J''_j N_j^{\boxdot} + \prod_{j=1}^{d} \left(\#J'_j \#J''_j\right).$$

In general, $\mathfrak{b}^{(j)}$ is only a frame. Therefore, a further orthonormalisation of $\mathfrak{b}^{(j)}$ may be desired. By (13.15b), the additional cost is

$$2d(r^2)^{d+1} + 2dn(r^2)^2 = 2dnr^4 + 2dr^{2d+2} \qquad \text{(cf. (13.15b))}.$$

13.5.3 Hierarchical Representation

Let $\mathbf{v} \in \mathcal{H}_{\mathfrak{r}'}(\mathbf{V})$ and $\mathbf{w} \in \mathcal{H}_{\mathfrak{r}''}(\mathbf{W})$ be two tensors described in two different hierarchical formats, but with the same dimension partition tree T_D. Since $\mathbf{v} = \sum_\ell c_\ell^{\prime(D)} \mathbf{b}_\ell^{\prime(D)}$ and $\mathbf{w} = \sum_k c_k^{\prime\prime(D)} \mathbf{b}_k^{\prime\prime(D)}$, one starts from

$$\mathbf{v} \boxdot \mathbf{w} = \sum_{\ell=1}^{r'_D} \sum_{k=1}^{r''_D} c_\ell^{\prime(D)} c_k^{\prime\prime(D)} \, \mathbf{b}_\ell^{\prime(D)} \boxdot \mathbf{b}_k^{\prime\prime(D)} \tag{13.20a}$$

and uses the recursion

$$\mathbf{b}_\ell^{\prime(\alpha)} \boxdot \mathbf{b}_k^{\prime\prime(\alpha)} = \sum_{i=1}^{r'_{\alpha_1}} \sum_{j=1}^{r'_{\alpha_2}} \sum_{i=1}^{r''_{\alpha_1}} \sum_{j=1}^{r''_{\alpha_2}} c_{ij}^{\prime(\alpha,\ell)} c_{mn}^{\prime\prime(\alpha,k)} \left(\mathbf{b}_i^{\prime(\alpha_1)} \boxdot \mathbf{b}_m^{\prime\prime(\alpha_1)}\right) \otimes \left(\mathbf{b}_j^{\prime(\alpha_2)} \boxdot \mathbf{b}_n^{\prime\prime(\alpha_2)}\right) \tag{13.20b}$$

(cf. (11.24)), which terminates at the leaves of T_D.

In the first approach, we accept the *frame* $\mathfrak{b}^{(\alpha)}$ consisting of the $r'_\alpha r''_\alpha$ vectors $\mathbf{b}_\ell^{\prime(\alpha)} \boxdot \mathbf{b}_k^{\prime\prime(\alpha)}$ $(1 \le \ell \le r'_\alpha, 1 \le k \le r''_\alpha)$ describing the subspace $\mathbf{U}_\alpha$. The computation of $\mathfrak{b}^{(j)}$, $1 \le j \le d$, costs $\sum_{j=1}^{d} r'_j r''_j N_j^{\boxdot}$ operations. Denote the elements of $\mathfrak{b}^{(\alpha)}$ by $\mathbf{b}_m^{(\alpha)}$ with $m \in J_\alpha := \{1, \ldots, r'_\alpha\} \times \{1, \ldots, r''_\alpha\}$, i.e., $\mathbf{b}_m^{(\alpha)} = \mathbf{b}_\ell^{\prime(\alpha)} \boxdot \mathbf{b}_k^{\prime\prime(\alpha)}$ if $m = (\ell, k)$. Then (13.20b) yields the relation

$$\mathbf{b}_m^{(\alpha)} = \sum_{p \in J_{\alpha_1}} \sum_{q \in J_{\alpha_2}} c_{pq}^{(\alpha,m)} \mathbf{b}_p^{(\alpha_1)} \otimes \mathbf{b}_q^{(\alpha_2)} \quad \text{with } c_{pq}^{(\alpha,m)} := c_{p_1 q_1}^{\prime(\alpha,\ell)} c_{p_2 q_2}^{\prime\prime(\alpha,k)} \tag{13.20c}$$

for $p=(p_1,p_2)\in J_{\alpha_1}, q=(q_1,q_2)\in J_{\alpha_2}$.The new coefficient matrix $C^{(\alpha,m)}$ is the Kronecker product

$$C^{(\alpha,m)}=C'^{(\alpha,\ell)}\otimes C''^{(\alpha,k)}\qquad\text{for } m=(\ell,k)$$

and can be obtained by $\#J_{\alpha_1}\#J_{\alpha_2}=r'_{\alpha_1}r''_{\alpha_1}r'_{\alpha_2}r''_{\alpha_2}$ multiplications.

Equation (13.20a) can be rewritten as

$$\mathbf{v}\boxdot\mathbf{w}=\sum_{m\in J_D}c_m^{(D)}\mathbf{b}_m^{(D)}\qquad\text{with } c_m^{(D)}:=c'^{(D)}_{m_1}c''^{(D)}_{m_2}\text{ for } m=(m_1,m_2)$$

involving $\#J_D=r'_D r''_D$ multiplications. The result $\mathbf{v}\boxdot\mathbf{w}$ is represented in $\mathcal{H}_\mathfrak{r}(\mathbf{X})$ with representation ranks $r_\alpha:=r'_\alpha r''_\alpha$. Altogether, the computational cost amounts to

$$N_\boxdot^{\mathcal{H}_\mathfrak{r}}=\sum_{j=1}^d r'_j r''_j N_j^\boxdot+\sum_{\alpha\in T_D\backslash\mathcal{L}(T_D)}r'_{\alpha_1}r''_{\alpha_1}r'_{\alpha_2}r''_{\alpha_2}+r'_D r''_D\le dr^2n+(d-1)\,r^4+1.$$

By (13.16b) with r replaced by r^2, an additional orthonormalisation of the frame requires $2dnr^4+4dr^8+(\text{lower order terms})$ operations.

13.6 Hadamard Product of Tensors

The Hadamard product defined in §4.6.4 is of the form (13.17). For $V_j=\mathbb{K}^{n_j}$ the number of arithmetical operations is given by $N_j^\odot=n_j$ replacing $N_j^\boxdot$. Therefore, the considerations in §13.5 yield the following costs for the different formats:

$$N_\odot^{\text{full}}=\prod_{j=1}^d n_j,\tag{13.21a}$$

$$N_\odot^{\mathcal{R}_r}=r_\mathbf{v}\cdot r_\mathbf{w}\sum_{j=1}^d n_j\le dr^2n\qquad\text{with } r:=\max\{r_\mathbf{u},r_\mathbf{v}\},\tag{13.21b}$$

$$N_\odot^{\mathcal{T}_\mathbf{r}}=\sum_{j=1}^d n_j\#J'_j\#J''_j+\prod_{j=1}^d\left(\#J'_j\#J''_j\right)\le dnr^2+r^{2d},\tag{13.21c}$$

$$N_\odot^{\mathcal{H}_\mathfrak{r}}=\sum_{j=1}^d r'_j r''_j n_j+\sum_{\alpha\in T_D\backslash\mathcal{L}(T_D)}r'_{\alpha_1}r''_{\alpha_1}r'_{\alpha_2}r''_{\alpha_2}+r'_\alpha r''_\alpha\le dr^2n+(d-1)\,r^4+r^2,\tag{13.21d}$$

where $n:=\max_j\{n_j\}$ and in (13.21c) $r:=\max_j\{\#J'_j,\#J''_j\}$. Concerning an additional orthonormalisation of the frames obtained for the formats $\mathcal{T}_\mathbf{r}$ and $\mathcal{H}_\mathfrak{r}$ compare the remarks in §13.5.2 and §13.5.3.

Above, we have considered the Hadamard product as an example of a binary operation $\odot : \mathbf{V} \times \mathbf{V} \to \mathbf{V}$. Consider $\mathbf{h} := \mathbf{g} \odot \mathbf{f}$ with *fixed* $\mathbf{g} \in \mathbf{V}$. Then $\mathbf{f} \mapsto \mathbf{h}$ is a linear mapping and $\mathbf{G} \in L(\mathbf{V}, \mathbf{V})$ defined by

$$\mathbf{G}(\mathbf{f}) := \mathbf{g} \odot \mathbf{f} \tag{13.22}$$

is a linear multiplication operator. On the level of matrices, $\mathbf{G}$ is the diagonal matrix formed from the vector $\mathbf{g}$:

$$\mathbf{G} := \mathrm{diag}\{\mathbf{g}_\mathbf{i} : \mathbf{i} \in \mathbf{I}\}.$$

Remark 13.10. If $\mathbf{g}$ is given in one of the formats $\mathcal{R}_r, \mathcal{T}_\mathbf{r}, \mathcal{H}_\mathfrak{r}$, matrix $\mathbf{G}$ has a quite similar representation in $\mathcal{R}_r,\ \mathcal{T}_\mathbf{r},\ \mathcal{H}_\mathfrak{r}$ with vectors replaced by diagonal matrices:

$$\mathbf{g} = \sum_i \bigotimes_{j=1}^d g_i^{(j)} \quad \Rightarrow \mathbf{G} = \sum_i \bigotimes_{j=1}^d G_i^{(j)} \quad \text{with } G_i^{(j)} := \mathrm{diag}\{g_i^{(j)}[\nu] : \nu \in I_j\},$$

$$\mathbf{g} = \sum_\mathbf{i} \mathbf{a}[\mathbf{i}] \bigotimes_{j=1}^d b_{i_j}^{(j)} \Rightarrow \mathbf{G} = \sum_\mathbf{i} \mathbf{a}[\mathbf{i}] \bigotimes_{j=1}^d B_{i_j}^{(j)} \text{ with } B_{i_j}^{(j)} := \mathrm{diag}\{b_{i_j}^{(j)}[\nu] : \nu \in I_j\},$$

and analogously for $\mathcal{H}_\mathfrak{r}$. Even the storage requirements are identical, if we exploit that diagonal matrices are characterised by the diagonal entries.

13.7 Convolution of Tensors

We assume that the convolution operations $\star : V_j \times V_j \to V_j$ are defined and satisfy (13.17). For $V_j = \mathbb{K}^{n_j}$ we expect $N_j^\star = O(n_j \log n_j)$ replacing $N_j^\boxdot$. A realisation of the convolution of functions with similar operation count (n_j: data size of the function representation) is discussed in [85]. The algorithms from §13.5 with $\star$ instead of $\boxdot$ require the following costs:

$$N_\star^{\mathrm{full}} \le O(dn^d \log n), \tag{13.23a}$$
$$N_\star^{\mathcal{R}_r} \le O(dr^2 n \log n), \tag{13.23b}$$
$$N_\star^{\mathcal{T}_\mathbf{r}} \le O(dr^2 n \log n) + r^{2d}, \tag{13.23c}$$
$$N_\star^{\mathcal{H}_\mathfrak{r}} \le O(dr^2 n \log n) + (d-1)\, r^4. \tag{13.23d}$$

The same comment as above applies to an orthonormalisation.

A cheaper performance of the convolution will be proposed in §14.3, where in suitable cases $N_j^\star = O(\log n_j)$ may hold.

13.8 Matrix-Matrix Multiplication

Let $\mathbf{V} := \mathcal{L}(\mathbf{R},\mathbf{S})$, $\mathbf{W} := \mathcal{L}(\mathbf{S},\mathbf{T})$, and $\mathbf{X} := \mathcal{L}(\mathbf{R},\mathbf{T})$ be matrix spaces with $\mathbf{R}=\bigotimes_{j=1}^d R_j$, $\mathbf{S}=\bigotimes_{j=1}^d S_j$, $\mathbf{T}=\bigotimes_{j=1}^d T_j$. The matrix-matrix multiplication is a binary operation satisfying (13.17). In the case of $R_j=\mathbb{K}^{n_j^R}$, $S_j=\mathbb{K}^{n_j^S}$, $T_j=\mathbb{K}^{n_j^T}$, the standard matrix-matrix multiplication of $A_j' \in \mathcal{L}(R_j,S_j)$ and $A_j'' \in \mathcal{L}(S_j,T_j)$ requires $N_j^{\boxdot}=2n_j^R n_j^S n_j^T$ arithmetical operations. This leads to the following costs:

$$\begin{aligned}
N_{\mathsf{MMM}}^{\mathrm{full}} &= 2\prod_{j=1}^d n_j^R n_j^S n_j^T,\\
N_{\mathsf{MMM}}^{\mathcal{R}_r} &= 2r_R \cdot r_S \sum_{j=1}^d n_j^R n_j^S n_j^T,\\
N_{\mathsf{MMM}}^{\mathcal{T}_\mathbf{r}} &= 2\sum_{j=1}^d \#J_j'\#J_j'' n_j^R n_j^S n_j^T + \prod_{j=1}^d \left(\#J_j'\#J_j''\right) \le 2dr^2n^3 + r^{2d},\\
N_{\mathsf{MMM}}^{\mathcal{H}_\mathfrak{r}} &= 2\sum_{j=1}^d r_j' r_j'' n_j^R n_j^S n_j^T + \sum_{\alpha\in T_D\backslash\mathcal{L}(T_D)} r_{\alpha_1}' r_{\alpha_1}'' r_{\alpha_2}' r_{\alpha_2}'' + r_\alpha' r_\alpha''\\
&\le dr^2n^3 + (d-1)\,r^4 + r^2.
\end{aligned}$$

Note, however, that the matrix-matrix multiplication of hierarchical matrices of size $n_j \times n_j$ requires only $N_j^{\mathsf{MMM}} = O(n\log^* n)$ operations (cf. [86, §7.8.3]).

Often, one is interested in symmetric $(\mathbb{K}=\mathbb{R})$ or Hermitean matrices $(\mathbb{K}=\mathbb{C})$:

$$\mathbf{M} = \mathbf{M}^{\mathsf{H}}. \tag{13.24}$$

Sufficient conditions are given in the following lemma.

Lemma 13.11. *(a) Format $\mathcal{R}_r$: $\mathbf{M}=\sum_i \bigotimes_{j=1}^d M_i^{(j)}$ satisfies (13.24), if $M_i^{(j)}=(M_i^{(j)})^{\mathsf{H}}$.*

(b) Format $\mathcal{T}_\mathbf{r}$: $\mathbf{M}=\sum_\mathbf{i} \mathbf{a}[\mathbf{i}] \bigotimes_{j=1}^d b_{i_j}^{(j)}$ satisfies (13.24), if $b_{i_j}^{(j)} = (b_{i_j}^{(j)})^{\mathsf{H}}$.

(c) Format $\mathcal{H}_\mathfrak{r}$: $\mathbf{M}\in\mathcal{H}_\mathfrak{r}$ satisfies (13.24), if the bases $B_j = (b_i^{(j)})_{1\le i\le r_j} \subset V_j$ in (11.28) consist of Hermitean matrices: $b_i^{(j)} = (b_i^{(j)})^{\mathsf{H}}$.

Proof. See Exercise 4.132. □

13.9 Matrix-Vector Multiplication

We distinguish the following cases:
(a) the matrix $\mathbf{A}\in\mathcal{L}(\mathbf{V},\mathbf{W})$ and the vector $\mathbf{v}\in\mathbf{V}$ are given in the same format,
(b) the vector $\mathbf{v}\in\mathbf{V}$ is given in one of the formats, while $\mathbf{A}$ is of special form:

$$\mathbf{A} = A^{(1)} \otimes I \otimes \ldots \otimes I + I \otimes A^{(2)} \otimes I \otimes \ldots \otimes I + \ldots + I \otimes \ldots \otimes I \otimes A^{(d)}, \tag{13.25a}$$

$$\mathbf{A} = \bigotimes_{j=1}^{d} A^{(j)}, \tag{13.25b}$$

$$\mathbf{A} = \sum_{i=1}^{p} \bigotimes_{j=1}^{d} A_i^{(j)} \tag{13.25c}$$

with $A^{(j)}, A_i^{(j)} \in \mathcal{L}(V_j, W_j)$ (where $V_j = W_j$ in the case of (13.25a)). We assume $n_j = \dim(V_j)$ and $m_j = \dim(W_j)$. A matrix like in (13.25a) occurs for separable differential operators and their discretisations (cf. Definition 9.36). (13.25b) describes a general elementary tensor, and (13.25c) is the general p-term format.

13.9.1 Identical Formats

The matrix-vector multiplication is again of the form (13.17). The standard cost of $A^{(j)} v^{(j)}$ is $2n_j m_j$ (for hierarchical matrices the computational cost can be reduced to $O((n_j + m_j)\log(n_j + m_j))$, cf. [86, Lemma 7.8.1]). §13.5 shows

$$\begin{aligned}
N_{\mathsf{MVM}}^{\mathrm{full}} &= 2\prod\nolimits_{j=1}^{d} n_j m_j, \\
N_{\mathsf{MVM}}^{\mathcal{R}_r} &= 2r_{\mathbf{v}} \cdot r_{\mathbf{w}} \sum\nolimits_{j=1}^{d} n_j m_j, \\
N_{\mathsf{MVM}}^{\mathcal{T}_{\mathbf{r}}} &= 2\sum\nolimits_{j=1}^{d} \#J_j' \#J_j'' n_j m_j + \prod\nolimits_{j=1}^{d} \left(\#J_j' \#J_j''\right) \leq 2dr^2 nm + r^{2d}, \\
N_{\mathsf{MVM}}^{\mathcal{H}_{\mathbf{r}}} &= 2\sum_{j=1}^{d} r_j' r_j'' n_j m_j + \sum_{\alpha \in T_D \backslash \mathcal{L}(T_D)} r_{\alpha_1}' r_{\alpha_1}'' r_{\alpha_2}' r_{\alpha_2}'' + r_D' r_D'' \\
&\leq 2dr^2 nm + (d-1)\, r^4 + 1,
\end{aligned}$$

where $n := \max_j n_j$ and $m := \max_j m_j$.

13.9.2 Separable Form (13.25a)

Let $\mathbf{v} \in \mathbf{V}$ be given in full format. $\mathbf{w} = (A^{(1)} \otimes I \otimes I \otimes \ldots \otimes I)\mathbf{v}$ has the explicit description $\mathbf{w}[i_1 \ldots i_d] = \sum_{k_1=1}^{n_1} A_{i_1 k_1}^{(1)} \mathbf{v}[k_1 i_2 \ldots i_d]$. Its computation for all $i_1, \ldots, i_d$ takes $2n_1^2 n_2 \cdots n_d$ operations. This proves

$$N_{(13.25a)}^{\mathrm{full}} = 2\left(\sum_{j=1}^{d} n_j\right) \leq 2dn^{d+1}. \tag{13.26a}$$

Next we consider the tensor $\mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)} \in \mathcal{R}_r$ in r-term format. Multiplication by $\mathbf{A}$ from (13.25a) leads to the following cost and representation rank:

$$N_{(13.25a)}^{\mathcal{R}_r} = r \sum_{j=1}^{d} (2n_j - 1)\, n_j \leq 2drn^2, \quad \mathbf{Av} \in \mathcal{R}_{d \cdot r}. \tag{13.26b}$$

In the tensor subspace case, $\mathbf{v} \in \bigotimes_{j=1}^{d} U_j$ is mapped into $\mathbf{w} = \mathbf{Av} \in \bigotimes_{j=1}^{d} Y_j$, where $U_j \subset V_j$ and $Y_j \subset W_j$. Its representation is $\mathbf{v} = \sum_{\mathbf{k} \in \mathbf{J}} \mathbf{a_k} \bigotimes_{j=1}^{d} b_{k_j}^{(j)}$. For $\mathbf{A}$ from (13.25a) the resulting subspaces are $Y_j = \operatorname{span}\{U_j, A^{(j)} U_j\}$, which can be generated by the frame $(b_k^{(j)}, A^{(j)} b_k^{(j)} : 1 \leq k \leq r_j)$ of size $r_j^{\mathbf{w}} := 2r_j$. Denote these vectors by $(b_{k,\mathbf{w}}^{(j)})_{1 \leq k \leq r_j^{\mathbf{w}}}$ with

$$b_{k,\mathbf{w}}^{(j)} := b_k^{(j)} \text{ and } b_{k+r_j,\mathbf{w}}^{(j)} := A^{(j)} b_k^{(j)} \qquad \text{for } 1 \leq k \leq r_j.$$

Then $\mathbf{w} = \sum_{\mathbf{k} \in \mathbf{J_w}} \mathbf{b_k} \bigotimes_{j=1}^{d} b_{k_j,\mathbf{w}}^{(j)}$ holds with

$$\begin{array}{lll} \mathbf{b}_{k_1 \cdots k_{j-1}, k_j + r_j, k_{j+1} \cdots k_d} & = \mathbf{a}_{k_1 \cdots k_{j-1}, k_j, k_{j+1} \cdots k_d} & \text{for } 1 \leq k_\ell \leq r_\ell,\ 1 \leq j, \ell \leq d, \\ \mathbf{b_k} & = 0 & \text{otherwise.} \end{array}$$

The only arithmetical computations occur for $b_{k+r_j,\mathbf{w}}^{(j)} := A^{(j)} b_k^{(j)}$ (cost: $(2n_j - 1) n_j$ operations), while $\mathbf{b_k}$ needs only copying of data. However, note that the size of the coefficient tensor $\mathbf{b}$ is increased by 2^d: the new index set $\mathbf{J_w}$ has the cardinality $\#\mathbf{J_w} = \prod_{j=1}^{d} r_j^{\mathbf{w}} = 2^d \prod_{j=1}^{d} r_j = 2^d \#\mathbf{J}$. We summarise:

$$N_{(13.25a)}^{\mathcal{T}_\mathbf{r}} = \sum\nolimits_{j=1}^{d} (2n_j - 1)\, n_j \leq 2drn^2,\ r_j^{\mathbf{w}} = 2r_j,\ \#\mathbf{J_w} = 2^d \#\mathbf{J}. \tag{13.26c}$$

For $\mathbf{v}$ given in the hierarchical format, we obtain

$$N_{(13.25a)}^{\mathcal{H}_\mathfrak{r}} \leq 2drn^2, \tag{13.26d}$$

as detailed for the case of (13.25b) above.

13.9.3 Elementary Kronecker Tensor (13.25b)

For $\mathbf{v}$ in full format, the multiplication of $\mathbf{A}$ from (13.25b) by $\mathbf{v}$ requires

$$N_{(13.25b)}^{\text{full}} = 2 \sum_{j=1}^{d} \left(\prod_{k=1}^{j} m_k \right) \left(\prod_{k=j}^{d} n_k \right) \leq 2dn^{d+1}, \tag{13.27a}$$

operations, where $n := \max_j \{n_j, m_j\}$.

The r-term format $\mathbf{v} = \sum_{i=1}^{r} \bigotimes_{j=1}^{d} v_i^{(j)} \in \mathcal{R}_r$ requires to compute $A^{(j)} v_i^{(j)}$ leading to

$$N_{(13.25b)}^{\mathcal{R}_r} = r \sum\nolimits_{j=1}^{d} (2n_j - 1)\, m_j \leq 2drn^2, \quad \mathbf{Av} \in \mathcal{R}_r. \tag{13.27b}$$

For the tensor subspace format $\mathbf{v} = \sum_{\mathbf{k}\in\mathbf{J}} \mathbf{a_k} \bigotimes_{j=1}^{d} b_{k_j}^{(j)} \in \mathcal{T}_\mathbf{r}$ we obtain $\mathbf{w} = \mathbf{Av} = \sum_{\mathbf{k}\in\mathbf{J_w}} \mathbf{a_k^w} \bigotimes_{j=1}^{d} b_{k_j,\mathbf{w}}^{(j)}$, where, as in §13.9.2,

$$N_{(13.25b)}^{\mathcal{T}_\mathbf{r}} = \sum_{j=1}^{d} (2n_j - 1)\, m_j \leq 2drn^2, \quad r_j^\mathbf{w} = r_j,\ \#\mathbf{J_w} = \#\mathbf{J}. \tag{13.27c}$$

Next, we consider in more detail the case of $\mathbf{v} \in \mathcal{H}_\mathfrak{r}$. Let $\mathbf{A}^{(\alpha)} := \bigotimes_{j\in\alpha} A^{(j)}$ be the partial products. Let $\mathbf{v} = \sum_{i=1}^{r_D} c_i^{(D)} \mathbf{b}_i^{(D)} \in \mathcal{H}_\mathfrak{r}$. The product $\mathbf{w} = \mathbf{Av} = \mathbf{A}^{(D)}\mathbf{v} = \sum_{i=1}^{r_D} c_i^{(D)} \mathbf{A}^{(D)} \mathbf{b}_i^{(D)}$ satisfies the recursion

$$\mathbf{A}^{(\alpha)} b_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)} \left(\mathbf{A}^{(\alpha_1)} \mathbf{b}_i^{(\alpha_1)}\right) \otimes \left(\mathbf{A}^{(\alpha_2)} \mathbf{b}_j^{(\alpha_2)}\right) \quad (\{\alpha_1, \alpha_2\} = S(\alpha))\,.$$

At the leaves, $A^{(j)} b_i^{(j)}$ is to be computed for all $1 \leq j \leq d$. Defining frames with $\mathbf{b}_{\ell,\mathbf{w}}^{(\alpha)} := \mathbf{A}^{(\alpha)} \mathbf{b}_\ell^{(\alpha)}$ for all $\alpha \in T_D$ and $1 \leq \ell \leq r_\alpha$, we obtain the representation $\mathbf{w} \in \mathcal{H}_\mathfrak{r}$ with *identical* coefficient matrices $C^{(\alpha,\ell)}$ and $c_i^{(D)}$. Note that the frame vectors $\mathbf{b}_{\ell,\mathbf{w}}^{(\alpha)}$ are to be computed for the leaves $\alpha \in \mathcal{L}(T_D)$ only, i.e., for $\alpha = \{j\}$, $1 \leq j \leq d$. Therefore, the computational work is

$$N_{(13.25b)}^{\mathcal{H}_\mathfrak{r}} = \sum\nolimits_{j=1}^{d} (2n_j - 1)\, m_j \leq 2drn^2, \tag{13.27d}$$

while the data size is unchanged.

13.9.4 Matrix in p-Term Format (13.25c)

The general case $\mathbf{A} = \sum_{i=1}^{p} \bigotimes_{j=1}^{d} A_i^{(j)}$ from (13.25c) requires the p-fold work compared with (13.25b) plus the cost for $p-1$ additions of vectors:

$$N_{(13.25c)}^{\text{full}} = pN_{(13.25b)}^{\text{full}} \leq 2pdn^{d+1}, \tag{13.28a}$$

$$N_{(13.25c)}^{\mathcal{R}_p} \leq 2prdn^2, \quad \mathbf{w} \in \mathcal{R}_{p\cdot r}, \tag{13.28b}$$

$$N_{(13.25c)}^{\mathcal{T}_\mathbf{r}} \leq 2prdn^2 + (p-1) N_+^{\mathcal{T}_\mathbf{r}}, \tag{13.28c}$$

$$N_{(13.25c)}^{\mathcal{H}_\mathfrak{r}} \leq 2pdrn^2 + (p-1) N_+^{\mathcal{H}_\mathfrak{r}}. \tag{13.28d}$$

In the first two cases the addition is either of lower order (full format) or free of cost. The values of $N_+^{\mathcal{T}_\mathbf{r}}$ and $N_+^{\mathcal{H}_\mathfrak{r}}$ depend on the choice frame versus basis. In the latter case, $N_+^{\mathcal{T}_\mathbf{r}} \leq 2dnr^2 + 2dr^{d+1}$ and $N_+^{\mathcal{H}_\mathfrak{r}} \leq 8dnr^2 + 8dr^4$.

13.10 Functions of Tensors, Fixed Point Iterations

Given a function $f : D \subset \mathbb{C} \to \mathbb{C}$ and a matrix A with spectrum in D, a matrix $f(A)$ can be defined.[8] Details about functions of matrices can be found in Higham [98] or Hackbusch [86, §13]. Examples of such matrix functions are $\exp(A)$, $\exp(tA)$, or $A^{1/2}$, but also the inverse A^{-1} (corresponding to $f(z) = 1/z$).

In particular cases, there are fixed point iterations converging to $f(A)$. In the case of A^{-1}, the Newton method yields the iteration

$$X_{m+1} := 2X_m - X_m A X_m, \tag{13.29}$$

which shows local, quadratic convergence. A possible starting value is $X_0 := I$.

Equation (13.29) is an example of a fixed point iteration. If the desired tensor satisfies $X^* = \Phi(X^*)$, the sequence

$$X_{m+1} := \Phi(X_m)$$

converges to X^*, if Φ is contractive. Assuming that the evaluation of Φ involves only the operations studied in this chapter, $\Phi(X_m)$ is available. However, since the operations cause an increase of the representation ranks, the next iteration must be preceded by a truncation:

$$\tilde{X}_{m+1} := T(\Phi(X_m)) \qquad (T\text{: truncation}).$$

The resulting iteration is called 'truncated iteration' and studied in Hackbusch-Khoromskij-Tyrtyshnikov [93] (see also Hackbusch [86, §14.3.2]). In essence, the error decreases as in the original iteration until the iterates reach an X^* neighbourhood of the size of the truncation error (X^*: exact solution).

For the particular iteration (13.29) converging to $X^* = A^{-1}$, a suitable modification is proposed by Oseledets-Tyrtyshnikov [157] (see also [156]). The iteration for H_k, Y_k, X_k is defined by

$$H_k := T_0(2I - Y_k), \quad Y_{k+1} := T_1(Y_k H_k), \quad X_{k+1} := T_1(X_k H_k)$$

and uses a standard truncation T_1 and a possibly rougher truncation T_0. In the exact case (no truncation), $H_k \to I$, $Y_k \to I$, $X_k \to A^{-1}$ holds.

We conclude that approximations to A^{-1} can be determined iteratively, provided we have a sufficient starting value.

For other functions like $A^{1/2}$ and $\exp(tA)$ we refer to [86, §14.3.1] and [86, §14.2.2.2], respectively.

A useful and seemingly simple (nonlinear) function is the maximum of a tensor $\mathbf{v} \in \mathbf{V} = \bigotimes_{j=1}^{d} \mathbb{R}^{I_j} \cong \mathbb{R}^{\mathbf{I}}$ ($\mathbf{I} = \times_{j=1}^{d} I_j$):

$$\max(\mathbf{v}) := \max\{\mathbf{v_i} : \mathbf{i} \in \mathbf{I}\}.$$

[8] In the case of a general function, A must be diagonalisable.

Since $\min(\mathbf{v}) = -\max(-\mathbf{v})$, this function allows us to determine the maximum norm $\|\mathbf{v}\|_\infty$ of a tensor. The implementation is trivial for an elementary tensor:

$$\max\left(\bigotimes_{j=1}^{d} v^{(j)}\right) = \prod_{j=1}^{d} \max\left(v^{(j)}\right);$$

however, the implementation for general tensors is not straightforward. A possible approach, already described in Espig [52] and also contained in [55, §4.1], is based on the reformulation as an eigenvalue problem. The tensor $\mathbf{v} \in \mathbf{V}$ corresponds to a multiplication operator $\mathbf{G}(\mathbf{v})$ defined in (13.22). Let $\mathbf{I}^* := \{\mathbf{i} \in \mathbf{I} : \max(\mathbf{v}) = \mathbf{v}_\mathbf{i}\}$ be the index subset where the maximum is attained. Then the eigenvalue problem

$$\mathbf{G}(\mathbf{v})\mathbf{u} = \lambda\mathbf{u} \qquad (0 \neq \mathbf{u} \in \mathbf{V}) \tag{13.30}$$

has the maximal eigenvalue $\lambda = \max(\mathbf{v})$. The eigenspace consists of all vectors $\mathbf{u}$ with support in $\mathbf{I}^*$. In particular, if $\mathbf{I}^* = \{\mathbf{i}^*\}$ is a singleton, the maximal eigenvalue is a simple one and the eigenvector is a multiple of the unit vector $\mathbf{e}^{(\mathbf{i}^*)}$, which has tensor rank 1. Using the simple vector iteration or more advanced methods, we can determine not only the maximum $\max(\mathbf{v})$, but also the corresponding index.

An interesting function in statistics is the characteristic function $\chi_{(a,b)} : \mathbb{R}^\mathbf{I} \to \mathbb{R}^\mathbf{I}$ of an interval $(a,b) \subset \mathbb{R}$ (including $a = \infty$ or $b = \infty$) with the pointwise definition

$$\left(\chi_{(a,b)}(\mathbf{v})\right)_\mathbf{i} := \left\{\begin{matrix} 1 & \mathbf{v}_\mathbf{i} \in (a,b) \\ 0 & \text{otherwise} \end{matrix}\right\} \quad \text{for } \mathbf{v} \in \mathbb{R}^\mathbf{I},\ \mathbf{i} \in \mathbf{I}.$$

This function can be derived from the sign function:

$$(\operatorname{sign}(\mathbf{v}))_\mathbf{i} := \left\{\begin{matrix} +1 & \mathbf{v}_\mathbf{i} > 0 \\ 0 & \mathbf{v}_\mathbf{i} = 0 \\ -1 & \mathbf{v}_\mathbf{i} < 0 \end{matrix}\right\} \quad \text{for } \mathbf{v} \in \mathbb{R}^\mathbf{I},\ \mathbf{i} \in \mathbf{I}.$$

In contrast to (13.30), the tensor $\mathbf{u} := \chi_{(a,b)}(\mathbf{v})$ may have large tensor rank, even for an elementary tensor $\mathbf{v}$. However, in cases of rare events, $\mathbf{u}$ is sparse (cf. Remark 7.2). In Espig et al. [55, §4.2] an iteration for computing $\operatorname{sign}(\mathbf{v})$ is proposed, using either

$$\mathbf{u}^k := T\left(\tfrac{1}{2}\mathbf{u}^{k-1} + (\mathbf{u}^{k-1})^{-1})\right) \qquad (T\text{: truncation}) \tag{13.31a}$$

or

$$\mathbf{u}^k := T\left(\tfrac{1}{2}\mathbf{u}^{k-1} \odot (3 \cdot \mathbf{1} - \mathbf{u}^{k-1} \odot \mathbf{u}^{k-1})\right) \tag{13.31b}$$

with the constant tensor $\mathbf{1}$ of value 1 ($\mathbf{1}_\mathbf{i} = 1$). Iteration (13.31a) requires a secondary iteration for the pointwise inverse

$$\left((\mathbf{u}^{k-1})^{-1}\right)_\mathbf{i} := 1/\mathbf{u}_\mathbf{i}^{k-1} \qquad \text{for } \mathbf{i} \in \mathbf{I}.$$

For numerical examples see [55, §6].

13.11 Example: Operations for Quantum Chemistry Applications

The stationary electronic Schrödinger equation

$$\mathcal{H}\varPsi := \left[-\frac{1}{2}\sum_{i=1}^{d} \varDelta_i - \sum_{k=1}^{M}\sum_{i=1}^{d} \frac{Z_k}{|\mathbf{x}_i - \mathbf{R}_k|} + \sum_{1\leq i<j\leq d} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} \right] \varPsi = \lambda\varPsi, \qquad (13.32)$$

is an eigenvalue problem for a normed 'wave' function[9] $\varPsi(\mathbf{x}_1, ..., \mathbf{x}_d) \in D(\mathcal{H}) \cap \mathfrak{A}_d(L_2(\mathbb{R}^{3d})$, which must be antisymmetric because of the Pauli principle. The quantity of utmost interest is the ground state energy, i.e., the lowest eigenvalue λ of $\mathcal{H}$ together with the eigenfunction

$$\varPsi = \operatorname{argmin}\{\langle \mathcal{H}\varPhi, \varPhi\rangle : \ \|\varPhi\|_{L^2} = 1, \varPhi \in D(\mathcal{H}) \cap \mathfrak{A}_d(L_2(\mathbb{R}^3)\}. \qquad (13.33)$$

Formulation (13.33) is a minimisation problem for the Rayleigh quotient of $\mathcal{H}$.

Usually, a direct solution of the linear eigenvalue equation (13.32) is not feasible except for sufficiently small molecules. Larger molecules, consisting of hundreds to thousands of electrons, are nowadays computed by means of single-particle models as the *Hartree-Fock* model and the *Kohn-Sham* model of density functional theory (DFT). These models compute an antisymmetric rank-one approximation

$$\varPsi_{SL}(\mathbf{x}_1, \ldots, \mathbf{x}_d) = \frac{1}{\sqrt{d!}} \det(\varphi_i(\mathbf{x}_j))_{i,j=1}^{d}$$

to the solution (13.33) of (13.32). This *Slater determinant* (cf. Lemma 3.72) is constituted in the closed-shell case by d orthonormal functions $\varphi_1, \ldots, \varphi_d \in L^2(\mathbb{R}^3)$ which now have to be computed. Minimisation of (13.33) over the set of rank-1 functions yields as a necessary condition for the minimiser $\varPhi = (\varphi_1, \ldots, \varphi_d)$ the nonlinear *Hartree-Fock equations*

$$\mathcal{F}_{\varPhi}\varphi_i(\mathbf{x}) = \lambda_i\, \varphi_i(\mathbf{x}) \qquad (1 \leq i \leq d), \qquad (13.34a)$$

where the *Hamilton-Fock operator* $\mathcal{F}_{\varPhi}$ depends on $\varPhi$ and is given by

$$\mathcal{F}_{\varPhi}\varphi(\mathbf{x}) := -\frac{1}{2}\varDelta\varphi(\mathbf{x}) + V_c(\mathbf{x})\,\varphi(\mathbf{x}) + V_H(\mathbf{x})\varphi(\mathbf{x}) + (\mathcal{K}\varphi)\,(\mathbf{x})\,, \qquad (13.34b)$$

with the *core potential* V_c, the *Hartree potential* V_H and the *exchange operator* $\mathcal{K}$ defined via[10]

[9] The dependency on the spin variables $s_i \in \{-1/2, 1/2\}$ is suppressed (closed-shell case). $D(\mathcal{H})$ is the domain of $\mathcal{H}$.

[10] M is the number of nuclei with charge Z_k and position $\mathbf{R}_k \in \mathbb{R}^3$. All integrations are taken over $\mathbb{R}^3$.

$$\rho(\mathbf{x},\mathbf{y}) = 2\sum_{i=1}^{d} \varphi_i(\mathbf{x})\varphi_i(\mathbf{y}), \qquad \mathbf{V}_c(\mathbf{x}) = \sum_{k=1}^{M} \frac{Z_k}{|\mathbf{x}-\mathbf{R}_k|},$$
$$V_H(\mathbf{x}) = \int \frac{\rho(\mathbf{y},\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,\mathrm{d}\mathbf{y}, \qquad (\mathcal{K}(\varphi))(\mathbf{x}) = -\tfrac{1}{2}\int \frac{\rho(\mathbf{x},\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,\varphi(\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

Note that $V_H(\mathbf{x}) = V_H[\varphi_1,\varphi_2,\varphi_3](\mathbf{x})$ depends quadratically on φ_i. The equations (13.34a) form a nonlinearly coupled system of eigenvalue problems, but differently from (13.32) only functions in $\mathbb{R}^3$ are to be determined. Since it is known that the solutions are exponentially decaying for $|\mathbf{x}| \to \infty$, the unboundedness of $\mathbb{R}^3$ is not a severe problem.

The success of tensor computations for this problem depends on the involved representation ranks and therefore on the data compressibility. For a theoretical discussion see Flad-Hackbusch-Schneider [61, 62].

A discretisation of (13.34a) by, e.g., a finite difference method on a regular grid in a finite box with zero boundary conditions replaces functions in $\mathbb{R}^3$ by grid functions in $\bigotimes_{j=1}^{3}\mathbb{R}^{n_j}$ with possibly large n_j. The evaluations of the potentials involve almost all operations studied in this chapter.

1. *Hadamard product*. The application of the core potential, $\varphi \mapsto V_c\varphi$ is an example of the Hadamard product. For this purpose, the function $1/|\bullet - \mathbf{R}_k|$ has to be approximated in one of the formats (see Item 2).

   Another Hadamard product occurs in the computation of the *electron density* $n(\mathbf{y}) := \rho(\mathbf{y},\mathbf{y})$ involved in the Hartree potential V_H, since it requires to square the functions $\varphi_i(\mathbf{y})$. Having computed V_H, we have to perform the Hadamard product $V_H \odot \varphi$ for all $\varphi = \varphi_i$ (cf. (13.34b)).

   A third Hadamard product appears in $\mathcal{K}(\varphi)$. Let $\varphi = \varphi_j$. Using the definition of ρ, we see that

$$\mathcal{K}(\varphi_j) = -\sum_{i=1}^{d} \varphi_i \int \frac{\varphi_i(\mathbf{y})\varphi_j(\mathbf{y})}{|\bullet - \mathbf{y}|}\,\mathrm{d}\mathbf{y}.$$

   The Hadamard product $\varphi_i \odot \varphi_j$ has to be determined for all $1 \le i < j \le 3$ (for $i = j$, the result is already known from $\rho(\mathbf{y},\mathbf{y})$). There is even a fourth Hadamard product $\varphi_i \odot \psi$, where ψ is the result of $\int \frac{\varphi_i(\mathbf{y})\varphi_j(\mathbf{y})}{|\bullet-\mathbf{y}|}\,\mathrm{d}\mathbf{y}$, which is discussed in Item 3. A related problem is the evaluation of two-electron integrals (cf. Benedikt-Auer-Espig-Hackbusch [12]).

2. *Representation of the Coulomb potential*. The techniques from §9.7.2.5.2 yield a k-term representation of the function $1/|\bullet|$. A simple substitution $x_i \to x_i - \mathbf{R}_{k,i}$ in this representation can be used for $1/|\bullet - \mathbf{R}_k|$. Since the representation in $\mathcal{R}_k$ can be transferred easily into any other format, the approximation of $1/|\bullet - \mathbf{R}_k|$ causes no problem.

   The Coulomb potential $1/|\bullet - \mathbf{y}|$ also appears in Item 3.

3. *Convolution.* The function V_H is the convolution result of the electron density $n(\mathbf{y}) = \rho(\mathbf{y},\mathbf{y})$ and the Coulomb potential $1/|\bullet - \mathbf{y}|$, i.e., $V_H = \frac{1}{|\bullet|} \star n$. Lemma 9.30 ensures that the approximation from Item 2 admits an accurate result of V_H. For the performance of the convolution see §13.7, but also the later §14.3.

   The function $\psi = \int \frac{\varphi_i(\mathbf{y})\varphi_j(\mathbf{y})}{|\bullet - \mathbf{y}|} \mathrm{d}\mathbf{y}$ mentioned in Item 1 requires the convolution $\frac{1}{|\bullet|} \star n_{ij}$, where $n_{ij} := \varphi_i \odot \varphi_j$ is determined in Item 1.

4. *Addition.* Finally, $V_c\varphi + V_H\varphi + \mathcal{K}(\varphi)$ has to be added.

5. *Laplace inverse.* An inverse iteration of the (nonlinear) eigenvalue problem (13.34a) starts from the reformulation of (13.34a,b) by

$$\varphi_i = 2\,(-\varDelta)^{-1}\,(\lambda_i\,\varphi_i - V_c\,\varphi_i - V_H\varphi_i - \mathcal{K}(\varphi_i))\,.$$

   As seen in §9.7.2.6 and, in particular, in Remark 9.35, the inverse matrix $(-\varDelta)^{-1}$ possesses a simple and accurate representation in $\mathcal{R}_k$. See also Khoromskij [116].

6. *Matrix-vector multiplication.* The (Kronecker) operator approximating $(-\varDelta)^{-1}$ has to be multiplied by $\lambda_i\,\varphi_i - V_c\,\varphi_i - V_H\varphi_i - \mathcal{K}(\varphi_i)$.

The *Kohn-Sham model* of density functional theory replaces the nonlocal exchange operator $\mathcal{K}$ in the Fock operator $\mathcal{F}_\Phi$ by an *exchange correlation potential* V_{xc} depending only on the *electron density*

$$n(\mathbf{x}) := \rho(\mathbf{x},\mathbf{x}).$$

Although this dependence is not known explicitly, several models developed in physics yield satisfactory results. One of these models uses or the 'local density approximation' (LDA) the so-called exchange part

$$\varepsilon_X(n(\mathbf{x})) = -\frac{3}{4}\left(\frac{3}{\pi}n(\mathbf{x})\right)^{1/3}$$

(cf. Koch-Holthausen [127, §6.4]), which gives rise to the next item.

7. *Function evaluation.* By definition, $n(\mathbf{x})$ is non-negative, so that the function $f(t) := t^{1/3}$ can be applied to each entry of the (grid) function. Here, the generalised cross approximation method can be applied (cf. §15.1.3) or an iterative approach as in §13.10.

A realisation of the computations using the formats R_r, $\mathcal{T}_\mathbf{r}$, and the hybrid format from §8.2.4 can be found in Khoromskij-Khoromskaia-Flad [121] (see also [113]). The tensorisation from §5.3 and §14 has been employed in Khoromskaia-Khoromskij-Schneider [114].

Chapter 14
Tensorisation

Abstract Tensorisation has been introduced by Oseledets [153] (applied to matrices instead of vectors). The tensorised version of a $\mathbb{K}^n$ vector can easily be truncated in a black-box fashion. Under suitable conditions, the data size reduces drastically. Operations applied to these tensors instead of the original vectors have a cost related to the (much smaller) tensor data size. *Section 14.1* describes the main principle, the hierarchical format $\mathcal{H}_{\boldsymbol{\rho}}^{\mathrm{tens}}$ corresponding to the TT format, operations with tensorised vectors, and the generalisation to matrices. The reason, why the data size can be reduced so efficiently is analysed in *Sect. 14.2*. Tensorisation mimics classical analytical approximations method which exploit the smoothness of a function to obtain an approximation with much less degrees of freedom. *Section 14.3* presents in detail the (exact) convolution of vectors performed by means of their tensorisations. It is shown that the cost corresponds to the data size of the tensors. *Section 14.4* is devoted to the tensorised counterpart of the fast Fourier transform (FFT). During the algorithm one has to insert truncation steps, since otherwise a maximal representation rank arises. While the original tensorisation technique applies to discrete data, *Sect. 14.5* generalises the approach to functions.

14.1 Basics

14.1.1 Notations, Choice of T_D

As introduced in §5.3, vectors can be rewritten as tensors. Here, we restrict our discussion to the case of vectors from $\mathbb{K}^n = \mathbb{K}^I$ with $I = \{0, 1, \ldots, n-1\}$ and

$$n = 2^d \tag{14.1a}$$

(cf. Remark 14.1 below). Let

$$\mathbf{V} := \bigotimes_{j=1}^{d} \mathbb{K}^2, \tag{14.1b}$$

where $\mathbb{K}^2 = \mathbb{K}^J$ with $J = \{0, 1\}$. The isomorphism $\Phi_n : \mathbf{V} \to \mathbb{K}^I$ is given by

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_14,

means of the binary integer representation $k = \sum_{j=1}^{d} i_j 2^{j-1}$ $(0 \leq i_j \leq 1)$:

$$\begin{aligned} \Phi_n : \mathbf{V} &\to \mathbb{K}^n \\ \mathbf{v} &\mapsto v \quad \text{with } v_k = \mathbf{v}[i_1 \cdots i_d] \text{ for } k = \sum_{j=1}^{d} i_j 2^{j-1}. \end{aligned} \tag{14.1c}$$

In special cases, the r-term format can be applied successfully to $\mathbf{v} \in \mathbf{V}$. An example is given in Remark 5.18, where the tensor has rank one, i.e., $\mathbf{v} \in \mathcal{R}_1$.

The tensor subspace format $\mathcal{T}_{\mathbf{r}}$ is not of much help for $\mathbf{V} = \otimes^d \mathbb{K}^2$. In general, the success of this format is caused by the fact that the subspace U_j has a much smaller dimension than V_j. Here, $\dim(V_j) = \dim(\mathbb{K}^2) = 2$ is already rather small. Subspaces of dimension $\dim(U_j) = 1$ rarely appear. If, however, $\dim(U_j) = 2$ holds, the tensor subspace format $\mathcal{T}_{(2,\ldots,2)}$ is identical to the full representation.

Therefore, we use the hierarchical format $\mathcal{H}_{\mathfrak{r}}$ from §11 for the representation of $\mathbf{v} \in \mathbf{V}$. Because of the special nature of $\mathbb{K}^2$, we modify the structure of $\mathcal{H}_{\mathfrak{r}}$ as follows. Let $\alpha = \{j\} \in \mathcal{L}(T_D)$ be a leaf-node. The subspace $\mathbf{U}_\alpha = U_j \subset V_j = \mathbb{K}^2$ is chosen as

$$U_j = V_j \qquad (j \in D = \{1, \ldots, d\}) \tag{14.2a}$$

with the *fixed* basis $\mathfrak{b}^{(j)}$ consisting of the unit vectors

$$b_1^{(j)} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right], \quad b_2^{(j)} = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] \qquad (1 \leq j \leq d). \tag{14.2b}$$

This simplification is justified by the fact that in most of the cases, the minimal subspace $U_j = U_j^{\min}(\mathbf{v})$ will be two-dimensional anyway. Fixing the basis by (14.2b) avoids the overhead for coding U_j. As a consequence,

$$r_j = 2 \qquad (1 \leq j \leq d) \tag{14.2c}$$

holds for the representation ranks $r_\alpha = r_j$ of $\mathfrak{r}$ characterising the format $\mathcal{H}_{\mathfrak{r}}$ at the leaves $\alpha = \{j\} \in \mathcal{L}(T_D)$.

Concerning the choice of the dimension partition tree T_D, it will turn out that the linear tree from Fig. 11.2 used for the TT format $\mathbb{T}_{\boldsymbol{\rho}}$ is a good choice:

$$\begin{aligned} &T_D^{\mathsf{TT}} = D \cup \{\{1, \ldots, k\} : k \in D\} \text{ and} \\ &\{1, \ldots, k\} \text{ has the first son } \{1, \ldots, k-1\} \text{ and the second son } \{k\}. \end{aligned} \tag{14.2d}$$

Another choice is a balanced tree T_D^{bal} like in (11.1) or (11.2).

A first argument favouring T_D^{TT} is that the tensor spaces $\mathbf{V}_{\{1,\ldots,k\}} := \bigotimes_{j=1}^{k} \mathbb{K}^2$ correspond to vector spaces $\mathbb{K}^{2^k}$ with a natural interpretation $\mathbb{K}^{2^k} \cong \mathbb{K}^{2^{k-1}} \otimes \mathbb{K}^2$, but also T_D^{bal} has realistic interpretations.

Another convincing argument is the *storage cost* (cf. Remark 11.22). The part $\rho_{\mathrm{mem}}^{\mathrm{HTR}}((\mathbf{B}_\alpha)_{\alpha \in \mathcal{L}(T_D)})$ vanishes because of the fixed choice of the basis (14.2b), while $\rho_{\mathrm{mem}}^{\mathrm{HTR}}(c^{(D)}) = r_D = 1$ can be neglected. It remains the term

$$N_{\mathrm{mem}}^{\mathrm{HTR}}((\mathbf{C}_\alpha)_{\alpha \in T_D \setminus \mathcal{L}(T_D)}) = \sum_{\alpha \in T_D \setminus \mathcal{L}(T_D)} r_\alpha r_{\alpha_1} r_{\alpha_2} \tag{14.3}$$

(α_1, α_2 sons of α). Assume $r_\alpha \le r$. In the case of T_D^{TT}, the property $\alpha_2 \in \mathcal{L}(T_D)$ implies $r_{\alpha_2} = 2$ and therefore

$$N_{\text{mem}}^{\text{HTR}}((\mathbf{C}_\alpha)) \le 2\,(d-1)\,r^2.$$

The balanced tree T_D^{bal} contains products $r_\alpha r_{\alpha_1} r_{\alpha_2}$ whose factors may all be of size[1] r. This yields a storage requirement $S(T_D^{\mathsf{bal}}) = O(dr^3)$ cubic in r.

We recall that $\mathbb{T}_{\boldsymbol{\rho}}$ with ranks $\boldsymbol{\rho} = (\rho_1 = 2, \rho_2, \ldots, \rho_d)$ is characterised algebraically by subspaces $\mathbf{U}_j \subset \otimes^j \mathbb{K}^2$ for $1 \le j \le d$ such that

$$\mathbf{U}_1 = \mathbb{K}^2, \tag{14.4a}$$

$$\mathbf{U}_j \subset \mathbf{U}_{j-1} \otimes \mathbb{K}^2 \text{ and } \dim(\mathbf{U}_j) = \rho_j, \qquad (2 \le j \le d), \tag{14.4b}$$

$$\mathbf{v} \in \mathbf{U}_d \tag{14.4c}$$

holds for the represented tensor $\mathbf{v} \in \otimes^d \mathbb{K}^2$.

Remark 14.1. The requirement $n = 2^d$ in (14.1a) is made for the sake of convenience. Otherwise, there are two remedies:
(i) Assume a decomposition $n = \prod_{j=1}^d p_j$ $(1 < p_j \in \mathbb{N}$, cf. §5.3) and use the isomorphism $\mathbb{K}^n \cong \bigotimes_{j=1}^d \mathbb{K}^{p_j}$. The advantage of tensorisation vanishes if there are too large factors p_j.
(ii) One might embed $\mathbb{K}^n$ into $\mathbb{K}^m$ $(m > n)$ with $m = 2^d$ (or modified according to (i)) by replacing $v \in \mathbb{K}^n$ by $\tilde{v} \in \mathbb{K}^m$ with $\tilde{v}_i = v_i$ $(1 \le i \le n)$ and $\tilde{v}_i = 0$ $(n < i \le m)$. Many operations can be performed with $\tilde{v}$ instead of v.

14.1.2 Format $\mathcal{H}_{\boldsymbol{\rho}}^{\text{tens}}$

Let $\boldsymbol{\rho} = (\rho_0 = 1, \rho_1 = 2, \rho_2, \ldots, \rho_{d-1}, \rho_d = 1)$. In the following we use a particular hierarchical format[2] $\mathcal{H}_{\boldsymbol{\rho}}^{\text{tens}}$ based on the linear tree T_D^{TT}, which is almost identical to $\mathbb{T}_{\boldsymbol{\rho}}$. The parameters of $\mathbf{v} \in \mathcal{H}_{\boldsymbol{\rho}}^{\text{tens}}$ are

$$\mathbf{v} = \rho_{\text{HTR}}^{\text{tens}}\big((\mathbf{C}_j)_{j=2}^d, c^{(D)}\big). \tag{14.5a}$$

The coefficients $(\mathbf{C}_j)_{j=1}^d$ with

$$\mathbf{C}_j = \left(C^{(j,\ell)}\right)_{1 \le \ell \le \rho_j} \qquad \text{and} \qquad C^{(j,\ell)} = \left(c_{ik}^{(j,\ell)}\right)_{\substack{1 \le i \le \rho_j \\ 1 \le k \le 2}} \in \mathbb{K}^{\rho_j \times 2}$$

define the basis vectors $\mathbf{b}_\ell^{(1,\ldots,j)}$ recursively:

$$\mathbf{b}_\ell^{(1)} = b_\ell^{(1)} \qquad \text{for } j = 1 \text{ and } 1 \le \ell \le 2, \tag{14.5b}$$

$$\mathbf{b}_\ell^{(j)} = \sum_{i=1}^{\rho_{j-1}} \sum_{k=1}^{2} c_{ik}^{(j,\ell)}\, \mathbf{b}_i^{(j-1)} \otimes b_k^{(j)} \quad (1 \le \ell \le \rho_j) \quad \text{for } j = 2, \ldots, d. \tag{14.5c}$$

[1] Of course, the maximum ranks appearing in T_K^{TT} and T_K^{bal} may be different.

[2] The TT format applied to tensorised quantities has also been called QTT ('quantised TT' or 'quantics TT' inspite of the meaning of quantics; cf. [114]).

Here, the bases $b_\ell^{(1)}$ and $b_k^{(j)}$ are the unit vectors from (14.2b). Finally, the tensor $\mathbf{v}$ is defined by[3]

$$\mathbf{v} = c^{(D)}\mathbf{b}_1^{(d)}. \tag{14.5d}$$

The differences to the usual hierarchical format are: (i) fixed tree T_D^{TT}, (ii) fixed unit bases (14.2b), (iii) vertex $\alpha = \{1, \ldots, j\}$ is abbreviated by j in the notations $\mathbf{b}_\ell^{(j)}$ (instead of $\mathbf{b}_\ell^{(\alpha)}$) and $c_{ik}^{(j,\ell)}$ (instead of $c_{ik}^{(\alpha,\ell)}$), (iv) $\rho_d = 1$ simplifies (14.5d).

The data $\mathbf{C}_j$ from (14.5a) almost coincide with those of the $\mathbb{T}_{\boldsymbol{\rho}}$ format (12.1a). For the precise relation we refer to (12.11).

Once, the data of $v \in \mathbb{K}^n$ are approximated by $\tilde{\mathbf{v}} \in \mathcal{H}_{\boldsymbol{\rho}}^{\text{tens}}$, the further operations can be performed within this format (cf. §14.1.3). The question remains, how $v \in \mathbb{K}^n$ can be transferred to $\mathbf{v} = \Phi_n^{-1}(v) \in \mathcal{H}_{\boldsymbol{\rho}}^{\text{tens}}$ and then approximated by some $\tilde{\mathbf{v}} \in \mathcal{H}_{\boldsymbol{\rho}}^{\text{tens}}$. An exact representation by $\mathbf{v} = \Phi_n^{-1}(v)$ is possible, but requires to touch all n data. Hence, the cost may be much larger than the later data size of $\tilde{\mathbf{v}}$. Nevertheless, this is the only way if we require an exact approximation error bound. A much cheaper, but heuristic approach uses the generalised cross approximation tools from §15.

14.1.3 Operations with Tensorised Vectors

In the sequel, we assume (14.1a,b) and represent $\mathbf{v} \in \otimes^d\mathbb{K}^2$ by the $\mathcal{H}_{\boldsymbol{\rho}}^{\text{tens}}$ format (14.5a) with representation ranks ρ_j. The family $\mathbf{C}_j$ of matrices in (14.5a) is assumed to be orthonormal with respect to the Frobenius norm implying that the bases in (14.5b,c) are orthonormal.

The *storage* size of $\mathbf{v}$ follows from (14.3) with $r_{\alpha_2} = 2$: $S = 2\sum_{j=2}^d \rho_j\rho_{j-1}$.

The *addition* of two tensors $\mathbf{v}, \mathbf{w} \in \otimes^d\mathbb{K}^2$ with identical data $(\mathbf{C}_j)_{2\leq j\leq d}$ is trivial. In the standard case, there are different data $\mathbf{C}_j^{\mathbf{v}}$ and $\mathbf{C}_j^{\mathbf{w}}$ and the procedure **JoinBases** is to be applied (cf. (11.71a) and Remark 11.67). The arithmetical cost $N_{\text{QR}}(r_{\alpha_1} \cdot r_{\alpha_2}, r'_\alpha + r''_\alpha)$ mentioned in Remark 11.67b becomes[4] $N_{\text{QR}}(2(\rho_{j-1}^{\mathbf{v}} + \rho_{j-1}^{\mathbf{w}}), \rho_j^{\mathbf{v}} + \rho_j^{\mathbf{w}}) \leq 8\rho^3$ for $\rho := \max_j\{\rho_j^{\mathbf{v}}, \rho_j^{\mathbf{w}}\}$.

The *entry-wise evaluation* of $\mathbf{v}$ from (14.5a) costs $2\sum_{j=2}^d \rho_{j-1}\rho_j$ operations as seen from (13.2). The latter computation uses the $\mathcal{H}_{\boldsymbol{\rho}}^{\text{tens}}$ data which are directly given by (12.11).

The *scalar product* $\langle \mathbf{u}, \mathbf{v}\rangle$ of two tensors with identical data $(\mathbf{C}_j)_{2\leq j\leq d}$ is trivial as seen from (13.9). Otherwise, we may apply the recursion from (13.11). Note that (13.11) simplifies because $\beta_2 \in \mathcal{L}(T_D^{\mathsf{TT}})$ implies $\left\langle \mathbf{b}_j'^{(\beta_2)}, \mathbf{b}_n''^{(\beta_2)} \right\rangle = \delta_{jn}$ (cf. (14.2b)). The cost of the recursion becomes $7\sum_{j=2}^d \rho_j^{\mathbf{v}}\rho_j^{\mathbf{w}}\rho_{j-1}^{\mathbf{v}}\rho_{j-1}^{\mathbf{w}}$. Alternatively, we may join the bases as for the addition above. In fact, this approach is cheaper, since it is only cubic in the ranks:

[3] Here, we make use of $\rho_d = 1$. For $\rho_d > 1$, one has $\mathbf{v} = \sum_{\ell=1}^{\rho_d} c_\ell^{(D)}\mathbf{b}_\ell^{(d)}$. In the latter case, several tensors can be based of the same parameters $(\mathbf{C}_j)_{j=2}^d$.

[4] The transformations from (11.70a-c) do not appear, since the bases B_j are fixed.

$$4\sum_{j=2}^{d}\left(\rho_j^{\mathbf{v}}+\rho_j^{\mathbf{w}}\right)^2\left(\rho_{j-1}^{\mathbf{v}}+\rho_{j-1}^{\mathbf{w}}\right). \tag{14.6}$$

A *binary operation* $\boxdot$ between tensors of $\mathbf{V}$ and $\mathbf{W}$ can be performed as in §13.5.3, however, there are two particular features. First, the ranks r_{α_2} for the second son—which is a leaf—is $r_{\alpha_2}=2$ (cf. (14.2c)). In the matrix case it may be $r_{\alpha_2}=4$ (cf. §14.1.6). Second, the results of $\mathbf{b}_j^{\prime(\alpha_2)}\boxdot\mathbf{b}_n^{\prime\prime(\alpha_2)}$ in (13.20b) are explicitly known. The basis vectors $\mathbf{b}_j^{\prime(\alpha_2)}$ are from the set $\{\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right],\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\}$ and their $\boxdot$ products are again either zero or belong to this set (at least for all $\boxdot$ considered here). An example is the Hadamard product $\boxdot=\odot$, where $\mathbf{b}_j^{\prime(\alpha_2)}\odot\mathbf{b}_n^{\prime\prime(\alpha_2)}=\delta_{jn}\mathbf{b}_n^{(\alpha_2)}$. Hence, the general frame $\mathfrak{b}^{(\alpha_2)}$ used in the algorithm of §13.5.3 can be replaced by $\{\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right],\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\}$. Correspondingly, the computational cost is reduced to

$$2\sum_{j=1}^{d-1}\rho_j^{\mathbf{v}}\rho_j^{\mathbf{w}}. \tag{14.7}$$

The *Hadamard product* is invariant with respect to the tensorisation (cf. (14.1c)):

$$\Phi_n\left(\mathbf{v}\odot\mathbf{w}\right)=\Phi_n(\mathbf{v})\odot\Phi_n(\mathbf{w}), \tag{14.8}$$

i.e., the tensorisation of the vector-wise Hadamard product $v\odot w$ is expressed by the tensor-wise Hadamard product $\mathbf{v}\odot\mathbf{w}$. This binary operation is already mentioned above.

As in §12.2.7, the HOSVD bases can be computed, on which the *truncation* is based. The corresponding computational cost described in (12.14) for the general case becomes

$$4\sum_{j=2}^{d}\rho_{j-1}^2\left(\rho_{j-2}+2\rho_j+\frac{4}{3}\rho_{j-1}\right)\leq\frac{52}{3}\left(d-1\right)\left(\max_j\rho_j\right)^3. \tag{14.9}$$

14.1.4 Application to Representations by Other Formats

The tensorisation procedure can be combined with other formats in various ways. In the following, we consider the tensor space $\mathbf{V}=\bigotimes_{j=1}^{d}V_j$ with $V_j=\mathbb{K}^{n_j}$ and assume for simplicity that $n_j=2^{\delta_j}$ ($\delta_j\in\mathbb{N}$). The tensorisation of the spaces $V_j\cong\mathbf{V}_j=\otimes^{\delta_j}\mathbb{K}^2$ leads to

$$\mathbf{V}=\bigotimes_{j=1}^{d}V_j\;\cong\;\hat{\mathbf{V}}:=\bigotimes_{j=1}^{d}\left(\bigotimes_{\kappa=1}^{\delta_j}\mathbb{K}^2\right). \tag{14.10}$$

14.1.4.1 Combination with r-Term Format

The r-term representation of $\mathbf{v}=\sum_{i=2}^{r}\bigotimes_{j=1}^{d}v_i^{(j)}\in\mathbf{V}$ is based on the vectors $v_i^{(j)}\in V_j$. Following the previous considerations, we replace the vectors $v_i^{(j)}$ by (approximate) tensors $\mathbf{v}_i^{(j)} \in \mathbf{V}_j = \otimes^{\delta_j}\mathbb{K}^2$. For the representation of $\mathbf{v}_i^{(j)}$ we use the $\mathcal{H}_{\boldsymbol{\rho}}^{\mathrm{tens}}$ format (14.5a) involving rank parameters $\boldsymbol{\rho}^{(j)} = (\rho_1^{(j)}, \ldots, \rho_{\delta_j}^{(j)})$.

For $n_j \leq n$, the storage requirement of the r-term format has been described by drn. If a sufficient approximation $\mathbf{v}_i^{(j)}$ of data size $O(\rho^2 \log(n_j)) = O(\rho^2\delta_j)$ and moderate $\rho = \max_\kappa \rho_\kappa^{(j)}$ exists, the storage can be reduced to $O(dr\rho^2 \log(n))$.

Similarly, the cost of operations can be decreased drastically. In Remark 7.12, the cost of the scalar product in V_j is denoted by N_j. In the standard case $V_j = \mathbb{K}^{n_j}$ we expect $N_j = 2n_j - 1$ arithmetical operations. Now, with $v_i^{(j)}$ and $w_i^{(j)}$ replaced by $\mathbf{v}_i^{(j)}, \mathbf{w}_i^{(j)} \in \mathbf{V}_j$, the cost of the scalar product is

$$N_j = O(\rho^3 \log(n_j)) \qquad \text{with} \quad \rho = \max_\kappa\{\rho_\kappa^{\mathbf{v},(j)}, \rho_\kappa^{\mathbf{w},(j)}\}$$

as can be seen from (14.6). Analogously, the cost of the further operations improves.

14.1.4.2 Combination with Tensor Subspace Format

The data of the tensor subspace format are the coefficient tensor $\mathbf{a}$ and the basis (frame) vectors $b_i^{(j)} \in V_j$ (cf. (8.6c)). Tensorisation can be applied to $b_i^{(j)} \in V_j$ with the same reduction of the storage for $b_i^{(j)}$ as above. Unfortunately, the coefficient tensor $\mathbf{a}$ which requires most of the storage, is not affected. The latter disadvantages can be avoided by using the hybrid format (cf. §8.2.4).

All operations with tensors from $\mathcal{T}_{\mathbf{r}}$ lead to various operations between the basis vectors from V_j. If $b_i^{(j)} \in V_j$ is expressed by the tensorised version $\mathbf{b}_i^{(j)} \in \otimes^{\delta_j}\mathbb{K}^2$, these operations may be performed much cheaper. For instance, the convolution in V_j can be replaced by the algorithm described in §14.3 below.

14.1.4.3 Combination with the Hierarchical Format

There are two equivalent ways of integration into the hierarchical format. First, we may replace all basis vectors in $B_j = \left[b_1^{(j)}, \ldots, b_{r_j}^{(j)}\right] \in (U_j)^{r_j}$ by their tensorised version $\mathbf{b}_i^{(j)} \in \otimes^{\delta_j}\mathbb{K}^2$ using the $\mathcal{H}_{\boldsymbol{\rho}}^{\mathrm{tens}}$ format (14.5a). Consequently, all operations involving $b_i^{(j)}$ are replaced by the corresponding tensor operations for $\mathbf{b}_i^{(j)}$.

The second, simpler interpretation extends the dimension partition tree T_D of the hierarchical format. Each leaf vertex $\{j\}$ is replaced by the root of the linear tree $T_{\Delta_j}^{\mathsf{TT}}$ used for the tensorisation, where $\Delta_j = \{1, \ldots, \delta_j\}$ (see Fig. 14.1). The resulting extended tree is denoted by T_D^{ext}. The set $\mathcal{L}(T_D^{\mathsf{ext}})$ of its leaves is the union $\bigcup_{j=1}^d \mathcal{L}(T_{\Delta_j}^{\mathsf{TT}})$ of the leaves of $T_{\Delta_j}^{\mathsf{TT}}$. Hence, $\dim(\mathbf{V}^{(\alpha)}) = 2$ holds for all $\alpha \in \mathcal{L}(T_D^{\mathsf{ext}})$. One may interpret T_D^{ext} as the dimension partition tree for the tensor space $\hat{\mathbf{V}}$ from (14.10), where a general (possibly balanced) tree structure is combined with the linear tree structure below the vertices $\alpha \in \mathcal{L}(T_D)$, which are now inner vertices of T_D^{ext}.

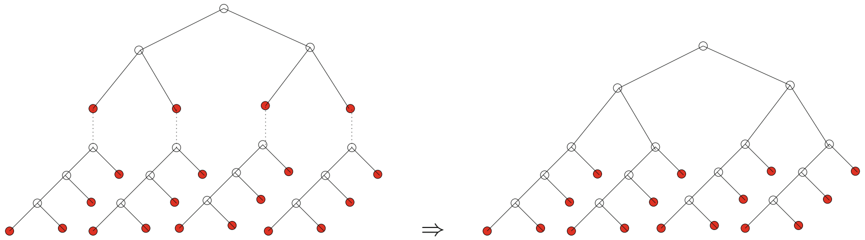


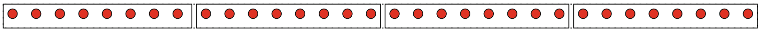
Fig. 14.1 *Left:* balanced tree with 4 leaves corresponding to $V_j = \mathbb{K}^{16}$. The isomorphic tensor spaces $\otimes^4\mathbb{K}^2$ are treated by the linear trees below. *Right:* Extended tree.

14.1.5 Matricisation

The vertices of the linear tree T_D^{TT} are $\{1,\ldots,j\}$ for $j = 1,\ldots,d$. In Definition 5.3 the matricisation $\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v})$ is defined. In this case, $\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v})$ can easily be described by means of the generating *vector* $v = \Phi_n(\mathbf{v}) \in \mathbb{K}^n$ $(n = 2^d)$:

$$\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v}) = \begin{bmatrix} v_0 & v_{2^j} & \cdots & v_{2^{d-1}} \\ v_1 & v_{2^j+1} & \cdots & v_{2^{d-1}+1} \\ \vdots & \vdots & & \vdots \\ v_{2^j-1} & v_{2^{j+1}-1} & \cdots & v_{2^d-1} \end{bmatrix}. \tag{14.11}$$

Hence, the columns of $\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v})$ correspond to blocks of the vector with block size 2^j. An illustration is given below for $M_3(\mathbf{v})$ in the case of $n = 32$. The columns of $M_3(\mathbf{v})$ consists of the four parts of the vector.



(14.12)

We recall that $\rho_j = \operatorname{rank}(\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v}))$.

14.1.6 Generalisation to Matrices

As mentioned above, the original description of the tensorisation technique by Oseledets [153] applies to matrices. Let M be a matrix of size $n \times n$ with $n = 2^d$. Since $\dim(\mathbb{K}^{n\times n}) = n^2 = (2^d)^2 = 4^d$, the matrix space $\mathbb{K}^{n\times n}$ is isomorphic to $\otimes^d\mathbb{K}^{2\times 2}$, the tensor product of 2×2 matrices. A possible isomorphism is given by the following counterpart of Φ_n from (14.1c):

$$\begin{aligned} \Phi_{n\times n} : \mathbf{M} \in \textstyle\bigotimes_{j=1}^d \mathbb{K}^{2\times 2} &\mapsto M \in \mathbb{K}^{n\times n} \\ M[\nu,\mu] &= \mathbf{M}[(\nu_1,\mu_1),\ldots,(\nu_d,\mu_d)] \\ \text{with } \nu &= \textstyle\sum_{j=1}^d \nu_j 2^{j-1},\ \mu = \sum_{j=1}^d \mu_j 2^{j-1}. \end{aligned}$$

The latter definition corresponds to the $(d-1)$-fold application of the Kronecker product (1.5) to 2×2 matrices.

Again, the hierarchical format with the linear tree T_D^{TT} can be used to represent tensors $\mathbf{M} \in \mathbf{V} := \otimes^d \mathbb{K}^{2\times 2}$. Differently from the definitions in (14.2b,c), we now have

$$r_j = 4,\ b_1^{(j)} = \left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right],\ b_2^{(j)} = \left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right],\ b_3^{(j)} = \left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right],\ b_4^{(j)} = \left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right].$$

This fact increases some constants in the storage cost, but does not change the format in principle.

Now, matricisation corresponds to a block representation of the matrix with block size $2^j \times 2^j$ as illustrated in Fig. 14.2. This means that the columns of the matrix $\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v})$ are formed by these subblocks.

Concerning operations, matrix operations are of particular interest. The multiplication $M'M''$ is a binary operation. The operation count (14.7) holds with a factor 4 instead of 2 because of $r_j = 4$. Also the matrix-vector multiplication Mv of a matrix $M \in \mathbb{K}^{n\times n}$ by $v \in \mathbb{K}^n$ using the tensor counterparts $\mathbf{M} \in \otimes^d \mathbb{K}^{2\times 2}$ and $\mathbf{v} \in \otimes^d \mathbb{K}^2$ is of the same kind.

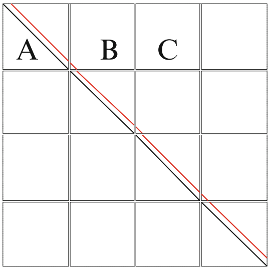


Fig. 14.2 Matricisation for $n = 32,\ j = 3$

Finally, we discuss the ranks ρ_j of certain Toeplitz matrices.[5] The identity matrix or any diagonal matrix with 2^j periodic data has rank $\rho_j = 1$, since the range of $\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v})$ is spanned by one diagonal block.

A banded upper triangular Toeplitz matrix consisting of the diagonal M_{ii} and off-diagonals $M_{i,i+\ell}$ $(1 \le \ell \le 2^j)$ has rank $\rho_j = 2$. The proof can be derived from Fig. 14.2. The $2^j \times 2^j$ blocks appearing in the matrix are either of type A and B, or zero (type C). A simple example of this kind are tridiagonal Toeplitz matrices, which appear as discretisations of one-dimensional differential equations with constant coefficients.

If, however, the band width increases and an off-diagonal $M_{i,i+\ell}$ with $2^j + 1 \le \ell \le 2 \cdot 2^j$ occurs, a new nonzero block appears at position C, leading to rank 3. Correspondingly, for a fixed $\ell \in \{3, \ldots, 2^j\}$, the rank becomes larger than $\rho_j = 2$ for sufficiently small block size $2^k \times 2^k$, when $2^k < \ell \le 2^j$.

A general banded Toeplitz matrix with diagonals $M_{i,i+\ell}$ for $-2^j \le \ell \le 2^j$ has rank $\rho_j = 3$, because a transposed version of block C appears in Fig. 14.2.

Explicit TT representations of Laplace and related matrices are given in [111].

Remark 14.2. A general, fully populated Toeplitz matrix satisfies

$$\rho_j \le \min\{2^{d-j+1}, 2^{j+1}\} - 1.$$

Proof. Because of the Toeplitz structure, there are at most $2^{d-j+1} - 1$ blocks of size $2^j \times 2^j$. Each block is characterised by data forming a vector in $\mathbb{K}^{2^{j+1}-1}$. This fact bounds the number of linearly independent blocks by $2^{j+1} - 1$. □

[5] A Toeplitz matrix M is defined by the property that the entries M_{ij} depend on $i - j$ only.

14.2 Approximation of Grid Functions

14.2.1 Grid Functions

In the following, we assume that the vector $v \in \mathbb{K}^n$, $n = 2^d$, is a grid function, i.e.,

$$v_k = f(a + kh) \quad \text{for } 0 \le k \le n-1, \quad h := \frac{b-a}{n}, \tag{14.13a}$$

where $f \in C([a,b])$ is sufficiently smooth. If $f \in C((a,b])$ has a singularity at $x = a$, the evaluation of $f(a)$ can be avoided by the definition

$$v_k = f(a + (k+1)h) \quad \text{for } 0 \le k \le n-1. \tag{14.13b}$$

Any approximation $\tilde{f}$ of f yields an approximation $\tilde{v}$ of v.

Remark 14.3. Let $f \in C([a,b])$. For $j \in D$ and $k = 0, \ldots, 2^{d-j} - 1$ consider the functions $f_{j,k}(\bullet) := f(a + k2^j h + \bullet) \in C([0, 2^j h])$ and the subspace

$$F_j := \operatorname{span}\{f_{j,k} : 0 \le k \le 2^{d-j} - 1\}.$$

Then $\operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v}) \le \dim(F_j)$ holds for $\mathbf{v} \in \otimes^d \mathbb{K}^2$ with $v = \Phi_n(\mathbf{v})$ satisfying (14.13a) or (14.13b).

Proof. The k-th columns of $\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v})$ in (14.11) are evaluations of $f_{j,k}$. Therefore, $\operatorname{rank}_{\{1,\ldots,j\}}(\mathbf{v}) = \operatorname{rank}(\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v}))$ cannot exceed $\dim(F_j)$. □

The tensorisation technique is not restricted to discrete grid functions. In §14.5, we shall describe a version for functions.

14.2.2 Exponential Sums

Here, we suppose that the function f can be approximated in $[a,b]$ by

$$f_r(x) := \sum_{\nu=1}^{r} a_\nu \exp(-\alpha_\nu x), \tag{14.14}$$

where $a_\nu, \alpha_\nu \in \mathbb{K}$. Examples with $a_\nu, \alpha_\nu > 0$ are given in §9.7.2.3 together with error estimates for the maximum norm $\|f - f_r\|_\infty$.

If f is periodic in $[a,b] = [0, 2\pi]$, the truncated Fourier sum yields (14.14) with imaginary $\alpha_\nu = [\nu - (r+1)/2]\,\mathrm{i}$ for odd $r \in \mathbb{N}$. Closely related are sine or cosine sums $\sum_{\nu=1}^{r} a_\nu \sin(\nu x)$ and $\sum_{\nu=0}^{r-1} a_\nu \cos(\nu x)$ in $[0, \pi]$.

An example of (14.14) with general complex exponents α_ν is mentioned in §9.7.2.4.

As seen in Remark 5.18, the grid function $v \in \mathbb{K}^n$ corresponding to f_r from (14.14) has a tensorised version $\mathbf{v} = \Phi_n^{-1}(v) \in \otimes^d \mathbb{K}^2$ in $\mathcal{R}_r$:

$$\mathbf{v} = \sum_{\nu=1}^{r} a_\nu \bigotimes_{j=1}^{d} \begin{bmatrix} 1 \\ \exp(-2^{j-1}\alpha_\nu) \end{bmatrix} \tag{14.15}$$

requiring $2rd = 2r \log_2 n$ data.[6]

14.2.3 *Polynomial Approximations for Asymptotically Smooth Functions*

Let $f \in C^\infty((0,1])$ be a function with a possible singularity at $x = 0$ and assume that the derivatives are bounded by

$$\left| f^{(k)}(x) \right| \le C k! x^{-k-a} \quad \text{for all } k \in \mathbb{N},\ 0 < x \le 1 \text{ and some } a > 0. \tag{14.16}$$

Because of a possible singularity at $x = 0$ we choose the setting (14.13b).

Exercise 14.4. Check condition (14.16) for $f(x) = 1/x,\ \ 1/x^2,\ \ x \log x,$ and x^x.

Functions f satisfying (14.16) are called *asymptotically smooth*. In fact, f is analytic in $(0,1]$. The Taylor series at $x_0 \in (0,1]$ has the convergence radius x_0. The remainder $\left|\sum_{k=N}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k\right|$ is bounded by

$$C\, x_0^{-a} \sum_{k=N}^{\infty} \left(1 - \frac{x}{x_0}\right)^k = C \frac{x_0^{1-a}}{x} \left(1 - \frac{x}{x_0}\right)^N \to 0.$$

Lemma 14.5. *Assume (14.16) and $\xi \in (0,1]$. Then there is a polynomial p of degree $N-1$ such that*

$$\|f - p\|_{[\xi/2,\xi],\infty} = \max_{\xi/2 \le x \le \xi} |f(x) - p(x)| \le \varepsilon_{N,\xi} := \frac{C}{2} \left(\frac{\xi}{4}\right)^{-a} 3^{1-a-N}.$$

Proof. Choose $x_0 = \frac{3}{4}\xi$ and set $p(x) := \sum_{k=0}^{N-1} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k$. The remainder in $[\xi/2, \xi]$ is bounded by $\varepsilon_{N,\xi}$ defined above. □

If an accuracy ε is prescribed, the number N_ε satisfying $\varepsilon_{N_\varepsilon,\xi} \le \varepsilon$ is asymptotically $N = (\log \frac{1}{\varepsilon} + a \log \frac{\xi}{4}) / \log \frac{1}{3}$.

Next, we define the following, piecewise polynomial function. Divide the interval $[\frac{1}{n}, 1]$ into the subintervals $[\frac{1}{n}, \frac{2}{n}] = [2^{-d}, 2^{1-d}]$, $(2^{1-d}, 2^{2-d}], \ldots, (\frac{1}{2}, 1]$. For each interval $(2^{-j}, 2^{1-j}]$ define a polynomial $p_j \in \mathcal{P}_{N-1}$ (cf. §10.4.2.1) according

[6] The factor in (14.15) can be integrated into the first factor. On the other hand, in the special case of (14.15) one need not store the number 1, so that only $r(d+1)$ data remain.

to Lemma 14.5. Altogether, we obtain a piecewise continuous function f_N (an hp-finite element approximation) defined on $[1/n, 1]$ satisfying the exponential decay

$$\|f - f_N\|_{[\frac{1}{n},1],\infty} \leq \varepsilon_N := 2^{-1+a(d+1)} 3^{1-a-N}.$$

Evaluation of f_N yields the vector entries $v_k = f\left((k+1)/n\right)$ and the tensorised version $\mathbf{v} \in \otimes^d \mathbb{K}^2$.

Proposition 14.6. *The tensor* $\mathbf{v}$ *constructed above possesses the* $\{1, \ldots, j\}$*-rank*

$$\rho_j = \operatorname{rank}_{\{1,\ldots,j\}} = \dim(\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v})) \leq N + 1.$$

Proof. For f_N define the functions $f_{N,j,k}$ from Remark 14.3. For $k \geq 1$, $f_{N,j,k}$ is a polynomial of degree $N-1$, only for $k = 0$, the function $f_{N,j,0}$ is *piecewise* polynomial. This proves $F_j \subset \mathcal{P}_{N-1} + \operatorname{span}\{f_{N,j,0}\}$ and $\dim(F_j) \leq N + 1$. The assertion follows from Remark 14.3. □

Since ρ_j is the (minimal) TT rank of $\mathbf{v}$ in the $\mathcal{H}_{\boldsymbol{\rho}}^{\mathrm{tens}}$ representation, the required storage is bounded by $2\,(d-1)\,(N+1)^2$.

A more general statement of a similar kind for functions with several singularities[7] is given by Grasedyck [74].

14.2.4 Multiscale Feature and Conclusion

Multiscale considerations of (grid) functions use different grid sizes $h_j = 2^j h$ and look for the behaviour in intervals of size h_j. A typical method exploiting these scales is the wavelet approach. Applying wavelet approximations to asymptotically smooth functions like in (14.16), one would need few wavelet levels on the right side of the interval, while the number of levels is increasing towards the singularity at $x = 0$. Again, one obtains estimates like in Proposition 14.6, showing that the advantages of the wavelet approach carry over to the hierarchical representation of $\mathbf{v} = \Phi_n^{-1}(v)$.

From (14.11) one sees that the subspace $U_{\{1,\ldots,j\}} = \operatorname{range}(\mathcal{M}_{\{1,\ldots,j\}}(\mathbf{v}))$ is connected to step size $2^j h$, i.e., to level $d - j$.

The approximation by exponentials, by hp-finite elements, or by wavelets helps to reduce the data size n of the uniformly discretised function v to a much smaller size, exploiting the regularity properties of the function. The tensorisation procedure has the same effect. The particular advantage is that the tensor approximation is a *black box* procedure using singular value decompositions, whereas the analytical methods mentioned above are chosen depending on the nature of the function and often require the computation of optimal coefficients (like in (14.14)).

[7] Then, in the right-hand side of (14.16), x is replaced by the distance of x to the next singularity.

14.2.5 Local Grid Refinement

Standard multiscale approaches apply local grid refinement in regions, where the approximation error is still too large. The tensorisation approach has fixed a step size $h = 1/n$. The data truncation discussed above can be related to grid *coarsening*. Nevertheless, also a grid refinement is possible. First, we discuss a prolongation from grid size h to $h/2$.

Remark 14.7 (prolongation). Consider a tensor $\mathbf{v} \in \mathbf{V} := \bigotimes_{j=1}^{d} \mathbb{K}^2$ corresponding to a vector $v \in \mathbb{K}^n$. We introduce a further vector space $V_0 := \mathbb{K}^2$ and define $\mathbf{v}^{\mathrm{ext}} \in \mathbf{V}^{\mathrm{ext}} := \bigotimes_{j=0}^{d} \mathbb{K}^2$ by

$$\mathbf{v}^{\mathrm{ext}} := \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] \otimes \mathbf{v}. \tag{14.17}$$

The prolongation $P : \mathbf{V} \to \mathbf{V}^{\mathrm{ext}}$ is defined by $\mathbf{v} \mapsto \mathbf{v}^{\mathrm{ext}}$ according to (14.17). $\mathbf{v}^{\mathrm{ext}} \in \mathbf{V}^{\mathrm{ext}}$ corresponds to $\Phi_{2n}^{-1}(\mathbf{v}^{\mathrm{ext}}) = v^{\mathrm{ext}} \in \mathbb{K}^{2n}$ via

$$\mathbf{v}^{\mathrm{ext}}[i_0, i_1, \ldots, i_d] = v^{\mathrm{ext}}\left[\sum_{i=0}^{d} i_j 2^j\right].$$

Furthermore, $v^{\mathrm{ext}}[i]$ represents the function value at grid point $i \cdot (h/2)$. The prolongation can be regarded as piecewise constant interpolation, since $v^{\mathrm{ext}}[2i] = v^{\mathrm{ext}}[2i+1]$.

The prolongation increases the data size by one tensor $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$. Note that the ranks ρ_j are not altered. Now, we redefine $\mathbf{V}^{\mathrm{ext}} = \bigotimes_{j=0}^{d} \mathbb{K}^2$ by $\mathbf{V}^{\mathrm{ext}} = \bigotimes_{j=1}^{d+1} \mathbb{K}^2$.

Let $\mathbf{v} \in \mathbf{V}^{\mathrm{ext}}$ be any tensor, e.g., in the image of the prolongation. Local refinement of the corresponding grid function in the subinterval

$$\left[\nu \cdot 2^{j^*} \cdot \frac{h}{2}, \left((\mu+1) \cdot 2^{j^*} - 1\right) \cdot \frac{h}{2}\right] \subset [0,1]$$

yields a tensor $\mathbf{v}'$, which satisfies the supposition of the next remark. If the refinement is really local, the level number j^* is not large.

Remark 14.8 (local change). Let $\mathbf{v}, \mathbf{v}' \in \otimes^{d+1}\mathbb{K}^2$ be two tensors such that $\Phi_{2n}^{-1}(\mathbf{v})$ and $\Phi_{2n}^{-1}(\mathbf{v}')$ differ only in the interval $\frac{h}{2}\left[\nu \cdot 2^{j^*}, (\mu+1) \cdot 2^{j^*} - 1\right]$ for some $1 \leq j^* \leq d$, $1 \leq \nu \leq \mu \leq 2^{d+1-j^*}$. Then the respective ranks ρ_j and ρ_j' satisfy

$$\begin{aligned} &\rho_j' \leq \rho_j + 1 && \text{for } j \geq j^*, \\ &\rho_j' \leq \min\{\rho_j + 2^{j^*-j}, 2^j\} && \text{for } 1 \leq j < j^*. \end{aligned}$$

Proof. For $j \geq j^*$, only one block in (14.12) is altered so that the rank increase is bounded by one. For $j < j^*$, at most 2^{j^*-j} blocks are involved so that $\rho_j' \leq \rho_j + 2^{j^*-j}$. On the other hand, $\rho_j' \leq 2^j$ holds for all tensors. □

14.3 Convolution

Operations like the Hadamard product can be applied to vectors $v = \Phi_n(\mathbf{v})$ as well as tensors $\mathbf{v}$, and identity (14.8) states that the results coincide. Similarly, the Euclidean scalar product $\langle v, w\rangle = \langle \Phi_n(\mathbf{v}), \Phi_n(\mathbf{w})\rangle$ of two vectors coincides with the scalar product $\langle \mathbf{v}, \mathbf{w}\rangle$ of the tensors. Such a property is not obvious for the convolution $v \star w$ of vectors which is discuss below (cf. Hackbusch [87]).

14.3.1 Notations

We consider vectors from $\mathbb{K}^n = \mathbb{K}^I$ with $I = \{0, 1, \ldots, n-1\}$. The convolution of $v, w \in \mathbb{K}^n$ is defined by[8]

$$u = v \star w \quad \text{with } u_k = \sum_{\ell=\max\{0,k+1-n\}}^{\min\{k,n-1\}} v_\ell w_{k-\ell} \quad (0 \le k \le 2n-2)\,. \tag{14.18a}$$

Note that the resulting vector u belongs to $\mathbb{K}^{2n-1}$, since for all $k \in \{0, \ldots, 2n-2\}$ the sum in (14.18) is non-empty.

An easier notation holds for vectors (infinite sequences) from $\ell_0 := \ell_0(\mathbb{N}_0)$ defined in (3.2):

$$\star : \ell_0 \times \ell_0 \to \ell_0, \quad u = v \star w \quad \text{with } u_k = \sum_{\ell=0}^{k} v_\ell w_{k-\ell} \quad \text{for all } k \in \mathbb{N} \tag{14.18b}$$

(check that the result belongs to ℓ_0, i.e., $u_k = 0$ for almost all $k \in \mathbb{N}$).

In the following, we embed $\mathbb{K}^n$ into ℓ_0 by identifying $v \in \mathbb{K}^n$ and $\hat{v} \in \ell_0$ with $\hat{v}_k = v_k$ for $0 \le k \le n-1$ and $\hat{v}_k = 0$ for $k \ge n$:

$$\mathbb{K}^n \subset \ell_0. \tag{14.19a}$$

A consequence of this identification is the embedding

$$\mathbb{K}^m \subset \mathbb{K}^n \qquad \text{for } 1 \le m \le n. \tag{14.19b}$$

Now we can rewrite (14.18a) as

$$u = v \star w \quad \text{with } u_k = \sum_{\ell=0}^{k} v_\ell w_{k-\ell} \qquad (0 \le k \le 2n-2)\,, \tag{14.18a'}$$

since the additional terms vanish.

[8] More generally, one may consider the convolution of $v \in \mathbb{K}^n$ and $w \in \mathbb{K}^m$ for different n, m. We avoid this trivial generalisation to reduce notational complications.

On ℓ_0 we define the degree

$$\deg(v) := \max\{k \in \mathbb{N} : v_k \neq 0\}.$$

Then $\mathbb{K}^n$ is identified with the subset $\{v \in \ell_0 : \deg(v) \leq n-1\}$.

Remark 14.9. Let $u = v \star w$ for $u, v, w \in \ell_0$. Then $\deg(u) = \deg(v) + \deg(w)$.

The reason for the name 'degree' becomes obvious from the following isomorphism.

Remark 14.10. Let $\mathcal{P}$ be the vector space of all polynomials (with coefficients in $\mathbb{K}$). Then $\mathcal{P}$ and ℓ_0 are isomorphic. The corresponding isomorphism is given by

$$\pi : \ell_0 \to \mathcal{P} \quad \text{with } \pi[v](x) := \sum_{k\in\mathbb{N}} v_k x^k. \tag{14.20}$$

The well-known connection of polynomials with the convolution is described by the property

$$u = v \star w \quad \text{for } u, v, w \in \ell_0 \qquad \text{if and only if} \qquad \pi[u] = \pi[v]\pi[w]. \tag{14.21}$$

Definition 14.11 (shift operator). For any $m \in \mathbb{Z}$, the shift operator $S^m : \ell_0 \to \ell_0$ is defined by

$$w = S^m(v) \text{ has entries } w_i = \left\{ \begin{matrix} v_{i-m} & \text{if } m \leq i \\ 0 & \text{otherwise} \end{matrix} \right\} \qquad \text{for } v \in \ell_0.$$

For $m \in \mathbb{N}_0$, S^m maps $(v_0, v_1, \ldots)$ into

$$(\underbrace{0, \ldots, 0}_{m \text{ positions}}, v_0, v_1, \ldots)$$

and has the left inverse S^{-m}, i.e., $S^{-m}S^m = id$.

The interaction of the shift operator and π is described by

$$\pi[S^m v](x) = x^m \cdot \pi[v](x) \qquad \text{for } m \in \mathbb{N}_0.$$

14.3.2 Preview and Motivation

Under the assumptions of §14.1.1, we rewrite the vectors $v, w \in \mathbb{K}^n$ as tensors $\mathbf{v}, \mathbf{w} \in \mathbf{V}$. To simplify the situation, assume that the tensors are elementary tensors: $\mathbf{v} = \bigotimes_{j=1}^d v^{(j)}$, $\mathbf{w} = \bigotimes_{j=1}^d w^{(j)}$ with $v^{(j)}, w^{(j)} \in \mathbb{K}^2$. We want to perform the composition of the following mappings:

$$\begin{aligned}(\mathbf{v},\mathbf{w})\in\mathbf{V}\times\mathbf{V}&\mapsto(v,w)\in\mathbb{K}^n\times\mathbb{K}^n &&\text{with } v=\Phi_n(\mathbf{v}),\, w=\Phi_n(\mathbf{w})\\ &\mapsto u:=v\star w\in\mathbb{K}^{2n} &&\text{(cf. (14.18a'))} \qquad (14.22)\\ &\mapsto \mathbf{u}:=\Phi_{2n}^{-1}(u)\in\otimes^{d+1}\mathbb{K}^2.\end{aligned}$$

We denote the mapping $(\mathbf{v},\mathbf{w})\mapsto\mathbf{u}$ from above shortly by

$$\mathbf{u}:=\mathbf{v}\star\mathbf{w}. \tag{14.23}$$

Note that the result $\mathbf{u}\in\mathbf{U}$ is a tensor of order $d+1$, since the corresponding vector $u\in\mathbb{K}^{2n-1}$ also belongs to $\mathbb{K}^{2n}$ (cf. (14.19b)) and $2n=2^{d+1}$.

In principle, the numerical realisation can follow definition (14.22). However, such an algorithm leads to at least $O(n)$ arithmetical operations, since $u:=v\star w$ is performed. Assuming that the data sizes of $\mathbf{v}$ and $\mathbf{w}$ are much smaller than $O(n)$, such an approach is too costly. Instead, the cost of $\mathbf{v}\star\mathbf{w}$ should be related to the data sizes of $\mathbf{v}$ and $\mathbf{w}$. An algorithm of this kind is obtained, if we perform the convolution *separately* in each direction:

$$\left(\bigotimes_{j=1}^{d}v^{(j)}\right)\star\left(\bigotimes_{j=1}^{d}w^{(j)}\right)=\bigotimes_{j=1}^{d}\left(v^{(j)}\star w^{(j)}\right), \tag{14.24}$$

provided that the right-hand side is a true description of $\mathbf{u}:=\mathbf{v}\star\mathbf{w}$. Unfortunately, this equation seems to be incorrect and even inconsistent, since the vector $v^{(j)}\star w^{(j)}$ belongs to $\mathbb{K}^3$ instead of $\mathbb{K}^2$.

Because of (14.1c), this difficulty is connected to a well-known problem arising for sums of integers in digital representation. When adding the decimal numbers 836 and 367, the place-wise addition of the digits yields

$$\begin{array}{c}8\;3\;6\\ 3\;6\;7\\ \hline (11)\;9\;(13)\end{array}$$

with the problem that 13 and 11 are no valid (decimal) digits. While this problem is usually solved by the carry-over, we can also allow a generalised decimal representation $(11)(9)(13)$ meaning $11\cdot10^2+9\cdot10^1+13\cdot10^0$, i.e., admitting all non-negative integers instead of the digits $\{0,\ldots,9\}$.

Such a 'generalised representation' for tensors will make use of the tensor space $\otimes^d\ell_0$ introduced in §14.3.3. Note that a vector from $\mathbb{K}^3$ appearing in the right-hand side of (14.24) is already considered as an element of ℓ_0. It will turn out that (14.24) has a correct interpretation in $\otimes^d\ell_0$. A corresponding 'carry-over' technique is needed to obtain a result in

$$\mathbf{U}=\otimes^{d+1}\mathbb{K}^2.$$

14.3.3 Tensor Algebra $\mathcal{A}(\ell_0)$

14.3.3.1 Definition and Interpretation in ℓ_0

The embedding $\mathbb{K}^2 \subset \ell_0$ leads to the embedding

$$\mathbf{V} = \otimes^d \mathbb{K}^2 = \bigotimes_{j=1}^{d} \mathbb{K}^2 \subset \otimes^d \ell_0 = \bigotimes_{j=1}^{d} \ell_0$$

(cf. Notation 3.23). The tensor algebra (cf. §3.4) is defined by

$$\mathfrak{A}(\ell_0) := \operatorname{span}\{\mathbf{a} \in \otimes^d \ell_0 : d \in \mathbb{N}\}.$$

Remark 14.12. A linear mapping $F : \mathfrak{A}(\ell_0) \to V$ into some vector space V is well-defined, if one of the following conditions holds:

(a) $F : \otimes^d \ell_0 \to V$ is defined as linear mapping for all $d \in \mathbb{N}$,

(b) the linear mapping F is defined for all elementary tensors $\bigotimes_{j=1}^d v^{(j)} \in \otimes^d \ell_0$ and for any $d \in \mathbb{N}$,

(c) the linear mapping F is defined for all elementary tensors $\bigotimes_{j=1}^d e^{(i_j)} \in \otimes^d \ell_0$ and for all $i_j \in \mathbb{N}_0$ and all $d \in \mathbb{N}$. Here, $e^{(\nu)} \in \ell_0$ is the unit vector with entries

$$e^{(\nu)}[k] = \delta_{k\nu} \qquad (k, \nu \in \mathbb{N}_0).$$

Proof. By definition, $\mathfrak{A}(\ell_0)$ is the *direct* sum of $\otimes^d \ell_0$, i.e., non-vanishing tensors of $\otimes^d \ell_0$ with different order d are linearly independent. This proves part (a). A linear mapping $F : \otimes^d \ell_0 \to V$ is well-defined by the images of the basis vectors $e^{(\mathbf{i})} := \bigotimes_{j=1}^d e^{(i_j)}$ for all $\mathbf{i} = (i_j)_{j=1,\ldots,d} \in \mathbb{N}_0^d$. □

In (14.1c), the isomorphism $\Phi_n : \mathbf{V} \to \mathbb{K}^n$ has been defined. We extend this mapping to $\mathfrak{A}(\ell_0)$ by

$$\begin{aligned} &\Phi : \mathfrak{A}(\ell_0) \to \ell_0\,, \\ &\mathbf{a} \in \otimes^d \ell_0 \mapsto v \in \ell_0 \\ &\quad \text{with } v_k = \sum_{i_1,\ldots,i_d \in \mathbb{N}_0 \text{ such that } k = \sum_{j=1}^d i_j 2^{j-1}} \mathbf{a}[i_1 i_2 \ldots i_d]. \end{aligned} \tag{14.25}$$

Remark 14.13. Since ℓ_0 is a subspace of $\mathfrak{A}(\ell_0)$ and $\Phi(v) = v$ holds for $v \in \ell_0$, the mapping Φ is a projection onto ℓ_0. Furthermore, the restriction of Φ to the tensor subspace $\mathbf{V} = \otimes^d \mathbb{K}^2 \subset \otimes^d \ell_0$ coincides with Φ_n from (14.1c) (hence, Φ is an extension of Φ_n).

14.3.3.2 Equivalence Relation and Polynomials

Definition 14.14. The equivalence relation $\sim$ on $\mathfrak{A}(\ell_0)$ is defined by

$$\mathbf{a} \sim \mathbf{b} \quad \Leftrightarrow \quad \Phi(\mathbf{a}) = \Phi(\mathbf{b}) \qquad (\mathbf{a}, \mathbf{b} \in \mathfrak{A}(\ell_0)).$$

Since Φ is a projection onto ℓ_0, we have in particular

$$\Phi(\mathbf{a}) \sim \mathbf{a} \qquad \text{for all } \mathbf{a} \in \mathfrak{A}(\ell_0). \tag{14.26}$$

The mapping $\pi : \ell_0 \to \mathcal{P}$ is defined in (14.20). We want to extend this mapping to $\mathfrak{A}(\ell_0) \supset \ell_0$ such that

$$\mathbf{a} \sim \mathbf{b} \quad \Leftrightarrow \quad \pi_{\mathfrak{A}}[\mathbf{a}] = \pi_{\mathfrak{A}}[\mathbf{b}]. \tag{14.27}$$

Definition 14.15. The extension[9] $\pi_{\mathfrak{A}} : \mathfrak{A}(\ell_0) \to \mathcal{P}$ of $\pi : \ell_0 \to \mathcal{P}$ from (14.20) is defined by

$$\pi_{\mathfrak{A}}\left[\bigotimes\nolimits_{j=1}^{d} a^{(j)}\right](x) := \prod\nolimits_{j=1}^{d} \pi[a^{(j)}]\left(x^{2^{j-1}}\right). \tag{14.28}$$

Lemma 14.16. *Mapping $\pi_{\mathfrak{A}}$ from Definition 14.15 is an extension of $\pi : \ell_0 \to \mathcal{P}$ and satisfies (14.27). Moreover,*

$$\pi[\Phi(\mathbf{a})] = \pi_{\mathfrak{A}}[\mathbf{a}]. \tag{14.29}$$

Proof. 1) $v \in \ell_0 = \otimes^1 \ell_0 \subset \mathfrak{A}(\ell_0)$ is an elementary tensor $\bigotimes_{j=1}^{1} v^{(j)}$ with the only factor $v^{(1)} = v$. Definition (14.28) yields $\pi_{\mathfrak{A}}[v](x) = \prod_{j=1}^{1} \pi[v^{(j)}](x^{2^{j-1}}) = \pi[v^{(1)}](x^1) = \pi[v](x)$, proving the extension property $\pi_{\mathfrak{A}}|_{\ell_0} = \pi$.

2) Let $\mathbf{e}^{(\mathbf{i})} := \bigotimes_{j=1}^{d} e^{(i_j)} \in \otimes^d \ell_0$ for some multi-index $\mathbf{i} = (i_j)_{j=1,\ldots,d} \in \mathbb{N}_0^d$. Since $\pi[e^{(i_j)}](x) = x^{i_j}$, definition (14.28) yields

$$\begin{aligned}\pi_{\mathfrak{A}}[\mathbf{e}^{(\mathbf{i})}](x) &= \prod\nolimits_{j=1}^{d} \pi[e^{(i_j)}](x^{2^{j-1}}) = \prod\nolimits_{j=1}^{d} (x^{2^{j-1}})^{i_j} = \prod\nolimits_{j=1}^{d} x^{i_j 2^{j-1}} \\ &= x^k \qquad \text{for } k := \sum\nolimits_{j=1}^{d} i_j 2^{j-1}. \end{aligned} \tag{14.30}$$

Definition (14.25) shows $\Phi(\mathbf{e}^{(\mathbf{i})}) = e^{(k)} \in \ell_0$ with k as above. Hence, $\pi[\Phi(\mathbf{e}^{(\mathbf{i})})] = x^k$ proves (14.29) for $\mathbf{a} = \mathbf{e}^{(\mathbf{i})}$. By Remark 14.12c, (14.29) follows for all $\mathbf{a} \in \mathfrak{A}(\ell_0)$.

3) The statement $\pi_{\mathfrak{A}}[\mathbf{a}] = \pi_{\mathfrak{A}}[\mathbf{b}] \Leftrightarrow \pi[\Phi(\mathbf{a})] = \pi[\Phi(\mathbf{b})]$ follows from (14.29). Since $\pi : \ell_0 \to \mathcal{P}$ is an isomorphism, also $\pi[\Phi(\mathbf{a})] = \pi[\Phi(\mathbf{b})] \Leftrightarrow \Phi(\mathbf{a}) = \Phi(\mathbf{b})$ holds. The latter equality is the definition of $\mathbf{a} \sim \mathbf{b}$. Hence, (14.27) is proved. □

Remark 14.17. $\mathbf{a} \otimes e^{(0)} \sim \mathbf{a}$ and $\pi_{\mathfrak{A}}[\mathbf{a} \otimes e^{(0)}] = \pi_{\mathfrak{A}}[\mathbf{a}]$ hold for all $\mathbf{a} \in \mathfrak{A}(\ell_0)$.

Proof. By Remark 14.12c, it suffices to consider the tensor $\mathbf{a} = \bigotimes_{j=1}^{d} e^{(i_j)}$. Then $\mathbf{b} := \mathbf{a} \otimes e^{(0)}$ equals $\bigotimes_{j=1}^{d+1} e^{(i_j)}$ with $i_{d+1} := 0$. By (14.28), $\pi_{\mathfrak{A}}[\mathbf{a}](x) = x^k$ holds with k as in (14.30), while $\pi_{\mathfrak{A}}[\mathbf{b}](x) = x^k \cdot \pi[e^{(0)}](x) = x^k \cdot 1 = x^k$. Now, (14.27) proves the assertion. □

[9] It may be more natural to define the mapping $\hat{\pi}_{\mathfrak{A}}$ into polynomials of all variables x_j $(j \in \mathbb{N})$ by $\hat{\pi}_{\mathfrak{A}}[\bigotimes_{j=1}^{d} v^{(j)}](\mathbf{x}) := \prod_{j=1}^{d} \pi[v^{(j)}]\,(x_j)$. Then the present value $\pi_{\mathfrak{A}}[\bigotimes_{j=1}^{d} v^{(j)}](x)$ results from the substitutions $x_j := x^{2^{j-1}}$.

14.3.3.3 Shift

Next, we extend the shift operator $S^m : \ell_0 \to \ell_0$ to $S^m_{\mathfrak{A}} : \mathfrak{A}(\ell_0) \to \mathfrak{A}(\ell_0)$ by means of [10]

$$S^m_{\mathfrak{A}}\left(\bigotimes_{j=1}^d v^{(j)}\right) := \left(S^m v^{(1)}\right) \otimes \bigotimes_{j=2}^d v^{(j)}. \tag{14.31}$$

Remark 14.18. $\Phi(S^m_{\mathfrak{A}}(\mathbf{a})) = S^m(\Phi(\mathbf{a}))$ holds for all $\mathbf{a} \in \mathfrak{A}(\ell_0)$. $S^m_{\mathfrak{A}}$ is an extension of S^m, since $S^m_{\mathfrak{A}}|_{\ell_0} = S^m$.

Proof. According to Remark 14.12c, we choose a tensor $\mathbf{a} = \mathbf{e}^{(\mathbf{i})} = \bigotimes_{j=1}^d e^{(i_j)} \in \otimes^d \ell_0$ with $\mathbf{i} = (i_j)_{j=1,\ldots,d} \in \mathbb{N}_0^d$. Since $\Phi(\mathbf{a}) = e^{(k)}$ holds with k defined in (14.30), the shift yields the result $S^m(\Phi(\mathbf{a})) = e^{(k+m)}$. On the other hand, (14.31) shows that $S^m_{\mathfrak{A}}(\mathbf{a}) = e^{(i_1+m)} \otimes \bigotimes_{j=2}^d e^{(i_j)}$ and $\Phi(S^m_{\mathfrak{A}}(\mathbf{a})) = e^{(k+m)}$, proving the first assertion $\Phi(S^m_{\mathfrak{A}}(\mathbf{a})) = S^m(\Phi(\mathbf{a}))$. The second one is trivial. □

On the right-hand side of (14.31), the shift operator S^m is applied to $v^{(1)}$ only. Next, we consider shifts of all $v^{(j)}$.

Lemma 14.19. *Let* $\mathbf{m} = (m_1, \ldots, m_d) \in \mathbb{N}_0^d$. *The operator* $\mathbf{S}^{(\mathbf{m})} := \bigotimes_{j=1}^d S^{m_j}$ *applied to* $\mathbf{v} \in \otimes^d \ell_0$ *yields*

$$\left.\begin{aligned} \pi_{\mathfrak{A}}[\mathbf{S}^{(\mathbf{m})}\mathbf{v}](x) &= x^m \pi_{\mathfrak{A}}[\mathbf{v}](x) \\ \mathbf{S}^{(\mathbf{m})}\mathbf{v} &\sim S^m(\Phi(\mathbf{v})) \end{aligned}\right\} \quad \text{with } m = \sum_{j=1}^d m_j 2^{j-1}. \tag{14.32}$$

Proof. 1) By Remark 14.12c, we may consider $\mathbf{v} = \mathbf{e}^{(\mathbf{i})} = \bigotimes_{j=1}^d e^{(i_j)}$. Set $i := \sum_{j=1}^d i_j 2^{j-1}$. Then, $\mathbf{S}^{(\mathbf{m})}\mathbf{e}^{(\mathbf{i})} = \bigotimes_{j=1}^d e^{(i_j+m_j)}$ yields

$$\pi_{\mathfrak{A}}[\mathbf{S}^{(\mathbf{m})}\mathbf{v}](x) = \pi_{\mathfrak{A}}\left[\bigotimes_{j=1}^d e^{(i_j+m_j)}\right](x) = \prod_{j=1}^d x^{(i_j+m_j)2^{j-1}} = x^{i+m},$$

which coincides with $x^m \pi_{\mathfrak{A}}[\mathbf{v}](x) = x^m \pi_{\mathfrak{A}}[\bigotimes_{j=1}^d e^{(i_j)}](x) = x^m x^i$. This proves the first part of (14.32).

2) $S^m(\Phi(\mathbf{v})) = \Phi(S^m_{\mathfrak{A}}(\mathbf{v}))$ holds by Remark 14.18. The definition of $S^m_{\mathfrak{A}}$ can be rewritten as $\mathbf{S}^{(\hat{\mathbf{m}})}$ with the multi-index $\hat{\mathbf{m}} = (m, 0, \ldots, 0)$. Statement (14.26) shows $\Phi(\mathbf{S}^{(\hat{\mathbf{m}})}(\mathbf{v})) \sim \mathbf{S}^{(\hat{\mathbf{m}})}(\mathbf{v})$, hence $S^m(\Phi(\mathbf{v})) \sim \mathbf{S}^{(\hat{\mathbf{m}})}(\mathbf{v})$. Since $\mathbf{S}^{(\hat{\mathbf{m}})}(\mathbf{v})$ and $\mathbf{S}^{(\mathbf{m})}\mathbf{v}$ have the identical image $x^m \pi_{\mathfrak{A}}[\mathbf{v}](x)$ under the mapping $\pi_{\mathfrak{A}}$, property (14.27) implies the second statement in (14.32). □

Corollary 14.20. Let $\mathbf{v} \in \otimes^d \ell_0$ and $\mathbf{m}, \mathbf{m}' \in \mathbb{N}_0^d$. Then

$$\sum_{j=1}^d m_j 2^{j-1} = \sum_{j=1}^d m'_j 2^{j-1} \quad \text{implies} \quad \mathbf{S}^{(\mathbf{m})}\mathbf{v} \sim \mathbf{S}^{(\mathbf{m}')}\mathbf{v}.$$

[10] Here, we make use of Remark 14.12b and restrict the definition to elementary tensors.

14.3.3.4 Multi-Scale Interpretation

The representation of a vector from ℓ_0 by means of $\Phi(\mathbf{a})$ with $\mathbf{a} \in \mathfrak{A}(\ell_0)$ has similarities to the multi-scale analysis of functions using a wavelet basis. A vector $v \in \ell_0$ is often viewed as the vector of grid values $v_k = f(k)$ of a (smooth) function f defined on $[0, \infty)$. Let $\mathbf{a} = \bigotimes_{\nu=1}^{d} a^{(\nu)} \in \otimes^d \ell_0$ and $j \in \{1, \ldots, d\}$. A shift in position j is described by

$$\mathbf{a} \mapsto \hat{\mathbf{a}} := a^{(1)} \otimes \ldots \otimes a^{(j-1)} \otimes (Sa^{(j)}) \otimes a^{(j+1)} \otimes \ldots \otimes a^{(d)}$$

and corresponds to $v = \Phi(\mathbf{a}) \mapsto \hat{v} := \Phi(\hat{\mathbf{a}})$ with $\hat{v} = S^{2^{j-1}} v$ (cf. (14.32)). The interpretation of v by $v_k = f(k)$ leads to $\hat{v}_\mu = \hat{f}(\mu)$ with the shifted function $\hat{f}(x) = f(x + 2^{j-1})$. On the other hand, a multi-scale basis at level $\ell = j - 1$ is given by $\{\psi_\nu\}$ with the shift property $\psi_\mu(x) = \psi_\nu(x + (\nu - \mu)\, 2^\ell)$. Hence, the shift

$$f = \sum_\nu c_\nu \psi_\nu \mapsto \hat{f} = \sum_\nu (Sc)_\nu\, \psi_\nu = \sum_\nu c_{\nu-1} \psi_\nu = \sum_\nu c_\nu \psi_{\nu+1}$$

also results in $\hat{f}(x) = f(x + 2^\ell)$.

14.3.3.5 Convolution

Finally, we define a convolution operation in $\mathfrak{A}(\ell_0)$. The following $\star$ operation will be different (but equivalent) to the $\star$ operation in (14.23). The former operation acts in $\mathfrak{A}(\ell_0) \times \mathfrak{A}(\ell_0)$ and yields results in $\mathfrak{A}(\ell_0)$, whereas the latter one maps $(\otimes^d \mathbb{K}^2) \times (\otimes^d \mathbb{K}^2)$ into $\otimes^{d+1} \mathbb{K}^2$.

For elementary tensors $\mathbf{a} = \bigotimes_{j=1}^{d} a^{(j)}$ and $\mathbf{b} = \bigotimes_{j=1}^{d} b^{(j)}$ from $\otimes^d \ell_0$ the obvious definition is

$$\left(\bigotimes_{j=1}^{d} a^{(j)} \right) \star \left(\bigotimes_{j=1}^{d} b^{(j)} \right) = \bigotimes_{j=1}^{d} \left(a^{(j)} \star b^{(j)} \right) \tag{14.33}$$

(cf. (14.24) and (4.74)). Since $\mathfrak{A}(\ell_0)$ contains tensors of different orders, we define more generally

$$\left(\bigotimes_{j=1}^{d_a} a^{(j)} \right) \star \left(\bigotimes_{j=1}^{d_b} b^{(j)} \right) = \bigotimes_{j=1}^{d_c} c^{(j)} \quad \text{with } d_c := \max\{d_a, d_b\}$$

$$\text{and} \quad \begin{cases} c^{(j)} := a^{(j)} \star b^{(j)} & \text{for } j \le \min\{d_a, d_b\}, \\ c^{(j)} := \begin{cases} a^{(j)} & \text{for } d_b < j \le d_c \text{ if } d_a = d_c, \\ b^{(j)} & \text{for } d_a < j \le d_c \text{ if } d_b = d_c. \end{cases} \end{cases} \tag{14.34}$$

Note that (14.34) coincides with (14.33) for $d_a = d_b$.

Corollary 14.21. Another interpretation of (14.34) follows. Assume $\mathbf{a} \in \otimes^{d_a}\ell_0$ and $\mathbf{b} \in \otimes^{d_b}\ell_0$ with $d_a < d_b$. Replace $\mathbf{a}$ by $\hat{\mathbf{a}} := \mathbf{a} \otimes \bigotimes_{j=d_a+1}^{d_c} e^{(0)} \in \otimes^{d_c}\ell_0$ and set $\mathbf{a} \star \mathbf{b} := \hat{\mathbf{a}} \star \mathbf{b}$, where the latter expression can be defined by (14.33) with $d = d_b$. As

$$e^{(0)} \star v = v \star e^{(0)} = v \quad \text{for all } v \in \ell_0,$$

the new definition of $\mathbf{a} \star \mathbf{b}$ coincides with (14.34).

Property (14.21) has a counterpart for the convolution in $\mathfrak{A}(\ell_0)$, which will be very helpful in §14.3.4.

Proposition 14.22. *(a)* $\Phi(\mathbf{a} \star \mathbf{b}) = \Phi(\mathbf{a}) \star \Phi(\mathbf{b})$ *holds for all* $\mathbf{a}, \mathbf{b} \in \mathfrak{A}(\ell_0)$, *where the second* $\star$ *operation is the convolution (14.18b) in* ℓ_0.
(b) The implication

$$\mathbf{c} \sim \mathbf{a} \star \mathbf{b} \quad \Leftrightarrow \quad \Phi(\mathbf{c}) = \Phi(\mathbf{a}) \star \Phi(\mathbf{b})$$

holds for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{A}(\ell_0)$.

Proof. We apply (14.21), which holds for the ℓ_0-convolution:

$$\Phi(\mathbf{c}) = \Phi(\mathbf{a}) \star \Phi(\mathbf{b}) \quad \Leftrightarrow \quad \pi[\Phi(\mathbf{c})] = \pi[\Phi(\mathbf{a})]\, \pi[\Phi(\mathbf{b})].$$

By (14.29), this is equivalent to $\pi_{\mathfrak{A}}[\mathbf{c}] = \pi_{\mathfrak{A}}[\mathbf{a}]\, \pi_{\mathfrak{A}}[\mathbf{b}]$. It suffices to consider elementary tensors $\mathbf{a} = \bigotimes_{j=1}^{d_a} a^{(j)}$ and $\mathbf{b} = \bigotimes_{j=1}^{d_b} b^{(j)}$ (extend Remark 14.12 to the bilinear mapping $\star$). First we assume $d_a = d_b =: d$. Then definition (14.33) yields

$$\begin{aligned}
\pi_{\mathfrak{A}}[\mathbf{a} \star \mathbf{b}](x) &= \pi_{\mathfrak{A}}\left[\bigotimes_{j=1}^{d}(a^{(j)} \star b^{(j)})\right](x) \underset{(14.28)}{=} \prod_{j=1}^{d} \pi[a^{(j)} \star b^{(j)}](x^{2^{j-1}}) \underset{(14.21)}{=} \\
&= \prod_{j=1}^{d} \left\{\pi[a^{(j)}](x^{2^{j-1}}) \cdot \pi[b^{(j)}](x^{2^{j-1}})\right\} \\
&= \left[\prod_{j=1}^{d} \pi[a^{(j)}](x^{2^{j-1}})\right] \cdot \left[\prod_{j=1}^{d} \pi[b^{(j)}](x^{2^{j-1}})\right] \underset{(14.28)}{=} \\
&= \pi_{\mathfrak{A}}[\mathbf{a}](x) \cdot \pi_{\mathfrak{A}}[\mathbf{b}](x).
\end{aligned}$$

This proves $\Phi(\mathbf{a} \star \mathbf{b}) = \Phi(\mathbf{a}) \star \Phi(\mathbf{b})$ for $\mathbf{a}, \mathbf{b} \in \otimes^d\ell_0$. For elementary tensors of different orders $d_v \neq d_w$ use the equivalent definition from Corollary 14.21. Since $\pi_{\mathfrak{A}}[\hat{\mathbf{a}}] = \pi_{\mathfrak{A}}[\mathbf{a}]$ (cf. Remark 14.17), assertion (a) follows from the previous result.

Because $\mathbf{c} \sim \mathbf{a} \star \mathbf{b}$ is equivalent to $\Phi(\mathbf{c}) = \Phi(\mathbf{a} \star \mathbf{b})$, Part (b) follows from (a). □

Exercise 14.23. Let $\mathbf{a}, \mathbf{a}' \in \otimes^d\ell_0$ and $\mathbf{b}, \mathbf{b}' \in \mathfrak{A}(\ell_0)$ with $\mathbf{a} \sim \mathbf{a}'$ and $\mathbf{b} \sim \mathbf{b}'$. Prove $\mathbf{a} \otimes \mathbf{b} \sim \mathbf{a}' \otimes \mathbf{b}'$.

14.3.3.6 Carry-over Procedure

In §14.3.2 we have pointed to the analogue of the 'carry-over' for sums of integers. In the present case, the 'carry-over' must change an element $\mathbf{a} \in \mathfrak{A}(\ell_0)$ into $\mathbf{a}' \in \mathfrak{A}(\ell_0)$ such that $\mathbf{a} \sim \mathbf{a}'$ and $\mathbf{a}' \in \otimes^d \mathbb{K}^2$ for a minimal d. Equivalence $\mathbf{a} \sim \mathbf{a}'$ ensures that both $\mathbf{a}$ and $\mathbf{a}'$ are (generalised) tensorisations of the same vector $\Phi(\mathbf{a}) = \Phi(\mathbf{a}') \in \ell_0$. The minimal d is determined by the inequality $2^{d-1} < \deg(\Phi(\mathbf{a})) \leq 2^d$. The following algorithm proceeds from Step 0 to Step $d-1$.

Step 0) Any element $\mathbf{a}$ is a finite sum $\sum_\nu \mathbf{a}_\nu$ of elementary tensors $\mathbf{a}_\nu \in \otimes^{d_\nu} \ell_0$.

Case a) If $d_\nu > d$, we may truncate to d as follows. Let $\mathbf{a}_\nu = \mathbf{a}'_\nu \otimes \mathbf{a}''_\nu$ with $\mathbf{a}'_\nu \in \otimes^d \ell_0$ and $\mathbf{a}''_\nu \in \otimes^{d_\nu - d} \ell_0$. Replace $\mathbf{a}_\nu$ by $\tilde{\mathbf{a}}_\nu := \lambda \mathbf{a}'_\nu \in \otimes^d \ell_0$, where $\lambda := (\mathbf{a}''_\nu)[0 \ldots 0]$ is the entry for $(i_1, \ldots, i_{d_\nu - d}) = (0, \ldots, 0)$. Then $\Phi(\mathbf{a}_\nu)$ and $\Phi(\tilde{\mathbf{a}}_\nu)$ have identical entries for the indices $0 \leq i \leq 2^d - 1$. The other may differ, but in the sum $\sum_\nu \mathbf{a}_\nu$ they must vanish, since $\deg(\Phi(\mathbf{a})) \leq 2^d$. Hence, $\sum_\nu \mathbf{a}_\nu \sim \sum_\nu \tilde{\mathbf{a}}_\nu$ are equivalent representations.

Case b) If $d_\nu < d$, replace $\mathbf{a}_\nu$ by $\tilde{\mathbf{a}}_\nu := \mathbf{a}_\nu \otimes \bigotimes_{j=d_\nu+1}^{d} e^{(0)}$. Remark 14.17 ensures $\mathbf{a}_\nu \sim \tilde{\mathbf{a}}_\nu$.

After these changes, the new $\tilde{\mathbf{a}} \sim \mathbf{a}$ belongs to $\mathbf{V} := \otimes^d \ell_0$.

Step 1) $\mathbf{a} \in \otimes^d \ell_0$ has the representation $\sum_\nu a_\nu^{(1)} \otimes a_\nu^{(>1)}$ with components $a_\nu^{(1)} \in \ell_0$ and $a_\nu^{(>1)} \in \bigotimes_{j=2}^{d} \ell_0$. In case of $\deg(a_\nu^{(1)}) > 2$, split $a := a_\nu^{(1)}$ into $a_\nu^{\prime(1)} + S^2 a_\nu^{\prime\prime(1)}$ with $a_\nu^{\prime(1)} \in \mathbb{K}^2$. For this purpose, set

$$a_\nu^{\prime(1)} := (a_0, a_1) \quad \text{and} \quad a_\nu^{\prime\prime(1)} := (a_3, a_4, \ldots) = S^{-2} a_\nu^{(1)} \in \ell_0.$$

Then (14.32) implies

$$a_\nu^{(1)} \otimes a_\nu^{(>1)} = a_\nu^{\prime(1)} \otimes a_\nu^{(>1)} + \left(S^2 a_\nu^{\prime\prime(1)}\right) \otimes a_\nu^{(>1)} \sim a_\nu^{\prime(1)} \otimes a_\nu^{(>1)} + a_\nu^{\prime\prime(1)} \otimes \left(S^1 a_\nu^{(>1)}\right).$$

Note that $\deg(a_\nu^{\prime\prime(1)}) = \deg(a_\nu^{(1)}) - 2$ has decreased. As long as $\deg(a_\nu^{\prime\prime(1)}) > 2$, this procedure is to be repeated.

At the end of Step 1, a new tensor $\tilde{\mathbf{a}} = \sum_\nu a_\nu^{(1)} \otimes a_\nu^{(>1)} \sim \mathbf{a}$ with $a_\nu^{(1)} \in \mathbb{K}^2$ is obtained.

Step 2) If $d > 2$, we apply the procedure of Step 1 to $a_\nu^{(>1)}$ in $\tilde{\mathbf{a}} = \sum_\nu a_\nu^{(1)} \otimes a_\nu^{(>1)}$ with d replaced by d–1. Each $a_\nu^{(>1)}$ is replaced by $\tilde{a}_\nu^{(>1)} = \sum_\mu a_{\nu\mu}^{(2)} \otimes a_{\nu\mu}^{(>2)} \sim a_\nu^{(>1)}$. By Exercise 14.23, $\tilde{\mathbf{a}} \sim \hat{\mathbf{a}} := \sum_{\nu,\mu} a_\nu^{(1)} \otimes a_{\nu\mu}^{(2)} \otimes a_{\nu\mu}^{(>2)}$ holds with $a_\nu^{(1)}, a_{\nu\mu}^{(2)} \in \mathbb{K}^2$. Take $\hat{\mathbf{a}}$ as new $\tilde{\mathbf{a}}$. Reorganisation of the sum $\sum\limits_{\nu,\mu}$ yields $\tilde{\mathbf{a}} = \sum_\nu a_\nu^{(1)} \otimes a_\nu^{(2)} \otimes a_\nu^{(>2)} \sim \mathbf{a}$ with $a_\nu^{(1)}, a_\nu^{(2)} \in \mathbb{K}^2$.

$\vdots$

Step $d-1$**)** The previous procedure yields $\tilde{\mathbf{a}} = \sum_\nu \bigotimes_{j=1}^{d} a_\nu^{(j)}$ with $a_\nu^{(j)} \in \mathbb{K}^2$ for $1 \leq j \leq d-1$. If $\deg(a_\nu^{(d)}) > 2$, split $a_\nu^{(d)}$ into $a_\nu^{\prime(d)} + S^2 a_\nu^{\prime\prime(d)}$ with $a_\nu^{\prime(d)} \in \mathbb{K}^2$. As in Case a) of Step 0), we conclude that $\sum_\nu \left(\bigotimes_{j=1}^{d-1} a_\nu^{(j)}\right) \otimes \left(S^2 a_\nu^{\prime\prime(d)}\right) = 0$. Hence, $\tilde{\mathbf{a}} = \sum_\nu \left(\bigotimes_{j=1}^{d-1} a_\nu^{(j)}\right) \otimes a_\nu^{\prime(d)} \sim \mathbf{a}$ is the desired representation in $\mathbf{V} := \otimes^d \mathbb{K}^2$.

14.3.4 Algorithm

14.3.4.1 Main Identities

A tensor $\mathbf{v} \in \bigotimes^d \mathbb{K}^2$ ($d \in \mathbb{N}$) possesses a unique decomposition[11]

$$\mathbf{v} = \mathbf{v}' \otimes \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] + \mathbf{v}'' \otimes \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] \quad \text{with } \mathbf{v}', \mathbf{v}'' \in \otimes^{d-1}\mathbb{K}^2.$$

The linear mappings $\mathbf{v} \mapsto \mathbf{v}'$ and $\mathbf{v} \mapsto \mathbf{v}''$ from $\otimes^d\mathbb{K}^2$ onto $\otimes^{d-1}\mathbb{K}^2$ are denoted by ϕ'_d and ϕ''_d, respectively. Their precise definition is

$$\phi'_d\left(\mathbf{v} \otimes \left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right]\right) = \alpha\mathbf{v}, \quad \phi''_d\left(\mathbf{v} \otimes \left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right]\right) = \beta\mathbf{v} \qquad \text{for } \mathbf{v} \in \otimes^{d-1}\mathbb{K}^2.$$

We start with the simple case of $d = 1$. The next lemma demonstrates how the 'carry-over' is realised.

Lemma 14.24. *The convolution of* $\left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right], \left[\begin{smallmatrix}\gamma\\\delta\end{smallmatrix}\right] \in \mathbb{K}^2 = \otimes^1\mathbb{K}^2$ *yields*

$$\left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right] \star \left[\begin{smallmatrix}\gamma\\\delta\end{smallmatrix}\right] = \varPhi(\mathbf{v}) \quad \textit{with } \mathbf{v} := \left[\begin{smallmatrix}\alpha\gamma\\\alpha\delta+\beta\gamma\end{smallmatrix}\right] \otimes \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] + \left[\begin{smallmatrix}\beta\delta\\0\end{smallmatrix}\right] \otimes \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] \in \otimes^2\mathbb{K}^2. \tag{14.35a}$$

Furthermore, the shifted vector $S^1\left(\left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right] \star \left[\begin{smallmatrix}\gamma\\\delta\end{smallmatrix}\right]\right)$ *has the tensor representation*

$$S^1\left(\left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right] \star \left[\begin{smallmatrix}\gamma\\\delta\end{smallmatrix}\right]\right) = \varPhi(\mathbf{v}) \textit{ with } \mathbf{v} := \left[\begin{smallmatrix}0\\\alpha\gamma\end{smallmatrix}\right] \otimes \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] + \left[\begin{smallmatrix}\alpha\delta+\beta\gamma\\\beta\delta\end{smallmatrix}\right] \otimes \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] \in \otimes^2\mathbb{K}^2. \tag{14.35b}$$

Proof. An elementary calculation yields $\left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right] \star \left[\begin{smallmatrix}\gamma\\\delta\end{smallmatrix}\right] = \begin{bmatrix}\alpha\gamma\\\alpha\delta+\beta\gamma\\\beta\delta\end{bmatrix} \in \mathbb{K}^3$, where the latter vector is identified with $(\alpha\gamma, \alpha\delta + \beta\gamma, \beta\delta, 0, 0, \ldots) \in \ell_0$. We split this vector into $\left[\begin{smallmatrix}\alpha\gamma\\\alpha\delta+\beta\gamma\end{smallmatrix}\right] + S^2\left[\begin{smallmatrix}\beta\delta\\0\end{smallmatrix}\right]$. From $\left[\begin{smallmatrix}\alpha\gamma\\\alpha\delta+\beta\gamma\end{smallmatrix}\right] \sim \left[\begin{smallmatrix}\alpha\gamma\\\alpha\delta+\beta\gamma\end{smallmatrix}\right] \otimes \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ (cf. Remark 14.17) and

$$S^2\left[\begin{smallmatrix}\beta\delta\\0\end{smallmatrix}\right] \sim \left(S^2\left[\begin{smallmatrix}\beta\delta\\0\end{smallmatrix}\right]\right) \otimes \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] \sim \left[\begin{smallmatrix}\beta\delta\\0\end{smallmatrix}\right] \otimes \left(S^1\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\right) = \left[\begin{smallmatrix}\beta\delta\\0\end{smallmatrix}\right] \otimes \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$$

we obtain the first assertion. The second one follows analogously from

$$S^1\left(\left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right] \star \left[\begin{smallmatrix}\gamma\\\delta\end{smallmatrix}\right]\right) = \begin{bmatrix}0\\\alpha\gamma\\\alpha\delta+\beta\gamma\\\beta\delta\end{bmatrix} \sim \left[\begin{smallmatrix}0\\\alpha\gamma\end{smallmatrix}\right] + S^2\left[\begin{smallmatrix}\alpha\delta+\beta\gamma\\\beta\delta\end{smallmatrix}\right]$$

finishing the proof. □

The basic identity is given in the next lemma, which shows how the convolution product of tensors of order $d-1$ can be used for tensors of order d. Note that $\mathbf{a}', \mathbf{a}''$ and $\mathbf{u}'$, $\mathbf{u}''$ can be expressed by $\phi'_d(\mathbf{a})$, $\phi''_d(\mathbf{a})$ and $\phi'_d(\mathbf{u})$, $\phi''_d(\mathbf{u})$, respectively.

[11] For $d = 1$, the tensors $\mathbf{v}', \mathbf{v}''$ degenerate to numbers from the field $\mathbb{K}$.

Lemma 14.25. *Let $d \geq 2$. Assume that for $\mathbf{v}, \mathbf{w} \in \otimes^{d-1}\mathbb{K}^2$ the equivalence*

$$\mathbf{v} \star \mathbf{w} \sim \mathbf{a} = \mathbf{a}' \otimes \begin{bsmallmatrix}1\\0\end{bsmallmatrix} + \mathbf{a}'' \otimes \begin{bsmallmatrix}0\\1\end{bsmallmatrix} \in \otimes^d\mathbb{K}^2 \tag{14.36a}$$

holds. Let the tensors $\mathbf{v} \otimes x, \mathbf{w} \otimes y \in \otimes^d\mathbb{K}^2$ be defined by $x = \begin{bsmallmatrix}\alpha\\\beta\end{bsmallmatrix}$, $y = \begin{bsmallmatrix}\gamma\\\delta\end{bsmallmatrix} \in \mathbb{K}^2$. Then,

$$\begin{aligned} (\mathbf{v} \otimes x) \star (\mathbf{w} \otimes y) &\sim \mathbf{u} = \mathbf{u}' \otimes \begin{bsmallmatrix}1\\0\end{bsmallmatrix} \; + \; \mathbf{u}'' \otimes \begin{bsmallmatrix}0\\1\end{bsmallmatrix} \in \otimes^{d+1}\mathbb{K}^2 \\ \textit{with } \mathbf{u}' &= \mathbf{a}' \otimes \begin{bsmallmatrix}\alpha\gamma\\\alpha\delta+\beta\gamma\end{bsmallmatrix} + \mathbf{a}'' \otimes \begin{bsmallmatrix}0\\\alpha\gamma\end{bsmallmatrix} \in \otimes^d\mathbb{K}^2 \\ \textit{and } \mathbf{u}'' &= \mathbf{a}' \otimes \begin{bsmallmatrix}\beta\delta\\0\end{bsmallmatrix} + \mathbf{a}'' \otimes \begin{bsmallmatrix}\alpha\delta+\beta\gamma\\\beta\delta\end{bsmallmatrix} \in \otimes^d\mathbb{K}^2. \end{aligned} \tag{14.36b}$$

Proof. Proposition 14.22 ensures that

$$(\mathbf{v} \otimes x) \star (\mathbf{w} \otimes y) \sim (\mathbf{v} \star \mathbf{w}) \otimes z \qquad \text{with } z := x \star y \in \mathbb{K}^3 \subset \ell_0.$$

Assumption (14.36a) together with $\mathbf{a}' \otimes \begin{bsmallmatrix}1\\0\end{bsmallmatrix} \sim \mathbf{a}'$ (cf. Remark 14.17) and

$$\mathbf{a}'' \otimes \begin{bsmallmatrix}0\\1\end{bsmallmatrix} = \mathbf{a}'' \otimes \left(S^1 \begin{bsmallmatrix}1\\0\end{bsmallmatrix}\right) \sim S_{\mathfrak{A}}^{2^{d-1}} \mathbf{a}''$$

(cf. Remark 14.18) yields

$$(\mathbf{v} \star \mathbf{w}) \otimes z \sim \left(\mathbf{a}' + S_{\mathfrak{A}}^{2^{d-1}} \mathbf{a}''\right) \otimes z.$$

Again, Remark 14.18 shows that

$$(S_{\mathfrak{A}}^{2^{d-1}} \mathbf{a}'') \otimes z = S_{\mathfrak{A}}^{2^{d-1}} (\mathbf{a}'' \otimes z) \sim \mathbf{a}'' \otimes (Sz).$$

Using (14.35a,b), we obtain

$$\begin{gathered} \mathbf{a}' \otimes z \sim \mathbf{a}' \otimes \begin{bsmallmatrix}\alpha\gamma\\\alpha\delta+\beta\gamma\end{bsmallmatrix} \otimes \begin{bsmallmatrix}1\\0\end{bsmallmatrix} + \mathbf{a}' \otimes \begin{bsmallmatrix}\beta\delta\\0\end{bsmallmatrix} \otimes \begin{bsmallmatrix}0\\1\end{bsmallmatrix}, \\ \left(S^{2^{\delta-1}} \mathbf{a}''\right) \otimes z \sim \mathbf{a}'' \otimes (Sz) \sim \mathbf{a}'' \otimes \begin{bsmallmatrix}0\\\alpha\gamma\end{bsmallmatrix} \otimes \begin{bsmallmatrix}1\\0\end{bsmallmatrix} + \mathbf{a}'' \otimes \begin{bsmallmatrix}\alpha\delta+\beta\gamma\\\beta\delta\end{bsmallmatrix} \otimes \begin{bsmallmatrix}0\\1\end{bsmallmatrix}. \end{gathered}$$

Summation of both identities yields the assertion of the lemma. □

If $x = \begin{bsmallmatrix}\alpha\\\beta\end{bsmallmatrix}$ and $y = \begin{bsmallmatrix}\gamma\\\delta\end{bsmallmatrix}$ are equal to any of the unit vectors $\begin{bsmallmatrix}1\\0\end{bsmallmatrix}$, $\begin{bsmallmatrix}0\\1\end{bsmallmatrix}$, the quantities

$$\begin{bsmallmatrix}\alpha\gamma\\\alpha\delta+\beta\gamma\end{bsmallmatrix}, \begin{bsmallmatrix}0\\\alpha\gamma\end{bsmallmatrix}, \begin{bsmallmatrix}\beta\delta\\0\end{bsmallmatrix}, \begin{bsmallmatrix}\alpha\delta+\beta\gamma\\\beta\delta\end{bsmallmatrix}$$

arising in (14.36b) are of the form $\begin{bsmallmatrix}0\\0\end{bsmallmatrix}$, $\begin{bsmallmatrix}1\\0\end{bsmallmatrix}$, or $\begin{bsmallmatrix}0\\1\end{bsmallmatrix}$.

Remark 14.26. Given $\mathbf{v}, \mathbf{w} \in \otimes^d\mathbb{K}^2$, Lemmata 14.24 and 14.25 allow to find the unique $\mathbf{u} \in \otimes^{d+1}\mathbb{K}^2$ with $\mathbf{v} \star \mathbf{w} \sim \mathbf{u}$. In the following, we write $\mathbf{v} \star \mathbf{w}$ instead of $\mathbf{u}$. This notation coincides with the definition in (14.23).

14.3.4.2 Realisation in Different Formats

First, we assume that $v, w \in \mathbb{K}^n$ are represented by $v = \varPhi(\mathbf{v})$ and $w = \varPhi(\mathbf{w})$ with elementary tensors $\mathbf{v} = \bigotimes_{j=1}^{d} v^{(j)}$ and $\mathbf{w} = \bigotimes_{j=1}^{d} w^{(j)}$. Already the convolution $v^{(1)} \star w^{(1)} \in \otimes^2\mathbb{K}^2$ yields a tensor of rank 2, as seen from (14.35a). Assume by induction that $\big(\bigotimes_{j=1}^{d-1} v^{(j)}\big)\star\big(\bigotimes_{j=1}^{d-1} w^{(j)}\big) \sim \mathbf{a} \in \otimes^d\mathbb{K}^2$ has a representation rank 2^{d-1}. Then (14.36b) yields a representation rank 2^d. Since $2^d = n$ is the bound of the maximal tensor rank in $\otimes^{d+1}\mathbb{K}^2$, the r-term representation may yield large representation ranks for $\mathbf{v} \star \mathbf{w}$, even if $\operatorname{rank}(\mathbf{v}) = \operatorname{rank}(\mathbf{w}) = 1$. Hence, the r-term format R_r is not a proper choice for the convolution.

Since the tensor subspace format $\mathcal{T}_\mathbf{r}$ is questionable anyway (see discussion in §14.1.1), we use the $\mathcal{H}_{\boldsymbol\rho}^{\text{tens}}$ format as described in §14.1.2. We recall the involved subspaces $\mathbf{U}_j \subset \otimes^j\mathbb{K}^2$ for $1 \le j \le d$ (cf. (14.4a-c)).

Theorem 14.27. *Let tensors* $\mathbf{v}, \mathbf{w} \in \otimes^d\mathbb{K}^2$ *be represented as* $\mathbf{v} \in \mathcal{H}_{\boldsymbol\rho'}^{\text{tens}}$ *and* $\mathbf{w} \in \mathcal{H}_{\boldsymbol\rho''}^{\text{tens}}$ *involving the respective subspaces* $\mathbf{U}'_j$ *and* $\mathbf{U}''_j$, $1 \le j \le d$, *i.e.,*

$$\begin{array}{llll} \mathbf{U}'_1 = \mathbb{K}^2, & \mathbf{U}'_j \subset \mathbf{U}'_{j-1} \otimes \mathbb{K}^2, & \dim(\mathbf{U}'_j) = \rho'_j, & \mathbf{v} \in \mathbf{U}'_d, \\ \mathbf{U}''_1 = \mathbb{K}^2, & \mathbf{U}''_j \subset \mathbf{U}''_{j-1} \otimes \mathbb{K}^2, & \dim(\mathbf{U}''_j) = \rho''_j, & \mathbf{w} \in \mathbf{U}''_d. \end{array} \tag{14.37a}$$

Then $\mathbf{v} \star \mathbf{w} \in \otimes^{d+1}\mathbb{K}^2$ *belongs to the format* $\mathcal{H}_{\boldsymbol\rho}^{\text{tens}}$ *with*

$$\rho_1 = \rho_{d+1} = 2, \qquad \rho_j \le 2\rho'_j\rho''_j \quad (1 \le j \le d) \tag{14.37b}$$

The involved subspaces

$$\mathbf{U}_j := \operatorname{span}\{\phi'_{j+1}(\mathbf{x} \star \mathbf{y}), \phi''_{j+1}(\mathbf{x} \star \mathbf{y}) : \mathbf{x} \in \mathbf{U}'_j,\ \mathbf{y} \in \mathbf{U}''_j\} \ (1 \le j \le d) \tag{14.37c}$$

with $\dim(\mathbf{U}_j) = \rho_j$ *satisfy again*

$$\mathbf{U}_1 = \mathbb{K}^2, \quad \mathbf{U}_j \subset \mathbf{U}_{j-1} \otimes \mathbb{K}^2 \ (2 \le j \le d+1), \quad \mathbf{v} \star \mathbf{w} \in \mathbf{U}_{d+1}. \tag{14.37d}$$

Proof. 1) By Lemma 14.24, $\mathbf{U}_1$ defined in (14.37c) equals $\mathbb{K}^2$ as required in (14.4a).

2) Let $j \in \{2, \ldots, d\}$. Because of $\mathbf{U}'_j \subset \mathbf{U}'_{j-1} \otimes \mathbb{K}^2$ and $\mathbf{U}''_j \subset \mathbf{U}''_{j-1} \otimes \mathbb{K}^2$, we have

$$\mathbf{U}_j \subset \operatorname{span}\left\{\phi'_{j+1}(\mathbf{x} \star \mathbf{y}), \phi''_{j+1}(\mathbf{x} \star \mathbf{y}) : \left\{\begin{array}{l} \mathbf{x} = \mathbf{v} \otimes x,\ \mathbf{v} \in \mathbf{U}'_{j-1},\ x \in \mathbb{K}^2 \\ \mathbf{y} = \mathbf{w} \otimes y,\ \mathbf{w} \in \mathbf{U}''_{j-1},\ y \in \mathbb{K}^2 \end{array}\right\}\right\}.$$

By (14.36a), $\mathbf{v} \star \mathbf{w} \in \operatorname{span}\{\mathbf{a}', \mathbf{a}''\} \otimes \mathbb{K}^2$ holds with $\mathbf{a}', \mathbf{a}'' \in \mathbf{U}_{j-1}$. The tensors $\mathbf{u}' = \phi'_{j+1}(\mathbf{x} \star \mathbf{y})$ and $\mathbf{u}'' = \phi''_{j+1}(\mathbf{x} \star \mathbf{y})$ from (14.36b) belong to $\mathbf{U}_{j-1} \otimes \mathbb{K}^2$, proving $\mathbf{U}_j \subset \mathbf{U}_{j-1} \otimes \mathbb{K}^2$. □

The fact that the ranks are squared is the usual consequence of binary operations (cf. §13.5.3). The factor 2 in $\rho_j \le 2\rho'_j\rho''_j$ is the carry-over effect. The exact computation using a frame $\mathfrak{b}^{(j)}$ of size ρ_j spanning $\mathbf{U}_j$ can be followed by an orthonormalisation and truncation.

14.4 Fast Fourier Transform

The fast Fourier algorithm uses the same hierarchical structure as the tensorisation. Therefore it is not surprising that the fast Fourier transform can be realised by means of the tensors without using the original vectors. After recalling the algorithm for vectors in §14.4.1, the tensorised version is derived in §14.4.2. The latter algorithm is studied by Dolgov et al. [49], which also describes the sine and cosine transforms.

14.4.1 FFT for $\mathbb{C}^n$ Vectors

Let $n = 2^d$ and $\omega_d := \exp(2\pi\mathrm{i}/n)$. The discrete Fourier transform (DFT) is the mapping $v \in \mathbb{C}^n$ into $\hat{v} \in \mathbb{C}^n$ defined by

$$\hat{v} = F_d v \quad \text{with } F_d = \tfrac{1}{\sqrt{n}}(\omega_d^{k\ell})_{k,\ell=0}^{n-1}.$$

The inverse Fourier transform $\hat{v} \mapsto v$ is described by F_d^{H} involving $\overline{\omega_d}$ instead of ω_d.

We recall the fast Fourier transform (FFT) in the case of $n = 2^d$. If $d = 0$, $\hat{v} = v$ holds. Otherwise, introducing $v^I = (v_k)_{k=0}^{n/2-1}$ and $v^{II} = (v_k)_{k=n/2}^{n-1}$, we observe that

$$\begin{aligned}
\hat{v}_{2k} &= \frac{1}{\sqrt{n}}\left[\sum_{\ell=0}^{\frac{n}{2}-1}\omega_d^{2k\ell}v_\ell + \sum_{\ell=\frac{n}{2}}^{n-1}\omega_d^{2k\ell}v_{\ell+\frac{n}{2}}\right] = \frac{1}{\sqrt{2}}F_{d-1}\big(v^I + v^{II}\big)\,,\\
\hat{v}_{2k+1} &= \frac{1}{\sqrt{n}}\left[\sum_{\ell=0}^{\frac{n}{2}-1}\omega_d^{2k\ell}\omega_d^{\ell}v_\ell - \sum_{\ell=\frac{n}{2}}^{n-1}\omega_d^{2k\ell}\omega_d^{\ell}v_{\ell+\frac{n}{2}}\right] = \tfrac{1}{\sqrt{2}}F_{d-1}\big(\varpi_d \odot \big(v^I - v^{II}\big)\big)
\end{aligned}$$

for $0 \le k \le n/2 - 1$. The last expression uses a Hadamard product with the vector

$$\varpi_d := (\omega_d^{\ell})_{\ell=0}^{n/2-1} \in \mathbb{C}^{n/2}.$$

For an algorithmic description one needs a function *Divide* with the property $v^I = Divide(v,1)$, $v^{II} = Divide(v,2)$, and a function[12] *Merge*, such that the arguments $u = (u_0, u_1, \ldots, u_{n/2-1})^{\mathsf{T}}$ and $v = (v_0, v_1, \ldots, v_{n/2-1})^{\mathsf{T}}$ are mapped into $w := Merge(u,v) \in \mathbb{C}^n$ with $w = (u_0, v_0, u_1, v_1, \ldots, u_{n/2-1}, v_{n/2-1})^{\mathsf{T}}$, i.e., $w_{2k} = u_k$ and $w_{2k+1} = v_k$. Then the discrete Fourier transform can be performed by the following recursive function:

```
function $DFT(v,d)$;                                        $\{v \in \mathbb{C}^n$ with $n = 2^d\}$
if $d = 0$ then $DFT := v$ else
begin $v^I := Divide(v,1)$; $v^{II} := Divide(v,2)$;                        (14.38)
      $DFT := \frac{1}{\sqrt{2}}Merge(DFT(v^I+v^{II}, d-1), DFT(\varpi_d \odot (v^I - v^{II}), d-1))$
end;
```

[12] After d merge steps the bit reversal is already performed.

14.4.2 FFT for Tensorised Vectors

The vectors v and $\hat{v} = F_d v$ correspond to tensors $\mathbf{v}, \hat{\mathbf{v}} \in \otimes^d \mathbb{C}^2$ with $\Phi_n(\mathbf{v}) = v$ and $\Phi_n(\hat{\mathbf{v}}) = \hat{v}$. Note that

$$\hat{\mathbf{v}} = \mathbf{F}_d \mathbf{v} \qquad \text{for } \mathbf{F}_d := \Phi_n^{-1} F_d \, \Phi_n. \tag{14.39}$$

To perform $\mathbf{F}_d$ directly, we have to rewrite the function DFT from (14.38) for the tensorised quantities.

Lemma 14.28. *Assume $n = 2^d$, $d \geq 1$. (a) The tensor $\mathbf{v}^I \in \otimes^{d-1}\mathbb{C}^2$ satisfying $\Phi_n(\mathbf{v}^I) = v^I = Divide(v, 1)$ is obtained by evaluating $\mathbf{v} = \Phi_n^{-1}(v)$ at $i_d = 0$, i.e.,*

$$\mathbf{v}^I[i_1, \ldots, i_{d-1}] := \mathbf{v}[i_1, \ldots, i_{d-1}, 0] \qquad \text{for all } 0 \leq i_j \leq 1, 1 \leq j \leq d-1.$$

Similarly, $\Phi_n(\mathbf{v}^{II}) = v^{II} = Divide(v, 2)$ holds for

$$\mathbf{v}^I[i_1, \ldots, i_{d-1}] := \mathbf{v}[i_1, \ldots, i_{d-1}, 1] \qquad \text{for all } 0 \leq i_j \leq 1, 1 \leq j \leq d-1.$$

(b) Let $w^I, w^{II} \in \mathbb{C}^{n/2}$ with $w := Merge(w^I, w^{II}) \in \mathbb{C}^n$. The tensorised quantities $\mathbf{w}^I = \Phi_{n/2}^{-1}(w^I)$, $\mathbf{w}^{II} = \Phi_{n/2}^{-1}(w^{II})$, and $\mathbf{w} = \Phi_n^{-1}(w)$ satisfy

$$\mathbf{w} = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] \otimes \mathbf{w}^I + \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] \otimes \mathbf{w}^{II}.$$

(c) $\boldsymbol{\varpi}_d = \bigotimes_{j=1}^{d-1} \begin{bmatrix} 1 \\ \omega_d^{2^{j-1}} \end{bmatrix}$ with $\Phi_{n/2}(\boldsymbol{\varpi}_d) = \varpi_d$ is an elementary tensor.

Proof. For (a), (b) use definition (14.1c). For (c) see Remark 5.18 and (14.8). □

By Lemma 14.28 there are direct realisations of $Divide$ and $Merge$ for the tensor counterpart, which we denote by $\mathbf{Divide}$ and $\mathbf{Merge}$. The tensorised version of DFT is $\hat{\mathbf{v}} = \mathbf{DFT}(\mathbf{v}, d)$ satisfying (14.39). The analogous algorithmic description is[13]

```
function $\mathbf{DFT}(\mathbf{v}, d)$;                                   $\{\mathbf{v} \in \otimes^d \mathbb{C}^2\}$
if $d = 0$ then $\mathbf{DFT} := \mathbf{v}$ else
begin $\mathbf{v}^I := \mathbf{Divide}(\mathbf{v}, 1)$; $\mathbf{v}^{II} := \mathbf{Divide}(\mathbf{v}, 2)$;
  $\mathbf{DFT} := \frac{1}{\sqrt{2}}\mathbf{Merge}(\mathbf{DFT}(\mathbf{v}^I + \mathbf{v}^{II}, d-1), \mathbf{DFT}(\boldsymbol{\varpi}_d \odot (\mathbf{v}^I - \mathbf{v}^{II}), d-1))$
end;
```

Define $\Omega_d \mathbf{v}$ by $\frac{1}{\sqrt{2}}\mathbf{Merge}(\mathbf{v}^I + \mathbf{v}^{II}, \boldsymbol{\varpi}_d \odot (\mathbf{v}^I - \mathbf{v}^{II}))$ and observe the identity $\mathbf{F}_d = (id \otimes \mathbf{F}_{d-1}) \, \Omega_d$. The d-fold recursion yields

$$\mathbf{F}_d = \Omega_1 \Omega_2 \cdots \Omega_{d-1} \Omega_d,$$

where Ω_j applies to the components $d - j + 1, \ldots, d$, while the components $1, \ldots, d - j$ remain unchanged.

[13] Instead of multiplying by $1/\sqrt{2}$ in each step, one can divide by $\sqrt{n}$ in the end.

In its exact form this procedure is unsatisfactory. Each Ω_j doubles the number of terms so that finally $n = 2^d$ terms are created. As a consequence, the amount of work is not better than the usual FFT for vectors. Instead, after each step a truncation is applied:

$$\mathbf{F}_d^{\mathsf{trunc}} = T_1\Omega_1 T_2\Omega_2 \cdots T_{d-1}\Omega_{d-1}T_d\Omega_d \qquad (T\text{: truncation}). \tag{14.40}$$

The (nonlinear) truncation operator T_j appearing in (14.40) can be based on a prescribed accuracy or a prescribed rank.

Lemma 14.29. *Given* $\mathbf{v}$, *set*

$$\begin{aligned} \mathbf{v}_{(d+1)} &:= \tilde{\mathbf{v}}_{(d+1)} := \mathbf{v} \qquad \textit{and} \\ \mathbf{v}_{(j)} &:= T_j\Omega_j\mathbf{v}_{(j+1)}, \qquad \tilde{\mathbf{v}}_{(j)} := T_j\Omega_j\tilde{\mathbf{v}}_{(j+1)} \qquad \textit{for } j = d, \ldots, 1. \end{aligned}$$

Then $\tilde{\mathbf{v}}_{(1)} = \mathbf{F}_d^{\mathsf{trunc}}\mathbf{v}$ *has to be compared with* $\mathbf{v}_{(1)} = \mathbf{F}_d\mathbf{v}$. *Assume that* T_j *is chosen such that*

$$\left\|T_j\left(\Omega_j\tilde{\mathbf{v}}_{(j+1)}\right) - \Omega_j\tilde{\mathbf{v}}_{(j+1)}\right\| \le \varepsilon\left\|\Omega_j\tilde{\mathbf{v}}_{(j+1)}\right\|$$

holds with respect to the Euclidean norm. Then the resulting error can be estimated by

$$\left\|\mathbf{F}_d^{\mathsf{trunc}}\mathbf{v} - \mathbf{F}_d\mathbf{v}\right\| \le \left[(1+\varepsilon)^d - 1\right]\|\mathbf{v}\| \approx d\varepsilon\,\|\mathbf{v}\|\,.$$

Proof. Set $\delta_j := \left\|\tilde{\mathbf{v}}_{(j)} - \mathbf{v}_{(j)}\right\|$. Note that $\delta_{(d+1)} = 0$. Since the operation Ω_j is unitary, the recursion

$$\begin{aligned} \delta_j &= \left\|\tilde{\mathbf{v}}_{(j)} - \mathbf{v}_{(j)}\right\| = \left\|\left[T_j\left(\Omega_j\tilde{\mathbf{v}}_{(j+1)}\right) - \Omega_j\tilde{\mathbf{v}}_{(j+1)}\right] + \Omega_j\left(\tilde{\mathbf{v}}_{(j+1)} - \mathbf{v}_{(j+1)}\right)\right\| \\ &\le \varepsilon\left\|\Omega_j\tilde{\mathbf{v}}_{(j+1)}\right\| + \left\|\Omega_j\left(\tilde{\mathbf{v}}_{(j+1)} - \mathbf{v}_{(j+1)}\right)\right\| \le \varepsilon\left\|\tilde{\mathbf{v}}_{(j+1)}\right\| + \delta_{j+1} \\ &\le \varepsilon\left\|\mathbf{v}_{(j+1)}\right\| + (1+\varepsilon)\,\delta_{j+1} = \varepsilon\,\|\mathbf{v}\| + (1+\varepsilon)\,\delta_{j+1} \end{aligned}$$

proves $\delta_j \le \left[(1+\varepsilon)^j - 1\right]\|\mathbf{v}\|\,.$ □

Next, we assume that the ranks $\rho_1, \ldots, \rho_d$ of $\tilde{\mathbf{v}}_{(j)} \in \mathcal{H}_{\boldsymbol{\rho}}^{\mathrm{tens}}$ are uniformly bounded by ρ. The main part of Ω_j is the Hadamard product with $\boldsymbol{\varpi}_j$. The corresponding cost (cf. (14.7)) is of lower order than the truncation cost $O(d\rho^3)$ (cf. (14.9)). Since $\mathbf{F}_d^{\mathsf{trunc}}\mathbf{v}$ is obtained after d steps, we obtain the following result about the computational cost.

Remark 14.30. If, using the $\mathcal{H}_{\boldsymbol{\rho}}^{\mathrm{tens}}$ format, all intermediate results have TT ranks bounded by ρ, the truncated FFT version $\mathbf{F}_d^{\mathsf{trunc}}\mathbf{v}$ costs $O(d^2\rho^3)$ operations.

Note that ρ is the maximum of the ranks of the input tensor $\mathbf{v}$ and of all intermediate results $\tilde{\mathbf{v}}_{(j)}$ including the final Fourier image $\hat{\mathbf{v}}$. Concerning numerical examples, we refer to [49].

14.5 Tensorisation of Functions

So far, tensorisation has been applied to vectors which might be viewed as a *grid function*. Now we use the same formalism for functions. This corresponds to the multiscale treatment of functions.

14.5.1 Isomorphism Φ_n^F

Consider a space $F((a,b])$ of functions defined on the interval $(a,b] \subset \mathbb{R}$. The norm (and, possibly, the scalar product) of $F((a,b])$ must be such that the norm is invariant with respect to a shift, i.e.,

$$\|f\|_{F((a,b])} = \|f(\cdot+\delta)\|_{F((a+\delta,b+\delta])} \,.$$

Furthermore, the function space $F((0,1])$ must allow discontinuities of the functions at ν/n, $1 \le \nu \le n-1$.

In the following we try to establish an isomorphism between $F((0,1])$ and

$$\mathbf{V}_n^F := F((0,1/n]) \otimes \bigotimes_{j=1}^{d} \mathbb{K}^2, \qquad \text{where } n = 2^d.$$

For a better understanding, we first introduce the intermediate tensor space

$$\mathbf{V}_n := F((0,1/n]) \otimes \mathbb{K}^n.$$

Definition 14.31. Define $\hat{\Phi}_n : \mathbf{V}_n \to F((0,1])$ by

$$f = \hat{\Phi}_n(\varphi \otimes v) \in F((0,1]) \text{ for } \varphi \in F((0,1/n]) \text{ and } v = (v_k)_{k=0}^{n-1} \in \mathbb{K}^n \quad (14.41)$$
$$\text{with } f(x) = v_k \cdot \varphi(x - \tfrac{k}{n}) \text{ for } \tfrac{k}{n} < x \le \tfrac{k+1}{n} \text{ and } k \in \{0,\dots,n-1\}.$$

For $v = e^{(k)}$ being the k-th unit vector, $f = \hat{\Phi}_n(\varphi \otimes v)$ can be regarded as φ shifted by k/n, i.e., $f(x+k/n) = \varphi(x)$ for $0 < x \le 1/n$ and $f = 0$ elsewhere.

In general, the function $\hat{\Phi}_n(\varphi \otimes v)$ from (14.41) will be discontinuous, since we do not require $f(\frac{k+1}{n}) = v_k\varphi(1/n) = v_{k+1}\varphi(0-0) = f(\frac{k+1}{n} - 0)$.

Using the isomorphism $\Phi_n : \bigotimes_{j=1}^d \mathbb{K}^2 \to \mathbb{K}^n$ for $n = 2^d$, we obtain $\mathbf{V}_n \cong \mathbf{V}$ and can define the final isomorphism

$$\Phi_n^F : \mathbf{V}_n^F = F((0,1/n]) \otimes \bigotimes_{j=1}^{d} \mathbb{K}^2 \to F((0,1]),$$
$$\Phi_n^F\left(\varphi \otimes \bigotimes_{j=1}^{d} v^{(j)}\right) = \hat{\Phi}_n\left(\varphi \otimes \Phi_n\left(\bigotimes_{j=1}^{d} v^{(j)}\right)\right).$$

Let $\varPsi_n^F$ be the inverse of $\varPhi_n^F$ which maps $F((0,1])$ into $\mathbf{V}_n^F$:

$$\varPsi_n^F : f \in F((0,1]) \mapsto \sum_{k=0}^{n-1} f_k \otimes \varPhi_n^{-1}(e^{(k)}) \quad \text{with } f_k := f(\cdot + \frac{k}{n}) \in F((0,1/n]).$$

Lemma 14.32. *a) For function spaces $F((0,1]) = L^p((0,1])$ $(1 \le p \le \infty)$, the mapping $\varPhi_n^F : \mathbf{V}_n^F \to F((0,1])$ is an isomorphism.*

b) For function spaces $F((0,1]) \subset C((0,1])$, $\varPsi_n^F$ maps $F((0,1])$ into a proper subspace of $\mathbf{V}_n^F$. If one extends $F((0,1])$ to the space

$$F_{\mathrm{pw}}((0,1]) = \bigtimes_{k=0}^{n-1} F\big((\tfrac{k}{n}, \tfrac{k+1}{n}]\big)$$

of piecewise *smooth functions, $\varPhi_n^F$ is again an isomorphism.*

For the representation of a tensor $\mathbf{v} \in \mathbf{V}_n^F = \bigotimes_{j=0}^d V_j$, a variant of $\mathcal{H}_{\boldsymbol{\rho}}^{\mathrm{tens}}$ can be used. Note that $D = \{0, 1, \ldots, d\}$ includes 0. T_D is again the linear tree. The spaces $V_j = \mathbb{K}^2$ for $1 \le j \le d$ are treated as before. In particular, the bases $\{b_1^{(j)}, b_2^{(j)}\}$ are fixed (cf. (14.2b)). Only for $j = 0$, we follow the general concept of $\mathcal{H}_{\mathfrak{r}}$ and choose a subspace $U_0 \subset V_0$ by means of a basis $\{b_i^{(0)} : 1 \le i \le \rho_0\}$.

A prominent example of U_0 is the subspace $\mathcal{P}_{\rho_0 - 1}$ of polynomials of degree $\rho_0 - 1$. A simple basis is $b_k^{(0)}(x) = x^{k-1}$. A more stable choice is are the orthogonal Legendre polynomial (mapped from $[-1,1]$ onto $[0,1/n]$).

14.5.2 Scalar Products

Assume that $V_0 = F((0,1/n])$ is a Hilbert space with the scalar product $\langle \cdot, \cdot \rangle_F$, while $V_j = \mathbb{K}^2$ $(1 \le j \le d)$ is equipped with the Euclidean scalar product $\langle v, w \rangle_2 = v_1 \overline{w_1} + v_2 \overline{w_2}$. For $\mathbf{V}_n^F = \bigotimes_{j=0}^d V_j$ we choose the induced scalar product (cf. (4.62) and Lemma 4.124).

Remark 14.33. For $F((0,1/n]) = L^2((0,1/n])$, the induced scalar product of

$$\mathbf{V}_n^F = F((0,1/n]) \otimes \bigotimes_{j=2}^{d} \mathbb{K}^2$$

coincides with that of $L^2((0,1])$.

For the practical implementation of the scalar product one needs the products $g_{ik} = \langle b_k^{(0)}, b_i^{(0)} \rangle_F$ of the basis functions spanning U_0, which are of course trivial in the case of an orthonormal basis.

14.5.3 Convolution

Let $U_0^{I}\otimes\bigotimes_{j=1}^{d}\mathbb{K}^2$ and $U_0^{II}\otimes\bigotimes_{j=1}^{d}\mathbb{K}^2$ be two subspaces of $\mathbf{V}_n^F$ with possibly different subspaces $U_0^I,U_0^{II}\subset F((0,1/n])$. The convolution of two tensors $\mathbf{v}\in U_0^{I}\otimes\bigotimes_{j=1}^{d}\mathbb{K}^2$ and $\mathbf{w}\in U_0^{II}\otimes\bigotimes_{j=1}^{d}\mathbb{K}^2$ is described in [87, §6]. Here, we suppose that the products $b_k^{I,(0)}\star b_i^{II,(0)}\in F((0,2/n])$ of all basis functions of U_0^I and U_0^{II} are known. The result belongs to $U_0^{III}\otimes\bigotimes_{j=1}^{d+1}\mathbb{K}^2$, where U_0^{III} is spanned by the restrictions $(b_k^{I,(0)}\star b_i^{II,(0)})|_{(0,1/n]}$ and $(b_k^{I,(0)}\star b_i^{II,(0)})(\bullet+1/n)|_{(0,1/n]}$. Note that the result of the convolution is exact.

14.5.4 Continuous Functions

As mentioned in Lemma 14.32, functions from $\mathbf{V}_n^F$ are, in general, discontinuous at $\frac{k}{n}$. This is a drawback for applications, when continuous or even $H^1([0,1])$ functions are required, i.e., $\Phi_n^F(\mathbf{V}_n^F)\subset F([0,1])\subset C([0,1])$.

A possible remedy uses a decomposition which is well-known from hp-finite element approaches. Choose a subspace $U_0\subset F([0,1/n])$ with zero boundary conditions: $f(0)=f(1/n)=0$. Therefore, the space

$$\mathbf{V}_n^I:=U_0\otimes\bigotimes_{j=1}^{d}\mathbb{K}^2$$

is isomorphic to the subspace[14]

$$\Phi_n^I(\mathbf{V}_n^I)\subset F_0[0,1]:=\left\{f\in F_{\mathrm{pw}}([0,1]):f(\tfrac{k}{n})=0\text{ for }0\le k\le n\right\}\subset C([0,1]).$$

Define the space of piecewise linear functions:

$$F_{\mathrm{pl}}[0,1]:=\left\{f\in F([0,1]):f(1)=0\text{ and }f|_{[\frac{k}{n},\frac{k+1}{n}]}\text{ linear for }0\le k\le n-1\right\}$$

and note that $F_{\mathrm{pl}}[0,1]$ is determined by the n nodal values $f(\frac{k}{n})$ for $0\le k\le n-1$. Hence, $F_{\mathrm{pl}}[0,1]$ is isomorphic to $\mathbb{K}^n$ and to

$$\mathbf{V}_n^{II}:=\bigotimes_{j=1}^{d}\mathbb{K}^2.$$

The sum of $F_0[0,1]$ and $F_{\mathrm{pl}}[0,1]$ yields all $f\in F([0,1])$ with the side condition $f(1)=0$. To avoid the latter restriction, a further one-dimensional space must be added.

As a consequence, there are two tensors $\mathbf{v}^I\in\mathbf{V}_n^I$ and $\mathbf{v}^{II}\in\mathbf{V}_n^{II}$ representing a function. Operations, which cause an interaction of both data, like the scalar product, are possible, but more involved (the matrices from Figure 14.1.6 come into play).

[14] For $F([0,1])=C([0,1])$ or $F([0,1])=H^1([0,1])$, $\Phi_n^I(\mathbf{V}_n^I)=F([0,1])$ holds. For smoother function spaces like $F([0,1])=C^1([0,1])$, the derivatives of $f\in\Phi_n^I(\mathbf{V}_n^I)$ are still discontinuous.

Chapter 15
Generalised Cross Approximation

Abstract An important feature is the computation of a tensor from comparably few tensor entries. The input tensor $\mathbf{v} \in \mathbf{V}$ is assumed to be given in a full functional representation so that, on request, any entry can be determined. This partial information can be used to determine an approximation $\tilde{\mathbf{v}} \in \mathbf{V}$. In the matrix case $(d=2)$ an algorithm for this purpose is well-known under the name 'cross approximation' or 'adaptive cross approximation' (ACA). The generalisation to the multi-dimensional case is not straightforward. We present a generalised cross approximation, which fits to the hierarchical format. If $\mathbf{v} \in \mathcal{H}_{\mathfrak{r}}$ holds with a known rank vector $\mathfrak{r}$, this tensor can be reproduced exactly, i.e., $\tilde{\mathbf{v}} = \mathbf{v}$. In the general case, the approximation is heuristic. Exact error estimates require either inspection of all tensor entries (which is practically impossible) or strong theoretical a priori knowledge.
There are many different applications, where tensors are given in a full functional representation. *Section 15.1* gives examples of multivariate functions constructed via integrals and describes multiparametric solutions of partial differential equations, which may originate from stochastic coefficients. *Section 15.2* introduces the definitions of fibres and crosses. The matrix case is recalled in *Sect. 15.3*., while the true tensor case $(d \geq 3)$ is considered in *Sect. 15.4*.

15.1 Approximation of General Tensors

In the following, we consider tensors from $\mathbf{V} = \bigotimes_{j=1}^{d} V_j$ with $V_j = \mathbb{K}^{I_j}$ and recall the *full functional representation* of tensors mentioned in §7.2. In that case, all entries $\mathbf{v}[i_1, \ldots, i_d]$ of $\mathbf{v} \in \mathbf{V}$ are available; however, they are not collectively stored. Instead, they are obtainable via a function $v(i_1, i_2, \ldots, i_d)$. This approach is the only possibility for standard functions, which can never be represented by the infinite number of their function values. In the discrete case, a tensor $\mathbf{v}$ can often be regarded as grid function:

$$\mathbf{v}[i_1, \ldots, i_d] = \varphi(i_1 h, \ldots, i_d h) \qquad (i_j \in I_j \subset \mathbb{Z}). \tag{15.1}$$

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_15,

If, e.g., $I_j = \{0, \ldots, n\}$ and $h = 1/n$, the tensor $\mathbf{v}$ is the restriction of the function $\varphi \in C([0,1]^d)$ to the uniform grid with step size h.

If $\mathbf{v}$ is given in some tensor format together with a procedure for its entrywise evaluation, the full functional representation is given. Note that in the latter case, we even do not need to know the kind of format.

The aim is to approximate the tensor in the hierarchical format $\mathcal{H}_{\mathfrak{r}}$.

Next, we give different examples for tensors in full functional representation.

15.1.1 Approximation of Multivariate Functions

Even if a multivariate function[1] $\varphi \in C([0,1]^d)$ is available, each call of φ might be rather costly. The goal is the computation of an approximation $\tilde{\mathbf{v}}$ to the tensor $\mathbf{v}$ from (15.1) in some tensor format. Then the evaluation of the tensor $\tilde{\mathbf{v}}[i_1, \ldots, i_d]$ approximates the grid value $\varphi(i_1 h, \ldots, i_d h)$. Depending on the representation ranks, the tensor evaluation may be much cheaper than the function evaluation of φ.

As illustration we give an example from boundary element applications. Here, the surface $\Gamma = \partial\Omega$ of a finite domain $\Omega \subset \mathbb{R}^3$ is covered by a triangulation $\mathcal{T}$ (set of triangles). For piecewise constant finite elements (simplest choice) the system matrix $M \in \mathbb{K}^{\mathcal{T}\times\mathcal{T}}$ is given by entries which are the following (four-dimensional) surface integrals:[2]

$$M_{\Delta'\Delta''} = \iint_{\Delta'} \iint_{\Delta''} \frac{\mathrm{d}\Gamma_{\mathbf{x}} \mathrm{d}\Gamma_{\mathbf{y}}}{\|\mathbf{x}-\mathbf{y}\|} \qquad (\Delta', \Delta'' \in \mathcal{T}). \quad (15.2)$$

If the triangles Δ' and Δ'' have a positive distance, the integrand $1/\|x-y\|$ is analytic and there are efficient ways to handle this part of the matrix (cf. [86, Satz 4.2.8]). However, for neighbouring triangles (i.e., $\overline{\Delta'} \cap \overline{\Delta''} \neq \emptyset$) the entries $M_{\Delta'\Delta''}$ must be computed. The required quadrature methods can be found in Sauter-Schwab [166]; nevertheless, the final four-fold Gauss quadrature is expensive and makes the generation of the system data to the costly part of the boundary element method.

Consider the situation of two triangles Δ', Δ'' intersecting in a common side. Since the kernel of (15.2) is rotational and shift-invariant and homogeneous, we may assume without loss of generality that the corners of Δ' are $(0,0,0)$, $(1,0,0)$, $(x,y,0)$, while those of Δ'' are $(0,0,0)$, $(1,0,0)$, (ξ,η,τ). Therefore, up to a scaling factor, all integrals (15.2) with Δ', Δ'' intersecting in a common side are given by the 5-variate function[3]

$$\varphi(x,y,\xi,\eta,\tau) = \iint_{\Delta\left(\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}x\\y\\0\end{bmatrix}\right)} \iint_{\Delta\left(\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}\xi\\\eta\\\tau\end{bmatrix}\right)} \frac{\mathrm{d}\Gamma_{\mathbf{x}} \mathrm{d}\Gamma_{\mathbf{y}}}{\|\mathbf{x}-\mathbf{y}\|}.$$

This function can be evaluated with any desired accuracy.

[1] For simplicity, we choose $[0,1]^d$. Generalisations to $[a_1,b_1] \times \ldots \times [a_d,b_d]$ are trivial.

[2] As a typical example the kernel of the single-layer potential is chosen; cf. Hsiao-Wendland [104].

[3] If the triangles (i) are identical, (ii) have one corner in common, (iii) are disjoint, then the arising integral is a function of 2 (case i), 8 (case ii), 11 (case iii) variables.

Remark 15.1. If the goal is to approximate the grid values (15.1) for later evaluations of the tensor representation of $\tilde{\mathbf{v}}$, the time needed for the computation of $\tilde{\mathbf{v}}$ is irrelevant. The format of the tensor representation should be chosen such that the evaluation[4] is as cheap as possible. The r-term representation is a favourable choice, since the evaluation cost is rd (cf. §13.2.1).

For φ from above and accuracies from 10^{-3} to 10^{-10}, the evaluation of the tensor representation turns out to be faster than quadrature by a factor 630 to 2800 (cf. Ballani [5, Tab. 2]).

In the previous example the parameters are connected with the integration domain. They may also appear in the integrand:

$$\varphi(p_1,\ldots p_d) := \iiint F(\mathbf{x}, p_1, \ldots p_d)\mathrm{d}\mathbf{x},$$

whose numerical evaluation may be rather involved.

15.1.2 Multiparametric Boundary Value Problems and PDE with Stochastic Coefficients

15.1.2.1 Formulation of the Problem

The multivariate function may be the solution of a boundary value problem:

$$\begin{aligned} &L(x,\mathbf{p},u)u = f(x,\mathbf{p}) \text{ in } \mathcal{D},\ B(x,\mathbf{p})u = g(x,\mathbf{p}) \text{ on } \partial\mathcal{D}, \\ &\text{where } x\in\mathcal{D},\ \mathbf{p}=(p_1,\ldots p_d),\ p_j\in P_j\ . \end{aligned} \tag{15.3}$$

Here, $\mathcal{D}\subset\mathbb{R}^m$ may be an arbitrary domain. $\mathcal{D}$ may even depend on the parameters $\mathbf{p}$. L is a second order elliptic differential operator, which may depend on u, so that the problem is nonlinear. B is a boundary operator, for instance the restriction to the boundary. Assuming solvability of the boundary value problem for all $\mathbf{p}\in\mathbf{P} := \times_{j=1}^{d} P_j$, we obtain solutions $u=u(x,\mathbf{p})=u(x,p_1,\ldots p_d)$. Formally, u depends on $d+1$ variables $(x,p_1,\ldots p_d)\in\mathcal{D}\times\mathbf{P}$.

After the discretisation, $\mathcal{D}$ is to be replaced by a set $\mathcal{D}_h$ of nodal points. The discrete solution is denoted by $u_h = u_h(x,\mathbf{p})$ for $(x,\mathbf{p})\in\mathcal{D}_h\times\mathbf{P}$. Now, the evaluation of u_h at a certain (grid) point in $\mathcal{D}_h\times\mathbf{P}$ requires the solution of the discrete boundary value problem for fixed parameters $\mathbf{p}\in\mathbf{P}$.

Next, we consider the boundary value problem

$$\operatorname{div} a(x,\omega)\operatorname{grad} u = f(x) \text{ in } \mathcal{D},\ u=0 \text{ on } \partial\Omega,$$

where the coefficients ('random field') a are defined on $\mathcal{D}\times\Omega$, and $\omega\in\Omega$ is a stochastic variable (even f and $\mathcal{D}$ may be stochastic). Ω is a probability space with a probability measure P. To ensure ellipticity, one has to assume the inequalities $0<a_-\le a(x,\omega)\le a_+<\infty$ for (almost all) $(x,\omega)\in\mathcal{D}\times\Omega$.

[4] In particular, in connection with Monte-Carlo methods, evaluations are also called 'samples'.

The Monte-Carlo method solves many (discretised) boundary value problems with different realisations of ω. Another approach is based on the Karhunen-Loève expansion. As shown next, this brings us back to the situation of a multiparametric problem (15.3).

15.1.2.2 Karhunen-Loève Expansion

It is convenient to split $a(x,\omega)$ into $a_0(x)+r(x,\omega)$, where $a_0(x)=\int_\Omega a(x,\omega)\mathrm{d}P(\omega)$ is the mean value, while $r(x,\omega)$ is called fluctuation. Under minimal conditions, $r \in L^2(\mathcal{D} \times \Omega)$ holds. This implies that the operator $\Phi : L^2(\Omega) \to L^2(\mathcal{D})$ defined by the kernel r is compact. By Theorem 4.114 and Corollary 4.115, an infinite singular value decomposition

$$a(x,\omega) = a_0(x) + \sum_{j=1}^{\infty} \sigma_j \phi_j(x) X_j(\omega) \tag{15.4}$$

holds with orthonormal systems $\{\phi_j\}$ and $\{X_j\}$. Since the mean value of X_j is zero, X_j are uncorrelated random variables. By historical reasons, (15.4) is called Karhunen-Loève expansion (cf. Karhunen [109], Loève [141, §37.5B]). Setting $\sigma_0 = 1$, $\phi_0 := a_0$, and $X_1 = P$, we may write $a(x,\omega) = \sum_{j=0}^{\infty} \sigma_j \phi_j(x) X_j(\omega)$. We recall that ϕ_j $(j \in \mathbb{N})$ are the eigenfunctions of $\mathcal{C} := \Phi\Phi^*$, whose kernel is the two-point correlation

$$C_r(x,x') := \int_\Omega r(x,\omega)\, r(x',\omega)\, \mathrm{d}P(\omega).$$

The critical question is how fast the singular values σ_j in (15.4) are decaying. Under suitable conditions, Todor-Schwab [171, 182] prove exponential decay. This allows us to truncate (15.4) and to replace $a(x,\omega)$ by

$$a_M(x,\omega) = a_0(x) + \sum_{j=1}^{M} \sigma_j \phi_j(x) X_j(\omega).$$

Assuming $\|X_j\|_\infty < \infty$, the sets $P_j := \mathrm{range}(X_j) \subset \mathbb{R}$ are bounded. We can substitute ω by $\mathbf{p} \in \mathbf{P} := \times_{j=1}^{M} P_j$ and $X_j(\omega)$ by $p_j \in P_j$. The multiparametric boundary value problem

$$\operatorname{div} a_M(x,\mathbf{p}) \operatorname{grad} u_M = f(x) \text{ in } \mathcal{D} \text{ for all } \mathbf{p} \in \mathbf{P},\ u_M = 0 \text{ on } \partial\Omega,$$
$$\text{with } a_M(x,\mathbf{p}) = a_0(x) + \sum_{j=1}^{M} \sigma_j \phi_j(x) p_j,$$

has a solution $u_M(x,\mathbf{p})$ for all $\mathbf{p} \in \mathbf{P}$. The solution $u_M(x,\omega)$ ('random field') can be obtained by the back-substitution $p_j \mapsto X_j(\omega)$ (cf. [182, Prop. 3.4]).

Tensor-based solutions of elliptic problems with multiparametric or stochastic coefficients can be found in Khoromskij-Schwab [124], Khoromskij-Oseledets [122], and Espig et al. [55]; in particular, we recommend the recent description Schwab-Gittelson [170].

15.1.3 Function of a Tensor

Let $f : \mathbb{K} \to \mathbb{K}$ be any function.[5] Given a tensor $\mathbf{v} \in \mathbf{V} = \bigotimes_{j=1}^{d} V_j$ with $V_j = \mathbb{K}^{I_j}$, the application of f to $\mathbf{v}$ is explained by

$$f(\mathbf{v}) \in \mathbf{V} \quad \text{and} \quad f(\mathbf{v})[i_1, \ldots, i_d] := f\left(\mathbf{v}[i_1, \ldots, i_d]\right). \tag{15.5a}$$

Even if $\mathbf{v}$ is an elementary tensor, there is no easy description of $f(\mathbf{v})$. On the other hand, if $\mathbf{v}$ is given in one of the formats, the evaluation of $\mathbf{v}[i_1, \ldots, i_d]$ is well-defined and therefore also the determination of $f(\mathbf{v}[i_1, \ldots, i_d])$. Hence, $f(\mathbf{v})$ is given in a full functional representation. Again, the goal is to determine an approximation $\mathbf{w} \in \mathbf{V}$ with $\mathbf{w} \approx f(\mathbf{v})$.

In (15.5a), an explicit function f is given. The tensor may also be determined implicitly:

$$\text{given } \mathbf{v} \in \mathbf{V} \text{ find } \mathbf{x} \in \mathbf{V} \text{ such that } g(\mathbf{x}[i_1, \ldots, i_d]) = \mathbf{v}[i_1, \ldots, i_d]. \tag{15.5b}$$

Here, for instance, a Newton iteration can be applied to solve for $\mathbf{x}[i_1, \ldots, i_d]$.

15.2 Notations

As usual, we write $\mathbf{I} = I_1 \times \ldots \times I_d$ and $\mathbf{I}_{[j]} = \times_{k \neq j} I_k$. The terms ‘fibre’ and ‘cross’ will become important. Fibres are a generalisation of rows and columns of matrices.

Definition 15.2 (fibre). For $1 \leq j \leq d$ and $\mathbf{i}_{[j]} = (i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_d) \in \mathbf{I}_{[j]}$ the j-th fibre of $\mathbf{v} \in \mathbf{V}$ at position $\mathbf{i}_{[j]}$ is the vector

$$\mathcal{F}(\mathbf{v}; j, \mathbf{i}_{[j]}) := \mathbf{v}(i_1, \ldots, i_{j-1}, \bullet, i_{j+1}, \ldots, i_d) \in V_j. \tag{15.6a}$$

The simpler notation $\mathcal{F}(\mathbf{v}; j, \mathbf{i})$ for $\mathbf{i} \in \mathbf{I}$ means $\mathcal{F}(\mathbf{v}; j, \mathbf{i}_{[j]})$, where $\mathbf{i}_{[j]}$ is obtained from $\mathbf{i}$ by dropping the j-th tuple element. The involved indices form the (index) fibre

$$\mathcal{F}(j, \mathbf{i}_{[j]}) := \mathcal{F}(j, \mathbf{i}) := \{\mathbf{k} \in \mathbf{I} : k_\ell = i_\ell \text{ for all } \ell \in \{1, \ldots, d\} \backslash \{j\}\}. \tag{15.6b}$$

The first fibre $\mathcal{F}(M; 1, j)$ of a matrix $M \in \mathbb{K}^{I_1 \times I_2}$ is the j-th column $M[\bullet, j] \in \mathbb{K}^{I_1}$, while $\mathcal{F}(M; 2, i)$ is the i-th row $M[i, \bullet] \in \mathbb{K}^{I_2}$.

[5] $f : D \to \mathbb{K}$ may be restricted to a subset $D \subset \mathbb{K}$, if all entries $\mathbf{v}[i_1, \ldots, i_d]$ belong to D.

Remark 15.3. The vector $\mathcal{F}(\mathbf{v};j,\mathbf{i}_{[j]})$ is the $\mathbf{i}_{[j]}$-th column of the matricisation $\mathcal{M}_j(\mathbf{v})\in\mathbb{K}^{I_j\times\mathbf{I}_{[j]}}$.

Definition 15.4 (cross). The following d-tuple from $V_1\times\ldots\times V_d$ is called the *cross* of $\mathbf{v}\in\mathbf{V}$ at $\mathbf{i}\in\mathbf{I}$:

$$\mathcal{C}(\mathbf{v};\mathbf{i}) := \big(\mathcal{F}(\mathbf{v};1,\mathbf{i}),\mathcal{F}(\mathbf{v};2,\mathbf{i}),\ldots,\mathcal{F}(\mathbf{v};d,\mathbf{i})\big). \tag{15.7a}$$

The involved indices form the (index) cross

$$\mathcal{C}(\mathbf{i}) := \bigcup_{j=1}^{d}\mathcal{F}(j,\mathbf{i}). \tag{15.7b}$$

Note that all vectors $\{\mathcal{F}(\mathbf{v};j,\mathbf{i}) : 1\le j\le d\}$ of a cross have the entry $\mathbf{v}[\mathbf{i}]$ in common. In the following example, the matrix cross $\mathcal{C}(\mathbf{v};\mathbf{i})$ at $(3,5)$ contains the fifth column and third row, while $\mathcal{C}(\mathbf{i})$ contains the indicated indices.

				M_{15} M_{25}	
M_{31}	M_{32}	M_{33}	M_{34}	M_{35}	M_{36}
				M_{45} M_{55}	

Lemma 15.5. *Let* $\mathbf{v}\in\mathbf{V}=\mathbb{K}^{\mathbf{I}}$ *and* $\mathbf{i}\in\mathbf{I}$ *satisfy* $\mathbf{v}[\mathbf{i}]\neq 0$. *The elementary tensor*

$$\mathcal{E}(\mathbf{v};\mathbf{i}) := (\mathbf{v}[\mathbf{i}])^{1-d}\bigotimes_{j=1}^{d}\mathcal{F}(\mathbf{v};j,\mathbf{i})\in\mathcal{R}_1 \tag{15.8a}$$

coincides with $\mathbf{v}$ *on the cross* $\mathcal{C}(\mathbf{i})$, *i.e.,*

$$\mathcal{E}(\mathbf{v};\mathbf{i})[\mathbf{k}] = \mathbf{v}[\mathbf{k}] \qquad \textit{for all } \mathbf{k}\in\mathcal{C}(\mathbf{i}). \tag{15.8b}$$

Proof. $\mathbf{k}=(k_1,i_2,\ldots,i_d)$ belongs to $\mathcal{C}(\mathbf{i})$ for all $k_1\in I_1$. Note that $\mathbf{v}[\mathbf{k}] = \mathcal{F}(\mathbf{v};1,\mathbf{i})[k_1]$ and

$$\mathcal{E}(\mathbf{v};\mathbf{i})[\mathbf{k}] = \frac{1}{\mathbf{v}[\mathbf{i}]^{d-1}}\prod_{j=1}^{d}\mathcal{F}(\mathbf{v};j,\mathbf{i})[k_j] = \frac{1}{\mathbf{v}[\mathbf{i}]^{d-1}}\underbrace{\mathcal{F}(\mathbf{v};1,\mathbf{i})[k_1])}_{=\mathbf{v}[\mathbf{k}]}\prod_{j=2}^{d}\mathcal{F}(\mathbf{v};j,\mathbf{i})[i_j]$$

because of $k_j=i_j$ for $j\ge 2$. Since $\mathcal{F}(\mathbf{v};j,\mathbf{i})[i_j]=\mathbf{v}[\mathbf{i}]$ for $j\ge 2$, the identity $\mathcal{E}(\mathbf{v};\mathbf{i})[\mathbf{k}]=\mathbf{v}[\mathbf{k}]$ follows. □

Property (15.8b) describes the *interpolation on the cross* $\mathcal{C}(\mathbf{i})$. The tensor $\mathcal{E}(\mathbf{v};\mathbf{i})$ can be used for approximation.

Corollary 15.6. Let $\mathbf{v}\in\mathcal{R}_s$ with $s\ge r$. An approximation $\mathbf{u}\in\mathcal{R}_r$ can be obtained by the following procedure:

```
$\mathbf{u} := 0$;  $\mathbf{d} := \mathbf{v}$;  for $\rho := 1$ to $r$ do                                  (15.9)
begin choose $\mathbf{i}\in\mathbf{I}$ with $\mathbf{d}[\mathbf{i}]\neq 0$; $\mathbf{u} := \mathbf{u}+\mathcal{E}(\mathbf{d};\mathbf{i})$; $\mathbf{d} := \mathbf{d}-\mathcal{E}(\mathbf{d};\mathbf{i})$ end;
```

Since $\|\mathbf{v}-\mathcal{E}(\mathbf{v};\mathbf{i})\|\ge\|\mathbf{v}\|$ may hold, the optimal scaling would help (cf. [54, Lemma 6.7]), but it requires a scalar product.

Remark 15.7. Let $\mathbf{v} \in \mathbf{V} = \mathbb{K}^{\mathbf{I}}$ and $\mathbf{i} \in \mathbf{I}$ with $\mathbf{v}[\mathbf{i}] \neq 0$. Set $\lambda := \frac{\langle \mathbf{v}, \mathcal{E}(\mathbf{v};\mathbf{i})\rangle}{\|\mathcal{E}(\mathbf{v};\mathbf{i})\|^2}$. Then $\lambda\mathcal{E}(\mathbf{v};\mathbf{i})$ is an approximation satisfying

$$\|\mathbf{v} - \lambda\mathcal{E}(\mathbf{v};\mathbf{i})\|^2 = \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v}, \mathcal{E}(\mathbf{v};\mathbf{i})\rangle|^2}{\|\mathcal{E}(\mathbf{v};\mathbf{i})\|^2}.$$

15.3 Properties in the Matrix Case

Cross approximation originates from the matrix case. The fibres are rows and columns. $\mathcal{E}(\mathbf{v};\mathbf{i})$ corresponds to the rank-1 matrix $E(M,i,j) \in \mathbb{K}^{I\times J}$ with entries $E_{\nu,\mu} = M_{i,\mu}M_{\nu,j}/M_{i,j}$ associated to the cross $\mathcal{C}(\mathbf{i})$ centred at $\mathbf{i} = (i,j) \in I \times J$. Using the row $M[i,\bullet]$ and the column $M[\bullet,j]$, we may write

$$E(M,i,j) = \tfrac{1}{M[i,j]} M[i,\bullet]M[\bullet,j].$$

Algorithm (15.9) from Corollary 15.6 can be rewritten for matrices:

```
$M \in \mathbb{K}^{I\times J}$ input matrix; $M_0 := M$; $R_0 := 0$;
for $\ell := 1$ to $r$ do
begin choose $(i_\ell, j_\ell) \in I \times J$ with $M[i_\ell, j_\ell] \neq 0$;
      $R_\ell := R_{\ell-1} + E(M_{\ell-1}, i, j)$;  $M_\ell := M_{\ell-1} - E(M_{\ell-1}, i, j)$
end;
```

(15.10)

Lemma 15.8. *(a) If* $r = \mathrm{rank}(M)$, *algorithm (15.10) is well-defined and results in* $R_r = M$ *and* $M_r = 0$.
(b) If the loop in (15.10) terminates at step ℓ *because of* $M_\ell[i,j] = 0$ *for all entries,* $\ell = \mathrm{rank}(M)$ *and* $R_\ell = M$ *hold.*

Proof. Prove that $\mathrm{rank}(M_\ell) = \mathrm{rank}(M_{\ell-1}) - 1$. This leads to statement (a), while (b) is equivalent to (a). □

Property (b) in Lemma 15.8 is called *rank revealing*.

Denote the cross centres (also called 'pivots') in (15.10) by (i_ℓ, j_ℓ), $1 \leq \ell \leq r$. It turns out that the result is independent of the order in which these centres are chosen. Set

$$\tau := \{i_1, \ldots, i_r\} \subset I, \quad \sigma := \{j_1, \ldots, j_r\} \subset J.$$

Then

$$R_r = M|_{I\times\sigma} \left(M|_{\tau\times\sigma}\right)^{-1} M|_{\tau\times J} \tag{15.11}$$

holds (cf. [86, §9.4], for the notation see §1.7).

A direct interpretation of the sets τ, σ with $\#\tau = \#\sigma = r := \mathrm{rank}(M)$ follows from Remark 2.1f: M possesses at least one regular $r\times r$ submatrix. The corresponding row and column indices form a possible choice of the index subsets τ, σ.

Another question is the *approximation* of $M \in \mathbb{K}^{I\times J}$ by matrices $R_r \in \mathcal{R}_r$. Assume that $M = M_r + S$, where $M_r \in \mathcal{R}_r$ is the desired rank-r matrix, while

S is a small perturbation. According to (15.10), we have to avoid pivots (i,j) with $M_r[i,j]=0$. Such indices yield $M[i,j]=S[i,j]$, which by assumption is small. Hence, the criterion $M_{i,j}\neq 0$ in (15.10) has to be replaced by '$M_{i,j}$ as large as possible'.[6] Alternatively, one may ask for subsets $\tau\subset I$, $\sigma\subset J$ in (15.11) with $\#\tau=\#\sigma=r$ such that $|\det(M|_{\tau\times\sigma})|$ is maximal. In fact, Goreinov-Tyrtyshnikov [70] prove that this choice is close to the optimal one (see also [71]). However, for large matrices it is infeasible to check all $M_{i,j}$ or even all block determinants $\det(M|_{\tau\times\sigma})$.

Because of the later application to tensors, we prefer the representation (15.11) and add a remark concerning the cheap recursive computation of $(M|_{\tau\times\sigma})^{-1}$.

Remark 15.9. Set $S_{\tau\times\sigma}:=M|_{\tau\times\sigma}$ and $T_{\tau\times\sigma}:=S_{\tau\times\sigma}^{-1}$. For $\#\tau=\#\sigma=1$, the computation of $T_{\tau\times\sigma}$ is trivial. Otherwise, let $\tau=\tau'\cup\{p\}$ and $\sigma=\sigma'\cup\{q\}$ and assume that $T_{\tau'\times\sigma'}$ is known. $S_{\tau\times\sigma}$ is of the form[7] $\begin{bmatrix} S_{\tau'\times\sigma'} & a \\ b^{\mathsf{T}} & c\end{bmatrix}$, where $\begin{bmatrix} a \\ c\end{bmatrix}=M[\bullet,q]$ and $\begin{bmatrix} b^{\mathsf{T}} & c\end{bmatrix}=M[p,\bullet]$. Then

$$T_{\tau\times\sigma}=\frac{1}{d}\begin{bmatrix} T_{\tau'\times\sigma'}\left(dI-ab^{\mathsf{T}}T_{\tau'\times\sigma'}\right) & -T_{\tau'\times\sigma'}a \\ b^{\mathsf{T}}T_{\tau'\times\sigma'} & 1\end{bmatrix} \quad \begin{array}{l}\text{with } d:= \\ c-b^{\mathsf{T}}T_{\tau'\times\sigma'}\,a.\end{array} \tag{15.12a}$$

The evaluation of $M-R_r=M-M|_{I\times\sigma}\left(M|_{\tau\times\sigma}\right)^{-1}M|_{\tau\times J}$ at $[i,j]$ with $i\notin\tau$ and $j\notin\sigma$ requires only the computation of $M|_{\{i\}\times\sigma}$ and $M|_{\tau\times\{j\}}$:

$$\left(M-R_r\right)[i,j]=M\,[i,j]-\sum_{i'\in\tau}\sum_{j'\in\sigma}M[i,j']\,T_{\tau\times\sigma}[j',i']\,M[i',j]. \tag{15.12b}$$

The *adaptive cross approximation* (ACA) tries to choose a new pivot i,j such that $|\left(M-R_{\ell-1}\right)[i,j]|$ is maximal. The search for i,j is, however, restricted to few crosses. We start with some (i_ℓ,j_ℓ) such that $\left(M-R_{\ell-1}\right)[i_\ell,j_\ell]\neq 0$ and try to improve the choice by iterating $(i_\ell,j_\ell):=\mathit{ImprovedPivot}(M-R_{\ell-1},(i_\ell,j_\ell))$:

```
function $\mathit{ImprovedPivot}(X,(i_\ell,j_\ell))$;
begin $j_\ell :=$ argmax$_j\,|X[i_\ell,j]|$ ; $i_\ell :=$ argmax$_i\,|X[i,j_\ell]|$ ;          (15.13a)
    $\mathit{ImprovedPivot} := (i_\ell,j_\ell)$
end;
```

The second line requires the evaluation of $M[\bullet,j_\ell]$ and $M[i_\ell,\bullet]$.

The adaptive cross approximation may be performed as follows:

```
$R_0 := 0$; $\tau := \sigma := \emptyset$;
for $\ell := 1$ to $r$ do
begin choose any $(i_\ell,j_\ell)$ such that $(M-R_{\ell-1})[i_\ell,j_\ell]\neq 0$;
    iterate: $(i_\ell,j_\ell) := \mathit{ImprovedPivot}(M-R_{\ell-1},(i_\ell,j_\ell))$;          (15.13b)
    $R_\ell := R_{\ell-1}+E(M-R_{\ell-1},i_\ell,j_\ell)$;
    $\tau := \tau\cup\{i_\ell\}$; $\sigma := \sigma\cup\{j_\ell\}$; determine $T_{\tau\times\sigma}$ from (15.12a)
end;
```

[6] This choice is also be recommended in the case of $\operatorname{rank}(M)=r$ because of numerical stability.

[7] The ordering of the indices in τ and σ is irrelevant.

The choice in line 3 can be made by random. Note that for $\ell > 1$ it is impossible to choose $(i_\ell, j_\ell) := (i_{\ell-1}, j_{\ell-1})$, since this index belongs to the cross, where $M - R_{\ell-1}$ is vanishing. More generally, $i_\ell \notin \{i_1, \ldots, i_{\ell-1}\}$ and $j_\ell \notin \{j_1, \ldots, j_{\ell-1}\}$ is required. The second last value $(i_{\ell-1}, j_{\ell-1})$ from the loop 4 in the previous step $\ell - 1$ may satisfy this requirement.

The loop in line 4 may be repeated, e.g., three times.

Line 5 contains the explicit definition of R_ℓ. Implicitly, $M - R_r$ is determined by (15.12b) from $T_{\tau \times \sigma}$ defined in line 6.

In the case of approximation, algorithm (15.10) is either performed with fixed r, or the termination depends on a suitable stopping criterion.

Remark 15.10. In (15.10) the number of evaluated matrix entries is bounded by $O(r(\#I + \#J))$. Also for large-scale matrices one assumes that computations and storage of size $O(\#I + \#J)$ are acceptable, while $O(\#I\#J)$ is too large. However, this statement holds for a successful application only.

There are cases where this heuristic approach fails. If a small low-rank matrix is embedded into a large zero matrix, $O(\min(\#I, \#J))$ crosses have to be tested before $(M - R_{\ell-1})\,[i_\ell, j_\ell] \neq 0$ occurs. A second difficulty arises for a block matrix $\begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}$, where the stars indicate nonzero blocks. If (i_ℓ, j_ℓ) is chosen from the first block, all later iterates determined by (15.13a) stay in this block. Only a random choice has a chance to enter the second diagonal block.

Descriptions of the adaptive cross approximation and variations can be found in Bebendorf [8, 9] and Börm-Grasedyck [22].

Finally, we remark that the same approach can be used for multivariate functions. In the case of a function Φ in two variables the analogue of the rank-1 matrix $E(M, i, j)$ is the rank-1 function $E(\Phi, \xi, \eta)(x, y) = \Phi(\xi, y)\Phi(x, \eta)/\Phi(\xi, \eta)$, which interpolates $\Phi(\cdot, \cdot)$ in the lines $x = \xi$ and $y = \eta$, i.e., $\Phi(\xi, y) = E(\Phi, \xi, \eta)(\xi, y)$ and $\Phi(x, \eta) = E(\Phi, \xi, \eta)(x, \eta)$.

15.4 Case $d \geq 3$

As stated in Lemma 15.5, the rank-1 tensor $\mathcal{E}(\mathbf{v}; \mathbf{i})$ interpolates $\mathbf{v}$ at the cross $\mathcal{C}(\mathbf{i})$, i.e., $\mathbf{v}' := \mathbf{v} - \mathcal{E}(\mathbf{v}; \mathbf{i})$ vanishes on $\mathcal{C}(\mathbf{i})$. However, when we choose another cross $\mathcal{C}(\mathbf{i}')$, the next tensor $\mathcal{E}(\mathbf{v}'; \mathbf{i}')$—differently from the matrix case $d = 2$—need not vanish on $\mathcal{C}(\mathbf{i})$ so that the next iterate $\mathbf{v}' - \mathcal{E}(\mathbf{v}'; \mathbf{i}')$ loses the interpolation property on $\mathcal{C}(\mathbf{i})$. As a consequence, the iteration (15.9) yields tensors of rank r, which do not satisfy the statement of Lemma 15.8. In the case of locally best rank-1 approximations, we have seen this phenomenon already in Remark 9.18b.

The properties observed above make it impossible to generalise the cross approximation to higher dimension. Another reason why Lemma 15.8 cannot hold for $d \geq 3$, is Proposition 3.34: tensor rank revealing algorithms must be NP hard.

15.4.1 Matricisation

There are several approaches to get tensor versions of the cross approximation (cf. Espig-Grasedyck-Hackbusch [53], Oseledets-Tyrtyshnikov [159], Bebendorf [10]). Here, we follow the algorithm of Ballani-Grasedyck-Kluge [7].

First, we are looking for the *exact* representation of $\mathbf{v}$ in the hierarchical format $\mathcal{H}_{\mathfrak{r}}$. We recall that the optimal ranks are $r_\alpha = \dim(\mathbf{U}_\alpha^{\min}(\mathbf{v}))$, where $\mathbf{U}_\alpha^{\min}(\mathbf{v}) = \mathrm{range}(\mathcal{M}_\alpha(\mathbf{v}))$ involves the matricisation of $\mathbf{v}$ (cf. (6.15)). $\mathcal{M}_\alpha(\mathbf{v})$ is a matrix from $\mathbb{K}^{\mathbf{I}_\alpha \times \mathbf{I}_{\alpha^c}}$ with the index sets $\mathbf{I}_\alpha := \times_{j\in\alpha} I_j$ and $\mathbf{I}_{\alpha^c} := \times_{j\in\alpha^c} I_j$ ($\alpha^c = D\backslash\alpha$). Theoretically, we may apply the methods of §15.3. Choose pivot subsets

$$P_\alpha = \{\mathbf{p}_1^{(\alpha)}, \ldots, \mathbf{p}_{r_\alpha}^{(\alpha)}\} \subset \mathbf{I}_\alpha, \qquad P_{\alpha^c} = \{\mathbf{p}_1^{(\alpha^c)}, \ldots, \mathbf{p}_{r_\alpha}^{(\alpha^c)}\} \subset \mathbf{I}_{\alpha^c} \tag{15.14a}$$

containing r_α indices such that $\mathcal{M}_\alpha(\mathbf{v})|_{P_\alpha \times P_{\alpha^c}}$ is regular. Then $\mathcal{M}_\alpha(\mathbf{v})$ is equal to

$$\mathcal{M}_\alpha(\mathbf{v})|_{\mathbf{I}_\alpha \times P_{\alpha^c}} \cdot \left(\mathcal{M}_\alpha(\mathbf{v})|_{P_\alpha \times P_{\alpha^c}}\right)^{-1} \cdot \mathcal{M}_\alpha(\mathbf{v})|_{P_\alpha \times \mathbf{I}_{\alpha^c}} \in \mathbb{K}^{\mathbf{I}_\alpha \times \mathbf{I}_{\alpha^c}} \tag{15.14b}$$

(cf. (15.11) and Lemma 15.8). The columns

$$\mathbf{b}_i^{(\alpha)} := \mathcal{M}_\alpha(\mathbf{v})[\bullet, \mathbf{p}_i^{(\alpha^c)}], \qquad 1 \le i \le r_\alpha,$$

form a basis of $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ (cf. (11.20b)). Similarly, $\mathbf{b}_j^{(\alpha^c)} := \mathcal{M}_\alpha(\mathbf{v})[\mathbf{p}_j^{(\alpha)}, \bullet]^\mathsf{T}$ yields a basis of $\mathbf{U}_{\alpha^c}^{\min}(\mathbf{v})$.

In general, $\mathbf{I}_\alpha$ and $\mathbf{I}_{\alpha^c}$ are huge sets, so that neither the matrix $\mathcal{M}_\alpha(\mathbf{v})|_{\mathbf{I}_\alpha \times P_{\alpha^c}}$ nor $\mathcal{M}_\alpha(\mathbf{v})|_{P_\alpha \times \mathbf{I}_{\alpha^c}}$ are practically available. This corresponds to the fact that the bases $\{\mathbf{b}_i^{(\alpha)}\}$ and $\{\mathbf{b}_i^{(\alpha^c)}\}$ need never be computed. The only practically computable quantity is the $r_\alpha \times r_\alpha$ matrix $S_\alpha := \mathcal{M}_\alpha(\mathbf{v})|_{P_\alpha \times P_{\alpha^c}}$ (we use S_α shortly for $S_{P_\alpha \times P_{\alpha^c}}$, which is introduced in Remark 15.9).

Evaluations at $\mathbf{p}_i^{(\alpha)} \in P_\alpha$ are particular functionals:

$$\boldsymbol{\varphi}_i^{(\alpha)} \in \mathbf{V}_\alpha' \quad \text{with } \boldsymbol{\varphi}_i^{(\alpha)}(\mathbf{v}_\alpha) := \mathbf{v}_\alpha[\mathbf{p}_i^{(\alpha)}] \quad \text{for } \mathbf{v}_\alpha \in \mathbf{V}_\alpha.$$

As explained in Notation 3.50b, we identify $\boldsymbol{\varphi}_i^{(\alpha)} \in \mathbf{V}_\alpha'$ and $\boldsymbol{\varphi}_i^{(\alpha)} \in L(\mathbf{V}, \mathbf{V}_{\alpha^c})$:

$$\boldsymbol{\varphi}_i^{(\alpha)}(\mathbf{v}) = \mathbf{v}[\mathbf{p}_i^{(\alpha)}, \bullet] \in \mathbf{V}_{\alpha^c} \quad \text{for } \mathbf{v} \in \mathbf{V}.$$

Similarly, evaluations at $\mathbf{p}_i^{(\alpha^c)} \in P_{\alpha^c}$ are particular functionals $\boldsymbol{\varphi}_i^{(\alpha^c)} \in \mathbf{V}_{\alpha^c}'$ defined by $\boldsymbol{\varphi}_i^{(\alpha^c)}(\mathbf{v}_{\alpha^c}) := \mathbf{v}_{\alpha^c}[\mathbf{p}_i^{(\alpha^c)}]$ for $\mathbf{v}_{\alpha^c} \in \mathbf{V}_{\alpha^c}$. Identification of $\boldsymbol{\varphi}_i^{(\alpha^c)} \in \mathbf{V}_{\alpha^c}'$ and $\boldsymbol{\varphi}_i^{(\alpha^c)} \in L(\mathbf{V}, \mathbf{V}_\alpha)$ yields

$$\boldsymbol{\varphi}_i^{(\alpha^c)}(\mathbf{v}) = \mathbf{v}[\bullet, \mathbf{p}_i^{(\alpha^c)}] \in \mathbf{V}_\alpha \quad \text{for } \mathbf{v} \in \mathbf{V}.$$

The bases $\{\mathbf{b}_i^{(\alpha)}\}$ and $\{\mathbf{b}_i^{(\alpha^c)}\}$ defined above, spanning $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ and $\mathbf{U}_{\alpha^c}^{\min}(\mathbf{v})$, take now the form

$$\mathbf{b}_i^{(\alpha)} = \boldsymbol{\varphi}_i^{(\alpha^c)}(\mathbf{v}), \quad \mathbf{b}_j^{(\alpha^c)} = \boldsymbol{\varphi}_j^{(\alpha)}(\mathbf{v}) \qquad (1 \le i \le r_\alpha)\,. \tag{15.15}$$

Since $\mathbf{v} \in \mathbf{U}_\alpha^{\min}(\mathbf{v}) \otimes \mathbf{U}_{\alpha^c}^{\min}(\mathbf{v})$, there are coefficients c_{ij} with

$$\mathbf{v} = \sum\nolimits_{i,j=1}^{r_\alpha} c_{ij}\, \mathbf{b}_i^{(\alpha)} \otimes \mathbf{b}_j^{(\alpha^c)}. \tag{15.16}$$

Application of $\boldsymbol{\varphi}_\nu^{(\alpha^c)}$ yields $\mathbf{b}_\nu^{(\alpha)} = \boldsymbol{\varphi}_\nu^{(\alpha^c)}(\mathbf{v}) = \sum_{i,j=1}^{r_\alpha} c_{ij}\, \mathbf{b}_i^{(\alpha)} \boldsymbol{\varphi}_\nu^{(\alpha^c)}(\mathbf{b}_j^{(\alpha^c)})$. Since $\mathbf{b}_j^{(\alpha^c)} = \boldsymbol{\varphi}_\nu^{(\alpha^c)}(\mathbf{v})$, the identity $\boldsymbol{\varphi}_\nu^{(\alpha^c)}(\mathbf{b}_j^{(\alpha^c)}) = (\boldsymbol{\varphi}_j^{(\alpha)} \otimes \boldsymbol{\varphi}_\nu^{(\alpha^c)})(\mathbf{v})$ holds. Hence, the matrix $C = (c_{ij}) \in \mathbb{K}^{r_\alpha \times r_\alpha}$ is the inverse $T_\alpha := S_\alpha^{-1}$ of

$$S_\alpha = \left((\boldsymbol{\varphi}_j^{(\alpha)} \otimes \boldsymbol{\varphi}_i^{(\alpha^c)})(\mathbf{v})\right)_{i,j=1}^{r_\alpha} = \left(\mathcal{M}_\alpha(\mathbf{v})[\mathbf{p}_j^{(\alpha)}, \mathbf{p}_i^{(\alpha^c)}]\right)_{i,j=1}^{r_\alpha}. \tag{15.17}$$

Equation (15.16) with $C = T_\alpha$ is the interpretation of (15.14b) in $\mathbf{V}$.

15.4.2 Nestedness

Let $\alpha \in T_D$ with sons α_1 and α_2. While the index sets P_α, P_{α^c} from (15.14a) are associated to α, there are some other index sets $P_{\alpha_1}, P_{\alpha_1^c}$ and $P_{\alpha_2}, P_{\alpha_2^c}$ related to α_1 and α_2. Their cardinalities are

$$\#P_{\alpha_\iota} = \#P_{\alpha_\iota^c} = r_{\alpha_\iota} := \dim(\mathbf{U}_{\alpha_\iota}^{\min}(\mathbf{v})) \qquad (\iota = 1, 2)\,.$$

The bases of $\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v})$ and $\mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})$ are given by

$$\mathbf{b}_i^{(\alpha_1)} = \boldsymbol{\varphi}_i^{(\alpha_1^c)}(\mathbf{v}) := \mathbf{v}[\bullet, \mathbf{p}_i^{(\alpha_1)}], \quad \mathbf{b}_j^{(\alpha_2)} = \boldsymbol{\varphi}_j^{(\alpha_2^c)}(\mathbf{v}) := \mathbf{v}[\bullet, \mathbf{p}_j^{(\alpha_2)}].$$

Since $\mathbf{U}_\alpha^{\min}(\mathbf{v}) \subset \mathbf{U}_{\alpha_1}^{\min}(\mathbf{v}) \otimes \mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})$ (cf. (11.16c)), there are coefficient matrices $C^{(\alpha,\ell)} = (c_{ij}^{(\alpha,\ell)})$ such that

$$\mathbf{b}_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)} \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} \tag{15.18}$$

(cf. (11.24)). For the determination of $c_{ij}^{(\alpha,\ell)}$ we make the ansatz

$$\mathbf{b}_\ell^{(\alpha)} = \sum_{i,\nu=1}^{r_{\alpha_1}} \sum_{j,\mu=1}^{r_{\alpha_2}} c_{i\nu}^{(\alpha_1)} c_{j\mu}^{(\alpha_2)} \left(\boldsymbol{\varphi}_\nu^{(\alpha_1)} \otimes \boldsymbol{\varphi}_\mu^{(\alpha_2)} \otimes \boldsymbol{\varphi}_\ell^{(\alpha^c)}\right)(\mathbf{v})\, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}. \tag{15.19}$$

Since the $r_{\alpha_1} r_{\alpha_2}$ functionals $\boldsymbol{\varphi}_\nu^{(\alpha_1)} \otimes \boldsymbol{\varphi}_\mu^{(\alpha_2)}$ are linearly independent on the space $\mathbf{U}_{\alpha_1}^{\min}(\mathbf{v}) \otimes \mathbf{U}_{\alpha_2}^{\min}(\mathbf{v})$ of dimension $r_{\alpha_1} r_{\alpha_2}$, equation (15.19) holds if and only if all images under $\boldsymbol{\varphi}_{\nu'}^{(\alpha_1)} \otimes \boldsymbol{\varphi}_{\mu'}^{(\alpha_2)}$ are equal. The left-hand side yields

$$\begin{aligned} \left(\boldsymbol{\varphi}_{\nu'}^{(\alpha_1)} \otimes \boldsymbol{\varphi}_{\mu'}^{(\alpha_2)}\right)\left(\mathbf{b}_\ell^{(\alpha)}\right) &\underset{(15.15)}{=} \left(\boldsymbol{\varphi}_{\nu'}^{(\alpha_1)} \otimes \boldsymbol{\varphi}_{\mu'}^{(\alpha_2)}\right)\left(\boldsymbol{\varphi}_\ell^{(\alpha^c)}(\mathbf{v})\right) \\ &= \left(\boldsymbol{\varphi}_{\nu'}^{(\alpha_1)} \otimes \boldsymbol{\varphi}_{\mu'}^{(\alpha_2)} \otimes \boldsymbol{\varphi}_\ell^{(\alpha^c)}\right)(\mathbf{v}), \end{aligned} \tag{15.20a}$$

while the right-hand side equals

$$\sum_{i,\nu=1}^{r_{\alpha_1}}\sum_{j,\mu=1}^{r_{\alpha_2}} c_{i\nu}^{(\alpha_1)} c_{j\mu}^{(\alpha_2)} \left(\boldsymbol{\varphi}_{\nu'}^{(\alpha_1)}\otimes\boldsymbol{\varphi}_{\mu'}^{(\alpha_2)}\otimes\boldsymbol{\varphi}_{\ell}^{(\alpha^c)}\right)(\mathbf{v})\ \boldsymbol{\varphi}_{\nu'}^{(\alpha_1)}(\mathbf{b}_i^{(\alpha_1)})\ \boldsymbol{\varphi}_{\mu'}^{(\alpha_2)}(\mathbf{b}_j^{(\alpha_2)}). \tag{15.20b}$$

As in (15.17), we have

$$S_{\alpha_1}=\left(\boldsymbol{\varphi}_{\nu}^{(\alpha_1)}\otimes\boldsymbol{\varphi}_{i}^{(\alpha_1^c)}\right)(\mathbf{v})=\left(\boldsymbol{\varphi}_{\nu}^{(\alpha_1)}(\mathbf{b}_i^{(\alpha_1)})\right)_{\nu,i=1}^{r_{\alpha_1}},\qquad S_{\alpha_2}=\left(\boldsymbol{\varphi}_{\mu}^{(\alpha_2)}(\mathbf{b}_j^{(\alpha_2)})\right)_{\mu,j=1}^{r_{\alpha_2}}$$

and $T_{\alpha_1}:=S_{\alpha_1}^{-1}$, $T_{\alpha_2}:=S_{\alpha_2}^{-1}$. Set $C_{\alpha_1}=(c_{i\nu}^{(\alpha_1)})_{i,\nu=1}^{r_{\alpha_1}}$ and $C_{\alpha_2}=(c_{j\mu}^{(\alpha_2)})_{j,\mu=1}^{r_{\alpha_2}}$. Then, (15.20b) becomes

$$\sum_{\nu=1}^{r_{\alpha_1}}\sum_{\mu=1}^{r_{\alpha_2}}(S_{\alpha_1}C_{\alpha_1})_{\nu',\nu}\ (S_{\alpha_2}C_{\alpha_2})_{\mu',\mu}\ \left(\boldsymbol{\varphi}_{\nu'}^{(\alpha_1)}\otimes\boldsymbol{\varphi}_{\mu'}^{(\alpha_2)}\otimes\boldsymbol{\varphi}_{\ell}^{(\alpha^c)}\right)(\mathbf{v}). \tag{15.20c}$$

Comparison of (15.20a) and (15.20c) shows that $C_{\alpha_1}=T_{\alpha_1}$ and $C_{\alpha_2}=T_{\alpha_2}$ yield the desired identity. Hence, the coefficients $c_{ij}^{(\alpha,\ell)}$ in (15.18) satisfy

$$\begin{aligned} c_{ij}^{(\alpha,\ell)} &= \sum_{\nu=1}^{r_{\alpha_1}}\sum_{\mu=1}^{r_{\alpha_2}} T_{\alpha_1}[i,\nu]\,T_{\alpha_2}[j,\mu]\ \left(\boldsymbol{\varphi}_{\nu'}^{(\alpha_1)}\otimes\boldsymbol{\varphi}_{\mu'}^{(\alpha_2)}\otimes\boldsymbol{\varphi}_{\ell}^{(\alpha^c)}\right)(\mathbf{v}) \\ &= \sum_{\nu=1}^{r_{\alpha_1}}\sum_{\mu=1}^{r_{\alpha_2}} T_{\alpha_1}[i,\nu]\,T_{\alpha_2}[j,\mu]\,\mathbf{v}[\mathbf{p}_{\nu}^{(\alpha_1)},\mathbf{p}_{\mu}^{(\alpha_2)},\mathbf{p}_{\ell}^{(\alpha^c)}]. \end{aligned}$$

For a matrix notation set $V_\ell^{(\alpha)}:=\left(\mathbf{v}[\mathbf{p}_{\nu}^{(\alpha_1)},\mathbf{p}_{\mu}^{(\alpha_2)},\mathbf{p}_{\ell}^{(\alpha^c)}]\right)_{\nu=1,\ldots,r_{\alpha_1},\ \mu=1,\ldots,r_{\alpha_2}}$ and

$$C^{(\alpha,\ell)}=T_{\alpha_1}\,V_\ell^{(\alpha)}\,T_{\alpha_2}^{\mathsf{T}}. \tag{15.20d}$$

We summarise the results.

Proposition 15.11. *Assume that $\mathbf{v}\in\mathcal{H}_{\mathfrak{r}}$ holds exactly. Then the parameters $C^{(\alpha,\ell)}$, $c^{(D)}$, and B_j of the hierarchical representation can be determined as follows.*
(a) For $\alpha=D$ with sons α_1 and α_2, $\mathbf{v}=\mathbf{b}_1^{(D)}=\sum_{i,j=1}^{r_{\alpha_1}}c_{ij}^{(D,1)}\mathbf{b}_i^{(\alpha_1)}\otimes\mathbf{b}_j^{(\alpha_2)}$ holds with $C^{(D,1)}=T_D$. Furthermore, $c^{(D)}=1\in\mathbb{K}$.
(b) Let $\alpha\in T_D\backslash(D\cup\mathcal{L}(T_D))$ with sons α_1 and α_2. For $\beta\in\{\alpha,\alpha_1,\alpha_2\}$ and $r_\beta:=\mathrm{rank}_\beta(\mathbf{v})$, there are row and column index sets P_β,P_{β^c} of $\mathcal{M}_\alpha(\mathbf{v})$ such that the $r_\beta\times r_\beta$ matrix $\mathcal{M}_\beta(\mathbf{v})|_{P_\beta\times P_{\beta^c}}$ is regular. For any such index sets, bases $\{\mathbf{b}_i^{(\beta)}:1\leq i\leq r_\beta\}$ are defined by $\mathbf{v}[\bullet,\mathbf{p}_i^{(\beta^c)}]$. The coefficient matrix $C^{(\alpha,\ell)}$ for the characteristic relation (15.18) is given by (15.20d).
(c) If in the cases (a,b) a son α_ι belongs to $\mathcal{L}(T_D)$, $\alpha_\iota=\{j\}$ holds for some $j\in D$, and $B_j=[\mathbf{b}_1^{(\alpha_\iota)}\ \mathbf{b}_2^{(\alpha_\iota)}\ \cdots\ \mathbf{b}_{r_j}^{(\alpha_\iota)}]$ is used as basis of $U_j^{\min}(\mathbf{v})\subset V_j$.
(d) The inverses T_α are computed via the recursion (15.12a).

Proof. Part (a) follows from (15.16). Part (b) is explained in §15.4.2. □

15.4.3 Algorithm

15.4.3.1 Provisional Form

Let a tensor $\mathbf{v}$ be given. We want to represent $\mathbf{v}$ (exactly) in the hierarchical format $\mathcal{H}_{\mathfrak{r}}$ with tree T_D and rank vector $\mathfrak{r} = (r_\alpha)_{\alpha \in T_D}$, $r_\alpha = \operatorname{rank}_\alpha(\mathbf{v})$. The first step can be performed independently for all α:

Step 1 for all $\alpha \in T_D$ determine index subsets $P_\alpha \subset \mathbf{I}_\alpha, P_{\alpha^c} \subset \mathbf{I}_{\alpha^c}$ with $\#P_\alpha = \#P_{\alpha^c} = r_\alpha$ (cf. (15.14a)) such that S_α from (15.17) is regular. T_α is computed via (15.12a).

Note that the subsets P_α, P_{α^c} are in general not unique. The bases of $U_j^{\min}(\mathbf{v})$ are determined in the second step:

Step 2 for all $1 \leq j \leq d$ set

$$b_i^{(j)} := \boldsymbol{\varphi}_i^{(\{j\}^c)}(\mathbf{v}) = \mathbf{v}[\bullet, \mathbf{p}_i^{(\{j\}^c)}] \qquad (1 \leq i \leq r_j).$$

The bases $\{\mathbf{b}_i^{(\alpha)} : 1 \leq i \leq r_\alpha\}$ of $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ for non-leaf vertices $\alpha \in T_D \backslash \mathcal{L}(T_D)$ do not enter the algorithm. Instead, the coefficient matrices $C^{(\alpha,\ell)}$ are determined:

Step 3a for all $\alpha \in T_D$ determine S_α from (15.17) and T_α from (15.12a),

Step 3b for all $\alpha \in T_D \backslash (D \cup \mathcal{L}(T_D))$ determine $V_\ell^{(\alpha)} := (\mathbf{v}[\mathbf{p}_\nu^{(\alpha_1)}, \mathbf{p}_\mu^{(\alpha_2)}, \mathbf{p}_\ell^{(\alpha^c)}])_{\nu,\mu}$ (α_1 and α_2 sons of α) and compute $C^{(\alpha,\ell)}$ from (15.20d).

The representation of $\mathbf{v}$ follows from (15.16) with α and α^c being sons of $D \in T_D$:

Step 4 for $\alpha = D$, the matrix $C^{(D,1)}$ is the inverse of S_D. Set $c_1^{(D)} := 1$.

Remark 15.12. (a) For Step 1, one may in principle apply the algorithm from (15.10) until $M_\ell = 0$ (cf. Lemma 15.8b). For theoretical considerations, regularity of S_α is sufficient. Practically, the determinant $|\det(S_\alpha)|$ should not be too large. Therefore, strategies for suitable subsets P_α, P_{α^c} will be discussed later.
(b) $b_i^{(j)} = \mathbf{v}[\bullet, \mathbf{p}_i^{(\{j\}^c)}]$ is the fibre $\mathcal{F}(\mathbf{v}; j, \mathbf{p}_i^{(\{j\}^c)})$ in direction j. Its computation requires $\dim(V_j)$ evaluations of $\mathbf{v}$.
(c) Matrix S_α requires r_α^2 evaluations of $\mathbf{v}$. The matrices $V_\ell^{(\alpha)}$ $(1 \leq \ell \leq r_\alpha)$ need $r_\alpha r_{\alpha_1} r_{\alpha_2}$ evaluations. Altogether,

$$\sum_{\alpha \in T_D \backslash \mathcal{L}(T_D)} r_\alpha r_{\alpha_1} r_{\alpha_2} + \sum_{\alpha \in T_D} r_\alpha^2 + \sum_{j=1}^{d} r_j \dim(V_j)$$

evaluations of $\mathbf{v}$ are required.

In the case of tensorisation with $\dim(V_j) = 2$, the leading term of evaluations is dr^3 ($r := \max_\alpha r_\alpha$), which is surprisingly small compared with the huge total number of tensor entries.

15.4.3.2 Choice of Index Sets

In (15.13a) we have described how to improve the choice of the two-dimensional pivot (i,j). Now we generalise this approach to general order $d \geq 2$. Consider a vertex $\alpha \in T_D / \mathcal{L}(T_D)$ with sons α_1 and α_2. We want to find the pivots for $\mathcal{M}_{\alpha_1}(\mathbf{v})$, whose rows and columns belong to $\mathbf{I}_{\alpha_1}$ and $\mathbf{I}_{\alpha_1^c} = \mathbf{I}_{\alpha_2} \cup \mathbf{I}_{\alpha^c}$. From the previous computation at α we have already a pivot index set $P_{\alpha^c} = \{\mathbf{p}_1^{(\alpha^c)}, \ldots, \mathbf{p}_{r_\alpha}^{(\alpha^c)}\} \subset \mathbf{I}_{\alpha^c}$ (cf. (15.14a)). To reduce the search steps, possible pivots $\mathbf{i} = (i_1, \ldots, i_d) \in \mathbf{I}$ are restricted to those with the property $(i_j)_{j\in\alpha^c} \in P_{\alpha^c}$ (note that there are only r_α tuples in P_{α^c}). The remaining components $(i_j)_{j\in\alpha}$ are obtained by maximisation along one fibre. Again, the starting index $\mathbf{i}$ must be such that $\mathbf{x}[\mathbf{i}] \neq 0$ for the tensor $\mathbf{x} = \mathbf{v} - \mathbf{v}_{\ell-1}$ ($\mathbf{v}_{\ell-1}$: actual approximation). The sentence following (15.13b) is again valid.

The index tuple $\mathbf{i} \in \mathbf{I}$ is input and return value of the following function. We split $\mathbf{i} = (i_1, \ldots, i_d)$ into $\mathbf{i}_\alpha := (i_j)_{j\in\alpha} \in \mathbf{I}_\alpha$ and $\mathbf{i}_{\alpha^c} := (i_j)_{j\in\alpha^c} \in \mathbf{I}_{\alpha^c}$. We write $\mathbf{i}(\mathbf{i}_\alpha, \mathbf{i}_{\alpha^c}) \in \mathbf{I}$ for the tuple $\mathbf{i}$ constructed from both $\mathbf{i}_\alpha$ and $\mathbf{i}_{\alpha^c}$ (note that the indices in α, α^c need not be numbered consecutively; e.g., $\alpha = \{1, 4\}$, $\alpha^c = \{2, 3\}$).

```
1 function $ImprovedPivot(\alpha, \mathbf{x}, \mathbf{i})$;                                        (15.21)
2 {input: $\alpha \in T_D$, $\mathbf{x} \in \mathbf{V}$, $\mathbf{i} \in \mathbf{I}$ satisfying $\mathbf{i}_{\alpha^c} \in P_{\alpha^c} \subset \mathbf{I}_{\alpha^c}$}
3 begin for all $j \in \alpha$ do $i_j := \mathrm{argmax}_{i\in I_j} |\mathbf{x}[i_1, \ldots, i_{j-1}, i, i_{j+1}, \ldots i_d]|$;
4     $\mathbf{i}_{\alpha^c} := \mathrm{argmax}_{\mathbf{i}_{\alpha^c}\in P_{\alpha^c}} |\mathbf{x}[\mathbf{i}(\mathbf{i}_\alpha, \mathbf{i}_{\alpha^c})]|$;
5     $ImprovedPivot := \mathbf{i}(\mathbf{i}_\alpha, \mathbf{i}_{\alpha^c})$
6 end;
```

A single step in line 3 maximises $|\mathbf{x}[i_1, \ldots, i_d]|$ on the fibre $\mathcal{F}(j, \mathbf{i})$. This requires the evaluation of the tensor along this fibre. Since j is restricted to α, line 3 defines the part $\mathbf{i}_\alpha \in \mathbf{I}_\alpha$. Having fixed $\mathbf{i}_\alpha$, we need only r_α evaluations for the maximisation over P_{α^c} in line 4. The parts $\mathbf{i}_\alpha, \mathbf{i}_{\alpha^c}$ define the return value in line 5. In total, one call of the function requires

$$r_\alpha + \sum_{j\in\alpha} \dim(V_j)$$

evaluation of $\mathbf{x}$.

The obtained index $\mathbf{i} \in \mathbf{I}$ will be split into $\mathbf{i}_{\alpha_1} := (i_j)_{j\in\alpha_1} \in \mathbf{I}_{\alpha_1}$ and $\mathbf{i}_{\alpha_1^c}$ for a son α_1 of α. Note that $\mathbf{i}_{\alpha_1^c}$ is formed by $\mathbf{i}_{\alpha_2}$ (with entries optimised in line 3) and $\mathbf{i}_{\alpha^c}$ from line 4.

According to Steps 1 and 3a from §15.4.3.1, the matrices S_{α_1} and its inverse T_{α_1} for the sons α_1 of $\alpha \in T_D$ are to be determined. In fact, T_{α_1} is determined explicitly, while S_{α_1} follows implicitly. The following procedure *ImprovedS* performs one iteration step $T_{\alpha_1} \in \mathbb{K}^{(r-1)\times(r-1)} \mapsto T_{\alpha_1} \in \mathbb{K}^{r\times r}$ together with the determination of the index sets $P_{\alpha_1}, P_{\alpha_1^c}$. These data determine the approximation $\mathbf{v}_r \in \mathbf{V}$ satisfying $\mathrm{rank}(\mathcal{M}_{\alpha_1}(\mathbf{v}_r)) = r$. The evaluation of $\mathbf{x} := \mathbf{v} - \mathbf{v}_{r-1}$ in *ImprovedPivot* needs a comment, since $\mathbf{v}_{r-1}$ is not given directly.[8] The inverse matrix $T_{\alpha_1} = S_{\alpha_1}^{-1} \in \mathbb{K}^{(r-1)\times(r-1)}$ is performed via (15.12b) and yields the representation

$$\mathbf{v}_{r-1}[\mathbf{i}_{\alpha_1}, \mathbf{i}_{\alpha_2}, \mathbf{i}_{\alpha^c}] = \sum_{i,j=1}^{r-1} \mathbf{v}[\mathbf{i}_{\alpha_1}, \mathbf{p}_i^{(\alpha_2)}, \mathbf{i}_{\alpha^c}]\, T_{\tau\times\sigma}[\mathbf{p}_i^{(\alpha_2)}, \mathbf{p}_j^{(\alpha_1)}]\, \mathbf{v}[\mathbf{p}_j^{(\alpha_1)}, \mathbf{i}_{\alpha_2}, \mathbf{i}_{\alpha^c}] \tag{15.22}$$

(cf. (15.12b)), where the summation involves $\mathbf{p}_j^{(\alpha_1)} \in P_{\alpha_1}$ and $\mathbf{p}_i^{(\alpha_2)} \in P_{\alpha_2}$.

Changing $\mathbf{i}$, we have to update $\mathbf{v}[\mathbf{i}_{\alpha_1}, \mathbf{p}_i^{(\alpha_2)}, \mathbf{i}_{\alpha^c}]$ and $\mathbf{v}[\mathbf{p}_j^{(\alpha_1)}, \mathbf{i}_{\alpha_2}, \mathbf{i}_{\alpha^c}]$.

```
1 procedure $ImproveS(\alpha, \alpha_1, \mathbf{v}, r, T_{\alpha_1}, P_{\alpha_1}, P_{\alpha_1^c}, P_{\alpha^c})$;
2 {input parameters: $\alpha_1$ son of $\alpha \in T_D$, $\mathbf{v} \in \mathbf{V}$, $P_{\alpha^c} \subset \mathbf{I}_{\alpha^c}$;
3  in- and output: $r \in \mathbb{N}_0$; $T_{\alpha_1} \in \mathbb{K}^{r\times r}$, $P_{\alpha_1} \subset \mathbf{I}_{\alpha_1}$, $P_{\alpha_1^c} \subset \mathbf{I}_{\alpha_1^c}$}
4 begin choose start indices $\mathbf{i}_\alpha \in \mathbf{I}_\alpha$ and $\mathbf{i}_{\alpha^c} \in P_{\alpha^c}$; $\mathbf{i} := \mathbf{i}(\mathbf{i}_\alpha, \mathbf{i}_{\alpha^c})$;
5     for $\chi := 1$ to $\chi_{\max}$ do $\mathbf{i} := ImprovedPivot(\alpha_1, \mathbf{v} - \mathbf{v}_{\ell-1}, \mathbf{i})$;
6     $P_{\alpha_1} := P_{\alpha_1} \cup \mathbf{i}_{\alpha_1}$; $P_{\alpha_1^c} := P_{\alpha_1^c} \cup \mathbf{i}_{\alpha_1^c}$;
7     $r := r + 1$; compute $T_{\alpha_1} \in \mathbb{K}^{r\times r}$ from (15.12a);
8 end;
```

Index $\mathbf{i}$ obtained in line 5 yields the largest value $|(\mathbf{v} - \mathbf{v}_{\ell-1})[\mathbf{i}]|$ among the fibres checked in $ImprovedPivot$. A typical value of $\chi_{\max}$ is 3. In line 6, the parts $\mathbf{i}_{\alpha_1} := (i_j)_{j\in\alpha_1} \in \mathbf{I}_{\alpha_1}$ and $\mathbf{i}_{\alpha_1^c}$ become new pivots in P_{α_1} and $P_{\alpha_1^c}$ (in (15.17) the entries are denoted by $\mathbf{p}_j^{(\alpha)}$ and $\mathbf{p}_j^{(\alpha^c)}$). The update of T_{α_1} in line 7 is the inverse of S_{α_1}.

The final algorithm for determining S_{α_1} (T_{α_1}) depends on the stopping criterion. If r_{α_1} is prescribed, the criterion is $r = r_{\alpha_1}$. If r_{α_1} should be determined adaptively, an error estimation can be applied.

```
procedure $DetS(\alpha, \alpha_1, \mathbf{v}, r, T_{\alpha_1}, P_{\alpha_1}, P_{\alpha_1^c}, P_{\alpha^c})$;
{input parameters: $\alpha_1$ son of $\alpha \in T_D$, $\mathbf{v} \in \mathbf{V}$, $P_{\alpha^c} \subset \mathbf{I}_{\alpha^c}$;
 output: $r \in \mathbb{N}_0$; $T_{\alpha_1} \in \mathbb{K}^{r\times r}$, $P_{\alpha_1} \subset \mathbf{I}_{\alpha_1}$; $P_{\alpha_1^c} \subset \mathbf{I}_{\alpha_1^c}$}
begin $r := 0$; $T_{\alpha_1} := 0$; $P_{\alpha_1} := P_{\alpha_1^c} := \emptyset$;
    repeat $ImproveS(\alpha, \alpha_1, \mathbf{v}, r, T_{\alpha_1}, P_{\alpha_1}, P_{\alpha_1^c}, P_{\alpha^c})$
    until criterion satisfied
end;
```

The call

$$HierAppr(T_D, D, \mathbf{v}, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}, (P_\alpha)_{\alpha\in T_D})$$

of the next procedure determines the parameters $\mathbf{C}_\alpha = (C^{(\alpha,\ell)})_{1\leq\ell\leq r_\alpha}$, $c^{(D)} \in \mathbb{K}$, and $(B_j)_{j\in D} \in (V_j)^{r_j}$ of the hierarchical approximation $\tilde{\mathbf{v}} = \rho_{\mathrm{HTR}}(\ldots) \approx \mathbf{v}$ from (11.28). Note that $\mathbf{v}$ is the given tensor in full functional representation (cf. (7.6)), while $\tilde{\mathbf{v}}$ is the approximation due to the chosen criterion. We recall that $\mathbf{v} = \tilde{\mathbf{v}}$ holds if the criterion prescribes the ranks $r_\alpha = \dim(M_\alpha(\mathbf{v}))$.

```
procedure $HierAppr(T_D, \alpha, \mathbf{v}, (\mathbf{C}_\alpha)_{\alpha\in T_D\backslash\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}, (P_\alpha)_{\alpha\in T_D})$;
{input: $T_D$, $\alpha \in T_D\backslash\mathcal{L}(T_D)$, $\mathbf{v} \in \mathbf{V}$; output: $\mathbf{C}_\alpha, c^{(D)}, B_j, P_\alpha$}
if $\alpha = D$ then
```

[8] Note the difference to the matrix case. In (15.12b) the rows $M[\bullet, j']$ and columns $M[i', \bullet]$ are already evaluated.

```
4  begin $c^{(D)} := 1$; determine sons $\alpha_1, \alpha_2$ of $D$; $C^{(D,1)} := T_{\alpha_1}$;
5     $DetS(\alpha, \alpha_1, \mathbf{v}, r_{\alpha_1}, T_{\alpha_1}, P_{\alpha_1}, P_{\alpha_2}, \emptyset)$;
6     $HierAppr(T_D, \alpha_1, \mathbf{v}, \mathbf{C}_\alpha, c^{(D)}, B_j, P_\alpha)$;
7     $HierAppr(T_D, \alpha_2, \mathbf{v}, \mathbf{C}_\alpha, c^{(D)}, B_j, P_\alpha)$
8  end else if $\alpha \notin \mathcal{L}(T_D)$ then
9  begin determine sons $\alpha_1, \alpha_2$ of $\alpha$;
10    $DetS(\alpha, \alpha_1, \mathbf{v}, r_{\alpha_1}, T_{\alpha_1}, P_{\alpha_1}, P_{\alpha_1^c}, P_{\alpha^c})$;
11    $DetS(\alpha, \alpha_2, \mathbf{v}, r_{\alpha_2}, T_{\alpha_2}, P_{\alpha_2}, P_{\alpha_2^c}, P_{\alpha^c})$;
12    for $\ell := 1$ to $r_\alpha$ do $C^{(\alpha,\ell)} := T_{\alpha_1} V_\ell^{(\alpha)} T_{\alpha_2}^\mathsf{T}$;($V_\ell^{(\alpha)}[\nu,\mu] := \mathbf{v}[\mathbf{p}_\nu^{(\alpha_1)}, \mathbf{p}_\mu^{(\alpha_2)}, \mathbf{p}_\ell^{(\alpha^c)}]$)
13    if $\alpha_1 \notin \mathcal{L}(T_D)$ then $HierAppr(T_D, \alpha_1, \mathbf{v}, \mathbf{C}_\alpha, c^{(D)}, B_j, P_\alpha)$
14       else $B_j := [b_1^{(j)} \cdots b_{r_j}^{(j)}]$ with $\alpha_1 = \{j\}$, $b_i^{(j)} := \mathbf{v}[\bullet, \mathbf{p}_i^{(\alpha_1^c)}]$, $\mathbf{p}_i^{(\alpha_1^c)} \in P_{\alpha_1^c}$;
15    if $\alpha_2 \notin \mathcal{L}(T_D)$ then $HierAppr(T_D, \alpha_2, \mathbf{v}, \mathbf{C}_\alpha, c^{(D)}, B_j, P_\alpha)$
15       else $B_j := [b_1^{(j)} \cdots b_{r_j}^{(j)}]$ with $\alpha_2 = \{j\}$, $b_i^{(j)} := \mathbf{v}[\bullet, \mathbf{p}_i^{(\alpha_2^c)}]$, $\mathbf{p}_i^{(\alpha_2^c)} \in P_{\alpha_2^c}$
17 end;
```

Lines 3-7 correspond to Proposition 15.11a. Lines 9-15 describe the case of Proposition 15.11b. For leaves α_1 or α_2 case (c) of Proposition 15.11 applies.

15.4.3.3 Cost

We have to distinguish the number N_a of arithmetical operations and the number N_e of evaluations of tensor entries. The value

$$N_a = O(dr^4 + N_e) \quad \text{with } r := \max\nolimits_{\alpha \in T_D} r_\alpha$$

is of minor interest (cf. [7]). More important is N_e because evaluations may be rather costly. Following Remark 15.12c, for fixed pivots we need $(d-1)\,r^3 + (2d-1)\,r^2 + drn$ evaluations, where $r := \max_{\alpha \in T_D} r_\alpha$ and $n := \max_j n_j$. Another source of evaluations is the pivot search by $ImprovedPivot(\alpha, \mathbf{x}, \mathbf{i})$ in (15.21), where $\#\alpha$ fibres and further $\#P_{\alpha^c}$ indices $\mathbf{i} = \mathbf{i}(\mathbf{i}_\alpha, \mathbf{i}_{\alpha^c})$ are tested with respect to the size of $|\mathbf{x}[\mathbf{i}]|$. Since $\mathbf{x} = \mathbf{v} - \mathbf{v}_{\ell-1}$, one has to evaluate $\mathbf{v}[\mathbf{i}]$ as well as $\mathbf{v}_{\ell-1}[\mathbf{i}]$. The latter expression is defined in (15.22). If a new index $i_j \in I_j$ belongs to direction $j \in \alpha_1$, the sum in (15.22) involves $r-1$ new values $\mathbf{v}[\mathbf{i}_{\alpha_1}, \mathbf{p}_i^{(\alpha_2)}, \mathbf{i}_{\alpha^c}]$. Similarly for $j \in \alpha_2$. This leads to $2(r-1)\sum_{j \in \alpha} n_j$ evaluations. Variation of $\mathbf{i}_{\alpha^c} \in P_{\alpha^c}$ causes only $2(r-1)\#P_{\alpha^c}$ evaluations. The procedure $ImproveS$ is called for $1 \leq r \leq r_{\alpha_1}$, the total number of evaluations due to the pivot choice is bounded by

$$\sum\nolimits_{\alpha \in T_D \backslash \mathcal{L}(T_D)} r_{\alpha_1}^2 \#\alpha \sum\nolimits_{j \in \alpha} n_j \leq d \cdot depth(T_D) \cdot r^2 \sum\nolimits_{j=1}^{d} n_j,$$

where $r := \max_{\alpha \in T_D} r_\alpha$. We recall that for a balanced tree T_D the depth of T_D equals $\lceil \log_2 d \rceil$, whereas $depth(T_D) = d - 1$ holds for the TT format (cf. Remark 11.5). In the first case, the total number of evaluations is

$$N_e = O\left((d-1)\,r^3 + d\log(d) r^2 n\right).$$

Chapter 16
Applications to Elliptic Partial Differential Equations

Abstract We consider elliptic partial differential equations in d variables and their discretisation in a product grid $\mathbf{I} = \times_{j=1}^{d} I_j$. The solution of the discrete system is a grid function, which can directly be viewed as a tensor in $\mathbf{V} = \bigotimes_{j=1}^{d} \mathbb{K}^{I_j}$. In *Sect. 16.1* we compare the standard strategy of local refinement with the tensor approach involving regular grids. It turns out that the tensor approach can be more efficient. In *Sect. 16.2* the solution of boundary value problems is discussed. A related problem is the eigenvalue problem discussed in *Sect. 16.3*.

We concentrate ourselves to elliptic boundary value problems of second order. However, elliptic boundary value problems of higher order or parabolic problems lead to similar results.

16.1 General Discretisation Strategy

The discretisation of partial differential equations leads to matrices whose size grows with increasing accuracy requirement. In general, simple discretisation techniques (Galerkin method or finite difference methods) using uniform grids yield too large matrices. Instead, adaptive discretisation techniques are used. Their aim is to use as few unknowns as possible in order to ensure a certain accuracy.

The first type of methods is characterised by the relation $\varepsilon = O(n^{\kappa/d})$, where ε is the accuracy of the approximation (in some norm), n number of degrees of freedom, κ the consistency order, and d the spatial dimension. The corresponding methods are Galerkin discretisations with polynomial ansatz functions of fixed degree. The relation $\varepsilon = O(n^{\kappa/d})$ is not always reached.[1] Only if the solution behaves uniformly regular in its domain of definition, also the uniform grid yields

$$\varepsilon = O(n^{\kappa/d}). \tag{16.1}$$

[1] For three-dimensional problems, edge singularities require stretched tetraeders. However, usual adaptive refinement strategies try to ensure form regularity (cf. [82]).

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_16,

In the standard case, however, the solution of partial differential equations has point singularities and—for 3D problems—edge singularities. This requires a concentration of grid points towards corners or edges.

A more efficient method is the hp finite element method in the case of piecewise analytic solutions. Here, the ideal relation between the error ε and the number n of unknowns is $\varepsilon = O(\exp(-\beta n^{\alpha}))$ for suitable $\alpha, \beta > 0$.

So far, for fixed ε, the strategy is to minimise the problem size n. In principle, this requires that also the generation of the system matrix and its solution is $O(n)$. This requirement can be relaxed for the hp finite element method. If $n = O(\log^{1/\alpha} \frac{1}{\varepsilon})$, even Gauss elimination with cost $O(\log^{3/\alpha} \frac{1}{\varepsilon})$ is only a redefinition of α by $\alpha/3$.

Tensor applications require a Cartesian grid $\mathbf{I} := I_1 \times \ldots \times I_d$ of unknowns. This does not mean that the underlying domain $\Omega \subset \mathbb{R}^d$ must be of product form. It is sufficient that Ω is the image of a domain $\Omega_1 \times \ldots \times \Omega_d$. For instance, Ω may be a circle, which is the image of the polar coordinates varying in $\Omega_1 \times \Omega_2$.

The use of a (uniform[2]) grid $\mathbf{I} = I_1 \times \ldots \times I_d$ seems to contradict the strategies from above. However, again the leading concept is: best accuracy for minimal cost, where the accuracy[3] is fixed by the grid $\mathbf{I}$. For simplicity, we assume $n_j = \#I_j = n = O(\varepsilon^{-\beta})$. The storage cost of the tensor formats is $O(r^* nd)$, where r^* indicates possible powers of some rank parameters. The approximation of the inverse by the technique from §9.7.2.6 costs $O(\log^2(\frac{1}{\varepsilon}) \cdot r^* nd)$. Comparing the storage and arithmetical cost with the accuracy, we see a relation like in (16.1), but the exponent κ/d is replaced by some $\beta > 0$ independent of d.

A second step is the tensorisation from §14. As described in §14.2.3, the complexity reached by tensorisation corresponds (at least) to the hp finite element approach. Therefore, in the end, the cost should be not worse than the best hp method, but independent of d.

16.2 Solution of Elliptic Boundary Value Problems

We consider a linear boundary value problem

$$Lu = f \text{ in } \Omega = \Omega_1 \times \ldots \times \Omega_d \subset \mathbb{R}^d, \qquad u = 0 \text{ on } \partial\Omega \tag{16.2}$$

with a linear differential operator of elliptic type (cf. [82, §5.1.1]). Concerning the product form see the discussion from above. The homogeneous Dirichlet condition $u = 0$ on $\partial\Omega$ may be replaced by other conditions like the Neumann condition.

The standard dimension $d = 3$ is already of interest for tensor methods. The other extreme are dimensions of the order $d = 1000$.

[2] The grids I_j need not be uniform, but for simplicity this is assumed.

[3] In the case of a point singularity r^{α} ($\alpha > 0$, $r = \|x - x_0\|$, x_0: corner point), the grid size h leading to an accuracy ε is of the form $h = O(\varepsilon^{\beta})$, $\beta > 0$. A uniform grid I_j needs $n_j = O(\varepsilon^{-\beta})$ grid points.

16.2.1 Separable Differential Operator

The most convenient form of L is the separable one (cf. (1.10a); Definition 9.36):

$$L = \sum_{j=1}^{d} L_j, \qquad L_j \text{ differential operator in } x_j, \tag{16.3a}$$

i.e., L_j contains derivatives with respect to x_j only and its coefficients depend only on x_j (in particular, L_j may have constant coefficients). In this case, we can write the differential operator as Kronecker product:

$$L = \sum_{j=1}^{d} I \otimes \ldots \otimes I \otimes L_j \otimes I \otimes \ldots \otimes I, \tag{16.3b}$$

where L_j is considered as one-dimensional differential operator acting on a suitable space V_j.

16.2.2 Discretisation

16.2.2.1 Finite Difference Method

Choose a uniform[4] grid with n_j (interior) grid points in direction j. The one-dimensional differential operator L_j from (16.3b) can be approximated by a difference operator Λ_j (see [82, §4.1] for details). The standard choice of finite differences leads to *tridiagonal* matrices Λ_j. Higher-order differences may produce more off-diagonals. The resulting system matrix of the difference method takes the form

$$\mathbf{A} = \sum_{j=1}^{d} I \otimes \ldots \otimes I \otimes \Lambda_j \otimes I \otimes \ldots \otimes I, \tag{16.4}$$

provided that (16.3b) holds.

16.2.2.2 Finite Element Method

The variational formulation of (16.2) is given by a $\left\{\begin{matrix} \text{bilinear} & \text{if } \mathbb{K} = \mathbb{R} \\ \text{sesquilinear} & \text{if } \mathbb{K} = \mathbb{C} \end{matrix}\right\}$ form $a(\cdot,\cdot)$ and the functional $f(v) = \int_\Omega f v \mathrm{d}x$:

$$\text{find } u \in H_0^1(\Omega) \text{ such that } a(u,v) = f(v) \text{ for all } v \in H_0^1(\Omega). \tag{16.5}$$

[4] Also non-uniform grids may be used. In this case, the difference formulae are Newton's first and second difference quotients (cf. [82, §4.1]).

According to the splitting (16.3a), the form $a(\cdot,\cdot)$ is a sum of products:

$$a(\cdot,\cdot) = \sum_{j=1}^{d} \left[a_j(\cdot,\cdot) \prod_{k\neq j} (\cdot,\cdot)_k \right], \tag{16.6}$$

where $a_j : H_0^1(\Omega_j) \times H_0^1(\Omega_j) \to \mathbb{K}$ and $(\cdot,\cdot)_k$ is the $L^2(\Omega_k)$ scalar product.

The (possibly non-uniform) intervals $[x_\nu^{(j)}, x_{\nu+1}^{(j)}]$, $0 \le \nu \le n_j$, of the one-dimensional grids in directions $1 \le j \le d$ form the cuboids $\tau_{\boldsymbol{\nu}} := \times_{j=1}^{d} [x_{\nu_j}^{(j)}, x_{\nu_j+1}^{(j)}]$ for multi-indices $\boldsymbol{\nu} \in \times_{j=1}^{d} \{0, \ldots, n_j\}$. Let $b_\nu^{(j)} \in H_0^1(\Omega_j)$ for $1 \le \nu \le n_j$ be the standard, one-dimensional, piecewise linear hat function: $b_\nu^{(j)}(x_\mu^{(j)}) = \delta_{\nu\mu}$. They span the subspace $V_j \subset H_0^1(\Omega_j)$. The final finite elements basis functions are

$$\mathbf{b}_{\boldsymbol{\nu}} := \bigotimes_{j=1}^{d} b_{\nu_j}^{(j)} \in H_0^1(\Omega) \quad \text{for } \boldsymbol{\nu} \in \mathbf{I} := \bigtimes_{j=1}^{d} I_j,\ I_j := \{1, \ldots, n_j\}.$$

Their span is the space $\mathbf{V} := \bigotimes_{j=1}^{d} V_j \subset H_0^1(\Omega)$. The finite element solution $\mathbf{u} \in \mathbf{V}$ is defined by

$$a(\mathbf{u}, \mathbf{v}) = f(\mathbf{v}) \text{ for all } \mathbf{v} \in \mathbf{V}. \tag{16.7}$$

The solution has a representation $\mathbf{u} = \sum \mathbf{x}_{\boldsymbol{\mu}} \mathbf{b}_{\boldsymbol{\mu}}$, where the coefficient vector $\mathbf{x} = (\mathbf{x}_{\boldsymbol{\mu}})$ is the solution of the linear system $\mathbf{A}\mathbf{x} = \boldsymbol{\phi}$. Here, the right-hand side $\boldsymbol{\phi}$ has the entries $\boldsymbol{\phi}_{\boldsymbol{\nu}} = f(\mathbf{b}_{\boldsymbol{\nu}})$. The finite element system matrix $\mathbf{A}$ is defined by

$$\mathbf{A}_{\boldsymbol{\nu}\boldsymbol{\mu}} := a(\mathbf{b}_{\boldsymbol{\mu}}, \mathbf{b}_{\boldsymbol{\nu}}).$$

From (16.6) one derives that

$$\mathbf{A} = \sum_{j=1}^{d} \left(\bigotimes_{k=1}^{j-1} M_k \right) \otimes A_j \otimes \left(\bigotimes_{k=j+1}^{d} M_k \right), \text{ where} \tag{16.8}$$
$$A_j[\nu,\mu] := a_j(b_\mu^{(j)}, b_\nu^{(j)}) \text{ and } M_j[\nu,\mu] := (b_\mu^{(j)}, b_\nu^{(j)})_j.$$

Note that the mass matrix M_k replaces the identity matrix in (16.4).

Remark 16.1. Let $\mathbf{M} := \bigotimes_{j=1}^{d} M_j$, and define $\boldsymbol{\Lambda} := \mathbf{M}^{-1}\mathbf{A}$. Then $\boldsymbol{\Lambda}$ takes the form (16.4) with $\Lambda_j := M_j^{-1} A_j$.

16.2.2.3 Treatment of Non-separable Differential Operators

The assumption of a separable L excludes not only mixed derivatives $\frac{\partial^2}{\partial x_i \partial x_j}$, but also coefficients depending on other x-components than x_j. As an example, we consider the first order term $L_{\text{first}} := c\nabla = \sum_{j=1}^{d} c_j(x) \frac{\partial}{\partial x_j}$ appearing in L with coefficients $c_j(x_1, \ldots, x_d)$ and discuss the definition of a tensor-based finite difference scheme.

The forward difference ∂^+ (defined by $(\partial^+\varphi)(\xi) = [\varphi(\xi + h) - \varphi(\xi)]/h$) or the backward difference ∂^- can be represented exactly in the format $\mathcal{H}_{\boldsymbol{\rho}}^{\mathrm{tens}}$ with ranks $\rho_k = 2$ (cf. §14.1.6). The central difference requires $\rho_k = 3$. Next, we apply the technique of §15.4 to construct a tensor $\mathbf{c}_j \in \mathbf{V}$ approximating the d-variate function c_j. According to Remark 13.10, we may define the multiplication operator $\mathbf{C}_j \in L(\mathbf{V}, \mathbf{V})$. Hence, the discretisation of $L_{\text{first}} = c\nabla$ is given by

$$\mathit{\Lambda}_{\text{first}} := \sum_{j=1}^{d} \mathbf{C}_j \partial_j^+.$$

Usually, it is not necessary to determine the operator $\mathit{\Lambda}_{\text{first}}$ explicitly, but in principle this can be done (cf. §13.8). Analogously, other parts of L can be treated.

16.2.3 Solution of the Linear System

In the following, we discuss the use of iterative schemes (details in Hackbusch [81, §3]). An alternative approach is mentioned in §17.2.1. Let $\mathbf{Ax} = \mathbf{b}$ be the linear system with $\mathbf{x}, \mathbf{b} \in \mathbf{V}$. The basic form of a linear iteration is

$$\mathbf{x}^{(m+1)} := \mathbf{x}^{(m)} - \mathbf{C}\left(\mathbf{Ax}^{(m)} - \mathbf{b}\right) \tag{16.9}$$

with some matrix $\mathbf{C}$ and any starting value $\mathbf{x}^{(0)}$. Then convergence $\mathbf{x}^{(m)} \to \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ holds if and only if the spectral radius (4.76) of $\mathbf{CA}$ satisfies $\rho(\mathbf{CA}) < 1$. For the efficient solution one needs $\rho(\mathbf{CA}) \leq \eta < 1$ with η independent of the grid size and of possible parameters appearing in the problem.

The (slow) convergence of non-efficient methods is often directly connected to the condition of the matrix $\mathbf{A}$. Assume for simplicity that $\mathbf{A} \in L(\mathbf{V}, \mathbf{V})$ for $\mathbf{V} = \bigotimes_{j=1}^{d} \mathbb{K}^{n_j}$ is of the form (16.4) with positive definite matrices $\mathit{\Lambda}_j$ possessing eigenvalues $\lambda_1^{(j)} \geq \ldots \geq \lambda_{n_j}^{(j)} > 0$. Then the condition of $\mathbf{A}$ equals

$$\operatorname{cond}(\mathbf{A}) = \sum_{j=1}^{d} \lambda_1^{(j)} \Big/ \sum_{j=1}^{d} \lambda_{n_j}^{(j)}.$$

In the simplest case of $\mathit{\Lambda}_1 = \mathit{\Lambda}_2 = \ldots = \mathit{\Lambda}_d$, the eigenvalues $\lambda_\nu^{(j)} = \lambda_\nu$ are independent and $\operatorname{cond}(\mathbf{A}) = \lambda_1/\lambda_n$ holds. This together with Exercise 4.57 proves the next remark.

Remark 16.2. The condition of the matrix $\mathbf{A}$ depends on the numbers $n_j = \#I_j$, but not on the dimension d. In particular, the following inequalities hold:

$$\min_j \operatorname{cond}(\mathit{\Lambda}_j) \leq \operatorname{cond}(\mathbf{A}) \leq \max_j \operatorname{cond}(\mathit{\Lambda}_j).$$

If $\mathbf{A}$ and $\mathbf{C} = \mathbf{B}^{-1}$ are positive definite, $\mathbf{C}$ is a suitable choice[5] if $\mathbf{A}$ and $\mathbf{B}$ are *spectrally equivalent*, i.e., $\frac{1}{c_1}(\mathbf{Ax}, \mathbf{x}) \le (\mathbf{Bx}, \mathbf{x}) \le c_2 (\mathbf{Ax}, \mathbf{x})$ for all $\mathbf{x} \in \mathbf{V}$ and constants $c_1, c_2 < \infty$. In the case of (not singularly degenerate) elliptic boundary value problems, $(\mathbf{A}\cdot, \cdot)$ corresponds to the H^1 norm. As a consequence, different system matrices $\mathbf{A}$ and $\mathbf{B}$ corresponding to H^1 coercive elliptic problems are spectrally equivalent. In particular, there are elliptic problems with separable $\mathbf{B}$ (e.g., the Laplace equation $-\Delta u = f$).

Given a positive definite and separable differential operator, its discretisation $\mathbf{B}$ satisfies the conditions of Proposition 9.34 (therein, $\mathbf{B}$ is called $\mathbf{A}$, while $\mathbf{C} = \mathbf{B}^{-1}$ is called $\mathbf{B}$). The matrices $M^{(j)}$ in Proposition 9.34 are either the identity (finite difference case) or the mass matrices (finite element case). As a result, a very accurate approximation of $\mathbf{C} = \mathbf{B}^{-1}$ can be represented in the format $\mathcal{R}_r$ (transfer into other formats is easy). We remark that the solution requires the matrix exponentials $\exp(-\alpha T)$ or $\exp(-\alpha M^{-1}T)$ (T: $n \times n$ triangular matrix). In the case of $\exp(-\alpha T)$, this can be performed exactly by means of a diagonalisation of T. In general, the technique of hierarchical matrices yields $\exp(-\alpha M^{-1}T)$ in almost linear cost $O(n \log^* n)$ (cf. Hackbusch [86, §13.2.2]).

Once, a so-called preconditioner $\mathbf{C}$ is found, we have to apply either iteration (16.9) or an accelerated version using conjugate gradients or GMRES. For simplicity, we assume $\rho(\mathbf{CA}) \le \eta < 1$ and apply (16.9). The representation ranks of $\mathbf{x}^{(m)}$ are increased first by the evaluation of the defect $\mathbf{Ax}^{(m)} - \mathbf{b}$ and second by multiplication by $\mathbf{C}$. Therefore, we have to apply the truncated iteration from §13.10:

$$\mathbf{x}^{(m+1)} := T\left[\mathbf{x}^{(m)} - \mathbf{C}\big(\mathbf{Ax}^{(m)} - \mathbf{b}\big)\right], \quad \text{or}$$

$$\mathbf{x}^{(m+1)} := T\left[\mathbf{x}^{(m)} - \mathbf{C}\big(T\big[\mathbf{Ax}^{(m)} - \mathbf{b}\big]\big)\right],$$

where T denotes a suitable truncation. See also Khoromskij [117].

If a very efficient $\mathbf{C}$ is required (i.e., $0 < \eta \ll 1$), $\mathbf{C}$ must be close to $\mathbf{A}^{-1}$. The fixed point iterations explained in §13.10 can be used to produce $\mathbf{C} \approx \mathbf{A}^{-1}$.

Another approach is proposed in Ballani-Grasedyck [6], where a projection method onto a subspace is used, which is created in a Krylov-like manner.

A well-known efficient iterative method is the *multi-grid iteration* (cf. [83]). For its implementation one needs a sequence of grids with decreasing grid size, prolongations and restrictions, and a so-called smoother. Since we consider uniform grids, the construction of a sequence of grids with grid width $h_\ell = 2^{-\ell}h_0$ (ℓ: level of the grid) is easy. The prolongations and the restrictions are elementary Kronecker tensors. For the solution on the coarsest grid (level $\ell = 0$) one of the aforementioned methods can be applied. As smoothing iteration one may choose the damped Jacobi iteration. The numerical examples in Ballani-Grasedyck [6] confirm a grid-independent convergence rate.

An alternative to the iteration (16.9) is the direct minimisation approach from §17.2.1.

[5] To be precise, $\mathbf{C}$ must be suitably scaled to obtain $\rho(\mathbf{CA}) \le \eta < 1$.

16.3 Solution of Elliptic Eigenvalue Problems

As already stated in the introduction (see page 11), an eigenvalue problem

$$Lu = \lambda u \text{ in } \Omega = \Omega_1 \times \ldots \times \Omega_d, \quad u = 0 \text{ on } \partial\Omega, \tag{16.10}$$

for a separable differential operator L (cf. (16.3a)) is trivial, since the eigenvectors are elementary tensors: $u \in \mathcal{R}_1$. The determination of u can be completely reduced to one-dimensional eigenvalue problems

$$L_j u^{(j)} = \mu u^{(j)}, \quad u^{(j)} \in V_j \backslash \{0\}.$$

In the following, we consider a linear, symmetric eigenvalue problem (16.10) discretised by[6]

$$Ax = \lambda x \tag{16.11a}$$

involving a symmetric matrix A. Regarding $x \in \mathbb{K}^N$ as tensor $\mathbf{x} \in \mathbf{V}$, we interpret A as a Kronecker product $\mathbf{A} \in L(\mathbf{V}, \mathbf{V})$. In general, due to truncations, $\mathbf{A}$ and $\mathbf{x}$ will be only approximations of the true problem

$$\mathbf{Ax} = \lambda \mathbf{x}. \tag{16.11b}$$

According to Lemma 13.11, we can ensure Hermitean symmetry $\mathbf{A} = \mathbf{A}^{\mathsf{H}}$ exactly.

16.3.1 Regularity of Eigensolutions

Since separable differential operators lead to rank-1 eigenvectors, one may hope that, in the general case, the eigenvector is well approximated in one of the formats. This property can be proved, e.g., under the assumption that the coefficients of L are analytic. The details can be found in Hackbusch-Khoromskij-Sauter-Tyrtyshnikov [91]. Besides the usual ellipticity conditions, all coefficients appearing in L are assumed to fulfil

$$\|\nabla^p c\|_{L^\infty(\Omega)} := \left\| \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d \text{ with } |\boldsymbol{\nu}|=p} \frac{p!}{\boldsymbol{\nu}!} \left| \left(\frac{\partial}{\partial x} \right)^{\boldsymbol{\nu}} u \right|^2 \right\|_{L^\infty(\Omega)}^{1/2} \le C_c \gamma^p p! \tag{16.12a}$$

for all $p \in \mathbb{N}_0$ and some $C_c, \gamma > 0$. Then for analytic Ω or for $\Omega = \mathbb{R}^d$, the eigensolutions u corresponding to an eigenvalue λ satisfy

$$\begin{aligned} \left\|\nabla^{p+2} u\right\|_{L^2(\Omega)} &:= \sqrt{\sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d \text{ with } |\boldsymbol{\nu}|=p+2} \frac{(p+2)!}{\boldsymbol{\nu}!} \left\| \left(\frac{\partial}{\partial x} \right)^{\boldsymbol{\nu}} u \right\|_{L^2(\Omega)}^2} \\ &\le C K^{p+2} \max\left\{p, \sqrt{|\lambda|}\right\} \quad \text{for all } p \in \mathbb{N}_0, \end{aligned} \tag{16.12b}$$

[6] A Galerkin discretisation leads to a generalised eigenvalue $Ax = \lambda M x$ with the mass matrix M.

where C and K depend only on the constants in (16.12a) and on Ω (cf. [91, Theorem 5.5]).

Because of these smoothness results, one obtains error bounds for polynomial interpolants. As shown in [91, Theorem 5.8], this implies that a polynomial $u_{\mathbf{r}} \in \mathcal{P}_{\mathbf{r}} \subset \mathcal{T}_{\mathbf{r}}$ exists with $\mathbf{r} = (r, \ldots, r)$ and

$$\begin{aligned} &\|u - u_{\mathbf{r}}\|_{H^1} \le CMr \log^d(r)\rho^{-r}, \quad \text{where} \\ &\rho := 1 + \frac{\hat{C}_d}{1 + \sqrt{|\lambda|}}, \quad M := \frac{C\tilde{C}_d}{\sqrt{2\pi}} \left(K(p + \sqrt{|\lambda|})\right)^{\lceil (d+1)/2 \rceil} \end{aligned}$$

with $\hat{C}_d, \tilde{C}_d > 0$ depending only on C, K from (16.12b) (cf. [91, Theorem 5.8]).

The latter approximation carries over to the finite element solution (cf. [91, Theorem 5.12]). This proves that, under the assumptions made above, the representation rank $\mathbf{r}$ depends logarithmically on the required accuracy. However, numerical tests with non-smooth coefficients show that even then good tensor approximations can be obtained (cf. [91, §6.2]).

The most challenging eigenvalue problem is the Schrödinger equation (13.32). This is a linear eigenvalue problem, but the requirement of an antisymmetric eigenfunction (Pauli principle) is not easily compatible with the tensor formats (see, e.g., Mohlenkamp et al. [16, 148, 150]). The alternative density functional theory (DFT) approach—a nonlinear eigenvalue problem—and its treatment are already explained in §13.11. Again the question arises whether the solution can be well approximated within one of the tensor formats. In fact, the classical approximation in quantum chemistry uses Gaussians[7]

$$\exp\{-\alpha_\nu \|\bullet - \mathbf{x}_\nu\|^2\} \qquad (\alpha_\nu > 0, \mathbf{x}_\nu \in \mathbb{R}^3 \text{ position of nuclei})$$

multiplied by suitable polynomials. Since the Gaussian function times a monomials is an elementary tensor in $\otimes^3 C(\mathbb{R})$, all classical approximations belong to format $\mathcal{R}_r$, more specifically, to the subset of $\mathcal{R}_r$ spanned by r Gaussians modulated by a monomial. This shows that methods from tensor calculus can yield results which are satisfactory for quantum chemistry purpose. Nevertheless, the number r of terms is large and increases with molecule size. For instance, for C_2H_5OH the number r is about 7000 (cf. [63]). Although Gaussian functions are suited for approximation, they are not the optimal choice. Alternatively, one can choose a sufficiently fine grid in some box $[-A, A]^3$ and search for approximations in $\mathcal{R}_r \subset \mathbf{V} = \otimes^3 \mathbb{R}^n$ corresponding to a grid width $2A/n$ (see concept in [60]). Such a test is performed in Chinnamsetty et al. [34, 35] for the electron density $n(\mathbf{y}) = \rho(\mathbf{y}, \mathbf{y})$ (see page 415) and shows a reduction of the representation rank by a large factor. A similar comparison can be found in Flad et al. [63]. See also Chinnamsetty et al. [36].

Theoretical considerations about the approximability of the wave function are subject of Flad-Hackbusch-Schneider [61, 62].

[7] There are two reasons for this choice. First, Gaussians approximate the solution quite well (cf. Kutzelnigg [135] and Braess [26]). Second, operations as mentioned in §13.11 can be performed analytically. The historical paper is Boys [23] from 1950.

16.3.2 Iterative Computation

Here, we follow the approach of Kressner-Tobler [130], which is based on the algorithm[8] of Knyazev [125] for computing the smallest eigenvalue of an eigenvalue problem (16.11a) with positive definite matrix A and suitable preconditioner B:

```
1  procedure LOBPCG($A, B, x, \lambda$);                                  (16.13)
2  {input: $A, B, x$ with $\|x\| = 1$; output: eigenpair $x, \lambda$}
3  begin $\lambda := \langle Ax, x\rangle$ ; $p := 0$;
4      repeat $r := B^{-1}(Ax - \lambda x);\ U := [x, r, p] \in \mathbb{C}^{N\times 3}$;
5          $\hat{A} := U^{\mathsf{H}} AU;\ \hat{M} := U^{\mathsf{H}} U$;
6          determine eigenpair $\left(y \in \mathbb{C}^3, \lambda\right)$ of $\hat{A}y = \lambda\hat{M}y$ with smallest $\lambda$;
7          $p := y_2 \cdot r + y_3 \cdot p;\ x := y_1 \cdot x + p;\ x := x/\left\|x\right\|$
8      until suitable stopping criterion satisfied
9  end;
```

$\|x\|^2 = \langle x, x\rangle$ is the squared Euclidean norm. The input of B is to be understood as a (preconditioning) method performing $\xi \mapsto B^{-1}\xi$ for $\xi \in \mathbb{C}^N$. The input value $x \in \mathbb{C}^N$ is the starting value. The desired eigenpair (x, λ) of (16.11a) with minimal λ is the output of the procedure. In line 5, $\hat{A}$ and $\hat{M}$ are positive semidefinite 3×3 matrices; hence, the computation of the eigenpair in line 6 is very cheap.

Next, we consider the tensor formulation (16.11b) of the eigenvalue problem. Then, procedure (16.13) becomes

```
1   procedure T-LOBPCG($\mathbf{A}, \mathbf{B}, \mathbf{x}, \lambda$);                       (16.14)
2   {input: $\mathbf{A}, \mathbf{B} \in L(\mathbf{V}, \mathbf{V}),\ \mathbf{x} \in \mathbf{V}$ with $\|\mathbf{x}\| = 1$; output: eigenpair $\mathbf{x}, \lambda$}
3   begin $\lambda := \langle \mathbf{Ax}, \mathbf{x}\rangle$ ; $\mathbf{p} := 0 \in \mathbf{V}$;
4       repeat $\mathbf{r} := T(\mathbf{B}^{-1}(\mathbf{Ax} - \lambda\mathbf{x}));\ \mathbf{u}_1 := \mathbf{x};\ \mathbf{u}_2 := \mathbf{r};\ \mathbf{u}_3 := \mathbf{p}$;
5a          for $i := 1$ to 3 do for $j := 1$ to $i$ do
5b          begin $\hat{A}_{ij} := \overline{\hat{A}_{ji}} := \langle \mathbf{Au}_j, \mathbf{u}_i\rangle$ ; $\hat{M}_{ij} := \overline{\hat{M}_{ji}} := \langle \mathbf{u}_j, \mathbf{u}_i\rangle$ end;
6           determine eigenpair $\left(y \in \mathbb{C}^3, \lambda\right)$ of $\hat{A}y = \lambda\hat{M}y$ with smallest $\lambda$;
7           $\mathbf{p} := T(y_2 \cdot \mathbf{r} + y_3 \cdot \mathbf{p});\ \mathbf{x} := T(y_1 \cdot \mathbf{x} + \mathbf{p});\ \mathbf{x} := \mathbf{x}/\left\|\mathbf{x}\right\|$
8       until suitable stopping criterion satisfied
9   end;
```

This procedure differs from the (16.13) in lines 4 and 7, where a truncation T to a suitable format is performed. The required tensor operations are (i) the matrix-vector multiplication $\mathbf{Au}_j$ for $1 \leq j \leq 3$ (note that $\mathbf{u}_1 = \mathbf{x}$), (ii) the scalar product in lines 3 and 5b, (iii) additions and scalar multiplications in lines 4 and 7, and (iv) the performance of $\mathbf{B}^{-1}$ in line 4. Here, we can apply the techniques from §16.2.3.

As pointed out in [130], the scalar products $\langle \mathbf{Au}_j, \mathbf{u}_i\rangle$ are to be computed exactly, i.e., no truncation is applied to $\mathbf{Au}_j$.

[8] LOBPCG means 'locally optimal block preconditioned conjugate gradient'. For simplicity, we consider only one eigenpair, i.e., the block is of size 1×1.

Algorithm (16.14) can be combined with any of the tensor formats. The numerical examples in [130] are based on the hierarchical format.

Ballani-Grasedyck [6] compute the minimal (or other) eigenvalues by means of the (shifted) inverse iteration. Here the arising linear problems are solved by the multi-grid iteration described in §16.2.3. Numerical examples can be found in [6].

16.3.3 Alternative Approaches

In the case of a positive definite matrix A, the minimal eigenvalue of $Ax = \lambda x$ is the minimum of the Rayleigh quotient:

$$\min\left\{\frac{\langle Ax, x\rangle}{\langle x, x\rangle} : x \neq 0\right\}.$$

This allows us to use the minimisation methods from §17.2. Also this approach is discussed in [130, §4].

Another approach is proposed by [67, §4], where the spectrum is recovered from the time-dependent solution $x(t)$ of $\dot{x}(t) = \mathrm{i}Ax(t)$.

16.4 On Other Types of PDEs

So far, we have only discussed partial differential operators of elliptic type. Nevertheless, the tensor calculus can also be applied to partial differential equations of other type. Section 17.3 is concerned with time-dependent problems. In the case of $\dot{\mathbf{v}}(t) = \mathbf{A}\mathbf{v}(t) + \mathbf{f}(t)$ with an elliptic operator $\mathbf{A}$, we obtain a *parabolic* partial differential equation. On the other hand, the wave equation is a prototype of a *hyperbolic* differential equation. For its solution by means of the retarded potential, tensor methods are used in Khoromskij-Sauter-Veit [123].

Chapter 17
Miscellaneous Topics

Abstract In this chapter we mention further techniques which are of interest for tensor calculations. The first two sections consider optimisation problems. *Section 17.1* describes iterative minimisation methods on a theoretical level of topological tensor spaces assuming exact tensor arithmetic. On the other hand, *Sect. 17.2* applies optimisation directly to the parameters of the tensor representation. *Section 17.3* is devoted to ordinary differential equations for tensor-valued functions. Here, the tangent space and the Dirac-Frenkel discretisation are explained. Finally, *Sect. 17.4* recalls the ANOVA decomposition ('analysis of variance').

17.1 Minimisation Problems on $\mathbf{V}$

In this section we follow the article of Falcó-Nouy [58]. We are looking for a minimiser $\mathbf{u} \in \mathbf{V}$ satisfying

$$J(\mathbf{u}) = \min_{\mathbf{v}\in\mathbf{V}} J(\mathbf{v}) \tag{17.1}$$

under certain conditions on $\mathbf{V}$ and J. For its solution a class of iterations producing a sequence $(\mathbf{u}_m)_{m\in\mathbb{N}}$ is described and convergence $\mathbf{u}_m \to \mathbf{u}$ is proved.

17.1.1 Algorithm

The usual approach is to replace the global minimisation by a finite or infinite sequence of simpler minimisations. For this purpose we define a set $\mathcal{S}$ for which we shall give examples below. Then the simplest version of the algorithm (called 'purely progressive PGD' in [58], where PGD means 'proper generalised decomposition') is the following iteration starting from some $\mathbf{u}_0 := 0$:

$$\mathbf{u}_m := \mathbf{u}_{m-1} + \mathbf{z}_m, \text{ where } J(\mathbf{u}_{m-1} + \mathbf{z}_m) = \min_{\mathbf{z}\in\mathcal{S}} J(\mathbf{u}_{m-1} + \mathbf{z}). \tag{17.2}$$

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6_17,

The set $\mathcal{S}$ must be rich enough. The precise conditions are:

$$\begin{array}{l} 0 \in \mathcal{S} \subset \mathbf{V}, \quad \mathcal{S} = \lambda\mathcal{S} := \{\lambda \mathbf{s} : \mathbf{s} \in \mathcal{S}\} \quad \text{for all } \lambda \in \mathbb{K}, \\ \operatorname{span}(\mathcal{S}) \text{ is dense in } \mathbf{V}, \quad \mathcal{S} \text{ is weakly closed.} \end{array} \tag{17.3}$$

Example 17.1. Sets $\mathcal{S}$ satisfying (17.3) are (a) $\mathcal{S} = \mathcal{R}_1$ (set of elementary tensors), (b) $\mathcal{S} = \mathcal{T}_\mathbf{r}$ with $\mathbf{r} \geq (1, \ldots, 1)$.

Proof. The first two conditions are obvious. By definition, $\operatorname{span}(\mathcal{R}_1) = \mathbf{V}_{\text{alg}}$ is dense in $\mathbf{V}$. As $\mathcal{R}_1 \subset \mathcal{T}_\mathbf{r}$, the same holds for $\mathcal{T}_\mathbf{r}$. By Lemma 8.6, $\mathcal{T}_\mathbf{r}$ and in particular $\mathcal{R}_1 = \mathcal{T}_{(1,\ldots,1)}$ are weakly closed. □

Iteration (17.2) can be improved by certain updates of the correction. Falcó-Nouy [58] propose two variants. In the following, $\mathbf{v} \mapsto \mathbf{U}(\mathbf{v})$ are mappings from $\mathbf{V}$ into the set of closed subspaces with the property $\mathbf{v} \in \mathbf{U}(\mathbf{v})$.

Update A Replace the iteration step (17.2) by

$$\begin{array}{ll} \hat{\mathbf{z}} \in \mathcal{S} & \text{with } J(\mathbf{u}_{m-1} + \hat{\mathbf{z}}) = \min_{\mathbf{z} \in \mathcal{S}} J(\mathbf{u}_{m-1} + \mathbf{z}), \\ \mathbf{u}_m \in \mathbf{U}(\mathbf{u}_{m-1} + \hat{\mathbf{z}}) & \text{with } J(\mathbf{u}_m) = \min_{\mathbf{v} \in \mathbf{U}(\mathbf{u}_{m-1} + \hat{\mathbf{z}})} J(\mathbf{v}). \end{array} \tag{17.4}$$

While $\min_{\mathbf{z} \in \mathcal{S}}$, in general, involves a non-convex set, minimisation over $\mathbf{U}(\mathbf{u}_{m-1} + \hat{\mathbf{z}})$ is of simpler kind. Next, $\mathbf{U}(\mathbf{u}_{m-1} + \hat{\mathbf{z}})$ is replaced by an affine subspace. For instance, $\mathbf{U}(\mathbf{v})$ may be chosen as $\mathbf{U}^{\min}(\mathbf{v}) := \bigotimes_{j=1}^d U_j^{\min}(\mathbf{v})$.

Update B Replace the iteration step (17.2) by

$$\begin{array}{ll} \hat{\mathbf{z}} \in \mathcal{S} & \text{with } J(\mathbf{u}_{m-1} + \hat{\mathbf{z}}) = \min_{\mathbf{z} \in \mathcal{S}} J(\mathbf{u}_{m-1} + \mathbf{z}), \\ \mathbf{u}_m \in \mathbf{u}_{m-1} + \mathbf{U}(\hat{\mathbf{z}}) & \text{with } J(\mathbf{u}_m) = \min_{\mathbf{v} \in \mathbf{U}(\hat{\mathbf{z}})} J(\mathbf{u}_{m-1} + \mathbf{v}). \end{array} \tag{17.5}$$

Concerning examples of $\mathbf{U}(\hat{\mathbf{z}})$ see [58, Example 4]. The choice of the subspaces may be different in any iteration.

Steps (17.4) and (17.5) are called 'updated progressive PGD'. The final iteration may choose for each index m one of the variants (17.2), (17.4), or (17.5).

17.1.2 Convergence

Because of the later property (17.7), the iterates $\mathbf{u}_m$ are bounded. To find a weakly convergent subsequence, we have to assume that

$$\mathbf{V} \text{ is a reflexive Banach tensor space.} \tag{17.6a}$$

The nonlinear functional J has to be sufficiently smooth:

$$J \text{ is Fréchet differentiable with Fréchet differential } J' : \mathbf{V} \to \mathbf{V}^*. \tag{17.6b}$$

J must satisfy the following ellipticity condition with constants $\alpha > 0$ and $s > 1$:

$$\langle J'(\mathbf{v}) - J'(\mathbf{w}), \mathbf{v} - \mathbf{w} \rangle \geq \alpha \left\| \mathbf{v} - \mathbf{w} \right\|^s . \tag{17.6c}$$

Here, $\langle \boldsymbol{\varphi}, \mathbf{v} \rangle := \boldsymbol{\varphi}(\mathbf{v})$ denotes the dual pairing in $\mathbf{V}^* \times \mathbf{V}$. Furthermore, one of the following two conditions (17.6d,e) are required:

$$J : \mathbf{V} \to \mathbb{R} \text{ is weakly sequentially continuous,} \tag{17.6d}$$

i.e., $J(\mathbf{u}_m) \to J(\mathbf{u})$ for sequences $\mathbf{u}_m \rightharpoonup \mathbf{u}$. Alternatively, $J' : \mathbf{V} \to \mathbf{V}^*$ may be assumed to be Lipschitz continuous on bounded sets, i.e.,

$$\left\| J'(\mathbf{v}) - J'(\mathbf{w}) \right\| \leq C_A \left\| \mathbf{v} - \mathbf{w} \right\| \text{ for } \mathbf{v}, \mathbf{w} \in A \text{ and bounded } A \subset \mathbf{V}. \tag{17.6e}$$

As shown in [58, Lemma 3], (17.6b) and (17.6c) imply that J is strictly convex, bounded from below, and satisfies

$$\lim_{\|\mathbf{v}\| \to \infty} J(\mathbf{v}) = \infty. \tag{17.7}$$

Under these conditions including (17.3), the minima in (17.2), (17.4), and (17.5) exist. The values $J(\mathbf{u}_m)$ decrease weakly:

$$J(\mathbf{u}_m) \leq J(\mathbf{u}_{m-1}) \qquad \text{for } m \geq 1.$$

If equality $J(\mathbf{u}_m) = J(\mathbf{u}_{m-1})$ occurs, $\mathbf{u} := \mathbf{u}_{m-1}$ is the solution of the original problem (17.1) (cf. [58, Lemma 8]). The following result is proved in [58, Thm. 4].

Proposition 17.2. *(a) Under the conditions (17.6a-d), all variants of the progressive PGD (17.2), (17.4), (17.5) converge to the solution* $\mathbf{u}$ *of (17.1):* $\mathbf{u}_m \to \mathbf{u}$.
*(b) The same statement holds under conditions (17.6a-*c,e*), if* $s \leq 2$ *in (17.6c).*

17.2 Solution of Optimisation Problems involving Tensor Formats

There are two different strategies in tensor calculus. The first one performs tensor operations (as described in §13) in order to calculate certain tensors (solutions of fixed point iterations, etc.). In this case, truncation procedures are essential for the practical application. The final representation ranks of the solution may be determined adaptively. The second strategy fixes the ranks and tries to optimise the *parameters* of the representation.

17.2.1 Formulation of the Problem

Many problems can be written as minimisation problems of the form

$$\text{find } \mathbf{x} \in \mathbf{V} \text{ such that } J(\mathbf{x}) = \min_{\mathbf{v}\in\mathbf{V}} J(\mathbf{v}). \tag{17.8}$$

Examples are linear systems $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{x}, \mathbf{b} \in \mathbf{V} := \bigotimes_{j=1}^{d} \mathbb{K}^{n_j}$ and $\mathbf{A} \in \mathbf{M} := \bigotimes_{j=1}^{d} \mathbb{K}^{n_j \times n_j}$. Here, J takes the form

$$J(\mathbf{v}) = \langle \mathbf{Av}, \mathbf{v} \rangle - \Re\mathrm{e}\, \langle \mathbf{b}, \mathbf{v} \rangle ,$$

if $\mathbf{A}$ is positive definite (cf. [81, §10.1.4]). Otherwise, use

$$J(\mathbf{v}) = \left\| \mathbf{Av} - \mathbf{b} \right\|^2 \;\text{ or }\; \left\| \mathbf{B}\left(\mathbf{Av} - \mathbf{b} \right) \right\|^2$$

with a preconditioning operator $\mathbf{B}$. The largest eigenvalue of a positive definite matrix $\mathbf{A}$ and the corresponding eigenvector can be determined from the Rayleigh quotient

$$J(\mathbf{v}) = \frac{\langle \mathbf{Av}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

Let $\mathcal{F} = \mathcal{F}(\mathbf{V})$ be any tensor format (e.g., $\mathcal{R}_r$, $\mathcal{T}_{\mathbf{r}}$, $\mathcal{H}_{\mathfrak{r}}$, etc.). Instead of Problem (17.8), we want to solve

$$\text{find } \mathbf{x}_{\mathcal{F}} \in \mathcal{F} \text{ such that } J(\mathbf{x}_{\mathcal{F}}) = \min_{\mathbf{v}\in\mathcal{F}} J(\mathbf{v}). \tag{17.9}$$

Definition 17.3. A mapping $J : \mathbf{V} \to \mathbb{R} \cup \{\infty\}$ is called *weakly sequentially lower semicontinuous* in $S \subset \mathbf{V}$, if

$$J(v) \leq \liminf_{n\to\infty} J(v_n) \quad \text{for all } v_n, v \in S \text{ with } v_n \rightharpoonup v.$$

The following conditions ensuring the existence of a minimiser of (17.9) can be found in [58, Theorems 1 and 2].

Proposition 17.4. *Problem (17.9) is solvable, if* $\mathbf{V}$ *is a reflexive Banach space,* $\mathcal{F}$ *is weakly closed,* J *is weakly sequentially lower semicontinuous, and either* $\mathcal{F}$ *is bounded or* $\lim_{\|\mathbf{v}\|\to\infty} J(\mathbf{v}) = \infty$.

$$\text{In the cases } \mathcal{F} = \left\{ \begin{array}{c} \mathcal{R}_r \\ \mathcal{T}_{\mathbf{r}} \\ \mathcal{H}_{\mathfrak{r}} \end{array} \right\} \text{ we may write } \mathbf{x}_{\mathcal{F}} = \left\{ \begin{array}{c} \mathbf{x}_r \\ \mathbf{x}_{\mathbf{r}} \\ \mathbf{x}_{\mathfrak{r}} \end{array} \right\} \text{, respectively.}$$

Remark 17.5. Let the minimisers $\mathbf{x}$ from (17.8) and $\mathbf{x}_{\mathcal{F}}$ from (17.9) exist and assume that J is continuous (condition (17.6b) is sufficient). Then the respective values $J(\mathbf{x}_r)$, $J(\mathbf{x}_{\mathbf{r}})$, or $J(\mathbf{x}_{\mathfrak{r}})$ converge to $J(\mathbf{x})$, provided that $r \to \infty$, $\min \mathbf{r} := \min_{1\leq j\leq d} r_j \to \infty$, or $\min \mathfrak{r} := \min_{\alpha\in T_D} r_\alpha \to \infty$, respectively.

Proof. By definition of $\mathbf{V}$, for any $\varepsilon > 0$ there are $\eta > 0$, $r \in \mathbb{N}_0$, and $\mathbf{x}_\varepsilon \in \mathcal{R}_r$ such that $\|\mathbf{x} - \mathbf{x}_\varepsilon\| \leq \eta$ and $J(\mathbf{x}_\varepsilon) - J(\mathbf{x}) \leq \varepsilon$. The optimal $\mathbf{x}_r \in \mathcal{R}_r$ satisfies $J(\mathbf{x}_r) - J(\mathbf{x}) \leq J(\mathbf{x}_\varepsilon) - J(\mathbf{x}) \leq \varepsilon$. This proves $J(\mathbf{x}_r) \to J(\mathbf{x})$ for the optimal $\mathbf{x}_r \in \mathcal{R}_r$. Setting $r := \min \mathbf{r}$, we derived from $\mathcal{R}_r \subset \mathcal{T}_\mathbf{r}$, that also $J(\mathbf{x}_\mathbf{r}) \to J(\mathbf{x})$. Similarly for $\mathbf{x}_\mathfrak{r} \in \mathcal{H}_\mathfrak{r}$, since $\mathcal{R}_r \subset \mathcal{H}_\mathfrak{r}$ for $r = \min \mathfrak{r}$. □

17.2.2 Reformulation, Derivatives, and Iterative Treatment

The general form of a format description $\mathbf{v} = \rho_\mathcal{S}(\ldots)$ is considered in §7.1.1. Particular examples are $\rho_\text{r-term}(r, (v_\nu^{(j)}))$ for the r-term format (7.7a), $\rho_\text{TS}(\mathbf{a}, (B_j))$ for the general subspace format in (8.6c), $\rho_\text{HTR}(T_D, (\mathbf{C}_\alpha), c^{(D)}, (B_j))$ for the hierarchical format in (11.28), etc. Discrete parameters like r in $\rho_\text{r-term}(r, (v_\nu^{(j)}))$ or T_D in $\rho_\text{HTR}(T_D, \ldots)$ are fixed. All other parameters are variable. Renaming the latter parameters

$$\mathbf{p} := (p_1, \ldots, p_m),$$

we obtain the description $\mathbf{v} = \rho_\mathcal{F}(\mathbf{p})$, where $\mathbf{p}$ varies in $\mathbf{P}$. The minimisation in (17.9) is equivalent to

$$\text{find } \mathbf{p} \in \mathbf{P} \text{ such that } J(\rho_\mathcal{F}(\mathbf{p})) = \min_{\mathbf{q} \in \mathbf{P}} J(\rho_\mathcal{F}(\mathbf{q})). \tag{17.10}$$

Iterative optimisation methods require at least parts of the derivatives in

$$\partial J(\rho_\mathcal{F}(\mathbf{p}))/\partial \mathbf{p} = \frac{\partial J}{\partial \mathbf{v}} \frac{\partial \rho_\mathcal{F}}{\partial \mathbf{p}}$$

or even second order derivatives like the Hessian. Since the mapping $\rho_\mathcal{F}(\mathbf{p})$ is multilinear in $\mathbf{p}$, the format-dependent part $\frac{\partial \rho_\mathcal{F}}{\partial \mathbf{p}}$ as well as higher derivatives are easy to determine.

Since, in general, the representations are non-unique (cf. §7.1.3), the Jacobi matrix $\frac{\partial \rho_\mathcal{F}}{\partial \mathbf{p}}$ does not have full rank. For instance, for $\rho_\text{r-term}(r, (v_\nu^{(j)}))$ the parameters are $p_1 := v_1^{(1)} \in V_1$, $p_2 := v_1^{(2)} \in V_2, \ldots$ The fact that $\rho_\mathcal{F}(sp_1, \frac{1}{s}p_2, p_3, \ldots)$ is independent of $s \in \mathbb{K}\backslash\{0\}$ leads to $\langle \frac{\partial \rho_\mathcal{F}}{\partial p_1}, v_1^{(1)} \rangle = \langle \frac{\partial \rho_\mathcal{F}}{\partial p_2}, v_1^{(2)} \rangle$. Hence $\frac{\partial \rho_\mathcal{F}}{\partial \mathbf{p}}$ has a nontrivial kernel. In order to avoid redundancies of the r-term format, one may, e.g., equi-normalise the vectors: $\|v_\nu^{(1)}\| = \|v_\nu^{(2)}\| = \ldots$ The problem of redundant parameters will be discussed in more detail in §17.3.1.

The usual iterative optimisation methods are alternating optimisations (ALS, see §9.5.2). A modification (MALS: modified alternating least squares) which often yields good results is the overlapping ALS, where optimisation is perform consecutively with respect to (p_1, p_2), (p_2, p_3), $(p_3, p_4), \ldots$ (cf. variant (δ) in §9.5.2.1). For particular quantum physics applications, this approach is called DMRG (density matrix renormalisation group, cf. [196, 197]). For a detailed discussion see Holtz-Rohwedder-Schneider [103] and Oseledets [154].

17.3 Ordinary Differential Equations

We consider initial value problems

$$\frac{d}{dt}\mathbf{v} = \mathbf{F}(t,\mathbf{v}) \text{ for } t \geq 0, \qquad \mathbf{v}(0) = \mathbf{v}_0, \tag{17.11}$$

where $\mathbf{v} = \mathbf{v}(t) \in \mathbf{V}$ belongs to a tensor space, while $\mathbf{F}(t,\cdot) : \mathbf{V} \to \mathbf{V}$ is defined for $t \geq 0$. $\mathbf{v}_0$ is the initial value. Since $\mathbf{V}$ may be a function space, $\mathbf{F}$ can be a differential operator. Then, parabolic problems $\frac{d}{dt}\mathbf{v} = \Delta\mathbf{v}$ or the instationary Schrödinger equation $\frac{d}{dt}\mathbf{v} = -\mathrm{i}H\mathbf{v}$ are included into the setting (17.11).

The discretisation with respect to time is standard. The unusual part is the discretisation with respect to a fixed format for the tensor $\mathbf{v}$. For this purpose we have to introduce the tangent space of a manifold.

17.3.1 Tangent Space

Let a format $\mathcal{F}$ be defined via $\mathbf{v} = \rho_\mathcal{F}(p_1,\ldots,p_m)$ (cf. §7.1.1), where $\rho_\mathcal{F}$ is differentiable with respect to the parameters p_i, $1 \leq i \leq m$. The set $\mathcal{F}$ forms a manifold parametrised by $p_1,\ldots,p_m \in \mathbb{K}$. The subscripts $r, \mathbf{r}$, and $\mathfrak{r}$ in $\mathcal{F} = \mathcal{R}_r$, $\mathcal{F} = \mathcal{T}_\mathbf{r}$, and $\mathcal{F} = \mathcal{H}_\mathfrak{r}$ indicate the fixed representation ranks.

Definition 17.6. Let $\mathbf{v} = \rho_\mathcal{F}(p_1,\ldots,p_m) \in \mathcal{F}$. The linear space

$$\mathcal{T}(\mathbf{v}) := \operatorname{span}\{\partial\rho_\mathcal{F}(p_1,\ldots,p_m)/\partial p_i : 1 \leq i \leq m\} \subset \mathbf{V}$$

is the *tangent space* at $\mathbf{v} \in \mathcal{F}$.

As observed above, the mapping $\rho_\mathcal{F}(p_1,\ldots,p_m)$ is, in general, not injective. Therefore, a strict inequality $m_\mathcal{T} := \dim(\mathcal{T}(\mathbf{v})) < m$ may hold. Instead of a bijective parametrisation, we use a basis for $\mathcal{T}(\mathbf{v})$. This will be exercised for the format $\mathcal{T}_\mathbf{r}$ in §17.3.3 and format $\mathcal{H}_\mathfrak{r}$ in §17.3.4.

17.3.2 Dirac-Frenkel Discretisation

The Galerkin method restricts $\mathbf{v}$ in (17.11) to a certain linear subspace. Here, we restrict $\mathbf{v}$ to the manifold $\mathcal{F}$, which is no subspace. We observe that any differentiable function $\mathbf{v}_\mathcal{F}(t) \in \mathcal{F}$ has a derivative $\frac{d}{dt}\mathbf{v}_\mathcal{F} \in \mathcal{T}(\mathbf{v}_\mathcal{F})$. Hence, the right-hand side $\mathbf{F}(t,\mathbf{v}_\mathcal{F})$ in (17.11) has to be replaced by an expression belonging to $\mathcal{T}(\mathbf{v}_\mathcal{F})$. The closest one is the orthogonal projection of $\mathbf{F}$ onto $\mathcal{T}(\mathbf{v}_\mathcal{F})$. Denoting the orthogonal projection onto $\mathcal{T}(\mathbf{v}_\mathcal{F})$ by $P(\mathbf{v}_\mathcal{F})$, we get the substitute

$$\begin{aligned} &\mathbf{v}_\mathcal{F}(t) \in \mathbf{F} \quad \text{with } \mathbf{v}_\mathcal{F}(0) = \mathbf{v}_{0\mathcal{F}} \in \mathbf{F} \quad \text{and} \\ &\tfrac{d}{dt}\mathbf{v}_\mathcal{F}(t) = P(\mathbf{v}_\mathcal{F}(t))\mathbf{F}(t,\mathbf{v}_\mathcal{F}(t)) \quad \text{for } t \geq 0, \end{aligned} \tag{17.12}$$

where the initial value $\mathbf{v}_{0\mathcal{F}}$ is an approximation of $\mathbf{v}_0$. The new differential equation is called the *Dirac-Frenkel discretisation* of (17.11) (cf. [48], [65]). The variational formulation of (17.12) is

$$\left\langle \frac{d}{dt}\mathbf{v}_{\mathcal{F}}(t) - \mathbf{F}(t, \mathbf{v}_{\mathcal{F}}(t)), \mathbf{t} \right\rangle = 0 \qquad \text{for all } \mathbf{t} \in \mathcal{T}(\mathbf{v}_{\mathcal{F}}).$$

Concerning an error analysis of this discretisation we refer to Lubich [143, 144] and Koch-Lubich [126].

For a concrete discretisation of (17.12), we may choose the explicit Euler scheme. The parameters of $\mathbf{v}_n \approx \mathbf{v}_{\mathcal{F}}(n \cdot \Delta t) \in \mathcal{F}$ are $\rho_{\mathcal{F}}(p_1^{(n)}, \ldots, p_m^{(n)})$. The vector $P(\mathbf{v}_n)\mathbf{F}(t, \mathbf{v}_n)$ from the tangent space leads to coordinates $\partial p_1^{(n)}, \ldots, \partial p_m^{(n)}$ (their explicit description in the case of $\mathcal{F} = \mathcal{T}_{\mathbf{r}}$ is given below in Lemma 17.8). Then, the Euler scheme with step size Δt produces the next approximation

$$\mathbf{v}_{n+1} := \rho_{\mathcal{F}}(p_1^{(n)} + \Delta t \partial p_1^{(n)}, \ldots, p_m^{(n)} + \Delta t \partial p_m^{(n)}).$$

17.3.3 Tensor Subspace Format $\mathcal{T}_{\mathbf{r}}$

Tensors from $\mathcal{F} = \mathcal{T}_{\mathbf{r}}$ are represented by

$$\mathbf{v} = \rho_{\mathrm{orth}}(\mathbf{a}, (B_j)) = \sum\nolimits_{\mathbf{i} \in \mathbf{J}} \mathbf{a}_{\mathbf{i}} \bigotimes\nolimits_{j=1}^{d} b_{i_j}^{(j)} = \mathbf{B}\,\mathbf{a}$$

(cf. (8.8b)), where without loss of generality we use orthonormal bases $B_j = (b_1^{(j)}, \ldots, b_{r_j}^{(j)}) \in V_j^{r_j}$. Note that $\mathbf{B} = \bigotimes_{j=1}^{d} B_j$. The set of matrices $B_j \in \mathbb{K}^{I_j \times r_j}$ with $B_j^{\mathsf{H}} B_j = I$ is called *Stiefel manifold.*

Lemma 17.7. *Let* $\mathbf{v} = \rho_{\mathrm{orth}}(\mathbf{a}, (B_j)) \in \mathcal{T}_{\mathbf{r}}$. *Every tangent tensor* $\mathbf{t} \in \mathcal{T}(\mathbf{v})$ *has a representation of the form*

$$\mathbf{t} = \mathbf{B}\mathbf{s} + \left(\sum_{j=1}^{d} B_1 \otimes \ldots \otimes B_{j-1} \otimes C_j \otimes B_{j+1} \otimes \ldots \otimes B_d \right) \mathbf{a}, \tag{17.13a}$$

where C_j *satisfies*

$$B_j^{\mathsf{H}} C_j = 0. \tag{17.13b}$$

C_j *and* $\mathbf{s}$ *are uniquely determined (see (17.13c,d)), provided that* $\mathrm{rank}_j(\mathbf{v}) = r_j$. *Vice versa, for any coefficient tensor* $\mathbf{s}$ *and all matrices* C_j *satisfying (17.13b) the right-hand side in (17.13a) belongs to* $\mathcal{T}(\mathbf{v})$.

Proof. 1a) Any $\mathbf{t} \in \mathcal{T}(\mathbf{v})$ is the limit of $\frac{1}{h}\left[\rho_{\mathrm{orth}}(\mathbf{a} + h\,\delta\mathbf{a}, (B_j + h\,\delta B_j)) - \mathbf{v}\right]$ as $h \to 0$ for some $\delta\mathbf{a}$ and δB_j. This limit equals

$$\mathbf{B}\,\delta\mathbf{a} + \left(\sum_{j=1}^{d} B_1 \otimes \ldots \otimes B_{j-1} \otimes \delta B_j \otimes B_{j+1} \otimes \ldots \otimes B_d \right) \mathbf{a}.$$

The first term is of the form $\mathbf{Bs}$. Since $B_j(h) := B_j + h\,\delta B_j$ must be orthogonal, i.e., $B_j(h)^{\mathsf{H}}B_j(h) = I$, it follows that $B_j^{\mathsf{H}}\,\delta B_j + \delta B_j^{\mathsf{H}}B_j = 0$. Split δB_j into $\delta B_j = \delta B_j^{I} + \delta B_j^{II}$ with $\delta B_j^{I} := \left(I - B_jB_j^{\mathsf{H}}\right)\delta B_j$ and $\delta B_j^{II} := B_jB_j^{\mathsf{H}}\,\delta B_j$. The derivative of $\mathbf{v}$ with respect to δB_j^{II} yields

$$\left(B_1 \otimes \ldots \otimes B_{j-1} \otimes \delta B_j^{II} \otimes B_{j+1} \otimes \ldots \otimes B_d\right)\mathbf{a} = \mathbf{B}\mathbf{a}'$$

with $\mathbf{a}' := \left(id \otimes \ldots \otimes id \otimes B_j^{\mathsf{H}}\,\delta B_j \otimes id \otimes \ldots \otimes id\right)\mathbf{a}$. Such a term can be expressed by $\mathbf{Bs}$ in (17.13a). Therefore, we can restrict δB_j to the part $\delta B_j^{I} =: C_j$, which satisfies (17.13b): $B_j^{\mathsf{H}}\,\delta B_j^{I} = B_j^{\mathsf{H}}\left(I - B_jB_j^{\mathsf{H}}\right)\delta B_j = 0$.

1b) Given $\mathbf{t} \in \mathcal{T}(\mathbf{v})$, we have to determine $\mathbf{s}$ and C_j. $\mathbf{B}^{\mathsf{H}}\mathbf{B} = \mathbf{I}$ and $B_j^{\mathsf{H}}C_j = 0$ imply that

$$\mathbf{s} = \mathbf{B}^{\mathsf{H}}\mathbf{t}. \tag{17.13c}$$

Hence, $\mathbf{t}' = \mathbf{t} - \mathbf{Bs}$ equals $\left(\sum_{j=1}^{d} B_1 \otimes \ldots \otimes B_{j-1} \otimes C_j \otimes B_{j+1} \otimes \ldots \otimes B_d\right)\mathbf{a}$. From (5.5) we conclude that $\mathcal{M}_j(\mathbf{t}') = C_j\mathcal{M}_j(\mathbf{a})\mathbf{B}_{[j]}^{\mathsf{T}}$. Let $\mathcal{M}_j(\mathbf{a}) = U_j\Sigma_jV_j^{\mathsf{T}}$ be the reduced singular value decomposition with $U_j, V_j \in V_j^{r_j}$ and $\Sigma_j \in \mathbb{R}^{r_j\times r_j}$. Thanks to $\mathrm{rank}_j(\mathbf{v}) = r_j$, Σ_j is invertible. This allows to solve for C_j:

$$C_j = \mathcal{M}_j(\mathbf{t}')\,\overline{\mathbf{B}_{[j]}}\,\overline{V_j}\,\Sigma_j^{-1}\,U_j^{\mathsf{H}}. \tag{17.13d}$$

2) Given $\mathbf{s}$ and matrices C_j satisfying (17.13b), the derivative

$$\mathbf{t} = \lim_{h\to 0}\frac{1}{h}\left[\rho_{\mathrm{orth}}(\mathbf{a} + h\,\mathbf{s}, (B_j + h\,C_j)) - \mathbf{v}\right] \in \mathcal{T}(\mathbf{v}) \quad \text{with } \mathbf{v} = \rho_{\mathrm{orth}}(\mathbf{a}, (B_j))$$

has the representation (17.13a). □

A more general problem is the description of the orthogonal projection on $\mathcal{T}(\mathbf{v})$. Given any $\mathbf{w} \in \mathbf{V}$, we need the parameters $\mathbf{s}$ and C_j of $\mathbf{t} \in \mathcal{T}(\mathbf{v})$ defined by $\mathbf{t} = P(\mathbf{v})\mathbf{w}$ and $\mathbf{v} = \rho_{\mathrm{orth}}(\mathbf{a}, (B_j)) \in \mathcal{T}_{\mathbf{r}}$. Another notation for $\mathbf{t}$ is $\|\mathbf{t} - \mathbf{w}\| = \min\{\|\tilde{\mathbf{t}} - \mathbf{w}\| : \tilde{\mathbf{t}} \in \mathcal{T}(\mathbf{v})\}$. First we split $\mathbf{w}$ into the orthogonal components $\mathbf{w} = \mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{w} + \mathbf{w}'$. Since $\mathbf{B}\mathbf{B}^{\mathsf{H}}\mathbf{w} \in \mathcal{T}(\mathbf{v})$, it remains to determine $\mathbf{t}'$ satisfying

$$\|\mathbf{t}' - \mathbf{w}'\| = \min\{\|\tilde{\mathbf{t}}' - \mathbf{w}'\| : \tilde{\mathbf{t}}' \in \mathcal{T}(\mathbf{v})\},$$

where $\mathbf{t}' = \sum_{j=1}^{d}\mathbf{t}_j'$ with $\mathbf{t}_j' := (\mathbf{B}_{[j]} \otimes C_j)\mathbf{a} \in \mathbf{U}_j$ and orthogonal subspaces $\mathbf{U}_j := U_j^{\perp} \otimes \bigotimes_{k\neq j} U_k$, $U_j := \mathrm{range}(B_j)$. Orthogonal projection onto $\mathbf{U}_j$ yields $\|\mathbf{t}_j' - \mathbf{B}_{[j]}\mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{w}'\| = \min\{\|\tilde{\mathbf{t}}_j' - \mathbf{B}_{[j]}\mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{w}'\| : \tilde{\mathbf{t}}_j' \in \mathcal{T}(\mathbf{v}) \cap \mathbf{U}_j\}$. Equivalent statements are

$$\begin{aligned}
&\mathbf{t}_j' = \mathbf{B}_{[j]}(C_j\mathbf{a}) \quad \text{with } \|C_j\mathbf{a} - \mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{w}'\| = \min\{\|\tilde{C}_j\mathbf{a} - \mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{w}'\| : B_j^{\mathsf{H}}\tilde{C}_j = 0\}\\
&\quad\Leftrightarrow \|\mathcal{M}_j(C_j\mathbf{a}) - \mathcal{M}_j(\mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{w}')\|_{\mathsf{F}} = \min_{B_j^{\mathsf{H}}\tilde{C}_j=0}\|\mathcal{M}_j(\tilde{C}_j\mathbf{a}) - \mathcal{M}_j(\mathbf{B}_{[j]}^{\mathsf{H}}\mathbf{w}')\|_{\mathsf{F}}\\
&\quad\Leftrightarrow \|C_j\mathcal{M}_j(\mathbf{a}) - \mathcal{M}_j(\mathbf{w}')\overline{\mathbf{B}_{[j]}}\|_{\mathsf{F}} = \min_{B_j^{\mathsf{H}}\tilde{C}_j=0}\|\tilde{C}_j\mathcal{M}_j(\mathbf{a}) - \mathcal{M}_j(\mathbf{w}')\overline{\mathbf{B}_{[j]}}\|_{\mathsf{F}},
\end{aligned}$$

since the Euclidean norm $\|\mathbf{x}\|$ of a tensor $\mathbf{x}$ is equal to the Frobenius norm of the matricisation $\mathcal{M}_j(\mathbf{x})$. By Exercise 2.12, the minimiser of the last formulation is

$$C_j = \mathcal{M}_j(\mathbf{w}')\,\overline{\mathbf{B}_{[j]}}\,M_j^{\mathsf{H}}\,(M_j M_j^{\mathsf{H}})^{-1} \quad \text{with } M_j := \mathcal{M}_j(\mathbf{a}).$$

The rank condition of Exercise 2.12 is equivalent to $\operatorname{rank}_j(\mathbf{v}) = r_j$. C_j satisfies (17.13b) because of $B_j^{\mathsf{H}}\mathcal{M}_j(\mathbf{w}') = \mathcal{M}_j(B_j^{\mathsf{H}}\mathbf{w}')$ and $B_j^{\mathsf{H}}\mathbf{w}' = B_j^{\mathsf{H}}(I - \mathbf{B})\mathbf{w} = 0$. Using the singular value decomposition $M_j = U_j \Sigma_j V_j^{\mathsf{T}}$, we may rewrite C_j as $\mathcal{M}_j(\mathbf{w}')\overline{\mathbf{B}_{[j]}V_j}\Sigma_j U_j^{\mathsf{H}}(U_j\Sigma_j^2 U_j^{\mathsf{H}})^{-1} = \mathcal{M}_j(\mathbf{w}')\overline{\mathbf{B}_{[j]}V_j}\Sigma_j^{-1}U_j^{\mathsf{H}}$. We summarise the result in the next lemma.

Lemma 17.8. *Let $\mathbf{v} \in \mathbf{V}$ with $\operatorname{rank}_j(\mathbf{v}) = r_j$ ($1 \le j \le d$). For any $\mathbf{w} \in \mathbf{V}$, the orthogonal projection onto $\mathcal{T}(\mathbf{v})$ is given by $\mathbf{t} = P(\mathbf{v})\mathbf{w}$ from (17.13a) with*

$$\mathbf{s} := \mathbf{B}^{\mathsf{H}}\mathbf{w}, \quad C_j := \mathcal{M}_j((\mathbf{I} - \mathbf{B}\mathbf{B}^{\mathsf{H}})\mathbf{w})\,\overline{\mathbf{B}_{[j]}}\,\overline{V_j}\,\Sigma_j^{-1}U_j^{\mathsf{H}}.$$

17.3.4 Hierarchical Format $\mathcal{H}_{\mathfrak{r}}$

We choose the HOSVD representation for $\mathbf{v} \in \mathcal{H}_{\mathfrak{r}}$:

$$\mathbf{v} = \rho_{\mathrm{HTR}}^{\mathrm{HOSVD}}\big(T_D, (\mathbf{C}_\alpha)_{\alpha\in T_D\setminus\mathcal{L}(T_D)}, c^{(D)}, (B_j)_{j\in D}\big)$$

(cf. Definition 11.3.3), i.e., all bases $\mathbf{B}_\alpha = \big[\mathbf{b}_1^{(\alpha)}, \ldots, \mathbf{b}_{r_\alpha}^{(\alpha)}\big]$ of $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ are HOSVD bases. They are characterised by the following properties of the coefficient matrices $C^{(\alpha,\ell)}$ in $\mathbf{C}_\alpha = (C^{(\alpha,\ell)})_{1\le\ell\le r_\alpha}$.

1) For the root D assume $r_D = 1$. Then

$$C^{(D,1)} = \Sigma_\alpha := \operatorname{diag}\{\sigma_1^{(D)}, \ldots\},$$

where $\sigma_i^{(D)}$ are the singular values of the matricisation $\mathcal{M}_{\alpha_1}(\mathbf{v})$ (α_1 son of D).

2) For non-leaf vertices $\alpha \in T_D$, $\alpha \ne D$, we have

$$\sum_{\ell=1}^{r_\alpha}(\sigma_\ell^{(\alpha)})^2\,C^{(\alpha,\ell)}C^{(\alpha,\ell)\mathsf{H}} = \Sigma_{\alpha_1}^2, \quad \sum_{\ell=1}^{r_\alpha}(\sigma_\ell^{(\alpha)})^2\,C^{(\alpha,\ell)\mathsf{T}}\overline{C^{(\alpha,\ell)}} = \Sigma_{\alpha_2}^2, \tag{17.14}$$

where α_1, α_2 are the first and second son of $\alpha \in T_D$, and Σ_{α_i} the diagonal containing the singular values of $\mathcal{M}_{\alpha_i}(\mathbf{v})$ (cf. Exercise 11.41).

Let $\mathbf{v}(t) \in \mathcal{H}_{\mathfrak{r}}$ be a differentiable function. We characterise $\dot{\mathbf{v}}(t)$ at $t = 0$ and abbreviate $\dot{\mathbf{v}} := \dot{\mathbf{v}}(0)$. The differentiation follows the recursion of the representation. Since $r_D = 1$, $\mathbf{v} = c_1^{(D)}\mathbf{b}_1^{(D)}$ yields

$$\dot{\mathbf{v}} = \dot{c}_1^{(D)}\mathbf{b}_1^{(D)} + c_1^{(D)}\dot{\mathbf{b}}_1^{(D)}. \tag{17.15a}$$

The differentiation of the basis functions follows from (11.24):

$$\dot{\mathbf{b}}_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} \dot{c}_{ij}^{(\alpha,\ell)}\, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} \tag{17.15b}$$
$$+ \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)}\, \dot{\mathbf{b}}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} + \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)}\, \mathbf{b}_i^{(\alpha_1)} \otimes \dot{\mathbf{b}}_j^{(\alpha_2)}$$

At the end of the recursion, $\dot{\mathbf{v}}$ is represented by the differentiated coefficients $\dot{c}_1^{(D)}$, $\dot{c}_{ij}^{(\alpha,\ell)}$ and the derivatives $\dot{b}_i^{(j)}$ of the bases at the leaves.

By the same argument as in Lemma 17.7, we may restrict the variations $\dot{\mathbf{b}}_\ell^{(\alpha)}$ to the orthogonal complement of $\mathbf{U}_\alpha^{\min}(\mathbf{v})$, i.e.,

$$\dot{\mathbf{b}}_\ell^{(\alpha)} \perp \mathbf{U}_\alpha^{\min}(\mathbf{v})\,. \tag{17.16}$$

We introduce the projections

$$P_\alpha : \mathbf{V}_\alpha \to \mathbf{U}_\alpha^{\min}(\mathbf{v}), \qquad P_\alpha = \sum_{\ell=1}^{r_\alpha} \left\langle \cdot, \mathbf{b}_\ell^{(\alpha)} \right\rangle \mathbf{b}_\ell^{(\alpha)}$$

onto $\mathbf{U}_\alpha^{\min}(\mathbf{v})$ and its complement $P_\alpha^\perp := I - P_\alpha$.

Next, we discuss the unique representation of the parameters. $\dot{c}_1^{(D)}$ is obtained as

$$\dot{c}_1^{(D)} = \left\langle \dot{\mathbf{v}}, \mathbf{b}_1^{(D)} \right\rangle. \tag{17.17a}$$

$\dot{\mathbf{b}}_1^{(D)}$ is the result of

$$\dot{\mathbf{b}}_1^{(D)} = \frac{1}{c_1^{(D)}} P_D^\perp \dot{\mathbf{v}} \tag{17.17b}$$

with $\|\mathbf{v}\| = |c_1^{(D)}|$. Note that $c_1^{(D)} \dot{\mathbf{b}}_1^{(D)}$ is the quantity of interest.

We assume by induction that $\dot{\mathbf{b}}_\ell^{(\alpha)}$ is known and use (17.15b):

$$\dot{c}_{ij}^{(\alpha,\ell)} = \left\langle \dot{\mathbf{b}}_\ell^{(\alpha)}, \mathbf{b}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)} \right\rangle. \tag{17.17c}$$

Set

$$\beta_\ell^{(\alpha)} := \left(P_{\alpha_1}^\perp \otimes id\right) \dot{\mathbf{b}}_\ell^{(\alpha)} = \sum_{i=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{ij}^{(\alpha,\ell)}\, \dot{\mathbf{b}}_i^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}.$$

The scalar product of $(\sigma_\ell^{(\alpha)})^2 \beta_\ell^{(\alpha)}$ and $\sum_k c_{i'k}^{(\alpha,\ell)} \mathbf{b}_k^{(\alpha_2)}$ with respect to $\mathbf{V}_{\alpha_2}$ is

$$\left\langle (\sigma_\ell^{(\alpha)})^2 \beta_\ell^{(\alpha)}, \sum_{k=1}^{r_{\alpha_2}} c_{ik}^{(\alpha,\ell)} \mathbf{b}_k^{(\alpha_2)} \right\rangle_{\alpha_2}$$
$$= \left\langle (\sigma_\ell^{(\alpha)})^2 \sum_{i'=1}^{r_{\alpha_1}} \sum_{j=1}^{r_{\alpha_2}} c_{i'j}^{(\alpha,\ell)}\, \dot{\mathbf{b}}_{i'}^{(\alpha_1)} \otimes \mathbf{b}_j^{(\alpha_2)}, \sum_{k=1}^{r_{\alpha_2}} c_{ik}^{(\alpha,\ell)} \mathbf{b}_k^{(\alpha_2)} \right\rangle_{\alpha_2}$$

$$= \sum_{i'=1}^{r_{\alpha_1}} \left\langle (\sigma_\ell^{(\alpha)})^2 \sum_{j=1}^{r_{\alpha_2}} c_{i'j}^{(\alpha,\ell)}\, \mathbf{b}_j^{(\alpha_2)}, \sum_{k=1}^{r_{\alpha_2}} c_{ik}^{(\alpha,\ell)} \mathbf{b}_k^{(\alpha_2)} \right\rangle \dot{\mathbf{b}}_{i'}^{(\alpha_1)}$$

$$= \sum_{i'=1}^{r_{\alpha_1}} \left(\sum_{j=1}^{r_{\alpha_2}} (\sigma_\ell^{(\alpha)})^2 c_{i'j}^{(\alpha,\ell)} \overline{c_{ij}^{(\alpha,\ell)}} \right) \dot{\mathbf{b}}_i^{(\alpha_1)} = \sum_{i'=1}^{r_{\alpha_1}} (\sigma_\ell^{(\alpha)})^2 \left(C^{(\alpha,\ell)} C^{(\alpha,\ell)\mathsf{H}} \right)_{i'i} \dot{\mathbf{b}}_{i'}^{(\alpha_1)}.$$

Summation over ℓ and identity (17.14) yield

$$\sum_{\ell=1}^{r_\alpha} \left\langle (\sigma_\ell^{(\alpha)})^2 \beta_\ell^{(\alpha)}, \sum_{k=1}^{r_{\alpha_2}} c_{ik}^{(\alpha,\ell)} \mathbf{b}_k^{(\alpha_2)} \right\rangle_{\alpha_2} = \sum_{i'=1}^{r_{\alpha_1}} \left(\sum_{\ell=1}^{r_\alpha} (\sigma_\ell^{(\alpha)})^2 C^{(\alpha,\ell)} C^{(\alpha,\ell)\mathsf{H}} \right)_{i'i} \dot{\mathbf{b}}_{i'}^{(\alpha_1)}$$

$$= \sum_{i'=1}^{r_{\alpha_1}} \left(\Sigma_{\alpha_1}^2 \right)_{i'i} \dot{\mathbf{b}}_{i'}^{(\alpha_1)} = (\sigma_i^{(\alpha_1)})^2\, \dot{\mathbf{b}}_i^{(\alpha_1)}. \tag{17.17d}$$

Similarly, $\gamma_\ell^{(\alpha)} := \left(id \otimes P_{\alpha_2}^{\perp} \right) \dot{\mathbf{b}}_\ell^{(\alpha)}$ holds and

$$\sum_{\ell=1}^{r_\alpha} \left\langle (\sigma_\ell^{(\alpha)})^2 \gamma_\ell^{(\alpha)}, \sum_{i=1}^{r_{\alpha_1}} c_{ij}^{(\alpha,\ell)} \mathbf{b}_i^{(\alpha_2)} \right\rangle_{\alpha_1} = (\sigma_j^{(\alpha_2)})^2\, \dot{\mathbf{b}}_j^{(\alpha_2)}. \tag{17.17e}$$

We summarise: Assume $\mathbf{v} \in \mathcal{H}_{\mathfrak{r}}$ and $\dim(\mathbf{U}_\alpha^{\min}(\mathbf{v})) = r_\alpha$ for $\alpha \in T_D$ (this implies $\sigma_i^{(\alpha)} > 0$ for $1 \leq i \leq r_\alpha$). Under condition (17.16), the tangential tensor $\dot{\mathbf{v}} \in \mathcal{H}_{\mathfrak{r}}$ has a unique description by $\dot{c}_1^{(D)}$, $\dot{\mathbf{b}}_\ell^{(\alpha)}$, and $\dot{c}_{ij}^{(\alpha,\ell)}$ characterised in (17.17a-e).

An investigation of the tangent space of the TT format is given by Holtz-Rohwedder-Schneider [102].

17.4 ANOVA

ANOVA is the abbreviation of ‘analysis of variance’. It uses a decomposition of functions into terms of different spatial dimensions. If contributions of high spatial dimension are sufficiently small, an approximation by functions of a smaller number of variables is possible.

17.4.1 Definitions

Consider a space $\mathbf{V}$ of functions in d variables. As example we choose

$$\mathbf{V} = C([0,1]^d) = {}_{\|\cdot\|_\infty}\bigotimes_{j=1}^{d} V_j \quad \text{with } V_j = C([0,1]).$$

We denote the function with constant value one by $1 \in V_k$. Functions which are constant with respect to the variable x_k are characterised by $U_k^{\min}(f) = \operatorname{span}\{1\}$

and can also be considered as elements of $\mathbf{V}_{[k]} = {}_{\|\cdot\|_\infty}\bigotimes_{j\in D\backslash\{k\}} V_j$. We may identify ${}_{\|\cdot\|_\infty}\bigotimes_{j\in t} V_j$ for any subset $t \subset D$ with

$$\mathbf{V}_t := {}_{\|\cdot\|_\infty}\bigotimes_{j\in t} V_j \cong {}_{\|\cdot\|_\infty}\bigotimes_{j=1}^{d} W_j \subset \mathbf{V} \quad \text{with } W_j := \begin{cases} V_j & \text{if } j \in t, \\ \operatorname{span}\{1\} & \text{if } j \notin t \end{cases}$$

(cf. Remark 3.25a). For instance, $f \in \mathbf{V}_\emptyset$ is a globally constant function, while $f \in \mathbf{V}_{\{1,3,4,\ldots\}}$ is constant with respect to x_2 so that $f(x_1, x_2, x_3, \ldots)$ can also be written as $f(x_1, x_3, x_4, \ldots)$.

Fix a functional $P_j \in V_j^*$ with $P_j 1 = 0$. We denote the mapping

$$f \in V_j \mapsto (P_j f) \cdot 1 \in V_j$$

by the same symbol P_j. In the second interpretation, $P_j \in \mathcal{L}(V_j, V_j)$ is a *projection* onto the subspace $\operatorname{span}\{1\} \subset V_j$. For each subset $t \subset D$, the product $\mathbf{P}_t := \prod_{j\in t} P_j$ defines a projection onto $\mathbf{V}_{t^c}$, where $t^c = D\backslash t$. Note that the order of its factors is irrelevant. $P_\emptyset = id$ holds for $t = \emptyset$. A tensor notation is

$$\mathbf{P}_t := \bigotimes_{j=1}^{d} \left\{ \begin{matrix} P_j & \text{if } j \in t \\ id & \text{if } j \notin t \end{matrix} \right\} \in \mathcal{L}(\mathbf{V}, \mathbf{V}_{t^c}).$$

The recursive definition

$$f_t := \mathbf{P}_{t^c} f - \sum\nolimits_{\tau \subsetneqq t} f_\tau \tag{17.18}$$

starts with $t = \emptyset$, since the empty sum in (17.18) leads to the constant function $f_\emptyset = \mathbf{P}_D f \in \mathbf{V}_\emptyset$. As $\mathbf{P}_{D^c} = \mathbf{P}_\emptyset$ is the identity, the choice $t = D$ in (17.18) yields the *ANOVA decomposition*

$$f = \sum_{t \subset D} f_t. \tag{17.19}$$

Note that f_t depends on (at most) $\#t$ variables.

17.4.2 Properties

Lemma 17.9. *(a) Let $s, t \subset D$. The Hadamard product of $f \in \mathbf{V}_s$ and $g \in \mathbf{V}_t$ belongs to $\mathbf{V}_{s\cup t}$.*

(b) $P_j f_t = 0$ holds for f_t from (17.19) with $j \in t$.

(c) If $s, t \subset D$ are different, the ANOVA components f_s and g_t of some functions f and g satisfy

$$\mathbf{P}_\tau (f_s \odot g_t) = 0 \quad \textit{for all } \tau \subset D \textit{ with } \tau \cap (s\backslash t \cup t\backslash s) \neq \emptyset.$$

Proof. Part (a) is trivial. For (b) we rewrite (17.18) as

$$\mathbf{P}_{t^c} f = \sum_{\tau \subset t} f_\tau, \tag{17.20a}$$

where the sum includes $\tau = t$. Split the sum into $\sum_{\tau\subset t \text{ with } j\in\tau}$ and $\sum_{\tau\subset t \text{ with } j\notin\tau}$. The second sum is identical to $\sum_{\tau\subset t\backslash\{j\}}$:

$$\mathbf{P}_{t^c} f = \sum_{\tau\subset t \text{ with } j\in\tau} f_\tau + \sum_{\tau\subset t\backslash\{j\}} f_\tau \underset{(17.20a)}{=} \sum_{\tau\subset t \text{ with } j\in\tau} f_\tau + \mathbf{P}_{(t\backslash\{j\})^c} f. \tag{17.20b}$$

Since $(t\backslash\{j\})^c = t^c \cup \{j\}$, the projection P_j satisfies the identity $P_j \circ \mathbf{P}_{t^c} = \mathbf{P}_{(t\backslash\{j\})^c} = P_j \circ \mathbf{P}_{(t\backslash\{j\})^c}$, and application of P_j to (17.20b) leads to

$$P_j\Bigg(\sum_{\tau\subset t \text{ with } j\in\tau} f_\tau\Bigg) = 0 \quad \text{for all } t\subset D \text{ with } j\in t. \tag{17.20c}$$

We use induction over $\#t$. The smallest set with $j\in t$ is $t=\{j\}$, for which (17.20c) becomes $P_j f_{\{j\}} = 0$. Assume that $P_j f_\sigma = 0$ holds for all $\sigma \subset D$ with $j\in\sigma$ and $\#\sigma \leq k < d$. Assume $j\in t$ and $\#t = k+1$. Identity (17.20c) shows $0 = \sum_{\tau\subset t \text{ with } j\in\tau} P_j f_\tau = P_j f_t$, since all other τ satisfy $\#\tau \leq k$.

For Part (c) choose some $j \in \tau \cap (s\backslash t \cup t\backslash s)$. Assume without loss of generality that $j\in t$, but $j\notin s$. Hence, f_s is constant with respect to x_j and, since Part (b) implies that $P_j g_t = 0$, also $P_j(f_s \odot g_t) = 0$ and $\mathbf{P}_\tau(f_s\odot g_t) = 0$ hold for any τ with $j\in\tau$. □

For the choice $V_j = L^2_{\rho_j}([0,1])$ of the Hilbert space of weighted L^2 functions, the scalar product is defined by $(f,g) = \int_0^1 f(x)\overline{g(x)}\mathrm{d}\rho_j(x)$. The weight should be scaled such that $\int_0^1 \mathrm{d}\rho_j = 1$. The induced scalar product of $\mathbf{V}$ is $\int_{[0,1]^d} f\overline{g}\,\mathrm{d}\rho$ with $\mathrm{d}\rho = \prod_{j=1}^d \mathrm{d}\rho_j$. Then the functional $P_j \in V_j^*$ from above can be defined by $P_j f := \int_0^1 f\mathrm{d}\rho_j$ and Lemma 17.9c takes the following form.

Remark 17.10. Let $\mathbf{V} = L^2_\rho([0,1]^d)$, $P_j\varphi := \int_0^1 \varphi\mathrm{d}\rho_j$, and $s,t\subset D$ with $s\neq t$. Then the ANOVA components f_s, g_t of any $f,g\in\mathbf{V}$ are orthogonal in $\mathbf{V}$.

In this case, $f = \sum_{t\subset D} f_t$ is an orthogonal decomposition. If the terms f_t decrease with increasing $\#t$, one may truncate the sum and introduce the approximation $f^{(k)} := \sum_{t:\#t\leq k} f_t$. It may even happen that f_t vanishes for $\#t > k$. An important example of $f_t = 0$ for $\#t > 2$ is the potential of the Schrödinger equation[1]

$$f(x_1,\ldots,x_d) = \sum_{1\leq i<j\leq d} \frac{1}{\|x_i - x_j\|} - \sum_{1\leq j\leq d}\sum_{1\leq i\leq A} \frac{Q_i}{\|a_i - x_j\|}$$

($x_j \in \mathbb{R}^3$, d: number of electrons, A: number of nuclei, a_i: positions of nuclei, Q_i: charges).

[1] The terms on the right-hand side are not yet the components $f_{\{j\}}$ and $f_{\{i,j\}}$, since these expressions are not orthogonal.

Since discrete weights (sums of Dirac distributions) are not excluded, functions may be replaced by $\mathbb{K}^n$ vectors. Therefore, the described analysis applies also to finite dimensional tensor spaces $\bigotimes_{j=1}^{d}\mathbb{K}^{n_j}$. Discrete weights are also introduced for the so-called 'anchored ANOVA' to simplify quadrature (cf. Griebel [77]).

17.4.3 Combination with Tensor Representations

So far, the components f_t are general $\#t$-variate functions (possibly with better regularity than f; cf. Griebel-Kuo-Sloan [78]). Any practical implementation has to introduce some discretisation in order to represent f_t. Depending on $\#t$ and the discretisation size, a representation of f_t by one of the tensor representations might be useful. For representation schemes, which have d as linear factor for the storage cost, the reduction from d to $\#t < d$ is of limited help, unless the representation ranks of f_t are much smaller than for f. On the other hand, the large number of terms in $\sum_{t\subset D} f_t$ or $\sum_{t:\#t\leq k} f_t$ is not encouraging the choice of ANOVA with tensor representations compared with an overall tensor representation. An exception will be discussed next.

17.4.4 Symmetric Tensors

A symmetric tensor $\mathbf{v} \in \mathfrak{S}_d(V) \subset \mathbf{V} = \otimes^d V$ may be represented by any $\mathbf{w} \in \mathbf{V}$ with the property $\mathbf{v} = P_{\mathfrak{S}}(\mathbf{w})$ (projection $P_{\mathfrak{S}}$ from (3.45)). However, if two tensors $\mathbf{v}', \mathbf{v}'' \in \mathfrak{S}_d(V)$ are represented via $\mathbf{w}', \mathbf{w}'' \in \mathbf{V}$, the computation of the scalar product $\langle \mathbf{v}', \mathbf{v}''\rangle$ is not easily described by means of $\mathbf{w}', \mathbf{w}''$.

The ANOVA decomposition (17.19) has components f_t such that $f_t \in \mathfrak{S}_{\#t}(V)$. Different t, t' of same cardinality lead to identical[2] functions $f_t = f_{t'}$. Therefore, a symmetric decomposition (17.19) requires only the data

$$f_\emptyset,\ f_{\{1\}},\ f_{\{1,2\}}, \ldots,\ f_{\{1,\ldots,d\}}.$$

Assume that P_j is related to the scalar product. Then, by Remark 17.10, the ANOVA decompositions $f = \sum_{t\subset D} f_t$ and $g = \sum_{t\subset D} g_t$ satisfy

$$\langle f, g\rangle = \sum_{t\subset D} \langle f_t, g_t\rangle = \sum_{k=0}^{d} \tbinom{k}{d} \left\langle f_{\{1,\ldots,k\}}, g_{\{1,\ldots,k\}} \right\rangle,$$

since there are $\binom{k}{d}$ different $t \subset D$ with $\#t = k$. This property favours the ANOVA representation with $f_{\{1,\ldots,k\}}$ in some tensor representation.

[2] The understanding of f_t is ambiguous. One may regard f_t as a function of $\#t$ variables. In this sense, $f_t(\xi_1, \ldots, \xi_{\#t}) = f_{t'}(\xi_1, \ldots, \xi_{\#t})$ holds. On the other hand, the univariate functions $f_{\{1\}^c} = f_{\{2\}^c}$ become different, when they are written as $f_{\{1\}^c}(x_1)$ and $f_{\{2\}^c}(x_2)$ involving two different independent variables (as it happens, e.g., in $f_{\{1\}^c}(x_1) + f_{\{2\}^c}(x_2)$).

References

1. Aleksandrov, A., Peller, V.: Functions of perturbed operators. C. R. Acad. Sci. Paris, Ser. I **347**, 483–488 (2009)
2. Almlöf, J.: Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach. Chem. Phys. Lett. **176**, 319–320 (1991)
3. Appellof, C.J., Davidson, E.R.: Strategies for analyzing data from video fluorometric monitoring of liquid-chromatographic effluents. Anal. Chem. **13**, 2053–2056 (1981)
4. Bader, B.W., Kolda, T.G.: MATLAB tensor toolbox, version 2.3. Tech. rep., http://csmr.ca.sandia.gov/˜tgkolda/TensorToolbox (2007)
5. Ballani, J.: Fast evaluation of BEM integrals based on tensor approximations. Preprint 77, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (2010)
6. Ballani, J., Grasedyck, L.: A projection method to solve linear systems in tensor format. Preprint 22, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (2010)
7. Ballani, J., Grasedyck, L., Kluge, M.: Black box approximation of tensors in hierarchical Tucker format. Linear Algebra Appl. (2011). To appear
8. Bebendorf, M.: Approximation of boundary element matrices. Numer. Math. **86**, 565–589 (2000)
9. Bebendorf, M.: Hierarchical matrices, *Lect. Notes Comput. Sci. Eng.*, vol. 63. Springer, Berlin (2008)
10. Bebendorf, M.: Adaptive cross approximation of multivariate functions. Constr. Approx. **34**, 149–179 (2011)
11. Bellman, R.: Adaptive control processes - a guided tour. Princeton University Press, New Jersey (1961)
12. Benedikt, U., Auer, A.A., Espig, M., Hackbusch, W.: Tensor decomposition in post-Hartree-Fock methods. I. Two-electron integrals and MP2. J. Chem. Phys. **134** (2011)
13. Bergman, G.M.: Ranks of tensors and change of base field. J. Algebra **11** 613–621 (1969)
14. Bernstein, S.N.: Leçons sur les proprietés extremales et la meilleure approximation des fonctions analytiques d'une variable réelle. Gauthier-Villars, Paris (1926)
15. Beylkin, G., Mohlenkamp, M.J.: Numerical operator calculus in higher dimensions. Proc. Natl. Acad. Sci. USA **99**, 10,246–10,251 (2002)
16. Beylkin, G., Mohlenkamp, M.J., Pérez, F.: Approximating a wavefunction as an unconstrained sum of Slater determinants. J. Math. Phys. **49**, 032,107 (2008)
17. Beylkin, G., Monzón, L.: On approximation of functions by exponential sums. Appl. Comput. Harmon. Anal. **19**, 17–48 (2005)
18. Beylkin, G., Monzón, L.: Approximation by exponential sums revisited. Appl. Comput. Harmon. Anal. **28**, 131–149 (2010)
19. Bini, D., Lotti, G., Romani, F.: Approximate solutions for the bilinear form computational problem. SIAM J. Comput. **9**, 692–697 (1980)
20. Björck, Å.: Numerical methods for least squares problems. SIAM, Philadelphia (1996)

W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series
in Computational Mathematics 42, DOI 10.1007/978-3-642-28027-6,

21. Börm, S.: Efficient numerical methods for non-local operators. EMS, Zürich (2010)
22. Börm, S., Grasedyck, L.: Hybrid cross approximation of integral operators. Numer. Math. **101**, 221–249 (2005)
23. Boys, S.F.: Electronic wave functions. I. A general method of calculation for stationary states of any molecular system. Proc. R. Soc. London Ser. A **200**, 542–554 (1950)
24. Brachat, J., Comon, P., Mourrain, B., Tsigaridas, E.: Symmetric tensor decomposition. Linear Algebra Appl. **433**, 1851–1872 (2010)
25. Braess, D.: Nonlinear approximation theory. Springer, Berlin (1986)
26. Braess, D.: Asymptotics for the approximation of wave functions by exponential-sums. J. Approx. Theory **83**, 93–103 (1995)
27. Braess, D., Hackbusch, W.: Approximation of $1/x$ by exponential sums in $[1,\infty)$. IMA J. Numer. Anal. **25**, 685–697 (2005)
28. Braess, D., Hackbusch, W.: On the efficient computation of high-dimensional integrals and the approximation by exponential sums. In: R.A. DeVore, A. Kunoth (eds.) Multiscale, nonlinear and adaptive approximation, pp. 39–74. Springer, Berlin (2009)
29. Bungartz, H.J., Griebel, M.: Sparse grids. Acta Numerica **13**, 147–269 (2004)
30. Carroll, J.D., Chang, J.J.: Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart-Young decomposition. Psychometrika **35**, 283–319 (1970)
31. Cattell, R.B.: Parallel proportional profiles and other principles for determining the choice of factors by rotation. Psychometrika **9**, 267–283 (1944)
32. Cayley, A.: Mémoire sur les hyperdéterminants. J. Reine Angew. Math. **30**, 1–37 (1846)
33. Cayley, A.: An introductory memoir on quantics. Philos. Trans. R. Soc. Lond. **144** (1854)
34. Chinnamsetty, S.R., Espig, M., Flad, H.J., Hackbusch, W.: Canonical tensor products as a generalization of Gaussian-type orbitals. Z. Phys. Chem. **224**, 681–694 (2010)
35. Chinnamsetty, S.R., Espig, M., Khoromskij, B., Hackbusch, W., Flad, H.J.: Tensor product approximation with optimal rank in quantum chemistry. J. Chem. Phys. **127**, 084,110 (2007)
36. Chinnamsetty, S.R., Luo, H., Hackbusch, W., Flad, H.J., Uschmajew, A.: Bridging the gap between quantum Monto Carlo and F12-methods. Chem. Phys. (2011). On-line published
37. Christoffel, E.B.: Über die Transformation der homogenen Differentialausdrücke zweiten Grades. J. Reine Angew. Math. **70**, 46–70 (1869)
38. Comon, P., ten Berge, J.M.F., De Lathauwer, L., Castaing, J.: Generic and typical ranks of multi-way arrays. Linear Algebra Appl. **430**, 2997–3007 (2009)
39. Comon, P., Golub, G.H., Lim, L.H., Mourrain, B.: Symmetric tensors and symmetric tensor rank. SIAM J. Matrix Anal. Appl. **30**, 1254–1279 (2008)
40. De Lathauwer, L.: Decompositions of a higher-order tensor in block terms - part II: definitions and uniqueness. SIAM J. Matrix Anal. Appl. **30** (2008)
41. De Lathauwer, L., De Moor, B., Vandewalle, J.: A multilinear singular value decomposition. SIAM J. Matrix Anal. Appl. **21**, 1253–1278 (2000)
42. De Lathauwer, L., De Moor, B., Vandewalle, J.: An introduction to independent component analysis. J. Chemometrics **14**, 123–149 (2000)
43. De Lathauwer, L., De Moor, B., Vandewalle, J.: On the best rank-1 and rank-(R_1,R_2,...,R_n) approximation of higher order tensors. SIAM J. Matrix Anal. Appl. **21**, 1324–1342 (2000)
44. De Silva, V., Lim, L.H.: Tensor rank and the ill-posedness of the best low-rank approximation problem. SIAM J. Matrix Anal. Appl. **30**, 1084–1127 (2008)
45. Defant, A., Floret, K.: Tensor methods and operator ideals. North-Holland, Amsterdam (1993)
46. DeVore, R.A., Lorentz, G.G.: Constructive approximation. Springer, Berlin (1993)
47. Dilworth, S.J., Kutzarova, D., Temlyakov, V.N.: Convergence of some greedy algorithms in Banach spaces. J. Fourier Anal. Appl. **8**, 489–505 (2002)
48. Dirac, P.A.M.: Note on exchange phenomena in the Thomas atom. Proc. Cambridge Phil. Soc. **26**, 376–385 (1930)
49. Dolgov, S., Khoromskij, B., Savostyanov, D.V.: Multidimensional Fourier transform in logarithmic complexity using QTT approximation. Preprint 18, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (2011)
50. Eckart, C., Young, G.: The approximation of one matrix by another of lower rank. Psychometrika **1**, 211–218 (1936)

51. Edelstein, M.: Weakly proximinal sets. J. Approx. Theory **18**, 1–8 (1976)
52. Espig, M.: Effiziente Bestapproximation mittels Summen von Elementartensoren in hohen Dimensionen. Dissertation, Universität Leipzig (2008)
53. Espig, M., Grasedyck, L., Hackbusch, W.: Black box low tensor-rank approximation using fiber-crosses. Constr. Approx. **30**, 557–597 (2009)
54. Espig, M., Hackbusch, W.: A regularized Newton method for the efficient approximation of tensors represented in the canonical tensor format. Preprint 78, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (2010)
55. Espig, M., Hackbusch, W., Litvinenko, A., Matthies, H.G., Zander, E.: Efficient analysis of high dimensional data in tensor formats. Preprint 62, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (2011)
56. Espig, M., Hackbusch, W., Rohwedder, T., Schneider, R.: Variational calculus with sums of elementary tensors of fixed rank. Numer. Math. (2012). To appear
57. Falcó, A., Hackbusch, W.: On minimal subspaces in tensor representations. Found. Comput. Math. (2011). To appear
58. Falcó, A., Nouy, A.: Proper generalized decomposition for nonlinear convex problems in tensor Banach spaces. Numer. Math. (2012). To appear
59. Feuersänger, C., Griebel, M.: Principal manifold learning by sparse grids. Computing **85**, 267–299 (2009)
60. Flad, H.J., Hackbusch, W., Khoromskij, B., Schneider, R.: Concept of data-sparse tensor-product approximation in many-particle modelling. In: V. Olshevsky, E.E. Tyrtyshnikov (eds.) Matrix methods - theory, algorithms, applications, pp. 313–343. World Scientific, Singapore (2010)
61. Flad, H.J., Hackbusch, W., Schneider, R.: Best N-term approximation in electronic structure calculations. I. One-electron reduced density matrix. M2AN **40**, 49–61 (2006)
62. Flad, H.J., Hackbusch, W., Schneider, R.: Best N-term approximation in electronic structure calculations. II. Jastrow factors. M2AN **41**, 261–279 (2007)
63. Flad, H.J., Khoromskij, B., Savostyanov, D.V., Tyrtyshnikov, E.E.: Verification of the cross 3D algorithm on quantum chemistry data. Contemp. Math. **23**, 329–344 (2008)
64. Floret, K.: Weakly compact sets, *Lect. Notes Math.*, vol. 119. Springer, Berlin (1980)
65. Frenkel, J.: Wave mechanics, advanced general theory. Clarendon Press, Oxford (1934)
66. Gavrilyuk, I.P., Hackbusch, W., Khoromskij, B.: $\mathcal{H}$-matrix approximation for the operator exponential with applications. Numer. Math. **92**, 83–111 (2002)
67. Gavrilyuk, I.P., Khoromskij, B.: Quantized-TT-Cayley transform for computing the dynamics and the spectrum of high-dimensional Hamiltonians. Comput. Meth. Appl. Math. **11**, 273–290 (2011)
68. Golub, G.H., Pereyra, V.: Separable nonlinear least squares: the variable projection method and its applications. Inverse Problems **19**, R1–R26 (2003)
69. Golub, G.H., Van Loan, C.F.: Matrix computations, 3rd edn. The Johns Hopkins University Press, Baltimore (1996)
70. Goreinov, S.A., Tyrtyshnikov, E.E.: The maximal-volume concept in approximation by low-rank matrices. Contemp. Math. **280**, 47–51 (2001)
71. Goreinov, S.A., Tyrtyshnikov, E.E., Zamarashkin, N.L.: A theory of pseudoskeleton approximations. Linear Algebra Appl. **261**, 1–22 (1997)
72. Grasedyck, L.: Existence and computation of a low Kronecker-rank approximant to the solution of a tensor system with tensor right-hand side. Computing **72**, 247–266 (2004)
73. Grasedyck, L.: Hierarchical singular value decomposition of tensors. SIAM J. Matrix Anal. Appl. **31**, 2029–2054 (2010)
74. Grasedyck, L.: Polynomial approximation in hierarchical Tucker format by vector-tensorization. Preprint 43, DFG-SPP 1324 (2010). http://www.dfg-spp1324.de
75. Grasedyck, L., Hackbusch, W.: An introduction to hierarchical ($\mathcal{H}$-)rank and TT-rank of tensors with examples. Comput. Meth. Appl. Math. **11**, 291–304 (2011)
76. Greub, W.H.: Multilinear algebra, 2nd edn. Springer, Berlin (1978)

77. Griebel, M.: Sparse grids and related approximation schemes for higher dimensional problems. In: L. Pardo, A. Pinkus, E. Süli, M.J. Todd (eds.) Foundations of computational mathematics (FoCM05), pp. 106–161. Cambridge University Press, Cambridge (2006)
78. Griebel, M., Kuo, F.Y., Sloan, I.H.: The smoothing effect of the ANOVA decomposition. J. Complexity **26**, 523–551 (2010)
79. Grothendieck, A.: Produits tensoriels topologiques et espaces nucléaires. Mem. Amer. Math. Soc. **16** (1955)
80. Grothendieck, A.: Résumé de la théorie métrique des produit tensoriels topologiques. Bol. Soc. Mat. São Paulo **8**, 1–79 (1956)
81. Hackbusch, W.: Iterative solution of large sparse system of equations. Springer, New York (1994)
82. Hackbusch, W.: Elliptic differential equations. Theory and numerical treatment, 2nd edn. Springer, Berlin (2003)
83. Hackbusch, W.: Multi-grid methods and applications, 2nd edn. Springer, Berlin (2003)
84. Hackbusch, W.: Entwicklungen nach Exponentialsummen. Techn. Bericht 4, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (2005)
85. Hackbusch, W.: Convolution of hp-functions on locally refined grids. IMA J. Numer. Anal. **29**, 960–985 (2009)
86. Hackbusch, W.: Hierarchische Matrizen - Algorithmen und Analysis. Springer, Berlin (2009)
87. Hackbusch, W.: Tensorisation of vectors and their efficient convolution. Numer. Math. **119**, 465–488 (2011)
88. Hackbusch, W., Khoromskij, B.: Low-rank Kronecker-product approximation to multi-dimensional nonlocal operators. Part I. Separable approximation of multi-variate functions. Computing **76**, 177–202 (2006)
89. Hackbusch, W., Khoromskij, B.: Low-rank Kronecker-product approximation to multi-dimensional nonlocal operators. Part II. HKT representation of certain operators. Computing **76**, 203–225 (2006)
90. Hackbusch, W., Khoromskij, B.: Tensor-product approximation to operators and functions in high dimensions. J. Complexity **23**, 697–714 (2007)
91. Hackbusch, W., Khoromskij, B., Sauter, S.A., Tyrtyshnikov, E.E.: Use of tensor formats in elliptic eigenvalue problems. Numer. Linear Algebra Appl. (2011). On-line published
92. Hackbusch, W., Khoromskij, B., Tyrtyshnikov, E.E.: Hierarchical Kronecker tensor-product approximations. J. Numer. Math. **13**, 119–156 (2005)
93. Hackbusch, W., Khoromskij, B., Tyrtyshnikov, E.E.: Approximate iterations for structured matrices. Numer. Math. **109**, 365–383 (2008)
94. Hackbusch, W., Kühn, S.: A new scheme for the tensor representation. J. Fourier Anal. Appl. **15**, 706–722 (2009)
95. Hamilton, W.R.: On quaternions, or on a new system of imaginaries in algebra. The London, Edinburgh and Dublin Philos. Mag. and J. of Science (3rd Series) **29**, 26–31 (1846)
96. Harshman, R.: Foundations of PARAFAC procedure: models and conditions for an "exploratory" multi-mode analysis. UCLA Working Papers in Phonetics **16**, 1–84 (1970)
97. Håstad, J.: Tensor rank is NP-complete. J. Algorithms **11**, 644–654 (1990)
98. Higham, N.J.: Functions of matrices, theory and computation. SIAM, Philadelphia (2008)
99. Hitchcock, F.L.: Multiple invariants and generalized rank of a p-way matrix or tensor. Journal of Mathematics and Physics **7**, 40–79 (1927)
100. Hitchcock, F.L.: The expression of a tensor or a polyadic as a sum of products. Journal of Mathematics and Physics **6**, 164–189 (1927)
101. Holmes, R.B.: A course on optimization and best approximation. Springer, Berlin (1980)
102. Holtz, S., Rohwedder, T., Schneider, R.: On manifolds of tensors of fixed TT-rank. Numer. Math. (2011). On-line published
103. Holtz, S., Rohwedder, T., Schneider, R.: The alternating linear scheme for tensor optimisation in the TT format. Preprint 71, DFG-SPP 1324 (2010). http://www.dfg-spp1324.de
104. Hsiao, G.C., Wendland, W.L.: Boundary integral equations. Springer, Berlin (2008)
105. Hübener, R., Nebendahl, V., Dür, W.: Concatenated tensor network states. New J. Phys. **12**, 025,004 (2010)

106. Ishteva, M., De Lathauwer, L., Absil, P.A., Van Huffel, S.: Differential-geometric Newton method for the best rank-$(R1, R2, R3)$ approximations of tensors. Numer. Algorithms **51**, 179–194 (2009)
107. Jemderson, H.V., Pukelsheim, F., Searle, S.R.: On the history of the Kronecker product. Linear Multilinear Algebra **14**, 113–120 (1983)
108. Johnson, W.B., Lindenstrauss, J.: Concepts in the geometry of Banach spaces. In: Handbook of the geometry of Banach spaces, vol. 1, pp. 1–84. North-Holland, Amsterdam (2001)
109. Karhunen, K.: Über lineare Methoden in der Wahrscheinlichkeitsrechnung. Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys. **37**, 1–79 (1947)
110. Kaup, W.: On Grassmannians associated with JB*-triples. Math. Z. **236**, 567–584 (2001)
111. Kazeev, V.A., Khoromskij, B.: On explicit QTT representation of Laplace and its inverse. Preprint 75, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (2010)
112. Keinert, F.: Uniform approximation to $|x|^\beta$ by Sinc functions. J. Approx. Theory **66**, 44–52 (1991)
113. Khoromskaia, V.: Computation of the Hartree-Fock exchange by the tensor-structured methods. Comput. Meth. Appl. Math. **10**, 204–218 (2010)
114. Khoromskaia, V., Khoromskij, B., Schneider, R.: QTT representation of the Hartree and exchange operators in electronic structure calculations. Comput. Meth. Appl. Math. **11**, 327–341 (2011)
115. Khoromskij, B.: Structured rank-$(r_1, ..., r_D)$ decomposition of function-related tensors in $\mathbb{R}^D$. Comput. Meth. Appl. Math. **6**, 194–220 (2006)
116. Khoromskij, B.: On tensor approximation of Green iterations for Kohn-Sham equations. Comput. Vis. Sci. **11**, 259–271 (2008)
117. Khoromskij, B.: Tensor-structured preconditioners and approximate inverse of elliptic operators in $\mathbb{R}^d$. Constr. Approx. **30**, 599–620 (2009)
118. Khoromskij, B.: $O(d \log N)$-quantics approximation of $N - d$ tensors in high-dimensional numerical modeling. Constr. Approx. (2011). To appear
119. Khoromskij, B., Khoromskaia, V.: Low rank Tucker-type tensor approximation to classical potentials. Cent. Eur. J. Math. **5**, 523–550 (2007)
120. Khoromskij, B., Khoromskaia, V.: Multigrid accelerated tensor approximation of function related multidimensional arrays. SIAM J. Sci. Comput. **31**, 3002–3026 (2009)
121. Khoromskij, B., Khoromskaia, V., Flad, H.J.: Numerical solution of the Hartree-Fock equation in multilevel tensor-structured format. SIAM J. Sci. Comput. **33**, 45–65 (2011)
122. Khoromskij, B., Oseledets, I.V.: Quantics-TT collocation approximation of parameter-dependent and stochastic elliptic PDEs. Comput. Meth. Appl. Math. **10**, 376–394 (2010)
123. Khoromskij, B., Sauter, S.A., Veit, A.: Fast quadrature techniques for retarded potentials based on TT/QTT tensor approximation. Comput. Meth. Appl. Math. **11**, 342–362 (2011)
124. Khoromskij, B., Schwab, C.: Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs. SIAM J. Sci. Comput. **33**, 364–385 (2011)
125. Knyazev, A.V.: Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method. SIAM J. Sci. Comput. **23**, 517–541 (2001)
126. Koch, O., Lubich, C.: Dynamical tensor approximation. SIAM J. Matrix Anal. Appl. **31**, 2360–2375 (2010)
127. Koch, W., Holthausen, M.C.: A chemist's guide to density functional theory. Wiley-VCH, Weinheim (2000)
128. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. SIAM Rev. **51**, 455–500 (2009)
129. Kolda, T.G., Sun, J.: Scalable tensor decompositions for multi-aspect data mining. In: 2008 Eighth IEEE international conference on data mining, pp. 363–372 (2008)
130. Kressner, D., Tobler, C.: Preconditioned low-rank methods for high-dimensional elliptic PDE eigenvalue problems. Comput. Meth. Appl. Math. **11**, 363–381 (2011)
131. Kressner, D., Tobler, C.: `htucker` - A MATLAB toolbox for tensors in hierarchical Tucker format. Tech. rep., Seminar for Applied Mathematics, ETH Zurich (2011)
132. Kreyszig, E.: Differentialgeometrie, 2nd edn. Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig (1968)

133. Kruskal, J.B.: Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. Linear Algebra Appl. **18**, 95–138 (1977)
134. Kruskal, J.B.: Rank, decomposition, and uniqueness for 3-way and N-way arrays. In: R. Coppi, S. Bolasco (eds.) Multiway data analysis, pp. 7–18. North-Holland, Amsterdam (1989)
135. Kutzelnigg, W.: Theory of the expansion of wave functions in a gaussian basis. Int. J. Quantum Chem. **51**, 447–463 (1994)
136. Landsberg, J.M., Qi, Y., Ye, K.: On the geometry of tensor network states. arXiv 1105.4449v1, math.AG (2011)
137. Langville, A.N., Stewart, W.J.: The Kronecker product and stochastic automata networks. J. Comput. Appl. Math. **167**, 429–447 (2004)
138. Lichtenberg, G., Eichler, A.: Multilinear algebraic Boolean modelling with tensor decomposition techniques. In: IFAC World Congress, vol. 18. IFAC (2011)
139. Light, W.A., Cheney, E.W.: Approximation theory in tensor product spaces, *Lect. Notes Math.*, vol. 1169. Springer, Berlin (1985)
140. Liu, J., Musialski, P., Wonka, P., Ye, J.: Tensor completion for estimating missing values in visual data. In: IEEE international conference on computer vision (2009)
141. Loève, M.: Probability theory II, 4th edn. Springer, New York (1978)
142. Lu, H., Plataniotis, K.N., Venetsanopoulos, A.N.: A survey of multilinear subspace learning for tensor data. Pattern Recognition **44**, 1540–1551 (2011)
143. Lubich, C.: On variational approximations in quantum molecular dynamics. Math. Comp. **74**, 765–779 (2005)
144. Lubich, C.: From quantum to classical molecular dynamics: reduced models and numerical analysis. EMS, Zürich (2008)
145. Meise, R., Vogt, D.: Introduction to functional analysis. Clarendon Press, Oxford (1997)
146. Melenk, J.M., Börm, S., Löhndorf, M.: Approximation of integral operators by variable-order interpolation. Numer. Math. **99**, 605–643 (2005)
147. Meyer, H.D., Gatti, F., Worth, G.A. (eds.): Multidimensional quantum dynamics. MCTDH theory and applications. Wiley-VCH, Weinheim (2009)
148. Mohlenkamp, M.J.: A center-of-mass principle for the multiparticle Schrödinger equation. J. Math. Phys. **51**, 022,112 (2010)
149. Mohlenkamp, M.J.: Musing on multilinear fitting. Linear Algebra Appl. (2011). To appear
150. Mohlenkamp, M.J.: Numerical implementation to approximate a wavefunction with an unconstrained sum of Slater determinants. Tech. rep., University Ohio (2011)
151. Mohlenkamp, M.J., Monzón, L.: Trigonometric identities and sums of separable functions. The Mathematical Intelligencer **27**, 65–69 (2005)
152. Oseledets, I.V.: A new tensor decomposition. Doklady Math. **80**, 495–496 (2009)
153. Oseledets, I.V.: Approximation of matrices using tensor decomposition. SIAM J. Matrix Anal. Appl. **31**, 2130–2145 (2010)
154. Oseledets, I.V.: DMRG approach to fast linear algebra in the TT-format. Comput. Meth. Appl. Math. **11**, 382–393 (2011)
155. Oseledets, I.V.: Tensor-train decomposition. SIAM J. Sci. Comput. **33**, 2295–2317 (2011)
156. Oseledets, I.V., Savostyanov, D.V., Tyrtyshnikov, E.E.: Linear algebra for tensor problems. Computing **85**, 169–188 (2009)
157. Oseledets, I.V., Tyrtyshnikov, E.E.: Approximate inversion of matrices in the process of solving a hypersingular integral equation. Comput. Math. Math. Phys. **45**, 302–313 (2005)
158. Oseledets, I.V., Tyrtyshnikov, E.E.: Tensor tree decomposition does not need a tree. Preprint 2009-08, RAS, Moskow (2009)
159. Oseledets, I.V., Tyrtyshnikov, E.E.: TT-cross approximation for multidimensional arrays. Linear Algebra Appl. **432**, 70–88 (2010)
160. Qi, L., Sun, W., Wang, Y.: Numerical multilinear algebra and its applications. Front. Math. China **2**, 501–526 (2007)
161. Quarteroni, A., Sacco, R., Saleri, F.: Numerical mathematics. Springer, New York (2000)
162. Remez, E.J.: Sur un procédé convergent d'approximations successives pour déterminer les polynômes d'approximation. Compt. Rend. Acad. Sc. **198**, 2063–2065 (1934)

163. Riesz, F., Sz.-Nagy, B.: Vorlesungen über Funktionalanalysis, 4th edn. VEB Deutscher Verlag der Wissenschaften, Berlin (1982)
164. Rivlin, T.J.: The Chebyshev polynomials. Wiley-Interscience, New York (1990)
165. Salmi, J., Richter, A., Koivunen, V.: Sequential unfolding SVD for tensors with applications in array signal processing. IEEE Trans. Signal Process. **57**, 4719–4733 (2009)
166. Sauter, S.A., Schwab, C.: Boundary element methods. Springer, Berlin (2011)
167. Schatten, R.: A theory of cross-spaces. University Press, Princeton (1950)
168. Schmidt, E.: Zur Theorie der linearen und nichtlinearen Integralgleichungen. I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener. Math. Ann. **63**, 433–476 (1907)
169. Schur, I.: Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. J. Reine Angew. Math. **141**, 1–28 (1911)
170. Schwab, C., Gittelson, C.J.: Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. Acta Numerica **20**, 291–467 (2011)
171. Schwab, C., Todor, R.A.: Karhunen-Loève approximation of random fields by generalized fast multipole methods. J. Comput. Phys. **217**, 100–122 (2006)
172. Simon, B.: Uniform crossnorms. Pacific J. Math. **46**, 555–560 (1973)
173. Smilde, A., Bro, R., Geladi, P.: Multi-way analysis. Applications in the chemical sciences. Wiley, West Sussex (2004)
174. Sprengel, F.: A class of periodic functions spaces and interpolation on sparse grids. Numer. Funct. Anal. Optim. **21**, 273–293 (2000)
175. Stegeman, A.: Degeneracy in Candecomp/Parafac explained for $p \times p \times 2$ arrays of rank $p + 1$ or higher. Psychometrika **71**, 483–501 (2006)
176. Stegeman, A., De Lathauwer, L.: A method to avoid diverging components in the CANDECOMP/PARAFAC model for generic $I \times J \times 2$ arrays. SIAM J. Matrix Anal. Appl. **30**, 1614–1638 (2009)
177. Stenger, F.: Numerical methods based of sinc and analytic functions. Springer, New York (1993)
178. Stoer, J.: Einführung in die Numerische Mathematik I, 8th edn. Springer, Berlin (1999)
179. Strassen, V.: Gaussian elimination is not optimal. Numer. Math. **13**, 354–356 (1969)
180. Strassen, V.: Rank and optimal computation of generic tensors. Linear Algebra Appl. **52**, 645–685 (1983)
181. Takatsuka, A., Ten-no, S., Hackbusch, W.: Minimax approximation for the decomposition of energy denominators in Laplace-transformed Møller-Plesset perturbation theories. J. Chem. Phys. **129**, 044,112 (2008)
182. Todor, R.A., Schwab, C.: Convergence rates for sparse chaos approximations of elliptic problems with stochastic coeffcients. IMA J. Numer. Anal. **27**, 232–261 (2007)
183. Trefethen, L.N.: Householder triangularization of a quasimatrix. IMA J. Numer. Anal. **30**, 887–897 (2010)
184. Tucker, L.R.: Some mathematical notes on three-mode factor analysis. Psychometrika **31**, 279–311 (1966)
185. Tyrtyshnikov, E.E.: Preservation of linear constraints in approximation of tensors. Numer. Math. Theory Methods Appl. **2**, 421–426 (2009)
186. Uschmajew, A.: Convex maximization problems on non-compact Stiefel manifolds with application to orthogonal tensor approximations. Numer. Math. **115**, 309–331 (2010)
187. Uschmajew, A.: Local convergence of the alternating least squares algorithm for canonical tensor approximation. SIAM J. Matrix Anal. Appl. (2012). To appear
188. Uschmajew, A.: Regularity of tensor product approximations to square integrable functions. Constr. Approx. **34**, 371–391 (2011)
189. Van Loan, C.F., Pitsianis, N.: Approximation with Kronecker products. In: M.S. Moonen, G.H. Golub (eds.) Linear algebra for large scale and real-time applications, *NATO Adv. Sci. Inst. Ser. E Appl. Sci.*, vol. 232, pp. 293–314. Kluwer Academic Publ., Dortrecht (1993)
190. Verstraete, F., Cirac, J.I.: Matrix product states represent ground states faithfully. Phys. Rev. B **73**, 094,423 (2006)
191. Vidal, G.: Efficient classical simulation of slightly entangled quantum computations. Phys. Rev. Letters **91**, 147,902 (2003)

192. Voigt, W.: Die fundamentalen physikalischen Eigenschaften der Krystalle in elementarer Darstellung. Veit & Comp., Leipzig (1898)
193. Wang, H., Ahuja, N.: Compact representation of multidimensional data using tensor rank-one decomposition. In: ICPR 2004 - Proceedings of the 17th International Conference on Pattern Recognition, vol. 1, pp. 44–47 (2004)
194. Wang, H., Thoss, M.: Multilayer formulation of the multiconfiguration time-dependent Hartree theory. J. Chem. Phys. **119**, 1289–1299 (2003)
195. Weidmann, J.: Lineare Operatoren in Hilberträumen, Teil 1. Teubner, Stuttgart (2000)
196. White, S.R.: Density matrix formulation for quantum renormalization groups. Phys. Rev. Letters **69**, 2863 (1992)
197. White, S.R., Martin, R.L.: Ab initio quantum chemistry using the density matrix renormalization group. J. Chem. Phys. **110**, 4127 (1999)
198. Yosida, K.: Functional analysis, 4th edn. Springer, Berlin (1974)
199. Yserentant, H.: Regularity and approximability of electronic wave functions, *Lect. Notes Math.*, vol. 2000. Springer, Berlin (2010)
200. Zehfuss, J.: Über eine gewisse Determinante. Z. für Math. und Phys. **3**, 298–301 (1858)
201. Zhang, T., Golub, G.H.: Rank-one approximation to high order tensors. SIAM J. Matrix Anal. Appl. **23**, 534–550 (2001)

List of authors involved in the references from above, but not placed as first author.

Absil, P.A. [106]
Ahuja, N. [193]
Auer, A.A. [12]
Bader, B.W. [128]
ten Bergen, J.M.F. [38]
Börm, S. [146]
Bro, R. [173]
Castaing, J. [38]
Chang, J.J. [30]
Cheney, E.W. [139]
Cirac, J.I. [190]
Comon, P. [24]
Davidson, E.R. [3]
De Lathauwer, L. [38, 106, 176]
De Moor, B. [41–43]
Dür, W. [105]
Eichler, A. [138]
Espig, M. [12, 34, 35]
Flad, H.J. [34–36, 121]
Floret, K. [45]
Gatti, F. [147]
Geladi, P. [173]
Gittelson, C.J. [170]
Golub, G.H. [39, 201]
Grasedyck, L. [6, 7, 22, 53]
Griebel, M. [29, 59]
Hackbusch, W. [12, 27, 28, 34–36, 53–57, 60–62, 66, 75, 92, 93, 181]
Holthausen, M.C. [127]
Khoromskaia, V. [119–121]
Khoromskij, B. [35, 49, 60, 63, 66, 67, 88–91, 111, 114]
Kluge, M. [7]
Koivunen, V. [165]
Kolda, T.G.: [4]
Kühn, S. [94]
Kuo, F.Y. [78]
Kutzarova, D. [47]
Lim, L.H. [39, 44]
Lindenstrauss, J. [108]
Litvinenko, A. [55]
Löhndorf, M. [146]
Lorentz, G.G. [46]
Lotti, G. [19]
Lubich, C. [126]
Luo, H. [36]
Martin, R.L. [197]
Matthies, H.G. [55]
Mohlenkamp, M.J. [15, 16]
Monzón, L. [17, 18, 151]
Mourrain, B. [24, 39]
Musialski, P. [140]
Nebendahl, V. [105]
Nouy, A. [58]
Oseledets, I.V. [122]
Peller, V. [1]
Pereyra, V. [68]
Pérez, F. [16]
Plataniotis, K.N. [142]
Qi, Y. [136]
Pitsianis, N. [189]
Pukelsheim, F. [107]
Richter, A. [165]
Rohwedder, T. [56, 102, 103]
Romani, F. [19]
Sauter, S.A. [91, 123]
Sacco, R. [161]
Saleri, F. [161]
Savotyanov, D.V. [49, 63, 156]
Schneider, R. [56, 60–62, 102, 103, 114]
Schwab, C. [124, 166, 182]
Searle, S.R. [107]
Sloan, I.H. [78]
Stewart, W.J. [137]
Sun, J. [129]
Sun, W. [160]
Sz.-Nagy, B. [163]
Temlyakov, V.N. [47]
ten Bergen, J.M.F. [38]
Ten-no, S. [181]
Thoss, M. [194]
Tobler, C. [130, 131]
Todor, R.A. [171]
Tsigaridas, E. [24]
Tyrtyshnikov, E.E. [63, 70, 71, 91–93, 156–159]
Uschmajew, A. [36]
Vandewalle, J. [41–43]
Van Huffel, S. [106]
Van Loan, C.F. [69]
Veit, A. [123]
Venetsanopoulos, A.N. [142]
Vogt, D. [145]
Wang, Y. [160]
Wendland, W.L. [104]
Wonka, P. [140]
Worth, G.H. [147]
Ye, J. [140]
Ye, K. [136]
Young, G. [50]
Zamarashkin, N.L. [71]
Zander, E. [55]

Index