

Claudia Zeilhofer

**Multi-dimensional B-spline Modeling
of Spatio-temporal Ionospheric Signals**

München 2008

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München



DGK Deutsche Geodätische Kommission
bei der Bayerischen Akademie der Wissenschaften

Reihe A

Theoretische Geodäsie

Heft Nr. 123

Claudia Zeilhofer

Multi-dimensional B-spline Modeling
of Spatio-temporal Ionospheric Signals

München 2008

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München

ISSN 0938-2836

ISBN 3 7696 8203 3

Adresse der Deutschen Geodätischen Kommission:



Deutsche Geodätische Kommission

Alfons-Goppel-Straße 11 • D – 80 539 München
Telefon +49 – 89 – 23 031 1113 • Telefax +49 – 89 – 23 031 - 1283/ - 1100
e-mail hornik@dgfi.badw.de • <http://www.dgk.badw.de>

Adressen der Autorin /
Addresses of the author

Dipl.-Math. Claudia Zeilhofer
Deutsches Geodätisches Forschungsinstitut
Alfons-Goppel-Straße 11
D – 80 539 München
zeilhofer@dgfi.badw.de

Diese Publikation ist als pdf-Dokument veröffentlicht im Internet unter der Adresse /

This volume is published in the internet

<http://dgk.badw.de>

© 2008 Deutsche Geodätische Kommission, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen

Abstract

Nowadays, ionospheric phenomena can be monitored globally from data of various satellite missions with an unprecedented accuracy. In many applications, like modeling the equatorial anomaly, regional processes are of special interest. To study regional phenomena multi-dimensional mathematical tools may be applied. In this report a regional multi-dimensional representation based on quadratic B-splines is derived and applied to spatio-temporal data sets.

At first the theory of a multiscale analysis is introduced, i.e. the decomposition of signals into a smoother version and a number of detail signals for the 1-dimensional case. To develop a multiscale analysis we introduce B-splines as scaling functions. The corresponding base functions, i.e. the B-spline wavelets can be computed via the so-called two-scale relations from the scaling functions. Series coefficients of the B-spline expansion can be computed by evaluating inner products using the dual base functions or by applying parameter estimation procedures. To generalize a 1-dimensional B-spline representation to the multi-dimensional case tensor product techniques are used. A data compression algorithm is derived from the multiscale analysis, i.e. we gain a tool to handle huge digital data sets.

In our applications we decompose ionospheric functions, e.g. the electron density or the vertical total electron content, into a reference part and an unknown correction term. We apply our approach regionally to the correction term, i.e. we expand it in a multi-dimensional series expansion in terms of B-splines. Since our observations are located rather unbalanced with respect to space and time finer structures are modelable only in regions with a sufficient number of observation sites. The multiscale analysis derived from the wavelet analysis allows monitoring the ionospheric signals at different resolution levels.

Zusammenfassung

Ionosphärische Phänomene können heutzutage mit Hilfe moderner Satellitenmissionen mit hoher Genauigkeit global beobachtet werden. In vielen Anwendungen, wie zum Beispiel der Erforschung und Modellierung der äquatorialen Anomalie, sind indes regionale Vorgänge von besonderem Interesse. Zur Untersuchung des raumzeitlichen Verhaltens derartiger regionaler Phänomene lassen sich verschiedene mehr-dimensionale mathematische Methoden nutzen. In dieser Arbeit wird eine regionale mehr-dimensionale Darstellung basierend auf quadratischen B-splines entwickelt und auf raum-zeitliche Datensätze angewandt.

Die Theorie der Multiskalen-Analyse, d.h. die Zerlegung eines Signals in eine geglättete Version und in eine bestimmte Anzahl von Detailsignalen, wird zunächst für den 1-dimensionalen Fall betrachtet. Die quadratischen B-Splines werden dabei als Skalierungsfunktionen der Multiskalen-Analyse eingeführt. Die dazugehörigen Basisfunktionen der Detailsignale, d.h. die B-spline Wavelets, können mit Hilfe der sogenannten Zwei-Skalen-Gleichungen aus den Skalierungsfunktionen berechnet werden. Die Reihenkoeffizienten der B-Spline-Entwicklungen kann man aus der Auswertung von Skalarprodukten, die die dualen Basisfunktionen enthalten, oder durch die Anwendung von Methoden der Parameterschätzung bestimmen. Um die ein-dimensionale Modellierung zu einer mehr-dimensionalen Modellierung zu verallgemeinern, wird in dieser Arbeit die Tensorprodukt-Darstellung genutzt. Die Multiskalen-Analyse liefert zudem einen äußerst effizienten Algorithmus zur Datenkompression, d.h. wir erhalten ein hervorragend geeignetes Verfahren zur Behandlung von sehr großen digitalen Datensätzen.

In den in dieser Arbeit vorgestellten Anwendungen werden ionosphärische Funktionen, z.B. die Elektronendichte oder der vertikale absolute Elektroneninhalt (engl.: vertical total electron content), in einen Referenzteil und einen unbekanntem Korrekturterm zerlegt. Letzterer wird regional als mehr-dimensionale Reihe in B-Splines entwickelt. Da die Beobachtungen i.Allg. jedoch sowohl räumlich als auch zeitlich unregelmäßig verteilt sind, kann man feine ionosphärische Strukturen nur in Regionen mit einer genügend großen Anzahl von Beobachtungsstationen modellieren. Die Multiskalen-Analyse, die man aus der Wavelet-Analyse ableitet, bedeutet dabei eine Möglichkeit ionosphärische Signale in verschiedenen Auflösungsstufen zu betrachten.

Preface

Nowadays, satellite systems have become almost indispensable in everyday life, e.g. for navigation and telecommunication. The ionosphere is defined as a thick shell of electrons and ions, which envelopes the Earth from about 60 to 1200 km height. A key point in correcting electromagnetic satellite measurements for ionospheric disturbances is the knowledge of the electron density. Hence, ionosphere modelling constitutes a very important task in many disciplines of geo-sciences at present and in the future.

This monograph has been developed during the beginning of my work at Deutsches Geodätisches Forschungsinstitut (DGFI). My thanks go to all members of DGFI. I would like to express my gratitude to PD Dr.-Ing. habil. Michael Schmidt for the setting of the problem. Especially I am grateful to him for the patience and constant attention to the work.

I would like to thank Prof. Dr. Walter Richert from the University of Munich (Ludwig-Maximilians-Universität) for many useful discussions.

May 2008

Claudia Zeilhofer

Contents

1	Introduction	11
2	Theory	14
2.1	Multiscale analysis	14
2.1.1	Scaling functions and scaling spaces	14
2.1.2	Wavelet functions and detail spaces	15
2.1.3	Two-scale relations	21
2.1.4	Decomposition and reconstruction	25
2.1.5	Multiscale analysis on a bounded interval	27
2.2	Definition of B-splines and some properties	29
2.3	1-dimensional B-spline model	32
2.3.1	Scaling functions and detail spaces	32
2.3.2	First two-scale relation for B-splines	35
2.3.3	Wavelets and detail spaces - the second two-scale relation	36
2.3.4	Construction of a decomposition relation and a reconstruction relation	37
2.3.5	Decomposition and reconstruction for the 1-dimensional B-spline model	38
2.4	n-dimensional B-spline model	41
2.4.1	n-dimensional tensor product scaling and wavelet functions	41
2.4.2	Decomposition and reconstruction in 4 dimensions	42
3	Application	47
3.1	Observations	47
3.1.1	Electron density from GNSS	47
3.1.2	VTEC from occultation measurements	48
3.1.3	VTEC from altimetry	48
3.1.4	Ionosonde measurements	49
3.2	Different options	49
3.2.1	2-dimensional approach	49
3.2.2	3-dimensional approach	51
3.2.3	4-dimensional approach	52
3.3	Selected applications	53
3.3.1	Electron density B-spline model for IRI	53
3.3.2	STEC model for GNSS observations	56
3.3.3	3-dimensional COSMIC VTEC modeling	69
3.4	Combinations	76
4	Summary and conclusions	78
A	L^2-theory of Fourier-series	80

B	Fourier transform in \mathbb{R}^n	81
C	Endpoint-interpolating quadratic B-spline wavelets	82
D	Statistical quantities	84

List of Figures

2.1	Semiorthogonal MSA.	28
2.2	B-spline scaling functions.	34
3.1	Occultation geometry.	48
3.2	Input strategies for the B-spline model.	53
3.3	Electron density estimation.	55
3.4	Height and temporal profiles.	56
3.5	IGS and regional GPS stations.	59
3.6	Influence of STEC observations on B-spline scaling coefficients.	60
3.7	Main steps of STEC modeling.	62
3.8	Location of the points P_1 and P_2	63
3.9	Ionosphere and plasmasphere.	63
3.10	Receiver positions, satellite orbits and selected signal-paths.	65
3.11	Results of STEC model.	67
3.12	Height profile for the STEC model.	68
3.13	Influence of VTEC observations on B-spline scaling coefficients.	70
3.14	Snapshots of residual VTEC estimation.	72
3.15	Observations for residual VTEC estimation.	72
3.16	COSMIC VTEC estimation.	73
3.17	Decomposition algorithm.	74
3.18	Reconstruction algorithm.	75
3.19	Results of data compression.	76
4.1	Digisonde network.	79
C.1	B-spline wavelet functions.	83

Chapter 1

Introduction

Since the beginning of the 1980s, the subject of "wavelet analysis" has drawn much attention from both, the mathematicians and the engineers. Particularly in geophysical problems engineers detect that the application of the classical methods, e.g., the Fourier transform or the Gabor transform, yields not in all cases satisfying results. Detailed introductions to the wavelet analysis can, e.g., be found in LOUIS ET AL. (1997) or CHUI (1992).

If we apply the Fourier analysis, for instance, the method describes the input signal in terms of frequency components. Fourier analysis is able to pick out the exact frequencies but it is unable to detect time-dependencies, i.e. it does not yield satisfying results for signals with time-varying amplitudes, phases and/or frequencies. There have been developed various methods such as the "windowed Fourier transform" to increase the applicability. Although these methods have achieved some success, they suffer from other deficiencies. Windowed transforms can localize simultaneously in time and in frequency, but the localization property in each dimension remains fixed. Only a narrow time window is needed to examine the high-frequency content while wide windows allow the investigation of low-frequency components. The most important advantage of the wavelet theory over the other methods is the time-frequency localization, since the localization in time and frequency is adapted automatically. With this property of wavelets an efficient representation and analysis of functions with discontinuities in derivatives, sharp spikes and discontinuities of the function itself is also possible. Thus, wavelet methods are useful in signal representation for a much broader class of functions; see OGDEN (1997).

Wavelets are intrinsically connected to the principle of a multiscale analysis (MSA). Similar to a microscope, with the MSA we can examine signals under different resolutions, i.e. levels. For lower level we can filter out coarse structures of the signal, while for higher levels we can examine the finer structures of the signal. This is called a "zoom-in, zoom-out" property. The MSA using wavelets is the analogon to the frequency analysis using Fourier decomposition. Basically, with the MSA it is possible to examine the features of a signal of any size by adjusting the level parameter - also called scaling parameter - in the analysis; see OGDEN (1997) and SCHMIDT (2001).

The focus of application of the wavelet analysis in geodetic problems is the adapted time-frequency localization, because, in general, phenomena with time- and space-dependent frequencies are investigated. Examples are the determination of the gravity field of the Earth as an application in physical geodesy or the analysis of the content of a digital image as an application in photogrammetry.

Nowadays, data from various satellite missions can be used to study geophysical phenomena globally, e.g. in climatology or oceanography. However, in many applications regional processes are of special interest, such as the ionospheric equatorial anomaly in space weather physics or sea level change in regions of post glacier rebound. In such cases appropriate multi-dimensional regional mathematical tools have to be chosen.

Global geophysical signals are traditionally represented by spherical harmonic expansions. For regional investigations, however, compactly supported base functions are more appropriate. The MSA is an appropriate tool for the consideration of huge data sets of modern measurement systems because its basic feature is to split a given input signal into a smoothed version of a specific resolution level and a certain number of detail signals by successive low-pass filtering. Each detail signal represents a band-passed filtered version of the input signal, related to a specific frequency band (resolution level). The higher the resolution level, the finer the detected structures. Since observations are frequently located rather unbalanced, finer structures of the signal under investigation are modelable just in regions with a sufficient number of observations.

As there are many applications of the Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS), the study of the most important effects on radio signal propagation has become more and more important. One of the main effects is caused by the ionosphere, i.e. the ionized layer of the atmosphere; it causes a delay in the electromagnetic signals that journey through the layer. The ionization of several molecular species in this layer is caused by the ultraviolet and X radiation emitted by the Sun. The maximal ionization in general takes place in a height of 200 to 400 km. The number of electrons is considered to be approximately equal to the ions. Hence there is the maximum of the electron density in these heights. Since the ionization depends on the activity of the Sun, there are diurnal variations. In this report, we consider the ionosphere to be a layer from 100 km to 1200 km height above the Earth. In the literature there can be found various height values. For more details see, e.g., GARCÍA-FERNÁNDEZ (2004).

In the 1-dimensional case a signal given within a finite time interval, can be modeled appropriately, for instance, by a series expansion in terms of endpoint-interpolating quadratic B-splines; for more details see e.g. STOLLNITZ ET AL. (1995). Furthermore, B-splines can favorably be used as scaling functions to establish a MSA. The corresponding base functions, i.e. the B-spline wavelet functions are computable via the so-called two-scale relations from the scaling functions. The series coefficients of the B-spline expansion can be computed by evaluating inner products using the dual base functions or by applying parameter estimation procedures, e.g. the least-squares method. That way data gaps can be handled appropriately by prior information. Data compression (reduction) and de-noising techniques can be applied; see e.g. SCHMIDT ET AL. (2007b).

Tensor-product techniques are used to generalize the 1-dimensional B-spline representation to the multi-dimensional case. The 2-dimensional approach, well-known in digital image processing can, e.g., also be applied to model ionospheric signals given on spherical regions. Spatial density distributions, such as the ionospheric electron density are examples for a 3-dimensional representation. Temporal variations of such a density distribution can be modeled as a 4-dimensional approach by introducing a fourth B-spline expansion for the time variable; for more details see, e.g., LYCHE AND SCHUMAKER (2001) and SCHMIDT ET AL. (2007b).

As already mentioned before GPS has become a promising tool for monitoring the electron density of the ionosphere in the last years. To determine the slant total electron content (STEC), i.e. the integral of the electron density along the ray-path between the transmitting satellite and the receiver, dual-frequency GPS observations can be used. By considering measurements from space-borne GPS receivers flying on low-Earth-orbiting (LEO) satellites such as the COSMIC/FORMOSAT-3 (Constellation Observing System for Meteorology, Ionosphere and Climate/Taiwan's FORMOSA SATellite Mission #3) six satellite constellation mission, CHAMP (Challenging Minisatellite Payload) or GRACE (Gravity Recovery And Climate Experiment) the insensitivity of ground-based GPS observations to the radial geometry of the ionosphere can be overcome. For more details on these satellite missions see, e.g. ROCKEN ET AL. (2000), REIGBER ET AL. (2000) or BEYERLE ET AL. (2005).

Various approaches can be used to establish a multi-dimensional model of the electron density. RIUS ET AL. (1997), for instance, combine GPS occultation data with ground-based GPS data to perform 4-dimensional ionospheric tomography with a substantial level of vertical resolution. Cells of constant electron density with a size adapted to the resolution of terrestrial data are used by HERNÁNDEZ-PAJARES ET AL. (1999). GARCÍA-FERNÁNDEZ (2004) presents an overview about various techniques for modeling the electron density is given. A model built by a combination of ground-based and space-based measurements can be found in DETTMERING (2003). In contrast to tomographic models it includes theoretical assumptions about the vertical structure of the ionosphere.

It is always common to subtract a reference model (background model) from the input data and to work with the residual data. In this report we use the International Reference Ionosphere (IRI) as reference model and we model corrections to the reference model. The IRI is an international project developed and improved by the Committee On SPace Research (COSPAR) and the International Union of Radio Science (URSI). In the late sixties a working group produced an empirical standard model of the ionosphere, based on all available data sources. Several improved editions of the model have been released from 1969 to present. For given location and time IRI describes, e.g., the electron density, electron temperature, ion temperature, and ion composition in the altitude range from about 50 km to about 2000 km; and also the electron content (see: <http://modelweb.gsfc.nasa.gov/ionos/iri.html>). For more details on the IRI see BILITZA (2001).

In this report a regional multi-dimensional representation based on quadratic B-splines is derived and applied to spatio-temporal ionospheric data sets. The report is organized as follows: In chapter 2 the theory is intro-

duced. In section 2.1 the definition of the 1-dimensional MSA is given. First the scaling functions and the scaling spaces are defined, then we introduce the detail spaces and prove the existence of orthogonal wavelets that span the detail spaces. We establish the two-scale relations wherein the two-scale coefficients play an important role. Later we generalize the concept to semiorthogonal wavelets and derive decomposition relations and a reconstruction procedure. Via the decomposition relations we perform a decomposition of a signal into a smoothed signal and a number of corresponding detail signals; from the smoothed signal and the corresponding details signals we can reconstruct the original signal via the reconstruction relation. In section 2.2 we introduce B-splines and give some properties of these functions. In section 2.3 the theory of the MSA is applied to the B-spline functions, i.e. we use the B-splines as scaling functions and construct semiorthogonal wavelet functions. With the two-scale coefficients we gain the decomposition and the reconstruction relations for the B-spline model. Finally in section 2.4 we extend the 1-dimensional B-spline model to the n-dimensional case.

Chapter 3 covers several applications. Section 3.1 gives an overview of ionospheric observations. We will consider the STEC from GPS, the electron density from occultation measurements, the vertical total electron content (VTEC) from altimetry and terrestrial measurements. The different options of our B-spline approach are repeated and summarized in section 3.2. In section 3.3 we study the evaluation of the observations in detail. In section 3.3.1 a simple application is considered: we model the electron density from IRI with our B-spline approach. Section 3.3.2 covers a more complicated application: we model the STEC for simulated GPS observations. The last application in section 3.3.3 is similar to the first application, but here we model the VTEC derived from COSMIC observations. The combination of different types of input data is treated in section 3.4.

Chapter 2

Theory

2.1 Multiscale analysis

In this first section we introduce the concept of a multiscale analysis and we gain the first two-scale relation; all the theory can be found in LOUIS ET AL. (1997), we made some changes in the notations.

For our investigations we will use the translated and dilated versions of a function $\phi \in L^2(\mathbb{R})$ as basis functions, these functions are defined in the following definition.

Definition 2.1. *The translated and dilated versions of a function $\phi \in L^2(\mathbb{R})$ are defined via*

$$\phi_{j,k}(x) := 2^{j/2} \phi(2^j x - k) \quad j, k \in \mathbb{Z}.$$

Definition 2.2. *The sequence space $l^2(\mathbb{Z})$ is defined as*

$$l^2(\mathbb{Z}) = \{(u_i)_{i \in \mathbb{Z}} \mid \sum_{i \in \mathbb{Z}} |u_i|^2 < \infty\}$$

wherein $u_i \in \mathbb{C}$.

2.1.1 Scaling functions and scaling spaces

Definition 2.3. *A multiscale analysis (MSA) of $L^2(\mathbb{R})$ is an increasing sequence of closed subspaces V_j of $L^2(\mathbb{R})$ fulfilling the following properties:*

$$\{0\} \dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots \subset L^2(\mathbb{R}) \quad (2.1)$$

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j}^{L^2(\mathbb{R})} = L^2(\mathbb{R}) \quad (2.2)$$

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\} \quad (2.3)$$

$$f(x) \in V_j \Leftrightarrow f(2^{-j}x) \in V_0, \quad j \in \mathbb{Z}. \quad (2.4)$$

Definition 2.4. *A function $\phi \in L^2(\mathbb{R})$ is called a **scaling function**, if the so-called **scaling spaces** $V_j := \overline{\text{span}\{\phi_{j,k} \mid k \in \mathbb{Z}\}}^{L^2(\mathbb{R})}$ with $\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k)$ (i.e. V_j are the closed subspaces generated by the translated and dilated versions of ϕ) satisfy the conditions (2.1), (2.2) and (2.4). Moreover the translates $\phi(x - k)$, $k \in \mathbb{Z}$, must form a Riesz-basis of V_0 , i.e.*

$$V_0 = \overline{\text{span}\{\phi(\cdot - k) \mid k \in \mathbb{Z}\}}^{L^2(\mathbb{R})} \quad (2.5)$$

and

$$A \sum_{k \in \mathbb{Z}} c_k^2 \leq \left\| \sum_{k \in \mathbb{Z}} c_k \phi(\cdot - k) \right\|_{L^2(\mathbb{R})}^2 \leq B \sum_{k \in \mathbb{Z}} c_k^2 \quad (2.6)$$

for all $\{c_k\}_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$; where A and B are positive constants.

We say the scaling function ϕ generates a MSA $\{V_j\}_{j \in \mathbb{Z}}$ of $L^2(\mathbb{R})$.

Remarks:

- Conditions (2.1) and (2.2) are satisfied by many families $\{V_j\}_{j \in \mathbb{Z}}$. Property (2.4) is the special feature of a MSA: all the spaces V_j are scaled versions of the central space V_0 , which is spanned by the translations of the scaling function ϕ . For $j \rightarrow \infty$ the spaces V_j contain the smaller structures, for $j \rightarrow -\infty$ the functions in V_j are dilated, i.e. their details are coarser. We can state this more precisely with the limits

$$\lim_{j \rightarrow \infty} \|P_j f - f\|_{L^2(\mathbb{R})} = 0 \text{ and}$$

$$\lim_{j \rightarrow -\infty} \|P_j f\|_{L^2(\mathbb{R})} = 0$$

wherein P_j denotes the orthogonal projector onto V_j .

- If a scaling function $\phi \in L^2(\mathbb{R})$ generates a MSA $\{V_j\}_{j \in \mathbb{Z}}$ of $L^2(\mathbb{R})$, then the sequence of nested subspaces necessarily satisfies the property (2.3); see CHUI (1992), p. 121.
- The relation (2.5) implies V_0 is invariant under integer translation

$$f \in V_0 \Leftrightarrow f(\cdot - k) \in V_0 \text{ for } k \in \mathbb{Z};$$

with (2.4) it follows that

$$f \in V_j \Leftrightarrow f(\cdot - 2^{-j}k) \in V_j \text{ for } k \in \mathbb{Z}.$$

- The functions $\phi_{j,k}$ all have the same L^2 -norm $\|\phi_{j,k}\|_{L^2(\mathbb{R})} = \|\phi\|_{L^2(\mathbb{R})}$.
- If $\phi_{0,k}$, $k \in \mathbb{Z}$, form a Riesz basis of V_0 then it follows, that $\phi_{j,k}$, $k \in \mathbb{Z}$, forms a Riesz basis of V_j for a $j \in \mathbb{Z}$ with the same Riesz bounds A and B ; see CHUI (1992).

Lemma 2.5. *The scaling function ϕ satisfies a scaling equation, i.e. there is a sequence $\{p_k\}_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$ with*

$$\phi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} p_k \phi(2x - k), \quad x \in \mathbb{R}. \quad (2.7)$$

*This equation is called first **two-scale relation** or **refinement equation**.*

Proof: Equation (2.7) follows from $\phi \in V_0 \subset V_1 = \overline{\text{span}\{2^{1/2} \phi(2x - k) \mid k \in \mathbb{Z}\}}^{L^2(\mathbb{R})}$; see e.g. CHUI (1992). \square

2.1.2 Wavelet functions and detail spaces

In this section we give the definition of the detail spaces and we will introduce the second two-scale relation. Here we restrict our investigations to orthogonal spaces, but in section 2.1.3 we will consider the more general concepts of biorthogonal and semiorthogonal spaces for more flexibility.

Definition 2.6. *For every $j \in \mathbb{Z}$ we define the orthogonal **detail space** W_{j-1} to be the orthogonal complement of V_{j-1} in V_j , i.e.*

$$V_j = V_{j-1} \oplus W_{j-1}, \quad V_{j-1} \perp W_{j-1} \quad (2.8)$$

and the operators P_j and Q_j to be the orthogonal projectors of $L^2(\mathbb{R})$ on V_j and W_j , respectively,

$$P_j = P_{j-1} + Q_{j-1}. \quad (2.9)$$

From (2.8) it follows

$$V_j = \bigoplus_{i \leq j-1} W_i, \quad (2.10)$$

where all these subspaces are orthogonal and so

$$L^2(\mathbb{R}) = \bigoplus_{i \in \mathbb{Z}} W_i. \quad (2.11)$$

The spaces W_j inherit the scaling property of V_j (see equation (2.4))

$$f(x) \in W_j \Leftrightarrow f(2^{-j}x) \in W_0, \quad j \in \mathbb{Z}.$$

According to equations (2.11) and (2.10) a function $f \in L^2(\mathbb{R})$ can be decomposed

$$f = \sum_{i \in \mathbb{Z}} Q_i f = \sum_{i \leq j-1} Q_i f + \sum_{i \geq j} Q_i f \quad (2.12)$$

$$= P_j f + \sum_{i=j}^{\infty} Q_i f. \quad (2.13)$$

This equality justifies the name MSA. $P_j f$ represents f on scale j , which corresponds to an application of a low pass filter that has an increasing bandwidth for increasing j . The remaining high-frequency part of f is split up into its frequency bands $Q_i f$, $j \leq i \leq \infty$, so $Q_i f$ only contains the details which distinguish $P_{i+1} f$ from $P_i f$, $Q_i = P_{i+1} - P_i$, see equation (2.9). Because of the uniqueness of the orthogonal projection and equation (2.10) we have $\sum_{i \leq j-1} Q_i = P_j$; see LOUIS ET AL. (1997).

The decompositions of signals $f_j \in V_j$ will be of importance and will appear later.

Now we will give the definition of a wavelet.

Definition 2.7. A function $\psi \in L^2(\mathbb{R})$ which satisfies the **admissibility condition**

$$0 < c_\psi := 2\pi \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (2.14)$$

is called a **wavelet**.

Let $\psi_{j,k}(x) := 2^{j/2} \psi(2^j x - k)$, $j, k \in \mathbb{Z}$, be the translated and dilated versions of ψ .

A wavelet ψ is called an **orthogonal wavelet** (in fact the wavelet is even orthonormal) if

$$\langle \psi_{j,k}, \psi_{i,l} \rangle_{L^2(\mathbb{R})} = \delta_{j,i} \delta_{k,l}, \quad j, i, k, l \in \mathbb{Z}.$$

An immediate consequence of the admissibility condition is the following: for a wavelet $\psi \in L^1(\mathbb{R})$: if $\psi \in L^1(\mathbb{R})$ then $\hat{\psi}$ is continuous. The equation (2.14) can then only be fulfilled if $\hat{\psi}(0) = 0$ and we get the property

$$\int_{\mathbb{R}} \psi(x) dx = 0 \quad (2.15)$$

see e.g. DAUBECHIES (1992). The name wavelet comes from equation (2.15), it means "small wave".

Now we want to prove the main result of this chapter:

Theorem 2.8. To every MSA there exists a wavelet ψ whose translated and dilated versions

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k) \quad j, k \in \mathbb{Z}$$

generate an orthonormal basis (ONB) of the space W_j for fixed $j \in \mathbb{Z}$. Further the wavelet can be explicitly constructed from the scaling function.

We first state the analogon to Lemma 2.5 for the wavelets.

Lemma 2.9. A wavelet function ψ whose translated and dilated versions $\psi_{0,k}$, $k \in \mathbb{Z}$ generate the detail space W_0 satisfies a scaling equation, i.e. there is a sequence $\{q_k\}_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$ with

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} q_k \phi(2x - k), \quad x \in \mathbb{R}. \quad (2.16)$$

This equation is the analogon to equation (2.7) and is also called **second two-scale relation** or **refinement equation**.

Proof: Equation (2.16) follows from $\psi \in W_0 \subset V_1 = \overline{\text{span}\{2^{1/2} \phi(2x - k) \mid k \in \mathbb{Z}\}}^{L^2(\mathbb{R})}$; see e.g. CHUI (1992). \square

We now want to prove Theorem 2.8. First we give an overview of the proof:

1. We state that for every scaling function ϕ there exists a function $\tilde{\phi}$ s.t. the translates $\tilde{\phi}_{0,k}$ generate an ONB of V_0 (Theorem 2.11).
2. We prove that under special assumptions a function $\phi \in L^2(\mathbb{R})$ generates a MSA of $L^2(\mathbb{R})$ (Theorem 2.13).
3. We prove that to an orthogonal scaling function that generates a MSA there exists a wavelet ψ (Theorem 2.15).
4. The results of the Theorems 2.11, 2.13 and 2.15 prove Theorem 2.8.

Definition 2.10. A scaling function $\phi \in L^2(\mathbb{R})$ generates an orthonormal system in V_0 if $\{\phi_{0,k} \mid k \in \mathbb{Z}\}$ is an orthonormal system, i.e.

$$\langle \phi(\cdot - k), \phi(\cdot - n) \rangle_{L^2(\mathbb{R})} = \delta_{k,n}. \quad (2.17)$$

If the scaling function ϕ of a MSA does not generate an ONB of V_0 we can orthogonalize the basis $\{\phi_{0,k} \mid k \in \mathbb{Z}\}$ in the following sense.

Theorem 2.11. Let $\phi \in L^2(\mathbb{R})$ be a scaling function and let positive constants A and B exist, where

$$A \leq \sum_{n \in \mathbb{Z}} |\hat{\phi}(w + 2\pi n)|^2 \leq B \text{ a.e.} \quad (2.18)$$

Then $\{\tilde{\phi}(x - k) \mid k \in \mathbb{Z}\}$ is an ONB of V_0 with

$$\hat{\tilde{\phi}}(w) = \frac{1}{\sqrt{2\pi}} \frac{\hat{\phi}(w)}{\sqrt{\sum_{n \in \mathbb{Z}} |\hat{\phi}(w + 2\pi n)|^2}}. \quad (2.19)$$

Proof: See LOUIS ET AL. (1997). □

Remark:

To each scaling function ϕ Theorem 2.11 guarantees the existence of a scaling function $\tilde{\phi}$ such that

$$\tilde{\phi}_{j,k}(x) = 2^{j/2} \tilde{\phi}(2^j x - k), \quad x \in \mathbb{R}, j, k \in \mathbb{Z}$$

generates an ONB of V_j .

The spaces $\{V_j\}_{j \in \mathbb{Z}}$ form a MSA if and only if they span the entire space $L^2(\mathbb{R})$, i.e. if $\overline{\bigcup_{j \in \mathbb{Z}} V_j}^{L^2(\mathbb{R})} = L^2(\mathbb{R})$ or equivalent, if for $f \in L^2(\mathbb{R})$

$$\lim_{j \rightarrow \infty} \|P_j f - f\|_{L^2(\mathbb{R})} = 0 \quad (2.20)$$

with

$$P_j f = \sum_{k \in \mathbb{Z}} \langle f, \tilde{\phi}_{j,k} \rangle_{L^2(\mathbb{R})} \tilde{\phi}_{j,k}$$

the orthogonal projection on V_j .

For the proof of theorem 2.13 we will use the following:

Theorem 2.12 (Pythagorean theorem). Let $\{x_1, x_2, \dots, x_n\}$ be an orthonormal system (ONS) in an inner product space V . Then for all $x \in V$,

$$\|x\|^2 = \sum_{k=1}^n |(x_k, x)|^2 + \|x - \sum_{k=1}^n (x_k, x) x_k\|^2. \quad (2.21)$$

Proof. See REED AND SIMON (1980). □

Theorem 2.13. *Let a function $\phi \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ be given that satisfies equation (2.18) and has a non-zero mean value*

$$\int_{\mathbb{R}} \phi(x) dx = \sqrt{2\pi} \hat{\phi}(0) > 0. \quad (2.22)$$

Further let

$$\hat{\phi}(2\pi k) = 0 \text{ for all } k \in \mathbb{Z} \setminus \{0\} \quad (2.23)$$

Then the spaces

$$V_j = \overline{\text{span}\{\phi_{j,k} \mid k \in \mathbb{Z}\}}^{L^2(\mathbb{R})}, \quad j \in \mathbb{Z}, \quad (2.24)$$

give an MSA of $L^2(\mathbb{R})$.

Proof: As we proved in theorem 2.11 we can orthogonalize ϕ . The calculated functions $\tilde{\phi}$ fulfill

$$\int_{\mathbb{R}} \tilde{\phi}(x) dx = \sqrt{2\pi} \hat{\tilde{\phi}}(0) \stackrel{(2.19)}{=} \sqrt{2\pi} \cdot \frac{1}{\sqrt{2\pi}} \cdot \frac{\hat{\phi}(0)}{\sqrt{\sum_{n \in \mathbb{Z}} |\hat{\phi}(2\pi n)|^2}} \stackrel{(2.23)}{=} \frac{\hat{\phi}(0)}{\sqrt{|\hat{\phi}(0)|^2}} = 1. \quad (2.25)$$

For simplification of the notation we set $\phi = \tilde{\phi}$ in the following. We mentioned before that we have to show $\overline{\bigcup_{j \in \mathbb{Z}} V_j}^{L^2(\mathbb{R})} = L^2(\mathbb{R})$ or equivalent the convergence (2.20) on a dense subset of $L^2(\mathbb{R})$. As a candidate for the dense subspace we choose the set

$$B = \{g \in L^1(\mathbb{R}) \mid \hat{g} \in C^\infty(\mathbb{R}), \hat{g} \text{ has compact support}\}$$

and verify (2.20) for $f \in B$ (We will not show that the set B is a dense subset in $L^2(\mathbb{R})$, but this set is used in LOUIS ET AL. (1997)). Let the support of \hat{f} be contained in $[-2^n\pi, 2^n\pi]$.

We get the equation

$$\begin{aligned} \|P_j f - f\|_{L^2(\mathbb{R})}^2 &\stackrel{(2.21)}{=} \|f\|_{L^2(\mathbb{R})}^2 - \sum_{k \in \mathbb{Z}} |\langle f, \phi_{j,k} \rangle_{L^2(\mathbb{R})}|^2 \\ &\stackrel{(B.6) \text{ and } (B.7)}{=} \|\hat{f}\|_{L^2(\mathbb{R})}^2 - \sum_{k \in \mathbb{Z}} |\langle \hat{f}, \hat{\phi}_{j,k} \rangle_{L^2(\mathbb{R})}|^2. \end{aligned}$$

The scalar products

$$\begin{aligned} \langle \hat{f}, \hat{\phi}_{j,k} \rangle_{L^2(\mathbb{R})} &= \int_{-2^n\pi}^{2^n\pi} \hat{f}(\omega) \overline{\hat{\phi}_{j,k}(\omega)} d\omega \\ &= 2^{j/2} \int_{-2^n\pi}^{2^n\pi} \hat{f}(\omega) \overline{\hat{\phi}(2^j\omega - k)} d\omega \\ &\stackrel{(B.4)}{=} 2^{j/2} \int_{-2^n\pi}^{2^n\pi} \hat{f}(\omega) e^{ik2^{-j}\omega} \overline{\hat{\phi}(2^j\omega)} d\omega \\ &= 2^{j/2} \int_{-2^n\pi}^{2^n\pi} \hat{f}(\omega) e^{ik2^{-j}\omega} \overline{\hat{\phi}\left(\frac{1}{2^{-j}}\omega\right)} d\omega \\ &\stackrel{(B.5)}{=} 2^{j/2} \int_{-2^n\pi}^{2^n\pi} \hat{f}(\omega) e^{ik2^{-j}\omega} \overline{2^{-j} \hat{\phi}(2^{-j}\omega)} d\omega \\ &= 2^{-j/2} \int_{-2^n\pi}^{2^n\pi} \hat{f}(\omega) e^{ik2^{-j}\omega} \overline{\hat{\phi}(2^{-j}\omega)} d\omega \\ &= \sqrt{2\pi} \sqrt{\frac{2^{-j-1}}{\pi}} \int_{-2^n\pi}^{2^n\pi} \hat{f}(\omega) \overline{\hat{\phi}(2^{-j}\omega)} e^{ik2^{-j}\omega} d\omega \end{aligned}$$

are $\sqrt{2\pi}$ -multiples of the Fourier coefficients of the product $\hat{f}(\omega)\overline{\hat{\phi}(2^{-j}\omega)}$ with respect to the ONB $\{\sqrt{2^{-j-1}}/\pi e^{ik2^{-j}\omega} \mid k \in \mathbb{Z}\}$ of $L^2([-2^j\pi, 2^j\pi])$. Since $L^2([-2^n\pi, 2^n\pi]) \subset L^2([-2^j\pi, 2^j\pi])$ for $j \geq n$ it holds

$$\sum_{k \in \mathbb{Z}} |\langle \hat{f}, \hat{\phi}_{j,k} \rangle_{L^2(\mathbb{R})}|^2 \stackrel{(A.3)}{=} 2\pi \|\hat{f}(\cdot)\overline{\hat{\phi}(2^{-j}\cdot)}\|_{L^2(\mathbb{R})}^2$$

and

$$\|P_j f - f\|_{L^2(\mathbb{R})}^2 = \|\hat{f}\|_{L^2(\mathbb{R})}^2 - 2\pi \|\hat{f}(\cdot)\overline{\hat{\phi}(2^{-j}\cdot)}\|_{L^2(\mathbb{R})}^2.$$

$\phi \in L^1(\mathbb{R})$, hence $\hat{\phi}$ is continuous. \hat{f} has compact support and it is $\hat{\phi}(0) = 1/\sqrt{2\pi}$ (compare equation (2.25)). Therefore $2\pi \|\hat{f}(\cdot)\overline{\hat{\phi}(2^{-j}\cdot)}\|_{L^2(\mathbb{R})}^2$ converges to $\|\hat{f}\|_{L^2(\mathbb{R})}^2$ for $j \rightarrow \infty$, and it follows that

$$\lim_{j \rightarrow \infty} \|P_j f - f\|_{L^2(\mathbb{R})}^2 = 0.$$

□

Lemma 2.14. *The coefficients of the two-scale relation (2.7) for an orthogonal scaling function satisfy the orthogonality relation:*

$$\sum_{k \in \mathbb{Z}} p_k \overline{p_{k-2j}} = \delta_{0,j}. \quad (2.26)$$

Proof: See LOUIS ET AL. (1997). □

Theorem 2.15. *Let $\{V_j\}_{j \in \mathbb{Z}}$ be an MSA generated by the orthogonal scaling function $\phi \in V_0$. The function $\psi \in V_1$, defined by the two-scale relation (2.16), i.e.*

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} q_k \phi(2x - k) = \sum_{k \in \mathbb{Z}} q_k \phi_{1,k}(x), \quad (2.27)$$

$$q_k = (-1)^k \overline{p_{1-k}}, \quad (2.28)$$

where $\{p_k\}_{k \in \mathbb{Z}}$ are the coefficients of the scaling equation (2.7), has the following properties

(i) $\{\psi_{j,k}(\cdot) = 2^{j/2} \psi(2^j \cdot - k) \mid k \in \mathbb{Z}\}$ is an ONB for W_j ,

(ii) $\{\psi_{j,k}(\cdot) \mid j, k \in \mathbb{Z}\}$ is an ONB for $L^2(\mathbb{R})$,

(iii) ψ is a wavelet with $c_\psi = 2\pi \int_{\mathbb{R}} |\omega|^{-1} |\hat{\psi}(\omega)|^2 d\omega = 2 \ln 2$, i.e. ψ fulfills the admissibility condition (2.14).

Remark:

For the proof we will use the following: To prove that $\{f_n \mid n \in \mathbb{Z}\}$ is an ONB of our scaling space $V_j = \overline{\text{span } \phi_{j,k} \mid k \in \mathbb{Z}}$, $j \in \mathbb{Z}$, it is sufficient to show the representability of $\phi_{j,0}$ in this basis. Because if $\phi_{j,0}(x) = \phi(2^j x) = \sum_{n \in \mathbb{Z}} \langle \phi_{j,0}(\cdot), f_n(\cdot) \rangle_{L^2(\mathbb{R})} f_n(x)$ holds it follows that $\phi_{j,k}(x) = \phi(2^j x - k) = \phi_{j,0}(x - \frac{k}{2^j}) = \sum_{n \in \mathbb{Z}} \langle \phi_{j,0}(\cdot), f_n(\cdot) \rangle_{L^2(\mathbb{R})} f_n(x - \frac{k}{2^j})$, $k \in \mathbb{Z}$ and it follows that $\phi_{j,k}(x) \in V_j$, $k \in \mathbb{Z}$.

Proof: In the first part we show that $\psi \in W_0 \subset V_1$. From (2.27) it is obvious that $\psi \in V_1$ and it follows that

$\psi_{0,k} \in V_1$, $k \in \mathbb{Z}$. Hence it remains to show that $\psi \perp \phi_{0,n}$, $n \in \mathbb{Z}$:

$$\begin{aligned}
\langle \psi(\cdot), \phi(\cdot - n) \rangle_{L^2(\mathbb{R})} &= 2 \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} q_k \overline{p_l} \langle \phi(2 \cdot - k), \phi(2 \cdot - 2n - l) \rangle_{L^2(\mathbb{R})} \\
&= \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} q_k \overline{p_l} \delta_{k, 2n+l} = \sum_{l \in \mathbb{Z}} q_{2n+l} \overline{p_l} \\
&\stackrel{(2.28)}{=} \sum_{l \in \mathbb{Z}} (-1)^l \overline{p_{1-2n-l}} \overline{p_l} \\
&= \sum_{l \in \mathbb{Z}} \overline{p_{1-2(n+l)}} \overline{p_{2l}} - \sum_{l \in \mathbb{Z}} \overline{p_{-2(n+l)}} \overline{p_{2l+1}} \\
&= \sum_{\lambda \in \mathbb{Z}} \overline{p_{1+2\lambda}} \overline{p_{-2(\lambda+n)}} - \sum_{l \in \mathbb{Z}} \overline{p_{-2(n+l)}} \overline{p_{2l+1}} \\
&= 0.
\end{aligned}$$

The orthogonality of the $\psi_{0,k}$, for all $k \in \mathbb{Z}$ to the space V_0 follows easily by $\langle \psi(\cdot - k), \phi(\cdot - n) \rangle_{L^2(\mathbb{R})} = \langle \psi(\cdot), \phi(\cdot + k - n) \rangle_{L^2(\mathbb{R})} = 0$. It follows $\psi_{0,k} \in W_0$.

The orthogonality of the family $\{\psi(\cdot - k) \mid k \in \mathbb{Z}\}$ can be proven by a similar calculation using (2.26).

To complete the proof for (i) and (ii) it remains to show the completeness of $\{\psi(\cdot - k) \mid k \in \mathbb{Z}\}$ in W_0 . We show the completeness of the orthonormal system $\{\phi(\cdot - k), \psi(\cdot - k) \mid k \in \mathbb{Z}\}$ in V_1 , as $V_0 \oplus W_0 = V_1$. It is sufficient to show the representability of $\phi_{1,0}$ by $\{\phi(\cdot - k), \psi(\cdot - k) \mid k \in \mathbb{Z}\}$ (compare the previous remark): we want to check equation (2.21) and therefore we calculate the sum of the coefficients of the orthogonal projection:

$$\begin{aligned}
&\sum_{k \in \mathbb{Z}} |\langle \phi_{1,0}(\cdot), \phi(\cdot - k) \rangle_{L^2(\mathbb{R})}|^2 + |\langle \phi_{1,0}(\cdot), \psi(\cdot - k) \rangle_{L^2(\mathbb{R})}|^2 \\
&2 \sum_{k \in \mathbb{Z}} |\langle \phi(2 \cdot), \phi(\cdot - k) \rangle_{L^2(\mathbb{R})}|^2 + |\langle \phi(2 \cdot), \psi(\cdot - k) \rangle_{L^2(\mathbb{R})}|^2 \text{ inserting equations (2.7) and (2.27)} \\
&= 4 \sum_{k \in \mathbb{Z}} \left(\left| \sum_{l \in \mathbb{Z}} \overline{p_l} \langle \phi(2 \cdot), \phi(2 \cdot - 2k - l) \rangle_{L^2(\mathbb{R})} \right|^2 + \left| \sum_{l \in \mathbb{Z}} \overline{q_l} \langle \phi(2 \cdot), \phi(2 \cdot - 2k - l) \rangle_{L^2(\mathbb{R})} \right|^2 \right) \\
&\stackrel{(2.28) \text{ and } (2.17)}{=} \sum_{k \in \mathbb{Z}} \left(\left| \sum_{l \in \mathbb{Z}} \overline{p_l} \delta_{0, 2k+l} \right|^2 + \left| \sum_{l \in \mathbb{Z}} (-1)^l \overline{p_{1-l}} \delta_{0, 2k+l} \right|^2 \right) \\
&= \sum_{k \in \mathbb{Z}} |\overline{p_{-2k}}|^2 + \sum_{k \in \mathbb{Z}} |p_{-2k+1}|^2 = \sum_{k \in \mathbb{Z}} |p_k|^2 \\
&= \sum_{k \in \mathbb{Z}} p_k \overline{p_k} \\
&\stackrel{(2.26)}{=} 1 = \|\phi_{1,0}\|_{L^2(\mathbb{R})}^2.
\end{aligned}$$

Inserting the result in (2.21) it follows

$$\|\phi_{1,0} - \sum_{k \in \mathbb{Z}} (\langle \phi_{1,0}(\cdot), \phi_{0,k}(\cdot) \rangle \phi_{0,k} + \langle \phi_{1,0}, \psi_{0,k}(\cdot) \rangle \psi_{0,k})\|_{L^2(\mathbb{R})} = 0.$$

Hence $\psi_{0,k}$, $k \in \mathbb{Z}$, form an ONB for W_0 .

We still have to show that $\psi_{j,k}$, $k \in \mathbb{Z}$, form an ONB for W_j , $j \in \mathbb{Z}$ but this follows analogous to the case $j = 0$ by a simple substitution of the coordinates. We will not repeat the proof here.

Part (ii) follows with equation (2.11).

Part (iii) follows with some properties on wavelet frames we will not discuss here. For more details see LOUIS ET AL. (1997). \square

Corollary 2.16. *The wavelet associated with a MSA is not unique, because*

$$q_k = (-1)^k \overline{p_{1+2l-k}} \text{ for an } l \in \mathbb{Z} \tag{2.29}$$

defines also a wavelet by (2.27), i.e. Theorem 2.15 also applies if (2.28) is replaced by (2.29).

Proof: For the proof there have to be made some changes in the proof of Theorem 2.15 and it will not be repeated here. \square

Now we want to give a short overview of the results on MSAs and wavelet bases. We can use an arbitrary function with non-zero mean value and satisfying the assumptions of Theorem 2.13 to generate a MSA. By calculating the two-scale coefficients in equation (2.7) we gain orthogonal wavelets with the equations (2.28) or (2.29). The crux of this method is the orthogonalization of the scaling function from Theorem 2.11. Apart from a few exceptions we cannot determine the orthogonal scaling function $\tilde{\phi}$.

Suppose we could determine the orthogonal scaling function, then another difficulty is the determination of the coefficients p_k from the scaling equation. If the scaling function $\tilde{\phi}$ has not compact support, then the sum (2.7) extends over \mathbb{Z} and we have to calculate infinitely many coefficients. Because the orthogonal scaling function $\tilde{\phi}$ is implicitly defined by its Fourier transform, this dilemma becomes the general rule; see LOUIS ET AL. (1997).

2.1.3 Two-scale relations

In chapter 2.1.2 we already mentioned that we will work with more general concepts than the orthogonal wavelets. The orthogonality puts a strong limitation on the construction of wavelets and we cannot achieve some properties that would be desirable.

Some important properties of wavelets are generally

- orthogonality
- compact support
- rational coefficients
- symmetry
- smoothness
- number of vanishing moments of the dual wavelets
- analytic expressions
- interpolation.

For more details on the advantages of these properties and for examples of wavelets fulfilling some of these properties see JAWERTH AND SWELDENS (1994).

Of course it is not possible to construct wavelets that have all the properties and there has to be a trade-off between them. For example the Daubechies wavelets are compactly supported and orthogonal; see e.g. JAWERTH AND SWELDENS (1994). Hence to gain more of the properties, e.g. symmetry and smoothness, we introduce a more flexible concept for constructing wavelets, i.e. we here introduce biorthogonal and semiorthogonal wavelets.

For simplification we will finally restrict ourself to the semiorthogonal wavelets because for the B-splines which we will use in section 2.3.1 as scaling functions we will construct semiorthogonal wavelets.

Definition 2.17. Given a scaling function ϕ and a dual scaling function $\tilde{\phi}$ (dual in the sense of $\langle \phi_{j,k}, \tilde{\phi}_{j,l} \rangle_{L^2(\mathbb{R})} = \delta_{k,l}$) generating two MSAs

$$\dots \subset V_0 \subset V_1 \subset V_2 \subset \dots$$

$$\dots \subset \tilde{V}_0 \subset \tilde{V}_1 \subset \tilde{V}_2 \subset \dots$$

with subspaces $V_j = \overline{\text{span}\{\phi_{j,k} \mid k \in \mathbb{Z}\}}^{L^2(\mathbb{R})}$ and $\tilde{V}_j = \overline{\text{span}\{\tilde{\phi}_{j,k} \mid k \in \mathbb{Z}\}}^{L^2(\mathbb{R})}$, respectively.

As in the orthogonal case, each of these sequences of scaling spaces has a sequence of successive detail spaces $\{W_j\}_{j \in \mathbb{Z}}$ and $\{\tilde{W}_j\}_{j \in \mathbb{Z}}$ but they are not necessarily orthogonal to V_j and \tilde{V}_j , respectively (i.e. $V_{j+1} = V_j \oplus W_j$ but the sum is no longer a direct sum). These detail spaces are generated by two wavelet functions ψ and $\tilde{\psi}$, which are dual to one another (in the sense of $\langle \psi_{j,k}, \tilde{\psi}_{i,l} \rangle_{L^2(\mathbb{R})} = \delta_{j,i} \delta_{k,l}$). The biorthogonality is expressed through relationships between the dual MSAs:

$$V_j \perp \tilde{W}_j, \quad j \in \mathbb{Z} \tag{2.30}$$

and

$$\tilde{V}_j \perp W_j, j \in \mathbb{Z}. \quad (2.31)$$

In terms of scaling and wavelet functions the biorthogonality (2.30) and (2.31) is expressed as

$$\langle \tilde{\psi}_{j,k}, \phi_{j,l} \rangle_{L^2(\mathbb{R})} = 0 \text{ and } \langle \tilde{\phi}_{j,k}, \psi_{j,l} \rangle_{L^2(\mathbb{R})} = 0 \quad (2.32)$$

and the duality is expressed as

$$\langle \phi_{j,l}, \tilde{\phi}_{j,k} \rangle_{L^2(\mathbb{R})} = \delta_{k,l} \text{ and } \langle \tilde{\psi}_{j,k}, \psi_{i,l} \rangle_{L^2(\mathbb{R})} = \delta_{j,i} \delta_{k,l}, \quad i, j, k, l \in \mathbb{Z}. \quad (2.33)$$

A wavelet ψ is called a biorthogonal wavelet if it satisfies the conditions (2.32) and (2.33). Note, nothing is said about the orthogonality of the functions $\{\phi_{j,k}\}$, $\{\tilde{\phi}_{j,k}\}$, $\{\psi_{j,k}\}$ and $\{\tilde{\psi}_{j,k}\}$ themselves (see OGDEN (1997)).

The semiorthogonal concept is a restriction of the biorthogonal concept and is defined in the following:

Definition 2.18. A biorthogonal wavelet function ψ is called a semiorthogonal wavelet if it satisfies the following concept: Associated with ϕ and ψ are the dual functions $\tilde{\phi}$ and $\tilde{\psi}$. The functions $\{\psi_{j,k} \mid k \in \mathbb{Z}\}$ form a Riesz basis for the detail space W_j , and just as in the orthogonal wavelet case,

$$V_j \perp W_j, j \in \mathbb{Z}.$$

Thus the W_j spaces are mutually orthogonal. In this formulation, both $\{\phi_{j,k} \mid k \in \mathbb{Z}\}$ and $\{\tilde{\phi}_{j,k} \mid k \in \mathbb{Z}\}$ are Riesz bases for the spaces V_j ; similarly, $\{\psi_{j,k} \mid k \in \mathbb{Z}\}$ and $\{\tilde{\psi}_{j,k} \mid k \in \mathbb{Z}\}$ are Riesz bases for W_j (i.e. $\tilde{W}_j = W_j$ and $\tilde{V}_j = V_j$).

Remarks:

- The dual MSA is not necessarily the same as the one generated by the original basis functions.
- The biorthogonal wavelets are more flexible than the orthogonal wavelets, hence we get less limitations on the construction of wavelets. An orthogonal wavelet is "self-dual" in the sense that $\tilde{\psi}_{j,k} = \psi_{j,k}$ for $j, k \in \mathbb{Z}$. Hence it fulfills of course the biorthogonal concept.
- The role of the basis (i.e. the ϕ and ψ) and the dual basis (i.e. the $\tilde{\phi}$ and $\tilde{\psi}$) can be interchanged.
- We consider here the semiorthogonal wavelets because in our applications we apply normalized endpoint interpolating B-splines which have semiorthogonal wavelets.

Here we will not consider the existence of dual functions $\tilde{\phi}$ and $\tilde{\psi}$ generating the dual bases $\tilde{\phi}_{j,k}$ and $\tilde{\psi}_{j,k}$. But for a proof of the existence of the dual functions see CHUI (1992), p. 77. Note, in the more general concepts there may exist dual bases $\tilde{\psi}_{j,k}$ that are not generated by a function $\tilde{\psi}$.

Due to the last item we suppose to have scaling functions ϕ and $\tilde{\phi}$ and wavelet functions ψ and $\tilde{\psi}$ that fulfill the concept of Definition 2.18. Note, we could also work with the more general concept of biorthogonal spaces but for simplification we will not use this concept and can therefore work without the dual spaces \tilde{V}_j and \tilde{W}_j .

The functions $\phi_{j,k}$ and $\tilde{\phi}_{j,k}$ both span the space V_j and are orthonormal to each other (i.e. $\langle \tilde{\phi}_{j,k}, \phi_{j,l} \rangle = \delta_{k,l}$). Then for every function $f \in V_j$ there exist sequences $\{c_{j,k}\}_{k \in \mathbb{Z}}$ and $\{d_{j,k}\}_{k \in \mathbb{Z}}$ of real numbers such that we can write the series expansions

$$f(x) = \sum_{k \in \mathbb{Z}} c_{j,k} \tilde{\phi}_{j,k}(x) \quad (2.34)$$

$$= \sum_{k \in \mathbb{Z}} d_{j,k} \phi_{j,k}(x) \text{ for } x \in \mathbb{R}. \quad (2.35)$$

To compute the coefficients $c_{j,m}$ we take the inner product of equation (2.34) with $\phi_{j,m}$ and obtain

$$\begin{aligned} \langle f, \phi_{j,m} \rangle_{L^2(\mathbb{R})} &= \left\langle \sum_{k \in \mathbb{Z}} c_{j,k} \tilde{\phi}_{j,k}, \phi_{j,m} \right\rangle_{L^2(\mathbb{R})} \\ &= \sum_{k \in \mathbb{Z}} c_{j,k} \delta_{k,m} \\ &= c_{j,m} \end{aligned}$$

and analogous taking the inner product with $\tilde{\psi}_{j,m}$ we can calculate the coefficients $d_{i,m}$ and get

$$\langle f, \tilde{\psi}_{j,m} \rangle_{L^2(\mathbb{R})} = d_{j,m},$$

see e.g. CHUI (1992).

Hence, we can rewrite the equations (2.34) and (2.35) as

$$f(x) = \sum_{k \in \mathbb{Z}} \langle f, \phi_{j,k} \rangle_{L^2(\mathbb{R})} \tilde{\phi}_{j,k}(x) \quad (2.36)$$

$$= \sum_{k \in \mathbb{Z}} \langle f, \tilde{\phi}_{j,k} \rangle_{L^2(\mathbb{R})} \phi_{j,k}(x). \quad (2.37)$$

If ϕ and $\tilde{\phi}$ are semiorthogonal in the sense of Definition 2.18 we can therefore write each signal $f \in V_j$ as one of the series expansions (2.36) or (2.37).

Two-scale relations

Now we construct more general two-scale relations similar to equations (2.7) and (2.16). It follows from the definition of a MSA that the $\phi_{j,k}$, $k \in \mathbb{Z}$, span the scaling space V_j . Since $V_j \subset V_{j+1}$ and $W_j \subset V_{j+1}$ the functions $\phi_{j,l} \in V_j$ and $\psi_{j,l} \in W_j$ can be written as series expansions analogous to (2.37), i.e.

$$\phi_{j,l}(x) = \sum_{k \in \mathbb{Z}} \langle \phi_{j,l}, \tilde{\phi}_{j+1,k} \rangle_{L^2(\mathbb{R})} \phi_{j+1,k}(x), \quad (2.38)$$

$$\psi_{j,l}(x) = \sum_{k \in \mathbb{Z}} \langle \psi_{j,l}, \tilde{\phi}_{j+1,k} \rangle_{L^2(\mathbb{R})} \phi_{j+1,k}(x) \quad (2.39)$$

or with

$$p_{k,l}^{j+1} := \langle \phi_{j,l}, \tilde{\phi}_{j+1,k} \rangle_{L^2(\mathbb{R})} \text{ and } q_{k,l}^{j+1} := \langle \psi_{j,l}, \tilde{\phi}_{j+1,k} \rangle_{L^2(\mathbb{R})}$$

we rewrite equation (2.38) and (2.39) as

$$\phi_{j,l}(x) = \sum_{k \in \mathbb{Z}} p_{k,l}^{j+1} \phi_{j+1,k}(x), \quad (2.40)$$

$$\psi_{j,l}(x) = \sum_{k \in \mathbb{Z}} q_{k,l}^{j+1} \phi_{j+1,k}(x). \quad (2.41)$$

These equations are called **two-scale relations** because the functions $\phi_{j,l}(x)$ and $\psi_{j,l}(x)$ refer to the scale j and the functions $\phi_{j+1,k}(x)$ refer to the scale $j+1$, the coefficients $p_{k,l}^{j+1}$ and $q_{k,l}^{j+1}$ are called **two-scale coefficients**. We get the following equations if we set $j=0$ and $l=0$:

$$\phi(x) = \sum_{k \in \mathbb{Z}} p_{k,0}^1 \phi_{1,k}(x) = 2^{1/2} \sum_{k \in \mathbb{Z}} p_{k,0}^1 \phi(2x - k) \quad (2.42)$$

$$\psi(x) = \sum_{k \in \mathbb{Z}} q_{k,0}^1 \phi_{1,k}(x) = 2^{1/2} \sum_{k \in \mathbb{Z}} q_{k,0}^1 \phi(2x - k); \quad (2.43)$$

the first equation (2.42) corresponds to the two-scale relation (2.7) the second equation corresponds to equation (2.16).

For the two-scale equations (2.38) and (2.39) we can compute the inner products as follows:

$$\begin{aligned} \langle \phi_{j,l}, \tilde{\phi}_{j+1,k} \rangle_{L^2(\mathbb{R})} &= 2^{j+1/2} \int_{-\infty}^{\infty} \phi(2^j x - l) \overline{\tilde{\phi}(2^{j+1} x - k)} dx \quad (\text{substitution: } s = 2^j x - l, dx = 2^{-j} ds) \\ &= 2^{1/2} \int_{-\infty}^{\infty} \phi(s) \overline{\tilde{\phi}(2s - (k - 2l))} ds \\ &= \langle \phi_{0,0}, \tilde{\phi}_{1,k-2l} \rangle_{L^2(\mathbb{R})} =: p_{k-2l} \end{aligned} \quad (2.44)$$

and analogous

$$\begin{aligned} \langle \psi_{j,l}, \tilde{\phi}_{j+1,k} \rangle_{L^2(\mathbb{R})} &= 2^{1/2} \int_{-\infty}^{\infty} \psi(s) \overline{\tilde{\phi}(2s - (k - 2l))} ds \\ &= \langle \psi_{0,0}, \tilde{\phi}_{1,k-2l} \rangle_{L^2(\mathbb{R})} =: q_{k-2l}. \end{aligned} \quad (2.45)$$

Note, now the two-scale coefficients p_{k-2l} and q_{k-2l} are independent of the resolution level j . And we can rewrite the two-scale equations (2.38) and (2.39)

$$\phi_{j,l}(x) = \sum_{k \in \mathbb{Z}} p_{k-2l} \phi_{j+1,k}(x) \quad (2.46)$$

$$\psi_{j,l}(x) = \sum_{k \in \mathbb{Z}} q_{k-2l} \phi_{j+1,k}(x) \quad x \in \mathbb{R}. \quad (2.47)$$

Analogous to equations (2.38) and (2.39) we can write for the dual MSA

$$\tilde{\phi}_{j,l}(x) = \sum_{k \in \mathbb{Z}} \langle \tilde{\phi}_{j,l}, \phi_{j+1,l} \rangle_{L^2(\mathbb{R})} \tilde{\phi}_{j+1,k}(x),$$

$$\tilde{\psi}_{j,l}(x) = \sum_{k \in \mathbb{Z}} \langle \tilde{\psi}_{j,l}, \phi_{j+1,k} \rangle_{L^2(\mathbb{R})} \tilde{\phi}_{j+1,k}(x)$$

and get analogous to p_{k-2l} and q_{k-2l} in equations (2.44) and (2.45) the dual two-scale coefficients \tilde{p}_{k-2l} and \tilde{q}_{k-2l}

$$\langle \tilde{\phi}_{j,l}, \phi_{j+1,k} \rangle_{L^2(\mathbb{R})} = \langle \tilde{\phi}_{0,0}, \phi_{1,k-2l} \rangle_{L^2(\mathbb{R})} = \tilde{p}_{k-2l} \quad \text{and} \quad (2.48)$$

$$\langle \tilde{\psi}_{j,l}, \phi_{j+1,k} \rangle_{L^2(\mathbb{R})} = \langle \tilde{\psi}_{0,0}, \phi_{1,k-2l} \rangle_{L^2(\mathbb{R})} = \tilde{q}_{k-2l}. \quad (2.49)$$

The two-scale coefficients p_{k-2l} , q_{k-2l} , \tilde{p}_{k-2l} and \tilde{q}_{k-2l} will later play an important role for the decomposition of signals $f_j \in V_j$. With those coefficients we will construct a decomposition relation and a reconstruction relation for signals $f_j \in V_j$.

If we establish an additional condition on ϕ and with equation (2.15) for ψ we gain more information on the coefficients $p_{k,0}^1$ and $q_{k,0}^1$.

Assume $\phi(x)$ is normalized, i.e.

$$\int_{-\infty}^{\infty} \phi(x) dx = 1,$$

then we apply equation (2.42) and get

$$2^{1/2} \sum_{k \in \mathbb{Z}} p_{k,0}^1 \int_{-\infty}^{\infty} \phi(2x - k) dx = 1.$$

And similarly with equation (2.15)

$$\int_{-\infty}^{\infty} \psi(x) dx = 0$$

we get from equation (2.43)

$$2^{1/2} \sum_{k \in \mathbb{Z}} q_{k,0}^1 \int_{-\infty}^{\infty} \phi(2x - k) dx = 0.$$

After substituting and evaluation of the integral we get

$$\sum_{k \in \mathbb{Z}} p_{k,0}^1 = 2^{1/2},$$

$$\sum_{k \in \mathbb{Z}} q_{k,0}^1 = 0$$

for the two-scale coefficients $p_{k,0}^1$ and $q_{k,0}^1$, $k \in \mathbb{Z}$. These equations are important for the constructions of scaling functions and wavelets. The two-scale coefficients $p_{k,0}^1$ and $q_{k,0}^1$ have to fulfill these equations. Note, besides these equations the two-scale coefficients have to fulfill more conditions, such that they can be determined uniquely.

2.1.4 Decomposition and reconstruction

Now we introduce a concept to decompose a function $f_j \in V_j$ into a coarser signal and the corresponding detail signals. These signals will be defined via coefficients $c_{k,l}^j$ and $d_{k,l}^j$, $j, k, l \in \mathbb{Z}$. Via those coefficients we will be able to perform a decomposition of the original function f_j in $f_{j-1} \in V_{j-1}$ and $g_{j-1} \in W_{j-1}$ and to reconstruct f_j from the signals f_{j-1} and g_{j-1} .

Decomposition

First we repeat the decomposition of a function f (compare equations (2.12) and (2.13)).

For $L^2(\mathbb{R}) = \bigoplus_{i \in \mathbb{Z}} W_i$ the function $f \in L^2(\mathbb{R})$ can be decomposed into a sum of detail signals $g_i \in W_i$:

$$f(x) = \sum_{i \in \mathbb{Z}} g_i(x), \quad x \in \mathbb{R}. \quad (2.50)$$

If we separate the sum into two parts we get

$$\begin{aligned} f(x) &= \sum_{i=-\infty}^{j-1} g_i(x) + \sum_{i=j}^{\infty} g_i(x) \\ &= f_j(x) + \sum_{i=j}^{\infty} g_i(x) \end{aligned}$$

with the smoothed signal $f_j(x)$ defined as

$$f_j(x) := \sum_{i=-\infty}^{j-1} g_i(x).$$

Since $f \in L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$ we can formulate the series expansion analogously to equation (2.37):

$$\begin{aligned} f(x) &= \sum_{i \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f, \tilde{\psi}_{i,k} \rangle_{L^2(\mathbb{R})} \psi_{i,k}(x) \\ &= \sum_{i=-\infty}^{j-1} \sum_{k \in \mathbb{Z}} \langle f, \tilde{\psi}_{i,k} \rangle_{L^2(\mathbb{R})} \psi_{i,k}(x) + \sum_{i=j}^{\infty} \sum_{k \in \mathbb{Z}} \langle f, \tilde{\psi}_{i,k} \rangle_{L^2(\mathbb{R})} \psi_{i,k}(x) \\ &= f_j(x) + \sum_{i=j}^{\infty} \sum_{k \in \mathbb{Z}} \langle f, \tilde{\psi}_{i,k} \rangle_{L^2(\mathbb{R})} \psi_{i,k}(x) \\ &= f_j(x) + \sum_{i=j}^{\infty} g_i(x) \end{aligned}$$

with

$$g_i(x) = \sum_{k \in \mathbb{Z}} \langle f, \tilde{\psi}_{i,k} \rangle_{L^2(\mathbb{R})} \psi_{i,k}(x) = \sum_{k \in \mathbb{Z}} c_{i,k} \psi_{i,k}(x), \quad \text{wherein } c_{i,k} = \langle f, \tilde{\psi}_{i,k} \rangle_{L^2(\mathbb{R})}, \quad i \in \mathbb{Z} \text{ and} \quad (2.51)$$

$$f_j(x) = \sum_{i=-\infty}^{j-1} g_i(x). \quad (2.52)$$

Since the wavelet functions $\psi_{i,k}$, $k \in \mathbb{Z}$, span the detail space W_i , the detail signals g_i as linear combinations of $\psi_{i,k}$ lie in W_i . The smoothed signal $f_j(x) = \sum_{i=-\infty}^{j-1} g_i(x)$ lies in the direct sum $\bigoplus_{i \leq j} W_i = V_j$. V_j is spanned by the functions $\phi_{j,k}$. Hence $f_j(x)$ can be written in a series expansion

$$f_j(x) = \sum_{k \in \mathbb{Z}} d_{j,k} \phi_{j,k}(x) \quad \text{wherein } d_{j,k} = \langle f, \tilde{\phi}_{j,k} \rangle_{L^2(\mathbb{R})}, \quad x \in \mathbb{R}. \quad (2.53)$$

The coefficients $d_{j,k}$ are called **scaling coefficients** because they define a signal in the scaling space ($f_j \in V_j$), the coefficients $c_{j,k}$ are called **detail coefficients** because they define a signal in the detail space ($g_j \in W_j$).

Decomposition relation

Now we construct the decomposition relation. With the decomposition relation we can perform a decomposition of the level j coefficients $d_{j,k}$, $k \in \mathbb{Z}$, into the level $j-1$ coefficients $d_{j-1,k}$ and $c_{j-1,k}$, $k \in \mathbb{Z}$. Therefore we can compute the signals f_{j-1} and g_{j-1} via the formulas (2.51) and (2.53).

Since we can decompose V_{j+1} into a direct sum of V_j and W_j , i.e. $V_{j+1} = V_j \oplus W_j$, we can formulate for $\phi_{j+1,l}(x) \in V_{j+1}$, $l \in \mathbb{Z}$ under consideration of (2.48) and (2.49):

$$\begin{aligned} \phi_{j+1,l}(x) &= \sum_{k \in \mathbb{Z}} \langle \phi_{j+1,l}, \tilde{\phi}_{j,k} \rangle_{L^2(\mathbb{R})} \phi_{j,k}(x) + \sum_{k \in \mathbb{Z}} \langle \phi_{j+1,l}, \tilde{\psi}_{j,k} \rangle_{L^2(\mathbb{R})} \psi_{j,k}(x) \\ &= \sum_{k \in \mathbb{Z}} \overline{\tilde{p}_{l-2k}} \phi_{j,k}(x) + \sum_{k \in \mathbb{Z}} \overline{\tilde{q}_{l-2k}} \psi_{j,k}(x). \end{aligned} \quad (2.54)$$

Recall, we get a similar equation if we interchange the roles of the basis and the dual basis.

We can write the signal $f_j(x)$ with $f_j \in V_j$ as follows (see equation (2.53))

$$f_j(x) = \sum_{k \in \mathbb{Z}} d_{j,k} \phi_{j,k}(x) \quad x \in \mathbb{R}. \quad (2.55)$$

Another representation of this function we gain from the relation $f_j(x) = f_{j+1}(x) - g_j(x)$ (see equation (2.52)) and insert the equation (2.54) in the following equation.

$$f_j(x) = \sum_{l \in \mathbb{Z}} d_{j+1,l} \phi_{j+1,l}(x) - \sum_{k \in \mathbb{Z}} c_{j,k} \psi_{j,k}(x) \quad (2.56)$$

$$= \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} d_{j+1,l} \overline{\tilde{p}_{l-2k}} \phi_{j,k}(x) + \sum_{k \in \mathbb{Z}} \left(\sum_{l \in \mathbb{Z}} d_{j+1,l} \overline{\tilde{q}_{l-2k}} - c_{j,k} \right) \psi_{j,k}(x) \quad (2.57)$$

If we compare the coefficients of equation (2.55) and equation (2.57) we get the recursion formulas

$$d_{j,k} = \sum_{l \in \mathbb{Z}} \overline{\tilde{p}_{l-2k}} d_{j+1,l} \quad \text{for } k \in \mathbb{Z}, \quad (2.58)$$

$$c_{j,k} = \sum_{l \in \mathbb{Z}} \overline{\tilde{q}_{l-2k}} d_{j+1,l} \quad \text{for } k \in \mathbb{Z}. \quad (2.59)$$

Note, \tilde{p}_{l-2k} and \tilde{q}_{l-2k} are independent of j . We define $\mathbf{d}_j = (d_{j,k})_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$ and $\mathbf{c}_j = (c_{j,k})_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$ with $(\mathbf{d}_j)_l := d_{j,l}$. The main result is that from the coefficients \mathbf{d}_j we can compute the coefficients \mathbf{d}_{j-1} and \mathbf{c}_{j-1} via equations (2.58) and (2.59). This means if we have given the coefficients \mathbf{d}_j corresponding to a signal $f_j \in V_j$ we can compute signals $f_{j-1} \in V_{j-1}$ and $g_{j-1} \in W_{j-1}$, i.e. with f_{j-1} we can compute a low pass filtered version of the signal and with g_{j-1} we can compute the remaining high-frequency part. This is demonstrated in the

decomposition scheme:

$$\begin{array}{ccccccc} \dots & \mathbf{d}_{j+1} & \rightarrow & \mathbf{d}_j & \rightarrow & \mathbf{d}_{j-1} & \rightarrow & \dots \\ & \searrow & & \searrow & & \searrow & & \\ & & & \mathbf{c}_j & & \mathbf{c}_{j-1} & & \dots \end{array} \quad (2.60)$$

Reconstruction relation

Now we derive a reconstruction relation. I.e. from the level $j-1$ coefficients \mathbf{d}_{j-1} and \mathbf{c}_{j-1} we want to compute the level j coefficients \mathbf{d}_j .

Since $V_{j+1} = V_j \oplus W_j$ we can write

$$\sum_{k \in \mathbb{Z}} d_{j+1,k} \phi_{j+1,k}(x) = \sum_{l \in \mathbb{Z}} d_{j,l} \phi_{j,l}(x) + \sum_{l \in \mathbb{Z}} c_{j,l} \psi_{j,l}(x)$$

inserting the two-scale equations (2.46) and (2.47) we get

$$\sum_{k \in \mathbb{Z}} d_{j+1,k} \phi_{j+1,k}(x) = \sum_{k \in \mathbb{Z}} \left(\sum_{l \in \mathbb{Z}} p_{k-2l} d_{j,l} + \sum_{l \in \mathbb{Z}} q_{k-2l} c_{j,l} \right) \phi_{j+1,k}(x);$$

comparing the coefficients we gain

$$d_{j+1,k} = \sum_{l \in \mathbb{Z}} p_{k-2l} d_{j,l} + \sum_{l \in \mathbb{Z}} q_{k-2l} c_{j,l}. \quad (2.61)$$

Note, the p_{k-2l} and q_{k-2l} are independent of the level j . With the reconstruction relation (2.61) we can compute the signal $f_j \in V_j$ from its low pass filtered version $f_{j-1} \in V_{j-1}$ and its detail signal $g_{j-1} \in W_{j-1}$. We can perform this reconstruction via the coefficient vectors \mathbf{d}_{j-1} , \mathbf{c}_{j-1} and \mathbf{d}_j . The procedure is demonstrated in the

reconstruction scheme:

$$\begin{array}{ccccccc} \dots & \mathbf{d}_j & \rightarrow & \mathbf{d}_{j+1} & \rightarrow & \mathbf{d}_{j+2} & \rightarrow & \dots \\ & \nearrow & & \nearrow & & \nearrow & & \\ \dots & \mathbf{c}_j & & \mathbf{c}_{j+1} & & \mathbf{c}_{j+2} & & \end{array} \quad (2.62)$$

In Figure 2.1 the main steps we have performed so far are summarized.

2.1.5 Multiscale analysis on a bounded interval

Since we want to work regionally we will work with a concept of a MSA on a bounded interval. We adopted the concept of a MSA on a bounded interval from CHUI AND QUAK (1992).

Now we want to restrict the concept of a MSA on $L^2(\mathbb{R})$ to the interval $[0, 1]$, i.e. we want to work with a MSA on $L^2([0, 1])$. To perform a MSA on a bounded interval, for example on $[0, 1]$, some changes in the concept have to be made. Adapted to a bounded interval $[0, 1]$, the sequence of nested closed subspaces V_j can no longer be bi-infinite. In a MSA on $L^2(\mathbb{R})$ the functions in the spaces V_j are dilated and their details are coarser for $j \rightarrow -\infty$, i.e. they contain the large-period structures. On the bounded interval $[0, 1]$ we have to consider an initial space $V_0^{[0,1]}$ that contains the functions with the coarsest structures on $[0, 1]$ and therefore we cannot have a bi-infinite sequence anymore. Hence there has to be an initial subspace $V_0^{[0,1]}$ and closed subspaces $V_j^{[0,1]}$ with $j \in \mathbb{N}$, satisfying the following properties

$$V_0^{[0,1]} \subset V_1^{[0,1]} \subset \dots \quad \text{and} \quad (2.63)$$

$$\overline{\bigcup_{j \geq 0} V_j^{[0,1]}}^{L^2([0,1])} = L^2([0, 1]). \quad (2.64)$$

Comparing with the MSA on $L^2(\mathbb{R})$ as defined in Definition 2.3 we see that the properties (2.3) and (2.4) are missing. Of course property (2.3) does not make sense in the MSA on a bounded interval (compare equation (2.63)). Although the property (2.4) gives us additional features for a MSA on $L^2(\mathbb{R})$ we cannot adopt it without some effort, but anyway we will not make use of the feature and we will disregard it.

The complementary orthogonal subspaces $W_j^{[0,1]}$ are defined via

$$V_j^{[0,1]} = V_{j-1}^{[0,1]} \oplus W_{j-1}^{[0,1]} \quad j \in \mathbb{N}_0,$$

hence

$$L^2[0, 1] = V_0^{[0,1]} \oplus \bigoplus_{j \in \mathbb{N}_0} W_j^{[0,1]}. \quad (2.65)$$

The behavior at the endpoints of the interval $[0, 1]$ must now be taken into account and special boundary functions (scaling functions and wavelets) have to be introduced, as the relevant families of functions in $V_j^{[0,1]}$ and $W_j^{[0,1]}$ can no longer be controlled just by dilation and translation alone. The aim is to produce Riesz bases for the spaces $V_j^{[0,1]}$ consisting of a finite set of suitable translates $\phi_{j,k}$ of the original scaling function and a finite set of specially constructed boundary scaling functions $\phi_{j,k}^b$. We have to construct bases of the complementary subspaces $W_j^{[0,1]}$ consisting of a finite set of translates $\psi_{j,k}$ of the wavelet function ψ and a finite set of special boundary wavelets $\psi_{j,k}^b$.

The theory can be applied to an arbitrary bounded interval $[a, b]$.

$$\begin{aligned}
V_j &= \overline{\text{span}\{\phi_{j,k} \mid k \in \mathbb{Z}\}} \\
&= \overline{\text{span}\{\tilde{\phi}_{j,k} \mid k \in \mathbb{Z}\}} \\
W_j &= \overline{\text{span}\{\psi_{j,k} \mid k \in \mathbb{Z}\}} \\
&= \overline{\text{span}\{\tilde{\psi}_{j,k} \mid k \in \mathbb{Z}\}} \text{ for } j \in \mathbb{Z}
\end{aligned}$$

Step 1:

We constitute the scaling space and the detail space.

We have 4 sets of functions:

- scaling functions,
- dual scaling functions,
- wavelets and
- dual wavelets.

$$\begin{aligned}
f \in V_j &\Rightarrow \begin{cases} \exists d_{j,k} \in \mathbb{C}, k \in \mathbb{Z} \text{ such that } f(x) = \sum_{k \in \mathbb{Z}} d_{j,k} \phi_{j,k}(x) \\ \exists \tilde{d}_{j,k} \in \mathbb{C}, k \in \mathbb{Z} \text{ such that } f(x) = \sum_{k \in \mathbb{Z}} \tilde{d}_{j,k} \tilde{\phi}_{j,k}(x) \end{cases} \\
g \in W_j &\Rightarrow \begin{cases} \exists c_{j,k} \in \mathbb{C}, k \in \mathbb{Z} \text{ such that } f(x) = \sum_{k \in \mathbb{Z}} c_{j,k} \psi_{j,k}(x) \\ \exists \tilde{c}_{j,k} \in \mathbb{C}, k \in \mathbb{Z} \text{ such that } f(x) = \sum_{k \in \mathbb{Z}} \tilde{c}_{j,k} \tilde{\psi}_{j,k}(x) \end{cases}
\end{aligned}$$

Step 2:

We develop series expansions for functions in the scaling spaces and analogous for functions in the detail spaces.

$$\begin{aligned}
\phi_{j,l}(x) &= \sum_{k \in \mathbb{Z}} p_{k-2l} \phi_{j+1,k}(x) \\
\psi_{j,l}(x) &= \sum_{k \in \mathbb{Z}} q_{k-2l} \phi_{j+1,k}(x) \\
\tilde{\phi}_{j,l}(x) &= \sum_{k \in \mathbb{Z}} \tilde{p}_{k-2l} \tilde{\phi}_{j+1,k}(x) \\
\tilde{\psi}_{j,l}(x) &= \sum_{k \in \mathbb{Z}} \tilde{q}_{k-2l} \tilde{\phi}_{j+1,k}(x)
\end{aligned}$$

Step 3:

We constitute the two-scale relations for the scaling functions, the wavelets, the dual scaling functions and the dual wavelets. The two-scale coefficients play an important role for the decomposition and reconstruction relations.

$$\begin{aligned}
d_{j,k} &= \sum_{l \in \mathbb{Z}} \overline{\tilde{p}_{l-2k}} d_{j+1,l}, \text{ for } k \in \mathbb{Z} \\
c_{j,k} &= \sum_{l \in \mathbb{Z}} \overline{\tilde{q}_{l-2k}} d_{j+1,l}, \text{ for } k \in \mathbb{Z}
\end{aligned}$$

Step 4:

We derive the decomposition relations.

The dual two-scale coefficients are used in these relations.

$$d_{j+1,k} = \sum_{l \in \mathbb{Z}} p_{k-2l} d_{j,l} + \sum_{l \in \mathbb{Z}} q_{k-2l} c_{j,l}$$

Step 5:

We derive the reconstruction relation.

The two-scale coefficients are used in this relation.

Figure 2.1: Main steps of the procedure for the semiorthogonal case.

2.2 Definition of B-splines and some properties

Later we want to use B-spline functions as scaling functions. Therefore in this section we define B-spline functions and give an overview of the main properties of these functions. We adopt the theory from Schumaker and omit the proofs of the statements but all of them can be found in SCHUMAKER (1981).

First we have to give some general definitions:

Definition 2.19. *We introduce the standard notations*

$$D_+f(x) = \lim_{h \downarrow 0} \frac{f(x+h) - f(x)}{h}$$

and

$$D_-f(x) = \lim_{h \downarrow 0} \frac{f(x) - f(x-h)}{h}.$$

When these limits exist they are called right and left derivatives of f at x , respectively. When both of them exist at a point x and are equal, then we write

$$Df(x) = D_-f(x) = D_+f(x).$$

A function f is said to be differentiable on the closed interval $[a, b]$ when $Df(x)$ exists for all $a < x < b$ and $D_+f(a)$ and $D_-f(b)$ exist. Then we write

$$Df(x) = \begin{cases} D_+f(x), & \text{if } x = a \\ D_-f(x), & \text{if } x = b. \end{cases}$$

Definition 2.20. *We define*

$$(x-y)_+^0 = \begin{cases} 1, & x \geq y \\ 0, & x < y, \end{cases}$$

and

$$(x-y)_+^{m-1} = \begin{cases} (x-y)^{m-1}, & x \geq y, \quad m > 1 \\ 0, & x < y. \end{cases}$$

Definition 2.21. *Let points y_1, \dots, y_{r+1} and a function f be given. We define the r th order divided difference of f over the points y_1, \dots, y_{r+1} by*

$$[y_1, \dots, y_{r+1}]f = \frac{\det \begin{bmatrix} 1 & y_1^1 & \dots & y_1^{r-1} & f(y_1) \\ 1 & y_2^1 & \dots & y_2^{r-1} & f(y_2) \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & y_{r+1}^1 & \dots & y_{r+1}^{r-1} & f(y_{r+1}) \end{bmatrix}}{\det \begin{bmatrix} 1 & y_1^1 & \dots & y_1^r \\ 1 & y_2^1 & \dots & y_2^r \\ \vdots & \vdots & & \vdots \\ 1 & y_{r+1}^1 & \dots & y_{r+1}^r \end{bmatrix}}.$$

In this definition we have tacitly assumed that the y 's are in increasing order (in order for the determinants to make sense).

With these definitions we can finally define the B-splines:

Definition 2.22. *Given a sequence of real-valued knots*

$$\dots \leq y_{-1} \leq y_0 \leq y_1 \leq y_2 \leq \dots$$

and $i \in \mathbb{Z}$ and $m > 0$. We define

$$Q_i^m(x) = \begin{cases} (-1)^m [y_i, \dots, y_{i+m}] (x - y)_+^{m-1}, & \text{if } y_i < y_{i+m} \\ 0, & \text{otherwise} \end{cases}$$

for all $x \in \mathbb{R}$. Q_i^m is called the m th order B-spline associated with the knots y_i, \dots, y_{i+m} .

For $m = 1$, $i \in \mathbb{Z}$ and $x \in \mathbb{R}$ the B-spline associated with $y_i < y_{i+1}$ is given by the piecewise constant function

$$Q_i^1(x) = \begin{cases} \frac{1}{y_{i+1} - y_i}, & y_i \leq x < y_{i+1} \\ 0, & \text{otherwise.} \end{cases}$$

The following theorem establishes a recursion formula for the m th order B-splines.

Theorem 2.23. *Let $m \geq 2$, $i \in \mathbb{Z}$ and suppose $y_i < y_{i+m}$. Then for all $x \in \mathbb{R}$ the following recursion relation holds*

$$Q_i^m(x) = \frac{(x - y_i) Q_i^{m-1}(x) + (y_{i+m} - x) Q_{i+1}^{m-1}(x)}{(y_{i+m} - y_i)}. \quad (2.66)$$

Definition 2.24. *We define the normalized B-spline N_i^m , $m > 0$ and $i \in \mathbb{Z}$, associated with the knots y_i, \dots, y_{i+m} as*

$$N_i^m(x) := (y_{i+m} - y_i) Q_i^m(x) \quad \text{for } x \in \mathbb{R}$$

For $m = 1$, the normalized B-spline associated with $y_i < y_{i+1}$ is given by the piecewise constant function

$$N_i^1(x) = \begin{cases} 1, & y_i \leq x < y_{i+1} \\ 0, & \text{otherwise.} \end{cases}$$

From the recursion relation (2.66) we gain another recursion relation for the normalized B-Splines: Let $m \geq 2$, $i \in \mathbb{Z}$ and suppose $y_i < y_{i+m-1}$ and $y_{i+1} < y_{i+m}$ then

$$N_i^m(x) = \frac{(x - y_i)}{y_{i+m-1} - y_i} N_i^{m-1}(x) + \frac{y_{i+m} - x}{y_{i+m} - y_{i+1}} N_{i+1}^{m-1}(x) \quad \text{for } x \in \mathbb{R} \quad (2.67)$$

Now we state that the normalized B-splines span specific spaces of polynomial splines. Therefore we will now give the exact definition of the spaces of polynomial splines.

Definition 2.25. *The space of polynomials of order m is defined as follows:*

$$\mathcal{P}_m = \{p(x) \mid p(x) = \sum_{i=1}^m c_i x^{i-1}, c_1, \dots, c_m, x \in \mathbb{R}\}.$$

Definition 2.26. *Let $[a, b]$ be a finite closed interval, and let*

$$\Delta = \{x_i\}_{i=1}^k \quad \text{with } a = x_0 < x_1 < x_2 < \dots < x_k < x_{k+1} = b$$

be a partition into $k + 1$ subintervals

$$I_i = [x_i, x_{i+1}), \quad i = 0, 1, \dots, k-1 \quad \text{and } I_k = [x_k, x_{k+1}).$$

Let m be a positive integer, and let $\mathfrak{M} = [m_1, \dots, m_k]'$ be a vector of integers with $1 \leq m_i \leq m$, $i = 1, 2, \dots, k$.

We call the space

$$\mathcal{S}(\mathcal{P}_m; \mathfrak{M}; \Delta) = \{s : \text{there exist polynomials } s_0, \dots, s_k \text{ in } \mathcal{P}_m \text{ such that } s(x) = s_i(x) \text{ for } x \in I_i, \\ i = 0, 1, \dots, k, \text{ and } D^j s_{i-1}(x_i) = D^j s_i(x_i) \text{ for } j = 0, 1, \dots, m-1-m_i, i = 1, \dots, k\}$$

the space of polynomial splines of order m with knots x_1, \dots, x_k of multiplicities m_1, \dots, m_k . We call $\mathfrak{M} = [m_1, \dots, m_k]'$ the multiplicity vector.

We have to interpret $m_i = m$ to mean that the two polynomial pieces s_{i-1} and s_i are unrelated to each other, i.e. there may be a jump discontinuity at x_i . If $m_i < m$ the two polynomial pieces s_{i-1} and s_i are tied together smoothly, i.e. the spline s and its first $m - 1 - m_i$ derivatives are all continuous across the knot.

Definition 2.27. Let m be a positive integer, $\Delta = \{x_i\}_{i=1}^k$ with $a < x_1 < x_2 < \dots < x_k < b$ and $\mathfrak{M} = (m_1, \dots, m_k)'$ with $1 \leq m_i \leq m$, $i = 1, 2, \dots, k$ be given. Let $K = \sum_{i=1}^k m_i$. Suppose

$$y_1 \leq y_2 \leq \dots \leq y_{2m+K}$$

is such that

$$y_1 \leq \dots \leq y_m \leq a, \quad b \leq y_{m+K+1} \leq \dots \leq y_{2m+K} \quad (2.68)$$

and

$$(y_{m+1}, \dots, y_{m+K}) = (\overbrace{x_1, \dots, x_1}^{m_1}, \dots, \overbrace{x_k, \dots, x_k}^{m_k}). \quad (2.69)$$

Then we call $\tilde{\Delta} = \{y_i\}_1^{2m+K}$ an extended partition associated with $\mathcal{S}(\mathcal{P}_m; \mathfrak{M}; \Delta)$.

Remark:

The points $\{y_i\}_{i=m+1}^{m+K}$ in an extended partition $\tilde{\Delta}$ associated with $\mathcal{S}(\mathcal{P}_m; \mathfrak{M}; \Delta)$ are uniquely determined (compare (2.69)). The first and last m points in $\tilde{\Delta}$ can be chosen arbitrarily, subject to (2.68).

Theorem 2.28. Let $\tilde{\Delta} = \{y_i\}_1^{2m+K}$ be an extended partition associated with $\mathcal{S}(\mathcal{P}_m; \mathfrak{M}; \Delta)$, and suppose $b < y_{2m+K}$. For $i = 1, 2, \dots, m + K$, let N_i^m be the normalized B-spline

$$N_i^m(x) = (-1)^m (y_{i+m} - y_i) [y_i, \dots, y_{i+m}] (x - y)_+^{m-1}, \quad a \leq x \leq b.$$

Then $\{N_i^m\}_{i=1}^{m+K}$ form a basis for $\mathcal{S}(\mathcal{P}_m; \mathfrak{M}; \Delta)$ with

$$N_i^m(x) = 0 \quad \text{for } x \notin [y_i, y_{i+m}]$$

and

$$N_i^m(x) > 0 \quad \text{for } x \in (y_i, y_{i+m}).$$

Moreover the normalized B-splines form a partition of unity, i.e.

$$\sum_{i=1}^{m+K} N_i^m(x) = 1 \quad \text{for all } a \leq x \leq b.$$

Remark

A corollary proves the statements of Theorem 2.28 for $b = y_{m+K+1} = \dots = y_{2m+K}$. In our application we will need this case but we will only refer to the theorem. For more details see SCHUMAKER (1981), Corollary 4.10.

The theorem states that the normalized B-splines $\{N_i^m\}_{i=1}^{m+K}$ associated with the knots in $\tilde{\Delta}$ form a basis for $\mathcal{S}(\mathcal{P}_m; \mathfrak{M}; \Delta)$. Furthermore the theorem shows that the normalized B-splines N_i^m , $i \in \mathbb{Z}$ and $m > 0$, are compactly supported and positive.

Note, for $m = 1$ the B-splines $Q_i^1(x)$ and $N_i^1(x)$ are piecewise constant functions, for $m = 2$ the B-splines $Q_i^2(x)$ and $N_i^2(x)$ are piecewise linear functions and so on. The notation here might be misleading. There are different notations in the literature.

Now we consider B-splines with equally spaced knots: We say that the set of knots $\dots, y_i, y_{i+1}, \dots$ is uniform with spacing $h \in \mathbb{R}$ provided

$$y_{i+1} - y_i = h \quad \text{for all } i \in \mathbb{Z}.$$

For uniformly spaced knots it turns out that any B-spline can be obtained from one basic B-spline by translation and scaling.

Definition 2.29. *Let*

$$Q^m(x) = \sum_{i=0}^m \frac{(-1)^i \binom{m}{i} (x-i)_+^{m-1}}{m!}, \quad x \in \mathbb{R}$$

(This is the usual B-spline associated with the simple knots $0, 1, \dots, m$). Associated with Q^m , we also introduce the normalized version

$$N^m(x) = m Q^m(x) \quad x \in \mathbb{R}.$$

The following theorem shows that any B-spline associated with uniformly spaced knots can be obtained from Q^m or N^m by a translation (and possible scaling). With this theorem we will later see that our inner scaling functions are generated by the function N^m .

Theorem 2.30. *Let $i \in \mathbb{Z}$, $m > 0$, $h \in \mathbb{R}$ and $h > 0$. Suppose y_i, \dots, y_{i+m} are uniformly spaced with spacing h . Then*

$$Q_i^m(x) = \frac{1}{h} Q^m\left(\frac{x - y_i}{h}\right) \quad x \in \mathbb{R}$$

and

$$N_i^m(x) = N^m\left(\frac{x - y_i}{h}\right) \quad x \in \mathbb{R}.$$

2.3 1-dimensional B-spline model

We introduce a 1-dimensional model as base for the multi-dimensional models.

Since in our applications we want to work regionally, we choose basis functions on the bounded interval $[0, 1]$. Then we have to take into account the behavior at the endpoints of the interval $[0, 1]$, i.e. we have to introduce special boundary functions (compare section 2.1.5). A big advantage of the spline approach is that it readily adapts to the case of the bounded interval $[0, 1]$ by introducing multiple knots at the endpoints.

2.3.1 Scaling functions and detail spaces

Now we intend to apply normalized B-Splines as basis functions $\phi_{j,k} = N_k^m(x)$ (the level $j \in \mathbb{N}_0$ will specify the knot sequence for the B-splines and we will also write $N_k^{m,j}$), we define $m_j := 2^j + (m - 1)$ (there will be m_j basis functions, hence our spaces V_j will have dimension m_j).

Our scaling spaces will be $V_j = \mathcal{S}(\mathcal{P}_3; \mathfrak{M}_j; \Delta_j)$ with $\Delta_j = \{\frac{1}{2^j}, \frac{2}{2^j}, \dots, \frac{2^j-1}{2^j}\}$ and the corresponding $2^j - 1 \times 1$ multiplicity vector $\mathfrak{M}_j = [1, \dots, 1]'$, $j \in \mathbb{N}_0$.

For a fixed j we want to define the basis functions $\phi_{j,k}$, $k = 0, \dots, m_j - 1$ (Note, from now on the notation $\phi_{j,k}$ does not mean only the translated and dilated versions of a function ϕ as in Definition 2.4. As introduced in chapter 2.1.5 for a MSA on a bounded interval we need additional boundary scaling functions that in our case cannot be derived by translation and scaling. We use this notation because for our basis functions $\phi_{j,k}$ we intend to adopt the concepts discussed in chapter 2):

As basis functions $\phi_{j,k}$ we apply a normalized B-spline $N_k^{m,j}(x)$ of order $m = 3$ and shift $k = 0, \dots, m_j - 1$ ($m_j = 2^j + (m - 1) = 2^j + 2$). The index j specifies the knot sequence for the B-splines. For a fixed $j \in \mathbb{N}_0$ let the knot sequence $t_0^j \leq t_1^j \leq \dots \leq t_{m_j+2}^j$ be given through

$$0 = t_0^j = t_1^j = t_2^j < t_3^j < t_4^j < \dots < t_{m_j-1}^j < t_{m_j}^j = t_{m_j+1}^j = t_{m_j+2}^j = 1 \quad (2.70)$$

with $t_l^j = (l - 2)h_j$, $h_j = \frac{1}{2^j}$ for $l = 3, \dots, m_j - 1$ (Note, this knot sequence forms an extended partition associated with $\mathcal{S}(\mathcal{P}_3; \mathfrak{M}_j; \Delta_j)$ (compare Definition 2.27), i.e. we may choose the first and last $m = 3$ knots arbitrarily, compare equation (2.68)).

We define our m_j basis function $\phi_{j,k}(x) := N_k^{3,j}(x)$, $k = 0, \dots, m_j - 1$ for $x \in [0, 1]$, to be the 3rd order normalized B-spline associated with the knots t_k^j, \dots, t_{k+3}^j , i.e.

$$N_k^{1,j}(x) := \begin{cases} 1, & t_k^j \leq x \leq t_{k+1}^j \\ 0, & \text{otherwise} \end{cases}$$

$$N_k^{m,j}(x) := \frac{x - t_k^j}{t_{k+m-1}^j - t_k^j} N_k^{m-1,j}(x) + \frac{t_{k+m}^j - x}{t_{k+m}^j - t_{k+1}^j} N_{k+1}^{m-1,j}(x) \text{ for } m \geq 2, k = 0, \dots, m_j - 1$$

for $x \in [0, 1]$. Note, the fractions in these equations are taken to be 0 when their denominators are 0, this concept we adopt from STOLLNITZ ET AL. (1995). And compare the recursion formula for normalized B-splines (2.67).

For $m = 3$ we have altogether $m_j = 2^j + 2$ basis functions $\phi_{j,k}(x)$, $k = 0, \dots, m_j - 1$, for a fixed $j \in \mathbb{N}_0$.

For $k = 2, \dots, m_j - 3$ the functions $\phi_{j,k}(x) = N_k^{3,j}(x)$ are just B-splines associated with the uniformly spaced knots t_k^j, \dots, t_{k+3}^j hence they are a translation and possible scaling of the cardinal B-spline N^3 (see Theorem 2.30), i.e. $\phi_{j,k}(x) = N_k^{3,j}(x) = N^3(2^j x - k)$ for $k = 2, \dots, m_j - 3$. For being the translated and dilated versions of a single function ϕ only the normalization constant $2^{j/2}$ is missing (compare the Definition 2.1), but in the construction of computational algorithms it is more convenient to drop it. This only changes the Riesz bounds by a factor of 2^{-j} .

For $k = 0, 1$ the resulting scaling functions $\phi_{j,k}(x) = N_k^{3,j}(x)$ form the boundary scaling functions for the endpoint 0 (the associated knots of the B-spline contain a multiple knot at 0), for $k = m_j - 2, m_j - 1$ the resulting scaling functions $\phi_{j,k}(x) = N_k^{3,j}(x)$ form the boundary scaling functions for the endpoint 1 (the associated knots of the B-spline contain a multiple knot at 1). These boundary scaling functions are not associated with uniformly spaced knots and therefore cannot be derived by translation and scaling of the B-spline N^3 .

We use the functions $\phi_{j,k}$, $k = 0, \dots, m_j - 1$, as basis functions for the scaling spaces V_j and therefore we will also use the denotation scaling functions. But we should keep in mind that not all of the $\phi_{j,k}$ are generated by a function ϕ . Figure 2.2 shows the scaling functions $\phi_{j,k}$, $k = 0, \dots, m_j - 1$, for the levels $j = 0, \dots, 3$. For level $j = 0, 1$ (first and second panel) all functions $\phi_{j,k}(x)$, $k = 0, \dots, m_j - 1$ are affected by the endpoint interpolating procedure due to the chose knot sequence (2.70). In the third and fourth panel we can see, that for the levels $j = 2$ and $j = 3$ only for $k = 0, 1$ and $k = m_j - 2, m_j - 1$ the functions $\phi_{j,k}(x)$ are affected by the endpoint interpolating procedure due to the knot sequence (2.70); in contrast the inner functions for $k = 3, \dots, m_j - 3$ are not affected by the endpoint interpolating procedure. The B-splines are real-valued, from now on we will only consider real valued functions and therefore we only work with real numbers.

Definition 2.31. We denote the closed spaces spanned by the m_j basis functions $\phi_{j,k}$, $k = 0, \dots, m_j - 1$, as V_j , i.e. $V_j := \overline{\text{span}\{\phi_{j,k} \mid k = 0, \dots, m_j - 1\}}^{L^2([0,1])}$.

As we work on the bounded interval we should use the notation $V_j^{[0,1]}$ but for simplification we will just write V_j , $j \in \mathbb{N}_0$.

According to Theorem 2.28 or STOLLNITZ ET AL. (1995) the functions $\phi_{j,k}$ form a basis of the space $\mathcal{S}(\mathcal{P}_3; \mathfrak{M}_j; \Delta_j)$ with $\Delta_j = \{\frac{1}{2^j}, \frac{2}{2^j}, \dots, \frac{2^j-1}{2^j}\}$ and the corresponding $2^j - 1 \times 1$ multiplicity vector $\mathfrak{M}_j = [1, \dots, 1]'$, i.e.

$$V_j = \mathcal{S}(\mathcal{P}_3; \mathfrak{M}_j; \Delta_j).$$

Now we check the conditions on V_j to form a MSA on the bounded interval $[0, 1]$:

- (1) Since $\Delta_j = \{\frac{1}{2^j}, \frac{2}{2^j}, \dots, \frac{2^j-1}{2^j}\} \subset \Delta_{j+1} = \{\frac{1}{2^{j+1}}, \frac{2}{2^{j+1}}, \dots, \frac{2^{j+1}-1}{2^{j+1}}\}$ it follows that $\mathcal{S}(\mathcal{P}_3; \mathfrak{M}_j; \Delta_j) \subset \mathcal{S}(\mathcal{P}_3; \mathfrak{M}_{j+1}; \Delta_{j+1})$ holds for $j \geq 0$ and the first condition of a MSA on a bounded interval is fulfilled, i.e. equation (2.63)

$$V_0 \subset V_1 \subset \dots \tag{2.71}$$

- (2) We have to check $\overline{\bigcup_{j \in \mathbb{N}_0} V_j}^{L^2([0,1])} = \overline{\bigcup_{j \in \mathbb{N}_0} \mathcal{S}(\mathcal{P}_3; \mathfrak{M}_j; \Delta_j)}^{L^2([0,1])} = L^2([0, 1])$.
Let χ_A be the characteristic function of the set A ,

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

We define the set of step functions \mathcal{E} defined on $[0,1]$

$$\mathcal{E} = \{f(x) = \sum_{k=0}^M a_k \chi_{A_k}(x) \text{ with } a_k \in \mathbb{R} \text{ for } k = 0, \dots, M \text{ and } \{A_k\}_{k=0}^M \text{ form a partition of the interval } [0, 1], x \in [0, 1]\}$$

and the sets

$$\mathcal{E}_j = \{f(x) = \sum_{k=0}^{2^j-1} b_k^j \chi_{B_k^j}(x) \text{ with } b_k^j \in \mathbb{R} \text{ for } k = 0, \dots, 2^j - 1 \text{ and } B_k^j = [\frac{k}{2^j}, \frac{k+1}{2^j}) \text{ for } k = 0, \dots, 2^j - 2, B_{2^j-1}^j = [\frac{k}{2^j}, \frac{k+1}{2^j}] \text{ for } k = 2^j - 1, x \in [0, 1]\}.$$

We can show that $\overline{\bigcup_{j \in \mathbb{N}_0} \mathcal{E}_j}^{L^2([0,1])} = \mathcal{E}$, we will not perform the proof (the main argument is that $|B_k^j| = h_j = \frac{1}{2^j} \xrightarrow{j \rightarrow \infty} 0$). The set \mathcal{E} lies dense in $L^2([0, 1])$, for the proof see, e.g., ALT (1999), therefore it holds $\overline{\bigcup_{j \in \mathbb{N}_0} \mathcal{E}_j}^{L^2([0,1])} = L^2([0, 1])$. Since $\mathcal{E}_j \subset \mathcal{S}(\mathcal{P}_3; \mathfrak{M}_j; \Delta_j) \subset L^2([0, 1])$ and $L^2([0, 1])$ is complete (see, e.g., WEIDMANN (1976), section 2.1) the assertion follows, i.e.

$$\overline{\bigcup_{j \in \mathbb{N}_0} V_j^{[0,1]}}^{L^2([0,1])} = \overline{\bigcup_{j \in \mathbb{N}_0} \mathcal{S}(\mathcal{P}_3; \mathfrak{M}_j; \Delta_j)}^{L^2([0,1])} = L^2([0, 1]). \quad (2.72)$$

(3) Since we have only a finite set of basis functions for the scaling spaces V_j , $j \in \mathbb{N}_0$ we will only have finite

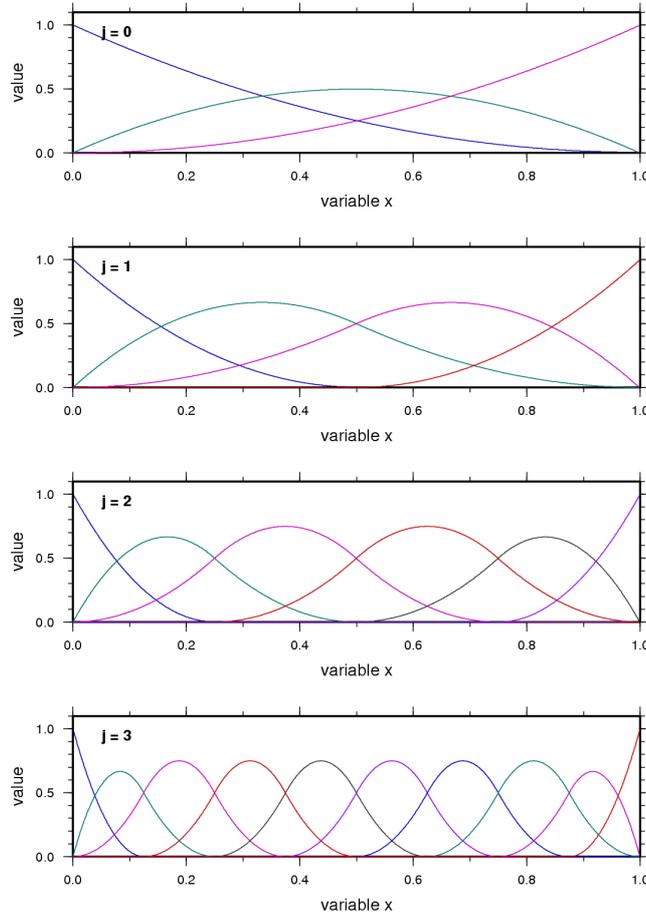


Figure 2.2: *B-spline scaling functions $\phi_{j,k}(x)$, $k = 0, \dots, m_j - 1$, for the levels $j = 0, \dots, 3$.*

the first column of the matrix is $p_{0,0}^j$, i.e. the indices are shifted.

The entries of this matrix $p_{k,l}^j$ could also be calculated from equation (2.40) with $p_{k,l}^j = \langle \phi_{j-1,l}, \tilde{\phi}_{j,k} \rangle_{L^2(\mathbb{R})}$. But we do not work with the dual functions $\tilde{\phi}_{j,k}$ and therefore we do not calculate the coefficients $p_{k,l}^j$ this way.

2.3.3 Wavelets and detail spaces - the second two-scale relation

The **detail spaces** W_{j-1} are defined as the orthogonal complement of the scaling space V_{j-1} in the scaling space V_j , i.e.

$$V_j = V_{j-1} \oplus W_{j-1}. \quad (2.75)$$

Now we intend to construct a set of basis functions $\psi_{j-1,k}(x)$ of the detail space W_{j-1} with $k = 0, \dots, n_{j-1} - 1$ and $n_{j-1} := m_j - m_{j-1}$ (i.e. $m_j = m_{j-1} + n_{j-1}$, the dimension of V_j is the sum of the dimensions of V_{j-1} and W_{j-1}). As we work on a bounded interval we have inner basis functions that can be controlled by dilation and translation of a single functions ψ and we have additional boundary basis functions that cannot be controlled by dilation and translation of the function ψ . Hence we cannot have a function ψ that generates all the basis functions $\psi_{j,k}$, $j \in \mathbb{N}_0$ and $k = 0, \dots, n_j - 1$. Therefore we will only presume that the weaker condition (2.15) of the admissibility condition (2.14) holds for each $\psi_{j,k}$, $j \in \mathbb{N}_0$ and $k = 0, \dots, n_j - 1$, i.e.

$$\int_0^1 \psi_{j,k}(x) dx = 0 \quad j \in \mathbb{N}_0, \quad k = 0, \dots, m_j - 1. \quad (2.76)$$

Now we want to construct basis functions $\psi_{j,k}$ that fulfill the condition (2.76) and we will call these functions wavelets.

Since $W_{j-1} \subset V_j$, we can represent the wavelet functions $\psi_{j-1,l} \in W_{j-1} \subset V_j$ with $l = 0, \dots, n_{j-1} - 1$ as a linear combination of the scaling functions $\phi_{j,k} \in V_j$ with $k = 0, \dots, m_j - 1$. I.e. we introduce the second **two-scale relation**

$$\psi_{j-1,l}(x) = \sum_{k=1}^{m_j-1} q_{k,l}^j \phi_{j,k}(x) \quad (2.77)$$

(compare equation (2.41)). With the $n_{j-1} \times 1$ vector

$$\Psi_{j-1}(x) = [\psi_{j-1,0}(x), \psi_{j-1,1}(x), \dots, \psi_{j-1,l}(x), \dots, \psi_{j-1,n_{j-1}-1}(x)]'$$

of resolution level $j - 1$ we rewrite the second two-scale relation (2.77) as the matrix equation

$$\Psi'_{j-1}(x) = \Phi'_j(x) \mathbf{Q}_j, \quad (2.78)$$

wherein $\mathbf{Q}_j = (q_{k,l}^j)$ is an $m_j \times n_{j-1}$ matrix with initially unknown elements $q_{k,l}^j$. Note, as for the matrix \mathbf{P}_j the indices are shifted and the entry in the first row and the first column of the matrix \mathbf{Q}_j is $q_{0,0}^j$. Hence, our next objective is the determination of the matrix \mathbf{Q}_j from the given matrix \mathbf{P}_j .

Since, as mentioned before, the wavelet functions $\psi_{j-1,k}(x)$ are assumed to be orthogonal to the scaling functions $\phi_{j-1,l}(x)$, their inner product vanishes, i.e.

$$\int_0^1 \phi_{j-1,l}(x) \psi_{j-1,k}(x) dx =: \langle \phi_{j-1,l}, \psi_{j-1,k} \rangle_{L^2([0,1])} = 0. \quad (2.79)$$

Note wavelet functions which fulfill this condition, i.e. they are orthogonal to the scaling functions but not orthogonal to each other fulfill the first condition of the definition of the semiorthogonal concept, i.e. $V_j \perp W_j$ (Definition 2.18). As mentioned before we will not work with the dual functions and therefore we will not check the existence of the dual functions.

Substituting the two-scale equation (2.77) into equation (2.79) yields

$$\sum_{m=0}^{m_j-1} q_{m,k}^j \langle \phi_{j-1,n}, \phi_{j,m} \rangle_{L^2([0,1])} = 0.$$

Writing this equation as a matrix equation gives us:

$$\mathbf{G}_j \mathbf{Q}_j = \mathbf{0} \quad (2.80)$$

wherein $\mathbf{0}$ is the $m_{j-1} \times n_{j-1}$ zero matrix, and $\mathbf{G}_j = (g_{m,n}^j)$ with $g_{m,n}^j = \langle \phi_{j-1,m}, \phi_{j,n} \rangle_{L^2([0,1])}$, $m = 0, \dots, m_{j-1} - 1$ and $n = 0, \dots, m_j - 1$. Hence again the indices are shifted, i.e. the entry in the first row and in the first column of the matrix is $g_{0,0}^j$.

A matrix equation with right hand-side of zero like this one is an homogeneous system of equations. The set of all possible solutions is called the null space of \mathbf{G}_j . The columns of \mathbf{Q}_j must form a basis of the nullspace. Since there are various bases of the nullspace of a matrix, there are various wavelet bases for a given detail space W_j . We impose further constraints in addition to the orthogonality requirement given above and equation (2.76). Finally we can determine the matrix \mathbf{Q}_j .

For our application we did not calculate the matrix \mathbf{Q}_j , $j \in \mathbb{N}$, but we worked with the matrix calculated in STOLLNITZ ET AL. (1995). The matrices can be found in section C. The constraint that has been introduced to derive a unique solution is explained in the following. To get wavelets with small supports there was required each column of \mathbf{Q}_j to have a minimal number of consecutive non-zeros. This constraint imposes a banded structure on \mathbf{Q}_j similar to that of \mathbf{P}_j .

From the structure of the matrix \mathbf{Q}_j we can see that only the wavelets $\psi_{j,k}$, $k = 0, 1, n_j - 2, n_j - 1$ are affected by the endpoint interpolating procedure, while the inner wavelets $\psi_{j,k}$, for $k = 3, \dots, n_j - 3$, are not affected by the endpoint interpolating procedure. The wavelet functions are shown in Figure C.1. One can check that the wavelet functions $\psi_{j,k}$ defined via the matrix \mathbf{Q}_j in fact fulfill the condition (2.76), for the inner wavelet functions this is an easy computation while for the boundary wavelets this is more complicated. We will not perform these computations here.

A construction of explicit formulas of the wavelet functions $\psi_{j,k}$ can be found in CHUI AND QUAK (1992).

2.3.4 Construction of a decomposition relation and a reconstruction relation

In the next step we want to derive the decomposition relation, which is required for the MSA. From the equation $V_j = V_{j-1} \oplus W_{j-1}$ it follows that there exist matrices of real coefficients $\overline{\mathbf{P}}_j$ and $\overline{\mathbf{Q}}_j$ such that

$$\Phi'_j(x) = \Phi'_{j-1}(x) \overline{\mathbf{P}}_j + \Psi'_{j-1}(x) \overline{\mathbf{Q}}_j \quad (2.81)$$

wherein $\overline{\mathbf{P}}_j$ and $\overline{\mathbf{Q}}_j$ are $m_{j-1} \times m_j$ and $n_{j-1} \times m_j$ unknown coefficient matrices, respectively. Inserting the two-scale equations (2.74) and (2.78) yields

$$\Phi'_j(x) = \Phi'_j(x) \mathbf{P}_j \overline{\mathbf{P}}_j + \Phi'_j(x) \mathbf{Q}_j \overline{\mathbf{Q}}_j \quad (2.82)$$

and therefore it follows

$$\mathbf{P}_j \overline{\mathbf{P}}_j + \mathbf{Q}_j \overline{\mathbf{Q}}_j = \begin{bmatrix} \mathbf{P}_j & \mathbf{Q}_j \end{bmatrix} \begin{bmatrix} \overline{\mathbf{P}}_j \\ \overline{\mathbf{Q}}_j \end{bmatrix} = \mathbf{I}.$$

The $m_j \times m_j$ matrix $\begin{bmatrix} \mathbf{P}_j & \mathbf{Q}_j \end{bmatrix}$ is of full rank it follows

$$\begin{bmatrix} \overline{\mathbf{P}}_j \\ \overline{\mathbf{Q}}_j \end{bmatrix} = \begin{bmatrix} \mathbf{P}_j & \mathbf{Q}_j \end{bmatrix}^{-1} \quad (2.83)$$

Similar to the general case we can decompose a function $f_J(x)$ with $f_J \in V_J$ into

$$\begin{aligned} f_J(x) &= f_0(x) + \sum_{i=0}^{J-1} g_i(x) \\ &= f_{j'}(x) + \sum_{i=j'}^{J-1} g_i(x) \end{aligned} \quad (2.84)$$

(compare equation (2.65)) wherein $f_i \in V_i$ and g_i in W_i are defined as

$$f_i(x) = \sum_{k=0}^{m_i-1} d_{i,k} \phi_{i,k}(x) = \Phi'_i(x) \mathbf{d}_i \quad (2.85)$$

and

$$g_i(x) = \sum_{k=0}^{n_i-1} c_{i,k} \psi_{i,k}(x) = \Psi'_i(x) \mathbf{c}_i \quad (2.86)$$

wherein $\mathbf{d}_i = [d_{i,0} \dots d_{i,m_i-1}]'$ and $\mathbf{c}_i = [c_{i,0} \dots c_{i,n_i-1}]'$.

Analogous to the general equations (2.51) and (2.53) we could compute the coefficients $d_{i,k}$ and $c_{i,k}$ as scalar products with the dual functions. But we will not use explicit formulas for the dual functions as already mentioned before and will therefore not use this technique.

Our main purpose is to derive the level $j-1$ coefficient vectors \mathbf{d}_{j-1} and \mathbf{c}_{j-1} from the level j coefficient vector \mathbf{d}_j , i.e. to derive decomposition relations similar to equation (2.58) and (2.59) and a reconstruction relation similar to equation (2.61).

Since the signal $f_j(x)$ can be expressed as $f_j(x) = f_{j-1}(x) + g_{j-1}(x)$ according to equation (2.84) it follows with equations (2.85) and (2.86)

$$f_j(x) = \Phi'_j(x) \mathbf{d}_j \quad (2.87)$$

$$= \Phi'_{j-1}(x) \mathbf{d}_{j-1} + \Psi'_{j-1}(x) \mathbf{c}_{j-1}. \quad (2.88)$$

Substituting the decomposition equation (2.81) for $\Phi'_j(x)$ into equation (2.87) yields

$$f_j(x) = \Phi'_{j-1}(x) \overline{\mathbf{P}}_j \mathbf{d}_j + \Psi'_{j-1}(x) \overline{\mathbf{Q}}_j \mathbf{d}_j. \quad (2.89)$$

Comparing the equations (2.88) and (2.89) yields us to the **decomposition relations**

$$\mathbf{d}_{j-1} = \overline{\mathbf{P}}_j \mathbf{d}_j \quad \text{and} \quad (2.90)$$

$$\mathbf{c}_{j-1} = \overline{\mathbf{Q}}_j \mathbf{d}_j. \quad (2.91)$$

If we rewrite these relations as

$$\begin{bmatrix} \mathbf{d}_{j-1} \\ \mathbf{c}_{j-1} \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{P}}_j \\ \overline{\mathbf{Q}}_j \end{bmatrix} \mathbf{d}_j \quad (2.92)$$

we get the **reconstruction relation** by solving the equation for \mathbf{d}_j under consideration of equation (2.83):

$$\mathbf{d}_j = \mathbf{P}_j \mathbf{d}_{j-1} + \mathbf{Q}_j \mathbf{c}_{j-1}. \quad (2.93)$$

Therefore we constructed the decomposition relations and the reconstruction relation for the B-spline model. With those equations we can now decompose signals $f_j \in V_j$ into a smoothed signal $f_{j-1} \in V_{j-1}$ and the corresponding detail signal $g_{j-1} \in W_{j-1}$. From the signals f_{j-1} and g_{j-1} we can reconstruct the signal f_j . Both the decomposition and the reconstruction is performed via the coefficient vectors.

2.3.5 Decomposition and reconstruction for the 1-dimensional B-spline model

Now we introduce the B-spline approach for signals $f \in L^2([0, 1])$.

For a signal $f \in L^2([0, 1])$ we want to compute an approach $f_J \in V_J$ of resolution level J . According to the equation $V_J = V_{J-1} \oplus W_{J-1}$ we decompose $f_J = f_{J-1} + g_{J-1}$ wherein the smoothed version $f_{J-1} \in V_{J-1}$ and the detail signal $g_{J-1} \in W_{J-1}$. As introduced in (2.85) and (2.86) the functions f_J and g_J are defined via their coefficient vectors \mathbf{d}_J and \mathbf{c}_J , respectively. With the decomposition relations (2.90) and (2.91) we compute the scaling coefficients of the lower levels and the corresponding detail coefficients down to resolution level 0. I.e. from the level J coefficient vector \mathbf{d}_J we compute the scaling coefficient vector \mathbf{d}_0 and the detail coefficient vectors $\mathbf{c}_0, \dots, \mathbf{c}_{J-1}$. Then we reconstruct the approach $f_J(t)$ from the coefficient vectors \mathbf{d}_0 and $\mathbf{c}_0, \dots, \mathbf{c}_{J-1}$ by the reconstruction relation (2.93).

$$\text{Suppose we have measurements } \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_n) \end{bmatrix} + \begin{bmatrix} e(t_1) \\ e(t_2) \\ \vdots \\ e(t_n) \end{bmatrix} \quad \text{of the signal } f(t) \text{ given at discrete points}$$

$t_1, t_2, \dots, t_n \in [0, 1]$.

- **computation of the scaling coefficients**

In the first step we estimate the scaling coefficient vector $\mathbf{d}_J = [d_{J,0}, \dots, d_{J,m_J-1}]'$ of resolution level J . Therefore we establish the following linear equation system

$$\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} \phi_{J,0}(t_1) & \phi_{J,1}(t_1) & \dots & \phi_{J,m_J-1}(t_1) \\ \phi_{J,0}(t_2) & \phi_{J,1}(t_2) & \dots & \phi_{J,m_J-1}(t_2) \\ \vdots & \vdots & & \vdots \\ \phi_{J,0}(t_n) & \phi_{J,1}(t_n) & \dots & \phi_{J,m_J-1}(t_n) \end{bmatrix} \begin{bmatrix} d_{J,0} \\ d_{J,1} \\ \vdots \\ d_{J,m_J-1} \end{bmatrix}. \quad (2.94)$$

The right hand-side contains the unknown scaling coefficients $d_{J,0}, \dots, d_{J,m_J-1}$. Each scaling coefficient $d_{J,k}$ corresponds to the compactly supported scaling function $\phi_{J,k}$, $k = 1, \dots, m_J - 1$. Of course a scaling coefficient is only computable if there is given at least one measurement in the support of the corresponding scaling function.

With the notations $\mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$, $\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$ and $\Phi_J = \begin{bmatrix} \phi_{J,0}(t_1) & \phi_{J,1}(t_1) & \dots & \phi_{J,m_J-1}(t_1) \\ \phi_{J,0}(t_2) & \phi_{J,1}(t_2) & \dots & \phi_{J,m_J-1}(t_2) \\ \vdots & \vdots & & \vdots \\ \phi_{J,0}(t_n) & \phi_{J,1}(t_n) & \dots & \phi_{J,m_J-1}(t_n) \end{bmatrix}$ we rewrite equation (2.94) in a short notation as

$$\mathbf{f} + \mathbf{e} = \Phi_J \mathbf{d}_J. \quad (2.95)$$

Hence we assume to have a so-called Gauss-Markov model

$$\mathbf{f} + \mathbf{e} = \Phi_J \mathbf{d}_J \text{ with } D(\mathbf{f}) = \sigma_{\mathbf{f}}^2 \mathbf{P}_{\mathbf{f}}^{-1} \quad (2.96)$$

wherein $D(\mathbf{f})$ is the covariance matrix given by the unknown variance factor $\sigma_{\mathbf{f}}^2$ and the given positive definite weight matrix $\mathbf{P}_{\mathbf{f}}$ of dimension $n \times n$.

The least squares method leads us to the following normal equation (for more details see e.g. KOCH (1999)):

$$(\Phi_J' D(\mathbf{f})^{-1} \Phi_J) \cdot \hat{\mathbf{d}}_J = \Phi_J' D(\mathbf{f})^{-1} \mathbf{y}. \quad (2.97)$$

Now we have to consider two cases for the $n \times m_J$ matrix Φ_J , with $n \geq m_J$:

- either the $n \times m_J$ matrix Φ_J is of full column rank, i.e. the column rank amounts m_J and therefore the matrix $(\Phi_J' D(\mathbf{f})^{-1} \Phi_J)$ in the normal equation (2.97) is invertible
- or the matrix Φ_J is not of full column rank, i.e. the column rank amounts less than m_J and therefore the matrix $(\Phi_J' D(\mathbf{f})^{-1} \Phi_J)$ is not invertible.

Since the scaling functions $\phi_{J,k}$, $k = 0, \dots, m_J - 1$ are linear independent (compare Theorem 2.28) a column rank deficiency can only be a result of the distribution of the data. I.e. the matrix will be of full rank if the t_1, \dots, t_n are evenly distributed such that at least one t_i , $i = 1, \dots, n$, lies in the support of each scaling function $\phi_{J,k}$, $k = 0, \dots, m_J - 1$, for more details see section 3.3.2. The matrix will not be of full column rank, if for at least one scaling function $\phi_{J,k}$, $k = 0, \dots, m_J - 1$, there is no observation given in its support, i.e. none of the t_i , $i = 1, \dots, n$, lies in the support of the scaling function. Hence in case of data gaps the matrix will not be of full column rank.

In the case that Φ_J is of full column rank we can invert the matrix $(\Phi_J' D(\mathbf{f})^{-1} \Phi_J)$ and therefore we estimate a unique least squares solution of the Gauss-Markov model (2.96) by

$$\hat{\mathbf{d}}_J = (\Phi_J' D(\mathbf{f})^{-1} \Phi_J)^{-1} \Phi_J' D(\mathbf{f})^{-1} \mathbf{f} \quad (2.98)$$

see e.g. KOCH (1999).

In the other case if the matrix Φ_J is not of full column rank we cannot invert the matrix $(\Phi_J' D(\mathbf{f})^{-1} \Phi_J)$ and hence a unique least squares solution of the Gauss-Markov model (2.96) does not exist. In this case none of the t_1, \dots, t_n lies in the support of a function $\phi_{J,k}$ and we cannot estimate the corresponding detail coefficient $d_{J,k}$. We may then cancel out the corresponding addends from equation (2.94) and finally achieve a reduced matrix Φ_J of full column rank and can proceed as in the first case. In our applications we will also apply another method: we introduce prior information to stabilize the estimation process, i.e. we define an additional linear model for the prior information. We do not discuss this method here but more details will be presented in section 3.3.2.

- **decomposition of the signal**

In the second step we want to compute the smoothed signal $f_J \in V_J$ of the original signal $f \in L^2([0, 1])$. With the coefficient vector \mathbf{d}_J we can define analogous to equation (2.85)

$$f_J(x) := \Phi'_J(x) \mathbf{d}_J,$$

wherein $\Phi_J(x) = [\phi_{J,0}(x), \phi_{J,1}(x), \dots, \phi_{J,m_{J-1}}(x)]'$. Clearly $f_J \in V_J$ and since $V_J \subset L^2([0, 1])$ we call it a level J approach of the original signal.

With the decomposition relations (2.90) and (2.91)

$$\mathbf{d}_{j-1} = \overline{\mathbf{P}}_j \mathbf{d}_j,$$

$$\mathbf{c}_{j-1} = \overline{\mathbf{Q}}_j \mathbf{d}_j$$

we can compute the level $J-1$ coefficient vectors \mathbf{d}_{J-1} and \mathbf{c}_{J-1} from the level J scaling coefficient vector \mathbf{d}_J . Therefore we can stepwise decompose the coefficients down to level j_0 , $0 \leq j_0 \leq J-1$, i.e. we can compute the coefficient vectors \mathbf{d}_{j_0} and $\mathbf{c}_{j_0}, \dots, \mathbf{c}_{J-1}$ (compare the decomposition scheme (2.60)). With these coefficients we can calculate the corresponding signals f_{j_0} and g_i , $i = j_0, \dots, J-1$ via equations (2.85) and (2.86) and get the decomposition

$$f_J(x) = f_{j_0}(x) + \sum_{i=j_0}^{J-1} g_i(x).$$

We just have to save the scaling coefficient vector of the lowest resolution level \mathbf{d}_{j_0} and all the detail coefficient vectors $\mathbf{c}_{j_0}, \dots, \mathbf{c}_{J-1}$; from those we can perform a reconstruction of the signal as described in the next step.

Note, the coefficient vector \mathbf{d}_J saves the structures of the signal $f_J \in V_J$. With the decomposition relation we get the coefficient vectors \mathbf{d}_{J-1} and \mathbf{c}_{J-1} . In the coefficient vector \mathbf{d}_{J-1} the coarser structures of the signal are saved while in the coefficient vector \mathbf{c}_{J-1} the fine structures that distinguish V_J from V_{J-1} are saved. Note, the number of coefficients in \mathbf{d}_{J-1} is less than the number of coefficients in \mathbf{d}_J , i.e. to save coarser structures clearly less coefficients are necessary than to save the fine structures. But of course the number of the coefficients in \mathbf{d}_J is equal to the sum of the number of coefficients in \mathbf{d}_{J-1} and the number of coefficients in \mathbf{c}_{J-1} , i.e. $m_J = m_{J-1} + n_{J-1}$.

- **reconstruction of the signal**

From the reconstruction relation (2.93)

$$\mathbf{d}_J = \mathbf{P}_J \mathbf{d}_{J-1} + \mathbf{Q}_J \mathbf{c}_{J-1}$$

we can calculate the level J scaling coefficient vector \mathbf{d}_J from the level $J-1$ scaling and wavelet coefficient vectors \mathbf{d}_{J-1} and \mathbf{c}_{J-1} .

Hence from our scaling coefficient vector \mathbf{d}_{j_0} and wavelet coefficient vectors $\mathbf{c}_{j_0}, \dots, \mathbf{c}_{J-1}$ we can reconstruct stepwise the scaling coefficient vectors $\mathbf{d}_{j_0+1}, \dots, \mathbf{d}_J$ and therefore we can reconstruct our approach $f_J \in V_J$ (compare the reconstruction scheme (2.62)).

- **data compression**

The concept of a MSA leads us to an effective method for data compression. Instead of saving the m_J coefficients in the vector \mathbf{d}_J for the approach $f_J \in V_J$ we may save the $m_{j_0} + n_{j_0} + \dots + n_{J-1} = m_J$ coefficients in the vectors \mathbf{d}_{j_0} and $\mathbf{c}_{j_0}, \dots, \mathbf{c}_{J-1}$. From this representation we can perform a data compression. We apply the simplest thresholding algorithm:

- (1) Fix a threshold $\epsilon > 0$
- (2) For $i = j_0, \dots, J-1$ if $|c_{i,k}| < \epsilon$ we set $c_{i,k} = 0$ for $k = 0, \dots, n_i - 1$. Let n_ϵ be the amount of coefficients set to 0.

From this algorithm we gain the new coefficient vectors $\overline{\mathbf{c}}_0, \dots, \overline{\mathbf{c}}_{J-1}$ where the nonsignificant structures are neglected. With the reconstruction relation (2.93) we gain a signal \overline{f}_J which is an approximation of the signal f_J . Information about the quality of approximation we may get from the root mean square (rms) value of the deviations and from the correlation between the signals f_J and \overline{f}_J , for the definitions of the rms value and the correlation compare definitions D.2 and D.3.

The compression rate ρ is given with $\rho = \frac{n_\epsilon}{m_i}$ wherein n_ϵ is the number of coefficients that have been neglected in $\overline{\mathbf{c}}_{j_0}, \dots, \overline{\mathbf{c}}_{J-1}$. It holds $0 \leq \rho \leq 1$, i.e. the more coefficients were neglected the higher becomes the compression rate ρ .

Data compression is an important tool in digital signal processing:

- With the data compression algorithm we can reduce storage space of digital data. Of course it is important that the influence on the quality of the data is small, i.e. no significant information should be lost.
- In real time data processing it is an important tool because the huge amount of data often exceeds the capacity of the transmission line.

Besides this simple algorithm for data compression there are other algorithms but we will not discuss them here. For more details see e.g. OGDEN (1997), chapter 8. In section 3.3.3 we will compare this data compression algorithm with a more flexible algorithm and we will see that we can achieve better results for more flexible algorithms.

2.4 n-dimensional B-spline model

For the n-dimensional model we will use tensor product techniques. For a short introduction to tensor products of Hilbert spaces see, e.g., WEIDMANN (1976).

2.4.1 n-dimensional tensor product scaling and wavelet functions

We intend to perform a MSA on $L^2([0, 1]^n)$, hence we have to choose a n-dimensional scaling function $\phi_{\mathbf{k}}^{\mathbf{j}}(\mathbf{x})$, with levels $\mathbf{j} = (j_1, \dots, j_n) \in \mathbb{N}_0^n$, shifts $\mathbf{k} = (k_1, \dots, k_n)$, $k_i = 0, \dots, m_{j_i} - 1$ for $i = 1, \dots, n$ and $\mathbf{x} = (x_1, \dots, x_n) \in [0, 1]^n$. Here we use a tensor product of our 1-dimensional scaling functions and define

$$\phi_{\mathbf{k}}^{\mathbf{j}}(\mathbf{x}) := \phi_{k_1}^{j_1}(x_1) \cdot \phi_{k_2}^{j_2}(x_2) \cdot \dots \cdot \phi_{k_n}^{j_n}(x_n)$$

as the n-dimensional tensor product of 1-dimensional normalized endpoint-interpolating B-splines with $k_i = 0, \dots, m_{j_i} - 1$, $i = 1, \dots, n$. Note, here for more clarity we use a superscript index \mathbf{j} for the levels and a subscript index \mathbf{k} for the shift of the scaling functions. We will adopt this notation also for the coefficient vectors and write $\mathbf{d}^{\mathbf{j}} = \left[d_{0, \dots, 0}^{j_1, \dots, j_n} \dots d_{m_{j_1}-1, \dots, m_{j_n}-1}^{j_1, \dots, j_n} \right]'$ and $\mathbf{c}^{\mathbf{j}} = \left[c_{0, \dots, 0}^{j_1, \dots, j_n} \dots c_{m_{j_1}-1, \dots, m_{j_n}-1}^{j_1, \dots, j_n} \right]'$.

The n-dimensional scaling space $V_{\mathbf{j}}$ is then a tensor product of 1-dimensional scaling spaces V_{j_1}, \dots, V_{j_n} , i.e.

$$V_{\mathbf{j}} = V_{j_1} \otimes V_{j_2} \otimes \dots \otimes V_{j_n}.$$

According to the 1-dimensional relation $V_j = V_{j-1} \oplus W_{j-1}$ we can decompose the scaling space $V_{\mathbf{j}}$ into

$$V_{\mathbf{j}} = (V_{j_1-1} \oplus W_{j_1-1}) \otimes (V_{j_2-1} \oplus W_{j_2-1}) \otimes \dots \otimes (V_{j_n-1} \oplus W_{j_n-1}). \quad (2.99)$$

Evaluating the products leads us to altogether 2^n n-dimensional tensor product spaces:

$$V_{\mathbf{j}-1} = V_{j_1-1} \otimes V_{j_2-1} \otimes \dots \otimes V_{j_n-1} \quad (2.100)$$

$$W_{\mathbf{j}-1}^1 = V_{j_1-1} \otimes \dots \otimes V_{j_{n-1}-1} \otimes W_{j_n-1} \quad (2.101)$$

$$W_{\mathbf{j}-1}^2 = V_{j_1-1} \otimes \dots \otimes V_{j_{n-2}-1} \otimes W_{j_{n-1}-1} \otimes V_{j_n-1} \quad (2.102)$$

⋮

$$W_{\mathbf{j}-1}^n = W_{j_1-1} \otimes V_{j_2-1} \otimes \dots \otimes V_{j_n-1} \quad (2.103)$$

$$W_{\mathbf{j}-1}^{n+1} = V_{j_1-1} \otimes \dots \otimes V_{j_{n-2}-1} \otimes W_{j_{n-1}-1} \otimes W_{j_n-1} \quad (2.104)$$

⋮

$$W_{\mathbf{j}-1}^{2^n-1} = W_{j_1-1} \otimes \dots \otimes W_{j_n-1} \quad (2.105)$$

With the corresponding scaling and wavelet functions of levels \mathbf{j}

$$\begin{aligned}
\phi_{\mathbf{k}}^{\mathbf{j}}(\mathbf{x}) &= \phi_{k_1}^{j_1}(x_1) \cdot \phi_{k_2}^{j_2}(x_2) \cdot \dots \cdot \phi_{k_n}^{j_n}(x_n) \\
\psi_{\mathbf{k}}^{\mathbf{j};1}(\mathbf{x}) &= \phi_{k_1}^{j_1}(x_1) \cdot \dots \cdot \phi_{k_{n-1}}^{j_{n-1}}(x_{n-1}) \cdot \psi_{k_n}^{j_n}(x_n) \\
\psi_{\mathbf{k}}^{\mathbf{j};2}(\mathbf{x}) &= \phi_{k_1}^{j_1}(x_1) \cdot \dots \cdot \phi_{k_{n-2}}^{j_{n-2}}(x_{n-2}) \cdot \psi_{k_{n-1}}^{j_{n-1}}(x_{n-1}) \cdot \phi_{k_n}^{j_n}(x_n) \\
&\vdots \\
\psi_{\mathbf{k}}^{\mathbf{j};n}(\mathbf{x}) &= \psi_{k_1}^{j_1}(x_1) \cdot \phi_{k_2}^{j_2}(x_2) \cdot \dots \cdot \phi_{k_n}^{j_n}(x_n) \\
\psi_{\mathbf{k}}^{\mathbf{j};n+1}(\mathbf{x}) &= \phi_{k_1}^{j_1}(x_1) \cdot \dots \cdot \phi_{k_{n-2}}^{j_{n-2}}(x_{n-2}) \cdot \psi_{k_{n-1}}^{j_{n-1}}(x_{n-1}) \cdot \psi_{k_n}^{j_n}(x_n) \\
&\vdots \\
\psi_{\mathbf{k}}^{\mathbf{j};2^n-1}(\mathbf{x}) &= \psi_{k_1}^{j_1}(x_1) \cdot \dots \cdot \psi_{k_n}^{j_n}(x_n)
\end{aligned}$$

with $k_i = 0, \dots, m_i - 1$ for a 1-dimensional scaling function $\phi_{k_i}^{j_i}$ on the i th position of the tensor product or $k_i = 0, \dots, n_i - 1$ for a 1-dimensional wavelet function $\psi_{k_i}^{j_i}$ on the i th position of the tensor product, $i = 1, \dots, n$.

From equation (2.99) we get

$$V_{\mathbf{j}} = V_{\mathbf{j}-\mathbf{1}} \oplus W_{\mathbf{j}-\mathbf{1}}^1 \oplus \dots \oplus W_{\mathbf{j}-\mathbf{1}}^{2^n-1} \quad (2.106)$$

by evaluating the products and inserting the notations (2.100) to (2.105) for scaling space and the detail spaces.

In the 1-dimensional model the $\phi_{j,k}$, $k = 0, \dots, m_j - 1$ form a basis of the scaling space V_j , we can extend this property to the n-dimensional tensor product model and get:

- the n-dimensional functions $\phi_{\mathbf{k}}^{\mathbf{j}}$, $k_i = 0, \dots, m_{j_i} - 1$ and $i = 0, \dots, n$ form a basis for the n-dimensional scaling space $V_{\mathbf{j}}$.
- the n-dimensional functions $\psi_{\mathbf{k}}^{\mathbf{j}}$, with $k_i = 0, \dots, m_{j_i} - 1$ or $k_i = 0, \dots, n_{j_i} - 1$ for a scaling function $\phi_{k_i}^{j_i}$ or a wavelet function $\psi_{k_i}^{j_i}$ on the i th position of the tensor product, respectively, and $i = 0, \dots, n$, form a basis of the n-dimensional detail spaces $W_{\mathbf{j}}$.

Now we want to extend the properties of a MSA on the bounded interval to the n-dimensional case.

From the equation (2.71) we get for the scaling spaces $V_{\mathbf{j}} = V_{j_1, \dots, j_n}$ the property

$$V_{j_1, \dots, j_n} \subseteq V_{j'_1, \dots, j'_n} \text{ if } j_1 \leq j'_1, j_2 \leq j'_2, \dots, j_n \leq j'_n. \quad (2.107)$$

Equality holds if $j_1 = j'_1, j_2 = j'_2, \dots, j_n = j'_n$. Hence we have an equation similar to equation (2.63).

The second property (2.64) we can easily extend to the n-dimensional case with equation (2.72) it follows

$$\overline{\bigcup_{j_1 \geq 0} \dots \bigcup_{j_n \geq 0} V_{j_1, \dots, j_n}}^{L^2([0,1]^n)} = L^2([0,1]^n). \quad (2.108)$$

2.4.2 Decomposition and reconstruction in 4 dimensions

For simplification we will explain the decomposition and reconstruction for the 4-dimensional case. The n-dimensional case can be constructed analogously for $n \in \mathbb{N}$.

Now we extend the B-spline approach for signals $f \in L^2([0,1])$ to a B-spline approach for signals $f \in L^2([0,1]^4)$. As in the 1-dimensional model we want to compute an approach $f_{\mathbf{J}} \in V_{\mathbf{J}}$ of highest resolution levels $\mathbf{J} = (J_1, J_2, J_3, J_4)$. We define $\mathbf{J} - \mathbf{1} = (J_1 - 1, J_2 - 1, J_3 - 1, J_4 - 1)$ and $\mathbf{J} - \mathbf{i}_0 = (J_1 - i_0, J_2 - i_0, J_3 - i_0, J_4 - i_0)$. According to equation (2.106) we can decompose

$$f_{\mathbf{J}} = f_{\mathbf{J}-\mathbf{1}} + g_{\mathbf{J}-\mathbf{1}}^1 + \dots + g_{\mathbf{J}-\mathbf{1}}^{2^4-1}$$

wherein the $f_{\mathbf{J}-1} \in V_{\mathbf{J}-1}$ and $g_{\mathbf{J}-1}^\nu \in W_{\mathbf{J}-1}^\nu$ for $\nu = 1, \dots, 2^4 - 1$. As in the 1-dimensional case these signals are defined via scaling coefficient vectors $\mathbf{d}^{\mathbf{J}-1}$ and $\mathbf{c}^{\mathbf{J}-1;\nu}$; the exact definitions will be presented later. We will extend the decomposition relations (2.90) and (2.91) and the reconstruction relation (2.93) to the 4-dimensional case and get a similar decomposition and reconstruction algorithm for 4-dimensional signals.

- **computation of the scaling coefficients**

We suppose to have given measurements $f(\mathbf{p}_i) + e(\mathbf{p}_i)$ at the discrete points $\mathbf{p}_1, \dots, \mathbf{p}_n \in [0, 1]^4$, with $\mathbf{p}_i = (x_i, y_i, z_i, t_i)$, $i = 1, \dots, n$. To compute the $m_{J_1} \cdot m_{J_2} \cdot m_{J_3} \cdot m_{J_4} \times 1$ scaling coefficient vector

$\mathbf{d}^{\mathbf{J}} = \begin{bmatrix} d_{0,0,0,0}^{\mathbf{J}} \\ \vdots \\ d_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1, m_{J_4}-1}^{\mathbf{J}} \end{bmatrix}$ of the highest resolution levels $\mathbf{J} = (J_1, J_2, J_3, J_4)$ we have to solve the linear equation system

$$\begin{bmatrix} f(\mathbf{p}_1) \\ f(\mathbf{p}_2) \\ \vdots \\ f(\mathbf{p}_n) \end{bmatrix} + \begin{bmatrix} e(\mathbf{p}_1) \\ e(\mathbf{p}_2) \\ \vdots \\ e(\mathbf{p}_n) \end{bmatrix} = \Phi^{\mathbf{J}} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_n \end{bmatrix} \mathbf{d}^{\mathbf{J}}$$

with the $n \times m_{J_1} \cdot m_{J_2} \cdot m_{J_3} \cdot m_{J_4}$ matrix

$$\Phi^{\mathbf{J}} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_n \end{bmatrix} = \begin{bmatrix} \phi_{0,0,0,0}^{\mathbf{J}}(\mathbf{p}_1) \cdots \cdots \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1, m_{J_4}-1}^{\mathbf{J}}(\mathbf{p}_1) \\ \vdots \\ \phi_{0,0,0,0}^{\mathbf{J}}(\mathbf{p}_n) \cdots \cdots \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1, m_{J_4}-1}^{\mathbf{J}}(\mathbf{p}_n) \end{bmatrix}.$$

Note, each column of the matrix $\Phi^{\mathbf{J}}$ refers to a scaling function $\phi_{k_1, k_2, k_3, k_4}^{\mathbf{J}}$ the order of these columns can be chosen arbitrarily but we have to choose the same order for the coefficients $d_{k_1, k_2, k_3, k_4}^{\mathbf{J}}$ in the vector $\mathbf{d}^{\mathbf{J}}$. From now on we assume that the order is always chosen consistently.

Analogously to the 1-dimensional case each scaling coefficient $d_{k_1, k_2, k_3, k_4}^{\mathbf{J}}$ corresponds to a compactly supported scaling function $\phi_{k_1, k_2, k_3, k_4}^{\mathbf{J}}$, $k_i = 0, \dots, m_{J_i} - 1$ and $i = 0, \dots, 4$. Hence it is only computable if an observation is given in the support of the scaling function. We can establish a Gauss-Markov model and solve it with the same strategies as in the 1-dimensional model introduced in chapter 2.3.5. We will not repeat the strategies here.

With the level \mathbf{J} scaling coefficient vector $\mathbf{d}^{\mathbf{J}}$ we can define the smoothed signal $f_{\mathbf{J}} \in V_{\mathbf{J}}$ via

$$f_{\mathbf{J}}(\mathbf{p}) = \Phi^{\mathbf{J}}(\mathbf{p}) \mathbf{d}^{\mathbf{J}} \quad (2.109)$$

wherein

$$\Phi^{\mathbf{J}}(\mathbf{p}) := \left[\phi_{0,0,0,0}^{\mathbf{J}}(\mathbf{p}) \cdots \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1, m_{J_4}-1}^{\mathbf{J}}(\mathbf{p}) \right]. \quad (2.110)$$

If we assume to have measurements on a fixed grid for $\mathbf{x} = (x_1, \dots, x_{n_x})$, $\mathbf{y} = (y_1, \dots, y_{n_y})$, $\mathbf{z} = (z_1, \dots, z_{n_z})$

and $\mathbf{t} = (t_1, \dots, t_{n_t})$. I.e. we have $n_x \cdot n_y \cdot n_z \cdot n_t$ measurements $\begin{bmatrix} (x_1, y_1, z_1, t_1) \\ \vdots \\ (x_{n_x}, y_{n_y}, z_{n_z}, t_{n_t}) \end{bmatrix}$ on the grid. Then

we can represent the $(n_x \cdot n_y \cdot n_z \cdot n_t) \times (m_{J_1} \cdot m_{J_2} \cdot m_{J_3} \cdot m_{J_4})$ matrix $\Phi^{\mathbf{J}} \left(\begin{bmatrix} (x_1, y_1, z_1, t_1) \\ \vdots \\ (x_{n_x}, y_{n_y}, z_{n_z}, t_{n_t}) \end{bmatrix} \right)$ as a

Kronecker product of the four matrices $\Phi^{J_1}(\mathbf{x})$, $\Phi^{J_2}(\mathbf{y})$, $\Phi^{J_3}(\mathbf{z})$ and $\Phi^{J_4}(\mathbf{t})$, i.e.

$$\Phi^{\mathbf{J}} \left(\begin{bmatrix} (x_1, y_1, z_1, t_1) \\ \vdots \\ (x_{n_x}, y_{n_y}, z_{n_z}, t_{n_t}) \end{bmatrix} \right) = \Phi^{J_4}(\mathbf{t}) \otimes \Phi^{J_3}(\mathbf{z}) \otimes \Phi^{J_2}(\mathbf{y}) \otimes \Phi^{J_1}(\mathbf{x})$$

wherein $\Phi^{J_1}(\mathbf{x}) = \begin{bmatrix} \phi_{J_1,0}(x_1) & \dots & \phi_{J_1,m_{J_1}-1}(x_1) \\ \vdots & & \vdots \\ \phi_{J_1,0}(x_{n_x}) & \dots & \phi_{J_1,m_{J_1}-1}(x_{n_x}) \end{bmatrix}$ and $\Phi^{J_2}(\mathbf{y})$, $\Phi^{J_3}(\mathbf{z})$ and $\Phi^{J_4}(\mathbf{t})$ are defined analogously. Less evaluations of the B-spline functions are necessary if we compute the matrix $\Phi^{\mathbf{J}} \left(\begin{bmatrix} (x_1, y_1, z_1, t_1) \\ \vdots \\ (x_{n_x}, y_{n_y}, z_{n_z}, t_{n_t}) \end{bmatrix} \right)$ as the Kronecker product of the 4 single matrices.

The order of the matrices in the Kronecker product determines the order of the elements of the product matrix. I.e. we have to adapt the order of the matrices in the product to the order of the coefficients in the coefficient vector $\mathbf{d}^{\mathbf{J}}$.

- **decomposition of the signal**

For a decomposition of the signal $f_{\mathbf{J}} \in V_{\mathbf{J}}$ we want to compute the level $\mathbf{J} - 1$ scaling and wavelet coefficient vectors $\mathbf{d}^{\mathbf{J}-1}$ and $\mathbf{c}^{\mathbf{J}-1;\nu}$, $\nu = 1, \dots, 2^4 - 1$ from the level \mathbf{J} scaling coefficient vector $\mathbf{d}^{\mathbf{J}}$, wherein $\mathbf{c}^{\mathbf{J}-1;1} = \left[c_{0,0,0,0}^{J_1-1, J_2-1, J_3-1, J_4-1;1} \dots c_{m_{J_1-1}-1, m_{J_2-1}-1, m_{J_3-1}-1, n_{J_4-1}-1}^{J_1-1, J_2-1, J_3-1, J_4-1;1} \right]'$ corresponding to the wavelet functions $\psi_{k_1, k_2, k_3, k_4}^{J_1-1, J_2-1, J_3-1, J_4-1;1}$, $k_1 = 0, \dots, m_{J_1-1} - 1$, $k_2 = 0, \dots, m_{J_2-1} - 1$, $k_3 = 0, \dots, m_{J_3-1} - 1$ and $k_4 = 0, \dots, n_{J_4-1} - 1$; for $\nu = 2, \dots, 2^4 - 1$ the coefficient vectors are defined analogously. We can extend the 1-dimensional decomposition equations (2.90) and (2.91), i.e.

$$\mathbf{d}_{j-1} = \overline{\mathbf{P}}_j \mathbf{d}_j$$

$$\mathbf{c}_{j-1} = \overline{\mathbf{Q}}_j \mathbf{d}_j$$

to the following equations

$$\mathbf{d}^{\mathbf{J}-1} = (\overline{\mathbf{P}}_{J_4} \otimes \overline{\mathbf{P}}_{J_3} \otimes \overline{\mathbf{P}}_{J_2} \otimes \overline{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.111)$$

$$\mathbf{c}^{\mathbf{J}-1;1} = (\overline{\mathbf{P}}_{J_4} \otimes \overline{\mathbf{P}}_{J_3} \otimes \overline{\mathbf{P}}_{J_2} \otimes \overline{\mathbf{Q}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.112)$$

$$\mathbf{c}^{\mathbf{J}-1;2} = (\overline{\mathbf{P}}_{J_4} \otimes \overline{\mathbf{P}}_{J_3} \otimes \overline{\mathbf{Q}}_{J_2} \otimes \overline{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.113)$$

$$\mathbf{c}^{\mathbf{J}-1;3} = (\overline{\mathbf{P}}_{J_4} \otimes \overline{\mathbf{Q}}_{J_3} \otimes \overline{\mathbf{P}}_{J_2} \otimes \overline{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.114)$$

$$\mathbf{c}^{\mathbf{J}-1;4} = (\overline{\mathbf{Q}}_{J_4} \otimes \overline{\mathbf{P}}_{J_3} \otimes \overline{\mathbf{P}}_{J_2} \otimes \overline{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.115)$$

$$\mathbf{c}^{\mathbf{J}-1;5} = (\overline{\mathbf{P}}_{J_4} \otimes \overline{\mathbf{P}}_{J_3} \otimes \overline{\mathbf{Q}}_{J_2} \otimes \overline{\mathbf{Q}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.116)$$

$$\mathbf{c}^{\mathbf{J}-1;6} = (\overline{\mathbf{P}}_{J_4} \otimes \overline{\mathbf{Q}}_{J_3} \otimes \overline{\mathbf{P}}_{J_2} \otimes \overline{\mathbf{Q}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.117)$$

$$\mathbf{c}^{\mathbf{J}-1;7} = (\overline{\mathbf{Q}}_{J_4} \otimes \overline{\mathbf{P}}_{J_3} \otimes \overline{\mathbf{P}}_{J_2} \otimes \overline{\mathbf{Q}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.118)$$

$$\mathbf{c}^{\mathbf{J}-1;8} = (\overline{\mathbf{P}}_{J_4} \otimes \overline{\mathbf{Q}}_{J_3} \otimes \overline{\mathbf{Q}}_{J_2} \otimes \overline{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.119)$$

$$\mathbf{c}^{\mathbf{J}-1;9} = (\overline{\mathbf{Q}}_{J_4} \otimes \overline{\mathbf{P}}_{J_3} \otimes \overline{\mathbf{Q}}_{J_2} \otimes \overline{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.120)$$

$$\mathbf{c}^{\mathbf{J}-1;10} = (\overline{\mathbf{Q}}_{J_4} \otimes \overline{\mathbf{Q}}_{J_3} \otimes \overline{\mathbf{P}}_{J_2} \otimes \overline{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.121)$$

$$\mathbf{c}^{\mathbf{J}-1;11} = (\overline{\mathbf{P}}_{J_4} \otimes \overline{\mathbf{Q}}_{J_3} \otimes \overline{\mathbf{Q}}_{J_2} \otimes \overline{\mathbf{Q}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.122)$$

$$\mathbf{c}^{\mathbf{J}-1;12} = (\overline{\mathbf{Q}}_{J_4} \otimes \overline{\mathbf{P}}_{J_3} \otimes \overline{\mathbf{Q}}_{J_2} \otimes \overline{\mathbf{Q}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.123)$$

$$\mathbf{c}^{\mathbf{J}-1;13} = (\overline{\mathbf{Q}}_{J_4} \otimes \overline{\mathbf{Q}}_{J_3} \otimes \overline{\mathbf{P}}_{J_2} \otimes \overline{\mathbf{Q}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.124)$$

$$\mathbf{c}^{\mathbf{J}-1;14} = (\overline{\mathbf{Q}}_{J_4} \otimes \overline{\mathbf{Q}}_{J_3} \otimes \overline{\mathbf{Q}}_{J_2} \otimes \overline{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.125)$$

$$\mathbf{c}^{\mathbf{J}-1;15} = (\overline{\mathbf{Q}}_{J_4} \otimes \overline{\mathbf{Q}}_{J_3} \otimes \overline{\mathbf{Q}}_{J_2} \otimes \overline{\mathbf{Q}}_{J_1}) \mathbf{d}^{\mathbf{J}} \quad (2.126)$$

wherein \otimes denotes the Kronecker product of matrices.

These extensions are not very hard; one just has to extend the equations in the derivation of the decomposition and the recursion relations to the n-dimensional case. This requires some work; one has to write

out all of the products and we will not discuss this here; for similar work see e.g. SCHMIDT (2001) section 4.4.4. or OGDEN (1997) chapter 9.

Now we can decompose the signal

$$\begin{aligned}
f_{\mathbf{J}}(\mathbf{p}) &= \Phi^{\mathbf{J}}(\mathbf{p}) \mathbf{d}^{\mathbf{J}} \\
&= \Phi^{\mathbf{J}-1}(\mathbf{p}) \mathbf{d}^{\mathbf{J}-1} + \sum_{\nu=1}^{2^4-1} \Psi^{\mathbf{J}-1;\nu}(\mathbf{p}) \mathbf{c}^{\mathbf{J}-1;\nu} \\
&= f_{\mathbf{J}-1}(\mathbf{p}) + \sum_{\nu=1}^{2^4-1} g_{\mathbf{J}-1}^{\nu}(\mathbf{p})
\end{aligned} \tag{2.127}$$

into the smoothed version

$$f_{\mathbf{J}-1}(\mathbf{p}) := \Phi^{\mathbf{J}-1}(\mathbf{p}) \mathbf{d}^{\mathbf{J}-1} \tag{2.128}$$

and the $2^4 - 1$ detail signals

$$g_{\mathbf{J}-1}^{\nu}(\mathbf{p}) := \Psi^{\mathbf{J}-1;\nu}(\mathbf{p}) \mathbf{c}^{\mathbf{J}-1;\nu}, \quad \nu = 1, \dots, 2^4 - 1, \tag{2.129}$$

wherein $\Psi^{\mathbf{J}-1;\nu}(\mathbf{p})$ is defined analogously to $\Phi^{\mathbf{J}-1}(\mathbf{p})$ (compare equation (2.110)), e.g.

$$\Psi^{\mathbf{J}-1;1}(\mathbf{p}) = \left[\psi_{0,0,0,0}^{\mathbf{J}-1;1}(\mathbf{p}) \dots \psi_{m_{J_1-1-1}, m_{J_2-1-1}, m_{J_3-1-1}, n_{J_4-1}}^{\mathbf{J}-1;1}(\mathbf{p}) \right]. \tag{2.130}$$

Note, the subscript indices depend on the definitions of the wavelets.

With the decomposition relations (2.111) to (2.126) we can compute the scaling coefficient vector $\mathbf{d}^{\mathbf{j}_0}$ of levels $\mathbf{j}_0 = (J_1 - i_0, J_2 - i_0, J_3 - i_0, J_4 - i_0)$ and the detail coefficient vectors $\mathbf{c}^{\mathbf{j}_0;\nu}, \dots, \mathbf{c}^{\mathbf{J}-1;\nu}$, $\nu = 0, \dots, 2^4 - 1$, for a $1 \leq i_0 \leq \min(J_1, J_2, J_3, J_4)$. With these coefficient vectors we can calculate analogously to the 1-dimensional case the corresponding signals $f_{\mathbf{j}_0}$ and $g_{\mathbf{j}_0}^{\nu}, \dots, g_{\mathbf{J}-1}^{\nu}$, $\nu = 1, \dots, 2^4 - 1$ (the definitions are given in equations (2.128) and (2.129)) and get the decomposition

$$f_{\mathbf{J}}(\mathbf{p}) = f_{\mathbf{j}_0}(\mathbf{p}) + \sum_{i=1}^{i_0} \sum_{\nu=1}^{2^4-1} g_{\mathbf{J}-i}^{\nu}(\mathbf{p})$$

- **reconstruction of the signal**

Now we again want to compute the level \mathbf{J} scaling coefficient vector $\mathbf{d}^{\mathbf{J}}$ from the level $\mathbf{J} - 1$ scaling and wavelet coefficient vectors $\mathbf{d}^{\mathbf{J}-1}$ and $\mathbf{c}^{\mathbf{J}-1;\nu}$ for $\nu = 1, \dots, 2^4 - 1$. The reconstruction relation in the 1-dimensional case $\mathbf{d}_{\mathbf{J}} = \mathbf{P}_{\mathbf{J}} \mathbf{d}_{\mathbf{J}-1} + \mathbf{Q}_{\mathbf{J}} \mathbf{c}_{\mathbf{J}-1}$ has to be extended to

$$\begin{aligned}
\mathbf{d}^{\mathbf{J}} &= (\mathbf{P}_{J_4} \otimes \mathbf{P}_{J_3} \otimes \mathbf{P}_{J_2} \otimes \mathbf{P}_{J_1}) \mathbf{d}^{\mathbf{J}-1} + \\
&\quad (\mathbf{P}_{J_4} \otimes \mathbf{P}_{J_3} \otimes \mathbf{P}_{J_2} \otimes \mathbf{Q}_{J_1}) \mathbf{c}^{\mathbf{J}-1;1} + \dots + (\mathbf{Q}_{J_4} \otimes \mathbf{Q}_{J_3} \otimes \mathbf{Q}_{J_2} \otimes \mathbf{Q}_{J_1}) \mathbf{c}^{\mathbf{J}-1;2^4-1}.
\end{aligned} \tag{2.131}$$

Note, in the reconstruction equation the Kronecker product corresponding to the coefficient vectors $\mathbf{c}^{\mathbf{J}-1;\nu}$ is analogous to the Kronecker product in equations (2.112) to (2.126) corresponding to the coefficient vectors $\mathbf{c}^{\mathbf{J}-1;\nu}$ only the matrices $\bar{\mathbf{P}}_{\mathbf{J}}$ and $\bar{\mathbf{Q}}_{\mathbf{J}}$ are interchanged with the matrices $\mathbf{P}_{\mathbf{J}}$ and $\mathbf{Q}_{\mathbf{J}}$, respectively. As for the decomposition equation the extension is not very hard and only the equations needed for the derivation of the relation have to be extended, we will not discuss this here.

- **data compression**

As in the 1-dimensional case we may perform the data compression algorithm. We can adopt the concept and will not repeat it here.

For the multi-dimensional model a new aspect arises: In the 4-dimensional model we may want to smooth the signal of levels $\mathbf{J} = (J_1, J_2, J_3, J_4)$ just in one dimension, e.g. we want to derive a signal of levels $\mathbf{J}' = (J_1 - 1, J_2, J_3, J_4)$. Here we will only discuss this case, we may also smooth in two or three dimensions. The concepts have to be adapted. In the decomposition relation we then have to exchange the matrices $\bar{\mathbf{P}}_{j_i}$ and $\bar{\mathbf{Q}}_{j_i}$, $i = 2, 3, 4$, corresponding to the dimensions we do not want to change in the coefficient equations (i.e. the matrices corresponding to the levels J_2, J_3 and J_4) with unit matrices, i.e.

$$\mathbf{d}^{J_1-1, J_2, J_3, J_4} = (\mathbf{I}_4 \otimes \mathbf{I}_3 \otimes \mathbf{I}_2 \otimes \bar{\mathbf{P}}_{J_1}) \mathbf{d}^{\mathbf{J}} \tag{2.132}$$

wherein \mathbf{I}_4 is a $m_{J_4} \times m_{J_4}$ unit matrix and \mathbf{I}_3 and \mathbf{I}_2 are defined analogously. Analogously we have to work for the decomposition relations for the detail coefficient vectors $\mathbf{c}^{(J_1-1, J_2, J_3, J_4); \nu}$ but some of the equations coincide and we have only $1 = 2^1 - 1$ resulting detail signal. This corresponds to the 1-dimensional decomposition where we get only 1 detail signal.

Analogous we then have to work for the reconstruction but we will not go into more details.

Chapter 3

Application

3.1 Observations

In the following we always work in an Earth-bounded coordinate system. The origin of the coordinate system is the center of the Earth, the xy -plane is the equatorial plane and the z -axis is perpendicular to the equator and therefore coincides approximately with the rotation axis of the Earth. We consider a receiver $R = (\lambda_R, \varphi_R, h_R)$ and a satellite position $S = (\lambda_S, \varphi_S, h_S)$ given in spherical coordinates $\lambda =$ longitude, $\varphi =$ latitude and $h =$ height over the spherical Earth with radius R_e (we assume the Earth is a sphere Ω_{R_e} with radius R_e). The corresponding geocentric position vectors are \mathbf{x}_R and \mathbf{x}_S . To be more specific, the geocentric position vector \mathbf{x} of a point $P = (\lambda, \varphi, h)$ is defined via

$$\mathbf{x} = |\mathbf{x}| \cdot [\cos \varphi \cos \lambda, \cos \varphi \sin \lambda, \sin \varphi]' = |\mathbf{x}| \cdot \mathbf{r}, \quad (3.1)$$

wherein $|\mathbf{x}| = R_e + h$ is the radial distance, \mathbf{r} denotes the corresponding unit vector.

3.1.1 Electron density from GNSS

Dual-frequency GPS (Global Positioning System) receivers register P-code and carrier phase measurements on the two frequencies $f_1 = 1.5$ GHz and $f_2 = 1.2$ GHz. We can formulate observation equations for both, P-code and carrier phase measurements. P-code observations are not ambiguous but have the disadvantage that the random noise is almost 100 times larger than for carrier phase measurements, hence we consider here the observation equation for carrier phase measurements.

We formulate the so-called geometry free linear combination of simultaneous phase observations $\phi_1(R, S, t)$ and $\phi_2(R, S, t)$ on the frequencies f_1 and f_2 as

$$\begin{aligned} \phi_4(R, S, t) &= \phi_1(R, S, t) - \phi_2(R, S, t) = \\ &= \alpha \cdot STEC(R, S, t) + \beta_R + \beta_S + \beta_{R,S} - e(R, S, t), \end{aligned} \quad (3.2)$$

see SCHMIDT ET AL. (2007b). In this equation β_R and β_S mean the inter-frequency differential delays of the receiver R and the satellite S (for more details on these terms see, e.g., BRUNINI ET AL. (2003)), $\beta_{R,S}$ is the combination of the carrier phase ambiguities on both frequencies f_1 and f_2 , α is a constant and $e(R, S, t)$ is the corresponding observation error. The **slant total electron content** (STEC) is defined as the integral of the electron density along the ray-path between receiver R and satellite S :

$$STEC(R, S, t) = \int_R^S N(\lambda, \varphi, h, t) ds. \quad (3.3)$$

$N(\lambda, \varphi, h, t)$ denotes the space- and time-dependent 4-dimensional electron density. Note, the ray-path of the signal actually is a path that starts at the satellite S and ends at the receiver R . But we will neglect the transmission time of the signal and therefore always consider ray-paths from receiver R to satellite S .

From equation (3.2) we gain information about the STEC along the ray-path between receiver R and satellite S . The ambiguity term $\beta_{R,S}$ is computable within a pre-processing step. The observation ϕ_4 contains neither the satellite-receiver geometry nor frequency-independent biases, because these terms cancel out in the subtraction of the simultaneous measurements ϕ_1 and ϕ_2 ; see SCHMIDT ET AL. (2007b).

3.1.2 VTEC from occultation measurements

COSMIC/FORMOSAT-3 is the Constellation Observing System for Meteorology, Ionosphere and Climate and Taiwan's FORMOSA SATellite Mission #3, a joint Taiwan-U.S. project. The six COSMIC satellites were successfully launched on April 14th, 2006. Over the first year, the satellites have been gradually boosted from their initial orbits of 400 km to their final orbits of 800 km. The COSMIC satellites are so-called low-Earth-orbiting (LEO) satellites. Each spacecraft is equipped with three instruments, namely a GPS occultation receiver, a tiny ionospheric photometer and a tri-band beacon; see LIN ET AL. (2007).

A GPS-LEO occultation geometry is shown in Figure 3.1. When a GPS satellite sets or rises behind the Earth's ionosphere as seen by a LEO satellite, the GPS radio signals are received by the LEO satellite. Each occultation therefore consists of a set of limb-viewing links with tangent points P ranging from the LEO satellite orbit altitude to the surface of the Earth. Two types of GPS observations can be used to apply the Abel inversion: the bending angle α and the STEC. It is possible to apply the Abel inversion on the STEC data in order to obtain the electron density expressed as function of height. These retrieval process generally assumes spherical symmetry of the electron density at the occultation position. It is obvious that this assumption is almost never true. An improved Abel transform has been developed by Tsai and Tsai; see TSAI AND TSAI (2004) and GARCÍA-FERNÁNDEZ (2004).

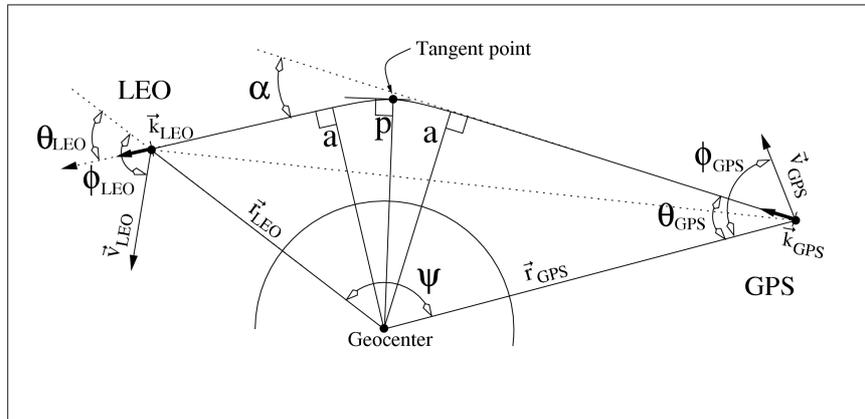


Figure 3.1: Scheme of the occultation geometry. The LEO satellite and the GPS satellite are shown. α denotes the bending angle; see GARCÍA-FERNÁNDEZ (2004).

3.1.3 VTEC from altimetry

We can also derive information about the ionosphere from satellite altimetry missions with double-frequency radar altimeter on-board, such as Topex/Poseidon and Jason-1. Topex/Poseidon was launched in August 1992 and was operational till October 2005, it observed the ocean circulation. The follow-on mission of Topex/Poseidon is Jason-1; it was launched in December 2001 and it has inherited the main features, i.e. orbit, instruments, measurement accuracy, etc. The orbit altitude of the two missions is 1336 km. The Topex/Poseidon and the Jason-1 mission are joint projects between NASA (National Aeronautics and Space Administration) and the French space agency CNES (Centre National d'Etudes Spatiales). The primary sensor of both satellites is the Nasa Radar Altimeter operating at 14.6 GHz (Ku band) and 5.3 GHz (C band) simultaneously. The measurements of the height of the satellite above the sea (satellite range) made at the two frequencies allow, among other information, to obtain the ionospheric vertical total electron content (VTEC).

To assess the precision of the VTEC from GNSS and from satellite altimetry these two techniques have often been compared. Generally, the agreement between GNSS and altimetry derived VTEC is good, but there are still some contradictions. Several studies have shown that the VTEC derived from Topex/Poseidon and Jason-1 are greater than the values from GNSS. This is a contradiction because in opposite to GNSS, the altimetry satellites do not sample the whole ionosphere due to their lower orbital altitude and therefore the VTEC values from satellite altimetry are expected to be lower. On the other hand, most of the ionospheric models from GNSS data are based on the single layer model, which does not account well for the ionospheric contribution above the altitude of the altimetry missions. For the single layer model, the STEC values derived from GNSS measurements have to be converted into VTEC by a mapping function, while the altimetry missions deliver

directly the VTEC values. The mapping function used for this conversion is also a potential error source for the GNSS VTEC estimates. Finally for comparing with altimetry VTEC, the values derived from GNSS have to be interpolated for regions far from the observing stations - that is above the oceans, i.e such comparison are performed in the worst scenario for GNSS; see TODOROVA ET AL. (2007).

3.1.4 Ionosonde measurements

Ionosondes are terrestrial instruments that observe the ionosphere. A ionosonde works with the radar principle. Short pulses in the frequency range from 1 to 40 MHz are vertically transmitted and received after ionospheric reflection. With increasing frequency, the reflection is weaker and therefore the pulses enter deeper in the ionosphere before they are reflected. The amplitude, runtime, doppler shift, polarization and angle of incidence can be derived from the ionospheric echo. A ionogram is the representation of the reflected signals versus time or frequency. From ionosondes we can derive profiles of the electron density for heights from 100 km to 250 km. We can, e.g., use ionosonde observations for a validation of GPS derived electron density profiles. For more details see www.iap-kborn.de, <http://de.wikipedia.org>, LIN ET AL. (2007) and Figure 4.1.

For the different observation types it is always common to subtract a reference model from the original observation, e.g. for a observation $y(R, S)$ we write $y(R, S) = y_{ref}(R, S) + \Delta y(R, S)$ wherein $y_{ref}(R, S)$ means the reference model and $\Delta y(R, S)$ is the correction term.

3.2 Different options

3.2.1 2-dimensional approach

In the 2-dimensional approach we work with measurements $f(\lambda, \varphi)$, wherein λ is the longitude and φ is the latitude. Let us assume to have given 2-dimensional observations

$$f(\lambda_i, \varphi_i) \text{ for } i = 1, \dots, M.$$

Since we want to work regionally our measurements (λ_i, φ_i) , $i = 1, \dots, M$, lie in a specified region $[\lambda_{min}, \lambda_{max}] \times [\varphi_{min}, \varphi_{max}]$. Now we want to apply our B-spline model as introduced in section 2.4 (we will adopt all the notations from this chapter and fix the dimension $n = 2$). To apply the B-spline model we have to evaluate the B-splines at the positions (λ, φ) . Since our B-splines are defined on the interval $[0, 1]$, we have to transform $[\lambda_{min}, \lambda_{max}] \times [\varphi_{min}, \varphi_{max}]$ to the unit square $[0, 1] \times [0, 1]$. Then we can assume our measurements to be given in the unit square $[0, 1] \times [0, 1]$ and can perform the evaluation of the B-splines. The transformation of the coordinates $(\lambda, \varphi) \in [\lambda_{min}, \lambda_{max}] \times [\varphi_{min}, \varphi_{max}]$ into the coordinates $(x, y) \in [0, 1] \times [0, 1]$ is defined via the equations

$$x = \frac{\lambda - \lambda_{min}}{\lambda_{max} - \lambda_{min}} \quad \text{and} \quad (3.4)$$

$$y = \frac{\varphi - \varphi_{min}}{\varphi_{max} - \varphi_{min}}. \quad (3.5)$$

In order to avoid confusion we will not distinguish between these two notations. For the evaluation of the B-splines we always mean the values in the unit square, i.e. we write $\phi_{k_1, k_2}^{J_1, J_2}(\lambda, \varphi)$ but we mean $\phi_{k_1, k_2}^{J_1, J_2}(x, y)$.

We assume $f(\lambda, \varphi) \in L^2([0, 1]^2)$.

Since the scaling spaces V_{J_1, J_2} lie dense in $L^2([0, 1]^2)$ (see equation (2.108), i.e. $\overline{\bigcup_{J_1 \geq 0, J_2 \geq 0} V_{J_1, J_2}} = L^2([0, 1]^2)$) and equation (2.107) (i.e. $V_{j_1, j_2} \subseteq V_{j'_1, j'_2}$ if $j_1 \leq j'_1$ and $j_2 \leq j'_2$) we can approximate the signal f arbitrarily well:

$$\forall \epsilon > 0 \exists J_1, J_2 \text{ such that } \exists f_{J_1, J_2} \in V_{J_1, J_2} \text{ with } f_{J_1, J_2}(\lambda, \varphi) = (\Phi^{J_1, J_2}(\lambda, \varphi))' \cdot \mathbf{d}^{J_1, J_2}$$

wherein $\mathbf{d}^{J_1, J_2} = [d_{0,0}^{J_1, J_2} \dots d_{m_{J_1}-1, m_{J_2}-1}^{J_1, J_2}]'$ and $\Phi^{J_1, J_2}(\lambda, \varphi) = [\phi_{0,0}^{J_1, J_2}(\lambda, \varphi) \dots \phi_{m_{J_1}-1, m_{J_2}-1}^{J_1, J_2}(\lambda, \varphi)]'$ (compare equation (2.109)) such that

$$\|f - f_{J_1, J_2}\|_{L^2([0,1]^2)} < \epsilon. \quad (3.6)$$

The functions $\phi_{k_1, k_2}^{J_1, J_2}$, $k_1 = 0, \dots, m_{J_1} - 1$ and $k_2 = 0, \dots, m_{J_2} - 1$, denote the 2-dimensional tensor product B-spline scaling functions with levels J_1 and J_2 defined on $[0, 1] \times [0, 1]$ (for the exact definition see section 2.4.1) the vector \mathbf{d}^{J_1, J_2} denotes the B-spline scaling coefficient vector.

Now we will repeat here the main steps of the B-spline model.

We fix the resolution levels J_1 and J_2 and want to estimate a level (J_1, J_2) approximation f_{J_1, J_2} of the signal f , i.e. we have to estimate the coefficient vector \mathbf{d}^{J_1, J_2} . Hence with the observations $f(\lambda_i, \varphi_i)$, $i = 1, \dots, M$, we establish the linear model

$$\begin{bmatrix} f(\lambda_1, \varphi_1) \\ \vdots \\ f(\lambda_M, \varphi_M) \end{bmatrix} + \begin{bmatrix} e(\lambda_1, \varphi_1) \\ \vdots \\ e(\lambda_M, \varphi_M) \end{bmatrix} = \begin{bmatrix} \phi_{0,0}^{J_1, J_2}(\lambda_1, \varphi_1) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1}^{J_1, J_2}(\lambda_1, \varphi_1) \\ \vdots & & \vdots \\ \phi_{0,0}^{J_1, J_2}(\lambda_M, \varphi_M) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1}^{J_1, J_2}(\lambda_M, \varphi_M) \end{bmatrix} \cdot \mathbf{d}^{J_1, J_2} \quad (3.7)$$

wherein \mathbf{d}^{J_1, J_2} contains the initially unknown B-spline scaling coefficients and $e(\lambda_i, \varphi_i)$, $i = 1, \dots, M$, are the measurement errors. With the covariance matrix $D(\mathbf{f}) = \sigma_{\mathbf{f}}^2 \mathbf{P}_{\mathbf{f}}^{-1}$ for $\mathbf{f} = [f(\lambda_1, \varphi_1) \dots f(\lambda_M, \varphi_M)]'$ this linear equation system corresponds to the Gauss-Markov model (2.96). $\sigma_{\mathbf{f}}^2$ denotes the unknown variance factor and $\mathbf{P}_{\mathbf{f}}$ is the given positive definite weight matrix of dimension $M \times M$. Hence the solvability of the

Gauss-Markov model depends on the column rank of the $M \times (m_{J_1} \cdot m_{J_2})$ matrix $\Phi^{J_1, J_2} \begin{pmatrix} \lambda_1, \varphi_1, \\ \vdots \\ \lambda_M, \varphi_M \end{pmatrix} =$

$\begin{bmatrix} \phi_{0,0}^{J_1, J_2}(\lambda_1, \varphi_1) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1}^{J_1, J_2}(\lambda_1, \varphi_1) \\ \vdots & & \vdots \\ \phi_{0,0}^{J_1, J_2}(\lambda_M, \varphi_M) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1}^{J_1, J_2}(\lambda_M, \varphi_M) \end{bmatrix}$ with $M \geq (m_{J_1} \cdot m_{J_2})$. For more details on the solvability of the

Gauss-Markov model compare section 2.3.5.

With the estimated coefficient vector \mathbf{d}^{J_1, J_2} we model the signal $f_{J_1, J_2}(\lambda, \varphi)$ of fixed resolution levels J_1 and J_2 (compare equation (2.109))

$$f_{J_1, J_2}(\lambda, \varphi) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} d_{k_1, k_2}^{J_1, J_2} \phi_{k_1, k_2}^{J_1, J_2}(\lambda, \varphi). \quad (3.8)$$

As explained in chapter 2.4 we can now perform a decomposition of the signal f_{J_1, J_2} into a smoothed version f_{J_1-1, J_2-1} and detail signals g_{J_1-1, J_2-1}^{ν} , $\nu = 1, \dots, 2^2 - 1$. For the 2-dimensional model there exist altogether $2^2 - 1 = 3$ detail signals. The signal f_{J_1, J_2} represents f on the scale (J_1, J_2) while the signal f_{J_1-1, J_2-1} represents f on the scale $(J_1 - 1, J_2 - 1)$ the detail signals g_{J_1-1, J_2-1}^{ν} , $\nu = 1, \dots, 3$, contain the structures that distinguishes f_{J_1, J_2} from f_{J_1-1, J_2-1} . The smaller the levels (J_1, J_2) , the coarser is the approximation f_{J_1, J_2} and the less coefficients are necessary to save the approximation f_{J_1, J_2} .

According to equations (2.128) and (2.129) in section 2.4.2 the formulas of the smoothed signal and the detail signals are the following:

$$f_{J_1-1, J_2-1}(\lambda, \varphi) = \sum_{k_1=0}^{m_{J_1-1}-1} \sum_{k_2=0}^{m_{J_2-1}-1} d_{k_1, k_2}^{J_1-1, J_2-1} \phi_{k_1}^{J_1-1}(\lambda) \phi_{k_2}^{J_2-1}(\varphi), \quad (3.9)$$

$$g_{J_1-1, J_2-1}^1(\lambda, \varphi) = \sum_{k_1=0}^{m_{J_1-1}-1} \sum_{k_2=0}^{n_{J_2-1}-1} c_{k_1, k_2}^{J_1-1, J_2-1; 1} \phi_{k_1}^{J_1-1}(\lambda) \psi_{k_2}^{J_2-1}(\varphi) \quad (3.10)$$

$$g_{J_1-1, J_2-1}^2(\lambda, \varphi) = \sum_{k_1=0}^{n_{J_1-1}-1} \sum_{k_2=0}^{m_{J_2-1}-1} c_{k_1, k_2}^{J_1-1, J_2-1; 2} \psi_{k_1}^{J_1-1}(\lambda) \phi_{k_2}^{J_2-1}(\varphi) \quad (3.11)$$

$$g_{J_1-1, J_2-1}^3(\lambda, \varphi) = \sum_{k_1=0}^{n_{J_1-1}-1} \sum_{k_2=0}^{n_{J_2-1}-1} c_{k_1, k_2}^{J_1-1, J_2-1; 3} \psi_{k_1}^{J_1-1}(\lambda) \psi_{k_2}^{J_2-1}(\varphi). \quad (3.12)$$

With the concept of a MSA we have now decomposed the signal f into a smoothed version of the signal f_{J_1-1, J_2-1} and the corresponding detail signals g_{J_1-1, J_2-1}^{ν} , $\nu = 1, \dots, 3$ these signals are defined via the coefficient vectors

$\mathbf{d}^{J_1-1, J_2-1}$ and $\mathbf{c}^{J_1-1, J_2-1; \nu}$, respectively; compare section 2.4.2. Now we can apply the data compression algorithm introduced in section 2.3.5. We neglect those coefficients in the vectors $\mathbf{c}^{J_1-1, J_2-1; \nu}$, $\nu = 1, 2, 3$, whose absolute value is lower than a specified threshold $\epsilon > 0$ and gain the coefficient vectors $\bar{\mathbf{c}}^{J_1-1, J_2-1; \nu}$, $\nu = 1, 2, 3$. That means we neglect the non-significant structures of the detail signals g_{J_1-1, J_2-1}^ν , $\nu = 1, \dots, 3$ and get the new detail signals $\bar{g}_{J_1-1, J_2-1}^\nu$, $\nu = 1, \dots, 3$. By applying the reconstruction algorithm we also get a new signal \bar{f}_{J_1, J_2} .

With the 2-dimensional approach we can, e.g., model the $VTEC(R, S, t)$. Working with the 2-dimensional model in longitude and latitude there is no consideration of the time, hence we have to perform a discretization in time and we have to estimate time-dependent B-spline scaling coefficients. In the following we will describe the 3-dimensional and the 4-dimensional approaches, these are analogous to the 2-dimensional approach and not all of the details are repeated.

3.2.2 3-dimensional approach

In the 3-dimensional approach we may consider either measurements in longitude, latitude and time or measurements in longitude, latitude and height. In the first case we may approximate, e.g., the $VTEC(R, S, t)$ as in the 2-dimensional approach but instead of a discretization in time, we have a third scaling function in the temporal direction and therefore we can estimate time-independent B-spline scaling coefficients. In the second case we approximate signals in the 3-dimensional space (longitude, latitude and height), e.g., the electron density $N(\mathbf{x}, t)$. Since there is again no consideration in time in this 3-dimensional model we have to choose a discretization in time and have to estimate time-dependent coefficients.

Let us assume we have given 3-dimensional observations in longitude, latitude and height

$$f(\lambda_i, \varphi_i, h_i) \text{ for } i = 1, \dots, M.$$

The observations $f(\lambda_i, \varphi_i, h_i)$, $i = 1, \dots, M$, are given in a specified region $[\lambda_{min}, \lambda_{max}] \times [\varphi_{min}, \varphi_{max}] \times [h_{min}, h_{max}]$. To apply the 3-dimensional B-spline model as introduced in section 2.4.2 we have to transform $[\lambda_{min}, \lambda_{max}] \times [\varphi_{min}, \varphi_{max}] \times [h_{min}, h_{max}]$ to $[0, 1] \times [0, 1] \times [0, 1]$. The transformation equations can be found analogously to the 2-dimensional approach (compare equations (3.4) and (3.5)). As before we will not distinguish between these two representations and for the evaluation of the B-splines we always mean the values in $[0, 1] \times [0, 1] \times [0, 1]$.

As in the 2-dimensional approach we suppose $f(\lambda, \varphi, h) \in L^2([0, 1]^3)$. Since the scaling spaces V_{J_1, J_2, J_3} lie dense in $L^2([0, 1]^3)$ and $V_{j_1, j_2, j_3} \subseteq V_{j'_1, j'_2, j'_3}$ if $j_1 \leq j'_1$, $j_2 \leq j'_2$, and $j_3 \leq j'_3$. (compare equations (2.107) and (2.108)) we can approximate the signal arbitrarily well by a signal $f_{J_1, J_2, J_3} \in V_{J_1, J_2, J_3}$ with $f_{J_1, J_2, J_3}(\lambda, \varphi, h) = [\phi_{0,0,0}^{J_1, J_2, J_3}(\lambda, \varphi, h) \dots \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(\lambda, \varphi, h)] \cdot \mathbf{d}^{J_1, J_2, J_3}$ and the level (J_1, J_2, J_3) scaling coefficient vector

$$\mathbf{d}^{J_1, J_2, J_3} = \begin{bmatrix} d_{0,0,0}^{J_1, J_2, J_3} \\ \vdots \\ d_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3} \end{bmatrix} \quad (3.13)$$

(we will not go into details here but compare with the 2-dimensional approach).

Now we fix the resolution levels J_1 , J_2 and J_3 and want to estimate a level (J_1, J_2, J_3) approximation of the signal f , i.e. we have to estimate the coefficient vector $\mathbf{d}^{J_1, J_2, J_3}$. Hence with our observations $f(\lambda_i, \varphi_i, h_i)$, $i = 1, \dots, M$, we establish the linear model

$$\begin{bmatrix} f(\lambda_1, \varphi_1, h_1) \\ \vdots \\ f(\lambda_M, \varphi_M, h_M) \end{bmatrix} + \begin{bmatrix} e(\lambda_1, \varphi_1, h_1) \\ \vdots \\ e(\lambda_M, \varphi_M, h_M) \end{bmatrix} = \Phi^{J_1, J_2, J_3} \cdot \mathbf{d}^{J_1, J_2, J_3}. \quad (3.14)$$

wherein

$$\Phi^{J_1, J_2, J_3} = \begin{bmatrix} \phi_{0,0,0}^{J_1, J_2, J_3}(\lambda_1, \varphi_1, h_1) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(\lambda_1, \varphi_1, h_1) \\ \vdots & & \vdots \\ \phi_{0,0,0}^{J_1, J_2, J_3}(\lambda_M, \varphi_M, h_M) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(\lambda_M, \varphi_M, h_M) \end{bmatrix}. \quad (3.15)$$

With the covariance matrix $D(\mathbf{f}) = \sigma_{\mathbf{f}}^2 \mathbf{P}_{\mathbf{f}}^{-1}$ for $\mathbf{f} = [f(\lambda_1, \varphi_1, h_1) \dots f(\lambda_M, \varphi_M, h_M)]'$ this model forms a Gauss-Markov model analogous to equation (2.96). The quantities $\sigma_{\mathbf{f}}^2$ and $\mathbf{P}_{\mathbf{f}}$ denote the unknown variance factor and the given positive definite weight matrix of dimension $M \times M$. For more details on the solvability of the Gauss-Markov model again see section 2.3.5.

With the estimated B-spline scaling coefficient vector $\mathbf{d}^{J_1, J_2, J_3}$ we model the signal $f_{J_1, J_2, J_3}(\lambda, \varphi, h)$ of fixed resolution levels J_1, J_2 and J_3 as

$$f_{J_1, J_2, J_3}(\lambda, \varphi, h) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} d_{k_1, k_2, k_3}^{J_1, J_2, J_3} \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, h). \quad (3.16)$$

As described in section 2.4.2 we can decompose the approximation $f_{J_1, J_2, J_3}(\lambda, \varphi, h)$ into a smoother version $f_{J_1-1, J_2-1, J_3-1}(\lambda, \varphi, h)$ of levels $(J_1 - 1, J_2 - 1, J_3 - 1)$ and $7 = 2^3 - 1$ corresponding detail signals $g_{J_1-1, J_2-1, J_3-1}^{\nu}(\lambda, \varphi, h)$, $\nu = 1, \dots, 2^3 - 1$, i.e.

$$f_{J_1, J_2, J_3}(\lambda, \varphi, h) = f_{J_1-1, J_2-1, J_3-1}(\lambda, \varphi, h) + \sum_{\nu=1}^{2^3-1} g_{J_1-1, J_2-1, J_3-1}^{\nu}(\lambda, \varphi, h). \quad (3.17)$$

The formulas for the smoothed signal and the detail signals are constructed analogously to equations (2.128) and (2.129), i.e. they form the 3-dimensional analogon to the equations (3.9) to (3.12). All of the signals are defined via their corresponding coefficient vectors, i.e. $\mathbf{d}^{J_1-1, J_2-1, J_3-1}$ and $\mathbf{c}^{J_1-1, J_2-1, J_3-1; \nu}$, $\nu = 1, \dots, 7$.

Now we again may apply the data compression algorithm introduced in section 2.3.5. We will not repeat this concept, for more details compare the 2-dimensional approach in section 3.2.1.

Note, in the approach explained before we used the height as a third coordinate but it can be interchanged with the time. I.e. we can either use the 3-dimensional approach as an approach in longitude, latitude and height and in this case have to choose a discretization in the time or we can use the 3-dimensional approach as an approach in longitude, latitude and time. We have to choose the appropriate model for the given observations.

3.2.3 4-dimensional approach

The 4-dimensional approach is used for 4-dimensional signals. Observations are given in a 4-dimensional space. In our applications we will consider spatio-temporal observations in longitude, latitude, height and time, hence, there is no need of a discretization in time. We suppose the observations at positions $(\lambda_i, \varphi_i, h_i, t_i)$, $i = 1, \dots, M$, lie in the specified region $[\lambda_{min}, \lambda_{max}] \times [\varphi_{min}, \varphi_{max}] \times [h_{min}, h_{max}] \times [t_{min}, t_{max}]$. Now we want to apply the 4-dimensional B-spline model and therefore we have to evaluate the B-splines. Hence we have to transform the region $[\lambda_{min}, \lambda_{max}] \times [\varphi_{min}, \varphi_{max}] \times [h_{min}, h_{max}] \times [t_{min}, t_{max}]$ to $[0, 1] \times [0, 1] \times [0, 1] \times [0, 1]$. The transformation equations can be formulated analogously to the 2- and 3-dimensional model (compare equations (3.4) and (3.5)). Note, for simplification we again do not distinguish between these two representations.

Let us assume we have given observations

$$f(\lambda_i, \varphi_i, h_i, t_i) \text{ for } i = 1, \dots, M$$

with $(\lambda_i, \varphi_i, h_i, t_i) \in [\lambda_{min}, \lambda_{max}] \times [\varphi_{min}, \varphi_{max}] \times [h_{min}, h_{max}] \times [t_{min}, t_{max}]$. As in the 2- and 3-dimensional approaches we assume $f \in L^2([0, 1]^4)$. Since the scaling spaces V_{J_1, J_2, J_3, J_4} lie dense in $L^2([0, 1]^4)$ and $V_{j_1, j_2, j_3, j_4} \subseteq V_{j'_1, j'_2, j'_3, j'_4}$ if $j_1 \leq j'_1, j_2 \leq j'_2, j_3 \leq j'_3$ and $j_4 \leq j'_4$ (compare equations (2.107) and (2.108)) we may approximate the signal f arbitrarily well by a signal $f_{J_1, J_2, J_3, J_4} \in V_{J_1, J_2, J_3, J_4}$.

To estimate the coefficient vector $\mathbf{d}^{J_1, J_2, J_3, J_4}$ for fixed resolution levels J_1, J_2, J_3 and J_4 we establish the linear model for the observations $f(\lambda_i, \varphi_i, h_i, t_i)$, $i = 1, \dots, M$,

$$\begin{bmatrix} f(\lambda_1, \varphi_1, h_1, t_1) \\ \vdots \\ f(\lambda_M, \varphi_M, h_M, t_M) \end{bmatrix} + \begin{bmatrix} e(\lambda_1, \varphi_1, h_1, t_1) \\ \vdots \\ e(\lambda_M, \varphi_M, h_M, t_M) \end{bmatrix} = \begin{bmatrix} \phi_{0,0,0,0}^{J_1, J_2, J_3, J_4}(\lambda_1, \varphi_1, h_1, t_1) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1, m_{J_4}-1}^{J_1, J_2, J_3, J_4}(\lambda_1, \varphi_1, h_1, t_1) \\ \vdots & & \vdots \\ \phi_{0,0,0,0}^{J_1, J_2, J_3, J_4}(\lambda_M, \varphi_M, h_M, t_M) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1, m_{J_4}-1}^{J_1, J_2, J_3, J_4}(\lambda_M, \varphi_M, h_M, t_M) \end{bmatrix} \cdot \mathbf{d}^{J_1, J_2, J_3, J_4}. \quad (3.18)$$

With the covariance matrix $D(\mathbf{f}) = \sigma_{\mathbf{f}}^2 \mathbf{P}_{\mathbf{f}}^{-1}$ for $\mathbf{f} = [f(\lambda_1, \varphi_1, h_1, t_1) \dots f(\lambda_M, \varphi_M, h_M, t_M)]'$ the model forms a Gauss-Markov model analogous to (2.96) and for more details on the solvability of the system we again refer to section 2.3.5. $\sigma_{\mathbf{f}}^2$ denotes the unknown variance factor and $\mathbf{P}_{\mathbf{f}}$ is the given positive definite weight matrix of dimension $M \times M$.

With the estimated coefficient vector $\mathbf{d}^{J_1, J_2, J_3, J_4}$ we may calculate the approximation $f_{J_1, J_2, J_3, J_4}(\lambda, \varphi, h, t)$ of the measurements of fixed resolution levels J_1, J_2, J_3 and J_4 in space and time. The approximation f_{J_1, J_2, J_3, J_4} is defined via

$$f_{J_1, J_2, J_3, J_4}(\lambda, \varphi, h, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} \sum_{k_4=0}^{m_{J_4}-1} d_{k_1, k_2, k_3, k_4}^{J_1, J_2, J_3, J_4} \phi_{k_1, k_2, k_3, k_4}^{J_1, J_2, J_3, J_4}(\lambda, \varphi, h, t). \quad (3.19)$$

Starting with the approach $f_{J_1, J_2, J_3, J_4}(\lambda, \varphi, h, t)$ we may decompose the signal with the concept of a MSA into a smoother signal $f_{J_1-1, J_2-1, J_3-1, J_4-1}$ and the corresponding detail signals $g_{J_1-1, J_2-1, J_3-1, J_4-1}^{\nu}$, $\nu = 1, \dots, 15$ (compare the 2-dimensional and the 3-dimensional approach).

Then we can again apply the data compression algorithm introduced in section 2.3.5. By neglecting the detail coefficients with absolute value lower than a specific threshold $\epsilon > 0$ we can save the data in a compressed way by neglecting the non-significant structures.

Figure 3.2 gives an overview of the input strategies for the 2-dimensional, the 3-dimensional and the 4-dimensional B-spline approaches. The multi-dimensional approach allows different input strategies: e.g. $VTEC(\lambda, \varphi)$, $VTEC(\lambda, \varphi, t)$, $N(\lambda, \varphi, h)$ and $N(\lambda, \varphi, h, t)$. From these input data we will subtract the corresponding values of a reference model and continue to work with the residual input data. Here we will only work regionally and apply our B-spline model. One can extend our model into a combined model with empirical orthogonal functions (EOFs) or Chapman functions and work either globally or regionally; see e.g. SCHMIDT ET AL. (2007b) and FELTENS (1998).

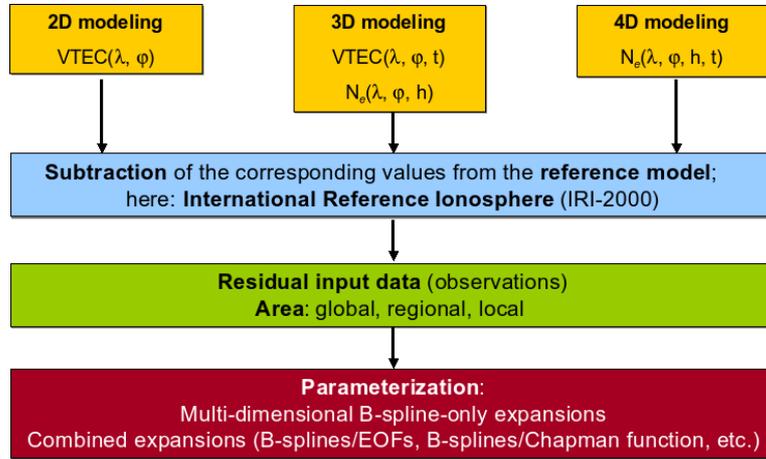


Figure 3.2: Overview of the input strategies of the multi-dimensional B-spline model ($N_e(\lambda, \varphi, h) \equiv N(\lambda, \varphi, h)$ and $N_e(\lambda, \varphi, h, t) \equiv N(\lambda, \varphi, h, t)$).

3.3 Selected applications

3.3.1 Electron density B-spline model for IRI

In order to demonstrate an application of the 3-dimensional B-spline approach we model the electron density over South America.

We consider a region over Central and South America between 250° and 340° in longitude, -60° and 30° in latitude (compare Figure 3.3) and between 100 km and 1200 km in height. We have given values of the electron

density $N(\lambda, \varphi, h, t)$ on a grid with a 2° spacing in longitude and latitude and a 25 km spacing in height on a specified day for each hour $t \in \{0, \dots, 23\}$. For this grid we have for a fixed $t \in \{0, \dots, 23\}$ altogether $M = 95220$ observations of the electron density $N(\lambda, \varphi, h, t)$ computed from the International Reference Ionosphere (IRI)

$$\text{which are saved in the observation vector } \mathbf{N} = \begin{bmatrix} N(\lambda_1, \varphi_1, h_1, t) \\ \vdots \\ N(\lambda_M, \varphi_M, h_M, t) \end{bmatrix}.$$

Now we apply our 3-dimensional B-spline model as explained in section 3.2.2. We want to model an approximation analogous to equation (3.16), i.e. we have to estimate a time-dependent coefficient vector $\mathbf{d}^{J_1, J_2, J_3} =$

$$\begin{bmatrix} d_{0,0,0}^{J_1, J_2, J_3}(t) \\ \vdots \\ d_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}(t) \end{bmatrix} \text{ compare equation (3.13) to compute the approximation } N_{J_1, J_2, J_3}(\lambda, \varphi, h, t) \text{ (note, the}$$

coefficient vector is dependent of the time but for our further calculations this does not make any difference). To estimate the coefficient vector $\mathbf{d}^{J_1, J_2, J_3}$ we establish the Gauss-Markov model analogous to equation (3.14)

$$\mathbf{N} + \mathbf{e} = \mathbf{\Phi}^{J_1, J_2, J_3} \cdot \mathbf{d}^{J_1, J_2, J_3} \text{ with } D(\mathbf{N}) = \sigma_{\mathbf{N}}^2 \mathbf{P}_{\mathbf{N}}^{-1} \quad (3.20)$$

wherein $\sigma_{\mathbf{N}}^2$ is the unknown variance factor and $\mathbf{P}_{\mathbf{N}}$ is the positive definite given $M \times M$ weight matrix of the

observations. $\mathbf{e} = \begin{bmatrix} e(\lambda_1, \varphi_1, h_1, t) \\ \vdots \\ e(\lambda_M, \varphi_M, h_M, t) \end{bmatrix}$ denotes the error vector and $\mathbf{\Phi}^{J_1, J_2, J_3}$ is the matrix wherein the B-spline

values are saved as defined in equation (3.15).

Since for $J_1 = J_2 = J_3 = 5$ it holds $M = 95220 > m_5 \cdot m_5 \cdot m_5 = 34^3 = 39304$ and the observations are evenly distributed the matrix $\mathbf{\Phi}^{5,5,5}$ has full column rank (the scaling functions $\phi_{5,k}$, $k = 0, \dots, m_5 - 1$ are linear independent; compare Theorem 2.28) and we can estimate the unique least squares solution according to equation (2.98). We do not have to introduce prior information. I.e. we gain the coefficient vector $\mathbf{d}^{J_1, J_2, J_3}$ and therefore can compute the approximation

$$\hat{N}_{J_1, J_2, J_3}(\lambda, \varphi, h, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} \hat{d}_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t) \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, h) \quad (3.21)$$

defined analogous to equation (3.16).

Figure 3.3 shows the results of the B-spline approach. The left column shows the electron density input data $N(\lambda_i, \varphi_i, h_i, t)$ computed from IRI, $i = 1, \dots, M$, for the specified time $t = 14 : 00$ UT and selected heights $h_i \in \{225, 250, 275, 300\}$ km, we computed the values on the grid in longitude and latitude, i.e. $\lambda_i \in \{250, 252, \dots, 340\}$ and $\varphi_i \in \{-60, -58, \dots, 30\}$. The mid column visualizes the corresponding level (5, 5, 5) approximations $\hat{N}_{5,5,5}(\lambda_i, \varphi_i, h_i, t)$, $i = 1, \dots, M$ and the right column shows the deviations $N(\lambda_i, \varphi_i, h_i, t) - \hat{N}_{5,5,5}(\lambda_i, \varphi_i, h_i, t)$. The root mean square (rms) values of the deviations are between 0.003×10^6 [electrons/cm³] and 0.007×10^6 [electrons/cm³].

Figure 3.4 a) shows a height profile of the electron density in Bogota in Colombia. The crosses are computed from the IRI, the solid curve is the level (5, 5, 5) approximation from the B-spline approach. The rms value of the deviations amounts 0.0013×10^6 [electrons/cm³]. The panels in b) show selected temporal profiles for 24 hours at heights $h = 300$ km and $h = 325$ km for Bogota (again: crosses = IRI model, line = level-(5,5,5) B-spline approximation). The rms values of the deviations amount 0.0042×10^6 [electrons/cm³] and 0.0019×10^6 [electrons/cm³], respectively.

We will use the coefficient vector $\mathbf{d}^{J_1, J_2, J_3}$ of levels $(J_1, J_2, J_3) = (4, 4, 4)$, $(3, 3, 3)$ and $(2, 2, 2)$ at time $t = 14 : 00$ UT for a simulation of the input data in the next section.

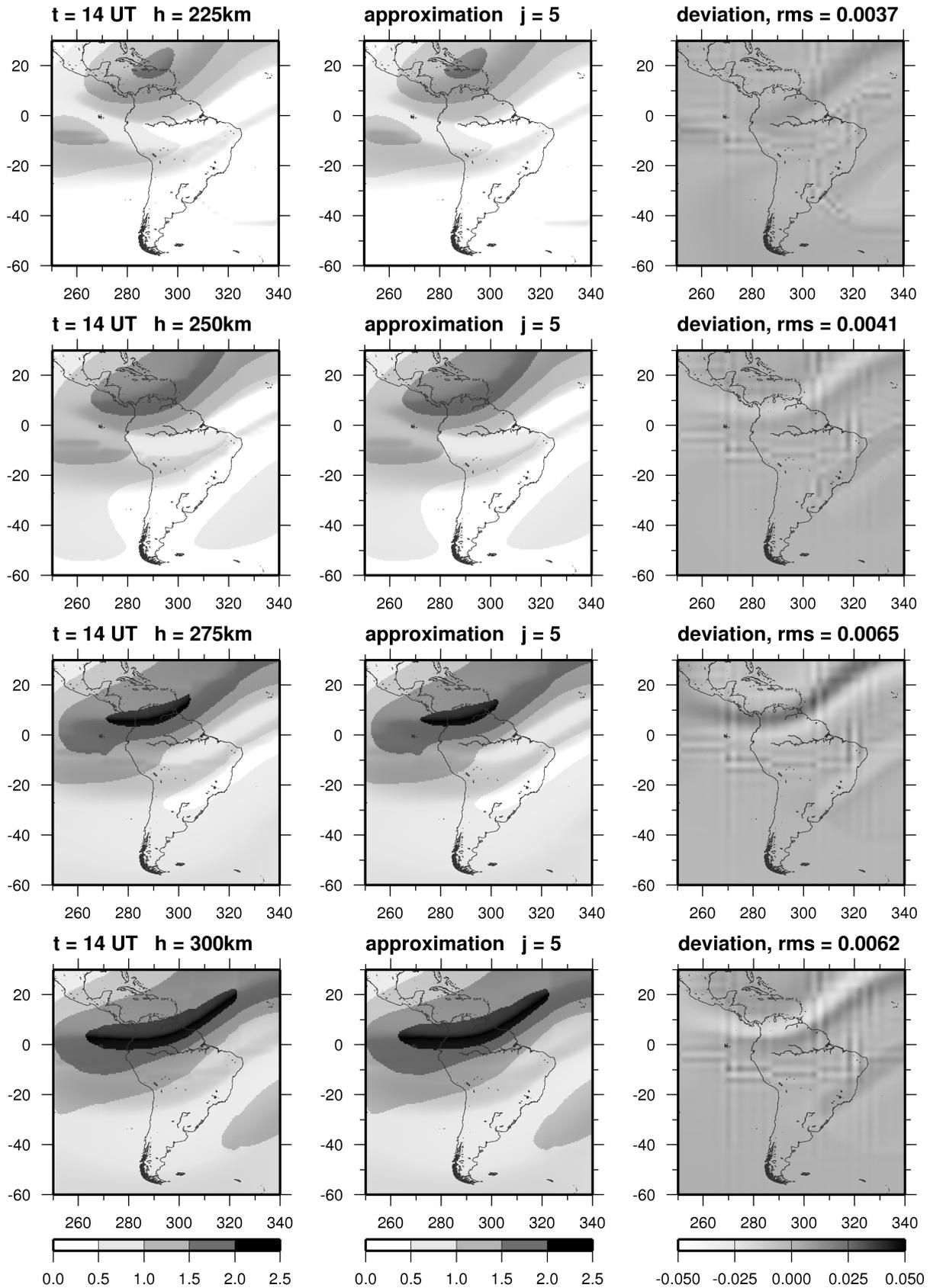


Figure 3.3: Electron density values from IRI (left panels), the corresponding B-spline approximations (mid column) and the differences (right panels) for selected heights $h = 225, 250, 275, 300$ km at time $t = 14$ UT; data sets in $[10^6 \text{ electrons/cm}^3]$; see SCHMIDT ET AL. (2007b).

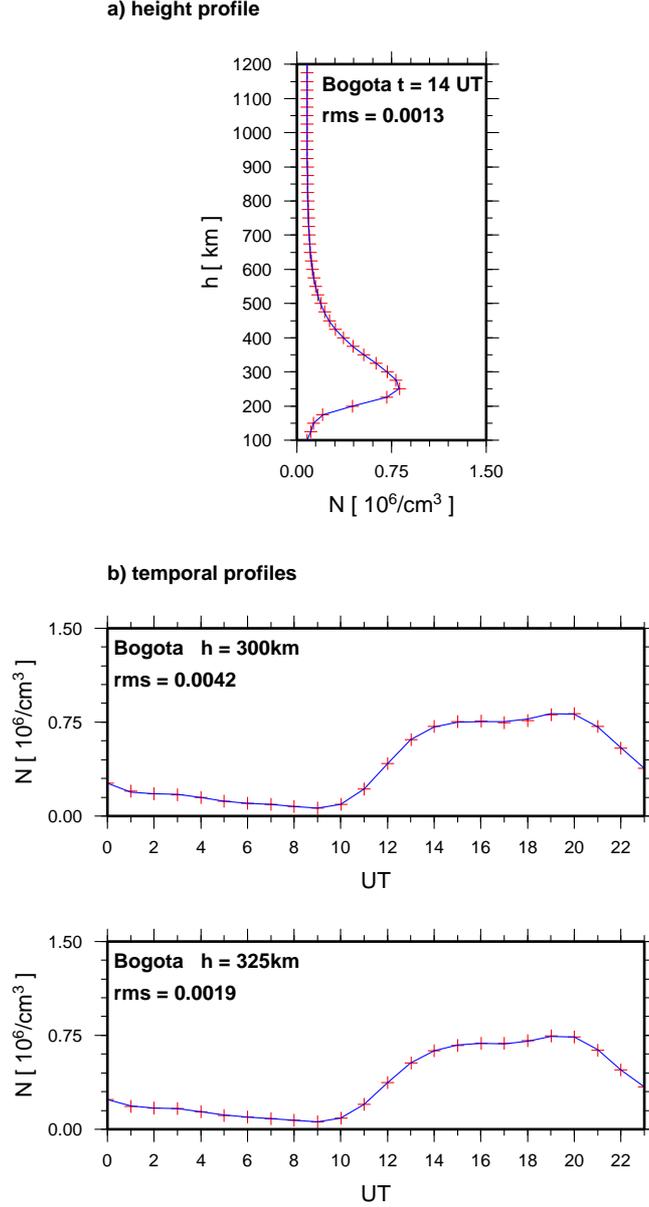


Figure 3.4: a) height profile of the electron density in Bogota in Colombia; b) temporal profiles of the electron density in Bogota for two heights $h = 300$ km and $h = 325$ km.

3.3.2 STEC model for GNSS observations

As introduced in equation (3.2) the GPS observation equation reads

$$\begin{aligned} \phi_4(R, S, t) &= \phi_1(R, S, t) - \phi_2(R, S, t) = \\ &= \alpha \cdot STEC(R, S, t) + \beta_R + \beta_S + \beta_{R,S} - e(R, S, t). \end{aligned} \quad (3.22)$$

The *STEC* is defined as the integral of the space and time-dependent electron density $N(\lambda, \varphi, h, t)$ along the ray-path between receiver R and satellite S, i.e.

$$STEC(R, S, t) = \int_R^S N(\lambda, \varphi, h, t) ds. \quad (3.23)$$

Our aim is to calculate a correction term for the 4-dimensional electron density, hence we decompose $N(\lambda, \varphi, h, t)$ into a reference model $N_{ref}(\lambda, \varphi, h, t)$ (e.g. computed from the IRI) and a correction term $\Delta N(\lambda, \varphi, h, t)$, i.e.

$$N(\lambda, \varphi, h, t) = N_{ref}(\lambda, \varphi, h, t) + \Delta N(\lambda, \varphi, h, t). \quad (3.24)$$

Inserting equation (3.24) in equation (3.23) yields

$$STEC(R, S, t) = STEC_{ref}(R, S, t) + \int_R^S \Delta N(\lambda, \varphi, h, t) ds \quad (3.25)$$

wherein $STEC_{ref}(R, S, t) = \int_R^S N_{ref}(\lambda, \varphi, h, t) ds$ is the approximate STEC value computed by means of the given reference model $N_{ref}(\lambda, \varphi, h, t)$.

To model the correction term $\Delta N(\lambda, \varphi, h, t)$ we work regionally by an expansion in B-spline functions as introduced in section 3.2.2. We could also work globally, e.g. by an expansion in spherical harmonics in combination with empirical orthogonal functions (EOFs), but we will not discuss this here. For more details see, e.g., SCHMIDT ET AL. (2007a) and SCHMIDT ET AL. (2007b) and the references therein.

In the following we work regionally and model the correction term with the 3-dimensional B-spline approach $\Delta N_{J_1, J_2, J_3}(\lambda, \varphi, h, t)$ as introduced in equation (3.16), i.e.

$$\Delta N_{J_1, J_2, J_3}(\lambda, \varphi, h, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} d_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t) \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, h) \quad (3.26)$$

with unknown coefficients $d_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t)$. Now we introduce the model (3.26) into equation (3.25) and obtain

$$STEC_{J_1, J_2, J_3}(R, S, t) = STEC_{ref}(R, S, t) + \frac{1}{\alpha} \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} d_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t) k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) \quad (3.27)$$

wherein

$$k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) = \alpha \int_R^S \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, h) ds. \quad (3.28)$$

Substituting the approach (3.27) into our observation equation (3.22) leads us to

$$\begin{aligned} \phi_4(R, S, t) - \beta_{R,S} - \alpha \cdot STEC_{ref}(R, S, t) + e(R, S, t) &= \\ &= \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} d_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t) k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) + \beta_R + \beta_S. \end{aligned} \quad (3.29)$$

We set $y(R, S, t) := \phi_4(R, S, t) - \beta_{R,S} - \alpha \cdot STEC_{ref}(R, S, t)$ and rewrite equation (3.29) as

$$y(R, S, t) + e(R, S, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} d_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t) k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) + \beta_R + \beta_S. \quad (3.30)$$

For the time-dependency of the coefficients $d_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t)$ we may perform a simple approach by stepwise functions. For a given time interval $[t_A, t_B]$ we define $t_i = t_A + i \cdot \Delta T$, $i = 0, \dots, I$, with I is the number of equally spaced intervals of length $\Delta T = \frac{t_B - t_A}{I}$ and let

$$\chi_i(t) = \begin{cases} 1, & \text{if } t_i \leq t < t_{i+1} \\ 0, & \text{else.} \end{cases} \quad (3.31)$$

We approximate $d_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t)$ by the step function $\sum_{i=0}^{I-1} d_{k_1, k_2, k_3; i}^{J_1, J_2, J_3} \chi_i(t)$. Now we rewrite the observation equation (3.30) as

$$y(R, S, t) + e(R, S, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} \sum_{i=0}^{I-1} d_{k_1, k_2, k_3; i}^{J_1, J_2, J_3} \chi_i(t) k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) + \beta_R + \beta_S. \quad (3.32)$$

The unknown terms are written on the right-hand side of the equation, i.e. the scaling coefficients $d_{k_1, k_2, k_3; i}^{J_1, J_2, J_3}$ and the inter-frequency delays β_R and β_S . For our simulation we suppose now the inter-frequency delays β_R and β_S

are computable (they can be downloaded, e.g., from www.aiub-download.unibe.ch), i.e. we suppose these terms to be known and bring them to the left hand-side of the equation. From now on we suppose $y(R, S, t)$ to include the inter-frequency delays and we will not consider them any more. As an alternative we can estimate these parameters together with the B-spline scaling coefficients, but we will not perform this computation. Therefore we can rewrite equation (3.32) as

$$y(R, S, t) + e(R, S, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} \sum_{i=0}^{I-1} d_{k_1, k_2, k_3; i}^{J_1, J_2, J_3} \chi_i(t) k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S). \quad (3.33)$$

Equation (3.33) is the observation equation for estimating the unknown coefficients of the electron density B-spline model.

Let R_1, \dots, R_n be a selection of altogether n receiver positions. Figure 3.5 shows International GNSS (Global Navigation Satellite Systems) Service (IGS) and regional GPS stations over Central and South America. And let S_1, \dots, S_m be a selection of altogether m satellite positions at a specified time $t \in [t_i, t_{i+1})$.

Our purpose is to estimate the $m_{J_1} \cdot m_{J_2} \cdot m_{J_3}$ coefficients $d_{k_1, k_2, k_3; i}^{J_1, J_2, J_3}$, $J_1, J_2, J_3 \in \mathbb{N}_0$, $k_1 = 0, \dots, m_{J_1}-1$, $k_2 = 0, \dots, m_{J_2}-1$ and $k_3 = 0, \dots, m_{J_3}-1$ for a fixed $i \in \{0, \dots, I-1\}$ from the observation equation (3.33).

We assume the observation vector $\mathbf{y} = \begin{bmatrix} y(R_1, S_1, t_{1,1}) \\ \vdots \\ y(R_n, S_m, t_{n,m}) \end{bmatrix}$ to be given for $t_{k,l} \in [t_i, t_{i+1})$, $k = 1, \dots, n$ and

$l = 1, \dots, m$. The value $y(R_k, S_l, t_{k,l})$ means the observations for the ray-path from the satellite position S_l to the receiver position R_k at time $t_{k,l}$.

Note, the receivers R_i , $i = 1, \dots, n$ may not be able to receive appropriate signals from all the satellite positions S_j , $j = 1, \dots, m$ (e.g. the zenith angle¹ of the ray-path has to be less than 80°). Therefore the observation vector \mathbf{y} may not include all combinations of receivers and satellites and contains less or equal than $n \cdot m$ observations.

With the observation vector \mathbf{y} we can then establish a linear equation system corresponding to equation (3.33):

$$\begin{bmatrix} y(R_1, S_1, t_{1,1}) \\ \vdots \\ y(R_n, S_m, t_{n,m}) \end{bmatrix} + \begin{bmatrix} e(R_1, S_1, t_{1,1}) \\ \vdots \\ e(R_n, S_m, t_{n,m}) \end{bmatrix} = \mathbf{K}^{J_1, J_2, J_3} \cdot \mathbf{d}_i^{J_1, J_2, J_3} \quad (3.34)$$

wherein $\mathbf{e} = \begin{bmatrix} e(R_1, S_1, t_{1,1}) \\ \vdots \\ e(R_n, S_m, t_{n,m}) \end{bmatrix}$ is the observation error vector,

$$\mathbf{K}^{J_1, J_2, J_3} := \begin{bmatrix} k_{0,0,0}^{J_1, J_2, J_3}(R_1, S_1) & \dots & k_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(R_1, S_1) \\ \vdots & & \vdots \\ k_{0,0,0}^{J_1, J_2, J_3}(R_n, S_m) & \dots & k_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(R_n, S_m) \end{bmatrix} \quad (3.35)$$

is the matrix wherein the integral values are saved.

$\mathbf{d}_i^{J_1, J_2, J_3} = \left[d_{0,0,0; i}^{J_1, J_2, J_3} \dots d_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1; i}^{J_1, J_2, J_3} \right]'$ is the $(m_{J_1} \cdot m_{J_2} \cdot m_{J_3}) \times 1$ B-spline scaling coefficient vector. Hence our aim is to estimate the unknown B-spline scaling coefficient vector $\mathbf{d}_i^{J_1, J_2, J_3}$.

We rewrite the linear equation system (3.34) in a short notation as

$$\mathbf{y} + \mathbf{e} = \mathbf{K} \mathbf{d} \quad (3.36)$$

¹The zenith angle is the angle between the zenith and the ray-path, e.g. for a vertical ray-path the zenith angle amounts 0° .

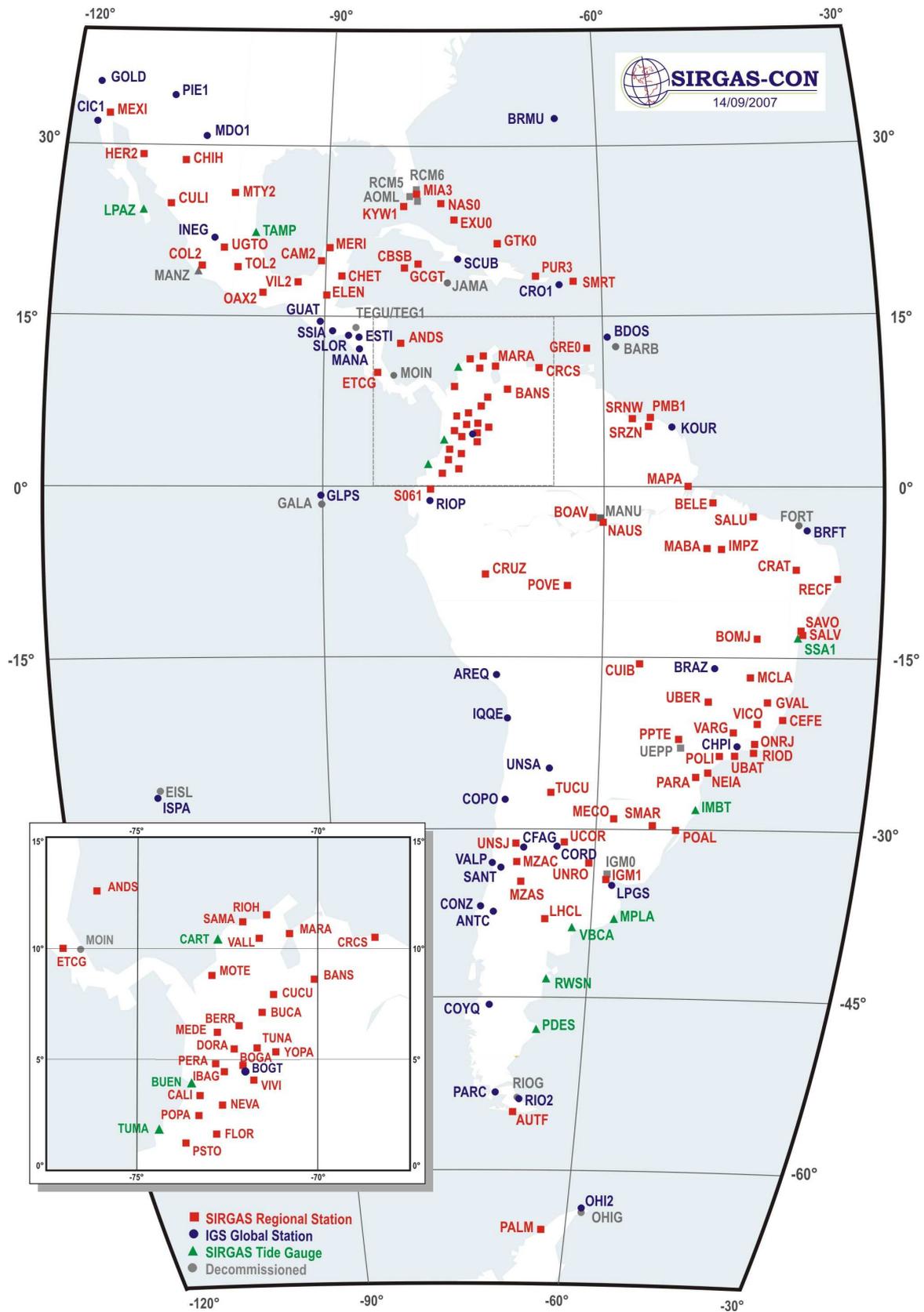


Figure 3.5: IGS and regional GPS stations; see www.sirgas.org.

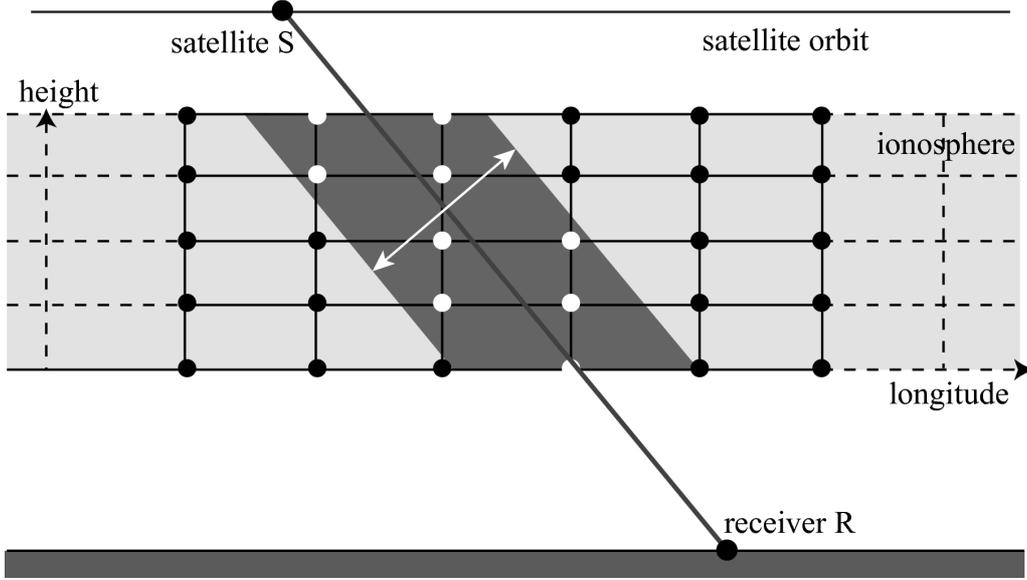


Figure 3.6: 2-dimensional problem in longitude and height. The dots indicate the centers (λ_{k_1}, h_{k_3}) of the 2-dimensional B-spline functions $\phi_{k_1, k_3}^{j_1, j_3}$. The white dots belong to those B-splines that have non-zero entries in the observation equation.

wherein $\mathbf{K} = [\mathbf{K}^{J_1, J_2, J_3}]$ and $\mathbf{d} = [\mathbf{d}_i^{J_1, J_2, J_3}]$. This leads us to the Gauss-Markov model analogous to equation (2.96)

$$\mathbf{y} + \mathbf{e} = \mathbf{K} \mathbf{d} \text{ with covariance matrix } D(\mathbf{y}) = \sigma_y^2 \mathbf{P}_y^{-1}. \quad (3.37)$$

The quantities σ_y^2 and \mathbf{P}_y are the unknown variance factor and the given positive definite weight matrix. According to equation (2.97) the resulting normal equation system reads

$$(\mathbf{K}' \mathbf{P}_y \mathbf{K}) \hat{\mathbf{d}} = \mathbf{K}' \mathbf{P}_y \mathbf{y}. \quad (3.38)$$

Now the solvability of the Gauss-Markov model depends on the column rank of the matrix \mathbf{K} (compare section 2.3.5).

Each scaling coefficient $d_{k_1, k_2, k_3; i}^{J_1, J_2, J_3}$ is related to the corresponding scaling function $\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}$ which has compact support and is located around a center point $P = (\lambda_{k_1}, \varphi_{k_2}, h_{k_3})$. Hence we can determine a coefficient only if observations are given that have ray-paths penetrating the support of the corresponding scaling function (only those coefficients have non-zero entries in the observation equation (3.34)). To be more specific, Figure 3.6 shows a 2-dimensional problem in longitude and height. The black and white dots indicate the center points $P = (\lambda_{k_1}, h_{k_3})$ of the 2-dimensional tensor product scaling functions. We see the ray-path between satellite S and receiver R . As mentioned before only those coefficients where the ray-path penetrates the support of the corresponding scaling function have non-zero entries in the observation equation (in the figure these are the B-splines related to the white dots, in contrast the B-splines related to the black dots have zero entries in the observation equation).

Since the input data, i.e. GPS observations, are generally scattered there may not all of the coefficients be computable and we can exclude the corresponding addends in the observation equation (3.33). If there are only few observations to compute a scaling coefficient we may introduce prior information in order to stabilize the estimation process, i.e. to perform a kind of regularization.

For our simulation we will introduce prior information. As already mentioned in section 2.3.5 we explain here how to introduce an additional linear model for the prior information.

With the prior information for the expectation vector $E(\mathbf{d}) = \boldsymbol{\mu}_d = [\mu_{0,0,0} \dots \mu_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}]'$ and the covariance matrix $D(\mathbf{d}) = \sigma_d^2 \mathbf{P}_d^{-1}$ for the unknown B-spline scaling coefficients we can formulate the additional linear model

$$\boldsymbol{\mu}_d + \mathbf{e}_d = \mathbf{d} \text{ with } D(\boldsymbol{\mu}_d) = \sigma_d^2 \mathbf{P}_d^{-1} \quad (3.39)$$

wherein \mathbf{e}_d is the error vector of the prior information, σ_d^2 is the unknown variance factor and \mathbf{P}_d is the given positive definite weight matrix.

The combination of the two linear models (3.37) and (3.39) yields the extended Gauss-Markov model

$$\begin{bmatrix} \mathbf{y} \\ \boldsymbol{\mu}_d \end{bmatrix} + \begin{bmatrix} \mathbf{e} \\ \mathbf{e}_d \end{bmatrix} = \begin{bmatrix} \mathbf{K} \\ \mathbf{I} \end{bmatrix} \mathbf{d} \text{ with } D \left(\begin{bmatrix} \mathbf{y} \\ \boldsymbol{\mu}_d \end{bmatrix} \right) = \begin{bmatrix} \sigma_y^2 \mathbf{P}_y^{-1} & \mathbf{0} \\ \mathbf{0} & \sigma_d^2 \mathbf{P}_d^{-1} \end{bmatrix} \quad (3.40)$$

wherein \mathbf{I} is a $(m_{J_1} \cdot m_{J_2} \cdot m_{J_3}) \times (m_{J_1} \cdot m_{J_2} \cdot m_{J_3})$ unit matrix and $\mathbf{0}$ are zero matrices with the appropriate dimensions. Solving the combined Gauss-Markov model (3.40) with the least squares method yields us to the extended normal equation

$$\left(\frac{1}{\sigma_y^2} \mathbf{K}' \mathbf{P}_y \mathbf{K} + \frac{1}{\sigma_d^2} \mathbf{P}_d \right) \hat{\mathbf{d}} = \frac{1}{\sigma_y^2} \mathbf{K}' \mathbf{P}_y \mathbf{y} + \frac{1}{\sigma_d^2} \mathbf{P}_d \boldsymbol{\mu}_d \quad (3.41)$$

or by multiplying equation (3.41) with σ_y^2 and introducing the so-called regularization parameter $\lambda = \frac{\sigma_y^2}{\sigma_d^2}$ we get the extended normal equation

$$(\mathbf{K}' \mathbf{P}_y \mathbf{K} + \lambda \mathbf{P}_d) \hat{\mathbf{d}} = \mathbf{K}' \mathbf{P}_y \mathbf{y} + \lambda \mathbf{P}_d \boldsymbol{\mu}_d \quad (3.42)$$

for the unknown parameter vector \mathbf{d} with the unknown variance components σ_y^2 and σ_d^2 . We will not explain the derivation of the extended normal equation for the combined model but for more details see KOCH (2000).

Figure 3.7 gives an overview of the most important steps we have performed so far.

As mentioned before we can exclude the coefficients for which there are no observations in the support of the corresponding scaling functions available. An alternative is to introduce prior information, i.e. to set the weight of the prior information $\mu_{k_1, k_2, k_3} = 0$ for those coefficients $d_{k_1, k_2, k_3; i}^{J_1, J_2, J_3}$ to infinity. However for numerical reasons we choose a high weight ω_0 for those coefficients $d_{k_1, k_2, k_3; i}^{J_1, J_2, J_3}$ for which no observations are given in the support of the corresponding scaling function $\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}$, while for the other weights we may choose a value reciprocal to the number M_{k_1, k_2, k_3} of ray-paths penetrating the support of the corresponding scaling function $\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}$. E.g. we set

$$\mathbf{P}_d = \begin{bmatrix} \omega_{0,0,0} & & & \\ & \ddots & & \\ & & \omega_{m_{J_1-1}, m_{J_2-1}, m_{J_3-1}} & \\ & & & \ddots \end{bmatrix}, \quad (3.43)$$

wherein $\omega_{k_1, k_2, k_3} = \frac{c}{M_{k_1, k_2, k_3}}$ if $M_{k_1, k_2, k_3} \neq 0$, and $\omega_{k_1, k_2, k_3} = \omega_0$ if $M_{k_1, k_2, k_3} = 0$, for $k_i = 0, \dots, m_{J_i} - 1$ and $i = 1, 2, 3$; c means an appropriate constant.

For solving the extended normal equation (3.42) we have to construct the matrix $\mathbf{K} = \mathbf{K}^{J_1, J_2, J_3}$ as defined in (3.35). Therefore we have to calculate the line integrals $\frac{1}{\alpha} k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) = \int_R^S \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, h) ds$ (see equation (3.28)). As mentioned in the introduction, we assume that the ionosphere is a thick shell of electrons between 100 km and 1200 km height above the spherical Earth with radius R_e and hence we do not integrate over the whole line from R to S , but only from P_1 to P_2 , the points where the ray-path leaves the ionosphere at 100 km $+R_e$ and enters the ionosphere at 1200 km $+R_e$ (the points P_1 and P_2 are shown in Figure 3.8). Below the ionosphere there are in fact no electrons, hence we do not have to neglect anything. In the plasmasphere above the ionosphere there are some variations of electron density but those we will neglect for our approach. The location of the plasmasphere is shown in Figure 3.9. As mentioned before we will not consider the transmission time and therefore we will perform the integration from P_1 to P_2 .

Hence from now on we have

$$\frac{1}{\alpha} k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) = \int_{P_1}^{P_2} \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, h) ds. \quad (3.44)$$

Note, for an evaluation of the B-splines we always have to work in longitude, latitude and height, hence $P_1 = (\lambda_1, \varphi_1, h_1)$ and $P_2 = (\lambda_2, \varphi_2, h_2)$. But for a parametrization of the line it is useful to consider the positions in cartesian coordinates, then we straightforward find the parametrization. Therefore we set $\mathbf{x}_{P_1} = (x_1, y_1, z_1)$ and $\mathbf{x}_{P_2} = (x_2, y_2, z_2)$ to be the geocentric position vectors of the points P_1 and P_2 , respectively, as introduced in section 3.1.

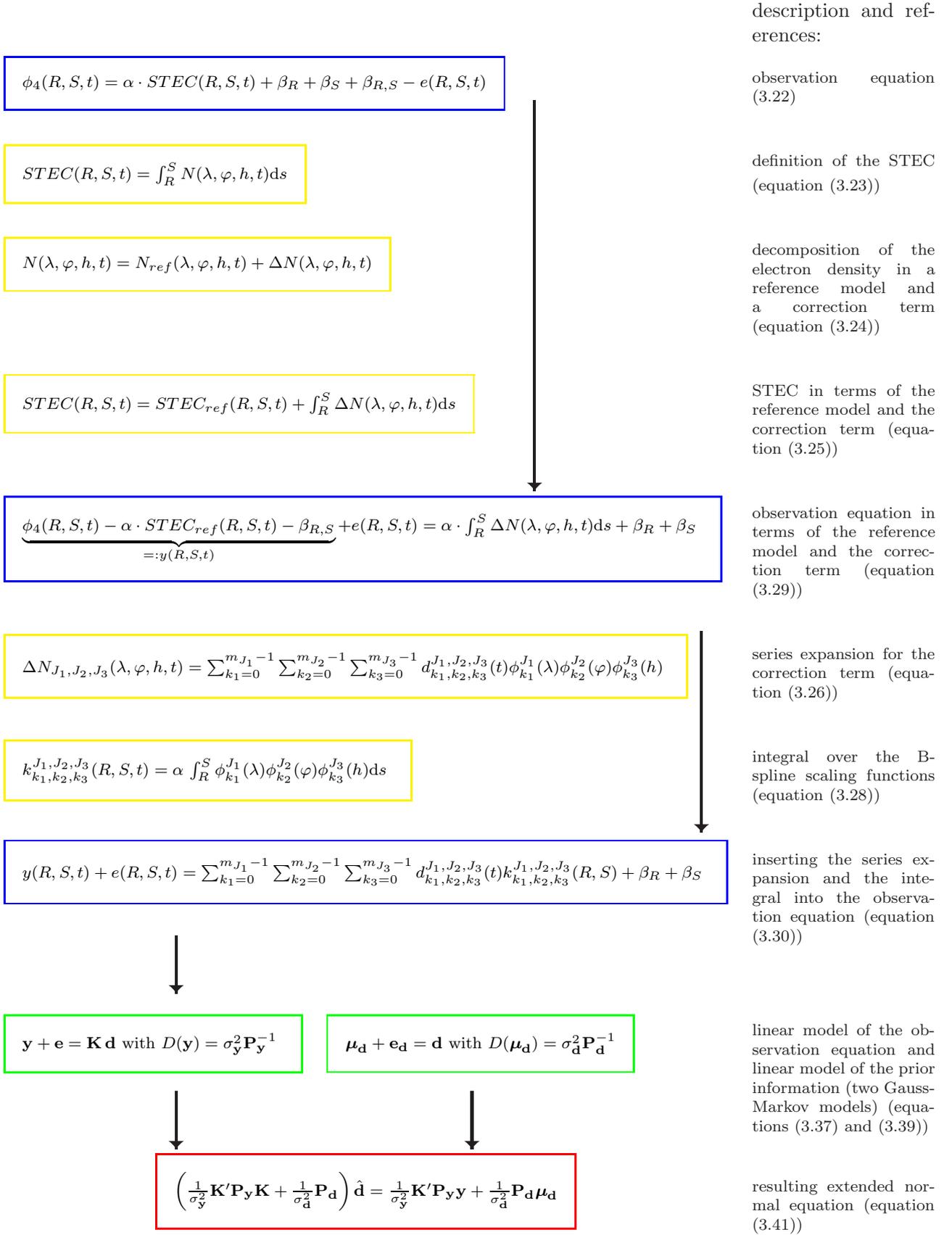


Figure 3.7: Flowchart of the main steps.

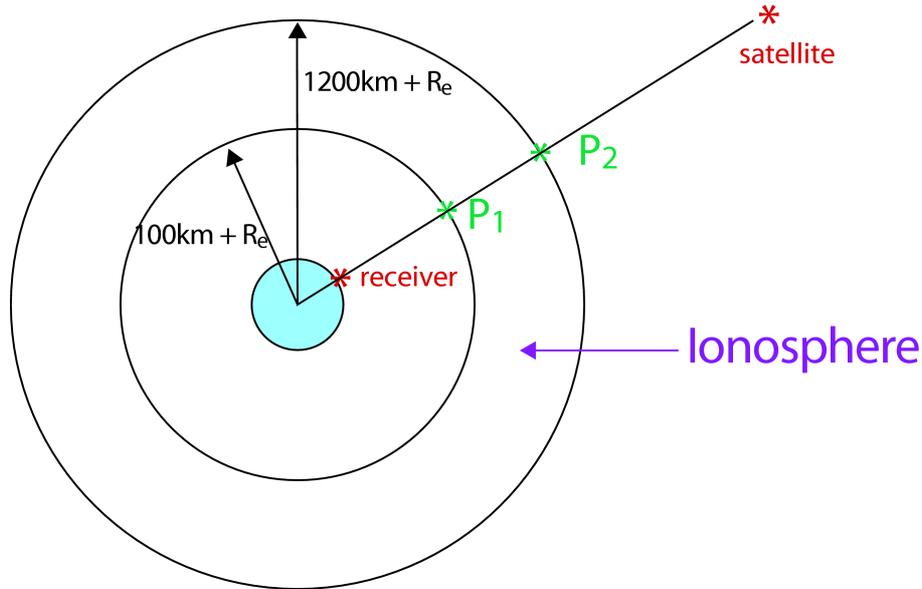


Figure 3.8: Location of P_1 and P_2 . The points P_2 and P_1 are defined as the points where the ray-path enters and leaves the ionosphere, respectively.

The transformation of the cartesian coordinates $\mathbf{x}_P = (x, y, z)$ into spherical coordinates $P = (\lambda, \varphi, h)$ is given with

$$u(\mathbf{x}_P) = u(x, y, z) := \left(\arctan\left(\frac{y}{x}\right), \arctan\left(\frac{z}{\sqrt{|x|^2 + |y|^2}}\right), \sqrt{|x|^2 + |y|^2 + |z|^2} \right); \quad (3.45)$$

see, e.g., DETTMERING (2003).

A B-spline $\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}$ in our model is defined for spherical coordinates (λ, φ, h) but with (3.45) we can also evaluate it for geocentric coordinates \mathbf{x} , i.e. suppose $P = (\lambda, \varphi, h)$ is a point on the ray-path and $\mathbf{x}_P = (x, y, z)$ is the corresponding geocentric position vector, then it holds $u(\mathbf{x}_P) = P$ and therefore

$$\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(P) = (\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3} \circ u)(\mathbf{x}_P). \quad (3.46)$$

With $\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3} \circ u$ we have now defined the B-spline functions on the geocentric coordinates and therefore we

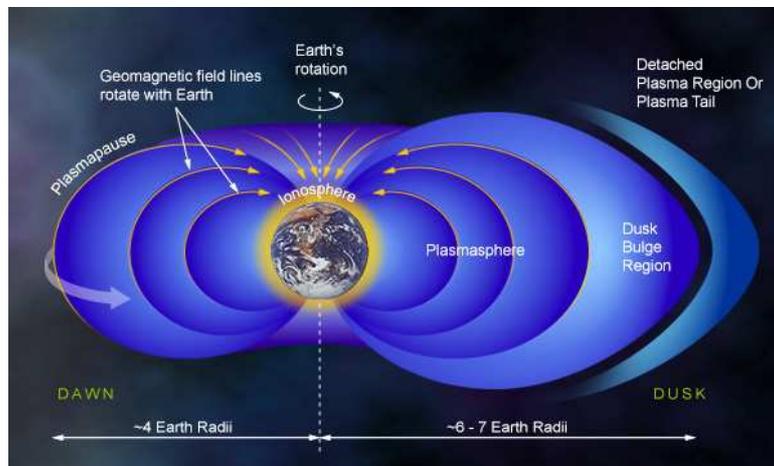


Figure 3.9: Location of the ionosphere and the plasmasphere (www.windows.ucar.edu). The effect of the plasmasphere on electromagnetic signals is estimated, e.g. by GARCÍA-FERNÁNDEZ (2004).

can evaluate the line integral in the geocentric coordinate system

$$\frac{1}{\alpha} k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) = \int_{\mathbf{x}_{P_1}}^{\mathbf{x}_{P_2}} (\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3} \circ u)(\mathbf{x}) \, ds \quad (3.47)$$

wherein the line coordinates are given in geocentric coordinates.

To solve this line integral we now have to find a parametrization of the line from \mathbf{x}_{P_1} to \mathbf{x}_{P_2} . This parametrization is given with

$$\gamma(\omega) = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \omega \cdot \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix}, \quad 0 \leq \omega \leq 1 \quad (3.48)$$

and

$$\dot{\gamma}(\omega) = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix} \quad \text{wherein } \dot{\gamma} = \frac{d\gamma}{d\omega}; \quad (3.49)$$

ω is the line coordinate. Then we can perform the transformation of the line integral:

$$\begin{aligned} \int_{\mathbf{x}_{P_1}}^{\mathbf{x}_{P_2}} (\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3} \circ u)(\mathbf{x}) \, ds &= \int_0^1 \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(u(\gamma(\omega))) |\dot{\gamma}(\omega)| \, d\omega \\ &= \int_0^1 \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(u(\gamma(\omega))) \cdot \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \, d\omega. \end{aligned} \quad (3.50)$$

To solve the integration we divide it into M subintegrals:

$$\begin{aligned} &\int_0^1 \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(u(\gamma(\omega))) \cdot \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \, d\omega = \\ &\int_0^{1/M} \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(u(\gamma(\omega))) \cdot \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \, d\omega \\ &+ \int_{1/M}^{2/M} \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(u(\gamma(\omega))) \cdot \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \, d\omega \\ &+ \dots \\ &+ \int_{(M-1)/M}^1 \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(u(\gamma(\omega))) \cdot \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \, d\omega. \end{aligned} \quad (3.51)$$

For our computations we approximate the integral by applying the Gauss-Legendre quadrature on each of the subintegrals. Therefore we have to transform the intervals of integration $[i/M, (i+1)/M]$ to $[-1, 1]$, $i = 0, \dots, M-1$.

Simulated numerical example

For our simulated numerical example we fix an area of consideration over Central and South America from 250° to 340° in longitude and from -60° to 30° in latitude, compare Figure 3.11. As mentioned before we consider heights from 100 km to 1200 km above the Earth and for a fixed day we choose the time interval $[t_i, t_{i+1}] = [14, 15]$ UT.

As mentioned before for our simulation we suppose the inter-frequency delays β_R and β_S are computable. We will not consider these terms for our simulation. Therefore we have to solve the linear equation system (3.36)

$$\mathbf{y} + \mathbf{e} = \mathbf{K} \mathbf{d} \quad (3.52)$$

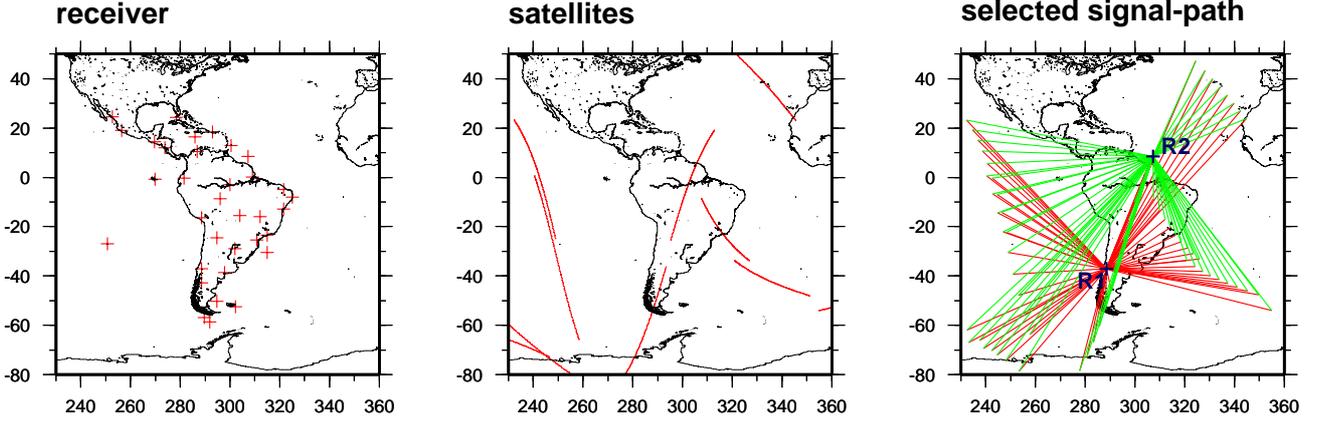


Figure 3.10: The receiver positions (left panel), the satellite orbits (mid panel) and the signal-paths for two selected receivers R_1 and R_2 (right panel).

wherein \mathbf{y} , \mathbf{e} , \mathbf{K} and \mathbf{d} were defined before. Now we want to simulate an observation vector \mathbf{y} . The observations in the vector \mathbf{y} have been defined as $y(R, S, t) := \phi_4(R, S, t) - \beta_{R,S} - \alpha \cdot STEC_{ref}(R, S, t)$. For our simulation we suppose $\beta_{R,S}$ to be computable in a pre-processing step and it will not be considered any more, we set the constant $\alpha = 1$ and gain

$$y(R, S, t) := \phi_4(R, S, t) - STEC_{ref}(R, S, t). \quad (3.53)$$

For the second term of the observation $y(R, S, t)$, i.e. the reference model $STEC_{ref}(R, S, t)$ we introduce here a level $(J_1, J_2, J_3) = (3, 3, 3)$ B-spline approximation $STEC_{3,3,3}(R, S, t)$ of $STEC(R, S, t)$ computed from IRI. To compute this approximation we use the level $(3, 3, 3)$ coefficient vector $\hat{\mathbf{d}}^{3,3,3}(t)$ computed in section 3.3.1 (compare equation (3.21)). We define the reference model $STEC_{ref} = STEC_{3,3,3}(R, S, t) = \int_R^S \hat{N}_{3,3,3}(\cdot, t) ds$ as the integral over the electron density approximation, i.e.

$$STEC_{ref}(R, S, t) = STEC_{3,3,3}(R, S, t) = \mathbf{K}^{3,3,3}(R, S) \cdot \hat{\mathbf{d}}^{3,3,3}(t) \quad (3.54)$$

wherein $\mathbf{K}^{3,3,3}(R, S) = \left[k_{0,0,0}^{3,3,3}(R, S) \dots k_{m_3-1, m_3-1, m_3-1}^{3,3,3}(R, S) \right]$ and $k_{k_1, k_2, k_3}^{3,3,3}(R, S)$, $k_i = 0, \dots, m_3 - 1$ for $i = 1, 2, 3$, is defined analogous to equation (3.28). For the first term of the observation $y(R, S, t)$, i.e. $\phi_4(R, S, t)$ we introduce $STEC_{4,4,4}(R, S, t)$ (computed analogous to equation (3.54)) and put some noise on this signal, i.e. we add a normal distributed random value with mean value 0 and a fixed standard deviation. Then we have simulated the observation vector \mathbf{y} and can proceed in our simulation.

For our simulation we choose 34 receivers R_k , $k = 1, \dots, 34$ and 69 satellite positions S_l , $l = 1, \dots, 69$, hence we get altogether $34 \cdot 69 = 2346$ signal-paths. As mentioned before the signal-paths with a zenith angle larger

than 80° cannot be used and have to be neglected and, thus 1676 signal-paths remain. I.e. $\begin{bmatrix} (R_1, S_1, t_{1,1}) \\ \vdots \\ (R_{34}, S_{69}, t_{34,69}) \end{bmatrix}$

from now on means the 1676×1 vector. We will adopt this notation for all vector and matrix notations. The receiver positions, the satellite orbits and the signal-paths for two selected receivers R_1 and R_2 are shown in Figure 3.10.

For all of these ray-path we calculate the simulated observation $y(R_k, S_l, t_{k,l})$ and we gain the 1676×1 obser-

$$\text{vation vector } \mathbf{y} = \begin{bmatrix} y(R_1, S_1, t_{1,1}) \\ \vdots \\ y(R_{34}, S_{69}, t_{34,69}) \end{bmatrix}.$$

We insert the vector \mathbf{y} in the Gauss-Markov model (3.37) to model a correction term $\Delta STEC_{J_1, J_2, J_3}(R, S, t)$ of levels $(J_1, J_2, J_3) = (4, 4, 4)$ to our reference model $STEC_{ref}(R, S, t) = STEC_{3,3,3}(R, S, t)$, i.e. we have to estimate the $(m_4 \cdot m_4 \cdot m_4) \times 1 = 18^3 \times 1 = 5832 \times 1$ coefficient vector \mathbf{d} .

We now have to solve the extended normal equation (3.42), i.e.

$$(\mathbf{K}' \mathbf{P}_y \mathbf{K} + \lambda \mathbf{P}_d) \hat{\mathbf{d}} = \mathbf{K}' \mathbf{P}_y \mathbf{y} + \lambda \mathbf{P}_d \boldsymbol{\mu}_d \quad (3.55)$$

wherein \mathbf{K} is a $1676 \times (m_{J_4} \times m_{J_4} \times m_{J_4}) = 1676 \times 5832$ matrix as defined in equation (3.35), \mathbf{P}_y is a given positive definite 1676×1676 weight matrix and \mathbf{P}_d is a given positive definite 5832×5832 weight matrix.

Since the elements in our observation vector \mathbf{y} represent noisy residual values (i.e. a noisy version of the difference between our simulated measurements $STEC(R, S, t)$ and a reference model $STEC_{ref}(R, S, t)$) we assume the values in the observation vector to be located around 0. Therefore we assume for the expectation vector of the coefficients $E(\mathbf{d}) = \boldsymbol{\mu}_d = \mathbf{0}$ and rewrite equation (3.55) as

$$(\mathbf{K}' \mathbf{P}_y \mathbf{K} + \lambda \mathbf{P}_d) \hat{\mathbf{d}} = \mathbf{K}' \mathbf{P}_y \mathbf{y}. \quad (3.56)$$

To estimate the regularization parameter λ we may apply a fast Monte-Carlo implementation of the iterative maximum-likelihood Variance Component Estimation, for more details see KOCH AND KUSCHE (2001) or SCHMIDT ET AL. (2007a). For our purpose we insert different values for λ and choose the result with the lowest rms values of the residuals ($\hat{\mathbf{e}} = \mathbf{K} \hat{\mathbf{d}} - \mathbf{y}$, see equation (3.52)).

For this simulation, however, we assume that all observations have the same weight, hence, we fix $\mathbf{P}_y = \mathbf{I}$ to be a 1676×1676 unit matrix. For the 5832×5832 weight matrix of the prior information \mathbf{P}_d we apply the weight matrix introduced in equation (3.43). Finally we have to solve the equation

$$(\mathbf{K}' \mathbf{K} + \lambda \mathbf{P}_d) \hat{\mathbf{d}} = \mathbf{K}' \mathbf{y}. \quad (3.57)$$

and can estimate the least squares solution for the coefficient vector $\hat{\mathbf{d}}$ uniquely for an appropriate λ .

With the estimated coefficient vector $\hat{\mathbf{d}} = \hat{\mathbf{d}}_i^{J_1, J_2, J_3}$ of levels $(J_1, J_2, J_3) = (4, 4, 4)$ we can now calculate the correction term $\widehat{\Delta STEC}_{J_1, J_2, J_3} = \widehat{\Delta STEC}_{4,4,4}$ to our reference model $STEC_{ref} = STEC_{3,3,3}$.

The modeled correction term $\widehat{\Delta STEC}_{4,4,4}$ is calculated with the estimated coefficient vector $\hat{\mathbf{d}}$ via

$$\widehat{\Delta STEC}_{4,4,4}(R, S, t) = \left[k_{0,0,0}^{4,4,4}(R, S) \dots k_{m_4-1, m_4-1, m_4-1}^{4,4,4}(R, S) \right] \cdot \hat{\mathbf{d}} \text{ for } t \in [t_i, t_{i+1}). \quad (3.58)$$

We can compare the values of the correction term $\Delta STEC(R, S, t) = STEC(R, S, t) - STEC_{ref}(R, S, t) = STEC_{4,4,4}(R, S, t) - STEC_{3,3,3}(R, S, t)$ (calculated by the coefficient vectors $\mathbf{d}^{4,4,4}(t)$ and $\mathbf{d}^{3,3,3}(t)$, which have been estimated in section 3.3.1, analogous to equation (3.54)) with our approximated correction term $\widehat{\Delta STEC}_{4,4,4}(R, S, t)$ (calculated by the estimated coefficient vector $\hat{\mathbf{d}}$). For that purpose we choose $t = 14$ UT. Since we cannot visualize the values of the STEC for different positions in longitude and latitude of the receiver R and satellite S we compute values for receiver positions R_k and satellite positions S_k with the same longitude and latitude. The integration that is performed in the $k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R_k, S_k)$ then represents an integration along a **vertical ray-path** from receiver position R_k to satellite position S_k ; i.e. $\Delta STEC = \Delta VTEC$.

To visualize the results of our computations we now lie a grid in longitude and latitude over our area of consideration, e.g. a grid from $[250^\circ, 340^\circ]$ in longitude and $[-60^\circ, 30^\circ]$ in latitude with a 3° spacing in longitude and latitude, i.e. we have 31 receiver positions and 31 satellite positions and therefore we have

altogether 961 grid points $\begin{bmatrix} (R_1, S_1) \\ \vdots \\ (R_{31}, S_{31}) \end{bmatrix}$. For each of these points on the grid we now calculate the values of $\Delta VTEC = \begin{bmatrix} \Delta VTEC(R_1, S_1, t) \\ \vdots \\ \Delta VTEC(R_{31}, S_{31}, t) \end{bmatrix}$ and the values of the modeled correction term $\widehat{\Delta VTEC}_{4,4,4} = \begin{bmatrix} \widehat{\Delta VTEC}_{4,4,4}(R_1, S_1, t) \\ \vdots \\ \widehat{\Delta VTEC}_{4,4,4}(R_{31}, S_{31}, t) \end{bmatrix}$.

Figure 3.11 shows the results. The left column shows the simulation as described before. The top left panel shows $\Delta VTEC$ wherein $\Delta VTEC(R_k, S_l, t) = VTEC(R_k, S_l, t) - VTEC_{ref}(R_k, S_l, t) = VTEC_{4,4,4}(R_k, S_l, t) - VTEC_{3,3,3}(R_k, S_l, t)$, i.e. our input data computed from IRI, the second left panel shows the estimated correction term $\widehat{\Delta VTEC}_{4,4,4}$ computed for the levels $(J_1, J_2, J_3) = (4, 4, 4)$ and the third left panel shows the deviations between $\Delta VTEC$ and the estimated correction term $\widehat{\Delta VTEC}_{4,4,4}$; the rms value of the deviations

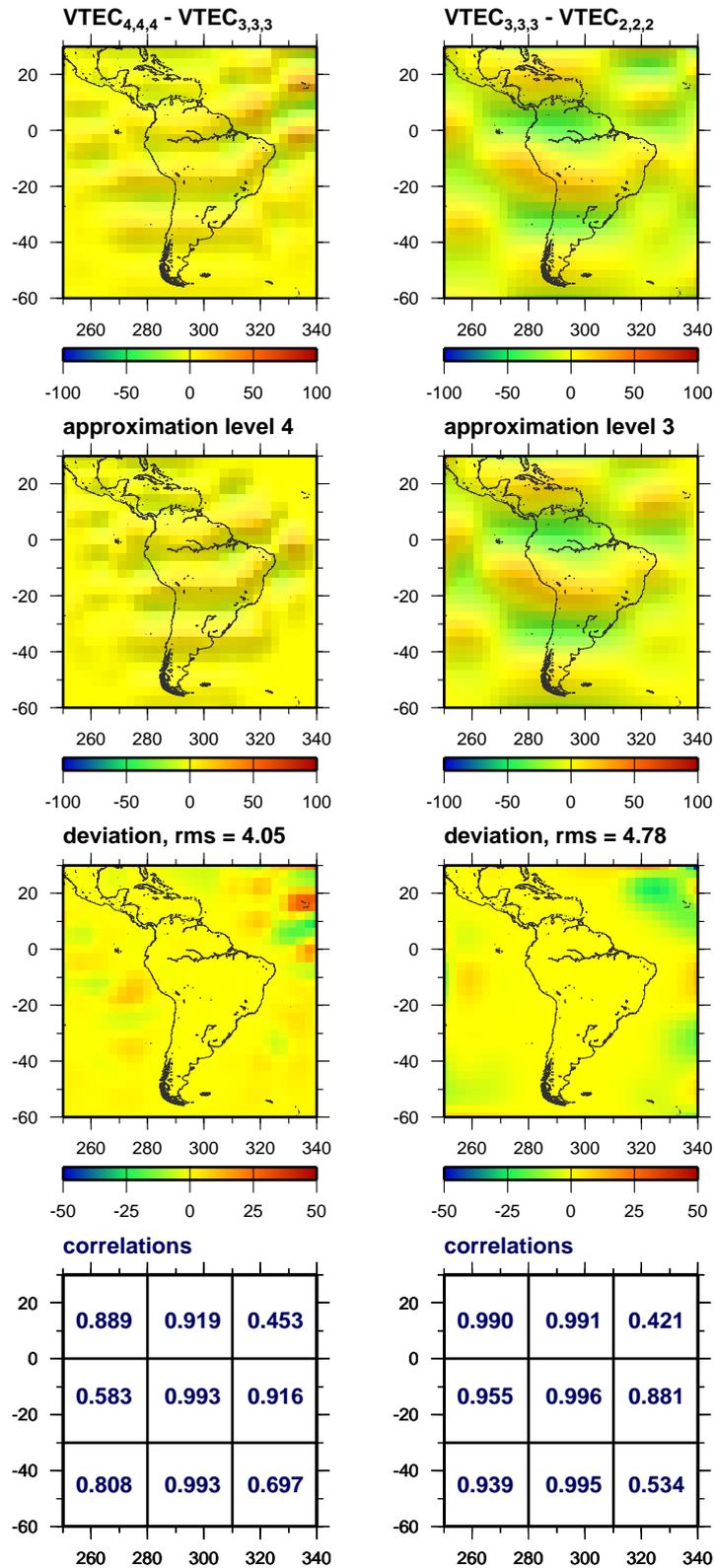


Figure 3.11: The top left (top right) panel shows our simulated residual $\Delta VTEC(R, S, t) = VTEC_{4,4,4}(R, S, t) - VTEC_{3,3,3}(R, S, t)$ ($\Delta VTEC(R, S, t) = VTEC_{3,3,3}(R, S, t) - VTEC_{2,2,2}(R, S, t)$), the second left (second right) panel shows the estimated correction term $\Delta \widehat{VTEC}_{4,4,4}$ ($\Delta \widehat{VTEC}_{3,3,3}$) and the third left (third right) panel shows the deviations; data sets in [TECU]. The fourth left (fourth right) panel shows the correlations of $\Delta VTEC(R, S, t)$ and $\Delta \widehat{VTEC}_{4,4,4}(R, S, t)$ (of $\Delta VTEC(R, S, t)$ and $\Delta \widehat{VTEC}_{3,3,3}(R, S, t)$) on the subareas.

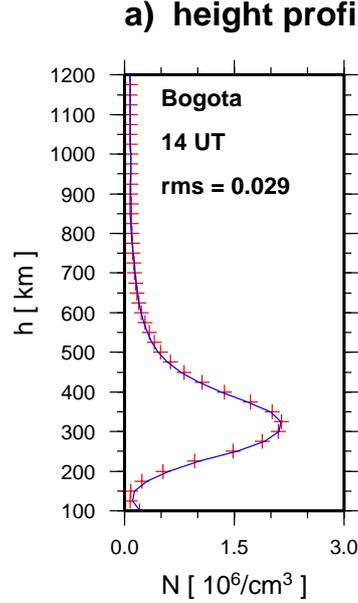


Figure 3.12: Height profile of the electron density in Bogota in Colombia.

amounts 4.05 TECU (total electron content unit, 1 TECU = 10^{16} electrons/m²). In order to prove that these deviations are mainly due to the uneven data distribution (cf. Figure 3.10) we divide the area under consideration into nine subareas and calculate the correlations between $\Delta VTEC$ and the estimated correction term $\Delta \widehat{VTEC}_{4,4,4}$ within these subareas. The last left panel shows the results. The correlations are higher in the areas where more observations are available and the correlations are lower in those areas where less observations are available.

The panels in the second column shows the same calculations for lower levels.

With the estimated coefficient vector $\hat{\mathbf{d}}$ and the reference model we can now compute a height profile of the electron density, e.g. for Bogota in Colombia for a specified time $t \in [t_i, t_{i+1})$. With equation (3.21) we compute the electron density of our reference model $N_{ref}(\lambda, \varphi, h, t) = N_{3,3,3}(\lambda, \varphi, h, t)$. With our estimated coefficient vector $\hat{\mathbf{d}}$ of levels $(J_1, J_2, J_3) = (4, 4, 4)$ we compute a correction term to the electron density $\Delta \hat{N}_{4,4,4}(\lambda, \varphi, h, t)$ by

$$\Delta \hat{N}_{4,4,4}(\lambda, \varphi, h, t) = \sum_{k_1=0}^{m_4-1} \sum_{k_2=0}^{m_4-1} \sum_{k_3=0}^{m_4-1} \hat{d}_{k_1, k_2, k_3; i}^{4,4,4} \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, h). \quad (3.59)$$

wherein $\hat{d}_{k_1, k_2, k_3; i}^{J_1, J_2, J_3}$ are the coefficients saved in the vector $\hat{\mathbf{d}}$. Figure 3.12 shows the electron density of our reference model $N_{3,3,3}(\lambda, \varphi, h, t)$ (crosses) and the corrected version of the electron density $N_{3,3,3}(\lambda, \varphi, h, t) + \Delta \hat{N}_{4,4,4}(\lambda, \varphi, h, t)$ (solid curve) for a fixed time t at Bogota for discrete heights h with spacing 25 km. The rms of the deviations amounts $0.029 \times [10^6/\text{cm}^3]$.

Extension in 4 dimensions

In the 3-dimensional approach which we used for modeling the electron density we had to perform a discretization in time. The scaling coefficients $d_{k_1, k_2, k_3}^{J_1, J_2, J_3}(t)$ are dependent of the time t . In order not to use a temporal discretization we may introduce the 4-dimensional B-spline approach for the electron density

$$\Delta N(\lambda, \varphi, h, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} \sum_{k_4=0}^{m_{J_4}-1} d_{k_1, k_2, k_3, k_4}^{J_1, J_2, J_3, J_4} \phi_{k_1}^{J_1}(\lambda) \phi_{k_2}^{J_2}(\varphi) \phi_{k_3}^{J_3}(h) \phi_{k_4}^{J_4}(t). \quad (3.60)$$

Substituting equation (3.60) instead of equation (3.26) in the observation equation (3.22) yields us to

$$y(R, S, t) + e(R, S, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} \sum_{k_4=0}^{m_{J_4}-1} d_{k_1, k_2, k_3, k_4}^{J_1, J_2, J_3, J_4} k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) \phi_{k_4}^{J_4}(t) + \beta_R + \beta_S \quad (3.61)$$

wherein $k_{k_1, k_2, k_3}^{J_1, J_2, J_3}(R, S) = \alpha \int_R^S \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, h) ds$.

Here we chose a smooth function for modeling the time-dependency. On the other hand large linear equation

systems have to be solved in order to estimate the $m_{J_1} \cdot m_{J_2} \cdot m_{J_3} \cdot m_{J_4}$ unknown 4-dimensional scaling coefficients $d_{k_1, k_2, k_3, k_4}^{J_1, J_2, J_3, J_4}$.

3.3.3 3-dimensional COSMIC VTEC modeling

As described in section 3.1.2 we can derive VTEC observations from COSMIC occultation measurements.

In our investigation the electron density was calculated in a pre-processing step from so-called compensated STEC values by the improved Abel transform. This work was done by the group of Prof. Tsai; for more details see TSAI AND TSAI (2004).

Consequently, we do not assume locally spherical symmetry within the ionosphere, but consider the effect of large-scale horizontal gradients and/or inhomogeneous electron density distribution.

According to $VTEC_{COSMIC} = VTEC_0 + VTEC_1$ a VTEC observation consists of two parts, namely

- $VTEC_0$ the integration of the calculated electron density along the vertical from bottom to orbital height, cf. equation (3.3).
- $VTEC_1$ an extrapolated model value for the range between orbital height and the top of the ionosphere (L.-C. Tsai, personal communication).

Again we divide $VTEC(\lambda, \varphi, t)$ into a reference model $VTEC_{ref}(\lambda, \varphi, t)$, here computed from IRI, and into a correction term $\Delta VTEC(\lambda, \varphi, t)$, i.e. $VTEC(\lambda, \varphi, t) = VTEC_{ref}(\lambda, \varphi, t) + \Delta VTEC(\lambda, \varphi, t)$. We suppose to have given COSMIC VTEC observations $VTEC_{COSMIC}(\lambda_i, \varphi_i, t_i)$ for $i = 1, \dots, M$ in a specified region. From those observations we estimate a correction term $\Delta VTEC(\lambda, \varphi, t)$ to our reference model $VTEC_{ref}(\lambda, \varphi, t)$. For that purpose we establish the linear equation system

$$\begin{bmatrix} VTEC_{COSMIC}(\lambda_1, \varphi_1, t_1) \\ \vdots \\ VTEC_{COSMIC}(\lambda_M, \varphi_M, t_M) \end{bmatrix} + \begin{bmatrix} e(\lambda_1, \varphi_1, t_1) \\ \vdots \\ e(\lambda_M, \varphi_M, t_M) \end{bmatrix} = \begin{bmatrix} VTEC_{ref}(\lambda_1, \varphi_1, t_1) \\ \vdots \\ VTEC_{ref}(\lambda_M, \varphi_M, t_M) \end{bmatrix} + \begin{bmatrix} \Delta VTEC(\lambda_1, \varphi_1, t_1) \\ \vdots \\ \Delta VTEC(\lambda_M, \varphi_M, t_M) \end{bmatrix} \quad (3.62)$$

For the 3-dimensional correction term $\Delta VTEC(\lambda, \varphi, t)$ we apply our 3-dimensional B-spline model introduced in section 3.2.2, i.e. we model $\Delta VTEC(\lambda, \varphi, t)$ with our B-spline approach (compare equation (3.16)), according to

$$\Delta VTEC_{J_1, J_2, J_3}(\lambda, \varphi, t) = \sum_{k_1=0}^{m_{J_1}-1} \sum_{k_2=0}^{m_{J_2}-1} \sum_{k_3=0}^{m_{J_3}-1} d_{k_1, k_2, k_3}^{J_1, J_2, J_3} \phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, t) \quad (3.63)$$

and therefore we can form a matrix equation for $\Delta VTEC_{J_1, J_2, J_3}(\lambda_i, \varphi_i, t_i)$, $i = 1, \dots, M$,

$$\begin{bmatrix} \Delta VTEC_{J_1, J_2, J_3}(\lambda_1, \varphi_1, t_1) \\ \vdots \\ \Delta VTEC_{J_1, J_2, J_3}(\lambda_M, \varphi_M, t_M) \end{bmatrix} = \begin{bmatrix} \phi_{0,0,0}^{J_1, J_2, J_3}(\lambda_1, \varphi_1, t_1) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(\lambda_1, \varphi_1, t_1) \\ \vdots & & \vdots \\ \phi_{0,0,0}^{J_1, J_2, J_3}(\lambda_M, \varphi_M, t_M) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(\lambda_M, \varphi_M, t_M) \end{bmatrix} \cdot \mathbf{d}^{J_1, J_2, J_3} \quad (3.64)$$

wherein $\mathbf{d}^{J_1, J_2, J_3}$ is defined analogously to equation (3.13). Note, here the time is respected in the 3-dimensional scaling function and therefore we do not have to perform a discretization of the time.

Inserting the series expansion for the correction term (3.64) into equation (3.62) yields

$$\begin{bmatrix} VTEC_{COSMIC}(\lambda_1, \varphi_1, t_1) \\ \vdots \\ VTEC_{COSMIC}(\lambda_M, \varphi_M, t_M) \end{bmatrix} + \begin{bmatrix} e(\lambda_1, \varphi_1, t_1) \\ \vdots \\ e(\lambda_M, \varphi_M, t_M) \end{bmatrix} = \begin{bmatrix} VTEC_{ref}(\lambda_1, \varphi_1, t_1) \\ \vdots \\ VTEC_{ref}(\lambda_M, \varphi_M, t_M) \end{bmatrix} + \begin{bmatrix} \phi_{0,0,0}^{J_1, J_2, J_3}(\lambda_1, \varphi_1, t_1) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(\lambda_1, \varphi_1, t_1) \\ \vdots & & \vdots \\ \phi_{0,0,0}^{J_1, J_2, J_3}(\lambda_M, \varphi_M, t_M) & \dots & \phi_{m_{J_1}-1, m_{J_2}-1, m_{J_3}-1}^{J_1, J_2, J_3}(\lambda_M, \varphi_M, t_M) \end{bmatrix} \cdot \mathbf{d}^{J_1, J_2, J_3}$$

or in a short way

$$\mathbf{VTEC}_{COSMIC} + \mathbf{e} = \mathbf{VTEC}_{ref} + \mathbf{\Phi}^{J_1, J_2, J_3} \cdot \mathbf{d}^{J_1, J_2, J_3}. \quad (3.65)$$

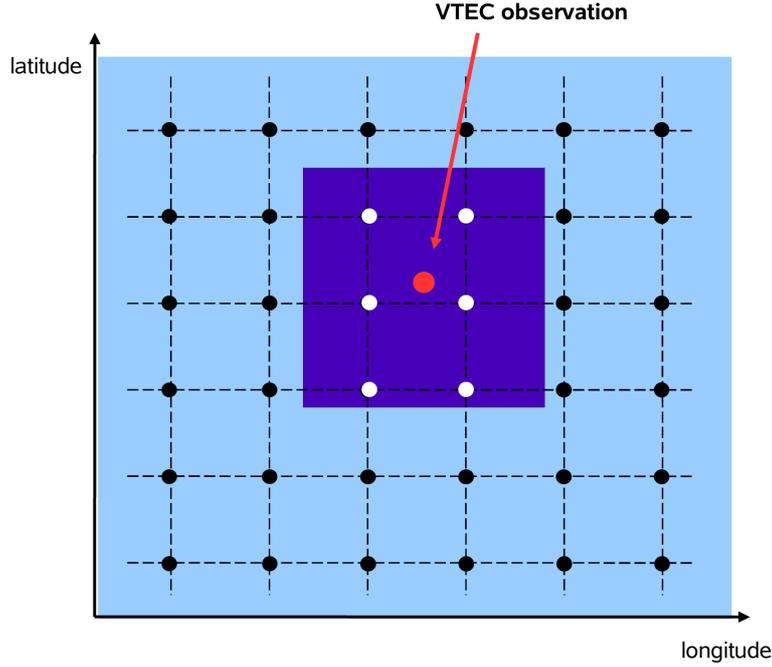


Figure 3.13: 2-dimensional problem in longitude and latitude. The dots correspond to the centers of the 2-dimensional B-spline scaling functions $\phi_{k_1, k_2}^{J_1, J_2}$. The white dots indicate the centers of those B-splines that have non-zero entries in the observation equation.

The short notations are defined as follows:

$$\mathbf{VTEC}_{COSMIC} = \left[VTEC_{COSMIC}(\lambda_1, \varphi_1, t_1) \dots VTEC_{COSMIC}(\lambda_M, \varphi_M, t_M) \right]',$$

$\mathbf{VTEC}_{ref} = \left[VTEC_{ref}(\lambda_1, \varphi_1, t_1) \dots VTEC_{ref}(\lambda_M, \varphi_M, t_M) \right]'$, $\mathbf{e} = \left[e(\lambda_1, \varphi_1, t_1) \dots e(\lambda_M, \varphi_M, t_M) \right]'$ and Φ^{J_1, J_2, J_3} is defined analogous to equation (3.15). From equation (3.65) we get the Gauss-Markov model analogous to equation (2.96)

$$\mathbf{y} + \mathbf{e} = \Phi^{J_1, J_2, J_3} \cdot \mathbf{d}^{J_1, J_2, J_3} \text{ with } D(\mathbf{y}) = \sigma_y^2 \mathbf{P}_y^{-1} \quad (3.66)$$

wherein $\mathbf{y} = \mathbf{VTEC}_{COSMIC} - \mathbf{VTEC}_{ref}$, $D(\mathbf{y})$ is the covariance matrix, σ_y^2 is the unknown variance factor and \mathbf{P}_y is the positive definite weight matrix of dimension $M \times M$. We solve the Gauss-Markov model analogous to section 2.3.5 by the least squares method and get the normal equation

$$\left((\Phi^{J_1, J_2, J_3})' \mathbf{P}_y \Phi^{J_1, J_2, J_3} \right) \mathbf{d}^{J_1, J_2, J_3} = (\Phi^{J_1, J_2, J_3})' \mathbf{P}_y \mathbf{y}. \quad (3.67)$$

The solvability of the model depends on the column rank of the matrix Φ^{J_1, J_2, J_3} , for more details see section 2.3.5.

As already explained in section 3.3.2 the 3-dimensional B-spline scaling functions $\phi_{k_1, k_2, k_3}^{J_1, J_2, J_3}(\lambda, \varphi, t)$ are related to grid points $(\lambda_{k_1}, \varphi_{k_2}, t_{k_3})$ within the unit cube, $k_i = 0, \dots, m_{J_i} - 1$, $i = 1, 2, 3$ for fixed J_1, J_2 and J_3 . For simplification Figure 3.13 shows a 2-dimensional problem by neglecting the time-dependency. The white and black dots indicate the centers $(\lambda_{k_1}, \varphi_{k_2})$ of the 2-dimensional B-spline scaling functions $\phi_{k_1, k_2}^{J_1, J_2}(\lambda, \varphi) = \phi_{k_1}^{J_1}(\lambda) \phi_{k_2}^{J_2}(\varphi)$ in the longitude-latitude plane (unit square). Due to their compact support only B-splines related to the white dots, have non-zero entries in the observation equation shown before. Thus, a scaling coefficient $d_{k_1, k_2, k_3}^{J_1, J_2, J_3}$ is computable only if observations are given close to the peak of the corresponding scaling function. Hence, in case of data gaps many scaling coefficients may not be calculable and the corresponding addends can be excluded from the observation equation or there has to be introduced prior information to estimate these coefficients. If just a few observations support the computation of a coefficient, we may also introduce prior information in order to stabilize the estimation process, i.e. to perform a regularization.

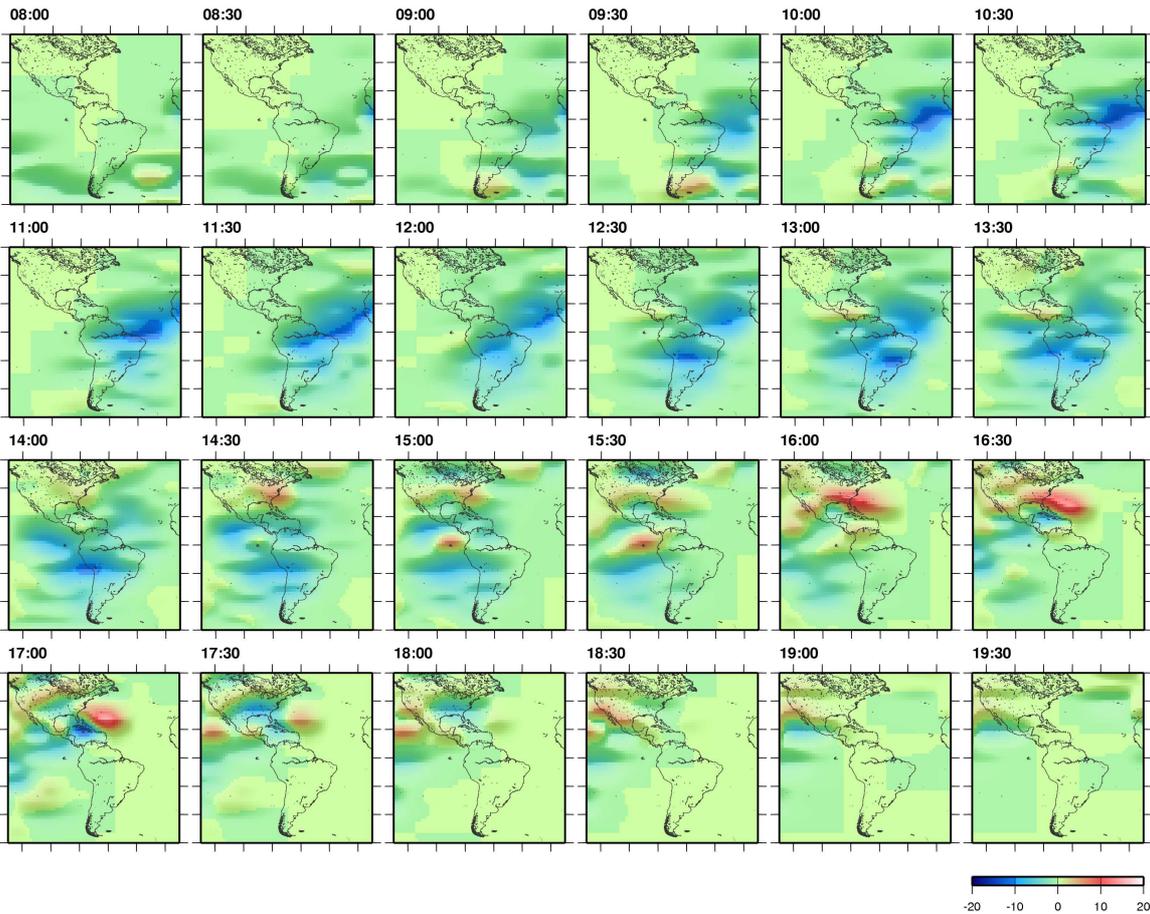


Figure 3.14: "Snapshots" of residual VTEC estimations between 8:00 and 19:30 UT; data sets in [TECU].

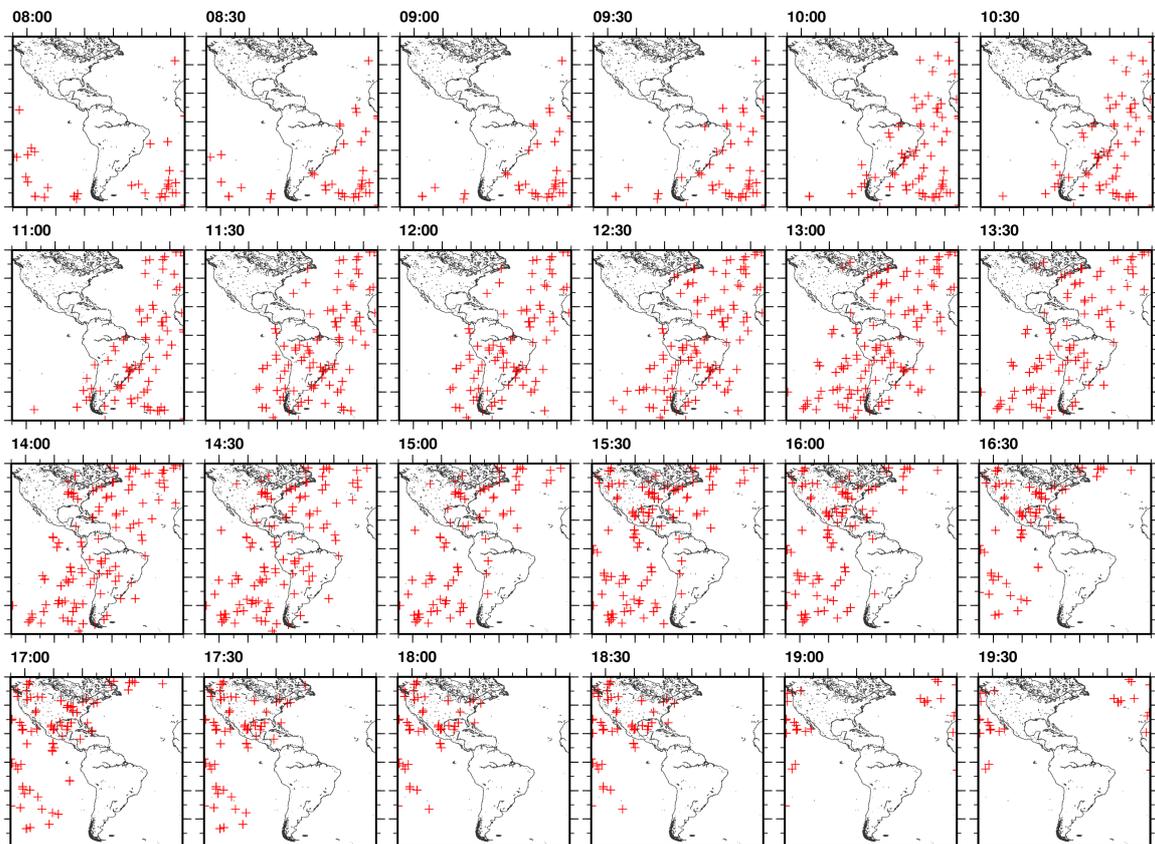


Figure 3.15: Observations that have influence on the computation of the corresponding residuals in Figure 3.14.

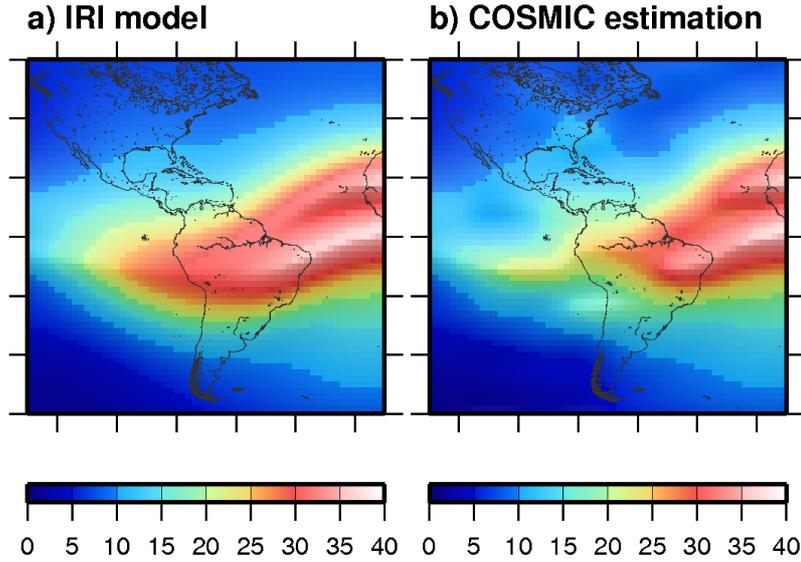


Figure 3.16: a) VTEC model data from IRI for July 21st, 14:00 UT; b) sum of the IRI input model (panel a)) and the residual VTEC estimation computed from equation (3.63); data sets in [TECU].

Figure 3.16 compares the VTEC data from IRI (i.e. $VTEC_{ref} = VTEC_{IRI}$) and the COSMIC estimation (i.e. $VTEC_{IRI} + \Delta VTEC_{J_1, J_2, J_3}$).

As mentioned in chapter 2.3.5 the MSA based on wavelet decomposition is an effective tool for data compression or data reduction. With the decomposition relations (2.111) to (2.126), which have to be adapted to the 3-dimensional model, we can decompose the estimated scaling coefficient vector $\hat{\mathbf{d}} = \mathbf{d}_{3,4,5}$ into the scaling coefficient vector of the lowest levels, i.e. $\mathbf{d}_{0,0,0}$ and the corresponding detail coefficient vectors $\mathbf{c}_{2,3,4}^{\nu_1}$, $\mathbf{c}_{1,2,3}^{\nu_1}$, $\mathbf{c}_{0,1,2}^{\nu_1}$, $\mathbf{c}_{0,0,1}^{\nu_2}$ and $\mathbf{c}_{0,0,0}^{\nu_3}$, wherein $\nu_1 = 7 = 2^3 - 1$ corresponds to the coefficients for which we decrease the levels in all 3 dimensions, $\nu_2 = 3 = 2^2 - 1$ corresponds to the coefficients for which we decrease the levels only in 2 dimensions and $\nu_3 = 1 = 2^1 - 1$ corresponds to the coefficients for which we decrease the level in only 1 dimension (Note, the decomposition relations analogous to (2.111) to (2.126) apply to the steps from levels (3, 4, 5) to levels (2, 3, 4), to levels (1, 2, 3) and to levels (0, 1, 2); for the steps (0, 1, 2) to (0, 0, 1) and to (0, 0, 0) we have to construct decomposition relations analogous to equation (2.132). For the reconstruction relations we will have to work analogously but will not discuss this concept.). Figure 3.17 visualizes the decomposition of the coefficients: we start with the coefficient vector $\mathbf{d}_{3,4,5}$ and compute stepwise the coefficient vectors of lower levels. For the reconstruction we only have to save the coefficient vectors $\mathbf{d}_{0,0,0}$, $\mathbf{c}_{0,0,1}^{\nu}$, $\mathbf{c}_{0,1,2}^{\nu}$, $\mathbf{c}_{1,2,3}^{\nu}$ and $\mathbf{c}_{2,3,4}^{\nu}$ for the corresponding values of ν .

Now we can apply the data compression algorithm as introduced in chapter 2.3.5. The coefficient vectors $\bar{\mathbf{c}}_{0,0,0}^{\nu}$, $\bar{\mathbf{c}}_{0,0,1}^{\nu}$, $\bar{\mathbf{c}}_{0,1,2}^{\nu}$, $\bar{\mathbf{c}}_{1,2,3}^{\nu}$ and $\bar{\mathbf{c}}_{2,3,4}^{\nu}$ are the compressed versions of the detail coefficient vectors wherein we neglect those coefficients in the detail coefficient vectors whose absolute values are lower than a specified threshold $\epsilon > 0$. Therefore the compressed version of the detail coefficients do not contain the non-significant structures. With the reconstruction relation (2.131), which has to be adapted to the 3-dimensional model, we can now perform a reconstruction of the scaling coefficient vector $\mathbf{d}_{3,4,5}$. The reconstruction algorithm is visualized in Figure 3.18. In the left panel we insert the coefficient vectors $\mathbf{d}_{0,0,0}$, $\mathbf{c}_{0,0,0}^{\nu}$, $\mathbf{c}_{0,0,1}^{\nu}$, $\mathbf{c}_{0,1,2}^{\nu}$, $\mathbf{c}_{1,2,3}^{\nu}$ and $\mathbf{c}_{2,3,4}^{\nu}$ (computed with the decomposition algorithm) into the reconstruction relation and derive the exact coefficient vector $\mathbf{d}_{3,4,5}$ while in the right panel we insert the compressed versions of the detail coefficient vectors ($\mathbf{d}_{0,0,0}$, $\bar{\mathbf{c}}_{0,0,0}^{\nu}$, $\bar{\mathbf{c}}_{0,0,1}^{\nu}$, $\bar{\mathbf{c}}_{0,1,2}^{\nu}$, $\bar{\mathbf{c}}_{1,2,3}^{\nu}$ and $\bar{\mathbf{c}}_{2,3,4}^{\nu}$) and derive a compressed version of the coefficient vector $\bar{\mathbf{d}}_{3,4,5}$.

Note from now on for simplification we omit the hat in the notations of the estimated correction terms and in the following always write $\Delta VTEC_{3,4,5}$ instead of $\Delta \widehat{VTEC}_{3,4,5}$.

Via the coefficient vector $\mathbf{d}_{3,4,5}$ the approximated correction term $\Delta VTEC_{3,4,5}$ is defined and via the coefficient vector $\bar{\mathbf{d}}_{3,4,5}$ the compressed version of the approximated correction term $\overline{\Delta VTEC}_{3,4,5}$ is defined, the two signals $\Delta VTEC_{3,4,5}$ and $\overline{\Delta VTEC}_{3,4,5}$ only differ in the non-significant structures which have been neglected

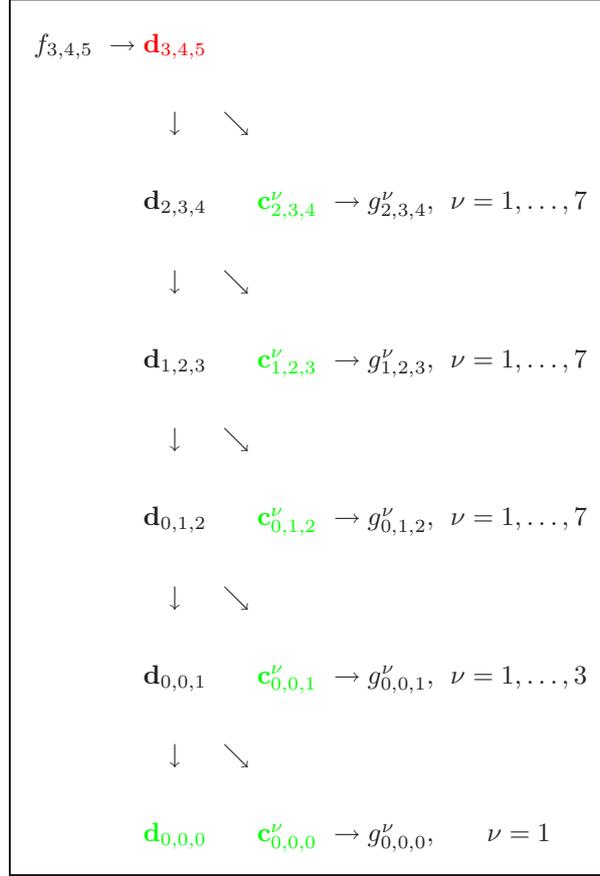


Figure 3.17: Decomposition algorithm for the coefficients.

in $\overline{\Delta VTEC}_{3,4,5}$. We can decompose the approximations

$$\Delta VTEC_{3,4,5} = \Delta VTEC_{0,0,0} + \sum_{\nu=1}^1 g_{0,0,0}^\nu + \sum_{\nu=1}^3 g_{0,0,1}^\nu + \sum_{\nu=1}^7 g_{0,1,2}^\nu + \sum_{\nu=1}^7 g_{1,2,3}^\nu + \sum_{\nu=1}^7 g_{2,3,4}^\nu \quad (3.71)$$

and

$$\overline{\Delta VTEC}_{3,4,5} = \Delta VTEC_{0,0,0} + \sum_{\nu=1}^1 \overline{g}_{0,0,0}^\nu + \sum_{\nu=1}^3 \overline{g}_{0,0,1}^\nu + \sum_{\nu=1}^7 \overline{g}_{0,1,2}^\nu + \sum_{\nu=1}^7 \overline{g}_{1,2,3}^\nu + \sum_{\nu=1}^7 \overline{g}_{2,3,4}^\nu. \quad (3.72)$$

We show the results of the data compression. Since the data compression algorithm does not affect the scaling coefficient vector $\mathbf{d}_{0,0,0}$ we do not have to take into account the part $\Delta VTEC_{0,0,0}$ in (3.71) and (3.72), therefore we consider the signals $\Delta VTEC_{3,4,5} - \Delta VTEC_{0,0,0}$ and $\overline{\Delta VTEC}_{3,4,5} - \Delta VTEC_{0,0,0}$, i.e. the sums of the detail signals in (3.71) and (3.72). The detail signals $g_{0,0,0}^\nu, g_{0,0,1}^\nu, g_{0,1,2}^\nu, g_{1,2,3}^\nu$ and $g_{2,3,4}^\nu$ and the compressed detail signals $\overline{g}_{0,0,0}^\nu, \overline{g}_{0,0,1}^\nu, \overline{g}_{0,1,2}^\nu, \overline{g}_{1,2,3}^\nu$ and $\overline{g}_{2,3,4}^\nu$ are represented by the coefficient vectors $\mathbf{c}_{0,0,0}^\nu, \mathbf{c}_{0,0,1}^\nu, \mathbf{c}_{0,1,2}^\nu, \mathbf{c}_{1,2,3}^\nu$ and $\mathbf{c}_{2,3,4}^\nu$ and the compressed coefficient vectors $\overline{\mathbf{c}}_{0,0,0}^\nu, \overline{\mathbf{c}}_{0,0,1}^\nu, \overline{\mathbf{c}}_{0,1,2}^\nu, \overline{\mathbf{c}}_{1,2,3}^\nu$ and $\overline{\mathbf{c}}_{2,3,4}^\nu$, respectively. The coefficient vectors contain altogether $m_5 \cdot m_4 \cdot m_3 - m_0 \cdot m_0 \cdot m_0 = 6120 - 27 = 6093$ detail coefficients. The compressed coefficient vectors contain altogether $m_5 \cdot m_4 \cdot m_3 - m_0 \cdot m_0 \cdot m_0 - n_\epsilon = 6120 - 27 - n_\epsilon$ detail coefficients, wherein n_ϵ is the number of coefficients that have been neglected. Therefore we here define the compression rate $\rho := \frac{n_\epsilon}{m_5 \cdot m_4 \cdot m_3 - m_0 \cdot m_0 \cdot m_0} = \frac{n_\epsilon}{6093}$. Note, as we do not take into account the smoothed signal $f_{0,0,0}$, we do not take into account the scaling coefficient vector $\mathbf{d}_{0,0,0}$.

In Figure 3.19 there are shown the signals $\Delta VTEC_{3,4,5} - \Delta VTEC_{0,0,0}$ and $\overline{\Delta VTEC}_{3,4,5} - \Delta VTEC_{0,0,0}$, i.e. Panel a) in the first row shows $\Delta VTEC_{3,4,5} - \Delta VTEC_{0,0,0}$ based on 6093 detail coefficients. Panel b) in the first row visualizes the result $\overline{\Delta VTEC}_{3,4,5} - \Delta VTEC_{0,0,0}$ after neglecting 4833 wavelet coefficients (threshold $\epsilon = 0.2$), i.e. the compression rate amounts $\rho = \frac{4833}{6093} = 79\%$. Panel c) shows the differences between panel a) and b); the rms value of the deviations amounts 0.749 TECU. The panels in the second row show the results

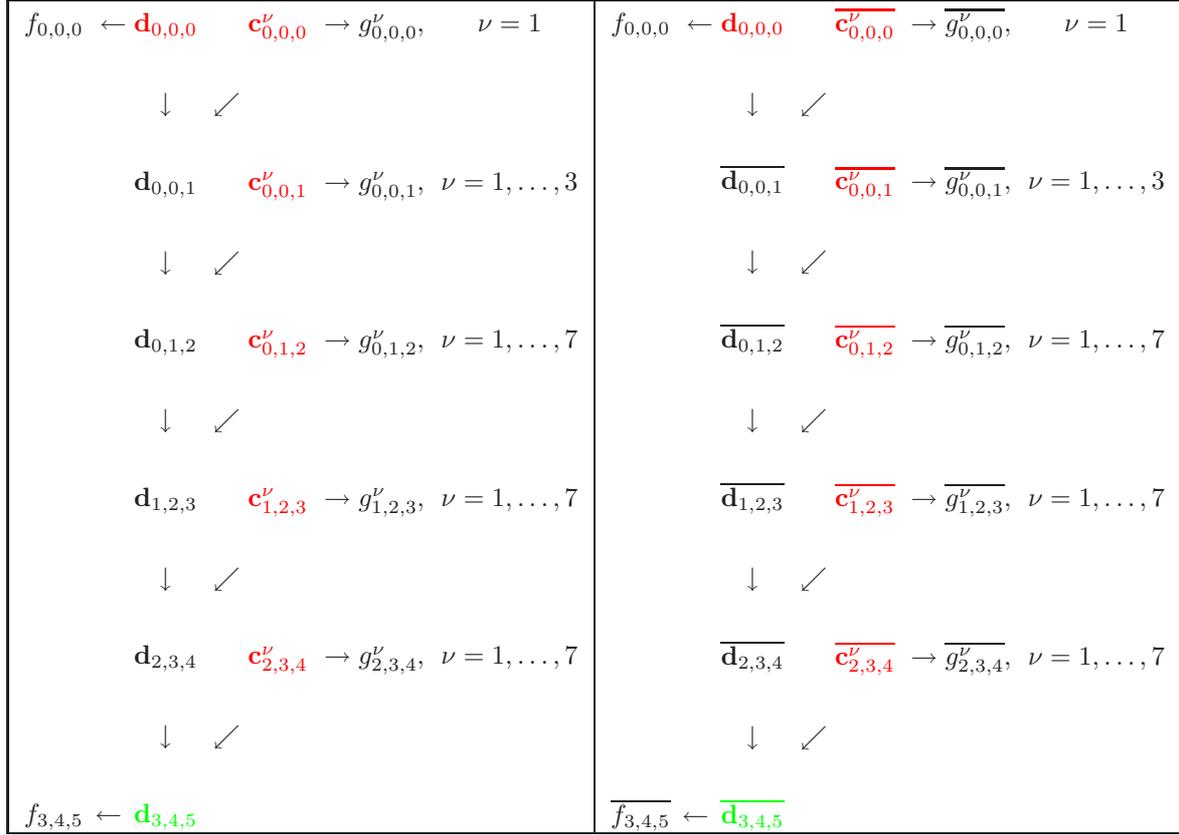


Figure 3.18: Reconstruction algorithm for the exact reconstruction and for a compressed reconstruction.

threshold ϵ	0.025	0.05	0.075	0.1	0.125	0.15	0.175	0.2
neglected coeff.	2491	3258	3715	4052	4314	4523	4690	4833
compression rate	40.9%	53.5%	61.0%	66.5%	70.8%	74.2%	77.0%	79.3%
rms	0.217	0.358	0.419	0.454	0.550	0.626	0.644	0.749
correlation	0.994	0.983	0.977	0.973	0.960	0.949	0.947	0.931
threshold ϵ	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6
neglected coeff.	5087	5222	5373	5477	5565	5634	5690	5738
compression rate	83.5%	85.7%	88.2%	89.9%	91.3%	92.5%	93.4%	94.2%
rms	0.918	1.084	1.062	1.083	1.037	1.064	1.071	1.092
correlation	0.899	0.865	0.867	0.870	0.872	0.870	0.869	0.860

Table 3.1: Results of the data compression: number of neglected coefficients, compression rate, rms values of the deviations and correlations for different thresholds $\epsilon > 0$. The rms values and the correlation refer to the estimation based on 6093 coefficients and the compressed estimation.

after neglecting 4052 coefficients (threshold $\epsilon = 0.1$). The rms values of the deviations and the correlations can be depicted from Table 3.1.

We have to make a trade-off between compression rate and rms value of the deviations and correlation. I.e if we want to derive a high compression rate we have higher rms values of the deviations and lower correlations. If we compute the rms value of the input data, we may choose, for instance, one third of the rms value as threshold for the rms value of the deviation ($\Delta VTEC_{3,4,5} - \overline{\Delta VTEC_{3,4,5}}$). Table 3.1 shows the results of the data compression for different thresholds ϵ .

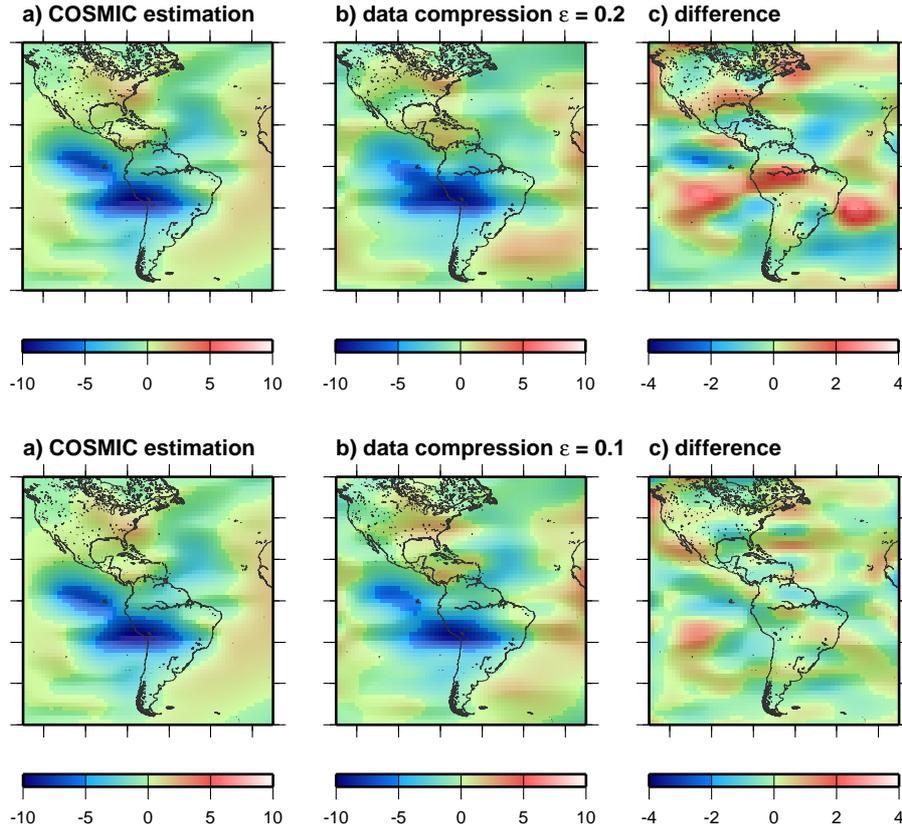


Figure 3.19: a) COSMIC estimation based on 6093 coefficients; b) estimation after applying the data compression algorithm for thresholds $\epsilon = 0.2$ and $\epsilon = 0.1$; c) deviations; data sets in [TECU].

By applying more flexible algorithms for data compression we may gain better results for the compressed signals. E.g. we may not use a fixed threshold ϵ for all coefficient vectors, but use different thresholds $\epsilon_{0,0,0}$, $\epsilon_{0,0,1}$, $\epsilon_{0,1,2}$, $\epsilon_{1,2,3}$ and $\epsilon_{2,3,4} > 0$ for the different levels. Here we neglect the coefficients in $\mathbf{c}_{0,0,0}'$ with absolute value lower than $\epsilon_{0,0,0}$, we neglect the coefficients in $\mathbf{c}_{0,0,1}'$ with absolute value lower than $\epsilon_{0,0,1}$, and so on. It is reasonable to use larger thresholds ϵ_{J_1, J_2, J_3} for higher levels (J_1, J_2, J_3) , since for high (J_1, J_2, J_3) the spaces V_{J_1, J_2, J_3} contain small structures and therefore we can neglect more coefficients.

To compare the results with the before applied data compression algorithm, we fixed an $\epsilon_{2,3,4} > 0$, and set $\epsilon_{1,2,3} = \frac{1}{2} \epsilon_{2,3,4}$, $\epsilon_{0,1,2} = \frac{1}{4} \epsilon_{2,3,4}$, $\epsilon_{0,0,1} = \frac{1}{8} \epsilon_{2,3,4}$ and $\epsilon_{0,0,0} = \frac{1}{16} \epsilon_{2,3,4}$. We fix $\epsilon_{2,3,4} = 0.45$ and neglect altogether 5344 coefficients (i.e. compression rate 87.7%) the rms value of the deviations amounts 0.714 TECU and the correlation amounts 0.934. Now we can compare these values with the values in Table 3.1 and see that for the columns in the table corresponding to $\epsilon = 0.2, 0.25$ and 0.3 there were less coefficients neglected, but the corresponding rms values of the deviations are higher and the correlation are lower. A data compression algorithm for which we can choose various thresholds for the different levels, therefore yields a more effective tool for data compression.

3.4 Combinations

The unknown B-spline scaling coefficients can be estimated from the combination of different observation types including altimetry and terrestrial measurements. For each of the different observation types it is possible to derive an observation equation for estimating the unknowns of the electron density B-spline model within a specific region. We suppose to have altogether p observation techniques. For each of the different techniques we have an observation vector \mathbf{y}_i , $i = 1, \dots, p$, e.g. \mathbf{y}_1 may contain all reduced GNSS geometry-free observations $y(R, S, t)$ (compare equation (3.30)) measured from ground stations within our specified region. Hence, for each observation vector \mathbf{y}_i we can establish a Gauss-Markov model

$$\mathbf{y}_i + \mathbf{e}_i = \mathbf{A}_i \boldsymbol{\beta} \text{ with } D(\mathbf{y}_i) = \sigma_{\mathbf{y}_i}^2 \mathbf{P}_{\mathbf{y}_i}^{-1} \quad (3.73)$$

wherein $\sigma_{\mathbf{y}_i}^2$ is the unknown variance factor and $\mathbf{P}_{\mathbf{y}_i}$ is the given positive definite weight matrix. The vector $\boldsymbol{\beta}$ consists of the unknown B-spline scaling coefficients and other auxiliary parameters. The combination of the p models yields the combined Gauss-Markov model

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_p \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_p \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_p \end{bmatrix} \boldsymbol{\beta} \text{ with } D \left(\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_p \end{bmatrix} \right) = \sigma_{\mathbf{y}_1}^2 \begin{bmatrix} \mathbf{P}_{\mathbf{y}_1}^{-1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + \dots + \sigma_{\mathbf{y}_p}^2 \begin{bmatrix} 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & \mathbf{P}_{\mathbf{y}_p}^{-1} \end{bmatrix} \quad (3.74)$$

for the unknown parameter vector $\boldsymbol{\beta}$ with unknown variance components $\sigma_{\mathbf{y}_i}^2$. According to equation (2.97) the corresponding normal equation system reads

$$\left(\sum_{i=1}^p \frac{1}{\sigma_{\mathbf{y}_i}^2} \mathbf{A}_i' \mathbf{P}_{\mathbf{y}_i} \mathbf{A}_i \right) \hat{\boldsymbol{\beta}} = \sum_{i=1}^p \frac{1}{\sigma_{\mathbf{y}_i}^2} \mathbf{A}_i' \mathbf{P}_{\mathbf{y}_i} \mathbf{y}_i. \quad (3.75)$$

To stabilize the estimation of the coefficient vector $\boldsymbol{\beta}$ we may have to introduce prior information, i.e. we introduce an additional linear model for the prior information analogous to (3.39) in section 3.3.2. Then we combine the models (3.74) and (3.39) and by solving the combined Gauss-Markov model with the least squares method we get the extended normal equation

$$\left(\sum_{i=1}^p \frac{1}{\sigma_{\mathbf{y}_i}^2} \mathbf{A}_i' \mathbf{P}_{\mathbf{y}_i} \mathbf{A}_i + \frac{1}{\sigma_{\boldsymbol{\beta}}^2} \mathbf{P}_{\boldsymbol{\beta}} \right) \hat{\boldsymbol{\beta}} = \sum_{i=1}^p \frac{1}{\sigma_{\mathbf{y}_i}^2} \mathbf{A}_i' \mathbf{P}_{\mathbf{y}_i} \mathbf{y}_i + \frac{1}{\sigma_{\boldsymbol{\beta}}^2} \mathbf{P}_{\boldsymbol{\beta}} \boldsymbol{\mu}_{\boldsymbol{\beta}}, \quad (3.76)$$

for more details see e.g. KOCH (2000).

The intention is to apply this combined model to the different observation types introduced in the sections 3.1.1, 3.1.2, 3.1.3 and 3.1.4. The different observation types have various advantages and disadvantages and with the combination of the different types the advantages of one type can be used to balance the disadvantages of another type. The different types may have different accuracies, therefore we consider different variance factors $\sigma_{\mathbf{y}_i}$ and we may introduce weight matrices $\mathbf{P}_{\mathbf{y}_i}$ with higher weights for the observations with high accuracy, while we may introduce lower weights for the observations with low accuracy. It is reasonable to consider combinations of different observation types, because, e.g., the GPS observations lie predominantly above the continent while altimetry yields observations above the sea, hence, the combination yields observations, both, above the sea and above the continent. From all different observation types some information can be used, different weights can be introduced to handle different accuracies of the observation types.

Chapter 4

Summary and conclusions

The first part of this report dealt with the derivation of the B-spline approach. In the section 2.1 we introduced the theory of a MSA in general but for the 1-dimensional case, i.e. we defined the scaling and detail spaces, introduced the scaling functions and the wavelets and established the two-scale relations. With the concept of a MSA we can decompose a signal in its low-frequency part and the corresponding detail signals. These signals are defined via coefficient vectors. The concept of a MSA leads us to a data compression technique.

In the next step we gave the definition of B-splines and some properties of these functions. Then we applied the theory of MSA to the B-splines, i.e. for the scaling functions we used B-splines and constructed semiorthogonal wavelet functions. With the two-scale coefficients we derived the decomposition and reconstruction relations for the B-spline model. The data compression algorithm can be applied and we can diminish the storage space of the signals. The 1-dimensional B-spline approach was extended via tensor product B-splines and wavelets to the multi-dimensional case, i.e. we established multi-dimensional approaches to model multi-dimensional signals. This way features like the equatorial anomaly and data gaps can be handled appropriately.

In section 2.4 we demonstrated how the multi-dimensional approaches, i.e. multi-dimensional series expansions in terms of the localizing base functions can be applied to ionospheric signals. The derived multi-dimensional B-spline approaches are based on Cartesian theory and therefore restricted to regional areas. In our first application we modeled the electron density from IRI by a 3-dimensional B-spline approach over Central and South America. From the electron density values we estimated a regional model of the electron density. The estimation of the scaling coefficients was performed by the least squares method. In the second application we simulated GPS observations to model the STEC by a 3-dimensional B-spline approach. Our intention was to estimate corrections to a reference model (IRI) from GPS observations. Here we introduced an additional linear model for the prior information of the coefficients and established a combined Gauss-Markov model. The combined model was solved with the least squares method with an appropriate regularization parameter. The third application treats 3-dimensional COSMIC VTEC observations. Again we estimated a correction term to the reference model (IRI). For this approach we present an application of the data compression algorithm: we performed the data compression for the correction term, i.e. we decomposed the modeled correction term into its low-frequency part and the corresponding detail signal and neglected the non-significant structures from the detail signals. The data compression is an important tool to handle huge amounts of observations.

We have seen that the multi-dimensional B-spline approach yields good results for regional observations; with an appropriate number of observations in a specified region we can estimate the unknown parameters. We achieved satisfying results for modeling signals of the electron density from IRI observations, the STEC from simulated GPS observations and the VTEC from COSMIC observations. The approach can therefore be compared with other techniques such as tomography. To extend our approach to global investigations we can introduce trigonometric B-spline functions (see LYCHE AND SCHUMAKER (2001)) or spherical harmonics (see SCHMIDT ET AL. (2007a)) and work with tensor product functions. Therefore the approach can also be applied to phenomena that have to be investigated globally. In order to take into account the height structure of the electron density we may introduce Chapman functions for the scaling functions with respect to the height (see, e.g., FELTENS (1998)).

Future work may focus on the combination of GNSS, COSMIC and altimetry measurements. This way we want to use information of all these different observation types to model the correction term. With different variance factors we can control the strength of the influence of the different observation types on the corrections.



Figure 4.1: Lowell Digisonde network; see <http://ulcar.uml.edu>.

With the combined model we can improve the corrections, i.e. we take advantage of the various properties (e.g. localization, accuracy) of the different observation types.

A validation of the estimated corrections, may be performed by using terrestrial observations from ionosondes, i.e. we can use electron density profiles derived from ionosondes for a validation. Figure 4.1 shows exemplary the Lowell Digisonde networks. In South America there are, e.g., 7 digisondes.

Appendix A

L^2 -theory of Fourier-series

We adopted the theory from www.mathematik.uni-muenchen.de/~lerdos/WS06/FA.

Theorem A.1. We define $e_n(x) = \frac{e^{inx}}{\sqrt{2\pi}}$, $n \in \mathbb{Z}$.
 $\{e_n\}_{n \in \mathbb{Z}}$ is an ONB in $L^2([0, 2\pi])$.

Then we can define

Definition A.2. For a function $f \in L^2([0, 2\pi])$ we define

$$c_n = \langle e_n, f \rangle_{L^2([0, 2\pi])} = \int_0^{2\pi} \frac{e^{-inx}}{\sqrt{2\pi}} f(x) dx \text{ for } n \in \mathbb{Z} \quad (\text{A.1})$$

these are the Fourier-coefficients of f .

Theorem A.3. Given a function $f \in L^2([0, 2\pi])$. Applying general basis decomposition we get the Fourier series of f

$$f = \sum_{n \in \mathbb{Z}} c_n e_n. \quad (\text{A.2})$$

And Parsevals identity holds

$$\|f\|_{L^2([0, 2\pi])}^2 = \sum_{n \in \mathbb{Z}} |c_n|^2. \quad (\text{A.3})$$

Appendix B

Fourier transform in \mathbb{R}^n

The theory can be found in www.mathematik.uni-muenchen.de/~lerdos/WS06/FA.

Definition B.1. Let $f \in L^1(\mathbb{R}^n)$, then its Fourier-transform is defined as

$$\hat{f}(k) := \frac{1}{2\pi^{n/2}} \int_{\mathbb{R}^n} e^{-ikx} f(x) dx \quad (\text{B.1})$$

In the exponent we have $kx = k \cdot x$ scalar product of two n -vectors, but for simplicity we will just write it as kx .

Definition B.2. The convolution of two functions f and g on \mathbb{R}^n is defined as

$$(f \star g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy. \quad (\text{B.2})$$

Note, it makes only sense for nice f and g (e.g. bounded, fastly decaying) then it holds $f \star g = g \star f$. In general if f, g are in some $L^p(\mathbb{R}^n)$ spaces then $f \star g$ may be meaningless.

Theorem B.3. (i) Let $f, g \in L^1(\mathbb{R}^n)$, then

$$\widehat{(f \star g)} = (2\pi)^{n/2} \hat{f} \hat{g}. \quad (\text{B.3})$$

(ii) [Translation]

$$\widehat{f(\cdot - h)} = e^{-ikh} \hat{f}(k). \quad (\text{B.4})$$

(iii) [Scaling]

$$\widehat{f(\cdot/\lambda)} = \lambda^n \hat{f}(\lambda k). \quad (\text{B.5})$$

Theorem B.4. Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then

- $\hat{f} \in L^2(\mathbb{R})$ and

$$\|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\mathbb{R})} \quad (\text{B.6})$$

also called Plancherel formula, similar to Parseval identity for Fourier-series.

- If $f, g \in L^2(\mathbb{R})$, then

$$\langle f, g \rangle_{L^2(\mathbb{R})} = \langle \hat{f}, \hat{g} \rangle_{L^2(\mathbb{R})}. \quad (\text{B.7})$$

I.e. the Fourier transform is an isometry on $L^2(\mathbb{R})$.

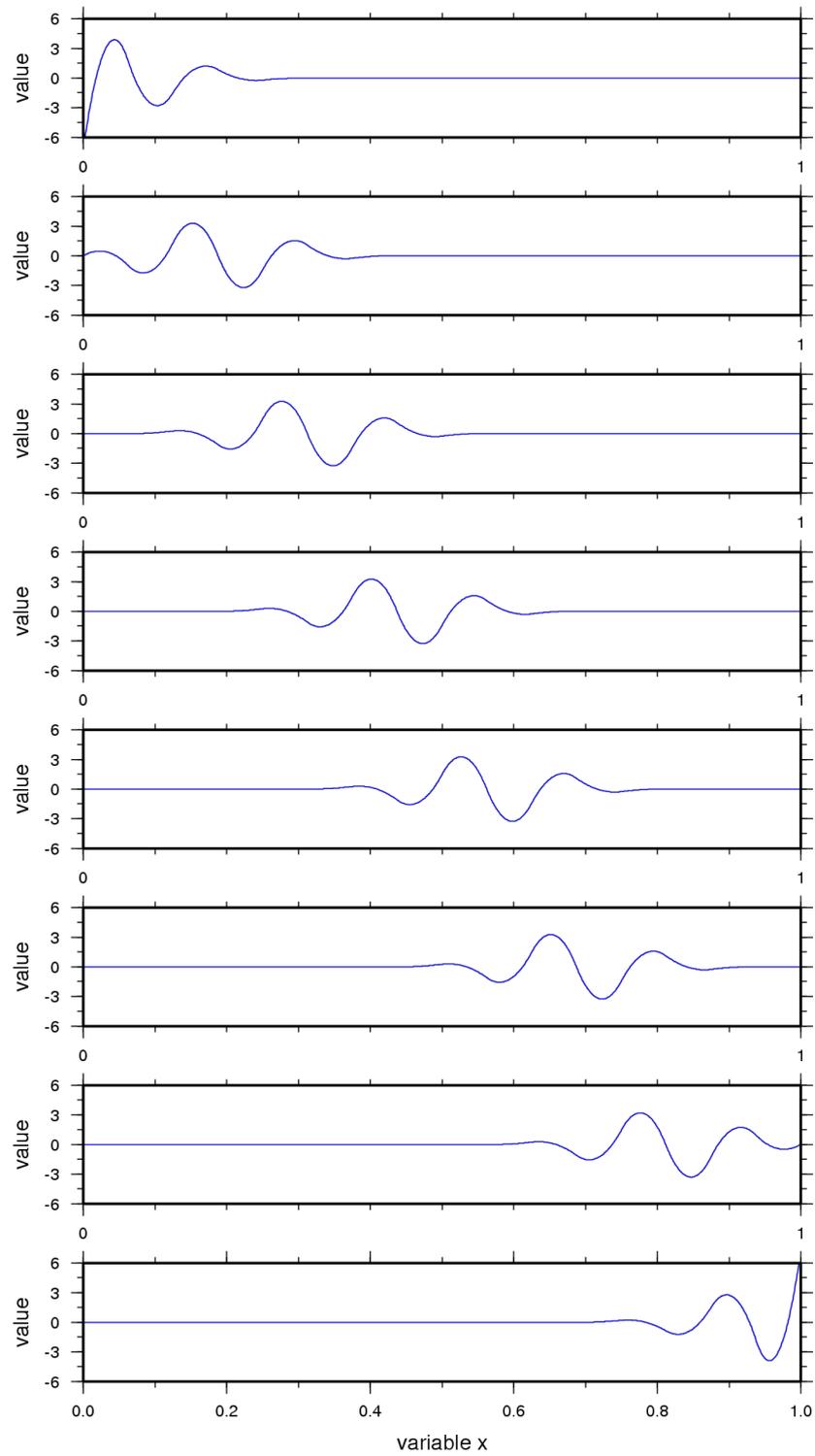


Figure C.1: Wavelets for level $j = 3$. The first and last two are affected by the endpoint interpolation. For the level $j = 3$ we have altogether $n_j = 8$ wavelets.

Appendix D

Statistical quantities

In this section we consider two vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$.

Definition D.1. The mean value \bar{x} of a vector \mathbf{x} is defined as

$$\bar{x} = \frac{x_1 + \dots + x_n}{n}. \quad (\text{D.1})$$

Definition D.2. The root mean square (rms) of a vector \mathbf{x} is defined as

$$\text{rms} = \sqrt{\frac{x_1^2 + \dots + x_n^2}{n}}. \quad (\text{D.2})$$

Definition D.3. The correlation between the two vectors \mathbf{x} and \mathbf{y} , i.e. $\text{Cor}(\mathbf{x}, \mathbf{y})$ is defined via

$$\text{Cor}(\mathbf{x}, \mathbf{y}) = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (\text{D.3})$$

wherein \bar{x} and \bar{y} are the mean values of \mathbf{x} and \mathbf{y} , respectively.

Bibliography

- [ALT (1999)] Alt, H. W. Lineare Funktionalanalysis (3. Auflage), Springer, Berlin.
- [BEYERLE ET AL. (2005)] Beyerle, G., Schmidt, T., Michalak, G., Heise, S., Wickert, J., Reigber, C. GPS radio occultation with GRACE: Atmospheric Profiling utilizing the zero difference technique. *Geophysical Research Letters*, Vol. 32, L13806, doi:10.1029/2005GL023109.
- [BILITZA (2001)] Bilitza, D. International Reference Ionosphere 2000. *Radio Science*, Vol. 36, Nr. 2, 261-275.
- [BRUNINI ET AL. (2003)] Brunini, C., Van Zele, M. A., Meza, A., Gende, M. Quiet and perturbed ionospheric representation according to the electron content from GPS signals. *Journal of Geophysical Research*, Vol. 108, A2, doi:10.1029/2002JA009346.
- [CHUI (1992)] Chui, C. K. *An Introduction to Wavelets*. Academic Press, Boston.
- [CHUI AND QUAK (1992)] Chui, C. K., Quak, E. Wavelets on a Bounded Interval. In: Braess, D., Schumaker L.L. (eds.). *Numerical Methods of Approximation Theory*, Vol. 9, 53-75, Birkhäuser, Basel.
- [DAUBECHIES (1992)] Daubechies, I. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Philadelphia.
- [DETTMERING (2003)] Dettmering, D. Die Nutzung des GPS zur dreidimensionalen Ionosphärenmodellierung. University of Stuttgart, Technical Reports Department of Geodesy and Geoinformatics, Report Nr. 2003.1.
- [FELTENS (1998)] Feltens, J. Chapman Profile Approach for 3-d Global TEC Representation. Proceedings of the 1998 IGS Analysis Centers Workshop, ESOC, Darmstadt, Germany, 285-297.
- [GARCÍA-FERNÁNDEZ (2004)] García-Fernández, M. Contributions to the 3D ionospheric sounding with GPS data. PhD thesis. Department of Applied Mathematics IV and Applied Physics, Technical University of Catalonia, Barcelona, Spain.
- [HERNÁNDEZ-PAJARES ET AL. (1999)] Hernández-Pajares, M., Juan, J. M., Sanz, J. New approaches in global ionospheric determination using ground GPS data. *Journal of Atmospheric and Solar Terrestrial Physics*, 61, 1237-1247.
- [JAWERTH AND SWELDENS (1994)] Jawerth, B., Sweldens, W. An Overview of Wavelet Based Multiresolution Analyses. *SIAM Review*, Vol. 36, No. 3, 377-412.
- [KOCH (1999)] Koch, K.-R. *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer, Berlin.
- [KOCH (2000)] Koch, K.-R. *Einführung in die Bayes-Statistik*. Springer, Berlin.
- [KOCH AND KUSCHE (2001)] Koch, K.-R., Kusche, J. Regularization of geopotential determination from satellite data by variance components. *Journal of Geodesy* 76 (5), 259-746.
- [LIN ET AL. (2007)] Lin, C. H., Liu, J. Y., Fang, T. W., Chang, P. Y., Tsai, H. F., Chen, C. H., Hsiao, C. C. Motions of the equatorial ionization anomaly crests imaged by FORMOSAT-3/COSMIC. *Geophysical Research Letters*, Vol. 34, L19101, doi:10.1029/2007GL030741.
- [LOUIS ET AL. (1997)] Louis, A. K., Maaß, P., Rieder, A. *Wavelets*. John Wiley & Sons Ltd.

- [LYCHE AND SCHUMAKER (2001)] Lyche, T., Schumaker, L. L. A Multiresolution Tensor Spline Method For Fitting Functions on the Sphere. *SIAM Journal of Scientific Computing* 22 (2), 724-746.
- [OGDEN (1997)] Ogden, R. T. *Essential Wavelets for Statistical Applications and Data Analysis*. Birkhäuser, Boston.
- [REED AND SIMON (1980)] Reed, M., Simon, B. *Methods of modern mathematical physics I Functional Analysis*. Academic Press, New York.
- [REIGBER ET AL. (2000)] Reigber, C., Lühr, H., Schwintzer, P. CHAMP mission status and perspectives. Supplement to EOS, Transactions, AGU 81(48), F307.
- [RIUS ET AL. (1997)] Rius, A., Ruffini, G., Cucurull, L. Improving the Vertical Resolution of Ionospheric Tomography with GPS Occultations. *Geophysical Research Letters*, 24 (18), 2291-2294.
- [ROCKEN ET AL. (2000)] Rocken, C., Kuo, Y.-H., Schreiner, W., Hunt, D., Sokolovskiy, S. COSMIC System Description. Special issue of *Terrestrial, Atmospheric and Oceanic Science*, 11 (1), 21-52.
- [SCHMIDT (2001)] Schmidt, M. *Grundprinzipien der Wavelet-Analyse und Anwendungen in der Geodäsie*. Shaker, Aachen.
- [SCHMIDT ET AL. (2007a)] Schmidt, M., Fengler, M., Mayer-Gürr, T., Eicker, A., Kusche, J., Sánchez, L., Han, S.-C. Regional gravity modeling in terms of spherical base functions. *Journal of Geodynamics* 81, 17-38, doi:10.1007/s00190-006-0101-5.
- [SCHMIDT ET AL. (2007b)] Schmidt, M., Bilitza, D., Shum, C. K., Zeilhofer, C. Regional 4-D modeling of the ionospheric electron density. *Journal of Advances in Space Research*, doi:10.1016/j.asr.2007.02.050.
- [SCHUMAKER (1981)] Schumaker, L. L. *Spline Functions Basic Theory*. Wiley.
- [STOLLNITZ ET AL. (1995)] Stollnitz, E. J., DeRose, T. D., Salesin, D. H. Wavelets for Computer Graphics: A Primer Part 2. *IEEE Computer Graphics and Applications*, 15(4), 75-85.
- [TODOROVA ET AL. (2007)] Todorova, S., Schuh, H., Hobinger, T., Hernández-Pajares, M. Global models of the ionosphere obtained by integration of GNSS and satellite altimetry data, *Österreichische Zeitschrift für Vermessung & Geoinformation*, 2, 80-89.
- [TSAI AND TSAI (2004)] Tsai, L.-C., Tsai, W.-H. Improvement of GPS/MET ionospheric profiling and validation using the Chung-Li ionosonde measurements and the IRI model TAO, Vol. 15, Nr. 4, 589-607.
- [WEIDMANN (1976)] Weidmann, J. *Lineare Operatoren in Hilberträumen*. Teubner, Stuttgart.