Reconstruction of Structured Functions from Sparse Fourier Data

vorgelegt von

Marius Wischerhoff

aus Duisburg

Göttingen, 2014

Betreuungsausschuss:

- Prof. Dr. Gerlind Plonka-Hoch, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

- Prof. Dr. Russell Luke, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission:

- Referentin:
  Prof. Dr. Gerlind Plonka-Hoch, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

- Korreferent:
  Prof. Dr. Jürgen Prestin, Institut für Mathematik, Universität zu Lübeck

- Weitere Mitglieder der Prüfungskommission:
  - Prof. Dr. Stephan Huckemann, Institut für Mathematische Stochastik, Georg-August-Universität Göttingen
  - Prof. Dr. Gert Lube, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen
  - Prof. Dr. Russell Luke, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen
  - Prof. Dr. Ingo Witt, Mathematisches Institut, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 14. Januar 2015

"The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with eager feet,
Until it joins some larger way
Where many paths and errands meet.
And wither then? I cannot say."

J. R. R. Tolkien, *The Fellowship of the Ring*, Book I, Chapter 1.

Acknowledgements

Contents

Notation

Number systems and spaces:

$\mathbb{C}$	complex numbers
$\mathbb{C}^n$	space of complex column vectors of length n ($n \in \mathbb{N}$)
$\mathbb{C}^{m \times n}$	space of m-by-n matrices with entries from $\mathbb{C}$, i.e. complex matrices with m rows and n columns ($m, n \in \mathbb{N}$)
$C_0(\mathbb{R}^d)$	space of continuous functions $f : \mathbb{R}^d \to \mathbb{C}$ vanishing at infinity
C_c^∞	see $C_c^\infty(\mathbb{R})$
$C_c^\infty(\mathbb{R}^d)$	space of infinitely differentiable functions $f : \mathbb{R}^d \to \mathbb{C}$ with compact support in $\mathbb{R}^d$
$L^1(\mathbb{R}^d)$	space of absolutely integrable functions $f : \mathbb{R}^d \to \mathbb{C}$
$\mathbb{N}$	natural numbers (excluding 0)
$\mathbb{N}_0$	natural numbers including 0
$\mathbb{R}$	real numbers
$\mathbb{R}^d$	space of real column vectors of length d ($d \in \mathbb{N}$)
$\mathcal{S}$	see $\mathcal{S}(\mathbb{R})$
$\mathcal{S}(\mathbb{R}^d)$	space of infinitely differentiable functions $f : \mathbb{R}^d \to \mathbb{C}$ which are, together with all their derivatives, rapidly decreasing; so-called Schwartz space
$\mathbb{Z}$	integer numbers

Mathematical constants:

π	constant pi
e	Euler's number, base of the natural exponential function
i	imaginary unit

Variables:

$\alpha_\eta^\nu, \lambda, \lambda_\ell$	coefficients
$\alpha_j, \widetilde{\alpha}_j, \beta_j, \widetilde{\beta}_j, \gamma_j$	components of points in the real plane
η, ν	index variables from $\mathbb{N}$
ϑ	parameter for angles that are given in radians
θ, ϱ, ψ	angles
μ	distribution
φ	function
$\xi_1, \xi_2, \omega, \omega_1, \omega_2$	real arguments of functions or components of real vectors
a_j	coefficient or polynomial
$\widetilde{a}_j, a_{\alpha_j}, a_{j,k}, a'_{j,k}, b_j, \widetilde{b}_j,$ $b_{\beta_j}, c_j, c_{\gamma_j}, \widetilde{c}_j, c_j^i,$ $c_{j,k}^{r_1,r_2}, d_m^\nu$	coefficients $(i \in \mathbb{N}_0)$
f, g	functions
h, h_1, h_2	step sizes with $h, h_1, h_2 \in \mathbb{R}$ and $h, h_1, h_2 > 0$
$j, k, k', \ell, m, m_i, n, r, r_i$	index variables from $\mathbb{N}$ $(i = 1, 2)$
l_j	parameterization of L_j or line through the origin in the plane
$l, l^\perp, \widetilde{l}, \widetilde{l}_j$	lines through the origin in the plane
m_j	slope
$\widetilde{n}$	outward pointing unit normal field
t, t_i, x, x_i	real arguments of functions $(i = 1, 2)$
z	complex argument of a function
z_j	zero of a polynomial
Λ, Φ	functions
C, T	real constants
D	polygonal domain
F_k	falling factorial (polynomial)
$G, \widetilde{G}, K, \widetilde{S}_2, V_1, V_2, \widetilde{V}_i$	sets $(i \in \mathbb{N})$
L_j	edge of a polygon
M, N, N_1, N_2, R, R_j	natural numbers
P, Q	functions

S_k	set, knot, or shift value
T_j	frequency in an exponential sum, knot, or shift value

Vectors and matrices:

$\mathbf{0}$	zero vector, i.e. column vector with zeros as entries
$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_d)$	multi-index, i.e. d-tuple of non-negative integers
$\boldsymbol{\lambda}, \boldsymbol{\lambda}', \boldsymbol{\xi}, \boldsymbol{\omega}, \boldsymbol{f}_j, \boldsymbol{t}, \boldsymbol{u}, \boldsymbol{u}^{\perp},$ $\widetilde{\boldsymbol{u}}, \boldsymbol{u}_j, \boldsymbol{x}$	column vectors
$\boldsymbol{e}_k$	unit vector in $\mathbb{R}^d$, $d \in \mathbb{N}$, such that $(\boldsymbol{e}_k)_k = 1$
$\boldsymbol{n}_j, \widetilde{\boldsymbol{n}}_j$	normal vector / unit normal vector (column vectors)
$\boldsymbol{p}, \boldsymbol{p}_{\text{new}}, \boldsymbol{p}_{l,\min}, \boldsymbol{p}_{l,\max}$	points in the real plane or in $\mathbb{R}^d$ (column vectors)
$\boldsymbol{v}, \boldsymbol{v}_j, \widetilde{\boldsymbol{v}}_j, \boldsymbol{w}$	shift vector or vertex of a polygon (column vectors)
$\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{G}, \boldsymbol{U}$	matrices
$\boldsymbol{D}_\varrho, \boldsymbol{D}_\psi$	rotation matrices
$\boldsymbol{H}_{N+1}, \boldsymbol{T}_{N+1}$	$(N+1)$-by-$(N+1)$ Hankel / Toeplitz matrix
$\boldsymbol{V}_{N,N+1}$	N-by-$(N+1)$ Vandermonde-type matrix

Functions and operators:

$\mathbf{1}_{[a,b)}$	characteristic function on the interval $[a, b)$
$\langle \boldsymbol{\omega}, \boldsymbol{x} \rangle$	standard inner product of real vectors $\boldsymbol{\omega}$ and $\boldsymbol{x}$: $\langle \boldsymbol{\omega}, \boldsymbol{x} \rangle = \boldsymbol{\omega}^{\mathrm{T}} \boldsymbol{x}$
$\langle \mu, \varphi \rangle$	evaluation of distribution μ with argument φ
$|\boldsymbol{\alpha}|$	the sum $\sum_{j=1}^{d} \alpha_j$ for the multi-index $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_d)$
$|x|, |z|$	absolute value of the real number x / the complex number z
$|S|$	cardinality of the set S
$\|\boldsymbol{\xi}\|_2$	euclidean norm of the vector $\boldsymbol{\xi}$
$\boldsymbol{\omega}^{\boldsymbol{\alpha}}$	the product $\prod_{j=1}^{d} \omega_j^{\alpha_j}$ for $\boldsymbol{\omega} \in \mathbb{R}^d$ and the multi-index $\boldsymbol{\alpha}$
f'	first derivative of the univariate function f
$f^{(k)}$	kth derivative of the univariate function f
$\widehat{f}$	Fourier transform of the function f
$\overline{z}$	complex conjugate of z (z: number, vector, matrix, function)
$\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{A}^{\mathrm{T}}$	transpose of the vector $\boldsymbol{x}$ and the matrix $\boldsymbol{A}$ respectively
$\boldsymbol{A}^*, \boldsymbol{A}^{-1}$	conjugate transpose / inverse of the matrix $\boldsymbol{A}$

$\frac{\partial^m}{\partial\omega^m}g$	higher-order partial derivative of g with respect to ω
∂D	boundary of the domain D
δ	Dirac delta distribution
$\tau_a, \tau_{a,b}$	translation operators with $a \in \mathbb{R}$ and $(a,b) \in \mathbb{R}^2$ respectively
$\mu_{(g)}$	distribution, identified with the function g
B_j^m	B-spline of order m determined by the knots $T_j, \dots, T_{j+m}$
$\mathrm{div}_{\boldsymbol{x}}$	divergence operator (with respect to $\boldsymbol{x}$)
D_ν	function, see p. 37
$\mathrm{D}\,\mu$	distributional derivative of μ
$\mathrm{D}^k\,\mu$	kth distributional derivative of μ
$\mathrm{D}^{(r_1,r_2)}\,\mu$	distributional derivative of μ ((r_1,r_2): multi-index)
$\mathrm{D}^{\boldsymbol{\alpha}}\,f$	mixed-partial derivative of the function $f : \mathbb{R}^d \to \mathbb{C}$ (p. 140)
$\mathrm{diag}(d_1, \dots, d_N)$	diagonal matrix with $d_1, d_2, \dots, d_N$ as main diagonal entries
$\exp$	natural exponential function
$\mathcal{F}\{f\}, \mathcal{F}\{\mu\}$	Fourier transform of the function f / the distribution μ
$\mathrm{grad}_{\boldsymbol{x}}$	gradient operator (with respect to $\boldsymbol{x}$)
$\mathrm{Im}(z)$	imaginary part of the complex number z
$\ln$	principal branch of the complex logarithm
$\min\{a,b\}$	minimum of the two expressions a and b
N_m	centred cardinal B-Spline
$\mathrm{P}_{l_k}(\boldsymbol{p})$	orthogonal projection of the point $\boldsymbol{p}$ onto the line l_k
$\mathrm{rank}(\boldsymbol{A})$	rank of the matrix $\boldsymbol{A}$
$\sin, \cos, \tan, \cot$	trigonometric functions sine, cosine, tangent, and cotangent

Miscellanea:

$\cdot, \cdot_1, \cdot_2$	place-holder for a variable, as in $f(\cdot)$ or $f(\cdot_1, \cdot_2)$, for example
$(T_j)_{j=1}^N$	column vector $(T_1, T_2, \dots, T_N)^{\mathrm{T}}$ or sequence $(T_1, T_2, \dots, T_N)$
$(a_{j,k})_{j=n_1,k=n_2}^N$	$(N - n_1 + 1)$-by-$(N - n_2 + 1)$ matrix with entries $a_{j,k}$
$(\boldsymbol{x}_j)_k, x_{j,k}$	notations for the kth component of a vector $\boldsymbol{x}_j$ ($k \in \mathbb{N}$)
$c_j^*, v_{j,k}^*, T_j^*$	numerically reconstructed values

List of Figures

List of Tables

1. Introduction

In several scientific areas, such as radio astronomy, computed tomography, and magnetic resonance imaging, [9], the reconstruction of structured functions from the knowledge of samples of their Fourier transform is a common problem. For the analysis of the examined object, it is important to reconstruct the underlying original signal as exactly as possible. In this dissertation, we aim to uniquely recover specially structured functions from a smallest possible set of Fourier data.

Frequently, special properties or structures of the functions to be reconstructed are exploited in order to reconstruct the functions from only a small set of sampling values. These can be function values or samples of a transformed version of the function, such as the Fourier transform.

Normally, one seeks to recover functions which possess sparse representations in a given basis or frame using only a small set of sampling values. In particular, there has been significant research activity in the area of *Compressed Sensing*, [18].

For instance, Candès et al. have shown in [14] that a discrete-time signal $f \in \mathbb{C}^n$ with $s \ll n$ non-zero entries can be recovered, under certain conditions, from a set of Fourier coefficients of size $\mathcal{O}(s \log(n))$. In [59], it has been shown that a trigonometric polynomial of degree N where only $s \ll N$ coefficients are non-zero can be reconstructed with probability at least $1 - \varepsilon$ from $\mathcal{O}(s \log((2N + 1)/\varepsilon))$ randomly chosen sampling points.

Usually, the reconstruction algorithms used in the framework of compressed sensing are based on l_1-minimization methods, and exact recovery is only attained with a certain probability.

In contrast to those minimization methods where the result is only guaranteed with a certain probability, also deterministic methods for the recovery of sparse trigonometric functions are available, which are based upon the *Prony method*, [33, pp. 457–462], or the *annihilating filter method*, [19, 65].

Within the past few years, the Prony method has been increasingly applied in the field of parameter estimation. It enables the determination of specific parameters of functions using only few sampling values if the functions to be recovered belong to the class of exponential sums.

Several numerically stable variants of the original Prony method, which is over 200 years old, [17], and suffers from numerical instabilities, have been derived. Potts

and Tasche have developed the *Approximate Prony method* in [55, 56], which works not only in the noiseless case but also in the case of noisy measurements. Another stabilization of the Prony method is proposed in [23], where the possibly perturbed sampling values are not used directly, but a windowed average of their autocorrelation sequence is used instead. This approach is motivated by the application of operators of the form $\sum_{k=0}^{N} g\left(\frac{k}{N}\right) \widehat{f}(k) \exp(\mathrm{i}\,k\cdot)$ with a suitable filter g in [45] and [46], where such operators have been applied in order to detect singularities of piecewise smooth functions f.

Recently, the reconstruction of piecewise smooth functions and shifts of several signals has also been studied in [4–6]. In these papers, the robustness of the involved Prony type systems is investigated too, see also [3].

There are also well-known parameter identification methods in signal processing, such as the *ESPRIT method*, [61], the *Matrix pencil method*, [36], and the *MUSIC method*, [62]. In [58], it has been pointed out that these methods are equivalent to the Prony method such that they can be seen as so-called *Prony-like methods*, see [48, Chapter 3].

The Prony method has also been generalized to the problem of reconstructing sparse Legendre expansions, [50], and, as the so-called *generalized Prony method*, to the problem of reconstructing sparse sums of eigenfunctions of linear operators, see [49] and the dissertation [48] of Thomas Peter.

Further, the Prony method is equivalent to the annihilating filter method. In [65], Vetterli et al. have introduced the concept of signals with finite rate of innovation. Such signals have a finite number of degrees of freedom per unit of time. Using the annihilating filter method, Vetterli at al. have shown that one can completely reconstruct signals of finite length which have a finite rate of innovation by applying a generalized Shannon sampling theorem although the signals considered are not band-limited.

In the dissertation on hand, we use the Prony method as key instrument in order to uniquely reconstruct structured, real-valued functions from sparse Fourier samples. We derive algorithms for unique recovery by means of a smallest possible set of Fourier data. First, we consider the univariate case, where we propose reconstruction methods for B-spline functions with non-uniform knots and linear combinations of non-uniform translates of a known low-pass filter function.

We can transfer the reconstruction results given in the univariate case to the bivariate case in a similar way if we consider separable functions such as tensor-products of non-uniform spline functions and non-uniform translates.

In case of non-separable functions, the problem gets more involved. We want to recover the unknown functions from only a small amount of Fourier samples taken on few distinct lines through the origin. In [44], a linear combination of N bivariate Diracs is recovered from $\mathcal{O}(N^2)$ samples of the signal by considering the Fourier series

coefficients of the signal and using the so-called *algebraic coupling of matrix pencils (ACMP) algorithm*, [64].

We will study linear combinations of non-uniform shifts of bivariate functions Φ of the form $f(x_1, x_2) := \sum_{j=1}^{N} c_j \Phi(x_1 - v_{j,1}, x_2 - v_{j,2})$ where we seek to recover the unknown parameters c_j, $v_{j,1}$, and $v_{j,2}$ from sparse Fourier data. As we will show, it is possible to uniquely recover these parameters by using only $3N+1$ Fourier samples on three lines through the origin. We consider two predetermined lines, while the third sampling line is chosen dependently on the results obtained by employing the samples from the first two lines. Moreover, if we want to use only predetermined sampling lines, we will show that $N+1$ lines and $\mathcal{O}(N^2)$ Fourier samples are required to ensure unique recovery.

Further, we propose a generalization to d dimensions with $d > 2$, where we need $d+1$ sampling lines for unique function reconstruction. Similarly to the bivariate case, we use d predetermined lines, and the $(d+1)$st line is chosen suitably.

Moreover, we examine the reconstruction of polygonal shapes in the real plane. In the *shape from moments problem*, [21, 27, 47], integral moments of the analytic power function $f(z) = z^k$ over the characteristic function of a polygonal domain D are taken, and the vertices of D are computed by applying Prony-like methods to these moments. We, on the other hand, will take a similar approach as in the case of bivariate non-uniform translates such that we are able to reconstruct concave polygonal domains D with N vertices by taking $3N$ samples of the Fourier transform of the characteristic function $\mathbf{1}_D$. Here, we use again two predetermined sampling lines and an appropriately chosen third line.

This dissertation is organized as follows: Chapter 2 provides an overview of the Prony method in such a formulation as we need later on for our proposed reconstruction methods in Chapters 3–6, where the focus lies on using as few Fourier samples as possible in order to uniquely recover functions of different classes. In Chapter 3, we discuss the one-dimensional case of reconstruction of structured functions from sparse Fourier data, that is, we consider univariate functions such as step functions, non-uniform spline functions, and non-uniform translates of low-pass filter functions. We transfer the results found in Chapter 3 to the two-dimensional setting in Chapter 4, where we deal with the reconstruction of tensor-products of non-uniform spline functions and non-uniform translates from sparse Fourier data. Further, we also consider non-uniform translates of bivariate functions where the variables are not separable. The results for the last mentioned case are then generalized to d dimensions with $d > 2$ in Chapter 5. Finally, in Chapter 6, we discuss the reconstruction of polygonal shapes in the real plane from sparse Fourier data. Some parts of this work have already been published in [54] and [67].

In several cases, we illustrate our reconstruction results with numerical experiments. All examples considered in the sections about numerical results have been computed

using MATLAB 7.11 with double precision arithmetic. For the computation of eigen-vectors, roots of polynomials, and least squares solutions to linear systems of equations, internal MATLAB routines have been used.

At the end, we collect some properties of the Fourier transform, which we use throughout this work, in Appendix A. Further, Appendix B contains the proofs of some lemmas for non-uniform spline representations which are needed for the recon-struction of non-uniform spline functions in Section 3.2. Some required lemmas for integration, needed in Appendix B, are given in Appendix C. The last chapter, Ap-pendix D, consists of the proofs of lemmas required for non-uniform tensor-product spline representations in Subsection 4.1.1.

2. Prony method

The Prony method is a method for the identification of parameters in exponential sums, that is, the Prony method can be used in order to determine frequencies T_j and corresponding coefficients c_j of a function f of the form

$$f(\omega) = c_1 \, e^{\omega T_1} + c_2 \, e^{\omega T_2} + \ldots + c_N \, e^{\omega T_N}$$

where function values of f are given on a set of at least $2N$ equispaced points, e.g. see [33, pp. 457–462]. This method dates back to 1795, when Gaspard Riche de Prony introduced it in [17] for solving an approximation problem about expansion characteristics of gases.

The Prony method consists of two separate steps. In the first step, the frequencies T_j are computed. Afterwards, using the knowledge of the found frequencies, the corresponding coefficients are determined. We employ the Prony method as a key instrument for our proposed approaches for the reconstruction of structured functions from sparse Fourier data in Chapters 3–6.

In the following two sections, we give an overview of the classical Prony method, which works in the case of exact data. We use a slightly different notation and formulation of the problem than in the literature about the Prony method and its variants from the past few years, e.g. [8, 23, 48, 51, 55–58], such that we can directly apply the method as described in Sections 2.1 and 2.2 to the problems dealt with in Chapters 3–6, that is, we already adjust the formulation of the Prony method in the following two sections to the settings considered later.

In the first of the two steps of the Prony method, an appropriate eigenvalue problem is solved, which yields the coefficients of a so-called *Prony polynomial*. By computing the roots of the Prony polynomial, we then obtain the frequencies T_j. In the second step, a linear system of equations is solved in order to determine the corresponding coefficients c_j.

The original Prony method is numerically unstable with respect to inexact measurements of the function values, see [42, pp. 275–280] for an example. Recently there have been made various efforts in order to derive stable variants of the Prony method such that it works also for noisy input data. A stable algorithm, the *Approximate Prony method*, has been derived in [51, 55, 56], based upon an approach from [8].

Moreover, as Potts and Tasche have pointed out in [58], there exist several *Prony-like methods*, that is, methods which solve the same kind of parameter estimation

problem as the original Prony method, such as the *ESPRIT method*, [61], and the *Matrix pencil method*, [36]. In [48], it is shown that also the *MUSIC method*, [62], is a Prony-like method. A nice overview and comparison of these methods can be found in [48, Chapter 3].

In [23], Filbir et al. have proposed a stabilization of the Prony method where, instead of using the (perturbed) function samples, a windowed average of their autocorrelation sequence is applied.

The performance of the Prony-like methods can be greatly improved if a larger number of sampling values of the function that has to be reconstructed is available, see [23, 51, 55], for example, where also error estimates in the case of noisy measurements are given. In order to obtain certain error estimates, the number M of measurements has to fulfil the condition $M > \frac{2\pi^2}{q}$ where q is the minimal separation distance between two frequencies. But this is only a theoretical result. Actually, the required number of function values is often considerably smaller, see [48, Section 3.4]. It is also concluded that the different Prony-like methods yield comparable reconstruction results, and that stable variants of the Prony method are given by them, see [48, Chapter 3].

Further, note that the Prony method is also equivalent to the *annihilating filter method* used in signal processing, see [19, 65], for instance. The relation between these two methods is shown in [53, Remark 2.4].

In the following, we concentrate on the classical Prony method for the noiseless case in order to understand the concept of this method, and we present it in an adjusted formulation such that we can apply the method as described in the following two sections to our reconstruction problems considered later.

2.1. Prony method for exponential sums with complex-valued coefficients

Consider a trigonometric function $P : \mathbb{R} \to \mathbb{C}$ of the special form

$$P(\omega) = \sum_{j=1}^{N} c_j \, \mathrm{e}^{-\mathrm{i}\,\omega T_j} \tag{2.1}$$

with $N \in \mathbb{N}$, non-zero coefficients $c_j \in \mathbb{C}$, and real-valued frequencies T_j in ascending order, i.e. $-\infty < T_1 < T_2 < \ldots < T_N < \infty$.

We want to compute the frequencies $T_1, \ldots, T_N$, and all coefficients $c_1, \ldots, c_N$ from the sampling values $P(\ell h)$ for $\ell = 0, \ldots, 2N$ where h is assumed to be a positive constant with $hT_j \in (-\pi, \pi]$ for all $j \in \{1, \ldots, N\}$. For this purpose, the Prony method can be applied as follows.

Let us consider the complex polynomial $\Lambda : \mathbb{C} \to \mathbb{C}$ defined by

$$\Lambda(z) := \prod_{j=1}^{N} \left(z - \mathrm{e}^{-\mathrm{i}\,hT_j} \right) = \sum_{k=0}^{N} \lambda_k z^k. \tag{2.2}$$

This polynomial, the so-called *Prony polynomial*, exactly possesses the values $\mathrm{e}^{-\mathrm{i}\,hT_j}$, $j = 1, \ldots, N$, with the unknown frequencies T_j from (2.1) as zeros. Here, λ_k are the coefficients of Λ in the monomial basis. Particularly, we have $\lambda_N = 1$ by definition of the Prony polynomial Λ.

Then we observe for $m = 0, \ldots, N$ that

$$
\begin{aligned}
\sum_{k=0}^{N} \lambda_k P\big(h(k+m)\big) &= \sum_{k=0}^{N} \lambda_k \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\,h(k+m)T_j} = \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\,hmT_j} \sum_{k=0}^{N} \lambda_k\, \mathrm{e}^{-\mathrm{i}\,hkT_j} \\
&= \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\,hmT_j}\, \Lambda(\mathrm{e}^{-\mathrm{i}\,hT_j}) \overset{(2.2)}{=} 0.
\end{aligned} \tag{2.3}
$$

Hence, the coefficient vector $\boldsymbol{\lambda} := (\lambda_0, \ldots, \lambda_N)^{\mathrm{T}}$ with $\lambda_N = 1$ is the solution of the linear system

$$\boldsymbol{H}_{N+1}\boldsymbol{\lambda} = \boldsymbol{0} \tag{2.4}$$

with the Hankel matrix[1]

$$\boldsymbol{H}_{N+1} := \Big(P\big(h(k+m)\big) \Big)_{m,k=0}^{N} \in \mathbb{C}^{(N+1)\times(N+1)},$$

which is given by the known sampling values $P(\ell h)$, $\ell = 0, \ldots, 2N$.

We can write this Hankel matrix as a product of two Vandermonde-type matrices and a diagonal matrix, which enables us to compute the rank of $\boldsymbol{H}_{N+1}$. We find

$$
\begin{aligned}
\boldsymbol{H}_{N+1} &= \Big(P\big(h(k+m)\big) \Big)_{m,k=0}^{N} = \left(\sum_{j=1}^{N} c_j \exp(-\mathrm{i}\,h(k+m)T_j) \right)_{m,k=0}^{N} \\
&= \left(\sum_{j=1}^{N} c_j \exp(-\mathrm{i}\,hmT_j) \cdot \exp(-\mathrm{i}\,hkT_j) \right)_{m,k=0}^{N} \\
&= \big(c_j \exp(-\mathrm{i}\,hmT_j) \big)_{m=0,j=1}^{N} \cdot \big(\exp(-\mathrm{i}\,hkT_j) \big)_{j=1,k=0}^{N} \\
&= \big(\exp(-\mathrm{i}\,hmT_j) \big)_{m=0,j=1}^{N} \cdot \mathrm{diag}(c_1, c_2, \ldots, c_N) \cdot \big(\exp(-\mathrm{i}\,hkT_j) \big)_{j=1,k=0}^{N} \\
&= \boldsymbol{V}_{N,N+1}^{\mathrm{T}} \cdot \boldsymbol{D} \cdot \boldsymbol{V}_{N,N+1}
\end{aligned}
$$

---

[1]A *Hankel matrix* is a matrix where the entries are constant along the diagonals parallel to the antidiagonal, see [34, pp. 27–28].

where $\boldsymbol{V}_{N,N+1}$ is a Vandermonde-type matrix and $\boldsymbol{D}$ a diagonal matrix with

$$\boldsymbol{V}_{N,N+1} := \left(\exp(-\,\mathrm{i}\,hkT_j)\right)_{j=1,k=0}^{N} \quad \text{and} \quad \boldsymbol{D} := \mathrm{diag}(c_1, c_2, \ldots, c_N).$$

Further, we know

$$\mathrm{rank}\!\left(\boldsymbol{V}_{N,N+1}\right) = \mathrm{rank}\!\left(\boldsymbol{V}_{N,N+1}^{\mathrm{T}}\right) = \mathrm{rank}(\boldsymbol{D}) = N \tag{2.5}$$

since $\boldsymbol{V}_{N,N+1}$ and $\boldsymbol{V}_{N,N+1}^{\mathrm{T}}$ are Vandermonde-type matrices where $T_j \neq T_k$ for $j \neq k$, and since $\boldsymbol{D}$ is a diagonal matrix with non-zero main diagonal entries[2].

In order to compute the rank of $\boldsymbol{H}_{N+1}$, we need some rank properties. The following lemma collects rank properties which are taken from [34, p. 13].

Lemma 2.1.

(a) *For a matrix $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ and a non-singular matrix $\boldsymbol{B} \in \mathbb{C}^{n \times n}$, one has the rank equality*

$$\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{AB}). \tag{2.6}$$

(b) *Let the matrices $\boldsymbol{A} \in \mathbb{C}^{m \times k}$ and $\boldsymbol{B} \in \mathbb{C}^{k \times n}$ be given. Then one has the rank inequality*

$$\left(\mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{B})\right) - k \leq \mathrm{rank}(\boldsymbol{A} \cdot \boldsymbol{B}) \leq \min\{\mathrm{rank}(\boldsymbol{A}), \mathrm{rank}(\boldsymbol{B})\}. \tag{2.7}$$

Using Lemma 2.1, we can now compute the rank of $\boldsymbol{H}_{N+1}$. Remember that we have

$$\boldsymbol{V}_{N,N+1}^{\mathrm{T}} \in \mathbb{C}^{N+1 \times N} \text{ and } \boldsymbol{D} \in \mathbb{C}^{N \times N} \text{ with } \mathrm{rank}(\boldsymbol{D}) \overset{(2.5)}{=} N.$$

First, we observe that

$$\mathrm{rank}\!\left(\boldsymbol{V}_{N,N+1}^{\mathrm{T}} \cdot \boldsymbol{D}\right) \overset{(2.6)}{=} \mathrm{rank}\!\left(\boldsymbol{V}_{N,N+1}^{\mathrm{T}}\right) \overset{(2.5)}{=} N. \tag{2.8}$$

Secondly, define $\boldsymbol{A} := \boldsymbol{V}_{N,N+1}^{\mathrm{T}} \cdot \boldsymbol{D} \in \mathbb{C}^{N+1 \times N}$ and $\boldsymbol{B} := \boldsymbol{V}_{N,N+1} \in \mathbb{C}^{N \times N+1}$. Then, using the equalities (2.5) and (2.8), we get the following by (2.7):

$$\mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{B}) - N \leq \mathrm{rank}(\boldsymbol{A} \cdot \boldsymbol{B}) \leq \min\{\mathrm{rank}(\boldsymbol{A}), \mathrm{rank}(\boldsymbol{B})\}$$
$$\Leftrightarrow \qquad N + N - N \leq \mathrm{rank}(\boldsymbol{A} \cdot \boldsymbol{B}) \leq \min\{N, N\}$$
$$\Leftrightarrow \qquad N \leq \mathrm{rank}(\boldsymbol{A} \cdot \boldsymbol{B}) \leq N.$$

Therefore, we have

$$\mathrm{rank}(\boldsymbol{A} \cdot \boldsymbol{B}) = \mathrm{rank}\!\left(\boldsymbol{V}_{N,N+1}^{\mathrm{T}} \cdot \boldsymbol{D} \cdot \boldsymbol{V}_{N,N+1}\right) = \mathrm{rank}\!\left(\boldsymbol{H}_{N+1}\right) = N.$$

---

[2]See [34, p. 29] and [34, p. 23] respectively.

Since $\mathrm{rank}(\boldsymbol{H}_{N+1}) = N$, the dimension of the eigenspace of $\boldsymbol{H}_{N+1}$ corresponding to the eigenvalue 0 is equal to one[3]. Thus, the eigenvector $\boldsymbol{\lambda}$ of $\boldsymbol{H}_{N+1}$ corresponding to the eigenvalue 0 is uniquely determined by (2.4) and $\lambda_N = 1$.

Knowing $\boldsymbol{\lambda}$, which is the coefficient vector of the polynomial Λ in (2.2), we can compute the zeros $z_j := \mathrm{e}^{-\mathrm{i}hT_j}$, $j = 1, \ldots, N$, of that polynomial Λ. Hence, using the assumptions about h (see also Remarks 2.3, 1.), we get the frequencies $T_1, \ldots, T_N$ by the computation

$$T_j = \frac{-\mathrm{Im}\big(\ln(z_j)\big)}{h}, \quad j = 1, \ldots, N,$$

where we consider the principal branch of the complex logarithm[4], i.e. $\ln(z) = \mathrm{i}\,\theta$ for $z = \mathrm{e}^{\mathrm{i}\theta}$ with $\theta \in (-\pi, \pi]$.

Finally, the coefficients c_j, $j = 1, \ldots, N$, are obtained from the linear, overdetermined Vandermonde-type system

$$P(\ell h) = \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\ell hT_j}, \quad \ell = 0, \ldots, 2N. \tag{2.9}$$

We summarize the algorithm to determine the frequencies and the coefficients in (2.1) as follows.

Algorithm 2.2 (Prony method for complex-valued coefficients).

- **Input:**
    - $P(\ell h)$, $\ell = 0, \ldots, 2N$;
    - step size h with $hT_j \in (-\pi, \pi]$ for all $j \in \{1, \ldots, N\}$.

- **Computation:**
    1. Construct the Hankel matrix
    $$\boldsymbol{H}_{N+1} := \Big(P\big(h(\ell - m)\big)\Big)_{m,\ell=0}^{N}.$$

    2. Solve the system $\boldsymbol{H}_{N+1}\boldsymbol{\lambda} = \boldsymbol{0}$ where $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_N)^{\mathrm{T}}$ with $\lambda_N = 1$.

    3. Consider the polynomial
    $$\Lambda(z) := \sum_{\ell=0}^{N} \lambda_\ell z^\ell,$$
    and compute all its zeros $z_j := \mathrm{e}^{-\mathrm{i}hT_j}$, $j = 1, \ldots, N$.

---

[3] Cf. [24, p. 317].
[4] See [68, p. 559], for example.

4. For $T_j = \frac{-\operatorname{Im}(\ln(z_j))}{h}$, $j = 1, \ldots, N$, compute the corresponding coefficients c_j, $j = 1, \ldots, N$, as least squares solution to the Vandermonde-type system

$$\sum_{j=1}^{N} c_j \, \mathrm{e}^{-\mathrm{i}\,\ell h T_j} = P(\ell h), \quad \ell = 0, \ldots, 2N.$$

(Here, ln is the principal branch of the complex logarithm.)

- **Output:** Frequencies T_j, $j = 1, \ldots, N$, and coefficients c_j, $j = 1, \ldots, N$, determining P in (2.1).

Remarks 2.3.

1. In order to compute the frequencies T_j for $j = 1, \ldots, N$ uniquely, we need to ensure that $hT_j \in (-\pi, \pi]$ since the function $\omega \mapsto \mathrm{e}^{-\mathrm{i}\omega}$ is 2π-periodic. Otherwise, we will not be able to extract the values T_j from the zeros $z_j = \mathrm{e}^{-\mathrm{i}hT_j}$ of Λ on the unit circle uniquely.

2. While the frequencies T_j are not known, we only need to find a suitable upper bound for $|T_j|$ in order to fix a suitable step size h.

3. In applications, also the number N of terms in (2.1) is usually unknown. Having given at least an upper bound $M \geq N$ and $2M + 1$ sampling values $P(\ell h)$ for $\ell = 0, \ldots, 2M$, we can also apply the above procedure (replacing N by M) and obtain N by examining the rank of $\boldsymbol{H}_{M+1}$ numerically. In this case, (2.4) cannot longer be solved uniquely, but each eigenvector corresponding to the eigenvalue 0 will serve for the determination of the zeros of Λ on the unit circle and hence of T_j, $j = 1, \ldots, N$, see [55], for example.

4. The sampling values taken here are the values $P(\ell h)$ for $\ell = 0, \ldots, 2N$. This is not the only possible choice. Observe that the equations in (2.3) are valid for all integers m. Thus, we can take also the $2N + 1$ sampling values $P(\ell h)$, $\ell = k, \ldots, 2N + k$ for some $k \in \mathbb{Z}$.

Remark 2.4.
Observe that we consider an eigenvalue problem in (2.4); i.e., we want to compute the eigenvector $\boldsymbol{\lambda}$ of the Hankel matrix $\boldsymbol{H}_{N+1}$ which corresponds to the eigenvalue 0. We know that this eigenvector is uniquely determined since we have $\operatorname{rank}(\boldsymbol{H}_{N+1}) = N$ and $\lambda_N = 1$. Moreover, this eigenvalue problem determines the number of sampling values of the trigonometric function P we need because the Hankel matrix $\boldsymbol{H}_{N+1}$ contains the function values $P(\ell h)$ for $\ell = 0, 1, \ldots, 2N$ as entries. But by taking a closer look at the system (2.4), we notice that we can change this eigenvalue problem into a linear

system of equations where we have to find N unknowns such that we need only $2N$ function values of P.

First, remember that the vector $\boldsymbol{\lambda}$ is the coefficient vector of the polynomial Λ in (2.2), and that we have $\lambda_N = 1$ due to the definition of Λ. Thus, we actually have N unknowns in the linear system (2.3), namely $\lambda_0, \ldots, \lambda_{N-1}$. By rearranging the equations in (2.3), we obtain the linear system

$$\sum_{k=0}^{N-1} \lambda_k P\big(h(k+m)\big) = -\lambda_N P\big(h(N+m)\big) = -P\big(h(N+m)\big), \quad m \in \mathbb{Z}.$$

Since we have to determine the N unknown parameters $\lambda_0, \ldots, \lambda_{N-1}$, it suffices to take N equations, e.g. the equations for $m = 0, \ldots, N-1$. Therefore, we need the function values $P(\ell h)$, $\ell = 0, 1, \ldots, 2N-1$, such that $2N$ sampling values of P are enough in order to determine the frequencies and coefficients of the trigonometric function P in (2.1).

2.2. Prony method for exponential sums with real-valued coefficients

Let us now consider an exponential sum where the coefficients are real numbers, i.e. a trigonometric function $P : \mathbb{R} \to \mathbb{C}$ of the form

$$P(\omega) = \sum_{j=1}^{N} c_j \, e^{-i\,\omega T_j} \tag{2.10}$$

with $N \in \mathbb{N}$, non-zero coefficients $c_j \in \mathbb{R}$, and real-valued frequencies T_j in ascending order, i.e. $-\infty < T_1 < T_2 < \ldots < T_N < \infty$.

In the present case, observe that we do not have complex-valued coefficients as in the exponential sum considered in (2.1) in the previous section but real-valued coefficients. Therefore, the function P in (2.10) has a conjugate symmetry, i.e.

$$P(-\omega) = \sum_{j=1}^{N} c_j \, e^{i\,\omega T_j} = \overline{P(\omega)}. \tag{2.11}$$

This enables us to identify the parameters of P, namely the frequencies T_j and the coefficients c_j ($j = 1, \ldots, N$), using fewer sampling values than in the case where the coefficients are complex-valued.

We still need $2N + 1$ function values of P for the actual computations. While we have taken the sampling values $P(\ell h)$ for $\ell = 0, \ldots, 2N$ in the previous section, we can also take the function values $P(\ell h)$ for $\ell = -N, \ldots, -1, 0, 1, \ldots, N$ as explained

in Remarks 2.3, 4. By using the conjugate symmetry (2.11), we can compute the values $P(\ell h)$ for $\ell = -N, \ldots, -1$ from the values $P(0), \ldots, P(Nh)$ such that the $N+1$ sampling values $P(\ell h)$ for $\ell = 0, \ldots, N$ suffice for the determination of the parameters of the exponential sum P.

Here, h is again assumed to be a positive constant satisfying $hT_j \in (-\pi, \pi]$ for all $j \in \{1, \ldots, N\}$. In detail, the Prony method applied to exponential sums with real-valued coefficients works as follows.

First, we again consider the Prony polynomial $\Lambda : \mathbb{C} \to \mathbb{C}$ defined in (2.2), i.e.

$$\Lambda(z) := \prod_{j=1}^{N} \left(z - \mathrm{e}^{-\mathrm{i}hT_j} \right) = \sum_{\ell=0}^{N} \lambda_\ell z^\ell, \tag{2.12}$$

which possesses the values $\mathrm{e}^{-\mathrm{i}hT_j}$ for $j = 1, \ldots, N$ with the unknown frequencies T_j from (2.10) as zeros. The leading coefficient of Λ in the monomial representation on the right-hand side of (2.12) is given as $\lambda_N = 1$ by definition of Λ.

As already mentioned, the equations in (2.3) hold for all integers m. Thus, we obtain the following for $m = 0, \ldots, N$:

$$\sum_{\ell=0}^{N} \lambda_\ell P\big(h(\ell - m)\big) = \sum_{\ell=0}^{N} \lambda_\ell \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}h(\ell-m)T_j} = \sum_{j=1}^{N} c_j\, \mathrm{e}^{\mathrm{i}hmT_j} \sum_{\ell=0}^{N} \lambda_\ell\, \mathrm{e}^{-\mathrm{i}h\ell T_j}$$

$$= \sum_{j=1}^{N} c_j\, \mathrm{e}^{\mathrm{i}hmT_j}\, \Lambda(\mathrm{e}^{-\mathrm{i}hT_j}) \overset{(2.12)}{=} 0, \tag{2.13}$$

which means that the coefficient vector $\boldsymbol{\lambda} := (\lambda_0, \ldots, \lambda_N)^{\mathrm{T}}$ is the solution of the linear system

$$\boldsymbol{T}_{N+1}\boldsymbol{\lambda} = \boldsymbol{0} \tag{2.14}$$

with the matrix

$$\boldsymbol{T}_{N+1} := \Big(P\big(h(\ell - m)\big) \Big)_{m,\ell=0}^{N} \in \mathbb{C}^{(N+1)\times(N+1)}.$$

Note that this matrix is no longer a Hankel matrix as in (2.4) but a Toeplitz matrix[5].

All $(N^2 + 2N + 1)$ entries of $\boldsymbol{T}_{N+1}$ are given by the $2N + 1$ function values $P(\ell h)$, $\ell = -N, \ldots, -1, 0, 1, \ldots, N$. Since

$$P(-\ell h) = \overline{P(\ell h)}, \quad \ell = 1, \ldots, N, \tag{2.15}$$

by the conjugate symmetry (2.11), the Toeplitz matrix $\boldsymbol{T}_{N+1}$ is completely determined by the sampling values $P(\ell h)$, $\ell = 0, \ldots, N$.

---

[5]A *Toeplitz matrix* is a matrix with constant entries along the diagonals which are parallel to the main diagonal, see [34, p. 27].

The further approach is based on the calculations in the previous section. First, we can write the Toeplitz matrix $\boldsymbol{T}_{N+1}$ as a product of two Vandermonde-type matrices and a diagonal matrix. Using the definitions

$$\boldsymbol{V}_{N,N+1} := \left(\exp(-\mathrm{i}\,hkT_j)\right)_{j=1,k=0}^{N}, \qquad \boldsymbol{D} := \operatorname{diag}(c_1, c_2, \ldots, c_N),$$

we find

$$
\begin{aligned}
\boldsymbol{T}_{N+1} &= \left(P\big(h(\ell - m)\big)\right)_{m,\ell=0}^{N} \\
&= \left(\sum_{j=1}^{N} c_j \exp(-\mathrm{i}\,h(\ell - m)T_j)\right)_{m,\ell=0}^{N} \\
&= \left(\sum_{j=1}^{N} c_j \exp(\mathrm{i}\,hmT_j) \cdot \exp(-\mathrm{i}\,h\ell T_j)\right)_{m,\ell=0}^{N} \\
&= \left(c_j \exp(\mathrm{i}\,hkT_j)\right)_{k=0,j=1}^{N} \cdot \left(\exp(-\mathrm{i}\,hkT_j)\right)_{j=1,k=0}^{N} \\
&= \left(\exp(\mathrm{i}\,hkT_j)\right)_{k=0,j=1}^{N} \cdot \operatorname{diag}(c_1, c_2, \ldots, c_N) \cdot \left(\exp(-\mathrm{i}\,hkT_j)\right)_{j=1,k=0}^{N} \\
&= \boldsymbol{V}_{N,N+1}^{*} \cdot \boldsymbol{D} \cdot \boldsymbol{V}_{N,N+1}.
\end{aligned}
$$

Similarly to (2.5), we have the equalities

$$\operatorname{rank}(\boldsymbol{V}_{N,N+1}) = \operatorname{rank}(\boldsymbol{V}_{N,N+1}^{*}) = \operatorname{rank}(\boldsymbol{D}) = N. \tag{2.16}$$

Thus, in complete analogy to the steps following Lemma 2.1, we obtain

$$\operatorname{rank}(\boldsymbol{T}_{N+1}) = N,$$

and the eigenvector $\boldsymbol{\lambda}$ of $\boldsymbol{T}_{N+1}$ corresponding to the eigenvalue 0 is uniquely determined by (2.14) and $\lambda_N = 1$.

The next step consists of the computation of the zeros $z_j := \mathrm{e}^{-\mathrm{i}\,hT_j}$, $j = 1, \ldots, N$, of the polynomial Λ, which is possible since we have determined the coefficient vector $\boldsymbol{\lambda}$ of Λ in the previous step. Then, using the assumptions about h (see also Remarks 2.3, 1.) and the principal branch of the complex logarithm, we get the frequencies $T_1, \ldots, T_N$ by the evaluation

$$T_j = \frac{-\operatorname{Im}\big(\ln(z_j)\big)}{h}, \quad j = 1, \ldots, N.$$

Finally, we obtain the coefficients c_j, $j = 1, \ldots, N$, from the linear Vandermonde-type system

$$P(\ell h) = \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\,\ell h T_j}, \quad \ell = 0, \ldots, N.$$

Observe here that this system consists only of $N+1$ equations, whereas we have $2N+1$ equations in the case of complex-valued coefficients, see (2.9).

In Summary, we have the following algorithm for the determination of the frequencies and coefficients in (2.10):

Algorithm 2.5 (Prony method for real-valued coefficients).

- **Input:**
  - $P(\ell h)$, $\ell = 0, \ldots, N$;
  - step size h with $hT_j \in (-\pi, \pi]$ for all $j \in \{1, \ldots, N\}$.

- **Computation:**
  1. Compute $P(-\ell h)$, $\ell = 1, \ldots, N$, using (2.15).
  2. Construct the Toeplitz matrix
  $$\boldsymbol{T}_{N+1} := \Big(P\big(h(\ell - m)\big) \Big)_{m,\ell=0}^N.$$
  3. Solve the system $\boldsymbol{T}_{N+1}\boldsymbol{\lambda} = \boldsymbol{0}$ where $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_N)^{\mathrm{T}}$ with $\lambda_N = 1$.
  4. Consider the polynomial
  $$\Lambda(z) := \sum_{\ell=0}^N \lambda_\ell z^\ell,$$
  and compute all its zeros $z_j := \mathrm{e}^{-\mathrm{i} h T_j}$, $j = 1, \ldots, N$.
  5. Determine the frequencies $T_1, \ldots, T_N$ by
  $$T_j = \frac{-\operatorname{Im}\big(\ln(z_j)\big)}{h}, \quad j = 1, \ldots, N.$$
  (Here, ln is the principal branch of the complex logarithm.)
  6. For the frequencies T_j, $j = 1, \ldots, N$, compute the corresponding coefficients c_j, $j = 1, \ldots, N$, as least squares solution to the Vandermonde-type system
  $$\sum_{j=1}^N c_j\, \mathrm{e}^{-\mathrm{i}\ell h T_j} = P(\ell h), \quad \ell = 0, \ldots, N.$$

- **Output:** Sequences $(T_j)_{j=1}^N$ and $(c_j)_{j=1}^N$, determining P in (2.10).

Remark 2.6.
In the present case of the Prony method applied to exponential sums with real-valued coefficients, we obtain a Toeplitz matrix in Step 2 of Algorithm 2.5, while we have a Hankel matrix in the case of complex-valued coefficients.

The different matrix structure in these two cases is due to the differences in the systems of equations (2.3) and (2.13). The latter system results from (2.3) by using the indices $m = -N, \ldots, 0$ instead of the values $m = 0, \ldots, N$, and by reversing the order of the equations.

Hence, the Hankel matrix in (2.4) turns into a Toeplitz matrix, and a further consequence of the reversed order of equations is that the rank of the Toeplitz matrix is also equal to N, which we have shown above in detail.

Remark 2.7.
In analogy to Remark 2.4, we can transfer the eigenvalue problem in (2.14) into a linear system of equations. But, in contrast to the situation described in Remark 2.4, this does not enable us to take fewer sampling values of the trigonometric function P, which is due to the different matrix structure in (2.13) compared to (2.3).

3. Recovery of special univariate functions from sparse Fourier samples

In this first chapter of the main part, we will discuss the reconstruction of structured, real-valued functions from sparse Fourier samples where the considered functions are univariate. For different cases, we want to answer the question of the number of needed Fourier samples in order to uniquely recover the original function.

The key instrument for the reconstruction will be the Prony method, which has been described in Chapter 2. The focus lies on using as few Fourier samples as possible where we use clean data. Further, for the Fourier transform $\widehat{f} : \mathbb{R} \to \mathbb{C}$ of a function $f \in L^1(\mathbb{R})$ we use the definition

$$\widehat{f}(\omega) := \int_{\mathbb{R}} f(x)\, e^{-i\,\omega x}\, dx.$$

Some important properties of the Fourier transform which we need throughout this work are collected in Appendix A.

The chapter on hand is organized as follows: In the first section, we will consider step functions. Afterwards, we will extend the theory from the first section to the reconstruction of non-uniform spline functions of higher order. The third section will deal with non-uniform translates of low-pass filter functions. In order to show the applicability of the presented approaches, we will conclude the chapter with a section in which we will conduct numerical experiments.

3.1. Step functions

Let us consider a step function with compact support of the form

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) := \sum_{j=1}^{N} c_j^0\, \mathbf{1}_{[T_j, T_{j+1})}(x) \tag{3.1}$$

where c_j^0, $j = 1, \ldots, N$, are real coefficients with $c_j^0 \neq c_{j+1}^0$ for all $j \in \{1, \ldots, N-1\}$. Further, T_j, $j = 1, \ldots, N+1$, are knots satisfying $-\infty < T_1 < T_2 < \ldots < T_{N+1} < \infty$.

Here, $\mathbf{1}_{[a,b)}$ denotes the characteristic function of the interval $[a, b)$, i.e.

$$\mathbf{1}_{[a,b)}(x) = \begin{cases} 1 & \text{if } x \in [a, b); \\ 0 & \text{else.} \end{cases}$$

We aim to recover the function f from sparse Fourier samples, that is, we want to recover f from as few Fourier samples as possible. The function f is completely determined by the N coefficients $c_1^0, \ldots, c_N^0$, and the knot sequence $(T_j)_{j=1}^{N+1}$.
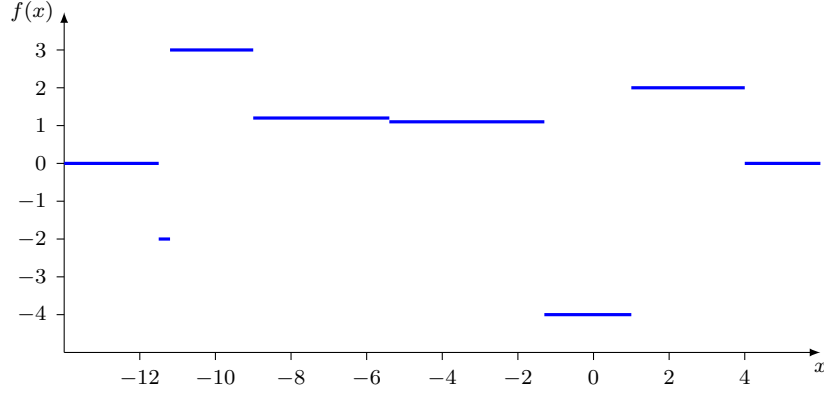
Figure 3.1. *Example for a function f of form* (3.1)*, completely determined by the knot sequence* $(-11.5, -11.2, -9, -5.4, -1.3, 1, 4)$ *and the coefficients* $-2, 3, 1.2, 1.1, -4, 2$.

In order to recover f, we need to know these $2N + 1$ function parameters. The following theorem shows that $N + 1$ Fourier samples suffice to compute all parameters.

Theorem 3.1 (Reconstruction of step functions).
Suppose that f has compact support, and that f is piecewise constant, that is, there exists a sequence $(T_j)_{j=1}^{N+1}$ of knots with

$$-\infty < T_1 < T_2 < \ldots < T_{N+1} < \infty,$$

and there exist real values c_j^0 for $j = 1, \ldots, N$ such that

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = \sum_{j=1}^{N} c_j^0 \, \mathbf{1}_{[T_j, T_{j+1})}(x). \tag{3.2}$$

The expression (3.2) *is supposed to be in its most simplified form, i.e. $c_j^0 \neq c_{j+1}^0$ for $j = 1, \ldots, N-1$, and $c_j^0 \neq 0$ for $j = 1, N$. Further, we assume that the constant $h > 0$ satisfies $hT_j \in (-\pi, \pi]$ for $j = 1, \ldots, N+1$. Then f can be completely recovered from the $N + 1$ Fourier samples $\widehat{f}(\ell h)$, $\ell = 1, \ldots, N+1$.*

Proof. We observe from (3.2) that the following holds for $\omega \neq 0$:

$$
\begin{aligned}
\widehat{f}(\omega) &\overset{\text{A.1}}{=} \int_{\mathbb{R}} f(x)\,\mathrm{e}^{-\mathrm{i}\,\omega x}\,\mathrm{d}x \\
&= \sum_{j=1}^{N} c_j^0 \int_{\mathbb{R}} \mathbf{1}_{[T_j,T_{j+1})}(x)\,\mathrm{e}^{-\mathrm{i}\,\omega x}\,\mathrm{d}x \\
&= \sum_{j=1}^{N} c_j^0 \int_{T_j}^{T_{j+1}} \mathrm{e}^{-\mathrm{i}\,\omega x}\,\mathrm{d}x \\
&= \frac{1}{\mathrm{i}\,\omega} \sum_{j=1}^{N} c_j^0 \left(\mathrm{e}^{-\mathrm{i}\,\omega T_j} - \mathrm{e}^{-\mathrm{i}\,\omega T_{j+1}} \right) \\
&= \frac{1}{\mathrm{i}\,\omega} \left[\sum_{j=1}^{N} c_j^0\,\mathrm{e}^{-\mathrm{i}\,\omega T_j} - \sum_{j=1}^{N} c_j^0\,\mathrm{e}^{-\mathrm{i}\,\omega T_{j+1}} \right] \\
&= \frac{1}{\mathrm{i}\,\omega} \left[\sum_{j=1}^{N} c_j^0\,\mathrm{e}^{-\mathrm{i}\,\omega T_j} - \sum_{j=2}^{N+1} c_{j-1}^0\,\mathrm{e}^{-\mathrm{i}\,\omega T_j} \right] \\
&= \sum_{j=1}^{N+1} \frac{c_j^1}{\mathrm{i}\,\omega}\,\mathrm{e}^{-\mathrm{i}\,\omega T_j}
\end{aligned}
\tag{3.3}
$$

with

$$
c_j^1 := c_j^0 - c_{j-1}^0, \quad j = 1, \ldots, N+1,
\tag{3.4}
$$

and with the convention that

$$
c_0^0 = c_{N+1}^0 = 0.
\tag{3.5}
$$

Observe that $c_j^1 \neq 0$ for $j = 1, \ldots, N+1$ by assumption. Otherwise, f in (3.2) can be simplified.

Since f is real-valued, i.e. $f(x) = \overline{f(x)}$, we have

$$
\widehat{f}(-\omega) = \overline{\widehat{f}(\omega)} \quad \text{for all } \omega \in \mathbb{R}
\tag{3.6}
$$

by Proposition A.2. Further, Equation (3.3) yields

$$
g(\omega) := (\mathrm{i}\,\omega)\widehat{f}(\omega) = \sum_{j=1}^{N+1} c_j^1\,\mathrm{e}^{-\mathrm{i}\,\omega T_j},
\tag{3.7}
$$

and this equation is also true for $\omega = 0$ since we have

$$\widehat{f}(0) = \sum_{j=1}^{N} c_j^0 \int_{T_j}^{T_{j+1}} 1 \, \mathrm{d}x < \infty \tag{3.8}$$

and

$$\sum_{j=1}^{N+1} c_j^1 \, \mathrm{e}^{-\mathrm{i} \cdot 0 \cdot T_j} = \sum_{j=1}^{N+1} c_j^0 - c_{j-1}^0 = c_{N+1}^0 - c_0^0 = 0. \tag{3.9}$$

Remember that $c_j^1 \neq 0$ for $j = 1, \ldots, N+1$. Moreover, for $h > 0$ we have the assumption $hT_j \in (-\pi, \pi]$ for $j = 1, \ldots, N+1$. Thus, the requirements for using the Prony method are fulfilled, and we can apply the Prony method as described in Section 2.2 to the function g, where we use the known values

$$g(\ell h) = (\mathrm{i}\, \ell h) \cdot \widehat{f}(\ell h), \quad \ell = 1, \ldots, N+1, \qquad \text{sampled Fourier data;}$$
$$g(-\ell h) = \overline{g(\ell h)}, \quad \ell = 1, \ldots, N+1, \qquad \qquad \text{by (3.7) and (3.6);}$$
$$g(0) = 0, \qquad \qquad \qquad \text{by (3.7), (3.8), and (3.9).}$$

In this way, we uniquely determine the knots T_j and the corresponding coefficients c_j^1 for $j = 1, \ldots, N+1$ in (3.7). Finally, the coefficients c_j^0, $j = 1, \ldots, N+1$, are obtained using the recursion

$$c_1^0 = c_1^1,$$
$$c_j^0 = c_{j-1}^0 + c_j^1, \quad j = 2, \ldots, N,$$

which is due to the definition of the coefficients c_j^1 in (3.4) and (3.5). □

The algorithm to compute the knot sequence and the corresponding coefficients in (3.2) from Fourier samples is based on Algorithm 2.5.

Algorithm 3.2 (Reconstruction of step functions).

- **Input:**
    - Step size $h > 0$ with $hT_j \in (-\pi, \pi]$ for $j = 1, \ldots, N+1$;
    - Fourier samples $\widehat{f}(\ell h)$, $\ell = 1, \ldots, N+1$.

- **Computation:**
    1. Compute $\widehat{f}(-\ell h)$, $\ell = 1, \ldots, N+1$, using (3.6).

2. Construct the Toeplitz matrix

$$\boldsymbol{G} := \Big(g\big(h(\ell - m)\big)\Big)_{m,\ell=0}^{N+1}$$

where $g(\omega) = (\mathrm{i}\,\omega)\widehat{f}(\omega)$ and $g(0) = 0$.

3. Solve the system $\boldsymbol{G}\boldsymbol{\lambda} = \boldsymbol{0}$ where $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_{N+1})^{\mathrm{T}}$ with $\lambda_{N+1} = 1$.

4. Consider the polynomial

$$\Lambda_{N+1}(z) := \sum_{k=0}^{N+1} \lambda_k z^k,$$

and compute all its zeros $z_j := \mathrm{e}^{-\mathrm{i}\,h T_j}$, $j = 1, \ldots, N+1$.

5. For $T_j = \frac{-\operatorname{Im}(\ln(z_j))}{h}$, $j = 1, \ldots, N+1$, compute the corresponding coefficients c_j^1, $j = 1, \ldots, N+1$, as least squares solution to the Vandermonde-type system

$$\sum_{j=1}^{N+1} c_j^1\, \mathrm{e}^{-\mathrm{i}\,\ell h T_j} = g(\ell h), \quad \ell = 1, \ldots, N+1.$$

(Here, ln is the principal branch of the complex logarithm.)

6. Compute c_j^0 for $j = 1, \ldots, N$ using the recursion

$$\begin{aligned}
c_1^0 &= c_1^1, \\
c_j^0 &= c_{j-1}^0 + c_j^1, \quad j = 2, \ldots, N.
\end{aligned}$$

- **Output:** Knot sequence $(T_j)_{j=1}^{N+1}$ and coefficients c_j^0, $j = 1, \ldots, N$, determining f in (3.2).

Remark 3.3.
The case described here is a special variant of non-uniform spline functions, which are dealt with in the following section. If the order m of the B-splines considered there is set to one, then we have a linear combination of characteristic functions, i.e. a step function as examined in the section on hand.

3.2. Non-uniform spline functions

In this section, we extend the theory from the previous section to the reconstruction of non-uniform spline functions of higher order. For this purpose, consider non-uniform spline functions $f : \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) := \sum_{j=1}^{N} c_j^0 B_j^m(x) \tag{3.10}$$

with non-zero, real-valued coefficients c_j^0 for $j = 1, \dots, N$. Here, B_j^m is the B-spline of order $m \in \mathbb{N}$ determined by the knots $T_j, \dots, T_{j+m} \in \mathbb{R}$. These knots do not have to be equispaced. The aim is to reconstruct functions f of the form (3.10) from sparse Fourier data. In order to use Fourier data for the determination of the parameters of f, we have to know a representation of the Fourier transform $\widehat{f}$. Then we can derive an algorithm to determine the parameters of f, namely the coefficients $c_1^0, \dots, c_N^0$, and the knots $T_1, \dots, T_{N+m}$.
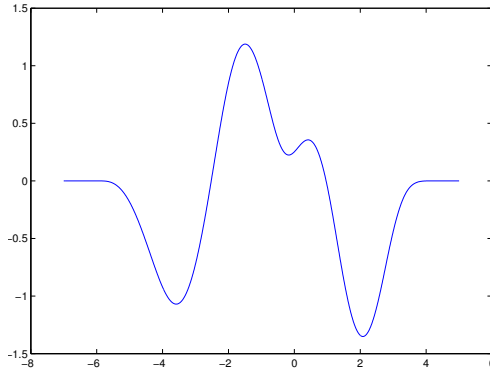
Figure 3.2. *Example for a function of the form* (3.1) *with* $N = 5, m = 5$*; determined by* $(T_j)_{j=1}^{10}$ *and* $(c_j^0)_{j=1}^{5}$ *given in* Table 3.1.

j	T_j	c_j^0
1	-6	-3.2
2	-5.8	3.1
3	-4	-0.8
4	-2.25	1.5
5	-0.6	-3
6	0	
7	1.3	
8	2.73	
9	3.5	
10	4.2	

Table 3.1. *Knots* T_j *and coefficients* c_j^0 *of the function displayed in* Figure 3.2.

First, we will take a look at some properties of the B-splines B_j^m. The B-spline B_j^1 is defined as the characteristic function of the interval $[T_j, T_{j+1})$, i.e.

$$B_j^1 := \mathbf{1}_{[T_j, T_{j+1})}. \tag{3.11}$$

B-splines of higher order m for $m > 1$ are obtained through the recurrence relation

$$B_j^m(x) = \frac{x - T_j}{T_{j+m-1} - T_j} B_j^{m-1}(x) + \frac{T_{j+m} - x}{T_{j+m} - T_{j+1}} B_{j+1}^{m-1}(x), \quad x \in \mathbb{R}, \tag{3.12}$$

see [16, pp. 89-90]. From (3.11) and (3.12) it follows directly that the B-splines are piecewise polynomial functions of degree $m - 1$ with the compact support $[T_j, T_{j+m}]$.

The smoothness of the function B_j^m is determined by the multiplicity of the knots $T_j, \ldots, T_{j+m}$, see [16, pp. 91, 99].

In the present case, we have B-splines of order m with simple knots, i.e. knots with multiplicity 1. This means that the splines are $(m-2)$-times continuously differentiable at the knots for $m \geq 3$, and for $m = 2$ the splines are only continuous at the knots but not differentiable; see [16, p. 99]. By definition, the splines of order one are only continuous from the right.

For $m \geq 3$, we can compute the first derivative of B_j^m and obtain (see [16, p. 115])

$$(B_j^m)'(x) = (m-1) \cdot \left(\frac{B_j^{m-1}(x)}{T_{j+m-1} - T_j} - \frac{B_{j+1}^{m-1}(x)}{T_{j+m} - T_{j+1}} \right). \tag{3.13}$$

Using this relation, we acquire the following lemma, see also [16, p. 117].

Lemma 3.4.
For a non-uniform spline function f of the form (3.10) with $m \geq 3$, we can compute the kth derivative $f^{(k)}$ for $k = 1, \ldots, m-2$ by

$$f^{(k)}(x) = \sum_{j=1}^{N+k} c_j^k B_j^{m-k}(x). \tag{3.14}$$

Here, the coefficients c_j^k for $j = 1, \ldots, N+k$ are recursively defined by

$$c_j^k := \left(\frac{m-k}{T_{j+m-k} - T_j} \right) \cdot \left(c_j^{k-1} - c_{j-1}^{k-1} \right) \tag{3.15}$$

with the convention

$$c_0^{k-1} = c_{N+k}^{k-1} = 0. \tag{3.16}$$

Note that c_j^k does not denote the kth power of c_j but the coefficient corresponding to the B-spline B_j^{m-k} of order $m-k$ in the linear combination (3.14).

Proof of Lemma 3.4. We will prove the statement by induction.

(BC) Base case: Using (3.13), we get the following for the first derivative of f:

$$f'(x) = \sum_{j=1}^{N} c_j^0 (B_j^m)'(x) = \sum_{j=1}^{N} c_j^0 \left[\frac{m-1}{T_{j+m-1} - T_j} B_j^{m-1}(x) - \frac{m-1}{T_{j+m} - T_{j+1}} B_{j+1}^{m-1}(x) \right]$$

$$= \sum_{j=1}^{N} c_j^0 \cdot \frac{m-1}{T_{j+m-1} - T_j} B_j^{m-1}(x) - \sum_{j=1}^{N} c_j^0 \cdot \frac{m-1}{T_{j+m} - T_{j+1}} B_{j+1}^{m-1}(x)$$

$$= \sum_{j=1}^{N} c_j^0 \cdot \frac{m-1}{T_{j+m-1} - T_j} B_j^{m-1}(x) - \sum_{j=2}^{N+1} c_{j-1}^0 \cdot \frac{m-1}{T_{j+m-1} - T_j} B_j^{m-1}(x)$$

$$= \sum_{j=1}^{N+1} \left(c_j^0 - c_{j-1}^0 \right) \cdot \frac{m-1}{T_{j+m-1} - T_j} B_j^{m-1}(x)$$

where we use the convention that $c_0^0 = c_{N+1}^0 = 0$. Finally, we have

$$f'(x) = \sum_{j=1}^{N+1} c_j^1 B_j^{m-1}(x)$$

with

$$c_j^1 := \left(c_j^0 - c_{j-1}^0 \right) \cdot \frac{m-1}{T_{j+m-1} - T_j} \quad \text{for } j = 1, \dots, N+1.$$

(IH) Induction hypothesis: The following is true for some $k \in \mathbb{N}$ with $k \geq 2$:

$$f^{(k-1)}(x) = \sum_{j=1}^{N+(k-1)} c_j^{k-1} B_j^{m-(k-1)}(x)$$

with

$$c_j^{k-1} = \left(c_j^{k-2} - c_{j-1}^{k-2} \right) \cdot \frac{m-(k-1)}{T_{j+m-(k-1)} - T_j}, \qquad c_0^{k-2} = c_{N+k-1}^{k-2} = 0.$$

(IS) Inductive step: $(k-1) \to k \leq m-1$

$$f^{(k)}(x) = \sum_{j=1}^{N+k-1} c_j^{k-1} \cdot \left(B_j^{m-(k-1)} \right)'(x) \qquad\qquad \text{by (IH)}$$

$$= \sum_{j=1}^{N+k} \left(c_j^{k-1} - c_{j-1}^{k-1} \right) \cdot \frac{m-(k-1)-1}{T_{j+m-(k-1)-1} - T_j} B_j^{m-(k-1)-1}(x) \qquad \text{by (BC)}$$

where we use the convention that $c_0^{k-1} = c_{N+k}^{k-1} = 0$. $\qquad\qquad\qquad\qquad\qquad\square$

In the case $k = m - 2$, the kth derivative $f^{(k)}$ of the spline function f is a linear combination of B-splines of order two, i.e. piecewise linear funtions. These splines are not differentiable at the knots. Therefore, we will consider distributions and generalized functions in order to be able to compute derivatives of B-splines of order two and one. This is discussed in detail in Appendix B.1 and leads to the following two lemmas.

Lemma 3.5.

Consider a generalized function f of the form (3.10) with $m \geq 2$. The $(m-1)$-th distributional derivative $D^{m-1}\mu_{(f)}$ is given by the generalized function

$$f^{(m-1)} : \mathbb{R} \ni x \mapsto \sum_{j=1}^{N+m-1} c_j^{m-1} B_j^1(x) \in \mathbb{R}$$

where the coefficients c_j^{m-1} for $j = 1, \ldots, N+m-1$ are recursively defined by

$$c_j^{m-1} := \frac{1}{T_{j+1} - T_j} \cdot (c_j^{m-2} - c_{j-1}^{m-2}) \tag{3.17}$$

with the convention

$$c_0^{m-2} = c_{N+m-1}^{m-2} = 0. \tag{3.18}$$

The proof of Lemma 3.5 is given in Section B.2. In Section B.3, we will prove Lemma 3.6.

Lemma 3.6.

The m-th distributional derivative $D^m\mu_{(f)}$ of the generalized function f of the form (3.10) with $m \geq 1$ is a linear combination of translated Dirac delta distributions, that is, it is given by the distribution

$$f^{(m)} := D^m \mu_{(f)} = \sum_{j=1}^{N+m} c_j^m (\tau_{T_j} \delta).$$

Here, δ denotes the Dirac delta distribution[1], and $\tau_{T_j}\delta$ denotes the translation of it[2] for the definition of the translate of a distribution. The coefficients c_j^m for $j = 1, \ldots, N+m$ are recursively defined by

$$c_j^m := c_j^{m-1} - c_{j-1}^{m-1} \tag{3.19}$$

with the convention

$$c_0^{m-1} = c_{N+m}^{m-1} = 0. \tag{3.20}$$

---

[1]The definition of the Dirac delta distribution is given in (B.12) on p. 146.
[2]See (B.14) on p. 147.

We want to have a representation of the Fourier transform of the function f in (3.10) in order to derive an algorithm for determining the function's parameters from sparse Fourier samples. Since the function f is not often enough differentiable in the classical sense, we had to introduce distributions and generalized functions. Therefore, we need the definition of the Fourier transform of a distribution, see Appendix B.4. Together with Lemma 3.6, this yields the following lemma which is proved in Section B.4.

Lemma 3.7.
The Fourier transform $\widehat{f}$ of the function f in (3.10) can be represented by

$$(\mathrm{i}\,\omega)^m \widehat{f}(\omega) = \sum_{j=1}^{N+m} c_j^m \, \mathrm{e}^{-\mathrm{i}\,\omega T_j}, \qquad \omega \in \mathbb{R},$$

where the coefficients c_j^m are defined by (3.19) and (3.20).

Theorem 3.8 (Reconstruction of non-uniform spline functions).
Let N and m be natural numbers. Suppose that there exist a knot sequence $(T_j)_{j=1}^{N+m}$, i.e. $-\infty < T_1 < T_2 < \ldots < T_{N+m} < \infty$, and non-zero, real values $c_j \in \mathbb{R}$ for $j = 1, \ldots, N$ such that

$$f(x) = \sum_{j=1}^{N} c_j^0 B_j^m(x) \tag{3.21}$$

is a linear combination of B-splines of order m. The B-splines B_j^m are defined by (3.11) and (3.12). Assume that the constant $h > 0$ satisfies $hT_j \in (-\pi, \pi]$ for all $j \in \{1, \ldots, N+m\}$. Then the spline function f in (3.21) can be completely recovered from the $N + m$ Fourier samples

$$\widehat{f}(\ell h), \quad \ell = 1, \ldots, N+m.$$

Proof. Due to Lemma 3.7, we obtain the representation

$$g(\omega) := (\mathrm{i}\,\omega)^m \widehat{f}(\omega) = \sum_{j=1}^{N+m} c_j^m \, \mathrm{e}^{-\mathrm{i}\,\omega T_j}, \quad \omega \in \mathbb{R}, \tag{3.22}$$

where $\widehat{f}$ is the Fourier transform of the function f of the form (3.21). Considering the assumptions about f and the constant h, we can uniquely compute the knots T_j and the coefficients c_j^m for $j = 1, \ldots, m$ by applying the Prony method for exponential sums with real-valued coefficients, see Section 2.2, to the function g.

For this purpose, we need the function values

$$g(\ell h), \quad \ell = -(N+m), \ldots, -1, 0, 1, \ldots, N+m.$$

Observe that the values $g(\ell h)$ for $\ell = 1, \ldots, N + m$ are given by the Fourier samples $\widehat{f}(\ell h)$, $\ell = 1, \ldots, N + m$, and the step size h.

These Fourier samples also provide the values $g(\ell h)$ for $\ell = -(N+m), \ldots, -1$. Since we have

$$\widehat{f}(-\omega) = \overline{\widehat{f}(\omega)} \quad \text{for all } \omega \in \mathbb{R}$$

by Proposition A.2, the values $g(\ell h)$ for $\ell = -(N + m), \ldots, -1$ can be computed by using the Fourier samples and the relation

$$g(-\omega) = (-\mathrm{i}\,\omega)^m \widehat{f}(-\omega) = \overline{\left((\mathrm{i}\,\omega)^m \widehat{f}(\omega)\right)} = \overline{g(\omega)}. \tag{3.23}$$

The last remaining value of g which we need is given as

$$g(0) = 0. \tag{3.24}$$

This can be seen as follows: The left-hand side of (3.22), i.e. $g(\omega) = (\mathrm{i}\,\omega)^m \widehat{f}(\omega)$, yields the value zero for $\omega = 0$ since

$$\left|\widehat{f}(0)\right| = \left|\int_{\mathbb{R}} f(x)\, \mathrm{e}^{-\mathrm{i}\cdot 0 \cdot x}\, \mathrm{d}x\right| < \infty$$

due to f being an almost everywhere continuous function with compact support. Further, the right-hand side of (3.22) also equals to zero for $\omega = 0$:

$$\sum_{j=1}^{N+m} c_j^m\, \mathrm{e}^{-\mathrm{i}\cdot 0 \cdot T_j} = \sum_{j=1}^{N+m} \left(c_j^{m-1} - c_{j-1}^{m-1}\right) = \sum_{j=1}^{N+m} c_j^{m-1} - \sum_{j=0}^{N+m-1} c_j^{m-1} = c_{N+m}^{m-1} - c_0^{m-1} = 0,$$

where we have used the definition of the coefficients c_j^m in (3.19) and (3.20). Thus, the representation (3.22) is also valid for $\omega = 0$, and the equality (3.24) holds.

The coefficients c_j^0 are computed by a recursion formula which is derived from the definitions of the coefficients c_j^k in Lemmas 3.4–3.6. Using (3.15)–(3.20), we obtain

$$c_j^{k-1} = \begin{cases} c_1^m & \text{for } k = m,\ j = 1, \\ c_j^m + c_{j-1}^{m-1} & \text{for } k = m,\ j = 2, \ldots, N + m - 1, \\ \left(\frac{T_{1+m-k} - T_1}{m-k}\right) c_1^k & \text{for } k = m-1, \ldots, 1,\ j = 1, \\ \left(\frac{T_{j+m-k} - T_j}{m-k}\right) c_j^k + c_{j-1}^{k-1} & \text{for } k = m-1, \ldots, 1,\ j = 2, \ldots, N + k - 1. \end{cases}$$

$\square$

Remarks 3.9.

1. The proof of Theorem 3.8 which is given above is constructive. In particular, if N is not known, but we have an upper bound $M \geq N$, then we will also find the correct knots T_j and the corresponding coefficients c_j^m by applying the Prony method to $M + m$ Fourier samples, cf. Remarks 2.3, 3. Observe that we need fewer sampling values than stated in the mentioned remarks since the present case deals with exponential sums with real-valued coefficients, cf. Section 2.2. By this way, the numerical procedure will also be more stable, see [23, 51, 55].

2. In the above proof, we rely upon the fact that $c_j^m \neq 0$ for $j = 1, \ldots, N + m$. If we have the situation that $c_{j_0}^m = 0$ for an index $j_0 \in \{1, \ldots, N+m\}$, then we will not be able to reconstruct the knot T_{j_0}. But this situation will only occur if the representation of f in (3.21) is redundant, that is to say, if f in (3.21) can be represented by less than N summands. Therefore, we will still be able to exactly recover the function f. Observe that the above procedure always results in the simplest representation of f such that the reconstructed representation of f of the form (3.21) does not possess redundant terms.

3. Considering the non-linear problem to approximate a continuous univariate function g from given samples by a spline function with free knots, we want to find optimal knots as well as optimal coefficients of the B-spline expansion f in (3.21) such that $g - f$ is small in a given norm. This problem is very challenging but of high interest for sparse signal approximation. While the above Theorem 3.8 yields a reconstruction of spline functions with free knots from Fourier samples, it would be highly desirable to have a method that, using sampling values of g itself, yields a non-linear approximation of g by constructing optimal knots and optimal coefficients of a spline function.

The algorithm to determine the knots T_j and the coefficients c_j^0 in (3.21) is similar to Algorithm 3.2 and based on Algorithm 2.5.

Algorithm 3.10 (Reconstruction of non-uniform spline functions).

- **Input:**
  - Step size $h > 0$ with $hT_j \in (-\pi, \pi]$ for $j = 1, \ldots, N + m$;
  - Fourier samples $\widehat{f}(\ell h)$, $\ell = 1, \ldots, N + m$.

- **Computation:**
  1. Compute $\widehat{f}(-\ell h)$, $\ell = 1, \ldots, N + 1$, using (3.23).

2. Construct the Toeplitz matrix

$$\boldsymbol{G} := \Big(g\big(h(\ell - k)\big)\Big)_{k,\ell=0}^{N+m}$$

where $g(\omega) = (\mathrm{i}\,\omega)\widehat{f}(\omega)$ and $g(0) = 0$.

3. Solve the system $\boldsymbol{G}\boldsymbol{\lambda} = \boldsymbol{0}$ where $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_{N+m})^{\mathrm{T}}$ with $\lambda_{N+m} = 1$.

4. Consider the polynomial

$$\Lambda_{N+m}(z) := \sum_{k=0}^{N+m} \lambda_k z^k,$$

and compute all its zeros $z_j := \mathrm{e}^{-\mathrm{i}\,h T_j}$, $j = 1, \ldots, N + m$.

5. For $T_j = \frac{-\operatorname{Im}(\ln(z_j))}{h}$, $j = 1, \ldots, N + m$, compute the corresponding coefficients c_j^m, $j = 1, \ldots, N + m$, as least squares solution to the Vandermonde-type system

$$\sum_{j=1}^{N+m} c_j^m\, \mathrm{e}^{-\mathrm{i}\,\ell h T_j} = g(\ell h), \quad \ell = 1, \ldots, N + m.$$

(Here, ln is the principal branch of the complex logarithm.)

6. Compute c_j^0 for $j = 1, \ldots, N$ using the recursion

$$c_j^{k-1} = \begin{cases} c_1^m & \text{for } k = m,\, j = 1, \\ c_j^m + c_{j-1}^{m-1} & \text{for } k = m,\, j = 2, \ldots, N + m - 1, \\ \left(\frac{T_{1+m-k}-T_1}{m-k}\right) c_1^k & \text{for } k = m-1, \ldots, 1,\, j = 1, \\ \left(\frac{T_{j+m-k}-T_j}{m-k}\right) c_j^k + c_{j-1}^{k-1} & \text{for } k = m-1, \ldots, 1,\, j = 2, \ldots, N + k - 1. \end{cases}$$

- **Output:** Sequence $(T_j)_{j=1}^{N+m}$ and coefficients $c_1^0, \ldots, c_N^0$ determining f in (3.21).

3.3. Non-uniform translates

As the last univariate case, we will examine functions $f : \mathbb{R} \to \mathbb{R}$ that have a sparse representation of the form

$$f(x) := \sum_{j=1}^{N} c_j \Phi(x - T_j) \tag{3.25}$$

with $c_j \in \mathbb{R} \setminus \{0\}$ for $j = 1, \ldots, N$, and a shift sequence $-\infty < T_1 < \ldots < T_N < \infty$. Further, we assume that $\Phi \in L^1(\mathbb{R})$ is a real low-pass filter function with a Fourier transform that is bounded away from zero, i.e. $|\widehat{\Phi}(\omega)| > C$ for $\omega \in (-T, T)$ for some constants $C > 0$ and $T > 0$. For example, as a low-pass filter function Φ we can take

- the centred cardinal B-spline of order m, $\Phi = N_m$, with

$$\widehat{N}_m(\omega) = \left(\mathrm{sinc} \left(\frac{\omega}{2} \right) \right)^m \neq 0 \quad \text{for all } \omega \in (-2\pi, 2\pi);$$

- the Gaussian function, $\Phi(x) = \exp \left(-\frac{x^2}{\sigma^2} \right)$, $\sigma > 0$, with

$$\widehat{\Phi}(\omega) = \sqrt{\pi} \cdot \sigma \cdot \exp \left(-\frac{\sigma^2 \omega^2}{4} \right) > 0 \quad \text{for all } \omega \in \mathbb{R};$$

- the Meyer window Φ, see [15, p. 137], which is defined by the Fourier transform

$$\widehat{\Phi}(\omega) = \begin{cases} 1 & \text{for } |\omega| \leq \frac{1}{3}, \\ \cos \left(\frac{\pi}{2}(3|\omega| - 1) \right) & \text{for } \frac{1}{3} < |\omega| \leq \frac{2}{3}, \\ 0 & \text{otherwise}; \end{cases}$$

- a real-valued Gabor function $\Phi(x) = \mathrm{e}^{-\alpha x^2} \cos(\beta x)$, $\alpha > 0$, $\beta > 0$, with

$$\widehat{\Phi}(\omega) = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}} \left(\exp \left(-\frac{(\beta - \omega)^2}{4\alpha} \right) + \exp \left(-\frac{(\omega + \beta)^2}{4\alpha} \right) \right) > 0 \quad \text{for all } \omega \in \mathbb{R}.$$

The Fourier transform of f is given by (cf. Definition A.1 and Proposition A.3)

$$\widehat{f}(\omega) = \left(\sum_{j=1}^{N} c_j \, \mathrm{e}^{-\mathrm{i}\omega T_j} \right) \widehat{\Phi}(\omega), \quad \omega \in \mathbb{R}. \tag{3.26}$$

Theorem 3.11 (Reconstruction of non-uniform translates I).
Let $-\infty < T_1 < \ldots < T_N < \infty$ be a real sequence and $c_j \in \mathbb{R} \setminus \{0\}$ for $j = 1, \ldots, N$. Further, let $\Phi \in L^1(\mathbb{R})$ be a given real-valued function with $|\widehat{\Phi}(\omega)| > C$ for $\omega \in (-T, T)$ for some constants $C > 0$ and $T > 0$. Assume that the constant $h > 0$ satisfies $hT_j \in (-\pi, \pi]$ for $j = 1, \ldots, N$ and $h < \frac{T}{N}$. Then the function f of the form (3.25) can be uniquely recovered from the Fourier samples $\widehat{f}(\ell h)$, $\ell = 0, \ldots, N$.

Proof. Equation (3.26) yields the following for $\omega \in (-T, T)$:

$$g(\omega) := \frac{\widehat{f}(\omega)}{\widehat{\Phi}(\omega)} = \sum_{j=1}^{N} c_j \, \mathrm{e}^{-\mathrm{i}\omega T_j} . \tag{3.27}$$

Since $\widehat{\Phi}(\ell h) \neq 0$ for $\ell = 0, \ldots, N$, and $hT_j \in (-\pi, \pi]$ for $j = 1, \ldots, N$ by the assumptions about h, we can apply the Prony method as described in Section 2.2 to the function g, where we use the known function values

$$g(\ell h) = \frac{\widehat{f}(\ell h)}{\widehat{\Phi}(\ell h)}, \quad \ell = 0, \ldots, N.$$

But we still need the function values $g(\ell h)$ for $\ell = -N, \ldots, -1$ in order to compute all shift values T_j and all coefficients c_j in (3.27) by applying the Prony method. These values can be computed using the already known function values $g(\ell h)$ for $\ell = 1, \ldots, N$ because we have

$$g(-\omega) = \frac{\widehat{f}(-\omega)}{\widehat{\Phi}(-\omega)} = \frac{\overline{\widehat{f}(\omega)}}{\overline{\widehat{\Phi}(\omega)}} = \overline{g(\omega)}$$

by Proposition A.2 since f and Φ are real-valued. $\qquad\square$

The above idea can be generalized to functions $f : \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) := \sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r} \Phi^{(r)}(x - T_j) \tag{3.28}$$

where $c_{j,r}$ are real-valued coefficients with $c_{j,R-1} \neq 0$ for all $j \in \{1, \ldots, N\}$, and where $(T_j)_{j=1}^{N}$ is again a real sequence of shift values. Further, we assume that the given real-valued low-pass filter function $\Phi \in L^1(\mathbb{R})$ is $(R-1)$-times differentiable for $R \in \mathbb{N}$. For $r = 0, \ldots, R-1$, the r-th derivative of Φ is denoted by $\Phi^{(r)}$ where we suppose that $\Phi^{(r)} \in L^1(\mathbb{R})$ for $r = 1, \ldots, R-1$. Using the propositions A.1, A.3, and A.4, we obtain the following representation of the Fourier transform of f:

$$\widehat{f}(\omega) = \left(\sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r}(\mathrm{i}\,\omega)^r \mathrm{e}^{-\mathrm{i}\,\omega T_j} \right) \widehat{\Phi}(\omega), \quad \omega \in \mathbb{R}. \tag{3.29}$$

Theorem 3.12 (Reconstruction of non-uniform translates II).
Let $-\infty < T_1 < \ldots < T_N < \infty$ be a real sequence and $c_{j,r} \in \mathbb{R}$ for $j \in \{1, \ldots, N\}$ and $r \in \{0, \ldots, R-1\}$ where we assume that $c_{j,R-1} \neq 0$ for all $j \in \{1, \ldots, N\}$. Further, let $\Phi \in L^1(\mathbb{R})$ be a given real-valued function which is $(R-1)$-times differentiable with $\Phi^{(r)} \in L^1(\mathbb{R})$ for $r = 1, \ldots, R-1$, and which fulfils $|\widehat{\Phi}(\omega)| > C$ for all $\omega \in (-T, T)$ for some constants $C > 0$ and $T > 0$. Assume that the constant $h > 0$ satisfies $hT_j \in (-\pi, \pi]$ for $j = 1, \ldots, N$ and $h < \frac{T}{N}$. Then the function f in (3.28) can be uniquely recovered from the Fourier samples $\widehat{f}(\ell h)$, $\ell = 0, \ldots, NR$.

Proof. Using (3.29) and the assumptions about $\widehat{\Phi}$, we get

$$Q(\omega) := \frac{\widehat{f}(\omega)}{\widehat{\Phi}(\omega)} = \sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r}(\mathrm{i}\,\omega)^r \mathrm{e}^{-\mathrm{i}\,\omega T_j} \tag{3.30}$$

for $\omega \in (-T, T)$. In order to compute the shift values T_j and the coefficients $c_{j,r}$, we want to apply the Prony method to the function Q. Note that the structure of Q is different from the function structure in (2.10). Functions Q of the form in (3.30) are called *extended exponential sums*, see [10, p. 169].

We consider now the polynomial

$$\Lambda(z) := \prod_{j=1}^{N} \left(z - \mathrm{e}^{-\mathrm{i}\,hT_j} \right)^R = \sum_{\ell=0}^{NR} \lambda_\ell z^\ell \tag{3.31}$$

where $T_1, \ldots, T_N$ are the unknown shift values in (3.28). Then we observe that the following holds for $m = 0, \ldots, NR$:

$$\begin{aligned}
\sum_{\ell=0}^{NR} \lambda_\ell \, Q(h(\ell - m)) &= \sum_{\ell=0}^{NR} \lambda_\ell \sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r} (\mathrm{i}\,h(\ell - m))^r \, \mathrm{e}^{-\mathrm{i}\,h(\ell - m)T_j} \\
&= \sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r} \, \mathrm{e}^{\mathrm{i}\,hmT_j} \sum_{\ell=0}^{NR} \lambda_\ell \, (\mathrm{i}\,h(\ell - m))^r \, \mathrm{e}^{-\mathrm{i}\,h\ell T_j} \\
&= \sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r} \, \mathrm{e}^{\mathrm{i}\,hmT_j} (\mathrm{i}\,h)^r \sum_{\nu=0}^{r} \binom{r}{\nu} (-m)^{r-\nu} \sum_{\ell=0}^{NR} \lambda_\ell \, \ell^\nu \, \mathrm{e}^{-\mathrm{i}\,h\ell T_j} \\
&= \sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r} \, \mathrm{e}^{\mathrm{i}\,hmT_j} (\mathrm{i}\,h)^r \sum_{\nu=0}^{r} \binom{r}{\nu} (-m)^{r-\nu} S_\nu \tag{3.32}
\end{aligned}$$

where we have used the binomial theorem[3] and

$$S_\nu := \sum_{\ell=0}^{NR} \lambda_\ell \, \ell^\nu \, \mathrm{e}^{-\mathrm{i}\,h\ell T_j} .$$

The derivatives of the polynomial Λ are computed as

$$\Lambda^{(\eta)}(z) = \sum_{\ell=\eta}^{NR} \lambda_\ell \left(\prod_{k=0}^{\eta-1} (\ell - k) \right) z^{\ell-\eta}, \quad \eta \in \mathbb{N}_0.$$

We will show later, in Lemma 3.17, that S_ν can be written as a linear combination of the derivatives $\Lambda^{(\eta)}$, $\eta = 0, \ldots, \nu$, that is, there exist coefficients $\alpha_\eta^\nu \in \mathbb{N}_0$ such that

$$S_\nu = \sum_{\eta=0}^{\nu} \alpha_\eta^\nu \Lambda^{(\eta)}(\mathrm{e}^{-\mathrm{i}\,hT_j}) \, \mathrm{e}^{-\mathrm{i}\,\eta hT_j} . \tag{3.33}$$

---

[3]See [31, p. 57, Satz 7.4], for instance.

By definition of Λ in (3.31), we get

$$\Lambda^{(\eta)}(\mathrm{e}^{-\mathrm{i}\,hT_j}) = 0 \quad \text{for } \eta = 0, \ldots, R-1.$$

Plugged into (3.33), this yields

$$S_\nu = 0 \quad \text{for } \nu = 0, \ldots, R-1. \tag{3.34}$$

Combining (3.32) and (3.34), we finally get

$$\sum_{\ell=0}^{NR} \lambda_\ell Q(h(\ell - m)) = 0.$$

Hence, the coefficient vector $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_{NR})^\mathrm{T}$ with $\lambda_{NR} = 1$ is the solution of the linear system

$$\boldsymbol{T}_{NR+1} \boldsymbol{\lambda} = \boldsymbol{0} \tag{3.35}$$

where $\boldsymbol{T}_{NR+1}$ is the Toeplitz matrix defined by

$$\boldsymbol{T}_{NR+1} := \Big(Q\big(h(\ell - m)\big) \Big)_{m,\ell=0}^{NR} \in \mathbb{C}^{(NR+1)\times(NR+1)}.$$

Note that we have the property

$$\mathrm{rank}(\boldsymbol{T}_{NR+1}) = NR$$

due to the assumptions about the coefficients $c_{j,r}$ and the values hT_j. This is shown in detail in Lemma 3.13. Hence, the eigenvector $\boldsymbol{\lambda}$ of $\boldsymbol{T}_{NR+1}$ corresponding to the eigenvalue $\boldsymbol{0}$ is uniquely determined by (3.35) and $\lambda_{NR} = 1$.

Since f and Φ are real-valued, we have

$$Q(-\omega) = \frac{\widehat{f}(-\omega)}{\widehat{\Phi}(-\omega)} = \frac{\overline{\widehat{f}(\omega)}}{\overline{\widehat{\Phi}(\omega)}} = \overline{Q(\omega)}$$

by Proposition A.2. Thus, all entries of $\boldsymbol{T}_{NR+1}$ are given by the $NR+1$ sampling values $\widehat{f}(\ell h)$ for $\ell = 0, \ldots, NR$.

The solution $\boldsymbol{\lambda}$ of (3.35) is the coefficient vector of Λ in (3.31). Therefore, we can compute all roots of Λ, which are given by the values $z_j := \mathrm{e}^{-\mathrm{i}\,hT_j}$, $j = 1, \ldots, N$. Note that all these roots are roots of multiplicity R. From the roots z_j we get the shift values T_j by

$$T_j = \frac{-\mathrm{Im}\big(\ln(z_j)\big)}{h}, \quad j = 1, \ldots, N.$$

As the last step, we compute the coefficients $c_{j,r}$ for $j = 1, \ldots, N$, $r = 0, \ldots, R-1$ as least squares solution to the linear system

$$\sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r} (\mathrm{i}\,\ell h)^r \, \mathrm{e}^{-\mathrm{i}\,\ell h T_j} = Q(\ell h), \quad \ell = 0, \ldots, NR.$$

$\square$

We now want to show that the rank of the data matrix $\boldsymbol{T}_{NR+1}$ considered in the proof of Theorem 3.12 is equal to NR such that the dimension of the eigenspace corresponding to the eigenvalue 0 is equal to one.

Lemma 3.13.
Let the assumptions from Theorem 3.12 be fulfilled. Then the rank of the matrix $\boldsymbol{T}_{NR+1}$ in (3.35) is equal to NR such that the solution $\boldsymbol{\lambda}$ to the eigenvalue problem (3.35) is uniquely determined by the normalization $\lambda_{NR} = 1$.

Proof. Consider the linear system

$$\sum_{\ell=0}^{NR} \lambda_\ell Q\big(h(\ell - m)\big) = 0, \quad m = 0, \ldots, NR.$$

An equivalent formulation for this system is

$$\boldsymbol{T}_{NR+1} \boldsymbol{\lambda} = \boldsymbol{0}$$

with the Toeplitz matrix

$$\boldsymbol{T}_{NR+1} = \left(Q\big(h(\ell - m)\big) \right)_{m,\ell=0}^{NR} \in \mathbb{C}^{(NR+1) \times (NR+1)}$$

and the coefficient vector $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_{NR})^{\mathrm{T}}$ with $\lambda_{NR} = 1$.

Reversing the order of the equations, we obtain the linear system

$$\boldsymbol{H}_{NR+1} \boldsymbol{\lambda}' = \boldsymbol{0}$$

with the Hankel matrix

$$\boldsymbol{H}_{NR+1} := \left(Q\big(h(-NR + m + \ell)\big) \right)_{m,\ell=0}^{NR} \in \mathbb{C}^{(NR+1) \times (NR+1)}$$

and the coefficient vector $\boldsymbol{\lambda}' := (\lambda_{NR}, \ldots, \lambda_0)^{\mathrm{T}}$ with $\lambda_{NR} = 1$.

The data matrix $\boldsymbol{H}_{NR+1}$ consists of sampling values of the signal

$$Q(\omega) = \sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r} (\mathrm{i}\,\omega)^r \, \mathrm{e}^{-\mathrm{i}\,\omega T_j} = \sum_{j=1}^{N} \sum_{r=0}^{R-1} c_{j,r} (\mathrm{i}\,\omega)^r \, \mathrm{e}^{-\mathrm{i}\,\frac{\omega}{h}(hT_j)} = \sum_{j=1}^{N} a_j(\omega) z_j^{\frac{\omega}{h}} \quad (3.36)$$

where we use the definitions

$$z_j := \mathrm{e}^{-\mathrm{i}hT_j} \qquad \text{and} \qquad a_j(\omega) := \sum_{r=0}^{R-1} c_{j,r}(\mathrm{i}\omega)^r.$$

The coefficients $a_j(\omega)$ in the sum in (3.36) are complex-valued polynomials of degree less than or equal to $R - 1$. Each of these polynomials can be represented as a linear combination of falling factorials. Using

$$F_k(x) := \begin{cases} 0 & \text{if } k < 0, \\ 1 & \text{if } k = 0, \\ \prod_{k'=0}^{k-1}(x - k') & \text{if } k > 0, \end{cases}$$

we get

$$a_j(x) = \sum_{k=0}^{R-1} a'_{j,k} F_k(x), \quad j = 1, \dots, N,$$

with coefficients $a'_{j,k} \in \mathbb{C}$, $j = 1, \dots, N$, $k = 0, \dots, R - 1$, see [2, II. A. and II. B.]. Thus, the signal Q can be represented by

$$Q(\omega) = \sum_{j=1}^{N} \sum_{k=0}^{R-1} a'_{j,k} F_k(\omega) z_j^{\frac{\omega}{h}}.$$

For the sampling locations $ht \in \{-NRh, \dots, -h, 0, h, \dots, NRh\}$, we have

$$\begin{aligned} Q(ht) &= \sum_{j=1}^{N} \sum_{k=0}^{R-1} a'_{j,k} F_k(ht) z_j^t \\ &= \sum_{j=1}^{N} \sum_{k=0}^{R-1} a_{j,k} F_k(ht) z_j^{t-k} \end{aligned} \tag{3.37}$$

with the falling factorials

$$F_k(ht) = ht \cdot (ht - 1) \cdot \ldots \cdot (ht - k + 1)$$

and the coefficients $a_{j,k} := a'_{j,k} z_j^k$, cf. [2, [II. A. and II. B.] and [7, Eq. (1.5)].

Using the representation (3.37), the data matrix $\boldsymbol{H}_{NR+1}$ can now be factorized as follows:

$$\boldsymbol{H}_{NR+1} = \boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^{\mathrm{T}}$$

with a so-called *confluent Vandermonde matrix* $\boldsymbol{U}$ of size $(NR + 1) \times NR$ and an $NR \times NR$ block diagonal matrix $\boldsymbol{B} := \mathrm{diag}(\boldsymbol{B}_1, \dots, \boldsymbol{B}_N)$ where each block $\boldsymbol{B}_j$ for

$j = 1, \ldots, N$ of size $R \times R$ is an upper antitriangular matrix such that $(\boldsymbol{B}_j)_{k,\ell} = 0$ for $\ell > k$, see [2, Proposition III.7] and [7, Lemma 2.6].

Further, we have

$$\text{rank}(\boldsymbol{U}) = NR$$

since the values $z_j = \mathrm{e}^{-\mathrm{i}hT_j}$, $j = 1, \ldots, N$, are pairwise different by assumption, cf. [2, Proposition III.6] and [7, Eq. (2.4)].

The block diagonal matrix $\boldsymbol{B}$ is non-singular if and only if

$$a_{j,R-1} \neq 0 \quad \text{for all } j \in \{1, \ldots, N\},$$

cf. [2, Corollary III.8] and [7, Theorem 2.7], which means that each polynomial $a_j(\cdot)$ is of exact degree $R - 1$. This is fulfilled by the assumption that $c_{j,R-1} \neq 0$ for all $j \in \{1, \ldots, N\}$.

Then we have (compare the explanations following Lemma 2.1)

$$\text{rank}(\boldsymbol{H}_{NR+1}) = NR$$

such that the eigenvector corresponding to the eigenvalue 0, i.e. the solution vector $\boldsymbol{\lambda'} = (\lambda_{NR}, \ldots, \lambda_0)^{\mathrm{T}}$, is uniquely determined by the normalization $\lambda_{NR} = 1$.

Since the system

$$\boldsymbol{H}_{NR+1}\boldsymbol{\lambda'} = \boldsymbol{0}$$

is obtained by reversing the order of the equations in the system

$$\boldsymbol{T}_{NR+1}\boldsymbol{\lambda} = \boldsymbol{0},$$

also the vector $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_{NR})^{\mathrm{T}}$ with $\lambda_{NR} = 1$ is uniquely determined. This completes the proof. $\qquad\square$

Remarks 3.14.

1. Observe that the described reconstruction method still works if the highest orders of the derivatives $\Phi^{(r)}(\cdot - T_j)$ in (3.28) for $j \in \{1, \ldots, N\}$ are different, that is, if we consider

$$f(x) = \sum_{j=1}^{N} \sum_{r=0}^{R_j-1} c_{j,r}\Phi^{(r)}(x - T_j).$$

Then we need $R + 1$ Fourier samples where $R := \sum_{j=1}^{N} R_j$. In this case, see [2] and [7] for the factorization of the data matrix occurring in a problem similar to (3.28).

2. Similarly as in (3.28), we can also generalize the method to sums of B-splines and their derivatives, that is to say, we can consider non-uniform translates of B-splines of different order as follows:

$$f(x) = \sum_{j=1}^{N} \sum_{r=1}^{m} c_{j,r} \, B_j^r(x).$$

Then $f(x)$ can be recovered from the Fourier samples $\widehat{f}(\ell h)$, $\ell = 1, \dots, (N+m)m$.

3. The special functions f regarded in Sections $3.1 - 3.3$ can also be seen as functions of finite rate of innovation as introduced in [65], that is, functions with a finite number of degrees of freedom per unit of time.

It still remains to show that the representation of S_ν in (3.33) is valid. In order to prove this, we need some further statements.

Lemma 3.15.
Define for $\nu \in \mathbb{N}$ with $\nu \geq 2$ the function $D_\nu : \mathbb{N}_0 \to \mathbb{N}$ by

$$D_\nu(\ell) := \ell^\nu - \prod_{k=0}^{\nu-1} (\ell - k) = \ell^\nu - \ell(\ell-1) \cdot \dots \cdot (\ell - \nu + 1)$$

$$= \begin{cases} \ell^\nu, & \ell < \nu, \\ \ell^\nu - \frac{\ell!}{(\ell-\nu)!}, & \nu \leq \ell. \end{cases}$$

(3.38)

Then the following is true:

1) $D_\nu(\ell) = \sum_{m=1}^{\nu-1} (-1)^{m+1} d_{\nu-m}^\nu \ell^{\nu-m}$

$$\text{with} \qquad \begin{aligned} d_1^2 &= 1, \\ d_1^\nu &= d_1^{\nu-1} \cdot (\nu - 1), & \nu \geq 3, \\ d_{\nu-m}^\nu &= d_{\nu-m}^{\nu-1} \cdot (\nu - 1) + d_{\nu-m-1}^{\nu-1}, & m = 2, \dots, \nu - 2, \ \nu \geq 4, \\ d_{\nu-1}^\nu &= d_{\nu-2}^{\nu-1} + (\nu - 1), & \nu \geq 3. \end{aligned}$$

2) $\sum_{m=1}^{\nu-1} (-1)^{m+1} d_{\nu-m}^\nu = 1, \quad \nu \geq 2.$

3) $d_{\nu-1}^\nu = \sum_{j=1}^{\nu-1} j = \frac{(\nu-1)\nu}{2}, \quad \nu \geq 3.$

Proof. We will prove the statements by induction.

1) (BC) Base case:

$$D_2(\ell) = \ell^2 - \ell(\ell - 1) = \ell,$$
$$D_3(\ell) = \ell^3 - \ell(\ell - 1)(\ell - 2) = \ell^3 - \ell(\ell^2 - 3\ell + 2) = 3\ell^2 - 2\ell,$$
$$D_4(\ell) = \ell^4 - \ell(\ell - 1)(\ell - 2)(\ell - 3) = 6\ell^3 - 11\ell^2 + 6\ell.$$

$$\Rightarrow \quad d_1^2 = 1;$$
$$\Rightarrow \quad d_1^3 = 2 = d_1^2 \cdot 2, \qquad d_2^3 = 3 = d_1^2 + 2;$$
$$\Rightarrow \quad d_1^4 = 6 = d_1^3 \cdot 3, \qquad d_2^4 = 11 = d_2^3 \cdot 3 + d_1^3, \qquad d_3^4 = 6 = d_2^3 + 3.$$

(IH) Induction hypothesis: The claim holds for some $\nu \in \mathbb{N}$ with $\nu \geq 4$.

(IS) Inductive step: $4 \leq \nu \to \nu + 1$

$$
\begin{aligned}
D_{\nu+1}(\ell) \;&=\; \ell^{\nu+1} - \prod_{k=0}^{\nu}(\ell - k) \;=\; \ell^{\nu+1} - (\ell - \nu) \cdot \prod_{k=0}^{\nu-1}(\ell - k) \\[2mm]
&=\; \ell \cdot \left[\ell^\nu - \prod_{k=0}^{\nu-1}(\ell - k) \right] + \nu \cdot \prod_{k=0}^{\nu-1}(\ell - k) \\[2mm]
&=\; \ell \cdot \left[D_\nu(\ell) \right] + \nu \cdot \left[\ell^\nu - \left(\ell^\nu - \prod_{k=0}^{\nu-1}(\ell - k) \right) \right] \\[2mm]
&=\; \ell \cdot \left[D_\nu(\ell) \right] + \nu \cdot \left[\ell^\nu - D_\nu(\ell) \right] \\[2mm]
&\overset{\text{(IH)}}{=}\; \ell \cdot \sum_{m=1}^{\nu-1}(-1)^{m+1} d_{\nu-m}^\nu \ell^{\nu-m} + \nu \cdot \left[\ell^\nu - \sum_{m=1}^{\nu-1}(-1)^{m+1} d_{\nu-m}^\nu \ell^{\nu-m} \right] \\[2mm]
&=\; \nu\ell^\nu + \sum_{m=1}^{\nu-1}(-1)^{m+1} d_{\nu-m}^\nu \ell^{(\nu+1)-m} + \sum_{m=1}^{\nu-1}(-1)^{m+2} d_{\nu-m}^\nu \cdot \nu\ell^{\nu-m} \\[2mm]
&=\; (d_{\nu-1}^\nu + \nu) \cdot \ell^\nu + \sum_{m=2}^{\nu-1}(-1)^{m+1} d_{\nu-m}^\nu \ell^{(\nu+1)-m} \\[2mm]
&\qquad + \sum_{m=1}^{\nu-2} \left[(-1)^{m+2} d_{\nu-m}^\nu \cdot \nu\ell^{\nu-m} \right] + (-1)^{\nu-1+2}(d_1^\nu \cdot \nu) \cdot \ell \\[2mm]
&=\; (d_{\nu-1}^\nu + \nu) \cdot \ell^\nu + \sum_{m=2}^{\nu-1}(-1)^{m+1} d_{\nu-m}^\nu \ell^{(\nu+1)-m} \\[2mm]
&\qquad + \sum_{m=2}^{\nu-1} \left[(-1)^{m+1} d_{\nu-(m-1)}^\nu \cdot \nu\ell^{\nu-(m-1)} \right] + (-1)^{\nu+1}(d_1^\nu \cdot \nu) \cdot \ell
\end{aligned}
$$

$$= (d_{\nu-1}^{\nu} + \nu) \cdot \ell^{\nu} + \sum_{m=2}^{\nu-1} (-1)^{m+1} \left[d_{\nu+1-m}^{\nu} \cdot \nu + d_{\nu-m}^{\nu} \right] \cdot \ell^{(\nu+1)-m}$$
$$+ (-1)^{\nu+1} (d_1^{\nu} \cdot \nu) \cdot \ell.$$

Hence, we have

$$d_1^{\nu+1} = d_1^{\nu} \cdot \nu,$$

$$d_{\nu+1-m}^{\nu+1} = d_{\nu+1-m}^{\nu} \cdot \nu + d_{\nu-m}^{\nu}, \quad m = 2, \dots, \nu-1,$$

$$d_{\nu}^{\nu+1} = d_{\nu-1}^{\nu} + \nu.$$

2) (BC) We know from part 1) that

$$d_1^2 = 1 \quad \text{and} \quad d_1^3 = 2, \ d_2^3 = 3.$$

Therefore, we compute

$$\sum_{m=1}^{\nu-1} (-1)^{m+1} d_{\nu-m}^{\nu} = 1 \quad \text{for } \nu = 2 \text{ and } \nu = 3.$$

(IH) The claim holds for some $\nu \in \mathbb{N}$ with $\nu \geq 3$.

(IS) $3 \leq \nu \to \nu + 1$

$$\sum_{m=1}^{\nu} (-1)^{m+1} d_{\nu+1-m}^{\nu+1} = d_{\nu}^{\nu+1} + \sum_{m=2}^{\nu-1} \left[(-1)^{m+1} d_{\nu+1-m}^{\nu+1} \right] + (-1)^{\nu+1} d_1^{\nu+1}$$

$$\overset{1)}{=} (d_{\nu-1}^{\nu} + \nu) + \sum_{m=2}^{\nu-1} \left[(-1)^{m+1} \left(d_{\nu+1-m}^{\nu} \cdot \nu + d_{\nu-m}^{\nu} \right) \right] + (-1)^{\nu+1} d_1^{\nu} \cdot \nu$$

$$= (-1)^{\nu+1} d_1^{\nu} \cdot \nu + \sum_{m=2}^{\nu-1} \left[(-1)^{m+1} d_{\nu+1-m}^{\nu} \cdot \nu \right] + \sum_{m=2}^{\nu-1} \left[(-1)^{m+1} d_{\nu-m}^{\nu} \right]$$
$$+ (d_{\nu-1}^{\nu} + \nu)$$

$$= \sum_{m=2}^{\nu} \left[(-1)^{m+1} d_{\nu+1-m}^{\nu} \cdot \nu \right] + \sum_{m=1}^{\nu-1} \left[(-1)^{m+1} d_{\nu-m}^{\nu} \right] + \nu$$

$$= \sum_{m=1}^{\nu-1} \left[(-1)^{m+2} d_{\nu-m}^{\nu} \cdot \nu \right] + \sum_{m=1}^{\nu-1} \left[(-1)^{m+1} d_{\nu-m}^{\nu} \right] + \nu$$

$$= (-\nu) \cdot \sum_{m=1}^{\nu-1} \left[(-1)^{m+1} d_{\nu-m}^{\nu} \right] + \sum_{m=1}^{\nu-1} \left[(-1)^{m+1} d_{\nu-m}^{\nu} \right] + \nu$$

$$\overset{\text{(IH)}}{=} (-\nu) \cdot 1 + 1 + \nu = 1.$$

3) (BC) We know from part 1) that $d_2^3 = 3$. Therefore, we compute

$$d_2^3 = 3 = \sum_{j=1}^{2} j = \frac{2 \cdot 3}{2}.$$

(IH) The claim holds for some $\nu \in \mathbb{N}$ with $\nu \geq 3$.

(IS) $3 \leq \nu \rightarrow \nu + 1$

$$d_\nu^{\nu+1} \overset{1)}{=} d_{\nu-1}^\nu + \nu \overset{\text{(IH)}}{=} \left(\sum_{j=1}^{\nu-1} j \right) + \nu = \sum_{j=1}^{\nu} j = \frac{\nu(\nu+1)}{2}$$

by Gauss's formula for sums of consecutive integers[4].

$\square$

Remark 3.16.

Observe that the coefficients $d_{\nu-m}^\nu$ in Lemma 3.15, 1) are the Stirling numbers of the first kind[5]. This can be seen as follows:

The Stirling numbers of the first kind are denoted by $\begin{bmatrix} \nu \\ k \end{bmatrix}$. With the help of these numbers, we can write the factorial power $\ell(\ell-1)\cdot\ldots\cdot(\ell-\nu+1)$ as a linear combination of ordinary powers:

$$D_\nu(\ell) = \ell^\nu - \ell(\ell-1) \cdot \ldots \cdot (\ell - \nu + 1)$$

$$= \ell^\nu - \sum_{k=0}^{\nu} (-1)^{\nu-k} \begin{bmatrix} \nu \\ k \end{bmatrix} \ell^k, \tag{3.39}$$

see [40, p. 67, Eq. (44)]. Since we have $\begin{bmatrix} \nu \\ \nu \end{bmatrix} = 1$ and $\begin{bmatrix} \nu \\ 0 \end{bmatrix} = 0$, see [40, p. 67, Eq. (48)] and [40, p. 68, Eq. (50)] respectively, (3.39) leads to

$$D_\nu(\ell) = \ell^\nu - 1 \cdot \ell^\nu - \sum_{k=0}^{\nu-1} (-1)^{\nu-k} \begin{bmatrix} \nu \\ k \end{bmatrix} \ell^k = \sum_{k=0}^{\nu-1} (-1)^{\nu-k+1} \begin{bmatrix} \nu \\ k \end{bmatrix} \ell^k$$

$$= (-1)^{\nu+1} \cdot 0 + \sum_{k=1}^{\nu-1} (-1)^{\nu-k+1} \begin{bmatrix} \nu \\ k \end{bmatrix} \ell^k$$

$$= \sum_{m=1}^{\nu-1} (-1)^{m+1} \begin{bmatrix} \nu \\ \nu-m \end{bmatrix} \ell^{\nu-m}.$$

---

[4]For example, see [31, p. 59, 7.7].

[5]For the definition of the Stirling numbers of the first kind, see [1, p. 88] or [40, pp. 66–68], for instance.

The last equality is due to reversing the order of summation. Hence, $d^{\nu}_{\nu-m} = \begin{bmatrix} \nu \\ \nu-m \end{bmatrix}$, that is, the coefficients $d^{\nu}_{\nu-m}$ of D_{ν} in Lemma 3.15, 1) are the Stirling numbers of the first kind.

Now we have all required statements to validate (3.33).

Lemma 3.17.
Consider

$$S_{\nu} := \sum_{\ell=0}^{NR} \lambda_{\ell}\, \ell^{\nu}\, \mathrm{e}^{-\mathrm{i}\,h\ell T_j}, \quad \nu = 0, \ldots, R-1,$$

where λ_{ℓ}, $\ell = 0, \ldots, NR$, are the coefficients of the polynomial Λ which is defined in (3.31). The values T_j are the unknown shifts in (3.28). Then we can express the sum S_{ν}, $\nu = 0, \ldots, R-1$, by

$$S_{\nu} = \sum_{\eta=0}^{\nu} \alpha^{\nu}_{\eta} \Lambda^{(\eta)}(\mathrm{e}^{-\mathrm{i}\,hT_j})\, \mathrm{e}^{-\mathrm{i}\,\eta h T_j} \tag{3.40}$$

with coefficients $\alpha^{\nu}_{\eta} \in \mathbb{N}_0$, $\eta = 0, \ldots, \nu$, which fulfil the following conditions:

$$\alpha^0_0 = 1,$$
$$\alpha^{\nu}_0 = 0, \qquad\qquad\qquad \nu \geq 1,$$
$$\alpha^{\nu}_1 = \alpha^{\nu}_{\nu} = 1, \qquad\qquad\qquad \nu \geq 1,$$
$$\alpha^{\nu}_{\nu-1} = \sum_{j=1}^{\nu-1} j = \frac{(\nu-1)\nu}{2}, \qquad\qquad \nu \geq 2,$$
$$\alpha^{\nu}_{\eta} = \alpha^{\nu-1}_{\eta-1} + \eta \cdot \alpha^{\nu-1}_{\eta}, \qquad\qquad \eta = 2, \ldots, \nu-2 \ (\nu \geq 4).$$

Moreover, the coefficients α^{ν}_{η} have the recursive representation

$$\alpha^{\nu}_{\eta} = \sum_{m=1}^{\nu-\eta} (-1)^{m+1} d^{\nu}_{\nu-m} \alpha^{\nu-m}_{\eta}, \quad \eta = 1, \ldots, \nu-1 \ (\nu \geq 1),$$

where $d^{\nu}_{\nu-m}$ are the coefficients of D_{ν}, see Lemma 3.15.

Proof. We will prove the claim of Lemma 3.17 by induction. First, remember that

$$S_{\nu} = \sum_{\ell=0}^{NR} \lambda_{\ell}\, \ell^{\nu}\, \mathrm{e}^{-\mathrm{i}\,h\ell T_j}, \quad \Lambda(z) = \sum_{\ell=0}^{NR} \lambda_{\ell} z^{\ell}, \quad \Lambda^{(\eta)}(z) = \sum_{\ell=\eta}^{NR} \lambda_{\ell} \left(\prod_{k=0}^{\eta-1} (\ell-k) \right) z^{\ell-\eta}, \ \eta \in \mathbb{N}_0.$$

(BC)

$$S_0 = \sum_{\ell=0}^{NR} \lambda_\ell \, \mathrm{e}^{-\mathrm{i}h\ell T_j} = \Lambda\big(\mathrm{e}^{-\mathrm{i}hT_j}\big)$$

$$S_1 = \sum_{\ell=0}^{NR} \lambda_\ell \ell \, \mathrm{e}^{-\mathrm{i}h\ell T_j} = \mathrm{e}^{-\mathrm{i}hT_j} \sum_{\ell=0}^{NR} \lambda_\ell \ell \, \mathrm{e}^{-\mathrm{i}h(\ell-1)T_j}$$

$$= \mathrm{e}^{-\mathrm{i}hT_j} \sum_{\ell=1}^{NR} \lambda_\ell \ell \, \mathrm{e}^{-\mathrm{i}h(\ell-1)T_j} = \mathrm{e}^{-\mathrm{i}hT_j} \Lambda^{(1)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big)$$

$$S_2 = \sum_{\ell=0}^{NR} \lambda_\ell \ell^2 \, \mathrm{e}^{-\mathrm{i}h\ell T_j} = \sum_{\ell=0}^{NR} \lambda_\ell [\ell(\ell-1) + \ell] \, \mathrm{e}^{-\mathrm{i}h\ell T_j}$$

$$= \mathrm{e}^{-2\mathrm{i}hT_j} \sum_{\ell=2}^{NR} \lambda_\ell \ell(\ell-1) \, \mathrm{e}^{-\mathrm{i}h(\ell-2)T_j} + \mathrm{e}^{-\mathrm{i}hT_j} \sum_{\ell=1}^{NR} \lambda_\ell \ell \, \mathrm{e}^{-\mathrm{i}h(\ell-1)T_j}$$

$$= \mathrm{e}^{-2\mathrm{i}hT_j} \Lambda^{(2)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big) + \mathrm{e}^{-\mathrm{i}hT_j} \Lambda^{(1)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big)$$

$$S_3 = \sum_{\ell=0}^{NR} \lambda_\ell \ell^3 \, \mathrm{e}^{-\mathrm{i}h\ell T_j} = \sum_{\ell=0}^{NR} \lambda_\ell [\ell(\ell-1)(\ell-2) + 3\ell^2 - 2\ell] \, \mathrm{e}^{-\mathrm{i}h\ell T_j}$$

$$= \mathrm{e}^{-3\mathrm{i}hT_j} \sum_{\ell=3}^{NR} \lambda_\ell \ell(\ell-1)(\ell-2) \, \mathrm{e}^{-\mathrm{i}h(\ell-3)T_j} + 3 \sum_{\ell=0}^{NR} \lambda_\ell \ell^2 \, \mathrm{e}^{-\mathrm{i}hT_j} - 2 \sum_{\ell=0}^{NR} \lambda_\ell \ell \, \mathrm{e}^{-\mathrm{i}hT_j}$$

$$= \mathrm{e}^{-3\mathrm{i}hT_j} \Lambda^{(3)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big) + 3S_2 - 2S_1$$

$$= \mathrm{e}^{-3\mathrm{i}hT_j} \Lambda^{(3)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big) + 3\mathrm{e}^{-2\mathrm{i}hT_j} \Lambda^{(2)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big) + 3\mathrm{e}^{-\mathrm{i}hT_j} \Lambda^{(1)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big)$$

$$- 2\mathrm{e}^{-\mathrm{i}hT_j} \Lambda^{(1)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big)$$

$$= \mathrm{e}^{-3\mathrm{i}hT_j} \Lambda^{(3)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big) + 3\mathrm{e}^{-2\mathrm{i}hT_j} \Lambda^{(2)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big) + \mathrm{e}^{-\mathrm{i}hT_j} \Lambda^{(1)}\big(\mathrm{e}^{-\mathrm{i}hT_j}\big)$$

$$S_4 = \sum_{\ell=0}^{NR} \lambda_\ell \ell^4 \, \mathrm{e}^{-\mathrm{i}h\ell T_j} = \sum_{\ell=0}^{NR} \lambda_\ell [\ell(\ell-1)(\ell-2)(\ell-3) + 6\ell^3 - 11\ell^2 + 6\ell] \, \mathrm{e}^{-\mathrm{i}h\ell T_j}$$

$$= \mathrm{e}^{-4\mathrm{i}hT_j} \sum_{\ell=4}^{NR} \lambda_\ell \ell(\ell-1)(\ell-2)(\ell-3) \, \mathrm{e}^{-\mathrm{i}h(\ell-4)T_j} + 6 \sum_{\ell=0}^{NR} \lambda_\ell \lambda^3 \, \mathrm{e}^{-\mathrm{i}hT_j}$$

$$- 11 \sum_{\ell=0}^{NR} \lambda_\ell \ell^2 \, \mathrm{e}^{-\mathrm{i}hT_j} + 6 \sum_{\ell=0}^{NR} \lambda_\ell \ell \, \mathrm{e}^{-\mathrm{i}hT_j}$$

$$
\begin{aligned}
&= \mathrm{e}^{-4\,\mathrm{i}\,hT_j}\, \Lambda^{(4)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) + 6S_3 - 11S_2 + 6S_1 \\
&= \mathrm{e}^{-4\,\mathrm{i}\,hT_j}\, \Lambda^{(4)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) + 6\,\mathrm{e}^{-3\,\mathrm{i}\,hT_j}\, \Lambda^{(3)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) + 18\,\mathrm{e}^{-2\,\mathrm{i}\,hT_j}\, \Lambda^{(2)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) \\
&\qquad + 6\,\mathrm{e}^{-\,\mathrm{i}\,hT_j}\, \Lambda^{(1)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) - 11\,\mathrm{e}^{-2\,\mathrm{i}\,hT_j}\, \Lambda^{(2)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) \\
&\qquad\qquad - 11\,\mathrm{e}^{-\,\mathrm{i}\,hT_j}\, \Lambda^{(1)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) + 6\,\mathrm{e}^{-\,\mathrm{i}\,hT_j}\, \Lambda^{(1)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) \\
&= \mathrm{e}^{-4\,\mathrm{i}\,hT_j}\, \Lambda^{(4)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) + 6\,\mathrm{e}^{-3\,\mathrm{i}\,hT_j}\, \Lambda^{(3)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) + 7\,\mathrm{e}^{-2\,\mathrm{i}\,hT_j}\, \Lambda^{(2)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) \\
&\qquad + \mathrm{e}^{-\,\mathrm{i}\,hT_j}\, \Lambda^{(1)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right)
\end{aligned}
$$

The above computations yield the expression

$$
S_\nu = \sum_{\eta=0}^{\nu} \alpha_\eta^\nu \Lambda^{(\eta)}\!\left(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\right) \mathrm{e}^{-\,\mathrm{i}\,\eta hT_j}
$$

for $\nu = 0, 1, 2, 3, 4$ with the following coefficients:

$$
\alpha_0^0 = 1;
$$

$$
\alpha_0^1 = 0, \quad \alpha_1^1 = 1;
$$

$$
\alpha_0^2 = 0, \quad \alpha_1^2 = 1 = \sum_{j=1}^{1} j, \quad \alpha_2^2 = 1;
$$

$$
\alpha_0^3 = 0, \quad \alpha_1^3 = 1, \quad \alpha_2^3 = 3 = \sum_{j=1}^{2} j, \quad \alpha_3^3 = 1;
$$

$$
\alpha_0^4 = 0, \quad \alpha_1^4 = 1, \quad \alpha_2^4 = 7 = \alpha_2^3 \cdot 2 + \alpha_1^3, \quad \alpha_3^4 = 6 = \sum_{j=1}^{3} j, \quad \alpha_4^4 = 1.
$$

We know from Lemma 3.15 that

$$
\begin{aligned}
d_1^2 &= 1, & d_1^3 &= 2, & d_1^4 &= 6, \\
& & d_2^3 &= 3, & d_2^4 &= 11, \\
& & & & d_3^4 &= 6.
\end{aligned}
$$

Therefore, we compute the following:

$$
\alpha_1^2 = 1 = 1 \cdot 1 = d_1^2 \cdot \alpha_1^1 = \sum_{m=1}^{1} (-1)^{m+1} d_{2-m}^2 \alpha_1^{2-m}.
$$

$$\alpha_1^3 = 1 = 3 \cdot 1 - 2 \cdot 1 = d_2^3 \alpha_1^2 - d_1^3 \alpha_1^1 = \sum_{m=1}^{2} (-1)^{m+1} d_{3-m}^3 \alpha_1^{3-m},$$

$$\alpha_2^3 = 3 = 3 \cdot 1 = d_2^3 \alpha_2^2 = \sum_{m=1}^{1} (-1)^{m+1} d_{3-m}^3 \alpha_2^{3-m}.$$

$$\alpha_1^4 = 1 = 6 \cdot 1 - 11 \cdot 1 + 6 \cdot 1 = d_3^4 \alpha_1^3 - d_2^4 \alpha_1^2 + d_1^4 \alpha_1^1 = \sum_{m=1}^{3} (-1)^{m+1} d_{4-m}^4 \alpha_1^{4-m},$$

$$\alpha_2^4 = 7 = 6 \cdot 3 - 11 \cdot 1 = d_3^4 \alpha_2^3 - d_2^4 \alpha_2^2 = \sum_{m=1}^{2} (-1)^{m+1} d_{4-m}^4 \alpha_2^{4-m},$$

$$\alpha_3^4 = 6 = 6 \cdot 1 = d_3^4 \alpha_3^3 = \sum_{m=1}^{1} (-1)^{m+1} d_{4-m}^4 \alpha_3^{4-m}.$$

$$\Rightarrow \quad \alpha_\eta^\nu = \sum_{m=1}^{\nu-\eta} (-1)^{m+1} d_{\nu-m}^\nu \alpha_\eta^{\nu-m} \quad \text{for } \eta = 1, \dots, \nu-1, \text{ and } \nu = 2, 3, 4.$$

(IH) The claim holds for some $\nu \in \mathbb{N}$ with $\nu \geq 4$.

(IS) $4 \leq \nu \to \nu + 1$

$$
\begin{aligned}
S_{\nu+1} &= \sum_{\ell=0}^{NR} \lambda_\ell \ell^{\nu+1} \, \mathrm{e}^{-\mathrm{i}\, h\ell T_j} \\
&= \sum_{\ell=0}^{NR} \lambda_\ell \left(\prod_{k=0}^{\nu} (\ell - k) \right) \mathrm{e}^{-\mathrm{i}\, h\ell T_j} + \sum_{\ell=0}^{NR} \lambda_\ell \left[\ell^{\nu+1} - \left(\prod_{k=0}^{\nu} (\ell - k) \right) \right] \mathrm{e}^{-\mathrm{i}\, h\ell T_j} \\
&\overset{(3.38)}{=} \mathrm{e}^{-(\nu+1)\mathrm{i}\, hT_j} \sum_{\ell=\nu+1}^{NR} \lambda_\ell \left(\prod_{k=0}^{\nu} (\ell - k) \right) \mathrm{e}^{-\mathrm{i}\, h(\ell-\nu-1)T_j} + \sum_{\ell=0}^{NR} \lambda_\ell D_{\nu+1}(\ell) \, \mathrm{e}^{-\mathrm{i}\, h\ell T_j} \\
&\overset{3.15}{=} \mathrm{e}^{-(\nu+1)\mathrm{i}\, hT_j} \Lambda^{(\nu+1)} \left(\mathrm{e}^{-\mathrm{i}\, hT_j} \right) \\
&\quad + \sum_{\ell=0}^{NR} \lambda_\ell \left(\sum_{m=1}^{\nu} (-1)^{m+1} d_{\nu+1-m}^{\nu+1} \ell^{\nu+1-m} \right) \mathrm{e}^{-\mathrm{i}\, h\ell T_j} \\
&= \mathrm{e}^{-(\nu+1)\mathrm{i}\, hT_j} \Lambda^{(\nu+1)} \left(\mathrm{e}^{-\mathrm{i}\, hT_j} \right) \\
&\quad + \sum_{m=1}^{\nu} (-1)^{m+1} d_{\nu+1-m}^{\nu+1} \left(\sum_{\ell=0}^{NR} \lambda_\ell \ell^{\nu+1-m} \, \mathrm{e}^{-\mathrm{i}\, h\ell T_j} \right)
\end{aligned}
$$

$$
= \mathrm{e}^{-(\nu+1)\,\mathrm{i}\,hT_j}\,\Lambda^{(\nu+1)}\big(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\big) + \sum_{m=1}^{\nu}(-1)^{m+1}d^{\nu+1}_{\nu+1-m}S_{\nu+1-m}
$$

$$
\overset{(\mathrm{IH})}{=} \mathrm{e}^{-(\nu+1)\,\mathrm{i}\,hT_j}\,\Lambda^{(\nu+1)}\big(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\big)
$$

$$
+ \sum_{m=1}^{\nu}(-1)^{m+1}d^{\nu+1}_{\nu+1-m}\left(\sum_{\eta=0}^{\nu+1-m}\alpha^{\nu+1-m}_\eta\,\Lambda^{(\eta)}\big(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\big)\,\mathrm{e}^{-\,\mathrm{i}\,\eta hT_j}\right).
$$

Since $\nu + 1 - m \in \{1,\ldots,\nu\}$ for the index m with $m = 1,\ldots,\nu$, we have $\alpha^{\nu+1-m}_0 = 0$ for $m = 1,\ldots,\nu$ by the induction hypothesis. Therefore, we find the expression

$$
S_{\nu+1} = \mathrm{e}^{-(\nu+1)\,\mathrm{i}\,hT_j}\,\Lambda^{(\nu+1)}\big(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\big)
$$

$$
+ \sum_{m=1}^{\nu}(-1)^{m+1}d^{\nu+1}_{\nu+1-m}\left(\sum_{\eta=1}^{\nu+1-m}\alpha^{\nu+1-m}_\eta\,\Lambda^{(\eta)}\big(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\big)\,\mathrm{e}^{-\,\mathrm{i}\,\eta hT_j}\right)
$$

$$
= \mathrm{e}^{-(\nu+1)\,\mathrm{i}\,hT_j}\,\Lambda^{(\nu+1)}\big(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\big)
$$

$$
+ \sum_{\eta=1}^{\nu}\left(\sum_{m=1}^{\nu+1-\eta}(-1)^{m+1}d^{\nu+1}_{\nu+1-m}\alpha^{\nu+1-m}_\eta\right)\Lambda^{(\eta)}\big(\mathrm{e}^{-\,\mathrm{i}\,hT_j}\big)\,\mathrm{e}^{-\,\mathrm{i}\,\eta hT_j}. \tag{3.41}
$$

Note that the last equality is due to the fact that

$$
\sum_{m=1}^{\nu}\sum_{\eta=1}^{\nu+1-m} a_{m,\eta} = \sum_{\eta=1}^{\nu}\sum_{m=1}^{\nu+1-\eta} a_{m,\eta} \quad \text{for } a_{m,\eta} \in \mathbb{C}.
$$

Equation (3.41) now yields

$$
\alpha^{\nu+1}_0 = 0, \qquad \alpha^{\nu+1}_{\nu+1} = 1,
$$

and

$$
\alpha^{\nu+1}_\eta = \sum_{m=1}^{\nu+1-\eta}(-1)^{m+1}d^{\nu+1}_{\nu+1-m}\alpha^{\nu+1-m}_\eta, \quad \eta = 1,\ldots,\nu. \tag{3.42}
$$

For the coefficients $\alpha^{\nu+1}_1$ and $\alpha^{\nu+1}_\nu$, we find the following from (3.42):

$$
\alpha^{\nu+1}_1 = \sum_{m=1}^{\nu}(-1)^{m+1}d^{\nu+1}_{\nu+1-m}\alpha^{\nu+1-m}_1 \overset{(\mathrm{IH})}{=} \sum_{m=1}^{\nu}(-1)^{m+1}d^{\nu+1}_{\nu+1-m}\overset{\underset{\mathrm{2)}}{3.15}}{=} 1,
$$

$$
\alpha^{\nu+1}_\nu = \sum_{m=1}^{1}(-1)^{m+1}d^{\nu+1}_{\nu+1-m}\alpha^{\nu+1-m}_\nu = d^{\nu+1}_\nu\alpha^\nu_\nu \overset{(\mathrm{IH})}{=} d^{\nu+1}_\nu\overset{\underset{\mathrm{3)}}{3.15}}{=}\sum_{j=1}^{\nu} j = \frac{\nu(\nu+1)}{2}.
$$

Hence, it remains to prove the equality

$$\alpha_\eta^{\nu+1} = \sum_{m=1}^{\nu+1-\eta} (-1)^{m+1} d_{\nu+1-m}^{\nu+1} \alpha_\eta^{\nu+1-m} = \alpha_\eta^\nu \cdot \eta + \alpha_{\eta-1}^\nu$$

for $\eta = 2, \ldots, \nu - 1$. First, we will examine the cases $\eta = \nu - 2$ and $\eta = \nu - 1$ for $\nu + 1 \geq 5$.

$$\begin{aligned}
\alpha_{\nu-2}^{\nu+1} &= \sum_{m=1}^{3} (-1)^{m+1} d_{\nu+1-m}^{\nu+1} \alpha_\eta^{\nu+1-m} \\
&= d_\nu^{\nu+1} \alpha_{\nu-2}^\nu - d_{\nu-1}^{\nu+1} \alpha_{\nu-2}^{\nu-1} + d_{\nu-2}^{\nu+1} \alpha_{\nu-2}^{\nu-2} \\
&= \left(d_{\nu-1}^\nu + \nu\right) \alpha_{\nu-2}^\nu - \left[d_{\nu-1}^\nu \cdot \nu + d_{\nu-2}^\nu\right] \alpha_{\nu-2}^{\nu-1} + \left[d_{\nu-2}^\nu \cdot \nu + d_{\nu-3}^\nu\right] \alpha_{\nu-2}^{\nu-2}
\end{aligned}$$

by using part 1) of Lemma 3.15. The induction hypothesis about $\alpha_{\nu-2}^\nu$ now yields

$$\begin{aligned}
\alpha_{\nu-2}^{\nu+1} &= d_{\nu-1}^\nu \cdot \left[\alpha_{\nu-2}^{\nu-1} \cdot (\nu - 2) + \alpha_{\nu-3}^{\nu-1}\right] + \nu \alpha_{\nu-2}^\nu - \left[d_{\nu-1}^\nu \cdot \nu + d_{\nu-2}^\nu\right] \alpha_{\nu-2}^{\nu-1} \\
&\quad + \left[d_{\nu-2}^\nu \cdot \nu + d_{\nu-3}^\nu\right] \alpha_{\nu-2}^{\nu-2} \\
&= (\nu - 2) d_{\nu-1}^\nu \alpha_{\nu-2}^{\nu-1} + \nu \alpha_{\nu-2}^\nu + d_{\nu-1}^\nu \alpha_{\nu-3}^{\nu-1} - \nu d_{\nu-1}^\nu \alpha_{\nu-2}^{\nu-1} - d_{\nu-2}^\nu \alpha_{\nu-2}^{\nu-1} \\
&\quad + \nu d_{\nu-2}^\nu \alpha_{\nu-2}^{\nu-2} + d_{\nu-3}^\nu \alpha_{\nu-2}^{\nu-2} \\
&= (\nu - 2) \left(\sum_{m=1}^{1} (-1)^{m+1} d_{\nu-m}^\nu \alpha_{\nu-2}^{\nu-m}\right) + \nu \alpha_{\nu-2}^\nu + \left(\sum_{m=1}^{1} (-1)^{m+1} d_{\nu-m}^\nu \alpha_{\nu-3}^{\nu-m}\right) \\
&\quad - d_{\nu-2}^\nu \alpha_{\nu-2}^{\nu-1} + d_{\nu-3}^\nu \alpha_{\nu-2}^{\nu-2} - \nu \left(d_{\nu-1}^\nu \alpha_{\nu-2}^{\nu-1} - d_{\nu-2}^\nu \alpha_{\nu-2}^{\nu-2}\right).
\end{aligned}$$

Since $\alpha_{\nu-2}^{\nu-2} = 1 = \alpha_{\nu-3}^{\nu-3}$ and $\alpha_{\nu-2}^{\nu-1} = \frac{(\nu-2)(\nu-1)}{2}$ by the induction hypothesis, we get

$$\begin{aligned}
\alpha_{\nu-2}^{\nu+1} &= (\nu - 2) \left(\sum_{m=1}^{1} (-1)^{m+1} d_{\nu-m}^\nu \alpha_{\nu-2}^{\nu-m}\right) + \nu \alpha_{\nu-2}^\nu + \left(\sum_{m=1}^{1} (-1)^{m+1} d_{\nu-m}^\nu \alpha_{\nu-3}^{\nu-m}\right) \\
&\quad - d_{\nu-2}^\nu \frac{(\nu-2)(\nu-1)}{2} + d_{\nu-3}^\nu \alpha_{\nu-3}^{\nu-3} - \nu \underbrace{\left(\sum_{m=1}^{2} (-1)^{m+1} d_{\nu-m}^\nu \alpha_{\nu-2}^{\nu-m}\right)}_{=\alpha_{\nu-2}^\nu \text{ by (IH)}} \\
&= (\nu - 2) \left(\sum_{m=1}^{1} (-1)^{m+1} d_{\nu-m}^\nu \alpha_{\nu-2}^{\nu-m}\right) + \left(\sum_{m=1}^{1} (-1)^{m+1} d_{\nu-m}^\nu \alpha_{\nu-3}^{\nu-m}\right) \\
&\quad - d_{\nu-2}^\nu \frac{(\nu-2)(\nu-1)}{2} + d_{\nu-3}^\nu \alpha_{\nu-3}^{\nu-3}.
\end{aligned} \tag{3.43}$$

Further, we have

$$d_{\nu-2}^{\nu}\frac{(\nu-2)(\nu-1)}{2} = d_{\nu-2}^{\nu}\left[\frac{(\nu-3)(\nu-2)}{2}+(\nu-2)\right]$$

$$= (\nu-2)d_{\nu-2}^{\nu}\cdot 1 + d_{\nu-2}^{\nu}\frac{(\nu-3)(\nu-2)}{2}$$

$$\overset{\text{(IH)}}{=} (\nu-2)d_{\nu-2}^{\nu}\alpha_{\nu-2}^{\nu-2} + d_{\nu-2}^{\nu}\alpha_{\nu-3}^{\nu-2}.$$

Together with (3.43), this yields

$$\alpha_{\nu-2}^{\nu+1} = (\nu-2)\left(\sum_{m=1}^{1}(-1)^{m+1}d_{\nu-m}^{\nu}\alpha_{\nu-2}^{\nu-m}\right) - (\nu-2)d_{\nu-2}^{\nu}\alpha_{\nu-2}^{\nu-2}$$

$$+ \left(\sum_{m=1}^{1}(-1)^{m+1}d_{\nu-m}^{\nu}\alpha_{\nu-3}^{\nu-m}\right) - d_{\nu-2}^{\nu}\alpha_{\nu-3}^{\nu-2} + d_{\nu-3}^{\nu}\alpha_{\nu-3}^{\nu-3}$$

$$= (\nu-2)\left(\sum_{m=1}^{2}(-1)^{m+1}d_{\nu-m}^{\nu}\alpha_{\nu-2}^{\nu-m}\right) + \left(\sum_{m=1}^{3}(-1)^{m+1}d_{\nu-m}^{\nu}\alpha_{\nu-3}^{\nu-m}\right).$$

Incorporating the induction hypothesis, we finally get

$$\alpha_{\nu-2}^{\nu+1} = (\nu-2)\alpha_{\nu-2}^{\nu} + \alpha_{\nu-3}^{\nu}.$$

The proof of the case $\eta = \nu - 1$ for $\nu + 1 \geq 5$ also makes use of Lemma 3.15 and the induction hypothesis.

$$\alpha_{\nu-1}^{\nu+1} = \sum_{m=1}^{2}(-1)^{m+1}d_{\nu+1-m}^{\nu+1}\alpha_{\nu-1}^{\nu+1-m}$$

$$= d_{\nu}^{\nu+1}\alpha_{\nu-1}^{\nu} - d_{\nu-1}^{\nu+1}\alpha_{\nu-1}^{\nu-1}$$

$$= \left[d_{\nu-1}^{\nu}+\nu\right]\alpha_{\nu-1}^{\nu} - \left[d_{\nu-1}^{\nu}\cdot\nu + d_{\nu-2}^{\nu}\right]\alpha_{\nu-1}^{\nu-1} \qquad \text{by 3.15, 1)}$$

$$= (\nu-1+1)\alpha_{\nu-1}^{\nu} + \left[d_{\nu-1}^{\nu}\alpha_{\nu-1}^{\nu} - \nu d_{\nu-1}^{\nu}\alpha_{\nu-1}^{\nu-1}\right] - d_{\nu-1}^{\nu}\alpha_{\nu-1}^{\nu-1}$$

$$= (\nu-1)\alpha_{\nu-1}^{\nu} + \left[\alpha_{\nu-1}^{\nu} + d_{\nu-1}^{\nu}\alpha_{\nu-1}^{\nu} - \nu d_{\nu-1}^{\nu}\right] - d_{\nu-2}^{\nu}\alpha_{\nu-2}^{\nu-2} \qquad (3.44)$$

since we have the following by the induction hypothesis:

$$\alpha_{\nu-1}^{\nu-1} = 1 = \alpha_{\nu-2}^{\nu-2}.$$

We denote the middle term in (3.44) by A, i.e.

$$A := \alpha_{\nu-1}^{\nu} + d_{\nu-1}^{\nu}\alpha_{\nu-1}^{\nu} - \nu d_{\nu-1}^{\nu}.$$

For the expression A, we have the following:

$$
\begin{aligned}
A &= \frac{(\nu-1)\nu}{2} + \left[\frac{(\nu-1)\nu}{2}\right]^2 - \nu \cdot \frac{(\nu-1)\nu}{2} && \text{by (IH) and 3.15, 1)} \\
&= \left[1 - \nu + \frac{(\nu-1)\nu}{2}\right]\frac{(\nu-1)\nu}{2} \\
&= \frac{\nu^2 - 3\nu + 2}{2} \cdot \frac{(\nu-1)\nu}{2} \\
&= \frac{(\nu-2)(\nu-1)}{2} \cdot d_{\nu-1}^{\nu} && \text{by 3.15, 3)} \\
&= \alpha_{\nu-2}^{\nu-1} d_{\nu-1}^{\nu} && \text{by (IH).}
\end{aligned}
$$

Putting this into (3.44), we find

$$
\begin{aligned}
\alpha_{\nu-1}^{\nu+1} &= (\nu-1)\alpha_{\nu-1}^{\nu} + d_{\nu-1}^{\nu}\alpha_{\nu-2}^{\nu-1} - d_{\nu-2}^{\nu}\alpha_{\nu-2}^{\nu-2} \\
&= (\nu-1)\alpha_{\nu-1}^{\nu} + \sum_{m=1}^{2}(-1)^{m+1}d_{\nu-m}^{\nu}\alpha_{\nu-2}^{\nu-m} \\
&= (\nu-1)\alpha_{\nu-1}^{\nu} + \alpha_{\nu-2}^{\nu} \quad \text{by (IH).}
\end{aligned}
$$

The last remaining case is the case $\eta = 2, \ldots, \nu - 3$ for $\nu + 1 \geq 6$. In this case, we have the following for the coefficients $\alpha_{\eta}^{\nu+1}$:

$$
\begin{aligned}
\alpha_{\eta}^{\nu+1} &= \sum_{m=1}^{\nu+1-\eta}(-1)^{m+1}d_{\nu+1-m}^{\nu+1}\alpha_{\eta}^{\nu+1-m}, \qquad \eta = 2,\ldots,\nu-3 \ (\nu \geq 5), \\
&= d_{\nu}^{\nu+1}\alpha_{\eta}^{\nu} + \left(\sum_{m=2}^{\nu-\eta-1}(-1)^{m+1}d_{\nu+1-m}^{\nu+1}\alpha_{\eta}^{\nu+1-m}\right) \\
&\quad + (-1)^{\nu+1-\eta}d_{\eta+1}^{\nu+1}\alpha_{\eta}^{\eta+1} + (-1)^{\nu+2-\eta}d_{\eta}^{\nu+1}\alpha_{\eta}^{\eta} \\
&\overset{3.15}{=} \left[d_{\nu-1}^{\nu} + \nu\right]\alpha_{\eta}^{\nu} + \left(\sum_{m=2}^{\nu-\eta-1}(-1)^{m+1}\left[d_{\nu+1-m}^{\nu} \cdot \nu + d_{\nu-m}^{\nu}\right]\alpha_{\eta}^{\nu+1-m}\right) \\
&\quad + (-1)^{\nu+1-\eta}\left[d_{\eta+1}^{\nu} \cdot \nu + d_{\eta}^{\nu}\right]\alpha_{\eta}^{\eta+1} + (-1)^{\nu+2-\eta}\left[d_{\eta}^{\nu} \cdot \nu + d_{\eta-1}^{\nu}\right]\alpha_{\eta}^{\eta} \\
&\overset{\text{(IH)}}{=} d_{\nu-1}^{\nu}\left[\alpha_{\eta}^{\nu-1} \cdot \eta + \alpha_{\eta-1}^{\nu-1}\right] + \nu\alpha_{\eta}^{\nu} + \nu\left(\sum_{m=2}^{\nu-\eta-1}(-1)^{m+1}d_{\nu+1-m}^{\nu}\alpha_{\eta}^{\nu+1-m}\right) \\
&\quad + \left(\sum_{m=2}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^{\nu}\left[\alpha_{\eta}^{\nu-m} \cdot \eta + \alpha_{\eta-1}^{\nu-m}\right]\right) \\
&\quad + (-1)^{\nu+1-\eta}\left[d_{\eta+1}^{\nu} \cdot \nu + d_{\eta}^{\nu}\right]\alpha_{\eta}^{\eta+1} + (-1)^{\nu+2-\eta}\left[d_{\eta}^{\nu} \cdot \nu + d_{\eta-1}^{\nu}\right]\alpha_{\eta}^{\eta}
\end{aligned}
$$

$$= \nu\alpha_\eta^\nu + \nu\left(\sum_{m=2}^{\nu-\eta-1}(-1)^{m+1}d_{\nu+1-m}^\nu\alpha_\eta^{\nu+1-m}\right) + d_{\nu-1}^\nu\left[\alpha_\eta^{\nu-1}\cdot\eta + \alpha_{\eta-1}^{\nu-1}\right]$$
$$+ \left(\sum_{m=2}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^\nu\left[\alpha_\eta^{\nu-m}\cdot\eta + \alpha_{\eta-1}^{\nu-m}\right]\right)$$
$$+ (-1)^{\nu+1-\eta}\left[d_{\eta+1}^\nu\cdot\nu + d_\eta^\nu\right]\alpha_\eta^{\eta+1} + (-1)^{\nu+2-\eta}\left[d_\eta^\nu\cdot\nu + d_{\eta-1}^\nu\right]\alpha_\eta^\eta.$$

In the first sum, we use an index transformation, and we can add the term $d_{\nu-1}^\nu\left[\alpha_\eta^{\nu-1}\cdot\eta + \alpha_{\eta-1}^{\nu-1}\right]$ to the second sum. Thus, we obtain

$$\alpha_\eta^{\nu+1} = \nu\alpha_\eta^\nu + \nu\left(\sum_{m=1}^{\nu-\eta-2}(-1)^{m+2}d_{\nu-m}^\nu\alpha_\eta^{\nu-m}\right)$$
$$+ \left(\sum_{m=1}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^\nu\left[\alpha_\eta^{\nu-m}\cdot\eta + \alpha_{\eta-1}^{\nu-m}\right]\right)$$
$$+ (-1)^{\nu+1-\eta}\left[d_{\eta+1}^\nu\cdot\nu + d_\eta^\nu\right]\alpha_\eta^{\eta+1} + (-1)^{\nu+2-\eta}\left[d_\eta^\nu\cdot\nu + d_{\eta-1}^\nu\right]\alpha_\eta^\eta.$$

The terms $(-1)^{\nu+1-\eta}d_{\eta+1}^\nu\cdot\nu\cdot\alpha_\eta^{\eta+1}$ and $(-1)^{\nu+2-\eta}d_\eta^\nu\cdot\nu\cdot\alpha_\eta^\eta$ can be fit into the first sum. Further, we split the second sum into two sums. This results in

$$\alpha_\eta^{\nu+1} = \nu\alpha_\eta^\nu - \nu\left(\sum_{m=1}^{\nu-\eta}(-1)^{m+1}d_{\nu-m}^\nu\alpha_\eta^{\nu-m}\right) + \eta\left(\sum_{m=1}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^\nu\alpha_\eta^{\nu-m}\right)$$
$$+ \left(\sum_{m=1}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^\nu\alpha_{\eta-1}^{\nu-m}\right) + (-1)^{\nu+1-\eta}\left[d_\eta^\nu\alpha_\eta^{\eta+1} - d_{\eta-1}^\nu\alpha_\eta^\eta\right].$$

Using the induction hypothesis, we find that

$$\left(\sum_{m=1}^{\nu-\eta}(-1)^{m+1}d_{\nu-m}^\nu\alpha_\eta^{\nu-m}\right) = \alpha_\eta^\nu$$

and

$$\alpha_\eta^{\eta+1} = \frac{\eta(\eta+1)}{2}, \qquad \alpha_\eta^\eta = \alpha_{\eta-1}^{\eta-1}.$$

Hence, we get

$$\alpha_\eta^{\nu+1} \overset{\text{(IH)}}{=} \nu\alpha_\eta^\nu - \nu\alpha_\eta^\nu + \eta\left(\sum_{m=1}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^\nu\alpha_\eta^{\nu-m}\right)$$
$$+ \left(\sum_{m=1}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^\nu\alpha_{\eta-1}^{\nu-m}\right) + (-1)^{\nu+1-\eta}d_\eta^\nu\frac{\eta(\eta+1)}{2} \qquad (3.45)$$
$$+ (-1)^{\nu+2-\eta}d_{\eta-1}^\nu\alpha_{\eta-1}^{\eta-1}.$$

By the induction hypothesis, we find

$$\frac{\eta(\eta+1)}{2} = \eta \cdot 1 + \frac{(\eta-1)\eta}{2} = \eta\alpha_\eta^\eta + \alpha_{\eta-1}^\eta.$$

Thus, it follows from (3.45) that

$$\alpha_\eta^{\nu+1} = \eta\left(\sum_{m=1}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^\nu\alpha_\eta^{\nu-m}\right) + \eta\cdot(-1)^{\nu+1-\eta}d_\eta^\nu\alpha_\eta^\eta$$
$$+ \left(\sum_{m=1}^{\nu-\eta-1}(-1)^{m+1}d_{\nu-m}^\nu\alpha_{\eta-1}^{\nu-m}\right) + (-1)^{\nu+1-\eta}d_\eta^\nu\alpha_{\eta-1}^\eta$$
$$+ (-1)^{\nu+2-\eta}d_{\eta-1}^\nu\alpha_{\eta-1}^{\eta-1}.$$

By adding the term $\eta\cdot(-1)^{\nu+1-\eta}d_\eta^\nu\alpha_\eta^\eta$ to the first sum, and by fitting the terms $(-1)^{\nu+1-\eta}d_\eta^\nu\alpha_{\eta-1}^\eta$ and $(-1)^{\nu+2-\eta}d_{\eta-1}^\nu\alpha_{\eta-1}^{\eta-1}$ into the second sum, we obtain

$$\alpha_\eta^{\nu+1} = \eta\left(\sum_{m=1}^{\nu-\eta}(-1)^{m+1}d_{\nu-m}^\nu\alpha_\eta^{\nu-m}\right) + \left(\sum_{m=1}^{\nu-\eta+1}(-1)^{m+1}d_{\nu-m}^\nu\alpha_{\eta-1}^{\nu-m}\right)$$
$$= \eta\alpha_\eta^\nu + \alpha_{\eta-1}^\nu$$

where the last equality is due to the induction hypothesis.

$\square$

Remark 3.18.
Note that the coefficients α_η^ν in Lemma 3.17 fulfil the initial conditions and the recurrence relation for the Stirling numbers of the second kind; for instance, see [1, p. 89] or [40, pp. 66–68]. Thus, the coefficients used in the representation (3.40) are Stirling numbers of the second kind.

3.4. Numerical results

In this section, we want to apply the described reconstruction methods to examples of step functions and non-uniform spline functions using simulated Fourier data.

First, Figure 3.3 presents a step function consisting of eight steps, i.e. seven jumps inside its support, that is determined by the knot sequence $(T_j)_{j=1}^9$ and the coefficient sequence $(c_j^0)_{j=1}^8$ given in Table 3.2. Note that this is actually a special case of non-uniform spline functions where the order of the B-splines is set to one. Further, observe that the considered example includes several difficulties. On the one hand, some knots are very close to each other, namely T_1 and T_2, T_6 and T_7, and the last two knots

T_8 and T_9, such that these jumps are barely visible. On the other hand, there are some successive coefficients with a rather small difference, namely c_3^0 and c_4^0, and c_7^0 and c_8^0. In order to show the exactness of the reconstruction, we give the absolute reconstruction errors $|T_j - T_j^*|$ and $|c_j^0 - c_j^*|$, see Table 3.2. Here, T_j^* and c_j^* denote the reconstructed knots and coefficients respectively.
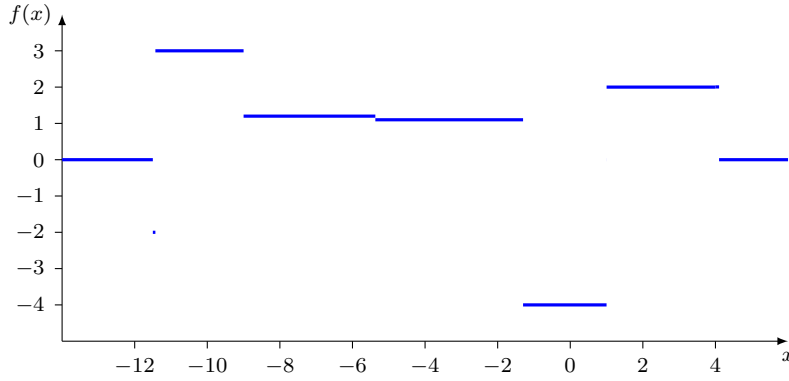
Figure 3.3. *Original function of the form* (3.1), *determined by the knot sequence* $(T_j)_{j=1}^9$ *and the coefficient sequence* $(c_j^0)_{j=1}^8$ *given in* Table 3.2.

j	T_j	$\lvert T_j - T_j^* \rvert$	c_j^0	$\lvert c_j^0 - c_j^* \rvert$
1	-11.5	$7.052 \cdot 10^{-13}$	-2	$3.962 \cdot 10^{-11}$
2	-11.43	$2.753 \cdot 10^{-13}$	3	$1.776 \cdot 10^{-14}$
3	-9	$3.553 \cdot 10^{-15}$	1.2	$1.998 \cdot 10^{-14}$
4	-5.37	$6.217 \cdot 10^{-15}$	1.1	$2.043 \cdot 10^{-14}$
5	-1.3	$4.885 \cdot 10^{-15}$	-4	$6.661 \cdot 10^{-15}$
6	1	$7.172 \cdot 10^{-9}$	0	$5.73 \cdot 10^{-5}$
7	1.001	$1.43 \cdot 10^{-8}$	2	$1.559 \cdot 10^{-13}$
8	4	$3.622 \cdot 10^{-11}$	2.005	$1.828 \cdot 10^{-12}$
9	4.1	$5.329 \cdot 10^{-15}$		

Table 3.2. *Parameters of the original function in* Figure 3.3 *and approximate reconstruction errors. The sampling step size used is* $h = 0.27$.

In the next two examples, see Figure 3.4 and 3.5, we consider the recovery of non-uniform spline functions of higher order, namely of order $m = 2$ and $m = 5$. The original parameters T_j and c_j^0 are listed in Tables 3.3 and 3.4, and we also compare them with the reconstructed values T_j^* and c_j^*.
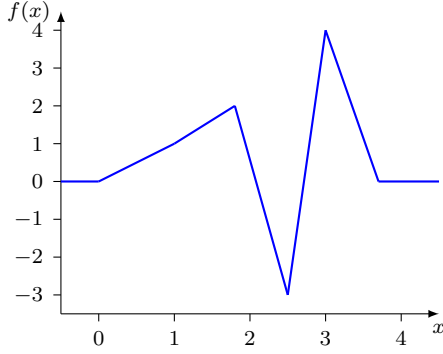
Figure 3.4. *Original function f of the form (3.10) where $m = 2$, determined by $(T_j)_{j=1}^6$ and $(c_j^0)_{j=1}^4$ given in Table 3.3.*

Figure 3.5. *Original function f of the form (3.10) where $m = 5$, determined by $(T_j)_{j=1}^{10}$ and $(c_j^0)_{j=1}^5$ given in Table 3.4.*

j	T_j	$|T_j - T_j^*|$	c_j^0	$|c_j^0 - c_j^*|$
1	0	0	1	$3.504 \cdot 10^{-12}$
2	1	$3.55 \cdot 10^{-12}$	2	$3.064 \cdot 10^{-12}$
3	1.8	$2.971 \cdot 10^{-13}$	-3	$1.06 \cdot 10^{-12}$
4	2.5	$2.132 \cdot 10^{-13}$	4	$1.678 \cdot 10^{-12}$
5	3	$2.665 \cdot 10^{-15}$		
6	3.7	$1.821 \cdot 10^{-14}$		

Table 3.3. *Parameters of the original function f in Figure 3.4 and approximate reconstruction errors ($h = 0.8$).*

j	T_j	$|T_j - T_j^*|$	c_j^0	$|c_j^0 - c_j^*|$
1	-6	0	-3.2	$8.66 \cdot 10^{-14}$
2	-5.8	$3.553 \cdot 10^{-15}$	3.1	$2.576 \cdot 10^{-14}$
3	-4	$8.882 \cdot 10^{-16}$	-0.8	$7.996 \cdot 10^{-13}$
4	-2.25	$1.332 \cdot 10^{-15}$	1.5	$2.783 \cdot 10^{-12}$
5	-0.6	$2.887 \cdot 10^{-15}$	-3	$5.799 \cdot 10^{-12}$
6	0	$1.337 \cdot 10^{-15}$		
7	1.3	$3.109 \cdot 10^{-15}$		
8	2.73	$4.441 \cdot 10^{-16}$		
9	3.5	$4.441 \cdot 10^{-15}$		
10	4.2	$4.441 \cdot 10^{-15}$		

Table 3.4. *Parameters of the original function f in Figure 3.5 and approximate reconstruction errors ($h = 0.5$).*

As we have mentioned in Remarks 3.9, 2., the reconstruction procedure for a non-uniform spline function f always results in the simplest representation of f such that the reconstructed representation of f of the form (3.21) does not possess redundant terms. We want to illustrate this by the next example.

Consider the non-uniform spline function f of order $m = 2$ given by

$$f(x) := \sum_{j=1}^4 c_j^{\mathrm{orig}} B_j^2, \quad x \in \mathbb{R}, \tag{3.46}$$

which is determined by the coefficients

$$c_1^{\mathrm{orig}} := 1, \quad c_2^{\mathrm{orig}} := 2, \quad c_3^{\mathrm{orig}} := 3, \quad c_4^{\mathrm{orig}} := 4, \tag{3.47}$$

and the knots

$$T_1^{\mathrm{orig}} := 1, \quad T_2^{\mathrm{orig}} := 2, \quad T_3^{\mathrm{orig}} := 3, \quad T_4^{\mathrm{orig}} := 4.5, \quad T_5^{\mathrm{orig}} := 5, \quad T_6^{\mathrm{orig}} := 6. \tag{3.48}$$

The representation in (3.46) with the coefficients and knots given in (3.47) and (3.48) is redundant. Using the definition of the B-splines and the recurrence relation, i.e. (3.11) and (3.12), we conclude that the function f considered above is the same as the function given in Figure 3.6, which is determined by the knot sequence $(T_j)_{j=1}^5$ and the coefficient sequence $(c_j)_{j=1}^3$ with

$$T_1 := 1, \quad T_2 := 3, \quad T_3 := 4.5, \quad T_4 := 5, \quad T_5 := 6, \tag{3.49}$$

and

$$c_1 := 2, \quad c_2 := 3, \quad c_3 := 4. \tag{3.50}$$

These parameters are approximately obtained by applying the reconstruction scheme for non-uniform spline functions, i.e. Algorithm 3.10, to sampling values of the Fourier transform of the function f in (3.46), see Table 3.5.
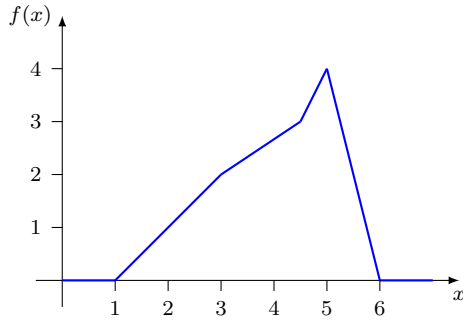
Figure 3.6. *Original function f of the form (3.10) where $m = 2$, determined by the knot sequence $(T_j)_{j=1}^5$ in (3.49) and the coefficient sequence $(c_j^0)_{j=1}^4$ in (3.50).*

j	$\|T_j - T_j^*\|$	$\|c_j - c_j^*\|$
1	$9.548 \cdot 10^{-15}$	$1.51 \cdot 10^{-14}$
2	$4.974 \cdot 10^{-14}$	$1.021 \cdot 10^{-13}$
3	$1.67 \cdot 10^{-13}$	$5.818 \cdot 10^{-14}$
4	$5.329 \cdot 10^{-14}$	
5	$4.441 \cdot 10^{-15}$	

Table 3.5. *Reconstruction results for f in (3.46). Note that the reconstruction of f does not possess redundant terms. The errors are obtained by comparison with the parameters in (3.49) and (3.50).*

4. Recovery of special bivariate functions from sparse Fourier samples

In this chapter, we want to extend the theory from the previous chapter to the bivariate case. On the one hand, we generalize the results for the reconstruction of non-uniform spline functions and non-uniform translates in the univariate case to tensor-products of such functions. On the other hand, we consider linear combinations of non-uniform translates of bivariate functions where the variables are not separable.

4.1. Tensor-products

4.1.1. Tensor-products of non-uniform spline functions

The first case of tensor-products where we want to give reconstruction results are non-uniform tensor-product spline representations of the form

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x_1, x_2) := \sum_{j=1}^{N_1} \sum_{k=1}^{N_2} c_{j,k}^{0,0} B_j^{m_1}(x_1) B_k^{m_2}(x_2) \tag{4.1}$$

with real-valued coefficients $c_{j,k}^{0,0} \neq 0$ for $j = 1, \ldots, N_1$, and $k = 1, \ldots, N_2$. As introduced in Section 3.2, $B_j^{m_1}$ and $B_k^{m_2}$ are B-splines of order m_1 and m_2 which are determined by the knot sequences $(T_j, \ldots, T_{j+m_1})$ and $(S_k, \ldots, S_{k+m_2})$ respectively.

As done similarly in the univariate case of non-uniform spline functions, we want to recover functions f of the form (4.1) from sparse Fourier samples, i.e. from as few Fourier samples as possible. Therefore, we need a representation of the Fourier transform of f. The approach to compute $\widehat{f}$ follows similar lines as the calculations in Section 3.2.

First, we can consider $f(\cdot, x_2)$ for all $x_2 \in \mathbb{R}$ as a univariate non-uniform spline function. In the case $m_1 \geq 3$, Lemma 3.4 yields

$$\frac{\partial^{r_1}}{\partial x_1^{r_1}} f(x_1, x_2) = \sum_{j=1}^{N_1+r_1} \sum_{k=1}^{N_2} c_{j,k}^{r_1,0} B_j^{m_1-r_1}(x_1) B_k^{m_2}(x_2)$$

for $r_1 = 1, \ldots, m_1 - 2$. The coefficients $c_{j,k}^{r_1,0}$ are defined by

$$c_{j,k}^{r_1,0} := \left(\frac{m_1 - r_1}{T_{j+m_1-r_1} - T_j} \right) \cdot \left(c_{j,k}^{r_1-1,0} - c_{j-1,k}^{r_1-1,0} \right), \qquad \begin{matrix} j = 1, \ldots, N_1 + r_1, \\ k = 1, \ldots, N_2, \end{matrix} \tag{4.2}$$

with the convention

$$c_{0,k}^{r_1-1,0} = c_{N_1+r_1,k}^{r_1-1,0} = 0, \qquad k = 1, \ldots, N_2. \tag{4.3}$$

Secondly, $\frac{\partial^{m_1-2}}{\partial x_1^{m_1-2}} f(x_1, \cdot)$ is a univariate non-uniform spline function for all $x_1 \in \mathbb{R}$. Therefore, we get the following for $m_2 \geq 3$ and $r_2 = 1, \ldots, m_2 - 2$ by Lemma 3.4:

$$\frac{\partial^{r_2}}{\partial x_2^{r_2}} \left(\frac{\partial^{m_1-2}}{\partial x_1^{m_1-2}} f(x_1, x_2) \right) = \sum_{j=1}^{N_1+m_1-2} \sum_{k=1}^{N_2+r_2} c_{j,k}^{m_1-2,r_2} B_j^2(x_1) B_k^{m_2-r_2}(x_2)$$

with the coefficients

$$c_{j,k}^{m_1-2,r_2} := \left(\frac{m_2 - r_2}{S_{k+m_2-r_2} - S_k} \right) \cdot \left(c_{j,k}^{m_1-2,r_2-1} - c_{j,k-1}^{m_1-2,r_2-1} \right) \tag{4.4}$$

for $k = 1, \ldots, N_2 + r_2$, and $j = 1, \ldots, N_1 + m_1 - 2$ where we have

$$c_{j,0}^{m_1-2,r_2-1} = c_{j,N_2+r_2}^{m_1-2,r_2-1} = 0, \qquad j = 1, \ldots, N_1 + m_1 - 2. \tag{4.5}$$

As an analogue to the univariate case, we have to work with distributions, generalized functions, and the Fourier transform of a distribution if $m_1, m_2 < 3$. For this purpose, the approach from the univariate case transfers to the bivariate case, which is explained in Appendix D and results in the following three lemmas.

Lemma 4.1.
Consider a generalized function f of the form (4.1) with $m_1, m_2 \geq 2$. The distributional derivative $\mathrm{D}^{(m_1-1,m_2-1)} \mu_{(f)}$, where $(m_1 - 1, m_2 - 1)$ denotes a multi-index, is given by the generalized function

$$f^{(m_1-1,m_2-1)} : \mathbb{R}^2 \ni (x_1, x_2)^{\mathrm{T}} \mapsto \sum_{j=1}^{N_1+m_1-1} \sum_{k=1}^{N_2+m_2-1} c_{j,k}^{m_1-1,m_2-1} B_j^1(x_1) B_k^1(x_2) \in \mathbb{R}.$$

The coefficients $c_{j,k}^{m_1-1,m_2-1}$ are defined by the following formulae:

$$c_{j,k}^{m_1-1,m_2-2} := \frac{c_{j,k}^{m_1-2,m_2-2} - c_{j-1,k}^{m_1-2,m_2-2}}{T_{j+1} - T_j}, \qquad \begin{matrix} j = 1, \ldots, N_1 + m_1 - 1, \\ k = 1, \ldots, N_2 + m_2 - 2, \end{matrix} \tag{4.6}$$

and

$$c_{j,k}^{m_1-1,m_2-1} := \frac{c_{j,k}^{m_1-1,m_2-2} - c_{j,k-1}^{m_1-1,m_2-2}}{S_{k+1} - S_k}, \qquad \begin{array}{l} k = 1, \dots, N_2 + m_2 - 1, \\ j = 1, \dots, N_1 + m_1 - 1, \end{array} \qquad (4.7)$$

where we use the conventions

$$c_{0,k}^{m_1-2,m_2-2} = c_{N_1-m_1-1,k}^{m_1-2,m_2-2} = 0, \qquad k = 1, \dots, N_2 + m_2 - 2, \qquad (4.8)$$

$$c_{j,0}^{m_1-1,m_2-2} = c_{j,N_2+m_2-1}^{m_1-1,m_2-2} = 0, \qquad j = 1, \dots, N_1 + m_1 - 1. \qquad (4.9)$$

Lemma 4.2.
The distributional derivative $\mathrm{D}^{(m_1,m_2)} \mu_{(f)}$ *of the generalized function f of the form* (4.1) *with $m_1, m_2 \geq 1$ is a linear combination of translated Dirac delta distributions. This means that* $\mathrm{D}^{(m_1,m_2)} \mu_{(f)}$ *is given by the distribution*

$$f^{(m_1,m_2)} := \mathrm{D}^{(m_1,m_2)} \mu_{(f)} = \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \big(\tau_{T_j,S_k} \delta \big)$$

where $\tau_{T_j,S_k} \delta$ *denotes the translation of the bivariate Dirac delta distribution[1]. The coefficients* $c_{j,k}^{m_1,m_2}$ *are defined by the following formulae:*

$$c_{j,k}^{m_1,m_2-1} := c_{j,k}^{m_1-1,m_2-1} - c_{j-1,k}^{m_1-1,m_2-1}, \qquad \begin{array}{l} j = 1, \dots, N_1 + m_1, \\ k = 1, \dots, N_2 + m_2 - 1, \end{array} \qquad (4.10)$$

where we use

$$c_{0,k}^{m_1-1,m_2-1} = c_{N_1+m_1,k}^{m_1-1,m_2-1} = 0, \qquad k = 1, \dots, N_2 + m_2 - 1; \qquad (4.11)$$

and

$$c_{j,k}^{m_1,m_2} := c_{j,k}^{m_1,m_2-1} - c_{j,k-1}^{m_1,m_2-1}, \qquad \begin{array}{l} k = 1, \dots, N_2 + m_2, \\ j = 1, \dots, N_1 + m_1, \end{array} \qquad (4.12)$$

with the convention

$$c_{j,0}^{m_1,m_2-1} = c_{j,N_2+m_2}^{m_1,m_2-1} = 0, \qquad j = 1, \dots, N_1 + m_1. \qquad (4.13)$$

---

[1]See (D.15) and (D.14) for the definitions of the bivariate Dirac delta distribution and the translate of a distribution in the bivariate case respectively.

Having given the distributional derivative $\mathrm{D}^{(m_1,m_2)}\mu_{(f)}$ of the function f considered in (4.1), we can apply the Fourier transform in order to get a representation of $\widehat{f}$.

Lemma 4.3.
The Fourier transform $\widehat{f}$ of the function f in (4.1) is represented by

$$(\mathrm{i}\,\omega_1)^{m_1}(\mathrm{i}\,\omega_2)^{m_2}\widehat{f}(\omega_1,\omega_2) = \sum_{j=1}^{N_1+m_1}\sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2}\, \mathrm{e}^{-\,\mathrm{i}\cdot(\omega_1 T_j+\omega_2 S_k)}$$

for $\boldsymbol{\omega}:=(\omega_1,\omega_2)^{\mathrm{T}}\in\mathbb{R}^2$ where the coefficients $c_{j,k}^{m_1,m_2}$ are defined by (4.10)–(4.13).

Knowing the structure of the Fourier transform of the non-uniform tensor-product spline representation in (4.1), we can now derive a theory for the reconstruction of such tensor-products from very few Fourier samples.

Theorem 4.4 (Reconstruction of tensor-product spline functions).
Let $m_1,m_2\in\mathbb{N}$ be given natural numbers, and let f be a real-valued, bivariate spline function of the form (4.1) with knot sequences $-\infty < T_1 < \ldots < T_{N_1+m_1} < \infty$ and $-\infty < S_1 < \ldots < S_{N_2+m_2} < \infty$, and real coefficients $c_{j,k}^{0,0}\neq 0$ for $j=1,\ldots,N_1$, and $k=1,\ldots,N_2$. Further, let h_1 and h_2 be two positive constants, and assume that they satisfy $h_1 T_j, h_2 S_k \in (-\pi,\pi]$ for all $j\in\{1,\ldots,N_1+m_1\}$ and $k\in\{1,\ldots,N_2+m_2\}$ respectively. Then f can be uniquely recovered from the $2(N_1+m_1)\cdot(N_2+m_2)$ Fourier samples $\widehat{f}(\eta h_1, \nu h_2)$ where $\eta = 1,\ldots,2(N_1+m_1)$ and $\nu=1,\ldots,N_2+m_2$.

Proof. Lemma 4.3 provides the following representation of the Fourier transform $\widehat{f}$:

$$
\begin{aligned}
(\mathrm{i}\,\omega_1)^{m_1}(\mathrm{i}\,\omega_2)^{m_2}\widehat{f}(\omega_1,\omega_2) &= \sum_{j=1}^{N_1+m_1}\sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2}\, \mathrm{e}^{-\,\mathrm{i}\cdot(\omega_1 T_j+\omega_2 S_k)}\\
&= \sum_{j=1}^{N_1+m_1}\left(\sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2}\, \mathrm{e}^{-\,\mathrm{i}\,\omega_2 S_k}\right)\mathrm{e}^{-\,\mathrm{i}\,\omega_1 T_j} \qquad (4.14)
\end{aligned}
$$

where the coefficients $c_{j,k}^{m_1,m_2}$ are defined by (4.10)–(4.13).

We have here an exponential sum. But we cannot directly apply the Prony method from Chapter 2 to this exponential sum since it as a bivariate one. In order to use the Prony method, we need a univariate function structure. Thus, in a first step, we restrict the function in (4.14) to the variable ω_1 in order to obtain the knots $T_1,\ldots,T_{N_1+m_1}$. Then we fix the variable ω_1 such that we get a univariate problem with respect to the variable ω_2. In this way, we can determine the knots $S_1,\ldots,S_{N_2+m_2}$, and all coefficients $c_{j,k}^{m_1,m_2}$ in (4.14). In the last step, we compute the original coefficients $c_{j,k}^{0,0}$, $j=1,\ldots,N_1$, $k=1,\ldots,N_2$, by using recursion formulae.

Part 1: First, we set

$$p_j(\omega_2) := \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} e^{-i\omega_2 S_k} \qquad \text{for } j = 1, \ldots, N_1 + m_1. \tag{4.15}$$

Then the equality (4.14) reads as

$$(i\,\omega_1)^{m_1}(i\,\omega_2)^{m_2}\widehat{f}(\omega_1,\omega_2) = \sum_{j=1}^{N_1+m_1} p_j(\omega_2)\,e^{-i\omega_1 T_j}. \tag{4.16}$$

We fix $\omega_2 := h_2$ and hence obtain a univariate function P with respect to the variable ω_1 which is defined by

$$P(\omega_1) := (i\,\omega_1)^{m_1}(i\,h_2)^{m_2}\widehat{f}(\omega_1,h_2) = \sum_{j=1}^{N_1+m_1} p_j(h_2)\,e^{-i\omega_1 T_j}. \tag{4.17}$$

In this univariate exponential sum, the coefficients $p_j(h_2)$ are complex-valued, see the definition of $p_j(h_2)$ in (4.15). Therefore, we apply the Prony method which is presented in Section 2.1 to the function P where we use the Fourier samples $\widehat{f}(\eta h_1, h_2)$ for $\eta = 1, \ldots, 2(N_1+m_1)$. Using these sampling values of $\widehat{f}$, we can compute the function values $P(\eta h_1)$ for $\eta = 1, \ldots, 2(N_1 + m_1)$, which we actually need for the application of the Prony method, by

$$P(\eta h_1) = (i\,\eta h_1)^{m_1}(i\,h_2)^{m_2}\widehat{f}(\eta h_1, h_2) \quad \text{for } \eta = 1, \ldots, 2(N_1 + m_1).$$

Observe that the still needed function value $P(0)$ is given by this computation too, i.e. $P(0) = 0$, since the equation (4.17) also holds for $\omega = 0$. This can be seen as follows: First, the left-hand side of (4.17), namely $P(\omega_1) = (i\,\omega_1)^{m_1}(i\,h_2)^{m_2}\widehat{f}(\omega_1, h_2)$, equals to zero for $\omega_1 = 0$ because we have $\left|\widehat{f}(0, h_2)\right| < \infty$ by definition of f as an almost everywhere continuous function with compact support. Secondly, also the right-hand side of (4.17) yields the value 0 if ω_1 is set to zero:

$$\sum_{j=1}^{N_1+m_1} p_j(h_2)\,e^{-i\cdot 0\cdot T_j} = \sum_{j=1}^{N_1+m_1}\sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2}\,e^{-i\omega_2 S_k}$$

$$= \sum_{k=1}^{N_2+m_2}\left(\sum_{j=1}^{N_1+m_1} c_{j,k}^{m_1,m_2}\right)e^{-i\omega_2 S_k}$$

$$\overset{(4.12)}{=} \sum_{k=1}^{N_2+m_2}\left(\sum_{j=1}^{N_1+m_1}\left(c_{j,k}^{m_1,m_2-1} - c_{j,k-1}^{m_1,m_2-1}\right)\right)e^{-i\omega_2 S_k}$$

$$\overset{(4.13)}{=} \sum_{k=1}^{N_2+m_2-1} \left(\sum_{j=1}^{N_1+m_1} c_{j,k}^{m_1,m_2-1} \right) \mathrm{e}^{-\mathrm{i}\,\omega_2 S_k}$$

$$- \sum_{k=2}^{N_2+m_2} \left(\sum_{j=1}^{N_1+m_1} c_{j,k-1}^{m_1,m_2-1} \right) \mathrm{e}^{-\mathrm{i}\,\omega_2 S_k}$$

$$= \sum_{k=1}^{N_2+m_2-1} \left(\sum_{j=1}^{N_1+m_1} c_{j,k}^{m_1,m_2-1} \right) \mathrm{e}^{-\mathrm{i}\,\omega_2 S_k}$$

$$- \sum_{k=1}^{N_2+m_2-1} \left(\sum_{j=1}^{N_1+m_1} c_{j,k}^{m_1,m_2-1} \right) \mathrm{e}^{-\mathrm{i}\,\omega_2 S_{k+1}}$$

$$= \sum_{k=1}^{N_2+m_2-1} \left(\sum_{j=1}^{N_1+m_1} c_{j,k}^{m_1,m_2-1} \right) \left[\mathrm{e}^{-\mathrm{i}\,\omega_2 S_k} - \mathrm{e}^{-\mathrm{i}\,\omega_2 S_{k+1}} \right]. \quad (4.18)$$

Further, we have

$$\sum_{j=1}^{N_1+m_1} c_{j,k}^{m_1,m_2-1} \overset{(4.10)}{=} \sum_{j=1}^{N_1+m_1} \left[c_{j,k}^{m_1-1,m_2-1} - c_{j-1,k}^{m_1-1,m_2-1} \right]$$

$$= c_{N_1+m_1,k}^{m_1-1,m_2-1} - c_{0,k}^{m_1-1,m_2-1}$$

$$\overset{(4.11)}{=} 0.$$

We insert this result into (4.18), which leads to

$$\sum_{j=1}^{N_1+m_1} p_j(h_2)\, \mathrm{e}^{-\mathrm{i}\cdot 0 \cdot T_j} = 0.$$

Thus, we obtain the knot sequence $(T_1, \ldots, T_{N_1+m_1})$ as well as the coefficients $p_j(h_2)$, $j = 1, \ldots, N_1 + m_1$, by using the Fourier samples $\widehat{f}(\eta h_1, h_2)$ for $\eta = 1, \ldots, 2(N_1 + m_1)$.

In the unlucky case that not all values $p_j(h_2)$ for $j \in \{1, \ldots, N_1 + m_1\}$ are non-zero, we will not find all parameters T_j of the exponential sum in (4.17) by this procedure. Hence, we may have to repeat the method for additional fixed values $\omega_2 = 2h_2$, $3h_2$ etc. in order to complete the knot sequence $(T_1, \ldots, T_{N_1+m_1})$. But it is always possible to complete the sequence in this way, see Remark 4.5.

Part 2: Next, knowing the knots $(T_1, \ldots, T_{N_1+m_1})$, we can compute the further coefficients $p_j(\nu h_2)$ for all $j \in \{1, \ldots, N_1 + m_1\}$ and $\nu \in \{2, \ldots, N_2 + m_2\}$ from the overdetermined linear system of equations in (4.16) using the Fourier samples $\widehat{f}(\eta h_1, \nu h_2)$, $\eta = 1, \ldots, 2(N_1 + m_1)$, $\nu = 1, \ldots, N_2 + m_2$.

Part 3: Up to now, we know the knot sequence $(T_1, \ldots, T_{N_1+m_1})$ and the values

$$p_j(\nu h_2) \qquad \text{for } j = 1, \ldots, N_1 + m_1, \text{ and } \nu = 1, \ldots, N_2 + m_2. \tag{4.19}$$

We still need to compute the knot sequence $(S_1, \ldots, S_{N_2+m_2})$ and the original coefficients of the function f, i.e. $c_{j,k}^{0,0}$ for $j = 1, \ldots, N_1$, $k = 1, \ldots, N_2$. This will be done by a further application of the Prony method and a recursion formula afterwards.

Consider the univariate exponential sum

$$p_1(\omega_2) = \sum_{k=1}^{N_2+m_2} c_{1,k}^{m_1,m_2} \, e^{-i \omega_2 S_k} \, .$$

We know the values

$$p_1(\nu h_2), \quad \nu = 1, \ldots, N_2 + m_2,$$

see (4.19), and p_1 is now an exponential sum with real-valued coefficients. Thus, we can use the Prony method as described in Section 2.2 in order to compute the knot sequence $(S_1, \ldots, S_{N_2+m_2})$ and the coefficients $c_{1,k}^{m_1,m_2}$, $k = 1, \ldots, N_2 + m_2$. Observe that the function value $p_1(0)$ is again given as zero since we have the following:

$$
\begin{aligned}
p_1(0) &= \sum_{k=1}^{N_2+m_2} c_{1,k}^{m_1,m_2} \, e^{-i \cdot 0 \cdot S_k} = \sum_{k=1}^{N_2+m_2} c_{1,k}^{m_1,m_2} \\
&\overset{(4.12)}{=} \sum_{k=1}^{N_2+m_2} \left[c_{1,k}^{m_1,m_2-1} - c_{1,k-1}^{m_1,m_2-1} \right] \\
&= c_{1,N_2+m_2}^{m_1,m_2-1} - c_{1,0}^{m_1,m_2-1} \\
&\overset{(4.13)}{=} 0.
\end{aligned}
$$

In case that $c_{1,k}^{m_1,m_2} = 0$ for some $k \in \{1, \ldots, N_2 + m_2\}$, we do not obtain all knots S_k and need to apply the Prony method also to $p_j(\omega_2)$ for $j = 2, 3$ etc. in order to complete the knot sequence $(S_1, \ldots, S_{N_2+m_2})$, cf. Remark 4.5.

Part 4: All further coefficients of the exponential sums in (4.15), i.e. $c_{j,k}^{m_1,m_2}$ for the indices $j = 2, \ldots, N_1 + m_1$, and $k = 1, \ldots, N_2 + m_2$, are determined from the linear systems

$$p_j(\nu h_2) = \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \, e^{-i \nu h_2 S_k}, \qquad \nu = 1, \ldots, N_2 + m_2,$$

for each $j \in \{2, 3, \ldots, N_1 + m_1\}$.

Part 5: Finally, we use recursion formulae in order to evaluate the original coefficients $c_{j,k}^{0,0}$ of the function f in (4.1) from the coefficients $c_{j,k}^{m_1,m_2}$, which we have just obtained. The definitions (4.2)–(4.13) result in the following recursions, which have to be executed one after the other:

- $j = 1, \ldots, N_1 + m_1$:

$$c_{j,k}^{m_1,m_2-1} = \begin{cases} c_{j,1}^{m_1,m_2} & \text{for } k = 1, \\ c_{j,k}^{m_1,m_2} + c_{j,k-1}^{m_1,m_2-1} & \text{for } k = 2, \ldots, N_2 + m_2 - 1; \end{cases}$$

- $k = 1, \ldots, N_2 + m_2 - 1$:

$$c_{j,k}^{m_1-1,m_2-1} = \begin{cases} c_{1,k}^{m_1,m_2-1} & \text{for } j = 1, \\ c_{j,k}^{m_1,m_2-1} + c_{j-1,k}^{m_1-1,m_2-1} & \text{for } j = 2, \ldots, N_1 + m_1 - 1; \end{cases}$$

- $j = 1, \ldots, N_1 + m_1 - 1$:

$$c_{j,k}^{m_1-1,m_2-2} = \begin{cases} (S_2 - S_1)c_{j,1}^{m_1-1,m_2-1}, & k = 1, \\ (S_{k+1} - S_k)c_{j,k}^{m_1-1,m_2-1} + c_{j,k-1}^{m_1-1,m_2-2}, & k = 2, \ldots, N_2 + m_2 - 2; \end{cases}$$

- $k = 1, \ldots, N_2 + m_2 - 2$:

$$c_{j,k}^{m_1-2,m_2-2} = \begin{cases} (T_2 - T_1)c_{1,k}^{m_1-1,m_2-2}, & j = 1, \\ (T_{j+1} - T_j)c_{j,k}^{m_1-1,m_2-2} + c_{j-1,k}^{m_1-2,m_2-2}, & j = 2, \ldots, N_1 + m_1 - 2; \end{cases}$$

- $r_2 = m_2 - 2, \ldots, 1$, and $j = 1, \ldots, N_1 + m_1 - 2$:

$$c_{j,k}^{m_1-2,r_2-1} = \begin{cases} \left(\frac{S_{1+m_2-r_2} - S_1}{m_2 - r_2}\right) c_{j,1}^{m_1-2,r_2}, & k = 1, \\ \left(\frac{S_{k+m_2-r_2} - S_k}{m_2 - r_2}\right) c_{j,k}^{m_1-2,r_2} + c_{j,k-1}^{m_1-2,r_2-1}, & k = 2, \ldots, N_2 + r_2 - 1; \end{cases}$$

- $r_2 = m_1 - 2, \ldots, 1$, and $k = 1, \ldots, N_2$:

$$c_{j,k}^{r_1-1,0} = \begin{cases} \left(\frac{T_{1+m_1-r_1} - T_1}{m_1 - r_1}\right) c_{1,k}^{r_1,0}, & j = 1, \\ \left(\frac{T_{j+m_1-r_1} - T_j}{m_1 - r_1}\right) c_{j,k}^{r_1,0} + c_{j-1,k}^{r_1-1,0}, & j = 2, \ldots, N_1 + r_1 - 1. \end{cases}$$

The last iteration of the last recursion formula now provides the coefficients $c_{j,k}^{0,0}$ for $j = 1, \ldots, N_1$, and $k = 1, \ldots, N_2$.

$\square$

Remark 4.5.

Observe that we usually need to apply the Prony method only twice in order to obtain the two knot sequences $(T_1, \ldots, T_{N_1+m_1})$ and $(S_1, \ldots, S_{N_2+m_2})$. All coefficients $c_{j,k}^{0,0}$ can be computed afterwards from linear systems of equations.

At the end of Part 1 of the proof of Theorem 4.4, we have encountered the problem that maybe not all values $p_j(h_2)$ for $j \in \{1, \ldots, N_1+m_1\}$ are non-zero such that we are

not able to find all parameters T_j. If this is the case, we repeat the method described in that part of the proof for additional values $\omega_2 = 2h_2$, $3h_2$ etc., that is, we consider $p_j(2h_2)$, $p_j(3h_2)$ etc. in order to complete the knot sequence $(T_1, \ldots, T_{N_1+m_1})$.

In this manner, we can always complete the knot sequence, which can be seen as follows:

Assume that the parameter T_{j_0} for some index $j_0 \in \{1, \ldots, N_1 + m_1\}$ is never found, i.e.

$$p_{j_0}(\nu h_2) = \sum_{k=1}^{N_2+m_2} c_{j_0,k}^{m_1,m_2} \big(\mathrm{e}^{-\mathrm{i}\,h_2 S_k}\big)^\nu = 0 \quad \text{for } \nu = 1, \ldots, N_2 + m_2.$$

Thus, we have the linear system

$$\begin{pmatrix} \mathrm{e}^{-\mathrm{i}\,h_2 S_1} & \cdots & \mathrm{e}^{-\mathrm{i}\,h_2 S_{N_2+m_2}} \\ \vdots & & \vdots \\ \big(\mathrm{e}^{-\mathrm{i}\,h_2 S_1}\big)^{N_2+m_2} & \cdots & \big(\mathrm{e}^{-\mathrm{i}\,h_2 S_{N_2+m_2}}\big)^{N_2+m_2} \end{pmatrix} \cdot \begin{pmatrix} c_{j_0,1}^{m_1,m_2} \\ \vdots \\ c_{j_0,N_2+m_2}^{m_1,m_2} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (4.20)$$

where we denote the system matrix by $\boldsymbol{V}$. Using simple rules for calculating determinants[2], we obtain

$$\det(\boldsymbol{V}) = \left(\prod_{k=1}^{N_2+m_2} \big(\mathrm{e}^{-\mathrm{i}\,h_2 S_k}\big)\right) \cdot \det\big(\widetilde{\boldsymbol{V}}\big) \quad (4.21)$$

with

$$\widetilde{\boldsymbol{V}} := \begin{pmatrix} 1 & \cdots & 1 \\ \mathrm{e}^{-\mathrm{i}\,h_2 S_1} & \cdots & \mathrm{e}^{-\mathrm{i}\,h_2 S_{N_2+m_2}} \\ \vdots & & \vdots \\ \big(\mathrm{e}^{-\mathrm{i}\,h_2 S_1}\big)^{N_2+m_2-1} & \cdots & \big(\mathrm{e}^{-\mathrm{i}\,h_2 S_{N_2+m_2}}\big)^{N_2+m_2-1} \end{pmatrix}.$$

Since $|\mathrm{e}^{-\mathrm{i}\,hS_k}| = 1$ for $k = 1, \ldots, N_2 + m_2$, the product $\prod_{k=1}^{N_2+m_2}\big(\mathrm{e}^{-\mathrm{i}\,h_2 S_k}\big)$ in (4.21) is unequal to zero. Further, the parameters S_k, $k = 1, \ldots, N_2 + m_2$, are pairwise different, and we have $h_2 S_k \in (-\pi, \pi]$ for $k = 1, \ldots, N_2 + m_2$. Thus, the values $\mathrm{e}^{-\mathrm{i}\,h_2 S_k}$, $k = 1, \ldots, N_2 + m_2$, are pairwise different such that the determinant of the Vandermonde-type matrix $\widetilde{\boldsymbol{V}}$ is also unequal to zero[3].

Hence, the matrix $\boldsymbol{V}$ is invertible, and the solution to the system in (4.20) is given by

$$c_{j_0,1}^{m_1,m_2} = \ldots = c_{j_0,N_2+m_2}^{m_1,m_2} = 0.$$

This means that the parameter T_{j_0} does not occur in the representation (4.14). Therefore, if a parameter T_{j_0} is necessary for the representation of f in (4.1), we will always be able to find it by using the approach described in the proof of Theorem 4.4.

---

[2] See [43, pp. 79–81, Lemma 7.11, (4) + Lemma 7.14, (1)], for example.
[3] Cf. [34, p. 29].

Remark 4.6.
As in the univariate case, see Remarks 3.9, 2., the problem of vanishing coefficients $c_{j,k}^{m_1,m_2}$ only occurs if the function f in (4.1) contains local redundancies. Then the approach presented above will result in the simplest representation of f, that is, the reconstructed representation of f will not possess redundant terms.

4.1.2. Tensor-products of non-uniform translates

In this subsection, we want to extend the theorem for the reconstruction of non-uniform translates of low-pass filter functions, i.e. Theorem 3.11, to the case of tensor-products in the bivariate case. Thus, we consider sparse tensor-product representations of the form

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x_1, x_2) := \sum_{j=1}^{N_1} \sum_{k=1}^{N_2} c_{j,k} \Phi_1(x_1 - T_j) \Phi_2(x_2 - S_k) \qquad (4.22)$$

with two low-pass filter functions Φ_1 and Φ_2 as introduced in Section 3.3; i.e., the functions $\Phi_\nu \in L^1(\mathbb{R})$, $\nu = 1, 2$, satisfy the condition $|\widehat{\Phi}_\nu(\omega)| > C$ for $\omega \in (-T, T)$ for some constants C_0 and $T > 0$ ($\nu = 1, 2$). Further, we have non-zero, real-valued coefficients $c_{j,k}$ for $j = 1, \ldots, N_1$, $k = 1, \ldots, N_2$, and the two real shift sequences $-\infty < T_1 < \ldots < T_{N_1} < \infty$ and $-\infty < S_1 < \ldots < S_{N_2} < \infty$.

Application of the Fourier transform yields[4]

$$\widehat{f}(\omega_1, \omega_2) = \left(\sum_{j=1}^{N_1} \sum_{k=1}^{N_2} c_{j,k} \, e^{-i\omega_1 T_j} \, e^{-i\omega_2 S_k} \right) \widehat{\Phi}_1(\omega_1) \widehat{\Phi}_2(\omega_2), \quad \boldsymbol{\omega} := (\omega_1, \omega_2)^{\mathrm{T}} \in \mathbb{R}^2. \quad (4.23)$$

Theorem 4.7 (Reconstruction of tensor-product non-uniform translates).
Let $-\infty < T_1 < \ldots < T_{N_1} < \infty$ and $-\infty < S_1 < \ldots < S_{N_2} < \infty$ be real sequences, and $c_{j,k} \in \mathbb{R} \setminus \{0\}$ for $j = 1, \ldots, N_1$, $k = 1, \ldots, N_2$. Further, let $\Phi_1, \Phi_2 \in L^1(\mathbb{R}^2)$ be given real-valued functions satisfying $|\widehat{\Phi}_\ell(\omega)| > C$ for $\omega \in (-T, T)$ for some constants $C > 0$ and $T > 0$ ($\ell = 1, 2$). Assume that the positive constants h_1 and h_2 satisfy $h_1 T_j, h_2 S_k \in (-\pi, \pi]$ for all $j \in \{1, \ldots, N_1\}$ and $k \in \{1, \ldots, N_2\}$, and assume that $h_1 < \frac{T}{N_1}$ and $h_2 < \frac{T}{N_2}$. Then the function f of the form (4.22) can be uniquely recovered from the $(2N_1 + 1) \cdot (N_2 + 1)$ Fourier samples $\widehat{f}(\eta h_1, \nu h_2)$ where $\eta = 0, 1, \ldots, 2N_1$ and $\nu = 0, 1, \ldots, N_2$.

Proof. The proof follows the same lines as the proof of Theorem 4.4. It consists of two major parts. First, we fix the variable ω_2 in the frequency domain such that we get a univariate problem to which we can apply the Prony method in order to obtain

---

[4]Note that we have used Definition and Proposition A.1, Proposition A.3, and Fubini's theorem, see [22, pp. 175–176, 2.1c)], in order to compute the Fourier transform.

the shift values $T_1, \ldots, T_{N_1}$. In the second part, we use a univariate problem with respect to the variable ω_2 and determine the shift values $S_1, \ldots, S_{N_2}$, and the original coefficients $c_{j,k}$, $j = 1, \ldots, N_1$, $k = 1, \ldots, N_2$.

Part 1: Considering (4.23), we obtain

$$\frac{\widehat{f}(\omega_1, \omega_2)}{\widehat{\Phi}_1(\omega_1)\widehat{\Phi}_2(\omega_2)} = \sum_{j=1}^{N_1}\sum_{k=1}^{N_2} c_{j,k}\, \mathrm{e}^{-\mathrm{i}\,\omega_1 T_j}\, \mathrm{e}^{-\mathrm{i}\,\omega_2 S_k} = \sum_{j=1}^{N_1}\left(\sum_{k=1}^{N_2} c_{j,k}\, \mathrm{e}^{-\mathrm{i}\,\omega_2 S_k}\right)\mathrm{e}^{-\mathrm{i}\,\omega_1 T_j}$$

for $\omega_1, \omega_2 \in (-T, T)$. Secondly, we set

$$p_j(\omega_2) := \sum_{k=1}^{N_2} c_{j,k}\, \mathrm{e}^{-\mathrm{i}\,\omega_2 S_k} \quad \text{for } j = 1, \ldots, N_1 \tag{4.24}$$

such that we get

$$\frac{\widehat{f}(\omega_1, \omega_2)}{\widehat{\Phi}_1(\omega_1)\widehat{\Phi}_2(\omega_2)} = \sum_{j=1}^{N_1} p_j(\omega_2)\, \mathrm{e}^{-\mathrm{i}\,\omega_1 T_j}. \tag{4.25}$$

We fix $\omega_2 := h_2$ in order to obtain a univariate exponential sum with respect to the variable ω_1. Define

$$P(\omega_1) := \frac{\widehat{f}(\omega_1, h_2)}{\widehat{\Phi}_1(\omega_1)\widehat{\Phi}_2(h_2)} = \sum_{j=1}^{N_1} p_j(h_2)\, \mathrm{e}^{-\mathrm{i}\,\omega_1 T_j}. \tag{4.26}$$

The coefficients $p_j(h_2)$ are complex-valued due to their definition in (4.24). Observe that we have $\widehat{\Phi}_1(\eta h) \neq 0$ for $\eta = 0, \ldots, 2N_1$, $hT_j \in (-\pi, \pi]$ for $j = 1, \ldots, N_1$, and $\widehat{\Phi}_2(h_2) \neq 0$ by the assumptions about h_1 and h_2. Therefore, using the Fourier samples $\widehat{f}(\eta h_1, h_2)$ for $\eta = 0, \ldots, 2N_1$, we can apply the Prony method as presented in Section 2.1 to the function P. For this purpose, we need the function values $P(\eta h_1)$ for $\eta = 0, \ldots, 2N_1$, which we can compute by

$$P(\eta h_1) = \frac{\widehat{f}(\eta h_1, h_2)}{\widehat{\Phi}_1(\eta h_1)\widehat{\Phi}_2(h_2)} \quad \text{for } \eta = 0, \ldots, 2N_1.$$

Using this approach, we obtain the shift sequence $(T_1, \ldots, T_{N_1})$ and the coefficients $p_j(h_2)$, $j = 1, \ldots, N_1$.

If it happens that not all values $p_j(h_2)$, $j = 1, \ldots, N_1$, are non-zero, we will not find all shift values T_j in (4.26) by using the Prony method. Thus, we my have to repeat the method for additional fixed values $\omega_2 = 2h_2$, $3h_2$ etc. such that we can complete the shift sequence $(T_1, \ldots, T_{N_1})$, cf. Remark 4.5.

Employing the found shift values $T_1, \ldots, T_{N_1}$, we can compute the coefficients $p_j(0)$ and $p_j(\nu h_2)$ for $j = 1, \ldots, N_1$, $\nu = 2, \ldots, N_2$ from the linear system of equations

that we obtain from (4.26) using the Fourier samples $\widehat{f}(\eta h_1, \nu h_2)$, $\eta = 0, \ldots, 2N_1$, $\nu = 0, \ldots, N_2$.

Part 2: Let us recall that we have computed the shift sequence $(T_1, \ldots, T_{N_1})$ and the values

$$p_j(\nu h_2) \quad j = 1, \ldots, N_1, \ \nu = 0, \ldots, N_2. \tag{4.27}$$

It remains to determine the shift sequence $(S_1, \ldots, S_{N_2})$ and the original coefficients of the function f, i.e. $c_{j,k}$ for $j = 1, \ldots, N_1$, $k = 1, \ldots, N_2$.

Consider the univariate exponential sum

$$p_1(\omega_2) = \sum_{k=1}^{N_2} c_{1,k} \, \mathrm{e}^{-\mathrm{i}\,\omega_2 S_k},$$

where we know the values

$$p_1(\nu h_2), \quad \nu = 0, \ldots, N_2,$$

see (4.27). Note that p_1 is an exponential sum with real-valued coefficients. Hence, by applying the Prony method from Section 2.2, we compute the shift sequence $(S_1, \ldots, S_{N_2})$ and the coefficients $c_{1,k}$, $k = 1, \ldots, N_2$.

As the last step, we have to compute the remaining original coefficients in (4.22), i.e. $c_{j,k}$ for the indices $j = 2, \ldots, N_1$, and $k = 1, \ldots, N_2$. They are determined by the linear systems

$$p_j(\nu h_2) = \sum_{k=1}^{N_2} c_{j,k} \, \mathrm{e}^{-\mathrm{i}\,\nu h_2 S_k}, \qquad \nu = 1, \ldots, N_2,$$

for each $j \in \{2, 3, \ldots, N_1\}$. This completes the proof. $\qquad\square$

Remark 4.8.
Most of the time, we need only two applications of the Prony method in order to establish the shift sequences $(T_1, \ldots, T_{N_1})$ and $(S_1, \ldots, S_{N_2})$, cf. Remark 4.5. The coefficients $c_{j,k}$ are determined from linear systems of equations.

4.2. Non-uniform translates of bivariate functions

In the previous subsection, we have generalized the results from Section 3.3 for the reconstruction of non-uniform translates to the bivariate case, namely to the case of tensor-products of low-pass filter functions. Now we want to consider a more general setting for non-uniform translates in the bivariate case.

Therefore, we assume the given bivariate function $\Phi \in L^1(\mathbb{R}^2)$ to have a bounded Fourier transform $\widehat{\Phi}$ which satisfies $|\widehat{\Phi}(\boldsymbol{\omega})| > C > 0$ for $\|\boldsymbol{\omega}\|_2 < T$ for some constant

$T > 0$, and we examine functions $f : \mathbb{R}^2 \to \mathbb{R}$ that have the sparse representation

$$f(\boldsymbol{x}) := \sum_{j=1}^{N} c_j \Phi(\boldsymbol{x} - \boldsymbol{v}_j) \tag{4.28}$$

where c_j and $\boldsymbol{v}_j$ for $j = 1, \dots, N$ are non-zero, real-valued coefficients and real shift vectors respectively.

Note that such a function Φ can, for example, be constructed with the help of a univariate low-pass filter function $\widetilde{\Phi}$ as given in Section 3.3 with

$$\Phi(\boldsymbol{x}) := \widetilde{\Phi}\big(\|\boldsymbol{x}\|_2\big).$$

Further, observe that we can also take radial basis functions such as *polynomials*, *thin plate splines*, *Gaussians*, *Multiquadrics* or *inverse Multiquadrics*[5] as the given function Φ. See Remark 4.10 for more details.

As before, we would like to answer the question of the number of needed Fourier samples in order to recover the function f if Φ and N are known. Additionally, we want to establish a reconstruction scheme for computing the real shift vectors $\boldsymbol{v}_j := (v_{j,1}, v_{j,2})^{\mathrm{T}}$ and the real-valued coefficients c_j $(j = 1, \dots, N)$ from these samples. Observe that this problem is completely different from the separable case considered in Subsection 4.1.2.

Application of the Fourier transform to the function f in (4.28) leads to

$$\widehat{f}(\boldsymbol{\omega}) = \left(\sum_{j=1}^{N} c_j \, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\omega}, \boldsymbol{v}_j \rangle} \right) \widehat{\Phi}(\boldsymbol{\omega}), \quad \boldsymbol{\omega} \in \mathbb{R}^2, \tag{4.29}$$

cf. Definition A.1 and Proposition A.3.

We will see later that there can be cancellation effects in the reconstruction scheme if the non-zero coefficients c_j are not restricted to having the same sign for all indices $j \in \{1, \dots, N\}$. Thus, we only consider the cases where all coefficients $c_1, \dots, c_N$ have the same sign. Without loss of generality, we assume $c_j > 0$ for all $j \in \{1, \dots, N\}$.

In the following, we will distinguish between *adaptive sampling* and *sampling on predetermined lines*. Adaptive sampling means that we use Fourier samples from the horizontal and vertical axis of the frequency domain and choose a third sampling line through the origin in the frequency domain where the choice of this third line is based upon the results we have obtained by using the data from the first two sampling lines. In the case of sampling on predetermined lines, we decide on $N + 1$ arbitrary sampling lines through the origin in the frequency domain, and we will show that fewer lines do not suffice in general. This may be of great interest for practical purposes, but we will need $N^2 - 2N$ samples more than we need by applying adaptive sampling in order to recover the N shift vectors and corresponding coefficients in the representation (4.28) if $N \geq 2$.

---

[5]For the definitions of these functions, see [37, p. 1], for example.

4.2.1. Adaptive sampling

Considering the assumptions mentioned at the beginning of the section on hand, the following theorem shows that $3N + 1$ Fourier samples suffice to uniquely reconstruct the function f in (4.28) if we use adaptive sampling.

Theorem 4.9 (Reconstruction of bivariate non-uniform translates).
Let $\Phi \in L^1(\mathbb{R}^2)$ be a given bivariate function with a Fourier transform satisfying $|\widehat{\Phi}(\boldsymbol{\omega})| > C$ for $\|\boldsymbol{\omega}\|_2 < T$ for some constants $C > 0$ and $T > 0$. Further, let f be a bivariate function with the sparse representation (4.28) where c_j and $\boldsymbol{v}_j = (v_{j,1}, v_{j,2})^{\mathrm{T}}$ for $j = 1, \ldots, N$ are positive real-valued coefficients and real shift vectors respectively. Assume that the constant $h > 0$ fulfils the conditions $h\|\boldsymbol{v}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N\}$ and $h < \frac{T}{N}$. Then we get the following reconstruction result:

The function f can be uniquely recovered from the $3N + 1$ Fourier samples which are given by the set

$$\left\{ \widehat{f}(0,0), \widehat{f}(\ell h, 0), \widehat{f}(0, \ell h), \widehat{f}(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h) \mid \ell = 1, \ldots, N \right\}$$

where $\vartheta \in (0,1) \setminus \{\frac{1}{2}\}$ needs to be chosen suitably.

Proof. Considering the assumption about the Fourier transform of the given function Φ, the equation (4.29) yields

$$g(\boldsymbol{\omega}) := \frac{\widehat{f}(\boldsymbol{\omega})}{\widehat{\Phi}(\boldsymbol{\omega})} = \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\omega}, \boldsymbol{v}_j \rangle} \tag{4.30}$$

for $\boldsymbol{\omega} := (\omega_1, \omega_2)^{\mathrm{T}} \in \mathbb{R}^2$ with $\|\boldsymbol{\omega}\|_2 = \left(\omega_1^2 + \omega_2^2\right)^2 < T$.

In order to recover the original function f from samples of its Fourier transform, we have to reconstruct the parameters in the exponential sum on the right-hand side of (4.30), namely the shift vectors $\boldsymbol{v}_j$ and the coefficients c_j. We cannot directly apply the Prony method from Chapter 2 since we have here a bivariate exponential sum. Therefore, we will restrict the function g to different lines through the origin in the frequency domain such that we obtain univariate exponential sums in each case.

Part 1: First, we restrict g to the ω_1-axis, that is, we set ω_2 to be zero. This produces

$$g(\omega_1, 0) = \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\,\omega_1 v_{j,1}} \,.$$

However, we are faced with the problem that two or more shift vectors $\boldsymbol{v}_j$ may possess the same first coordinate $v_{j,1}$. Let us assume that the set $\{v_{1,1}, \ldots, v_{N,1}\}$ of

the true first coordinates contains $N_1 \leq N$ distinct values $\widetilde{v}_{1,1} < \widetilde{v}_{2,1} < \ldots < \widetilde{v}_{N_1,1}$. Then we find

$$g(\omega_1, 0) = \sum_{j=1}^{N_1} c_j^1 \, \mathrm{e}^{-\mathrm{i}\,\omega_1 \widetilde{v}_{j,1}} \tag{4.31}$$

where c_j^1 is the sum of all coefficients belonging to shift vectors with the same first coordinate $\widetilde{v}_{j,1}$, i.e.

$$c_j^1 := \sum_{k:\, v_{k,1} = \widetilde{v}_{j,1}} c_k, \quad j = 1, \ldots, N_1. \tag{4.32}$$

If we would allow the coefficients $c_1, \ldots, c_N$ to have different signs, it may happen that $c_j^1 = 0$ for some indices $j \in \{1, \ldots, N_1\}$. But this would mean that the terms $\mathrm{e}^{-\mathrm{i}\omega_1 \widetilde{v}_{j,1}}$ for these indices j do not occur in the exponential sum in (4.31) such that we would not be able to recover the coordinate $\widetilde{v}_{j,1}$. Thus, we restrict the coefficients $c_1, \ldots, c_N$ to having the same sign and assume $c_j > 0$ for all $j \in \{1, \ldots, N\}$ without loss of generality. Hence, we have $c_j^1 > 0$ for $j = 1, \ldots, N_1$.

Observe that the coefficients c_j^1 for $j = 1, \ldots, N_1$ are real-valued. Further, we have $h|\widetilde{v}_{j,1}| \leq h\|\widetilde{\boldsymbol{v}}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N_1\}$ by assumption. Hence, we can apply the Prony method as described in Section 2.2, see Algorithm 2.5. For this purpose, we need the function values

$$g(\ell h, 0), \quad \ell = 0, \ldots, N,$$

which are given by the definition of g in (4.30) and the $N+1$ Fourier samples

$$\widehat{f}(\ell h, 0), \quad \ell = 0, \ldots, N.$$

Observe that the fractions

$$g(0,0) = \frac{\widehat{f}(0,0)}{\widehat{\Phi}(0,0)}, \quad g(h,0) = \frac{\widehat{f}(h,0)}{\widehat{\Phi}(h,0)}, \quad \ldots, \quad g(Nh,0) = \frac{\widehat{f}(Nh,0)}{\widehat{\Phi}(Nh,0)}$$

are well defined since the assumptions $h < \frac{T}{N}$ and $|\widehat{\Phi}(\boldsymbol{\omega})| > C > 0$ for $\|\boldsymbol{\omega}\|_2 < T$ ensure that $\widehat{\Phi}(\ell h, 0) \neq 0$ for all $\ell \in \{0, \ldots, N\}$.

The application of the Prony method then provides the set

$$\widetilde{V}_1 := \{\widetilde{v}_{1,1}, \ldots, \widetilde{v}_{N_1,1}\}$$

and corresponding coefficients c_j^1, $j = 1, \ldots, N_1$. Note that these coefficients are not the coefficients that we need for the reconstruction of f in (4.28) if $N_1 < N$. Those coefficients have to be determined later.

Actually, it would suffice to use the $N_1 + 1$ Fourier samples $\widehat{f}(\ell h, 0)$, $\ell = 0, \ldots, N_1$. But since we do not know the value N_1, we have to assume that the coordinate values $v_{1,1}, \ldots, v_{N,1}$ are distinct and hence take $N+1$ Fourier samples.

Part 2: Now we restrict g to the ω_2-axis and apply the above-described procedure to the univariate function $g(0, \cdot)$ defined by

$$g(0, \omega_2) = \sum_{j=1}^{N} c_j\, e^{-i\omega_2 v_{j,2}}\,.$$

We assume that the set $\{v_{1,2}, \ldots, v_{N,2}\}$ of the true second coordinates contains N_2 distinct values $\widetilde{v}_{1,2} < \widetilde{v}_{2,2} < \ldots < \widetilde{v}_{N_2,2}$. Setting

$$c_j^2 := \sum_{k:\, v_{k,2}=\widetilde{v}_{j,2}} c_k, \quad j = 1, \ldots, N_2,$$

where we have $c_j^2 > 0$ for $j = 1, \ldots, N_2$ by the assumption about the coefficients c_j $(j = 1, \ldots, N)$, we get

$$g(0, \omega_2) = \sum_{j=1}^{N_2} c_j^2\, e^{-i\omega_2 \widetilde{v}_{j,2}}\,. \tag{4.33}$$

Using the Fourier samples $\widehat{f}(0, \ell h)$ where $\ell = 0, \ldots, N$, we apply the Prony method in form of Algorithm 2.5 to the function g in (4.33). As a result, we obtain the set

$$\widetilde{V}_2 := \{\widetilde{v}_{1,2}, \ldots, \widetilde{v}_{N_2,2}\}$$

and corresponding coefficients c_j^2, $j = 1, \ldots, N_2$. Like in the first univariate problem, these coefficients are not the coefficients of f in (4.28) if $N_2 < N$.

Part 3: So far, we know the first coordinates $\widetilde{v}_{1,1}, \ldots, \widetilde{v}_{N_1,1}$, and the second coordinates $\widetilde{v}_{1,2}, \ldots, \widetilde{v}_{N_2,2}$ of the true shift vectors $\boldsymbol{v}_j$, $j = 1, \ldots, N$. Thus, the true shift vectors have to be elements of the set

$$K := \{\boldsymbol{v} := (v_1, v_2)^T \mid v_1 \in \widetilde{V}_1,\ v_2 \in \widetilde{V}_2\}, \tag{4.34}$$

which contains $N_1 N_2 \leq N^2$ *candidate vectors* for the true shift vectors.

In order to determine the N true shift vectors in the set K, we use a third univariate problem to which we apply the Prony method. This means that we need further Fourier samples from a third sampling line.

For this purpose, we choose a parameter $\vartheta \in (0, 1) \setminus \{\frac{1}{2}\}$ such that the orthogonal projections of all candidate vectors in K onto the line $x_2 = \tan(\vartheta\pi)x_1$ are pairwise different, that is to say that the values $\big(\cos(\vartheta\pi)v_1 + \sin(\vartheta\pi)v_2\big)$ where $\boldsymbol{v} = (v_1, v_2)^T \in K$ are distinct, see Figure 4.1. Such a line can always be found since we consider linear combinations with a finite number of summands in (4.28). Thus, the candidate set K contains finitely many points.

The parameter ϑ is then used to determine the third sampling line in the frequency domain. The angle between the positive ω_1-axis and the part of the line

$\omega_2 = \tan(\vartheta\pi)\omega_1$ which lies in the upper half space is given by the value $\vartheta\pi$ rad. We use equispaced sampling locations on this third sampling line $\omega_2 = \tan(\vartheta\pi)\omega_1$ with step size h, see Figure 4.1, and take the N Fourier samples

$$\widehat{f}\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big), \quad \ell = 1, \dots, N.$$

**: original shift vectors   **: further candidates in K $\boldsymbol{v} := (v_1, v_2)^{\mathrm{T}}$ $v_3 = \cos(\vartheta\pi)v_1 + \sin(\vartheta\pi)v_2$
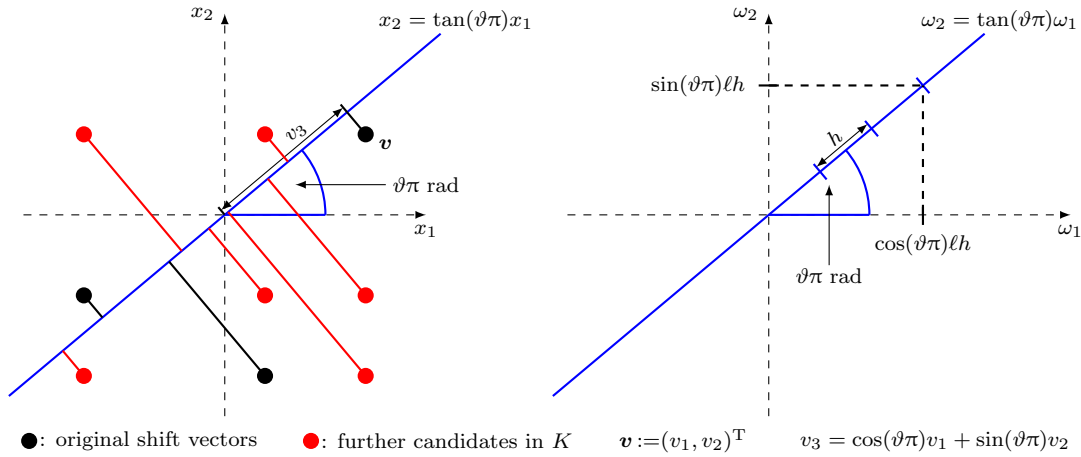
Figure 4.1. Left: *Determination of the parameter ϑ in the time domain.* Right: *Third sampling line in the frequency domain with sampling locations (displayed example: $\ell = 5$).*

<u>*Part 4:*</u> Now we consider the univariate function $g\big(\cos(\vartheta\pi)\cdot, \sin(\vartheta\pi)\cdot\big)$ given by

$$g\big(\cos(\vartheta\pi)\omega_1, \sin(\vartheta\pi)\omega_1\big) = \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\,\omega_1[\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}]} = \sum_{j=1}^{N} c_j\, \mathrm{e}^{-\mathrm{i}\,\omega_1 v_{j,3}} \quad (4.35)$$

with

$$v_{j,3} := \cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}, \quad j = 1, \dots, N.$$

Observe that the values $v_{1,3}, \dots, v_{N,3}$ are coordinate values of the true shift vectors with respect to the line $x_2 = \tan(\vartheta\pi)x_1$, and that these values are distinct since all possible shift vectors in K yield different orthogonal projections onto the line $x_2 = \tan(\vartheta\pi)x_1$, which is due to the choice of ϑ. This also means that there are no summation effects for the coefficients such that the coefficients $c_1, \dots, c_N$ in (4.35) are the true coefficients, which we seek to recover.

Let us order the coordinate values $v_{j,3}$ such that we obtain

$$
\begin{aligned}
g\big(\cos(\vartheta\pi)\omega_1, \sin(\vartheta\pi)\omega_1\big) &= \frac{\widehat{f}\big(\cos(\vartheta\pi)\omega_1, \sin(\vartheta\pi)\omega_1\big)}{\widehat{\Phi}\big(\cos(\vartheta\pi)\omega_1, \sin(\vartheta\pi)\omega_1\big)} \\
&= \sum_{j=1}^{N} c_j^3 \, \mathrm{e}^{-\mathrm{i}\,\omega_1 \widetilde{v}_{j,3}}
\end{aligned}
\tag{4.36}
$$

with $\widetilde{v}_{1,3} < \widetilde{v}_{2,3} < \ldots < \widetilde{v}_{N,3}$ and corresponding coefficients $c_1^3, \ldots, c_N^3$ where the set $\{\widetilde{v}_{1,3}, \ldots, \widetilde{v}_{N,3}\}$ is a permuted version of $\{v_{1,3}, \ldots, v_{N,3}\}$, and the same permutation maps $\{c_1, \ldots, c_N\}$ onto $\{c_1^3, \ldots, c_N^3\}$.

We now apply the Prony method as described in Section 2.2 to the function g in (4.36) where we use the Fourier samples $\widehat{f}\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big)$ for $\ell = 0, \ldots, N$. For this purpose, note that all coefficients c_j^3 are real-valued, and that we have the assumption $h\|\boldsymbol{v}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N\}$. The rotated vectors

$$
\begin{pmatrix} \cos(\vartheta\pi) & \sin(\vartheta\pi) \\ -\sin(\vartheta\pi) & \cos(\vartheta\pi) \end{pmatrix} \boldsymbol{v}_j, \quad j = 1, \ldots, N,
$$

fulfil the same norm condition. Thus, we obtain

$$
\begin{aligned}
h|v_{j,3}| = h\big|\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\big| &= h\sqrt{\big(\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\big)^2} \\
&\leq h\sqrt{\big(\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\big)^2 + \big(-\sin(\vartheta\pi)v_{j,1} + \cos(\vartheta\pi)v_{j,2}\big)^2} \\
&= \left\| \begin{pmatrix} \cos(\vartheta\pi) & \sin(\vartheta\pi) \\ -\sin(\vartheta\pi) & \cos(\vartheta\pi) \end{pmatrix} \boldsymbol{v}_j \right\|_2 < \pi
\end{aligned}
$$

for $j = 1, \ldots, N$. This means that $h|\widetilde{v}_{j,3}| < \pi$ holds for all $j \in \{1, \ldots, N\}$. Further, we have

$$
\left\|\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big)^{\mathrm{T}}\right\|_2 = \Big(\big(\cos^2(\vartheta\pi) + \sin^2(\vartheta\pi)\big)\ell^2 h^2\Big)^{\frac{1}{2}} = |\ell h| < T
$$

for all $\ell \in \{0, \ldots, N\}$ by using the assumption $h < \frac{T}{N}$. This ensures, together with the assumption about $\widetilde{\Phi}$, that the fractions in (4.36) are well defined for $\omega_1 = \ell h$, $\ell = 0, \ldots, N$.

The application of the Prony method then produces the set

$$
\widetilde{V}_3 := \{\widetilde{v}_{1,3}, \ldots, \widetilde{v}_{N,3}\}
$$

and corresponding coefficients c_j^3, $j = 1, \ldots, N$.

Part 5: Now, as the final step, we can compute the true shift vectors in the set K of candidate shift vectors by comparison of the set K with the set $\widetilde{V}_3$. For this purpose, we

determine all vectors $(v_1, v_2)^{\mathrm{T}}$ in the set K for which there exist indices $j \in \{1, \ldots, N\}$ such that $\cos(\vartheta\pi)v_1 + \sin(\vartheta\pi)v_2 = \widetilde{v}_{j,3}$. Then the set

$$G := \big\{ \boldsymbol{v} = (v_1, v_2)^{\mathrm{T}} \mid v_1 \in \widetilde{V}_1,\ v_2 \in \widetilde{V}_2,\ \cos(\vartheta\pi)v_1 + \sin(\vartheta\pi)v_2 \in \widetilde{V}_3 \big\}$$

contains only the N true shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$, and each of the vectors $\boldsymbol{v} \in G$ is uniquely associated with a coefficient c_j^3 for $j \in \{1, \ldots, N\}$. Thus, we have

$$G = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\} \tag{4.37}$$

with corresponding coefficients

$$c_1, c_2, \ldots, c_N. \tag{4.38}$$

The parameters of the function f in (4.28), which we wanted to recover, are now given by the elements of G in (4.37) and the corresponding coefficients in (4.38). $\quad\square$

Remark 4.10.
We can also take radial basis functions such as *polynomials*, *thin plate splines*, *Gaussians*, *Multiquadrics* or *inverse Multiquadrics* as the given function Φ. Observe that only the Gaussians and the inverse Multiquadrics have Fourier transforms in the classical sense. In all other cases, we have to consider the generalized Fourier transform. The generalized Fourier transform of the mentioned radial basis functions is positive on $\mathbb{R}^2 \setminus \{\boldsymbol{0}\}$ and has at most a pole at the origin, see [37, pp. 17–19]. Thus, we cannot take the function value $\widehat{\Phi}(\boldsymbol{0})$ in all cases, which means that we will not be able to consider the function value of g in (4.30) for $\boldsymbol{\omega} = \boldsymbol{0}$. Hence, we cannot apply the Prony method from Section 2.2 to the three univariate problems (4.31), (4.33), and (4.36) because the function value $g(\boldsymbol{0})$ is essential in order to use the conjugate symmetry of g, being an exponential sum with real-valued coefficients. Instead, we can consider $2N + 1$ sampling values of $\widehat{f}$ on each sampling line, e.g. on the right-hand side of the origin, in such cases where we cannot take the value $\widehat{\Phi}(\boldsymbol{0})$ and apply the Prony method from Section 2.1 to each of the three univariate problems.

Remark 4.11.
The reconstruction scheme given above uses three univariate problems, (4.31), (4.33), and (4.36), in order to recover the shift vectors and the coefficients in the original bivariate problem (4.30). Each of these three problems is solved by using the Prony method from Section 2.2, which yields coordinate values of the shift vectors and corresponding coefficients. But, as we have seen in the proof of Theorem 4.9, we do not need the coefficients which are results of the application of the Prony method to the first two problems (4.31) and (4.33). Thus, we do not need to compute those coefficients, that is, we only need the computation steps 1–5 in Algorithm 2.5.

We summarize the scheme for the reconstruction of non-uniform translates of bivariate functions in the following algorithm where we assume that the real-valued coefficients c_j, $j = 1, \ldots, N$, in the representation (4.28) are positive.

Algorithm 4.12 (Reconstruction of bivariate non-uniform translates).

- **Input:**
    - $\Phi \in L^1(\mathbb{R}^2)$ with $|\widehat{\Phi}(\boldsymbol{\omega})| > C$ for $\|\boldsymbol{\omega}\|_2 < T$ for some $C > 0$ and $T > 0$;
    - step size $h > 0$ with $h\|\boldsymbol{v}_j\|_2 < \pi$ for $j = 1, \ldots, N$ and $h < \frac{T}{N}$;
    - $3N + 1$ Fourier samples, given by the set
    $$\left\{ \widehat{f}(0,0), \widehat{f}(\ell h, 0), \widehat{f}(0, \ell h), \widehat{f}(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h) \mid \ell = 1, \ldots, N \right\}$$
    where $\vartheta \in (0,1) \setminus \{\frac{1}{2}\}$ needs to be chosen suitably, see Step 6 of the computation.

- **Computation:**
    1. Compute
    $$g(\ell h, 0) = \frac{\widehat{f}(\ell h, 0)}{\widehat{\Phi}(\ell h, 0)}, \quad \ell = 0, \ldots, N.$$

    2. Use Algorithm 2.5, Steps 1–5 in order to compute the coordinate values $\widetilde{v}_{j,1}$, $j = 1, \ldots, N_1$, in (4.31).

    3. Compute
    $$g(0, \ell h) = \frac{\widehat{f}(0, \ell h)}{\widehat{\Phi}(0, \ell h)}, \quad \ell = 1, \ldots, N.$$

    4. Use Algorithm 2.5, Steps 1–5 in order to compute the coordinate values $\widetilde{v}_{j,2}$, $j = 1, \ldots, N_2$, in (4.33).

    5. Compute the Cartesian product of the two sets $\widetilde{V}_1 = \{\widetilde{v}_{1,1} \ldots, \widetilde{v}_{N_1,1}\}$ and $\widetilde{V}_2 = \{\widetilde{v}_{1,2}, \ldots, \widetilde{v}_{N_2,2}\}$ as the set of possible candidates for the true shift vectors, i.e.
    $$K := \left\{ \boldsymbol{v} := (v_1, v_2)^{\mathrm{T}} \mid v_1 \in \widetilde{V}_1, \ v_2 \in \widetilde{V}_2 \right\}.$$

    6. Choose a parameter $\vartheta \in (0,1) \setminus \{\frac{1}{2}\}$ such that the orthogonal projections of all candidate vectors in K onto the line $x_2 = \tan(\vartheta\pi)x_1$ are distinct.

    7. Acquire the N Fourier samples $\widehat{f}(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h)$, $\ell = 1, \ldots, N$. Then compute
    $$g(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h) = \frac{\widehat{f}(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h)}{\widehat{\Phi}(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h)}, \quad \ell = 1, \ldots, N.$$

8. Use Algorithm 2.5 in order to compute the coordinate values $\widetilde{v}_{j,3}$ and the corresponding coefficients c_j^3, $j = 1, \ldots, N$, in (4.36).

9. Determine the true shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ in K by comparison with the set $\widetilde{V}_3 = \{\widetilde{v}_{1,3}, \ldots, \widetilde{v}_{N,3}\}$, that is, determine all vectors $(v_1, v_2)^{\mathrm{T}}$ in K for which there exist indices $j \in \{1, \ldots, N\}$ such that $\cos(\vartheta\pi)v_1 + \sin(\vartheta\pi)v_2 = \widetilde{v}_{j,3}$.

Thus, we obtain the set

$$G = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$$

with corresponding coefficients $c_1, c_2, \ldots, c_N$.

- **Output:** Shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ with corresponding coefficients $c_1, c_2, \ldots, c_N$ determining f in (4.28).

The above-given reconstruction scheme relies upon the assumption that all coefficients $c_1, \ldots, c_N$ in the representation (4.28) have the same sign. Otherwise, there may occur cancellation effects such that we would not be able to recover all coordinate values of the shift vectors $\boldsymbol{v}_j$, $j = 1 \ldots, N$.

Example 4.13 (Possible cancellation effects).
Let us consider the following example:

$$f(\boldsymbol{x}) := \sum_{j=1}^{4} c_j \Phi(\boldsymbol{x} - \boldsymbol{v}_j), \quad \boldsymbol{x} \in \mathbb{R}^2,$$

with

$$\boldsymbol{v}_1 := (1, 1)^{\mathrm{T}}, \qquad \boldsymbol{v}_2 := (-1, 1)^{\mathrm{T}}, \qquad \boldsymbol{v}_3 := (-1, -1)^{\mathrm{T}}, \qquad \boldsymbol{v}_4 := (1, -1)^{\mathrm{T}};$$
$$c_1 := -1, \qquad c_2 := 1, \qquad c_3 := -1, \qquad c_4 := 1,$$

see Figure 4.2.

Now assume that the Fourier transform $\widehat{\Phi}(\boldsymbol{\omega})$ does not vanish for $\|\boldsymbol{\omega}\|_2 < T$ for some constant $T > 0$. Then the application of the Fourier transform yields

$$g(\boldsymbol{\omega}) := \frac{\widehat{f}(\boldsymbol{\omega})}{\widehat{\Phi}(\boldsymbol{\omega})} = \sum_{j=1}^{4} c_j \, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\omega}, \boldsymbol{v}_j \rangle}, \quad \boldsymbol{\omega} := (\omega_1, \omega_2)^{\mathrm{T}} \in \mathbb{R}^2,$$

cf. (4.29). First, by restricting the function g to the ω_1-axis, we have

$$\begin{aligned}
g(\omega_1, 0) &= c_1 \, \mathrm{e}^{-\mathrm{i}\omega_1} + c_2 \, \mathrm{e}^{\mathrm{i}\omega_1} + c_3 \, \mathrm{e}^{\mathrm{i}\omega_1} + c_4 \, \mathrm{e}^{-\mathrm{i}\omega_1} \\
&= (c_1 + c_4) \, \mathrm{e}^{-\mathrm{i}\omega_1} + (c_2 + c_3) \, \mathrm{e}^{\mathrm{i}\omega_1} \qquad\qquad (4.39) \\
&= 0.
\end{aligned}$$

Secondly, restriction to the ω_2-axis results in

$$
\begin{aligned}
g(0, \omega_2) &= c_1\, \mathrm{e}^{-\mathrm{i}\,\omega_2} + c_2\, \mathrm{e}^{-\mathrm{i}\,\omega_2} + c_3\, \mathrm{e}^{\mathrm{i}\,\omega_2} + c_4\, \mathrm{e}^{\mathrm{i}\,\omega_2} \\
&= (c_1 + c_2)\, \mathrm{e}^{-\mathrm{i}\,\omega_2} + (c_3 + c_4)\, \mathrm{e}^{\mathrm{i}\,\omega_2} \\
&= 0.
\end{aligned}
\tag{4.40}
$$

Thus, we are not able to recover anything by applying the Prony method to the two univariate problems (4.39) and (4.40), meaning that we are not able to determine a candidate set in order to choose a third sampling line appropriately. This means that we have to choose two new, different sampling lines. But it can still happen that these new sampling lines are chosen inappropriately such that there are again cancellation effects.

Figure 4.2. *Illustration of the original shift vectors with corresponding coefficients considered in* Example 4.13. *Red, dashed lines indicate that the connected vectors yield the same projections onto the lines $x_2 = 0$ or $x_1 = 0$.*

Remark 4.14.
We will generalize the proposed approach for the determination of the shift vectors and the corresponding coefficients of a linear combination of translates of a given function to $d > 2$ dimensions in the following chapter.

In [12], Buhmann and Pinkus consider a similar problem of determining the shift vectors and the corresponding coefficients of a function f of the form (4.28) in $\mathbb{R}^d$. But they assume that the function f is completely known such that they can compute the Fourier transform of f, i.e. a function g of the form (4.29). In contrast, we take only a small number of sampling values of the Fourier transform $\widehat{f}$.

Buhmann and Pinkus then use a differential operator and values of higher-order directional derivatives of g in order to determine the shift vectors and the corresponding coefficients. But they are also faced with the problem that they only know the

projections of the shift vectors onto lines given by vectors that determine the direction of the derivatives, and that those projections have to be distinct for the unique determination of the shift vectors. In order to obtain always distinct projections, they consider altogether values of $\frac{1}{2}N(N-1)(d-1)+d$ directional derivatives of g.

By transferring this approach to our setting of sampling the Fourier transform of f, we would need sampling values on $\frac{1}{2}N(N-1)(d-1)+d$ lines. But, as we will show in the following chapter, it suffices to consider $d+1$ sampling lines by generalizing our approach of adaptive sampling to d dimensions.

4.2.2. Sampling on predetermined lines

In the reconstruction scheme given in Theorem 4.9, the third sampling line in the frequency domain is chosen dependently on the set K in (4.34), i.e. dependently on the candidates for shift vectors in K which are found using Fourier data from the first two sampling lines. For practical purposes, it would be of great interest to compute the true shift vectors and coefficients in (4.28) from Fourier samples taken beforehand independently from the shift vectors.

Using three sampling lines, each of the candidates in K has not only three coordinate values but also three corresponding coefficients. These coordinate values and corresponding coefficients are the results from the Prony method applied to the three univariate problems (4.31), (4.33), and (4.36), which are associated with the Fourier data from the three sampling lines.

The proof of Theorem 4.9 shows that we do not need the coefficients obtained by applying the Prony method to the problems (4.31) and (4.33), see also Remark 4.11. Thus, the computation of those coefficients is not part of Algorithm 4.12.

Now the question arises if the function f in (4.28) can by uniquely reconstructed by taking three predetermined lines where we also exploit the knowledge of the corresponding coefficients from all three univariate problems (4.31), (4.33), and (4.36). Is it possible to employ all these corresponding coefficients, namely c_j^1, c_j^2, and c_j^3, in order to acquire conditions of the form (4.32) (or similar conditions) such that we can find the true shift vectors in K by using these conditions? First, let us consider the

Example 4.15 (Using a linear system to determine the coefficients).
Consider the linear combination of non-uniform translates given by

$$f(\boldsymbol{x}) := \sum_{j=1}^{4} c_j \Phi(\boldsymbol{x} - \boldsymbol{v}_j), \quad \boldsymbol{x} \in \mathbb{R}^2,$$

with

$$\boldsymbol{v}_1 := (1,1)^{\mathrm{T}}, \qquad \boldsymbol{v}_2 := (-1,1)^{\mathrm{T}}, \qquad \boldsymbol{v}_3 := (-1,-1)^{\mathrm{T}}, \qquad \boldsymbol{v}_4 := (1,-1)^{\mathrm{T}},$$

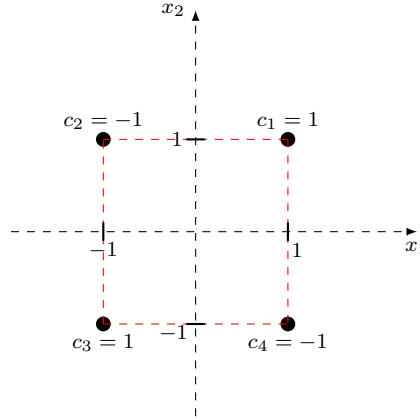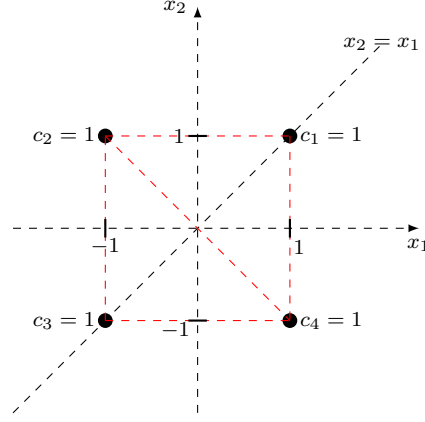**Figure 4.3.** *Illustration of the original shift vectors with corresponding coefficients. Red, dashed lines indicate that the connected vectors yield the same projections onto the lines $x_2 = 0$, $x_1 = 0$, or $x_2 = x_1$.*

and corresponding coefficients c_1, c_2, c_3, and c_4 as given in Figure 4.3. Actually, this is an adjusted version of Example 4.13 such that there are no cancellation effects.

Again, we assume that the Fourier transform $\widehat{\Phi}(\boldsymbol{\omega})$ does not vanish for $\|\boldsymbol{\omega}\|_2 < T$ for some constant $T > 0$. Then we have

$$g(\boldsymbol{\omega}) := \frac{\widehat{f}(\boldsymbol{\omega})}{\widehat{\Phi}(\boldsymbol{\omega})} = \sum_{j=1}^{4} c_j \, \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\omega}, \boldsymbol{v}_j \rangle}, \quad \boldsymbol{\omega} := (\omega_1, \omega_2)^{\mathrm{T}} \in \mathbb{R}^2.$$

We consider the ω_1-axis, the ω_2-axis, and the line $\omega_2 = \omega_1$ (i.e. the parameter ϑ for the third sampling line as introduced in the previous subsection is $\vartheta = \frac{\pi}{4}$) as sampling lines in the frequency domain. Restriction of g to the ω_1-axis yields

$$\begin{aligned} g(\omega_1, 0) &= (c_1 + c_4) \, \mathrm{e}^{-\,\mathrm{i}\,\omega_1} + (c_2 + c_3) \, \mathrm{e}^{\mathrm{i}\,\omega_1} \\ &= 2 \cdot \mathrm{e}^{-\,\mathrm{i}\,\omega_1} + 2 \cdot \mathrm{e}^{\mathrm{i}\,\omega_1}, \end{aligned} \tag{4.41}$$

and by restricting g to the ω_2-axis, we obtain

$$\begin{aligned} g(0, \omega_2) &= (c_1 + c_2) \, \mathrm{e}^{-\,\mathrm{i}\,\omega_2} + (c_3 + c_4) \, \mathrm{e}^{\mathrm{i}\,\omega_2} \\ &= 2 \cdot \mathrm{e}^{-\,\mathrm{i}\,\omega_2} + 2 \cdot \mathrm{e}^{\mathrm{i}\,\omega_2}. \end{aligned} \tag{4.42}$$

Next, we restrict g to the line $\omega_2 = \omega_1$ and get

$$g\big(\cos(\vartheta\pi)\omega_1, \sin(\vartheta\pi)\omega_1\big) = \sum_{j=1}^{4} c_j \, \mathrm{e}^{-\,\mathrm{i}\,\omega_1 \cdot \frac{1}{\sqrt{2}} \cdot (v_{j,1} + v_{j,2})} \tag{4.43}$$

since it is well known[6] that

$$\cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}} \qquad \text{and} \qquad \sin\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}.$$

Equation (4.43) now leads to

$$g\big(\cos(\vartheta\pi)\omega_1, \sin(\vartheta\pi)\omega_1\big) = c_1\,\mathrm{e}^{-\mathrm{i}\,\omega_1\cdot\frac{2}{\sqrt{2}}} + (c_2 + c_4)\,\mathrm{e}^{-\mathrm{i}\,\omega_1\cdot\frac{1}{\sqrt{2}}\cdot 0} + c_3\,\mathrm{e}^{-\mathrm{i}\,\omega_1\cdot\frac{(-2)}{\sqrt{2}}}$$

$$= 1\cdot\mathrm{e}^{-\mathrm{i}\,\omega_1\cdot\frac{2}{\sqrt{2}}} + 2\cdot\mathrm{e}^{-\mathrm{i}\,\omega_1\cdot\frac{1}{\sqrt{2}}\cdot 0} + 1\cdot\mathrm{e}^{-\mathrm{i}\,\omega_1\cdot\frac{(-2)}{\sqrt{2}}}.$$

Using Fourier samples from the three lines $\omega_2 = 0$, $\omega_1 = 0$, $\omega_2 = \omega_1$, and by applying the Prony method to the three univariate problems (4.41), (4.42), and (4.43), we obtain the candidate vectors

$$\widetilde{\boldsymbol{v}}_1 := (1,1)^{\mathrm{T}}, \quad \widetilde{\boldsymbol{v}}_2 := (-1,1)^{\mathrm{T}}, \quad \widetilde{\boldsymbol{v}}_3 := (-1,-1)^{\mathrm{T}}, \quad \widetilde{\boldsymbol{v}}_4 := (1,-1)^{\mathrm{T}},$$

which have to fulfil the following system of linear equations for the corresponding coefficients $\widetilde{c}_1$, $\widetilde{c}_2$, $\widetilde{c}_3$, and $\widetilde{c}_4$:

$$
\begin{array}{lll}
\widetilde{c}_1 + \widetilde{c}_2 = 2, & \widetilde{c}_1 + \widetilde{c}_4 = 2, & \widetilde{c}_1 = 1, \\
\widetilde{c}_3 + \widetilde{c}_4 = 2, & \widetilde{c}_2 + \widetilde{c}_3 = 2, & \widetilde{c}_2 + \widetilde{c}_4 = 2, \\
& & \widetilde{c}_3 = 1.
\end{array}
$$

This system is uniquely solvable such that we have the corresponding coefficients

$$\widetilde{c}_1 = \widetilde{c}_2 = \widetilde{c}_3 = \widetilde{c}_4 = 1,$$

associated with $\widetilde{\boldsymbol{v}}_1$, $\widetilde{\boldsymbol{v}}_2$, $\widetilde{\boldsymbol{v}}_3$, and $\widetilde{\boldsymbol{v}}_4$ respectively.

Example 4.15 illustrates the approach of using conditions on the coefficients associated with the shift vectors in order to determine the coefficients by using a linear system of equations. Unfortunately, such an approach is not always successful since we can always find counterexamples of sets of shift vectors with special symmetry properties where a complete reconstruction of f is not possible if we only use three predetermined sampling lines, see Example 4.17. We will show that for N shifts to be recovered $N+1$ predetermined sampling lines are always sufficient in order to uniquely recover the N shift vectors and corresponding coefficients in (4.28).

During the reconstruction process, we compute coordinate values of the true shift vectors with respect to certain lines through the origin in the time domain. If we consider Fourier samples from the horizontal or vertical axis of the frequency domain,

---

[6]See [68, p. 59], for example.

we obtain coordinate values with respect to the horizontal and vertical axis of the time domain respectively. On the other hand, if we use Fourier samples from the line $\omega_2 = \tan(\vartheta\pi)\omega_1$ with $\vartheta \in (0,1) \setminus \{\frac{1}{2}\}$ in the frequency domain, this yields coordinate values with respect to the line $x_2 = \tan(\vartheta\pi)x_1$ in the time domain, see also the proof of Theorem 4.9.

Actually, by considering Fourier samples from the three lines $\omega_2 = 0$, $\omega_1 = 0$, and $\omega_2 = \tan(\vartheta\pi)\omega_1$, we obtain the orthogonal projections of the true shift vectors onto the lines $x_2 = 0$, $x_1 = 0$, and $x_2 = \tan(\vartheta\pi)x_1$ respectively, see Figure 4.4. Thus, the question of the number of the needed predetermined sampling lines in the frequency domain can be reduced to the question of the number of needed lines in the time domain on which we have to know the orthogonal projections of a set of vectors in order to uniquely determine this set of vectors.
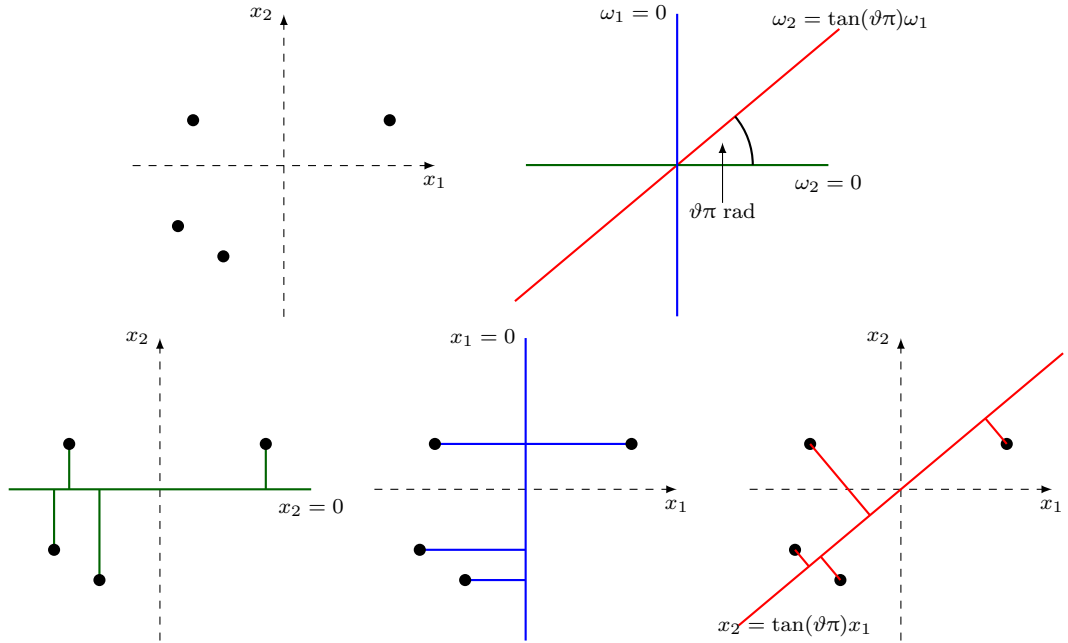
Figure 4.4. Top left: *Original shift vectors in the time domain.* Top right: *Sampling lines in the frequency domain.* Bottom left: *Using Fourier samples from the line $\omega_2 = 0$, we obtain the orthogonal projections of the shift vectors onto the line $x_2 = 0$.* Bottom middle: *By using Fourier samples from the line $\omega_1 = 0$, we find the orthogonal projections of the shift vectors onto the line $x_1 = 0$.* Bottom right: *Using Fourier samples from the line $\omega_2 = \tan(\vartheta\pi)\omega_1$, we get the orthogonal projections of the shift vectors onto the line $x_2 = \tan(\vartheta\pi)x_1$.*

These considerations lead us to the article [60] by A. Rényi. In the second chapter of that article, discrete mass distributions on the real plane which consist of N mass points with positive masses are examined, and Theorem 3 of that work states that

such distributions are uniquely determined if the orthogonal projections of all mass points onto $N + 1$ lines through the origin are known. By transferring this theorem to our setting, Theorem 3 of [60] reads as follows:

> **Proposition 4.16 (Determination of a set of vectors by projections).**
> *A set G of N vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N$ is uniquely determined if the orthogonal projections of all vectors in G onto $N + 1$ distinct lines through the origin are known.*

Since we have to understand the procedure of determining the vectors using the known projections in order to come up with a reconstruction scheme for our case of shift vectors and corresponding coefficients, we will prove this proposition by relying on the proof from [60, Chapter 2, Proof of Theorem 3].

Throughout this proof, we will use the following terminology (see also Figure 4.5):

- *Projection*: One projection onto a straight line through the origin consists of the N orthogonal projections of the vectors in G onto this line.

- *Projecting line*: A projecting line is a line which is perpendicular to a line on which the orthogonal projections of all vectors in G are taken, and which goes through at least one vector $\boldsymbol{v} \in G$.

- *Extreme points*: For each of the $N+1$ projections, we consider two specific points which are called extreme points of the projection: Let $l := \{\lambda \boldsymbol{u} \,|\, \lambda \in \mathbb{R}\}$ be one of the lines on which the orthogonal projections are considered. Further, let the orthogonal projection of an element $\boldsymbol{v} \in G$ onto this line l be denoted by $\mathrm{P}_l(\boldsymbol{v})$. Then take the two points

$$\boldsymbol{p}_{l,\min} := \operatorname*{argmin}_{\boldsymbol{p} \in \{\mathrm{P}_l(\boldsymbol{v}_1), \dots, \mathrm{P}_l(\boldsymbol{v}_N)\}} \langle \boldsymbol{p}, \boldsymbol{u} \rangle \quad \text{and} \quad \boldsymbol{p}_{l,\max} := \operatorname*{argmax}_{\boldsymbol{p} \in \{\mathrm{P}_l(\boldsymbol{v}_1), \dots, \mathrm{P}_l(\boldsymbol{v}_N)\}} \langle \boldsymbol{p}, \boldsymbol{u} \rangle.$$

These points are defined as the extreme points of the projection considered on the line l.

- *Extreme projecting lines*: For each projection, consider the projecting lines which go through the two extreme points of this projection. These two lines are called extreme projecting lines.

Proof of Proposition 4.16. For $N = 1$, the assertion is obviously true since the projections of $\boldsymbol{v}_1$ onto two distinct lines through the origin define $\boldsymbol{v}_1$ uniquely. Consider now the case $N \geq 2$.

Assume that we know $N + 1$ projections of the set G. We obtain two extreme points for at least N projections. Only for one projection it is possible that the projection
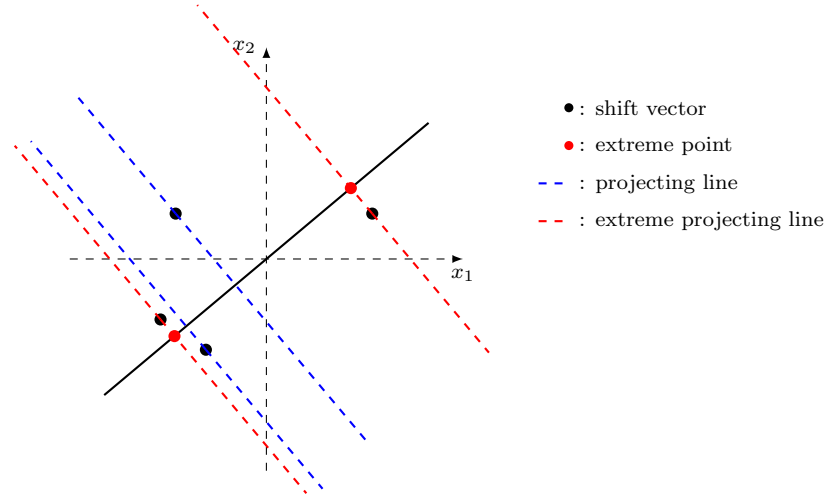
Figure 4.5. *Illustration of the used concepts in the proof of* Proposition 4.16.

consists of one point, which is the case if all vectors $\boldsymbol{v} \in G$ lie on the same projecting line. Thus, at least $2N + 1$ extreme projecting lines are given.

Further, we observe that each extreme projecting line passes through at least one vector $\boldsymbol{v} \in G$. We have N vectors and at least $2N + 1$ extreme projecting lines. Thus, there is at least one vector $\boldsymbol{v} \in G$ through which three or more extreme projecting lines are passing.

All vectors $\boldsymbol{v} \in G$ are located in one of the two closed half-planes which are determined by every extreme projecting line. Consider now the case that $r \geq 3$ extreme projecting lines pass through a point $\boldsymbol{p}$ of the plane. The before mentioned observations yield that the plane is divided into $2r$ angular domains by these r extreme projecting lines, see Figure 4.6 for an illustration. This can be seen by mathematical induction: every additional extreme projecting line through $\boldsymbol{p}$ divides two of the already existing angular domains where each of these two domains is divided in two parts. Thus the number of domains increases by 2.

Moreover, all vectors $\boldsymbol{v} \in G$ must lie in the interior or on the boundary of one of these angular domains. Let us denote this domain by D. The boundary of D is formed by two extreme projecting lines. This means that the other $r - 2$ extreme projecting lines which are passing through $\boldsymbol{p}$ can only have the point $\boldsymbol{p}$ in common with the set G. Since $r \geq 3$, at least one such extreme projecting line is given. But we know that every extreme projecting line passes through at least one vector $\boldsymbol{v} \in G$. Thus, we deduce that $\boldsymbol{p}$ is an element of G.

Summarizing this, we have shown the following: On the one hand, there exists at least one vector $\boldsymbol{v} \in G$ through which three or more extreme projecting lines are

passing. On the other hand, every point of the plane which is common to at least three extreme projecting lines is an element of G. Thus, we can find at least one element of G by considering only the extreme projecting lines.

Now we omit the orthogonal projections of the vector which we have just found. Hence, we know the orthogonal projections of the remaining $N-1$ vectors onto $N+1$ lines, and we apply the described procedure again to these projections. In this manner, we find step by step all vectors $\boldsymbol{v} \in G$. $\qquad\square$
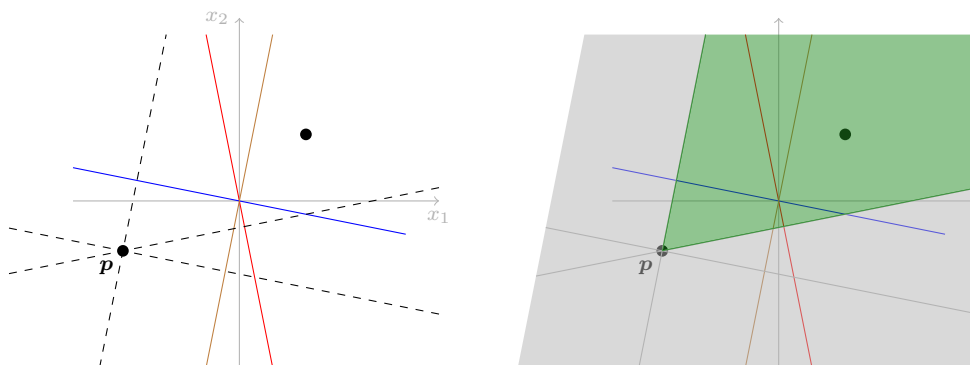
Figure 4.6. Left: *Set G of two points in the plane (black dots) and three lines on which the orthogonal projections of these points are taken (solid, coloured lines). The point $\boldsymbol{p}$ is common to three extreme projecting lines (dashed lines).* Right: *These extreme projecting lines divide the plane into six angular domains (grey domains and the green domain, in which all points of G are located).*

Example 4.17 (Non-unique recovery of shift vectors and coefficients).
Let us consider the set of original shift vectors given in Figure 4.7, i.e.

$$\boldsymbol{v}_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{v}_2 := \begin{pmatrix} \frac{3}{2} - 2\tan(\frac{\pi}{6}) \\ 1 - \cos(\frac{\pi}{6}) \end{pmatrix}, \quad \boldsymbol{v}_3 := \begin{pmatrix} \frac{3}{2} \\ -1 \end{pmatrix}.$$

Further, we assume that the corresponding coefficients c_1, c_2, c_3 for the representation (4.28) are all equal to two.

We will show that this set of vectors with corresponding coefficients is not uniquely recoverable by using Fourier samples from the three predetermined sampling lines $\omega_2 = 0$, $\omega_1 = 0$, and $\omega_2 = \tan(\frac{\pi}{6})\omega_1$.

Using Fourier samples from these three lines in the frequency domain and applying the Prony method to the problems (4.31), (4.33), and (4.36), we obtain the orthogonal projections of the original shift vectors onto the lines $x_2 = 0$, $x_1 = 0$, and $x_2 = \tan(\frac{\pi}{6})x_1$ in the time domain. The orthogonal projections are coordinate values of

the original shift vectors with respect to the these three lines. Combination of these coordinate values to points in the plane produces the candidate set $\{\widetilde{\boldsymbol{v}}_1, \widetilde{\boldsymbol{v}}_2, \widetilde{\boldsymbol{v}}_3, \widetilde{\boldsymbol{v}}_4, \widetilde{\boldsymbol{v}}_5, \widetilde{\boldsymbol{v}}_6\}$ with

$$\widetilde{\boldsymbol{v}_1} := \boldsymbol{v}_1, \qquad\qquad \widetilde{\boldsymbol{v}_2} := \boldsymbol{v}_2, \qquad\qquad \widetilde{\boldsymbol{v}_3} := \boldsymbol{v}_3,$$

$$\widetilde{\boldsymbol{v}}_4 := \begin{pmatrix} \frac{3}{2} \\ 1 - \cos(\frac{\pi}{6}) \end{pmatrix}, \quad \widetilde{\boldsymbol{v}}_5 := \begin{pmatrix} \frac{3}{2} - 2\tan(\frac{\pi}{6}) \\ 1 \end{pmatrix}, \quad \widetilde{\boldsymbol{v}}_6 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

since the vectors $\widetilde{\boldsymbol{v}}_4$, $\widetilde{\boldsymbol{v}}_5$, and $\widetilde{\boldsymbol{v}}_6$ yield the same orthogonal projections onto the lines $x_2 = 0$, $x_1 = 0$, and $x_2 = \tan(\frac{\pi}{6})x_1$ as the vectors $\widetilde{\boldsymbol{v}}_1$, $\widetilde{\boldsymbol{v}}_2$, and $\widetilde{\boldsymbol{v}}_3$, see Figure 4.7. Note that each point of the candidate set is an intersection point of three projection lines.

Now we have to determine the true shift vectors in this candidate set. But also considering the corresponding coefficients which we can compute for the problems (4.31), (4.33), and (4.36) does not lead to a unique solution. In particular, the unknown corresponding coefficients $\widetilde{c}_j$, $j = 1, \ldots, 6$, must solve the following linear system of equations:

$$\begin{aligned} \widetilde{c}_1 + \widetilde{c}_6 = 2, \qquad & \widetilde{c}_2 + \widetilde{c}_5 = 2, \qquad & \widetilde{c}_3 + \widetilde{c}_4 = 2, \\ \widetilde{c}_1 + \widetilde{c}_5 = 2, \qquad & \widetilde{c}_2 + \widetilde{c}_4 = 2, \qquad & \widetilde{c}_3 + \widetilde{c}_6 = 2, \\ \widetilde{c}_1 + \widetilde{c}_4 = 2, \qquad & \widetilde{c}_2 + \widetilde{c}_6 = 2, \qquad & \widetilde{c}_3 + \widetilde{c}_5 = 2, \end{aligned}$$

which is given by conditions of the form (4.32). This system has infinitely many solutions. For all $\alpha \in [0, 2]$, the set $\{\widetilde{\boldsymbol{v}}_1, \widetilde{\boldsymbol{v}}_2, \widetilde{\boldsymbol{v}}_3, \widetilde{\boldsymbol{v}}_4, \widetilde{\boldsymbol{v}}_5, \widetilde{\boldsymbol{v}}_6\}$ with corresponding coefficients

$$\widetilde{c}_1 = \widetilde{c}_2 = \widetilde{c}_3 = \alpha \quad \text{and} \quad \widetilde{c}_4 = \widetilde{c}_5 = \widetilde{c}_6 = 2 - \alpha$$

is a possible solution. Special cases are

- the set $\{\widetilde{\boldsymbol{v}}_1, \widetilde{\boldsymbol{v}}_2, \widetilde{\boldsymbol{v}}_3, \widetilde{\boldsymbol{v}}_4, \widetilde{\boldsymbol{v}}_5, \widetilde{\boldsymbol{v}}_6\}$ where all corresponding coefficients $\widetilde{c}_j$ ($j = 1, \ldots, 6$) are equal to one (i.e. $\alpha = 1$);

- the set $\{\widetilde{\boldsymbol{v}}_1, \widetilde{\boldsymbol{v}}_2, \widetilde{\boldsymbol{v}}_3\}$ with corresponding coefficients $\widetilde{c}_1 = \widetilde{c}_2 = \widetilde{c}_3 = 2$ (i.e. $\alpha = 2$);

- the set $\{\widetilde{\boldsymbol{v}}_4, \widetilde{\boldsymbol{v}}_5, \widetilde{\boldsymbol{v}}_6\}$ with corresponding coefficients $\widetilde{c}_4 = \widetilde{c}_5 = \widetilde{c}_6 = 2$ (i.e. $\alpha = 0$).

Example 4.17 illustrates for $N = 3$ that N lines on which the orthogonal projections of a set of N vectors are taken do not always suffice to uniquely determine this set of vectors. As Rényi states in [60], such examples can always be constructed in the following way:
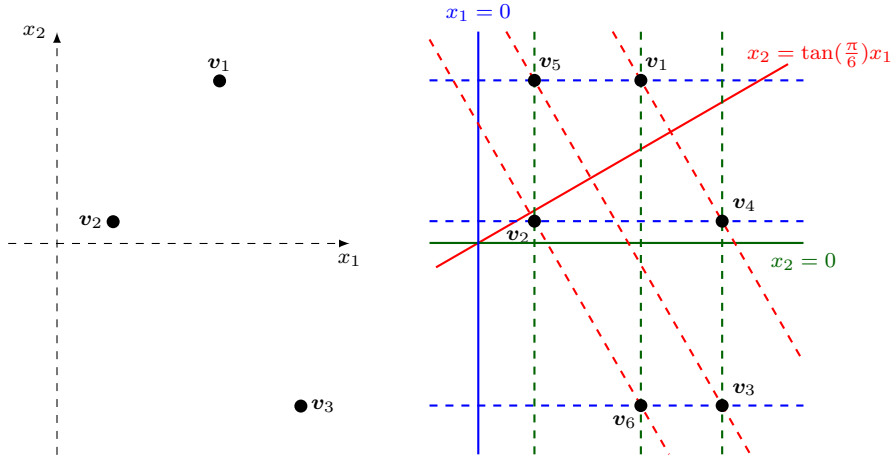
Figure 4.7. Left: *Original shift vectors in time domain.* Right: *The set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ yields the same orthogonal projections onto the lines $x_2 = 0$, $x_1 = 0$, and $x_2 = \tan(\frac{\pi}{6})x_1$ as the set $\{\boldsymbol{v}_4, \boldsymbol{v}_5, \boldsymbol{v}_6\}$.*

Procedure 4.18. (Construction of a set of N vectors being not uniquely determined by orthogonal projections onto N lines — according to [60]).

Consider a convex polygon of $2N$ sides with anticlockwise numbered vertices $\boldsymbol{v}_j$ for $j = 1, \ldots, 2N$ such that the following condition is satisfied:

> The line through the vertices $\boldsymbol{v}_j$ and $\boldsymbol{v}_k$ is parallel to the line through the vertices $\boldsymbol{v}_\ell$ and $\boldsymbol{v}_m$ in the case

$$j + k \equiv 1 \ (\mathrm{mod}\ 2) \quad \text{and} \quad j + k \equiv \ell + m \ (\mathrm{mod}\ 2N). \tag{4.44}$$

There are N distinct sets of such parallel lines, see Figure 4.8. Denote these sets by S_j, $j = 1, \ldots, N$. Further, denote the line through the origin which is perpendicular to the lines in S_j by l_j, and take the sets

$$V_1 := \{\boldsymbol{v}_{2j+1} \mid j \in \{0, 1, 2, \ldots, N-1\}\} \quad \text{and} \quad V_2 := \{\boldsymbol{v}_{2j} \mid j \in \{1, 2, \ldots, N\}\}.$$

Then the orthogonal projections of the elements in V_1 onto the lines $l_1, \ldots, l_N$ are the same as the orthogonal projections of the vectors in V_2.

Thus, in order to be always able to uniquely recover a set of N vectors from given orthogonal projections of these vectors onto predetermined lines through the origin, we need $N+1$ lines through the origin. In our reconstruction scheme, we obtain these orthogonal projections by using Fourier samples from corresponding $N + 1$ lines in
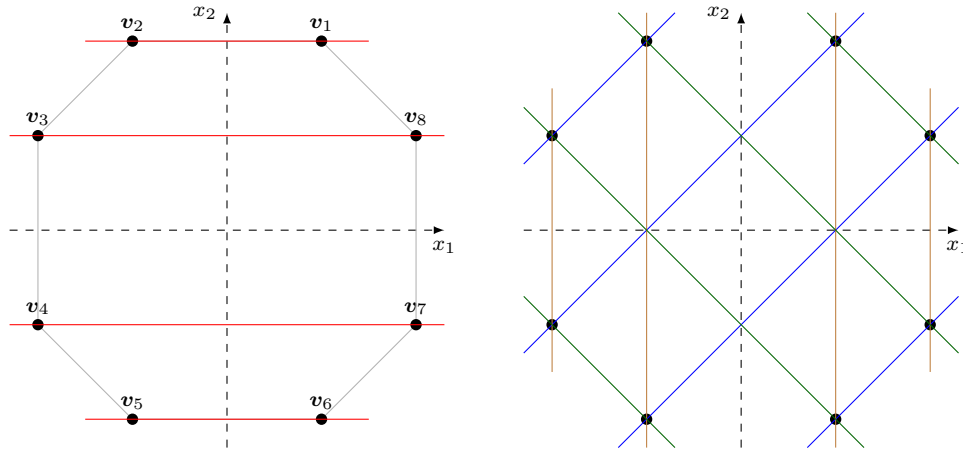
Figure 4.8. Left: *Starting with an arbitrary vertex of the grey polygon, and connecting it with its successor (here, starting with $\boldsymbol{v}_1$), we obtain a set of N parallel lines (here, N = 4) fulfilling (4.44). Right: Moving in anticlockwise direction, we find further $N - 1$ sets of parallel lines (blue, brown, and green lines) fulfilling (4.44).*

the frequency domain where we take $N + 1$ samples on each line. Remember that the origin is always taken as a sampling location. Hence, we need

$$(N + 1) \cdot (N + 1) - N = N \cdot (N + 1) + 1 = N^2 + N + 1$$

Fourier samples. This means that we need $N^2 - 2N$ sampling values more than we need in the case of adaptive sampling, where we only need $3N + 1$ Fourier samples.

Let us turn back to our original problem, i.e. the recovery of shift vectors and corresponding positive coefficients of a function of the form (4.28) using Fourier samples from predetermined lines through the origin in the frequency domain.

First, we take $N+1$ arbitrary distinct lines $\widetilde{l}_j$, $j = 1, \ldots, N+1$, through the origin in the frequency domain such that we obtain $N + 1$ univariate problems similar to (4.31) and (4.35). On each line, we use equispaced sampling locations with step size h where we always take the origin as a sampling location. Each of the lines $\widetilde{l}_j$ corresponds to a line l_j through the origin in the time domain ($j = 1, \ldots, N + 1$). Applying the Prony method, see Algorithm 2.5, to each of the $N + 1$ univariate problems, we obtain the orthogonal projections of all shift vectors onto the lines $l_1, \ldots, l_{N+1}$ with corresponding coefficients.

Secondly, the proof of Proposition 4.16 now provides a reconstruction scheme to recover the original shift vectors from these known orthogonal projections. Further, we have to compute the original corresponding coefficients. Remember that the corresponding coefficients which we obtain by applying the Prony method may be sums of coefficients that correspond to original shift vectors with same coordinate values.

Assume that the coefficients $c_1, \ldots, c_N$ in (4.28) are positive, and that we have given the orthogonal projections of the original shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ onto the lines $l_1, \ldots, l_{N+1}$, i.e. $v_{j,k}$ for $j \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, N+1\}$ with corresponding coefficients. Then we can compute the original shift vectors and the original corresponding coefficients by using the following algorithm:

Algorithm 4.19. (Reconstruction of shift vectors and corresponding coefficients using sampling on predetermined lines).

- **Input:**

  Orthogonal projections of the original shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ onto the lines $l_1, \ldots, l_{N+1}$ with corresponding coefficients. Thus, for each line l_k ($k \in \{1, \ldots, N+1\}$) the distinct coordinate values of the original shift vectors with respect to the lines $l_1, \ldots, l_{N+1}$ are given by

  $$\widetilde{V}_k := \big\{ \widetilde{\boldsymbol{v}}_{1,k}, \widetilde{\boldsymbol{v}}_{2,k}, \ldots, \widetilde{\boldsymbol{v}}_{N_k,k} \big\}$$

  with corresponding coefficients

  $$c_j^k := \sum_{\ell : v_{\ell,k} = \widetilde{v}_{j,k}} c_\ell, \quad j = 1, \ldots, N_k.$$

- **Computation:**

  1. Compute the set K of the original shift vectors, i.e.

     $$K := \big\{ \boldsymbol{v} \mid \mathrm{P}_{l_k}(\boldsymbol{v}) \in \widetilde{V}_k \text{ for } k = 1, \ldots, N+1 \big\}.$$

  2. For each of the lines $l_1, \ldots, l_{N+1}$, consider the associated extreme projecting lines.

  3. Determine a point $\boldsymbol{p}$ which is common to $r \geq 3$ extreme projecting lines and set

     $$\boldsymbol{v}_1 := \boldsymbol{p}.$$

     (We know from the proof of Proposition 4.16 that there exists at least one such point.)

  4. Among these r extreme projecting lines, find the line which goes only through $\boldsymbol{v}_1$ but not through other points in K. This line is perpendicular to a line l_{k_0} for an index $k_0 \in \{1, \ldots, N+1\}$. Determine the index $j_0 \in \{1, \ldots, N_k\}$ such that $\mathrm{P}_{l_{k_0}}(\boldsymbol{v}_1) = \widetilde{\boldsymbol{v}}_{j_0,k_0}$. Then set

     $$c_1 := c_{j_0}^{k_0}.$$

     (There exists at least one extreme projecting line which goes only through $\boldsymbol{v}_1$, see the proof of Proposition 4.16. Since we have found an extreme projecting line which goes only through $\boldsymbol{v}_1$, there are no summation effects such that $c_{j_0}^{k_0}$ is the original corresponding coefficient.)

5. Update the set K, the sets $\widetilde{V}_k$, $k = 1, \ldots, N+1$, and the corresponding coefficients c_j^k as follows:

   a) Set $K := K \setminus \{\boldsymbol{v}_1\}$.

   b) Set $\widetilde{V}_{k_0} := \widetilde{V}_{k_0} \setminus \{\widetilde{\boldsymbol{v}}_{j_0,k_0}\}$.

   c) For each $k \in \{1, \ldots, N+1\} \setminus \{k_0\}$, consider the set $\widetilde{V}_k$:

      – Determine the index $j_k \in \{1, \ldots, N_k\}$ such that $\mathrm{P}_{l_k}(\boldsymbol{v}_1) = \widetilde{\boldsymbol{v}}_{j_k,k}$.

      – If the projecting line which is associated with the line l_k and $\boldsymbol{v}_1$ contains no other point from K, then set

      $$\widetilde{V}_k := \widetilde{V}_k \setminus \{\widetilde{\boldsymbol{v}}_{j_k,k}\}.$$

      Otherwise, update the coefficient $c_{j_k}^k$ corresponding to $\widetilde{\boldsymbol{v}}_{j_k,k}$ by

      $$c_{j_k}^k := c_{j_k}^k - c_1.$$

6. Repeat the computation steps 1–5 until K is an empty set.

- **Output:** Shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ with corresponding coefficients $c_1, c_2, \ldots, c_N$ determining f in (4.28).

Remark 4.20.
In many practical cases, it is sufficient to take fewer than $N + 1$ sampling lines in order to uniquely determine the N shift vectors and the corresponding coefficients since the cases where N sampling lines do not suffice are sets of vectors with a very special structure, compare Procedure 4.18. Such an approach is used in [57], where in the case of parameter estimation for bivariate exponential sums at most $N + 1$ sampling lines are considered for the estimation of the frequency vectors $\boldsymbol{f}_j$ and the corresponding coefficients c_j in an exponential sum of the form

$$g(\boldsymbol{x}) := \sum_{j=1}^{N} c_j \, \mathrm{e}^{\mathrm{i}\langle \boldsymbol{f}_j, \boldsymbol{x}\rangle}, \quad \boldsymbol{x} \in \mathbb{R}^2.$$

First, the axes in the real plane are used as sampling lines. With regard to these two lines, two problems of parameter estimation for univariate exponential sums are obtained. The solutions to these two problems yield a set of feasible frequency vectors similar to the candidate set K we consider in Step 5 of Algorithm 4.12.

Then further sampling lines are considered in order to test if a feasible frequency vector is an actual frequency vector of g. But it is not directly clear if a third line suffices, or if a fourth, fifth, sixth line and so on is needed. This is only found out in the last

step of the computation of the coefficients, which is done by an overdetermined linear system. If the rank of the involved data matrix is not equal to the number of feasible frequency vectors, another sampling line is needed. This may happen repeatedly such that up to at most $N + 1$ sampling lines are needed.

Hence, also the method presented in [57] uses an adaptive sampling scheme in the sense that, due to the solvability of a linear system, it is decided if further sampling lines are needed. Those additional sampling lines are not chosen dependently on the already found set of feasible frequency vectors. Therefore, it may well happen that much more sampling values are needed than in our proposed reconstruction scheme in Subsection 4.2.1, where we choose the third sampling line dependently on the results obtained by using the sampling values from the first two lines. In this way, it is ensured that we do not need more than a third sampling line. Instead, a third line always suffices regardless of the structure of the set of the actual frequency or shift vectors.

4.3. Numerical results

Now we want to apply the reconstruction method using adaptive sampling from Subsection 4.2.1 in order to recover non-uniform translates of bivariate functions, i.e. functions f of the form (4.28), where we want to uniquely determine the shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$, and the coefficients $c_1, \ldots, c_N$.

First, remember Algorithm 4.12. In the first four steps of this algorithm, we restrict the function g defined in (4.30), which we consider for the reconstruction of f, to the ω_1- and the ω_2-axis in the frequency domain. Thus, we obtain two univariate problems. We apply the Prony method as described in Section 2.2 to these problems in order to obtain the sets $\widetilde{V}_1 = \{\widetilde{v}_{1,1}, \ldots, \widetilde{v}_{N_1,1}\}$ and $\widetilde{V}_2 = \{\widetilde{v}_{1,2}, \ldots, \widetilde{v}_{N_2,2}\}$ of the first and second coordinates of the true shift vectors respectively. Then we know that the true shift vectors are contained in the candidate set

$$K = \big\{ \boldsymbol{v} := (v_1, v_2)^{\mathrm{T}} \mid v_1 \in \widetilde{V}_1, \ v_2 \in \widetilde{V}_2 \big\},$$

and we determine a line $x_2 = \tan(\vartheta\pi)x_1$ with $\vartheta \in (0,1) \setminus \{\frac{1}{2}\}$ such that the orthogonal projections of all candidate vectors in K onto this line are distinct. We denote this line by l, and we have

$$l = \{\lambda\boldsymbol{u} \mid \lambda \in \mathbb{R}\}$$

with the normalized direction vector

$$\boldsymbol{u} := \begin{pmatrix} \cos(\vartheta\pi) \\ \sin(\vartheta\pi) \end{pmatrix}.$$

The parameter ϑ also determines the third sampling line in the frequency domain, i.e. $\omega_2 = \tan(\vartheta\pi)\omega_1$, to which we restrict the function g in the next step such that we

get a third univariate problem. Applying again the Prony method, we finally find the true shift vectors by comparing the results from the third problem with the candidate set, and we also compute the corresponding coefficients.

In order to improve the robustness of the reconstruction method, the parameter ϑ of the line l should be chosen in such a way that the minimal distance between two orthogonal projections of candidate vectors from K onto the line l is maximized. The orthogonal projection $P_l(\boldsymbol{v})$ of $\boldsymbol{v} \in K$ onto the line l is given by[7]

$$P_l(\boldsymbol{v}) = \langle \boldsymbol{v}, \boldsymbol{u} \rangle \boldsymbol{u} = \big(\cos(\vartheta\pi)v_1 + \sin(\vartheta\pi)v_2 \big) \boldsymbol{u},$$

and the distance between two projections $P_l(\boldsymbol{v})$ and $P_l(\boldsymbol{w})$, $\boldsymbol{v}, \boldsymbol{w} \in K$, $\boldsymbol{v} \neq \boldsymbol{w}$ is given by $\|P_l(\boldsymbol{v}) - P_l(\boldsymbol{w})\|_2$, for which we have

$$\begin{aligned}
\|P_l(\boldsymbol{v}) - P_l(\boldsymbol{w})\|_2^2 &= \big\| \langle \boldsymbol{v}, \boldsymbol{u} \rangle \boldsymbol{u} - \langle \boldsymbol{w}, \boldsymbol{u} \rangle \boldsymbol{u} \big\|_2^2 = \big\| \langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{u} \rangle \boldsymbol{u} \big\|_2^2 \\
&= \big(\langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{u} \rangle \big)^2 \|\boldsymbol{u}\|_2^2 \\
&= \big(\langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{u} \rangle \big)^2
\end{aligned}$$

since $\|\boldsymbol{u}\|_2 = 1$. Thus, in order to choose ϑ as mentioned, we need to maximize the minimal distance between two projections with respect to ϑ, that is, we have to solve the max-min problem

$$\max_{\vartheta \in (0,1) \setminus \{\frac{1}{2}\}} \min_{\substack{\boldsymbol{v}, \boldsymbol{w} \in K \\ \boldsymbol{v} \neq \boldsymbol{w}}} \big(\langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{u} \rangle \big)^2. \tag{4.45}$$

In the following numerical examples, this is done by considering a discrete set of values $\vartheta\pi$ rad for the angle between the positive ω_1-axis and the part of the line $\omega_2 = \tan(\vartheta\pi)\omega_1$ which lies in the upper half space. In particular, we consider the parameter ϑ for

$$\frac{1}{18} \leq \vartheta \leq \frac{4}{9} \quad \text{and} \quad \frac{5}{9} \leq \vartheta \leq \frac{17}{18} \tag{4.46}$$

in increments of $\frac{1}{720}$, that is, we consider angles from $10°$ to $80°$ and from $100°$ to $170°$ in $0.25°$ increments.

Now we have to take Fourier samples from the third sampling line. Note that we use clean data by simulating the Fourier data of the function f which we want to reconstruct by using Fourier samples. Since we use a discrete grid setting of size 128×128 such that we have given function values on all grid points, and since we use the discrete Fourier transform in order to obtain the simulated Fourier data, we are able to take Fourier samples on the lines $\omega_2 = 0$ and $\omega_1 = 0$ in the frequency domain. But, using this approach, we do not exactly have the Fourier data on the third sampling line $\omega_2 = \tan(\vartheta\pi)\omega_1$, where we want to consider equispaced sampling

---

[7]Cf. [39, p. 118, (1.88)], for example.

locations with step size h. Therefore, we use another approach to simulate the Fourier data of the third sampling line.

First, consider the case of a continuous function argument. The rotation around the fixed origin is a linear transformation. It is given by the rotation matrix

$$\boldsymbol{D}_\psi := \begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{pmatrix}$$

where ψ is the angle of the rotation, and the rotation is an orthogonal transformation (note $\boldsymbol{D}^{\mathrm{T}} = \boldsymbol{D}^{-1}$ and hence the preservation of the inner product, cf. [63, p. 134]). Thus, we can interchange the application of the Fourier transform and the rotation, see [63, p. 135, Theorem 1.1].

Let us turn back to the discrete setting. We consider the values of the function f on a square grid of size 128×128 where the values of the first and the second coordinate are ranging from -64 to 63 and from -63 to 64 respectively. Actually, we consider a matrix of size 128×128 such that the entries of the matrix are the function values of f at the grid points. Here, the first coordinate corresponds to the column index of the matrix, and the second coordinate corresponds to the row index. The origin of the coordinate system then corresponds with the matrix entry in the 65th row and 65th column of the matrix.

Aside from the horizontal and vertical axis in the frequency domain, the Fourier transform $\widehat{f}$ has to be sampled on a third, radial line. Since we also use a discrete setting in the frequency domain, we are not able to consider exact radial lines. In order to obtain exactly the data from the radial line $\omega_2 = \tan(\vartheta\pi)\omega_1$, we rotate the original function f, and afterwards we apply the Fourier transform. Actually, we use a so-called *passive rotation* in the time domain, meaning that we rotate the coordinate system anticlockwise such that the line $x_2 = \tan(\vartheta\pi)x_1$ in the original Cartesian coordinate system becomes the horizontal axis in the rotated coordinate system. By applying the Fourier transform afterwards, we can take Fourier samples on the horizontal axis in the frequency domain (i.e. 65th row in the discrete setting in the frequency domain) in order to obtain the Fourier data from the original third sampling line $\omega_2 = \tan(\vartheta\pi)\omega_1$.

We have explained that this is possible in the case of a continuous function argument. Now we use this approach in the discrete setting. For this purpose, we use the restriction that the shift vectors of the original function f are only lying in the disc determined by the *incircle of the discrete grid*. By this circle, we mean the incircle of the square domain which is the convex hull of the discrete grid.

For the following examples, we have taken the radial function

$$\Phi(x_1, x_2) := \exp\bigl(-\alpha \cdot (x_1^2 + x_2^2)\bigr)$$

with $\alpha := 0.05$. The original parameters are always denoted by $v_{j,1}$, $v_{j,2}$, and c_j for the coordinates of the shift vectors and the corresponding coefficients respectively. Parameters with the superscript * denote reconstructed parameters.

First, we have taken an original function which consists only of four summands, but where three shift vectors are lying closely to each other on the same vertical line, see Figure 4.9. Visually, one can only make out two shift vectors. The determining parameters for the displayed original function are listed in Table 4.1. In addition, also the absolute reconstruction errors between the original parameters and the reconstructed parameters are given in Table 4.1. We have used these parameters to evaluate the reconstructed function on the discrete grid and to compare it with the original function on this grid. In this way, we find a maximal absolute error between the original and reconstructed function of approximately $1.175 \cdot 10^{-8}$.
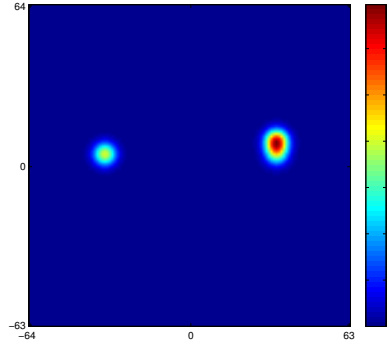
Figure 4.9. *Original function of form* (4.28), *determined by* $(\boldsymbol{v}_j)_{j=1}^4$, $(c_j)_{j=1}^4$ *in Table* 4.1.

j	$v_{j,1}$	$|v_{j,1} - v_{j,1}^*|$	$v_{j,2}$	$|v_{j,2} - v_{j,2}^*|$	c_j	$|c_j - c_j^*|$
1	34	0	5	$1.194 \cdot 10^{-11}$	3	$1.862 \cdot 10^{-10}$
2	-34	0	5	$1.194 \cdot 10^{-11}$	4	$1.408 \cdot 10^{-12}$
3	34	0	10	$1.779 \cdot 10^{-8}$	2	$5.141 \cdot 10^{-7}$
4	34	0	10.25	$7.993 \cdot 10^{-9}$	4	$5.143 \cdot 10^{-7}$

Table 4.1. *Parameters of the original function in* Figure 4.9 *and approximate reconstruction errors. The measure of the angle between the first sampling line and the adaptively determined third sampling line is* $80°$, *i.e.* $\vartheta = \frac{4}{9}$.

Secondly, we consider a linear combination of seven translates of $\varPhi$, see Figure 4.10. Again, some shift vectors are only slightly different, see Table 4.2, such that one can identify only five shift vectors visually. Evaluation of the reconstructed function on the grid yields a maximal absolute error of approximately $1.152 \cdot 10^{-7}$.

In the last example, see Figure 4.11, we have a function which has an 8-fold rotation symmetry if we only consider the shift vectors but not the coefficients. Thus, it is important to use adaptive sampling. By comparing the original function and the reconstructed function on the grid, we find a maximal absolute reconstruction error of approximately $3.338 \cdot 10^{-10}$.
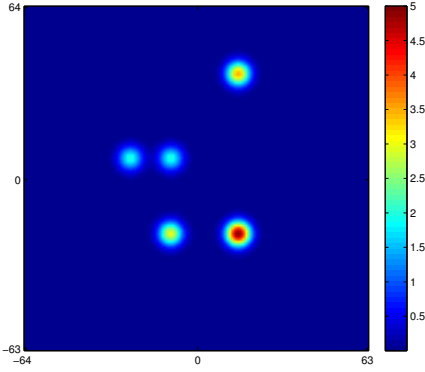
Figure 4.10. *Original function of the form* (4.28), *determined by* $(\boldsymbol{v}_j)_{j=1}^7$ *and* $(c_j)_{j=1}^7$ *given in* Table 4.2.
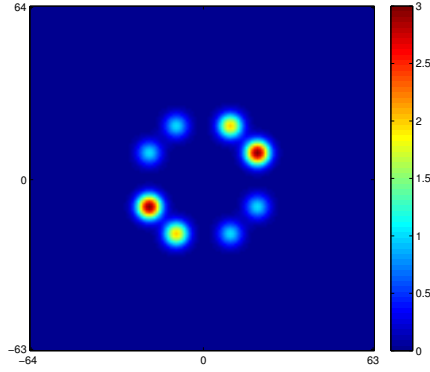
Figure 4.11. *Original function of the form* (4.28), *determined by* $(\boldsymbol{v}_j)_{j=1}^8$ *and* $(c_j)_{j=1}^8$ *given in* Table 4.3.

j	$v_{j,1}$	$|v_{j,1} - v_{j,1}^*|$	$v_{j,2}$	$|v_{j,2} - v_{j,2}^*|$	c_j	$|c_j - c_j^*|$
1	-25	$3.908 \cdot 10^{-14}$	8	0	2	$9.952 \cdot 10^{-13}$
2	-10.2	$2.959 \cdot 10^{-12}$	-19.9	$9.366 \cdot 10^{-11}$	1	$2.717 \cdot 10^{-6}$
3	-10	$1.847 \cdot 10^{-13}$	-20	$1.416 \cdot 10^{-11}$	2	$2.717 \cdot 10^{-6}$
4	-10	$1.847 \cdot 10^{-13}$	8	0	2	$5.64 \cdot 10^{-11}$
5	15	$5.329 \cdot 10^{-15}$	-20	$1.416 \cdot 10^{-11}$	5	$5.033 \cdot 10^{-11}$
6	15	$5.329 \cdot 10^{-15}$	39.25	$1.421 \cdot 10^{-12}$	3	$8.635 \cdot 10^{-11}$
7	15	$5.329 \cdot 10^{-15}$	40	$8.05 \cdot 10^{-12}$	0.5	$8.635 \cdot 10^{-11}$

Table 4.2. *Parameters of the original function in* Figure 4.10 *and approximate reconstruction errors. The measure of the angle between the first sampling line and the adaptively determined third sampling line is* $45°$, *i.e.* $\vartheta = \frac{1}{4}$.

j	$v_{j,1}$	$|v_{j,1} - v_{j,1}^*|$	$v_{j,2}$	$|v_{j,2} - v_{j,2}^*|$	c_j	$|c_j - c_j^*|$
1	-10	$1.776 \cdot 10^{-15}$	20	$1.776 \cdot 10^{-14}$	1	$2.847 \cdot 10^{-10}$
2	10	$1.954 \cdot 10^{-14}$	10	$1.776 \cdot 10^{-14}$	2	$1.004 \cdot 10^{-10}$
3	20	$1.066 \cdot 10^{-14}$	10	$1.03 \cdot 10^{-13}$	3	$3.338 \cdot 10^{-10}$
4	20	$1.066 \cdot 10^{-14}$	-10	$5.329 \cdot 10^{-15}$	1	$8.909 \cdot 10^{-12}$
5	10	$1.954 \cdot 10^{-14}$	-20	$2.842 \cdot 10^{-14}$	1	$8.074 \cdot 10^{-11}$
6	-10	$1.776 \cdot 10^{-15}$	-20	$2.842 \cdot 10^{-14}$	2	$5.079 \cdot 10^{-11}$
7	-20	$3.553 \cdot 10^{-15}$	-10	$5.329 \cdot 10^{-15}$	3	$1.29 \cdot 10^{-10}$
8	-20	$3.553 \cdot 10^{-15}$	10	$1.03 \cdot 10^{-13}$	1	$4.498 \cdot 10^{-11}$

Table 4.3. *Parameters of the original function in* Figure 4.11 *and approximate reconstruction errors. The measure of the angle between the first sampling line and the adaptively determined third sampling line is* $78.75°$, *i.e.* $\vartheta = 0.4375$.

5. Generalization to d-variate functions with $d > 2$

Now we want to generalize the theory of the reconstruction of non-uniform translates of bivariate functions to the d-variate case for $d \in \mathbb{N}$ with $d > 2$. In particular, we extend the approach from Subsection 4.2.1 to the d-dimensional setting, that is, we use adaptive sampling on $d + 1$ lines through the origin in the frequency domain in order to uniquely determine the shift vectors and coefficients of linear combinations of non-uniform translates. Here, adaptive sampling on $d + 1$ lines means that we use Fourier samples from the d axes in the frequency domain and from one further line which is chosen dependently on the results obtained by using the Fourier samples from the d axes.

Note that we could also generalize the sampling approach from Subsection 4.2.2, that is, we could use sampling on the d axes in the frequency domain and on some further predetermined lines. Suppose that we have a linear combination of N non-uniform translates. Then, by using the approach where the further sampling lines are not chosen dependently on the results obtained before, one may need up to $N-1$ additional sampling lines for each dimension $r = 2, \ldots, d$, see Remark 4.20 and [57, Section 6]. Using the method of adaptive sampling instead, we only need $d + 1$ sampling lines altogether as shown in the following section.

5.1. Reconstruction of non-uniform translates of d-variate functions using adaptive sampling

We consider functions $f : \mathbb{R}^d \to \mathbb{R}$ which have a sparse representation

$$f(\boldsymbol{x}) := \sum_{j=1}^{N} c_j \Phi(\boldsymbol{x} - \boldsymbol{v}_j) \tag{5.1}$$

with non-zero, real-valued coefficients c_j and real shift vectors $\boldsymbol{v}_j$ $(j = 1, \ldots, N)$. Further, we assume that the given d-variate function $\Phi \in L^1(\mathbb{R}^d)$ has a Fourier transform satisfying the condition $|\widehat{\Phi}(\boldsymbol{\omega})| > C > 0$ for $\|\boldsymbol{\omega}\|_2 < T$ for some positive constant T.

By applying the Fourier transform, we obtain

$$\widehat{f}(\boldsymbol{\omega}) = \left(\sum_{j=1}^{N} c_j \, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\omega}, \boldsymbol{v}_j \rangle} \right) \widehat{\Phi}(\boldsymbol{\omega}), \quad \boldsymbol{\omega} \in \mathbb{R}^d. \tag{5.2}$$

Using an analogous approach as in Subsection 4.2.1, we will show that we are able to uniquely determine the shift vectors $\boldsymbol{v}_j := (v_{j,k})_{k=1}^{d}$ as well as the coefficients c_j $(j = 1, \ldots, N)$ from $(d+1)N + 1$ Fourier samples. These samples are taken on $d + 1$ lines through the origin in the frequency domain where we use the d axes of the Cartesian coordinate system and one further sampling line. This $(d+1)$st sampling line is computed dependently on the results which are obtained by using the Fourier samples from the d axes.

Like in the bivariate case, we assume that the coefficients $c_1, \ldots, c_N$ have the same sign such that we avoid the problem of cancellation effects. We suppose that $c_j > 0$ for all $j \in \{1, \ldots, N\}$.

> **Theorem 5.1 (Reconstruction of d-variate non-uniform translates).**
> *Let $\Phi \in L^1(\mathbb{R}^d)$ be a given d-variate function with a Fourier transform satisfying $|\widehat{\Phi}(\boldsymbol{\omega})| > C$ for $\|\boldsymbol{\omega}\|_2 < T$ for some constants $C > 0$ and $T > 0$. Further, let f be a function with the sparse representation (5.1) where c_j and $\boldsymbol{v}_j = (v_{j,k})_{k=1}^{d}$ for $j = 1, \ldots, N$ are positive real-valued coefficients and real shift vectors respectively. Assume that the constant $h > 0$ fulfils the conditions $h\|\boldsymbol{v}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N\}$ and $h < \frac{T}{N}$. Then we get the following reconstruction result:*
>
> > *The function f can be uniquely recovered from the $(d+1)N + 1$ Fourier samples which are given by the set*
> >
> > $$\left\{ \widehat{f}(\boldsymbol{0}), \; \widehat{f}(\ell h \boldsymbol{e}_1), \ldots, \widehat{f}(\ell h \boldsymbol{e}_d), \; \widehat{f}(\ell h \boldsymbol{u}) \mid \ell = 1, \ldots, N \right\}$$
> >
> > *where $\boldsymbol{u}$ is the normalized direction vector of a line through the origin in the frequency domain which has to be chosen suitably.*
>
> *Here, $\boldsymbol{e}_k$, $k = 1, \ldots, d$, denote the kth unit vectors such that $(\boldsymbol{e}_k)_k = 1$.*

Proof. By considering (5.2) and the assumption about $\widehat{\Phi}$, we obtain

$$g(\boldsymbol{\omega}) := \frac{\widehat{f}(\boldsymbol{\omega})}{\widehat{\Phi}(\boldsymbol{\omega})} = \sum_{j=1}^{N} c_j \, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\omega}, \boldsymbol{v}_j \rangle}$$

for $\boldsymbol{\omega} := (\omega_k)_{k=1}^{d} \in \mathbb{R}^d$ with $\|\boldsymbol{\omega}\|_2 < T$.

In the first part of the proof, we restrict the function g to each of the axes of the Cartesian coordinate system. Thus, we get d univariate problems to which we can

apply the Prony method. Then, based on the results obtained by the application of the Prony method to these d problems, we choose another sampling line such that we have another univariate problem by restricting g to the $(d+1)$st sampling line. Finally, we can compute the original shift vectors and coefficients by using Fourier samples from this adaptively determined sampling line.

Part 1: For each $k \in \{1, \ldots, d\}$, consider the univariate function $g_k : \mathbb{R} \to \mathbb{C}$ defined by

$$g_k(\omega_k) := \frac{\widehat{f}(\omega_k \boldsymbol{e}_k)}{\widehat{\Phi}(\omega_k \boldsymbol{e}_k)} = \sum_{j=1}^{N} c_j \, \mathrm{e}^{-\mathrm{i}\,\omega_k v_{j,k}}, \quad \omega_k \in \mathbb{R},$$

where $\boldsymbol{e}_k$ denotes the kth unit vector such that $(\boldsymbol{e}_k)_k = 1$.

Assume that the set $\{v_{1,k}, \ldots, v_{N,k}\}$ of true shift coordinates contains $N_k \leq N$ distinct values $\widetilde{v}_{1,k} < \widetilde{v}_{2,k} < \ldots < \widetilde{v}_{N_k,k}$ $(k = 1, \ldots, d)$. This leads to

$$g_k(\omega_k) = \sum_{j=1}^{N_k} c_j^k \, \mathrm{e}^{-\mathrm{i}\,\omega_k \widetilde{v}_{j,k}} \quad \text{for } k = 1, \ldots, d \tag{5.3}$$

where we use the following definition for $k = 1, \ldots, d$:

$$c_j^k := \sum_{\ell:\, v_{\ell,k} = \widetilde{v}_{j,k}} c_\ell, \quad j = 1, \ldots, N_k.$$

Since the coefficients c_j^k, $j = 1, \ldots, N_k$, in each of the d problems in (5.3) are real-valued, we can apply the Prony method from Section 2.2, in particular the steps 1–5 in Algorithm 2.5, to each of these problems in order to obtain the sets

$$\widetilde{V}_k := \{\widetilde{v}_{1,k}, \ldots, \widetilde{v}_{N_k,k}\}, \quad k = 1, \ldots, d,$$

of the original shift coordinates. For this purpose, we use the sets

$$\{\widehat{f}(\ell h \boldsymbol{e}_k) \mid \ell = 0, \ldots, N\}, \quad k = 1, \ldots, d,$$

of Fourier samples from which we can compute the function values

$$g_k(0), \ g_k(h), \ g_k(2h), \ \ldots, \ g_k(Nh)$$

for $k = 1, \ldots, d$.

Observe that the condition for the application of the Prony method is fulfilled for all $k \in \{1, \ldots, d\}$ since we have $h|\widetilde{v}_{j,k}| \leq h\|\widetilde{\boldsymbol{v}}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N_k\}$ by assumption. Further, the fractions

$$g_k(\ell h) = \frac{\widehat{f}(\ell h \boldsymbol{e}_k)}{\widehat{\Phi}(\ell h \boldsymbol{e}_k)}, \quad \ell = 0, \ldots, N,$$

are well defined for all $k \in \{1, \ldots, d\}$ since we have $\widehat{\Phi}(\ell h \boldsymbol{e}_k) \neq 0$ for $\ell = 0, \ldots, N$ by the assumptions $h < \frac{T}{N}$ and $|\widehat{\Phi}(\boldsymbol{\omega})| > C > 0$ for $\|\boldsymbol{\omega}\|_2 < T$.

Part 2: Knowing the sets $\{\widetilde{v}_{1,k}, \ldots, \widetilde{v}_{N_k,k}\}$, $k = 1, \ldots, d$, of shift coordinates, we can compute the Cartesian product

$$K := \{\boldsymbol{v} := (v_k)_{k=1}^d \mid v_k \in \widetilde{V}_k, \ k = 1, \ldots, d\}, \tag{5.4}$$

which contains all true shift vectors and altogether $\prod_{k=1}^d N_k \leq N^d$ *candidate vectors* for the true shift vectors.

In the next part of the proof, we want to determine the true shift vectors in this candidate set and the corresponding coefficients. For this purpose, we choose a line in the time domain, denoted by $l := \{\lambda \boldsymbol{u} \mid \lambda \in \mathbb{R}\}$ with $\boldsymbol{u} \in \mathbb{R}^d$ where $\|\boldsymbol{u}\|_2 = 1$, such that the orthogonal projections of all candidate vectors in K onto this line l are pairwise different, see Figure 5.1. Then the direction vector $\boldsymbol{u}$ determines the $(d+1)$st sampling line in the frequency domain, that is, we take the Fourier samples

$$\widehat{f}(\ell h \boldsymbol{u}), \quad \ell = 0, 1, \ldots, N.$$

Part 3: Now we consider the univariate function $g_{d+1} : \mathbb{R} \to \mathbb{C}$ defined by

$$g_{d+1}(\omega) := \frac{\widehat{f}(\omega \boldsymbol{u})}{\widehat{\Phi}(\omega \boldsymbol{u})} = \sum_{j=1}^N c_j \, \mathrm{e}^{-\mathrm{i}\,\omega \langle \boldsymbol{u}, \boldsymbol{v}_j \rangle} = \sum_{j=1}^N c_j \, \mathrm{e}^{-\mathrm{i}\,\omega v_{j,d+1}}, \quad \omega \in \mathbb{R},$$

with

$$v_{j,d+1} := \langle \boldsymbol{u}, \boldsymbol{v}_j \rangle, \quad j = 1, \ldots, N.$$

Note that these values are coordinate values of the true shift vectors with respect to the line $l = \{\lambda \boldsymbol{u} \mid \lambda \in \mathbb{R}\}$, and that these values are pairwise different because all possible shift vectors in K yield, due to the choice of $\boldsymbol{u}$, different orthogonal projections onto the line l. Thus, there are no summation effects for the coefficients.

Let us order the coordinate values $v_{j,d+1}$ such that we obtain

$$g_{d+1}(\omega \boldsymbol{u}) = \sum_{j=1}^N c_j^{d+1} \, \mathrm{e}^{-\mathrm{i}\,\omega \widetilde{v}_{j,d+1}} \tag{5.5}$$

with

$$\widetilde{v}_{1,d+1} < \widetilde{v}_{2,d+1} < \ldots < \widetilde{v}_{N,d+1}$$

and corresponding coefficients $c_1^{d+1}, \ldots, c_N^{d+1}$ where $\{\widetilde{v}_{1,d+1}, \ldots, \widetilde{v}_{N,d+1}\}$ is a permuted version of $\{v_{1,d+1}, \ldots, v_{N,d+1}\}$, and the same permutation maps $\{c_1, \ldots, c_N\}$ onto $\{c_1^{d+1}, \ldots, c_N^{d+1}\}$.
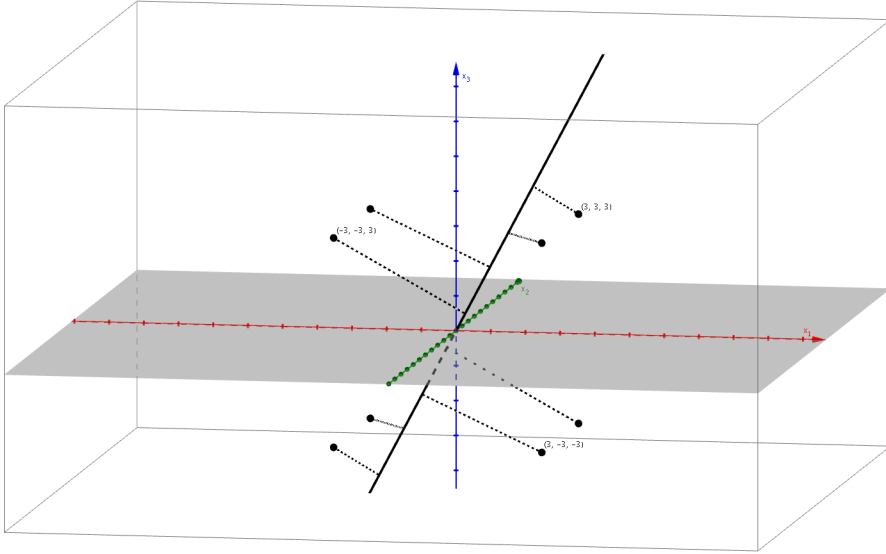
Figure 5.1. *Example for the determination of the line $l = \{\lambda \boldsymbol{u} \mid \lambda \in \mathbb{R}\}$ (black, solid line) in the time domain (displayed example: space $\mathbb{R}^3$, i.e. $d = 3$). The points $\boldsymbol{v}_1 := (3,3,3)^{\mathrm{T}}$, $\boldsymbol{v}_2 := (-3,-3,3)^{\mathrm{T}}$, and $\boldsymbol{v}_3 := (3,-3,-3)^{\mathrm{T}}$ are original shift vectors, while the other points (black dots) are further candidates in K. The direction vector $\boldsymbol{u}$ then determines the $(d+1)$st sampling line in the frequency domain.*

Then, by applying of the Prony method as described in Section 2.2 to the function g_{d+1}, we find the set

$$\widetilde{V}_{d+1} := \{\widetilde{v}_{1,d+1}, \dots, \widetilde{v}_{N,d+1}\}$$

and corresponding coefficients c_j^{d+1}, $j = 1, \dots, N$. Observe here that the conditions for the Prony method are fulfilled analogously to *Part 1* of the proof on hand.

Part 4: Now we have to find the true shift vectors in the candidate set K in (5.4). This can be done by comparison of the set K with the set $\widetilde{V}_{d+1}$. We determine all vectors $\boldsymbol{v} = (v_k)_{k=1}^d$ in K for which there exist indices $j \in \{1, \dots, N\}$ such that $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \widetilde{v}_{j,d+1}$. Then the set

$$G := \Big\{ \boldsymbol{v} = (v_k)_{k=1}^d \mid v_k \in \widetilde{V}_k \text{ for } k = 1, \dots, d, \ \langle \boldsymbol{u}, \boldsymbol{v} \rangle \in \widetilde{V}_{d+1} \Big\}$$

contains only the N true shift vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N$, and each of the vectors $\boldsymbol{v} \in G$ is uniquely associated with a coefficient c_j^{d+1} for $j \in \{1, \dots, N\}$. Thus, we have

$$G = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_N\} \tag{5.6}$$

with corresponding coefficients

$$c_1, c_2, \ldots, c_N. \tag{5.7}$$

The parameters of the function f in (5.1), which we wanted to recover, are now given by the elements of G in (5.6) and the corresponding coefficients in (5.7). $\square$

Summarizing the steps in the above proof, we obtain, as a generalization of Algorithm 4.12, the following algorithm for the reconstruction of non-uniform translates of d-variate functions where we assume that the coefficients c_j, $j = 1, \ldots, N$, in (5.1) are positive.

Algorithm 5.2 (Reconstruction of d-variate non-uniform translates).

- **Input:**
  - $\Phi \in L^1(\mathbb{R}^d)$ with $|\widehat{\Phi}(\boldsymbol{\omega})| > C$ for $\|\boldsymbol{\omega}\|_2 < T$ for some $C > 0$ and $T > 0$;
  - step size $h > 0$ with $h\|\boldsymbol{v}_j\|_2 < \pi$ for $j = 1, \ldots, N$ and $h < \frac{T}{N}$;
  - $(d + 1)N + 1$ Fourier samples, given by the set

  $$\left\{ \widehat{f}(\boldsymbol{0}),\ \widehat{f}(\ell h \boldsymbol{e}_1), \ldots, \widehat{f}(\ell h \boldsymbol{e}_d),\ \widehat{f}(\ell h \boldsymbol{u}) \mid \ell = 1, \ldots, N \right\}$$

  where $\boldsymbol{u}$ is the normalized direction vector of a line through the origin in the frequency domain which has to be chosen suitably, see Step 4 of the computation.

- **Computation:**
  1. Compute the function values

  $$g_k(0),\ g_k(h),\ g_k(2h),\ \ldots,\ g_k(Nh)$$

  for each $k \in \{1, \ldots, d\}$.

  2. Use Algorithm 2.5, Steps 1–5 in order to compute the coordinate values $\widetilde{v}_{j,k}$, $j = 1, \ldots, N_k$, in each of the problems in (5.3).

  3. Compute the Cartesian product of the sets $\widetilde{V}_k = \{\widetilde{v}_{1,k} \ldots, \widetilde{v}_{N_k,k}\}$ where $k = 1, \ldots, d$, as the set of possible candidates for the true shift vectors, i.e.

  $$K := \left\{ \boldsymbol{v} := (v_k)_{k=1}^d \mid v_k \in \widetilde{V}_k,\ k = 1, \ldots, d \right\}.$$

  4. Choose a normalized direction vector $\boldsymbol{u} \in \mathbb{R}^d$ such that the orthogonal projections of all candidate vectors in K onto the line $l = \{\lambda \boldsymbol{u} \mid \lambda \in \mathbb{R}\}$ are distinct.

5. Acquire the N Fourier samples $\widehat{f}(\ell h \boldsymbol{u})$, $\ell = 1, \ldots, N$. Then compute

$$g(\ell h \boldsymbol{u}) = \frac{\widehat{f}(\ell h \boldsymbol{u})}{\widehat{\Phi}(\ell h \boldsymbol{u})}, \quad \ell = 1, \ldots, N.$$

6. Use Algorithm 2.5 in order to compute the coordinate values $\widetilde{v}_{j,d+1}$ and the corresponding coefficients c_j^{d+1}, $j = 1, \ldots, N$, in (5.5).

7. Determine the true shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ in K by comparison with the set $\widetilde{V}_{d+1} = \{\widetilde{v}_{1,d+1}, \ldots, \widetilde{v}_{N,d+1}\}$, that is, determine all vectors $\boldsymbol{v} = (v_k)_{k=1}^d$ in K for which there exist indices $j \in \{1, \ldots, N\}$ such that $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \widetilde{v}_{j,d+1}$.

   Thus, we obtain the set $G = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$ with corresponding coefficients $c_1, c_2, \ldots, c_N$.

- **Output:** Shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ with corresponding coefficients $c_1, c_2, \ldots, c_N$ determining f in (5.1).

5.2. Numerical results

In analogy to the bivariate case, we want to show some numerical examples for the recovery of non-uniform translates of d-variate functions where we consider the case $d = 3$, that is, we consider functions f of the form (5.1), and we aim to uniquely determine the shift vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$, and the coefficients $c_1, \ldots, c_N$ in a three-dimensional setting. The Fourier data are simulated using an approach similar to the one used in the bivariate case.

Remember that we use Fourier samples from the ω_1-, ω_2-, and ω_3-axis in the frequency domain in order to obtain sets of distinct values for the coordinates of the true shift vectors. By considering the Cartesian product of these sets, we construct a set K of candidates for the true shift vectors. Then the fourth sampling line $\widetilde{l} := \{\lambda \boldsymbol{u} \mid \lambda \in \mathbb{R}\}$ in the frequency domain is chosen in such a way that the minimal distance between two orthogonal projections of candidate vectors onto the line $l := \{\lambda \boldsymbol{u} \mid \lambda \in \mathbb{R}\}$ in the time domain is maximized. As we have seen in the bivariate case, see Section 4.3, this leads to the problem of maximizing

$$\min_{\substack{\boldsymbol{v}, \boldsymbol{w} \in K \\ \boldsymbol{v} \neq \boldsymbol{w}}} \left(\langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{u} \rangle \right)^2.$$

Here, the direction vector $\boldsymbol{u}$ is given by

$$\boldsymbol{u} := \left(\cos(\psi) \sin(\theta), \, \sin(\psi) \sin(\theta), \, \cos(\theta) \right)^{\mathrm{T}} \tag{5.8}$$

with

$$(\psi, \theta) \in \left(\left[-\tfrac{\pi}{2}, \tfrac{\pi}{2} \right] \times (0, \pi) \right) \setminus \left\{ \left(0, \tfrac{\pi}{2} \right), \left(\tfrac{\pi}{2}, \tfrac{\pi}{2} \right) \right\} =: S.$$

Observe that $\boldsymbol{u}$ is given in spherical coordinates, and the restrictions for the angles ψ and θ are, on the one hand, due to the convention that the direction vector $\boldsymbol{u}$ should have a non-negative first coordinate. On the other hand, the fourth sampling line should be really a fourth line and not one of the axes of the coordinate system, which are already considered as sampling lines. Thus, we obtain the max-min problem

$$\max_{(\psi,\theta)\in S} \min_{\substack{\boldsymbol{v},\boldsymbol{w}\in K \\ \boldsymbol{v}\neq\boldsymbol{w}}} \left(\langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{u} \rangle \right)^2$$

for the determination of the fourth sampling line.

Since we use a discrete grid setting of size $128 \times 128 \times 128$ for our numerical examples, we are not able to exactly sample on radial lines in the frequency domain. Therefore, in order to obtain exactly the Fourier data from the fourth sampling line $\widetilde{l} = \{\lambda \boldsymbol{u} \mid \lambda \in \mathbb{R}\}$, we rotate the coordinate system in the time domain such that the line $l = \{\lambda \boldsymbol{u} \mid \lambda \in \mathbb{R}\}$ in the original coordinate system becomes the x_1-axis in the rotated coordinate system. After applying the Fourier transform, we can take Fourier samples on the ω_1-axis in the frequency domain in order to acquire the Fourier data from the original fourth sampling line $\widetilde{l}$.

We use the following approach for the rotation of the coordinate system in the time domain:

- Using the cross product, determine the unit vector $\boldsymbol{v}$ which is orthogonal to the plane spanned by $\boldsymbol{e}_1$ and $\boldsymbol{u}$. The direction of $\boldsymbol{v}$ is then given by the right-hand rule[1]. Remember the convention that the first coordinate of $\boldsymbol{u}$ is non-negative. In particular, we compute

$$\boldsymbol{u}^{\perp} := \boldsymbol{e}_1 \times \boldsymbol{u} = \begin{pmatrix} 0 \\ -u_3 \\ u_2 \end{pmatrix}$$

  and

$$\widetilde{\boldsymbol{u}} := \frac{\boldsymbol{u}^{\perp}}{\|\boldsymbol{u}^{\perp}\|_2}.$$

- Determine the angle between the vectors[2] $\boldsymbol{u}$ and $\boldsymbol{e}_1$, i.e.

$$\varrho := \arccos(u_1) \in \left(0, \tfrac{\pi}{2} \right].$$

---

[1] See [43, pp. 167–168], for example.
[2] See [43, p. 158, Definition 12.7, (1)], for example.

- Rotate the original coordinate system around the line $l^\perp := \{\lambda \widetilde{\boldsymbol{u}} \,|\, \lambda \in \mathbb{R}\}$ by the angle ϱ, and hence obtain a new coordinate system where the x_1-axis is now the line $l = \{\lambda \boldsymbol{u} \,|\, \lambda \in \mathbb{R}\}$ from the original coordinate system. Each point $\boldsymbol{p}_{\text{new}}$ of the rotated system is given by

$$\boldsymbol{p}_{\text{new}} = \boldsymbol{D}_\varrho \boldsymbol{p}$$

where $\boldsymbol{p}$ is the corresponding point in the original system. Further, $\boldsymbol{D}_\varrho$ denotes the rotation matrix for the described rotation and is given by[3]

$$\boldsymbol{D}_\varrho := \begin{pmatrix} D_{\varrho,1} & D_{\varrho,2} & D_{\varrho,3} \end{pmatrix}$$

with

$$D_{\varrho,1} := \begin{pmatrix} \widetilde{u}_1^2\big(1 - \cos(\varrho)\big) + \cos(\varrho) \\ \widetilde{u}_2\widetilde{u}_1\big(1 - \cos(\varrho)\big) + \widetilde{u}_3\sin(\varrho) \\ \widetilde{u}_3\widetilde{u}_1\big(1 - \cos(\varrho)\big) - \widetilde{u}_2\sin(\varrho) \end{pmatrix},$$

$$D_{\varrho,2} := \begin{pmatrix} \widetilde{u}_1\widetilde{u}_2\big(1 - \cos(\varrho)\big) - \widetilde{u}_3\sin(\varrho) \\ \widetilde{u}_2^2\big(1 - \cos(\varrho)\big) + \cos(\varrho) \\ \widetilde{u}_3\widetilde{u}_2\big(1 - \cos(\varrho)\big) + \widetilde{u}_1\sin(\varrho) \end{pmatrix},$$

and

$$D_{\varrho,3} := \begin{pmatrix} \widetilde{u}_1\widetilde{u}_3\big(1 - \cos(\varrho)\big) + \widetilde{u}_2\sin(\varrho) \\ \widetilde{u}_2\widetilde{u}_3\big(1 - \cos(\varrho)\big) - \widetilde{u}_1\sin(\varrho) \\ \widetilde{u}_3^2\big(1 - \cos(\varrho)\big) + \cos(\varrho) \end{pmatrix}.$$

In analogy to the bivariate case, we consider linear combinations of translates of the radial function

$$\varPhi(x,y,z) := \exp\big(-\alpha \cdot (x^2 + y^2 + z^2)\big)$$

with $\alpha := 0.05$. The original parameters are denoted by $v_{j,1}$, $v_{j,2}$, $v_{j,3}$, and c_j for the coordinates of the shift vectors and the corresponding coefficients respectively, while parameters with the superscript * denote reconstructed parameters. We evaluate the reconstructed function on the discrete grid such that we can compute a maximal absolute reconstruction error by comparison with the original function on the grid. Observe that in the second and third example we consider functions where some shift vectors are lying closely to each other.

---

[3]Cf. [39, pp. 447–448].
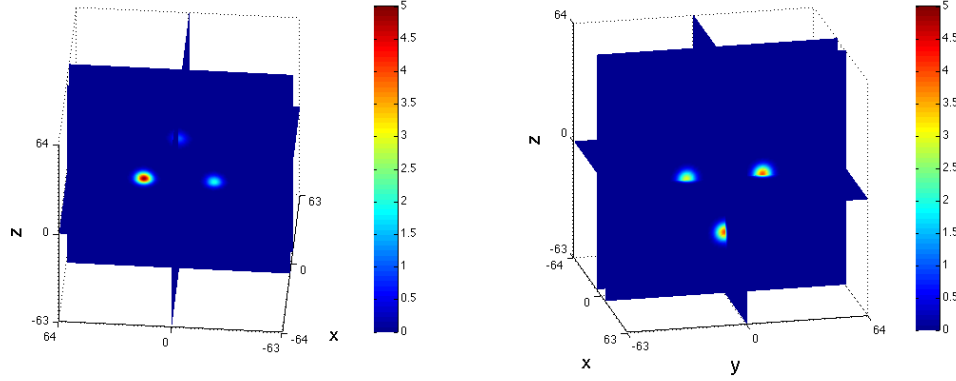
Figure 5.2. *Original function of the form* (5.1), *determined by* $(\boldsymbol{v}_j)_{j=1}^6$ *and* $(c_j)_{j=1}^6$ *given in* Table 5.1.

j	$v_{j,1}$	$v_{j,2}$	$v_{j,3}$	c_j	$\|v_{j,1} - v_{j,1}^*\|$	$\|v_{j,2} - v_{j,2}^*\|$	$\|v_{j,3} - v_{j,3}^*\|$	$\|c_j - c_j^*\|$
1	10	-20	0	3	$1.954 \cdot 10^{-14}$	$1.066 \cdot 10^{-14}$	$4.533 \cdot 10^{-15}$	$1.023 \cdot 10^{-11}$
2	10	20	0	4	$1.954 \cdot 10^{-14}$	$3.553 \cdot 10^{-15}$	$4.533 \cdot 10^{-15}$	$5.085 \cdot 10^{-12}$
3	-10	-20	0	2	$5.329 \cdot 10^{-15}$	$1.066 \cdot 10^{-14}$	$4.533 \cdot 10^{-15}$	$9.700 \cdot 10^{-12}$
4	-10	20	0	5	$5.329 \cdot 10^{-15}$	$3.553 \cdot 10^{-15}$	$4.533 \cdot 10^{-15}$	$5.600 \cdot 10^{-12}$
5	-10	0	30	1	$5.329 \cdot 10^{-15}$	$2.072 \cdot 10^{-14}$	$7.105 \cdot 10^{-15}$	$5.818 \cdot 10^{-14}$
6	10	0	-30	4	$1.954 \cdot 10^{-14}$	$2.072 \cdot 10^{-14}$	$3.553 \cdot 10^{-15}$	$7.416 \cdot 10^{-14}$

Table 5.1. *Parameters of the original function in* Figure 5.2 *and approximate reconstruction errors. Maximal absolute reconstruction error:* $1.023 \cdot 10^{-11}$. *The fourth sampling line is given by the direction vector* $\boldsymbol{u}$ *of form* (5.8) *where the measure of the angle* ψ *is* $-64°$, *and the measure of the angle* θ *is* $29°$.
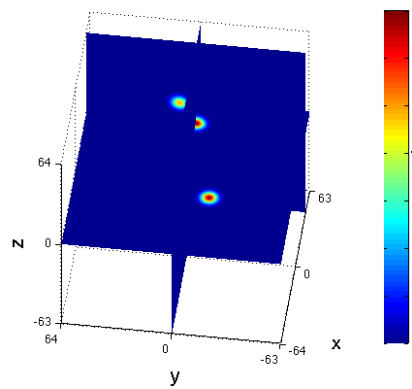
Figure 5.3. *Original function of the form* (5.1), *determined by* $(\boldsymbol{v}_j)_{j=1}^4$ *and* $(c_j)_{j=1}^4$ *given in* Table 5.2.

j	$v_{j,1}$	$v_{j,2}$	$v_{j,3}$	c_j	$\lvert v_{j,1} - v_{j,1}^* \rvert$	$\lvert v_{j,2} - v_{j,2}^* \rvert$	$\lvert v_{j,3} - v_{j,3}^* \rvert$	$\lvert c_j - c_j^* \rvert$
1	45	0	0	4	$1.269 \cdot 10^{-11}$	$7.208 \cdot 10^{-9}$	$1.16 \cdot 10^{-11}$	$6.838 \cdot 10^{-6}$
2	45	0.25	-0.5	3	$1.269 \cdot 10^{-11}$	$9.806 \cdot 10^{-9}$	$3.882 \cdot 10^{-11}$	$7.222 \cdot 10^{-6}$
3	45	10	15	5	$1.269 \cdot 10^{-11}$	$1.737 \cdot 10^{-12}$	$3.908 \cdot 10^{-14}$	$1.558 \cdot 10^{-7}$
4	-15	-15	0	7	$5.511 \cdot 10^{-11}$	$1.954 \cdot 10^{-13}$	$1.16 \cdot 10^{-11}$	$2.284 \cdot 10^{-7}$

Table 5.2. *Parameters of the original function in* Figure 5.3 *and approximate reconstruction errors. Maximal absolute reconstruction error:* $9.866 \cdot 10^{-7}$. *The fourth sampling line is given by the direction vector $\boldsymbol{u}$ of form* (5.8) *where the measure of the angle ψ is* $79°$, *and the measure of the angle θ is* $44°$.
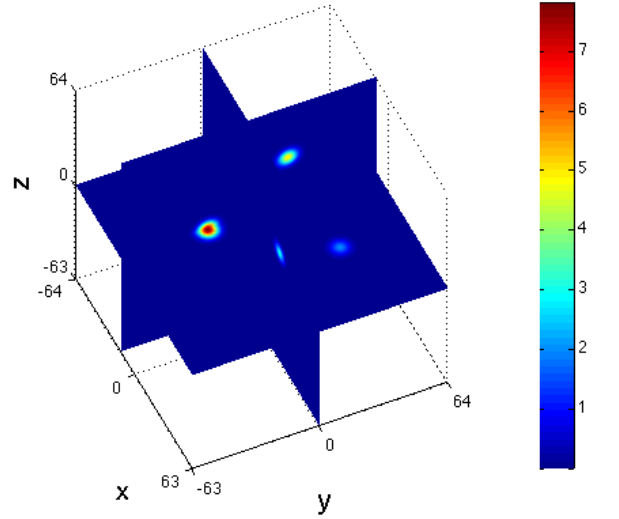
Figure 5.4. *Original function of the form* (5.1), *determined by $(\boldsymbol{v}_j)_{j=1}^5$ and $(c_j)_{j=1}^5$ given in* Table 5.3.

j	$v_{j,1}$	$v_{j,2}$	$v_{j,3}$	c_j	$\lvert v_{j,1} - v_{j,1}^* \rvert$	$\lvert v_{j,2} - v_{j,2}^* \rvert$	$\lvert v_{j,3} - v_{j,3}^* \rvert$	$\lvert c_j - c_j^* \rvert$
1	20	0	10	3	$3.553 \cdot 10^{-15}$	$1.852 \cdot 10^{-14}$	$1.421 \cdot 10^{-14}$	$7.389 \cdot 10^{-10}$
2	20	30	0	2	$3.553 \cdot 10^{-15}$	$2.487 \cdot 10^{-14}$	$2.328 \cdot 10^{-14}$	$5.677 \cdot 10^{-12}$
3	-15	20	30	5	$5.347 \cdot 10^{-13}$	$2.842 \cdot 10^{-14}$	$3.553 \cdot 10^{-15}$	$6.188 \cdot 10^{-12}$
4	-15	-20	0	4	$5.347 \cdot 10^{-13}$	$1.776 \cdot 10^{-14}$	$2.328 \cdot 10^{-14}$	$1.318 \cdot 10^{-7}$
5	-14	-20	0	4	$8.207 \cdot 10^{-13}$	$1.776 \cdot 10^{-14}$	$2.328 \cdot 10^{-14}$	$1.311 \cdot 10^{-7}$

Table 5.3. *Parameters of the original function in* Figure 5.4 *and approximate reconstruction errors. Maximal absolute reconstruction error:* $2.515 \cdot 10^{-8}$. *The fourth sampling line is given by the direction vector $\boldsymbol{u}$ of form* (5.8) *where the measure of the angle ψ is* $-58°$, *and the measure of the angle θ is* $78°$.

6. Reconstruction of polygonal shapes from sparse Fourier samples

In this last chapter of the main part, we want to consider a further reconstruction problem in two dimensions which is completely different from the two-dimensional problems dealt with in Chapter 4. Here, we want to reconstruct polygonal shapes in the real plane from as few Fourier samples as possible, that is, we want to recover an original polygonal domain D with N vertices by using sparse sampling values of the Fourier transform of the characteristic function of the polygonal domain. We only consider *simply-connected polygons*, i.e. polygons with non-intersecting edges. For this purpose, we need to reconstruct the vertices of the polygon, and we need to know what the interior of the polygon looks like, that is, we have to reconstruct the order of the vertices in order to know which vertices are connected by line segments. This is a very important step, especially in the case of non-convex polygons.

Considering convex polygons, it is sufficient to know the vertices in order to obtain a reconstruction of the polygon. But in the case of concave polygons with at least four vertices, there are always several distinct polygons which have the same vertices. Thus, we use an approach consisting of two steps. First, we compute the vertices by applying the Prony method from Section 2.2 to Fourier samples taken on three lines through the origin in the frequency domain. Here, we use, similarly to the approach in Subsection 4.2.1, adaptive sampling. This means that the third sampling line is chosen dependently on the resulting candidate set obtained from processing the Fourier data from the first two sampling lines. The application of the Prony method does not only yield the vertices but also corresponding coefficients which we use in the second step of our approach to determine the order of the vertices.

A similar problem of reconstructing polygons from some given data has been examined in [21, 27, 47]. In these papers, the given data are complex moments, i.e. integral moments of the analytic power function $f(z) = z^k$ over the characteristic function of the polygonal domain D with $z = x + \mathrm{i}\, y$, $x, y \in \mathbb{R}$. The given data are of the form

$$k(k-1) \int_D z^{k-2}\, \mathrm{d}(x,y) = \sum_{j=1}^{N} a_j z_j^k, \quad k = 0, \ldots, 2N-1, \tag{6.1}$$

where z_j are the vertices of the polygon in the complex plane, and Prony-like methods are employed in order to determine the vertices. In [27, 47], it is discussed that, even

if besides the vertices also the coefficients a_j in (6.1) are given, there are some cases such that the interior of the polygon is not uniquely determined. Also in cases of a unique solution, the step of ordering the vertices to determine the interior of the polygon yields the problem of deciding on the right configuration of up to at most 2^{N-1} possible scenarios for the sides of the polygon, see [27, Subsection 3.2], that is, it leads to a problem in computational geometry and graph theory, see [20].

In the following subsection, we will show that our proposed approach for the reconstruction of polygons from sparse Fourier samples always leads to the unique determination of the polygon by using adaptive sampling on three lines in the frequency domain, that is, we need $3N$ Fourier samples in order to determine the N vertices and their right order.

6.1. Reconstruction of polygonal shapes in the space $\mathbb{R}^2$

Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$ of the special form

$$f(\boldsymbol{x}) := \mathbf{1}_D(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^2, \tag{6.2}$$

where $\mathbf{1}_D$ is the characteristic function of the domain $D \subset \mathbb{R}^2$. Here, D is a polygonal domain determined by the N vertices $\boldsymbol{v}_j \in \mathbb{R}^2$, $j = 1, \ldots, N$, which are numbered anticlockwise. In the following, we will refer to the function f as *unit-height polygon*.
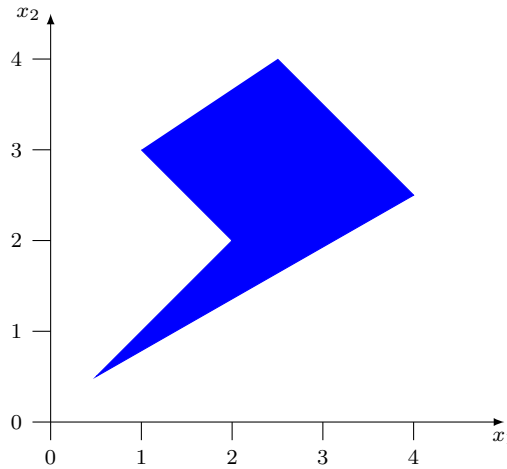
Figure 6.1. *Example for a function of the form* (6.2) *with* 5 *vertices.*

We gain to reconstruct the domain D from sparse Fourier samples. The Fourier transform of f is given by

$$\widehat{f}(\boldsymbol{\xi}) = \int\limits_{\mathbb{R}^2} \mathbf{1}_D(\boldsymbol{x}) \, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle} \, \mathrm{d}\boldsymbol{x} = \int\limits_D \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle} \, \mathrm{d}\boldsymbol{x}, \qquad \boldsymbol{\xi} \in \mathbb{R}^2. \tag{6.3}$$

In order to reconstruct the polygonal domain from Fourier samples, we have to extract information about D from these samples. Therefore, we need a representation of $\widehat{f}$ where the vertices $\boldsymbol{v}_j$ occur because the vertices and the order of the vertices, that is the information which vertices are connected by line segments, are the important data which we need for the reconstruction of D.

Such a formula has been derived by Komrska in [41]. There, the Fraunhofer diffraction at polygonal apertures is examined, which is described by an integral similar to the one in (6.3). Komrska then evaluates that integral by showing that it is a linear combination of exponential terms corresponding to the vertices of the polygonal aperture. We can transfer this approach to our setting and hence obtain

Proposition 6.1 (Fourier transform of a unit-height polygon).
Let D be a polygonal domain in $\mathbb{R}^2$ with the N vertices $\boldsymbol{v}_j = (v_{j,1}, v_{j,2})^{\mathrm{T}}$, $j = 1, \ldots, N$, which are numbered anticlockwise with $\boldsymbol{v}_{N+1} := \boldsymbol{v}_1$. For $\boldsymbol{\xi} \neq \boldsymbol{0}$, the Fourier transform of the function $f : \mathbb{R}^2 \ni \boldsymbol{x} \mapsto \mathbf{1}_D(\boldsymbol{x})$ is given by

$$\widehat{f}(\boldsymbol{\xi}) = \int_D \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{x}\rangle} \, \mathrm{d}\boldsymbol{x} = \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \frac{\langle\boldsymbol{\xi},\boldsymbol{n}_j\rangle}{\langle\boldsymbol{\xi},\boldsymbol{v}_{j+1} - \boldsymbol{v}_j\rangle} \left(\mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{v}_j\rangle} - \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{v}_{j+1}\rangle} \right) \tag{6.4}$$

with

$$\boldsymbol{n}_j := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\boldsymbol{v}_{j+1} - \boldsymbol{v}_j), \quad j = 1, \ldots, N,$$

and the convention

$$\frac{\langle\boldsymbol{\xi},\boldsymbol{n}_j\rangle}{\langle\boldsymbol{\xi},\boldsymbol{v}_{j+1} - \boldsymbol{v}_j\rangle} \left(\mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{v}_j\rangle} - \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{v}_{j+1}\rangle} \right) = \mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{n}_j\rangle \, \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{v}_j\rangle}$$

if $\langle\boldsymbol{\xi},\boldsymbol{v}_{j+1} - \boldsymbol{v}_j\rangle = 0$ for some $j \in \{1, \ldots, N\}$.

The proof which we will give here is based on the approach employed in [41].

Proof of Proposition 6.1. According to (6.3), the Fourier transform of f is given by

$$\widehat{f}(\boldsymbol{\xi}) = \int_D \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{x}\rangle} \, \mathrm{d}\boldsymbol{x}, \qquad \boldsymbol{\xi} \in \mathbb{R}^2. \tag{6.5}$$

The integrand $\exp\left(-\mathrm{i}\langle\boldsymbol{\xi}, \boldsymbol{x}\rangle\right)$ satisfies the following property:

$$\begin{aligned}
\mathrm{div}_{\boldsymbol{x}}\left(\mathrm{grad}_{\boldsymbol{x}}\left(\mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{x}\rangle}\right)\right) &= \frac{\partial^2}{\partial x_1^2} \mathrm{e}^{-\mathrm{i}(\xi_1 x_1 + \xi_2 x_2)} + \frac{\partial^2}{\partial x_2^2} \mathrm{e}^{-\mathrm{i}(\xi_1 x_1 + \xi_2 x_2)} \\
&= \left[(-\mathrm{i}\,\xi_1)^2 + (-\mathrm{i}\,\xi_2)^2\right] \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{x}\rangle} \\
&= -\|\boldsymbol{\xi}\|_2^2 \, \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\xi},\boldsymbol{x}\rangle}
\end{aligned}$$

with $\boldsymbol{\xi} := (\xi_1, \xi_2)^{\mathrm{T}}$, $\boldsymbol{x} := (x_1, x_2)^{\mathrm{T}}$. Together with (6.5), this yields

$$\widehat{f}(\boldsymbol{\xi}) = -\frac{1}{\|\boldsymbol{\xi}\|_2^2} \int_D \operatorname{div}_{\boldsymbol{x}} \Big(\operatorname{grad}_{\boldsymbol{x}} \big(\mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle} \big) \Big) \, \mathrm{d}\boldsymbol{x} \quad \text{for } \boldsymbol{\xi} \neq \mathbf{0}. \tag{6.6}$$

The polygonal domain D is a compact subset of $\mathbb{R}^2$ and has a piecewise smooth boundary ∂D. Further, the vector field $\operatorname{grad}_{\boldsymbol{x}}(\mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle})$ is continuously differentiable. Let $\widetilde{n}$ be the outward pointing unit normal field of $S := \partial D$. Application of the divergence theorem[1] to the integral in (6.6) results in

$$\widehat{f}(\boldsymbol{\xi}) = -\frac{1}{\|\boldsymbol{\xi}\|_2^2} \int_{\partial D} \big\langle \operatorname{grad}_{\boldsymbol{x}}\big(\mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle} \big), \widetilde{n}(\boldsymbol{x}) \big\rangle \, \mathrm{d}S(\boldsymbol{x}), \quad \boldsymbol{\xi} \neq \mathbf{0}, \tag{6.7}$$

where we have

$$\begin{aligned}
\operatorname{grad}_{\boldsymbol{x}}\Big(\mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle} \Big) &= \begin{pmatrix} \frac{\partial}{\partial x_1} \mathrm{e}^{-\mathrm{i}(\xi_1 x_1 + \xi_2 x_2)} \\ \frac{\partial}{\partial x_2} \mathrm{e}^{-\mathrm{i}(\xi_1 x_1 + \xi_2 x_2)} \end{pmatrix} = \begin{pmatrix} -\mathrm{i}\,\xi_1 \cdot \mathrm{e}^{-\mathrm{i}(\xi_1 x_1 + \xi_2 x_2)} \\ -\mathrm{i}\,\xi_2 \cdot \mathrm{e}^{-\mathrm{i}(\xi_1 x_1 + \xi_2 x_2)} \end{pmatrix} \\
&= -\mathrm{i} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \cdot \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle} = -\mathrm{i}\,\boldsymbol{\xi}\, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{x} \rangle}.
\end{aligned} \tag{6.8}$$

Since the domain D has a polygonal form, we get

$$\partial D = \bigcup_{i=1}^{N} L_j \tag{6.9}$$

where $L_1, \ldots, L_N$ are the edges of the polygon, with the parameterizations

$$l_j(t) := \boldsymbol{v_j} + t(\boldsymbol{v_{j+1}} - \boldsymbol{v_j}), \quad t \in [0, 1], \tag{6.10}$$

for $j = 1, \ldots, N$. Remember that we number the vertices $\boldsymbol{v}_j$ anticlockwise, and that we use the convention $\boldsymbol{v}_{N+1} = \boldsymbol{v}_1$. Then the outward pointing unit normal vector $\widetilde{\boldsymbol{n}}_j$ to the edge L_j of the polygon D is given by

$$\widetilde{\boldsymbol{n}}_j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \frac{\boldsymbol{v}_{j+1} - \boldsymbol{v}_j}{\|\boldsymbol{v}_{j+1} - \boldsymbol{v}_j\|_2} = \frac{1}{\|\boldsymbol{v}_{j+1} - \boldsymbol{v}_j\|_2} \cdot \begin{pmatrix} v_{j+1,2} - v_{j,2} \\ -[v_{j+1,1} - v_{j,1}] \end{pmatrix} \tag{6.11}$$

such that $\widetilde{\boldsymbol{n}}_j$ is normalized and perpendicular to $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)$:

$$\langle \widetilde{\boldsymbol{n}}_j, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle = \frac{(v_{j+1,2} - v_{j,2}) \cdot (v_{j+1,1} - v_{j,1}) - (v_{j+1,1} - v_{j,1}) \cdot (v_{j+1,2} - v_{j,2})}{\|\boldsymbol{v}_{j+1} - \boldsymbol{v}_j\|_2} = 0.$$

We still need to verify that the direction of $\widetilde{\boldsymbol{n}}_j$ is correct. Note that the vertices of the polygon D are numbered anticlockwise, and that $\widetilde{\boldsymbol{n}}_j$, being the unit normal vector to the edge L_j which is pointing outward with respect to the polygon, needs to point to the right-hand side of the direction vector $\boldsymbol{v}_{j+1} - \boldsymbol{v}_j$. By considering the following cases, we show that the direction of $\widetilde{\boldsymbol{n}}_j$ as given in (6.11) is always correct.

---

[1] See [25, pp. 182–183, Satz 3].

- $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)_1 \geq 0$, $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)_2 \geq 0$:
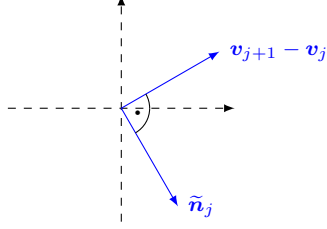
- $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)_1 < 0$, $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)_2 \geq 0$:

$$\Rightarrow \quad (\widetilde{\boldsymbol{n}}_j)_1 \geq 0,\ (\widetilde{\boldsymbol{n}}_j)_2 \leq 0. \qquad\qquad \Rightarrow \quad (\widetilde{\boldsymbol{n}}_j)_1 \geq 0,\ (\widetilde{\boldsymbol{n}}_j)_2 > 0.$$

- $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)_1 \leq 0$, $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)_2 < 0$:
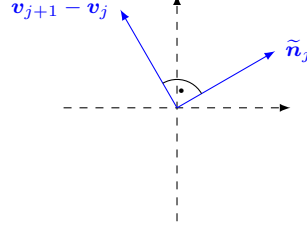
- $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)_1 > 0$, $(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j)_2 < 0$:
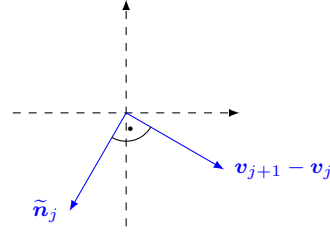
$$\Rightarrow \quad (\widetilde{\boldsymbol{n}}_j)_1 < 0,\ (\widetilde{\boldsymbol{n}}_j)_2 \geq 0. \qquad\qquad \Rightarrow \quad (\widetilde{\boldsymbol{n}}_j)_1 < 0,\ (\widetilde{\boldsymbol{n}}_j)_2 < 0.$$

Taking (6.8)–(6.11) into consideration, we evaluate the line integral[2] in (6.7) for $\boldsymbol{\xi} \neq \boldsymbol{0}$ as follows:

$$\widehat{f}(\boldsymbol{\xi}) = -\frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \int_0^1 (-\mathrm{i})\langle \boldsymbol{\xi}, \widetilde{\boldsymbol{n}}_j \rangle\, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, l_j(t) \rangle} \left\| \frac{\partial}{\partial t} l_j(t) \right\|_2 \mathrm{d}t$$

$$= \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \int_0^1 \mathrm{i}\langle \boldsymbol{\xi}, \widetilde{\boldsymbol{n}}_j \rangle\, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, l_j(t) \rangle} \|\boldsymbol{v}_{j+1} - \boldsymbol{v}_j\|_2\, \mathrm{d}t$$

$$= \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \int_0^1 \mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle\, \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j + t(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j) \rangle}\, \mathrm{d}t \qquad (6.12)$$

with

$$\boldsymbol{n}_j := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot (\boldsymbol{v}_{j+1} - \boldsymbol{v}_j). \qquad (6.13)$$

---

[2]See [32, pp. 376–377, Satz 180.6] and [32, p. 391, Satz 184.1] for the computation rules used for the evaluation of the line integral.

In the case $\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle \neq 0$ for all $j \in \{1, \ldots, N\}$, we compute

$$\widehat{f}(\boldsymbol{\xi}) = \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \int_0^1 \mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle \, \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j + t(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j) \rangle} \, \mathrm{d}t$$

$$= \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle \left[\frac{\mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j + t(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j) \rangle}}{(-\,\mathrm{i})\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} \right]_0^1$$

$$= \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} \left(\mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} - \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} \rangle} \right), \quad \boldsymbol{\xi} \neq \boldsymbol{0}.$$

If $\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle = 0$ for some $j \in \{1, \ldots, N\}$, we have to take again a look at the integral in (6.12). Then we have

$$\int_0^1 \mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle \, \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j + t(\boldsymbol{v}_{j+1} - \boldsymbol{v}_j) \rangle} \, \mathrm{d}t = \int_0^1 \mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle \, \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} \, \mathrm{d}t = \mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle \, \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} \,.$$

Altogether, we obtain

$$\widehat{f}(\boldsymbol{\xi}) = \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} \left(\mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} - \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} \rangle} \right), \quad \boldsymbol{\xi} \neq \boldsymbol{0},$$

with the convention

$$\frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} \left(\mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} - \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} \rangle} \right) = \mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle \, \mathrm{e}^{-\,\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle}$$

if $\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle = 0$ for some $j \in \{1, \ldots, N\}$. This completes the proof. $\qquad \square$

Now we have a representation of the Fourier transform of the unit-height polygon where the vertices of the polygon occur. We aim to derive a theory for the reconstruction of the polygonal domain D from sparse Fourier samples.

First, we make the assumption that no edge of the polygon D is parallel to the x_1-axis or the x_2-axis in the plane. Then it holds that

$$\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle \neq 0 \quad \text{for all } j \in \{1, \ldots, N\}$$

for vectors $\boldsymbol{\xi}$ of the form

$$\boldsymbol{\xi} = (\xi_1, 0)^{\mathrm{T}} \quad \text{or} \quad \boldsymbol{\xi} = (0, \xi_2)^{\mathrm{T}} \quad \text{with} \quad \xi_1, \xi_2 \neq 0,$$

and we can use the expression in (6.4), which we can rephrase by taking the assumption into consideration. Then we obtain

$$
\widehat{f}(\boldsymbol{\xi}) = \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} \left(\mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} - \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} \rangle} \right)
$$

$$
= \frac{1}{\|\boldsymbol{\xi}\|_2^2} \left[\sum_{j=1}^{N} \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} - \sum_{j=2}^{N+1} \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_{j-1} \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_j - \boldsymbol{v}_{j-1} \rangle} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} \right]
$$

$$
= \frac{1}{\|\boldsymbol{\xi}\|_2^2} \left[\sum_{j=1}^{N} \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} - \sum_{j=2}^{N} \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_{j-1} \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_j - \boldsymbol{v}_{j-1} \rangle} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} \right.
$$

$$
\left. - \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_N \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{N+1} - \boldsymbol{v}_N \rangle} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_{N+1} \rangle} \right].
$$

Remember that $\boldsymbol{v}_{N+1} = \boldsymbol{v}_1$. The additional conventions

$$
\boldsymbol{v}_0 := \boldsymbol{v}_N \quad \text{and} \quad \boldsymbol{n}_0 := \boldsymbol{n}_N
$$

yield

$$
\widehat{f}(\boldsymbol{\xi}) = \frac{1}{\|\boldsymbol{\xi}\|_2^2} \sum_{j=1}^{N} \left(\frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} - \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_{j-1} \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_j - \boldsymbol{v}_{j-1} \rangle} \right) \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} \tag{6.14}
$$

for vectors $\boldsymbol{\xi}$ of the form $\boldsymbol{\xi} = (\xi_1, 0)^{\mathrm{T}}$ or $\boldsymbol{\xi} = (0, \xi_2)^{\mathrm{T}}$ with $\xi_1, \xi_2 \neq 0$ (under the assumption that no edge of the polygon is parallel to the x_1- or x_2-axis).

By considering the function

$$
g(\boldsymbol{\xi}) := \|\boldsymbol{\xi}\|_2^2 \widehat{f}(\boldsymbol{\xi}), \tag{6.15}
$$

we get a bivariate exponential sum with the vertices $\boldsymbol{v}_j$ of the polygon D as exponents. Making the further assumptions that the vertex coordinates $v_{j,1}$, $j = 1, \ldots, N$, as well as the coordinates $v_{j,2}$, $j = 1, \ldots, N$, are pairwise different, and using a similar idea as in the bivariate case of non-uniform translates, cf. Subsection 4.2.1, we are able to compute the vertices by applying the Prony method to sampling values from three straight lines through the origin where the third line is determined adaptively.

The assumption that the vertex components $v_{j,1}$ and $v_{j,2}$ are distinct in the x_1- and x_2-direction of the Cartesian coordinate system respectively also ensures that no edge of the polygon is parallel to the x_1- or x_2-axis.

However, it does not suffice to only know the vertices, see Figure 6.2, since the polygon can be concave. Thus, we have to establish the true order of the vertices. This means that we need to determine which vertices have to be connected by line segments in order to reconstruct the original polygonal domain.
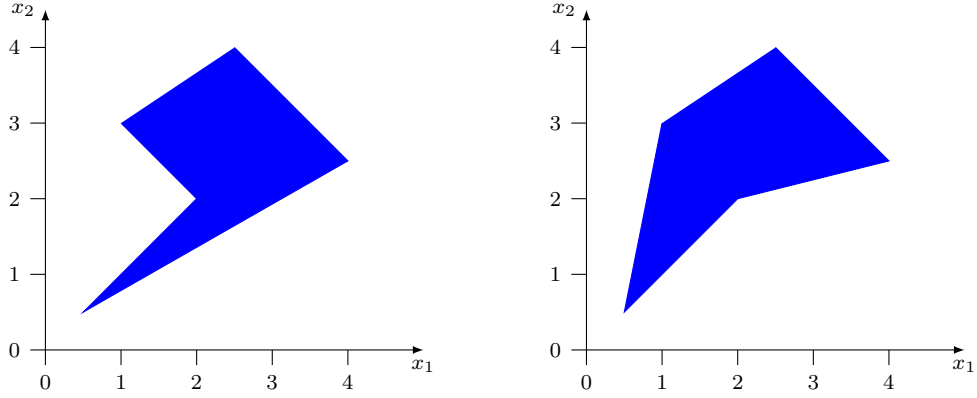
Figure 6.2. *Example for two different unit-height polygons with the same 5 vertices.*

The ordering of the computed vertices can be done with the help of the coefficients which are obtained by applying the Prony method to three univariate problems, i.e. to the sampling values from three straight lines in the plane.

We will collect the above observations in the following theorem, where the proof is constructive. Note that the considered polygons are always assumed to be non-degenerated such that three neighbouring vertices do not lie on the same line.

Theorem 6.2 (Reconstruction of polygonal domains).
Let f be a unit-height polygon as defined in (6.2), i.e. $f : \mathbb{R}^2 \to \mathbb{R}$ with

$$f(\boldsymbol{x}) := \mathbf{1}_D(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^2,$$

where D is a polygon in $\mathbb{R}^2$ with the N vertices $\boldsymbol{v}_j = (v_{j,1}, v_{j,2})^{\mathrm{T}}$ for $j = 1, \ldots, N$, which are numbered anticlockwise. Additionally, the vertex components $v_{j,1}$ and $v_{j,2}$ are distinct in the x_1- and x_2-direction of the Cartesian coordinate system respectively. Further, assume that the constant $h > 0$ satisfies the condition $h\|\boldsymbol{v}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N\}$. Then we get the following reconstruction result:

The polygon D can be uniquely recovered from the $3N$ Fourier samples $\widehat{f}(\boldsymbol{\xi})$ for

$$\boldsymbol{\xi}^{\mathrm{T}} \in \left\{ \big(\ell h, 0\big), \big(0, \ell h\big), \big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big) \mid \ell = 1, \ldots, N \right\}$$

where $\vartheta \in (0,1) \setminus \{\frac{1}{2}\}$ needs to be chosen suitably.

Proof. The proof consists of five major parts. In the first two parts, we use the Prony method (applied to sampling values of $\widehat{f}$ from two straight lines in the fre-

quency domain) in order to compute the coordinate values $v_{j,1}$ and $v_{j,2}$ of the vertices $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$.

But we are faced with the problem of combining these coordinate values to points in the plane which are the original vertices. Thus, in the next part, we determine a *candiate set* of points that are possible vertices. Dependently on this candidate set, we decide on a third sampling line such that we can use the results of the Prony method applied to sampling values from this third line in order to uniquely determine the original vertices in the fourth part of the proof.

In the fifth and last part, we establish the right order of the computed vertices, for which we use the coefficients that we have obtained by the application of the Prony method.

Part 1:
Due to the assumption that the vertex coordinates $v_{j,1}$, $j = 1, \ldots, N$, as well as the coordinates $v_{j,2}$, $j = 1, \ldots, N$, are pairwise different, it is not possible that any edge of the polygon D is parallel to the x_1- or x_2-axis of the Cartesian coordinate system. Therefore, we can consider the function g defined in (6.15). The representation of the Fourier transform of f in (6.14) then yields

$$g(\boldsymbol{\xi}) = \sum_{j=1}^{N} \left(\frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_j \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle} - \frac{\langle \boldsymbol{\xi}, \boldsymbol{n}_{j-1} \rangle}{\langle \boldsymbol{\xi}, \boldsymbol{v}_j - \boldsymbol{v}_{j-1} \rangle} \right) \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\xi}, \boldsymbol{v}_j \rangle} \tag{6.16}$$

with the conventions

$$\boldsymbol{v}_{N+1} = \boldsymbol{v}_1, \quad \boldsymbol{v}_0 = \boldsymbol{v}_N, \quad \boldsymbol{n}_0 = \boldsymbol{n}_N. \tag{6.17}$$

Taking vectors $\boldsymbol{\xi}$ of the form $\boldsymbol{\xi} = (\xi_1, 0)^{\mathrm{T}}$, $\xi_1 \neq 0$, we obtain

$$g(\xi_1, 0) = \sum_{j=1}^{N} \left(\frac{\xi_1 n_{j,1}}{\xi_1 (v_{j+1,1} - v_{j,1})} - \frac{\xi_1 n_{j-1,1}}{\xi_1 (v_{j,1} - v_{j-1,1})} \right) \mathrm{e}^{-\mathrm{i}\xi_1 v_{j,1}},$$

where

$$n_{j,1} = v_{j+1,2} - v_{j,2}, \quad j = 0, \ldots, N,$$

see (6.13). Thus, we have the univariate exponential sum

$$g(\xi_1, 0) = \sum_{j=1}^{N} a_j \, \mathrm{e}^{-\mathrm{i}\xi_1 v_{j,1}} \tag{6.18}$$

with the coefficients

$$a_j := \frac{v_{j+1,2} - v_{j,2}}{v_{j+1,1} - v_{j,1}} - \frac{v_{j,2} - v_{j-1,2}}{v_{j,1} - v_{j-1,1}}, \quad j = 1, \ldots, N. \tag{6.19}$$

Observe that the coefficient a_j for $j \in \{1, \ldots, N\}$ is the difference of the slopes of the two line segments connecting the vertex $\boldsymbol{v}_j$ with $\boldsymbol{v}_{j-1}$ and $\boldsymbol{v}_{j+1}$. Since we have made the assumption that the polygon D is non-degenerated, a_j is non-zero. Further, the value a_j is well defined because no edge of the polygon is parallel to the x_2-axis. Hence, the values of the slopes of the edges are finite.

In order to obtain the vertex coordinates $v_{j,1}$ and the corresponding coefficients a_j, $j = 1, \ldots, N$, we apply the Prony method to the univariate function $g(\cdot, 0)$ in (6.18). Note that we have

$$g(0,0) = \sum_{j=1}^{N} a_j = \sum_{j=1}^{N} \left(\frac{v_{j+1,2} - v_{j,2}}{v_{j+1,1} - v_{j,1}} - \frac{v_{j,2} - v_{j-1,2}}{v_{j,1} - v_{j-1,1}} \right)$$

$$= \frac{v_{N+1,2} - v_{N,2}}{v_{N+1,1} - v_{N,1}} - \frac{v_{1,2} - v_{0,2}}{v_{1,1} - v_{0,1}} = 0 \tag{6.20}$$

by (6.17). Further, the coefficients a_j are real-valued, and we have $h|v_{j,1}| \leq h\|\boldsymbol{v}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N\}$ by assumption. Hence, we can use the Prony method as described in Section 2.2. For this purpose, we need the function values

$$g(\ell h, 0), \quad \ell = 0, \ldots, N,$$

which are given by (6.20), the definition of g in (6.15), i.e.

$$g(\ell h, 0) = |\ell h|^2 \widehat{f}(\ell h, 0), \quad \ell = 1, \ldots, N,$$

and the N Fourier samples

$$\widehat{f}(\ell h, 0), \quad \ell = 1, \ldots, N.$$

The application of the Prony method then yields frequency values

$$\alpha_1 < \alpha_2 < \ldots < \alpha_N$$

and corresponding coefficients $a_{\alpha,1}, \ldots, a_{\alpha,N}$ as a result where the set $\{\alpha_1, \ldots, \alpha_N\}$ contains all vertex coordinates $v_{1,1}, \ldots, v_{N,1}$, and the set $\{a_{\alpha,1}, \ldots, a_{\alpha,N}\}$ consists of all coefficients $a_1, \ldots, a_N$. But we do not know the order of these coordinate values and coefficients; that is, the permutation which maps $(\alpha_1, \ldots, \alpha_N)$ onto $(v_{1,1}, \ldots, v_{N,1})$ and $(a_{\alpha,1}, \ldots, a_{\alpha,N})$ onto $(a_1, \ldots, a_N)$ is unknown. This has to be determined later.

Part 2:
In an analogous way, we compute a set of values containing all vertex coordinates $v_{j,2}$ and corresponding coefficients.

By considering vectors $\boldsymbol{\xi}$ of the form $\boldsymbol{\xi} = (0, \xi_2)^{\mathrm{T}}$, $\xi_2 \neq 0$, the representation (6.16) results in

$$g(0, \xi_2) = \sum_{j=1}^{N} \left(\frac{\xi_2 n_{j,2}}{\xi_2(v_{j+1,2} - v_{j,2})} - \frac{\xi_2 n_{j-1,2}}{\xi_2(v_{j,2} - v_{j-1,2})} \right) \mathrm{e}^{-\mathrm{i}\xi_2 v_{j,2}}$$

where the components $n_{j,2}$ of the normal vectors are given by (6.13), i.e.

$$n_{j,2} = v_{j,1} - v_{j+1,1}, \quad j = 0, \ldots, N.$$

Hence, the univariate function $g(0, \cdot)$ is given by

$$g(0, \xi_2) = \sum_{j=1}^{N} b_j \, e^{-i \xi_2 v_{j,2}} \tag{6.21}$$

where the coefficients b_j are defined as follows:

$$b_j := \frac{v_{j,1} - v_{j+1,1}}{v_{j+1,2} - v_{j,2}} - \frac{v_{j-1,1} - v_{j,1}}{v_{j,2} - v_{j-1,2}}, \quad j = 1, \ldots, N. \tag{6.22}$$

The value b_j is the difference of slopes of two neighbouring line segments where the parameterizations of the line segments are considered as functions of x_2. Similarly to the case of the coefficients a_j, see *Part 1* of this proof, we conclude that the coefficients b_j, $j = 1, \ldots, N$, are well defined and non-zero.

Observe that $h|v_{j,2}| \leq h\|\boldsymbol{v}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N\}$ by assumption, and that all coefficients b_j are real-valued such that we can use the Prony method from Section 2.2 in order to determine the values $v_{j,2}$ and the coefficients b_j for all $j \in \{1, \ldots, N\}$ in (6.21). The required function values

$$g(0, \ell h), \quad \ell = 0, \ldots, N,$$

are computed by using the N Fourier samples

$$\widehat{f}(0, \ell h), \quad \ell = 1, \ldots, N,$$

and

$$g(0, \ell h) = |\ell h|^2 \widehat{f}(0, \ell h), \quad \ell = 1, \ldots, N,$$

which is due to the definition of g in (6.15). Moreover, the function value $g(0, 0)$ is equal to zero since (6.21) and (6.17) yield

$$g(0, 0) = \sum_{j=1}^{N} b_j = \sum_{j=1}^{N} \left(\frac{v_{j,1} - v_{j+1,1}}{v_{j+1,2} - v_{j,2}} - \frac{v_{j-1,1} - v_{j,1}}{v_{j,2} - v_{j-1,2}} \right) = \frac{v_{N,1} - v_{N+1,1}}{v_{N+1,2} - v_{N,2}} - \frac{v_{0,1} - v_{1,1}}{v_{1,2} - v_{0,2}} = 0.$$

Using the approach described above, we obtain values

$$\beta_1 < \beta_2 < \ldots < \beta_N$$

with corresponding coefficients $b_{\beta,1}, \ldots, b_{\beta,N}$ where we have

$$\{\beta_1, \ldots, \beta_N\} = \{v_{1,2}, \ldots, v_{N,2}\} \quad \text{and} \quad \{b_{\beta,1}, \ldots, b_{\beta,N}\} = \{b_1, \ldots, b_N\}.$$

Comparably to the case of the vertex coordinates $v_{j,1}$ and the coefficients a_j, we do not yet know the permutation which maps the sequences $(\beta_1, \ldots, \beta_N)$ and $(b_{\beta,1}, \ldots, b_{\beta,N})$ onto $(v_{1,2}, \ldots, v_{N,2})$ and $(b_1, \ldots, b_N)$ respectively.

Part 3:
Now we have to compute the original vertices $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$. In the previous steps, we have obtained the values of the vertex coordinates $v_{j,1}$ and $v_{j,2}$ for $j = 1, \ldots, N$. But we do not know their order such that we do not know how to combine these coordinates to points in the plane which are the original vertices. Thus, we can only consider the Cartesian product of the sets $\{\alpha_1, \ldots, \alpha_N\}$ and $\{\beta_1, \ldots, \beta_N\}$ as a set of *candidate points* for the original vertices. Let us denote this set by K:

$$K := \left\{ (\alpha_k, \beta_\ell)^{\mathrm{T}} : \quad k = 1, \ldots, N, \ \ell = 1, \ldots, N \right\}.$$

In order to determine the N original vertices in the set K of N^2 candidate points, we use a third univariate problem to which we apply the Prony method. This means that we need further Fourier data, namely Fourier samples from a third sampling line.

For this purpose, we choose a parameter $\vartheta \in (0,1) \setminus \{\frac{1}{2}\}$ such that the orthogonal projections of all candidate points in K onto the line $x_2 = \tan(\vartheta\pi)x_1$ are pairwise different, see Figure 6.3. This parameter ϑ is then used to determine the third sampling line in the frequency domain. The angle between the positive ω_1-axis and the part of the line $\omega_2 = \tan(\vartheta\pi)\omega_1$ which lies in the upper half space is given by the value $\vartheta\pi$ rad. We use equispaced sampling locations on this third sampling line $\omega_2 = \tan(\vartheta\pi)\omega_1$ with step size h, see Figure 6.3, and take the N Fourier samples

$$\widehat{f}\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big), \quad \ell = 1, \ldots, N.$$

Since the orthogonal projections of all points in K onto the line $x_2 = \tan(\vartheta\pi)x_1$ are distinct, no possible edge of the polygon D, i.e. possible connection between two points in K, is perpendicular to the line $x_2 = \tan(\vartheta\pi)x_1$. Thus, we have

$$\langle \boldsymbol{\xi}, \boldsymbol{v}_{j+1} - \boldsymbol{v}_j \rangle \neq 0 \quad \text{for all } j \in \{1, \ldots, N\} \tag{6.23}$$

for vectors $\boldsymbol{\xi} = \big(\cos(\vartheta\pi)\xi_1, \sin(\vartheta\pi)\xi_1\big)^{\mathrm{T}}$ with $\xi_1 \neq 0$, and we can use the representation of $\widehat{f}$ in (6.14) and the function g in (6.15). Hence, we obtain the following for $\xi_1 \neq 0$:

$$g\big(\cos(\vartheta\pi)\xi_1, \sin(\vartheta\pi)\xi_1\big)$$
$$= \sum_{j=1}^{N} \left[\frac{\xi_1\big(\cos(\vartheta\pi)n_{j,1} + \sin(\vartheta\pi)n_{j,2}\big)}{\xi_1\big(\cos(\vartheta\pi)(v_{j+1,1} - v_{j,1}) + \sin(\vartheta\pi)(v_{j+1,2} - v_{j,2})\big)} \right.$$
$$\left. - \frac{\xi_1\big(\cos(\vartheta\pi)n_{j-1,1} + \sin(\vartheta\pi)n_{j-1,2}\big)}{\xi_1\big(\cos(\vartheta\pi)(v_{j,1} - v_{j-1,1}) + \sin(\vartheta\pi)(v_{j,2} - v_{j-1,2})\big)} \right] \mathrm{e}^{-\mathrm{i}\xi_1[\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}]}$$

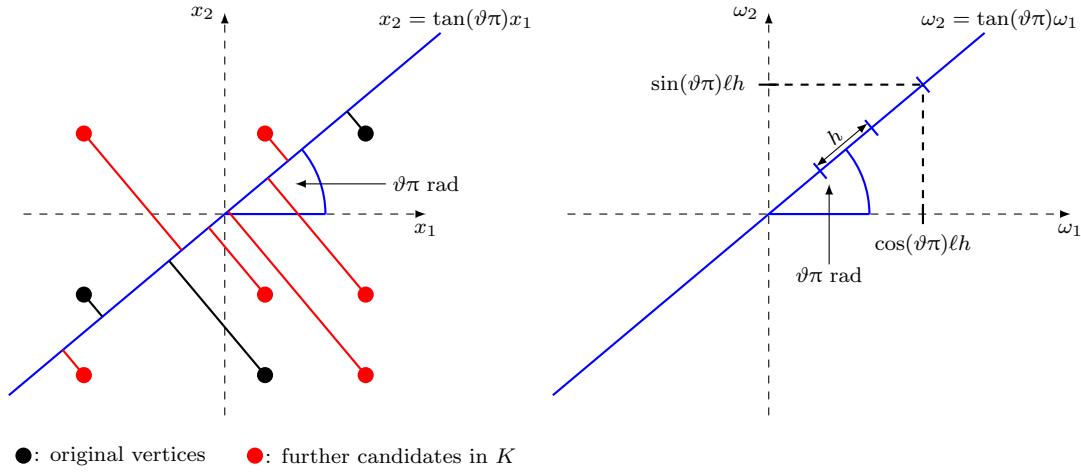● : original vertices   ● : further candidates in K

Figure 6.3. Left: *Determination of the parameter ϑ in the time domain.* Right: *Third sampling line in the frequency domain with sampling locations (displayed example: $\ell = 3$).*

where the components $n_{j,1}$ and $n_{j,2}$ of the normal vectors are given by (6.13), i.e.

$$n_{j,1} = v_{j+1,2} - v_{j,2}, \qquad j = 0, \ldots, N.$$
$$n_{j,2} = v_{j,1} - v_{j+1,1},$$

This leads to the univariate function $g\big(\cos(\vartheta\pi)\cdot, \sin(\vartheta\pi)\cdot\big)$ defined by

$$g\big(\cos(\vartheta\pi)\xi_1, \sin(\vartheta\pi)\xi_1\big) = \sum_{j=1}^{N} c_j\, e^{-\mathrm{i}\xi_1\big(\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\big)} \qquad (6.24)$$

with the coefficients

$$c_j := \frac{\cos(\vartheta\pi)(v_{j+1,2} - v_{j,2}) + \sin(\vartheta\pi)(v_{j,1} - v_{j+1,1})}{\cos(\vartheta\pi)(v_{j+1,1} - v_{j,1}) + \sin(\vartheta\pi)(v_{j+1,2} - v_{j,2})}$$
$$- \frac{\cos(\vartheta\pi)(v_{j,2} - v_{j-1,2}) + \sin(\vartheta\pi)(v_{j-1,1} - v_{j,1})}{\cos(\vartheta\pi)(v_{j,1} - v_{j-1,1}) + \sin(\vartheta\pi)(v_{j,2} - v_{j-1,2})}, \qquad j = 1, \ldots, N. \qquad (6.25)$$

Note that these coefficients are not only well defined due to (6.23), but they are also non-zero. This can be seen similarly as in the cases of the coefficients a_j and b_j in *Parts 1* and *2* of this proof since the coefficient c_j for $j \in \{1, \ldots, N\}$ is also a difference of slopes of two neighbouring line segments. Here, the parameterizations of the line segments are considered as functions with respect to a rotated coordinate system which results from the rotation of the original Cartesian coordinate system by the angle $\vartheta\pi$ rad.

In order to determine the values $\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}$ and the coefficients c_j for all $j \in \{1, \ldots, N\}$ in (6.24), we apply the Prony method. First, note that all coefficients c_j are real-valued. Secondly, the assumption that $h\|\boldsymbol{v}_j\|_2 < \pi$ for all $j \in \{1, \ldots, N\}$ is satisfied. The rotated vectors

$$\begin{pmatrix} \cos(\vartheta\pi) & \sin(\vartheta\pi) \\ -\sin(\vartheta\pi) & \cos(\vartheta\pi) \end{pmatrix} \boldsymbol{v}_j, \quad j = 1, \ldots, N,$$

fulfil the same norm condition. Thus, we have

$$h\big|\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\big|$$
$$= h\sqrt{\big(\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\big)^2}$$
$$\leq h\sqrt{\big(\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\big)^2 + \big(-\sin(\vartheta\pi)v_{j,1} + \cos(\vartheta\pi)v_{j,2}\big)^2}$$
$$= \left\|\begin{pmatrix} \cos(\vartheta\pi) & \sin(\vartheta\pi) \\ -\sin(\vartheta\pi) & \cos(\vartheta\pi) \end{pmatrix} \boldsymbol{v}_j\right\|_2 < \pi$$

for $j = 1, \ldots, N$. Hence, we can apply the Prony method as described in Section 2.2, where we use the function values

$$g\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big), \quad \ell = 0, \ldots, N.$$

In the case of $\ell = 1, \ldots, N$, these values are given by the N Fourier samples

$$\widehat{f}\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big), \quad \ell = 1, \ldots, N,$$

and the definition of g in (6.15), i.e.

$$g\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big) = |\ell h|^2 \widehat{f}\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big), \quad \ell = 1, \ldots, N.$$

Further, we have $g(0,0) = 0$ since (6.24) and (6.17) produce

$$g(0,0) = \sum_{j=1}^{N} c_j = \sum_{j=1}^{N} \left(\frac{\cos(\vartheta\pi)(v_{j+1,2} - v_{j,2}) + \sin(\vartheta\pi)(v_{j,1} - v_{j+1,1})}{\cos(\vartheta\pi)(v_{j+1,1} - v_{j,1}) + \sin(\vartheta\pi)(v_{j+1,2} - v_{j,2})} \right.$$
$$\left. - \frac{\cos(\vartheta\pi)(v_{j,2} - v_{j-1,2}) + \sin(\vartheta\pi)(v_{j-1,1} - v_{j,1})}{\cos(\vartheta\pi)(v_{j,1} - v_{j-1,1}) + \sin(\vartheta\pi)(v_{j,2} - v_{j-1,2})} \right)$$
$$= \frac{\cos(\vartheta\pi)(v_{N+1,2} - v_{N,2}) + \sin(\vartheta\pi)(v_{N,1} - v_{N+1,1})}{\cos(\vartheta\pi)(v_{N+1,1} - v_{N,1}) + \sin(\vartheta\pi)(v_{N+1,2} - v_{N,2})}$$
$$- \frac{\cos(\vartheta\pi)(v_{1,2} - v_{0,2}) + \sin(\vartheta\pi)(v_{0,1} - v_{1,1})}{\cos(\vartheta\pi)(v_{1,1} - v_{0,1}) + \sin(\vartheta\pi)(v_{1,2} - v_{0,2})} \bigg) = 0.$$

Application of the Prony method then results in frequency values

$$\gamma_1 < \gamma_2 < \ldots < \gamma_N$$

with corresponding coefficients $c_{\gamma,1}, \ldots, c_{\gamma,N}$. Similarly to the previous two univariate problems, we have the set relations

$$\{\gamma_1, \ldots, \gamma_N\} = \left\{ \left(\cos(\vartheta\pi)v_{1,1} + \sin(\vartheta\pi)v_{1,2}\right), \ldots, \left(\cos(\vartheta\pi)v_{N,1} + \sin(\vartheta\pi)v_{N,2}\right) \right\}$$

and

$$\{c_{\gamma,1}, \ldots, c_{\gamma,N}\} = \{c_1, \ldots, c_N\},$$

cf. (6.24), where $\left(\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\right)$ for $j = 1, \ldots, N$ are the coordinate values of the original vertices with respect to the line $x_2 = \tan(\vartheta\pi)x_1$.

Part 4:
The previous results now enable us to compute the original vertices $\boldsymbol{v}_j$, $j = 1, \ldots, N$, by comparison of the set K of candidate points with the set $\{\gamma_1, \ldots, \gamma_N\}$. For this purpose, we determine all points $(\alpha_k, \beta_\ell)^{\mathrm{T}}$ in the set K for which there exist indices $j \in \{1, \ldots, N\}$ such that $\cos(\vartheta\pi)\alpha_k + \sin(\vartheta\pi)\beta_\ell = \gamma_j$. Then the set

$$\widetilde{G} := \left\{ (\alpha_k, \beta_\ell)^{\mathrm{T}} \in K \mid \exists\, j \in \{1, \ldots, N\} : \ \cos(\vartheta\pi)\alpha_k + \sin(\vartheta\pi)\beta_\ell = \gamma_j \right\}$$

contains all N original vertices $\boldsymbol{v}_j$, $j = 1, \ldots, N$, of the polygon D and it holds that

$$\left|\widetilde{G}\right| = N;$$

that is, the set $\widetilde{G}$ contains only the vertices $\boldsymbol{v}_j$ and no other elements. We can sort the elements of $\widetilde{G}$ in an arbitrary order such that we have

$$\widetilde{G} = \left\{ \widetilde{\boldsymbol{v}}_1, \ldots, \widetilde{\boldsymbol{v}}_N \right\} \quad \text{with} \quad \widetilde{\boldsymbol{v}}_j = \left(\widetilde{\alpha}_j, \widetilde{\beta}_j \right)^{\mathrm{T}} \text{ for } j = 1, \ldots, N.$$

Further, each element $\widetilde{\boldsymbol{v}}_j$ of $\widetilde{G}$ has three corresponding coefficient values

$$\widetilde{a}_j, \quad \widetilde{b}_j, \quad \text{and} \quad \widetilde{c}_j,$$

which are the coefficients obtained by applying the Prony method to the three problems (6.18), (6.21), and (6.24). These coefficients are corresponding to the orthogonal projections of the vertex $\widetilde{\boldsymbol{v}}_j$ onto the x_1-axis, x_2-axis, and the line $x_2 = \tan(\vartheta\pi)x_1$ respectively. With the help of these coefficient values, we can determine the correct order of the vertices of D, or equivalently, which elements of $\widetilde{G}$ have to be connected by line segments such that we can uniquely reconstruct the polygonal domain D.

Part 5:

In this part, we establish the right order of the vertices as mentioned above. Using the coefficient values $\widetilde{a}_j$, $\widetilde{b}_j$, and $\widetilde{c}_j$, we can compute the slopes of the polygon's edges such that we are able to determine a predecessor and a successor for each vertex in the set $\widetilde{G}$.

Part 5a: First, let us take a look at the original edges $L_1, \ldots, L_N$ of the polygon D in order to understand how we can compute the slopes of the polygon's edges. Define

$$m_j := \frac{v_{j+1,2} - v_{j,2}}{v_{j+1,1} - v_{j,1}}, \quad j = 1, \ldots, N.$$

This value describes the slope of the line which contains the line segment L_j. Observe that m_j is well defined since the vertex coordinates $v_{j,1}$, $j = 1, \ldots, N$ are distinct. Moreover, we have

$$m_j \neq 0 \quad \text{for all } j \in \{1, \ldots, N\}$$

because we have made the assumption that no edge of the polygon is parallel to the x_1-axis. Using this definition of m_j, the coefficients a_j, b_j, and c_j, see (6.19), (6.22), and (6.25) respectively, which correspond to the vertex v_j for $j = 1, \ldots, N$, can be written as follows:

$$a_j = \frac{v_{j+1,2} - v_{j,2}}{v_{j+1,1} - v_{j,1}} - \frac{v_{j,2} - v_{j-1,2}}{v_{j,1} - v_{j-1,1}} = m_j - m_{j-1}, \tag{6.26}$$

$$\begin{aligned} b_j &= \frac{v_{j,1} - v_{j+1,1}}{v_{j+1,2} - v_{j,2}} - \frac{v_{j-1,1} - v_{j,1}}{v_{j,2} - v_{j-1,2}} \\ &= -\left(\frac{v_{j+1,1} - v_{j,1}}{v_{j+1,2} - v_{j,2}} \right) + \frac{v_{j,1} - v_{j-1,1}}{v_{j,2} - v_{j-1,2}} = -\frac{1}{m_j} + \frac{1}{m_{j-1}}, \end{aligned} \tag{6.27}$$

$$\begin{aligned} c_j &= \frac{\cos(\vartheta\pi)(v_{j+1,2} - v_{j,2}) + \sin(\vartheta\pi)(v_{j,1} - v_{j+1,1})}{\cos(\vartheta\pi)(v_{j+1,1} - v_{j,1}) + \sin(\vartheta\pi)(v_{j+1,2} - v_{j,2})} \\ &\quad - \frac{\cos(\vartheta\pi)(v_{j,2} - v_{j-1,2}) + \sin(\vartheta\pi)(v_{j-1,1} - v_{j,1})}{\cos(\vartheta\pi)(v_{j,1} - v_{j-1,1}) + \sin(\vartheta\pi)(v_{j,2} - v_{j-1,2})}. \end{aligned}$$

Using the definitions

$$c_j^{(\text{N1})} := \cos(\vartheta\pi)(v_{j+1,2} - v_{j,2}) + \sin(\vartheta\pi)(v_{j,1} - v_{j+1,1}),$$

$$c_j^{(\text{D1})} := \cos(\vartheta\pi)(v_{j+1,1} - v_{j,1}) + \sin(\vartheta\pi)(v_{j+1,2} - v_{j,2}),$$

$$c_j^{(\text{N2})} := \cos(\vartheta\pi)(v_{j,2} - v_{j-1,2}) + \sin(\vartheta\pi)(v_{j-1,1} - v_{j,1}),$$

$$c_j^{(\text{D2})} := \cos(\vartheta\pi)(v_{j,1} - v_{j-1,1}) + \sin(\vartheta\pi)(v_{j,2} - v_{j-1,2}),$$

we obtain

$$c_j = \frac{c_j^{(N1)} \cdot c_j^{(D2)} - c_j^{(N2)} \cdot c_j^{(D1)}}{c_j^{(D1)} \cdot c_j^{(D2)}}. \tag{6.28}$$

Simplification of the numerator in (6.28) leads to

$$\begin{aligned}
c_j^{(N1)} &\cdot c_j^{(D2)} - c_j^{(N2)} \cdot c_j^{(D1)} \\
&= \cos^2(\vartheta\pi) \cdot (v_{j+1,2} - v_{j,2})(v_{j,1} - v_{j-1,1}) + \sin^2(\vartheta\pi) \cdot (v_{j,1} - v_{j+1,1})(v_{j,2} - v_{j-1,2}) \\
&\quad + \sin(\vartheta\pi)\cos(\vartheta\pi)\big[(v_{j+1,2} - v_{j,2})(v_{j,2} - v_{j-1,2}) + (v_{j,1} - v_{j+1,1})(v_{j,1} - v_{j-1,1})\big] \\
&\quad - \cos^2(\vartheta\pi) \cdot (v_{j+1,1} - v_{j,1})(v_{j,2} - v_{j-1,2}) - \sin^2(\vartheta\pi) \cdot (v_{j-1,1} - v_{j,1})(v_{j+1,2} - v_{j,2}) \\
&\quad - \sin(\vartheta\pi)\cos(\vartheta\pi)\big[(v_{j-1,1} - v_{j,1})(v_{j+1,1} - v_{j,1}) + (v_{j,2} - v_{j-1,2})(v_{j+1,2} - v_{j,2})\big] \\
&= \cos^2(\vartheta\pi) \cdot \big[(v_{j+1,2} - v_{j,2})(v_{j,1} - v_{j-1,1}) - (v_{j+1,1} - v_{j,1})(v_{j,2} - v_{j-1,2})\big] \\
&\quad + \sin^2(\vartheta\pi) \cdot \big[(v_{j,1} - v_{j-1,1})(v_{j+1,2} - v_{j,2}) - (v_{j+1,1} - v_{j,1})(v_{j,2} - v_{j-1,2})\big] \\
&= (v_{j+1,2} - v_{j,2})(v_{j,1} - v_{j-1,1}) - (v_{j+1,1} - v_{j,1})(v_{j,2} - v_{j-1,2}),
\end{aligned}$$

where we use the well-known Pythagorean trigonometric identity[3].

The denominator in (6.28) can be expressed as

$$\begin{aligned}
c_j^{(D1)} \cdot c_j^{(D2)} &= \big[\cos(\vartheta\pi)(v_{j+1,1} - v_{j,1}) + \sin(\vartheta\pi)(v_{j+1,2} - v_{j,2})\big] \\
&\qquad \cdot \big[\cos(\vartheta\pi)(v_{j,1} - v_{j-1,1}) + \sin(\vartheta\pi)(v_{j,2} - v_{j-1,2})\big] \\
&= (v_{j+1,1} - v_{j,1})\left[\cos(\vartheta\pi) + \sin(\vartheta\pi)\frac{(v_{j+1,2} - v_{j,2})}{(v_{j+1,1} - v_{j,1})}\right] \\
&\qquad \cdot (v_{j,1} - v_{j-1,1})\left[\cos(\vartheta\pi) + \sin(\vartheta\pi)\frac{(v_{j,2} - v_{j-1,2})}{(v_{j,1} - v_{j-1,1})}\right] \\
&= (v_{j+1,1} - v_{j,1})\left[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_j\right] \\
&\qquad \cdot (v_{j,1} - v_{j-1,1})\left[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_{j-1}\right].
\end{aligned}$$

We insert the expressions which we have just computed into (6.28). This results in

$$\begin{aligned}
c_j &= \frac{(v_{j+1,2} - v_{j,2})(v_{j,1} - v_{j-1,1}) - (v_{j+1,1} - v_{j,1})(v_{j,2} - v_{j-1,2})}{(v_{j+1,1} - v_{j,1})\left[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_j\right] \cdot (v_{j,1} - v_{j-1,1})\left[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_{j-1}\right]} \\
&= \frac{\dfrac{(v_{j+1,2} - v_{j,2})}{(v_{j+1,1} - v_{j,1})} - \dfrac{(v_{j,2} - v_{j-1,2})}{(v_{j,1} - v_{j-1,1})}}{\left[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_j\right] \cdot \left[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_{j-1}\right]} \\
&= \frac{m_j - m_{j-1}}{\left[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_j\right] \cdot \left[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_{j-1}\right]}. \tag{6.29}
\end{aligned}$$

---

[3]See [31, p. 276, (48.20)], for example.

The representations (6.26), (6.27), and (6.29) can now be used to compute the slope m_j if the coefficients a_j, b_j, and c_j are known. For this purpose, we obtain the following system of equations:

$$a_j = m_j - m_{j-1} \tag{6.30}$$

$$b_j = -\frac{1}{m_j} + \frac{1}{m_{j-1}} \tag{6.31}$$

$$c_j = \frac{m_j - m_{j-1}}{[\cos(\vartheta\pi) + \sin(\vartheta\pi)m_j] \cdot [\cos(\vartheta\pi) + \sin(\vartheta\pi)m_{j-1}]}. \tag{6.32}$$

Multiplication of (6.31) with $m_j m_{j-1}$, and multiplication of (6.32) with the denominator on the right-hand side and division by c_j results in

$$m_j - m_{j-1} = a_j \tag{6.33}$$

$$m_j m_{j-1} b_j = m_j - m_{j-1} \tag{6.34}$$

$$\frac{m_j - m_{j-1}}{c_j} = \cos^2(\vartheta\pi) + (m_j + m_{j-1})\sin(\vartheta\pi)\cos(\vartheta\pi) + m_j m_{j-1}\sin^2(\vartheta\pi). \tag{6.35}$$

Using (6.33) in (6.34) as well as (6.33) and (6.34) in the third equation (6.35), we get

$$m_j - m_{j-1} = a_j \tag{6.36}$$

$$m_j m_{j-1} = \frac{a_j}{b_j} \tag{6.37}$$

$$\frac{a_j}{c_j} = \cos^2(\vartheta\pi) + (m_j + m_{j-1})\sin(\vartheta\pi)\cos(\vartheta\pi) + \frac{a_j}{b_j}\sin^2(\vartheta\pi). \tag{6.38}$$

Here, the first equation (6.36) yields

$$m_j = a_j + m_{j-1} \quad \text{and} \quad m_{j-1} = m_j - a_j,$$

which, together with the third equation (6.38) of the system given above, produces

$$\frac{a_j}{c_j} = \cos^2(\vartheta\pi) + (a_j + m_{j-1} + m_{j-1})\sin(\vartheta\pi)\cos(\vartheta\pi) + \frac{a_j}{b_j}\sin^2(\vartheta\pi)$$

$$\Leftrightarrow \quad 2m_{j-1}\sin(\vartheta\pi)\cos(\vartheta\pi) = \frac{a_j}{c_j} - \frac{a_j}{b_j}\sin^2(\vartheta\pi) - a_j\sin(\vartheta\pi)\cos(\vartheta\pi) - \cos^2(\vartheta\pi)$$

and

$$\frac{a_j}{c_j} = \cos^2(\vartheta\pi) + (m_j + m_j - a_j)\sin(\vartheta\pi)\cos(\vartheta\pi) + \frac{a_j}{b_j}\sin^2(\vartheta\pi)$$

$$\Leftrightarrow \quad 2m_j\sin(\vartheta\pi)\cos(\vartheta\pi) = \frac{a_j}{c_j} - \frac{a_j}{b_j}\sin^2(\vartheta\pi) + a_j\sin(\vartheta\pi)\cos(\vartheta\pi) - \cos^2(\vartheta\pi).$$

Finally, we get

$$m_{j-1} = \frac{a_j}{2c_j \sin(\vartheta\pi)\cos(\vartheta\pi)} - \frac{a_j}{2b_j}\tan(\vartheta\pi) - \frac{a_j}{2} - \cot(\vartheta\pi) \tag{6.39}$$

and

$$m_j = \frac{a_j}{2c_j \sin(\vartheta\pi)\cos(\vartheta\pi)} - \frac{a_j}{2b_j}\tan(\vartheta\pi) + \frac{a_j}{2} - \cot(\vartheta\pi). \tag{6.40}$$

Observe that the slopes m_{j-1} and m_j, i.e. the slopes of the line segments connecting $\boldsymbol{v}_j$ with its predecessor and successor respectively, can be computed by using only the coefficient values corresponding to the considered vertex $\boldsymbol{v}_j$.

Further, note that the computation of m_{j-1} and m_j in (6.39) and (6.40) is well defined since the values b_j and c_j are non-zero, see *Parts 2* and *3* of this proof. There we have seen that we have the equivalences

$$b_j = 0 \quad \Leftrightarrow \quad m_j = m_{j-1}$$

and

$$c_j = 0 \quad \Leftrightarrow \quad m_j = m_{j-1}.$$

But the case $m_j = m_{j-1}$ would mean that the edges L_j and L_{j-1}, i.e. the line segments connecting the vertex $\boldsymbol{v}_j$ with $\boldsymbol{v}_{j-1}$ and $\boldsymbol{v}_{j+1}$, are line segments on the same line such that we would have a degenerate case, that is, $\boldsymbol{v}_j$ would actually not be a vertex of the polygon D. But such a degenerate case has been excluded. Thus, the coefficients b_j and c_j are not equal to zero.

Moreover, observe that we have $\sin(\vartheta\pi) \neq 0$ and $\cos(\vartheta\pi) \neq 0$ since $\vartheta \in (0,1) \setminus \{\frac{1}{2}\}$. Therefore, also $\tan(\vartheta\pi)$ and $\cot(\vartheta\pi)$ are well defined.

Part 5b: Let us turn back to the set

$$\widetilde{G} = \{\widetilde{\boldsymbol{v}}_1, \ldots, \widetilde{\boldsymbol{v}}_N\} \tag{6.41}$$

of the vertices of the polygonal domain D, where we have to establish which elements have to be connected by line segments in order to uniquely reconstruct D. Since we have three corresponding values $\widetilde{a}_j$, $\widetilde{b}_j$, and $\widetilde{c}_j$ for each element $\widetilde{\boldsymbol{v}}_j$ in $\widetilde{G}$, we can use the formulae (6.39) and (6.40) in order to determine a predecessor and a successor for each vertex. Again, we use the usual conventions

$$\widetilde{\boldsymbol{v}}_0 = \widetilde{\boldsymbol{v}}_N \quad \text{and} \quad \widetilde{\boldsymbol{v}}_{N+1} = \widetilde{\boldsymbol{v}}_1,$$

and the same applies to the coefficients $\widetilde{a}_j$, $\widetilde{b}_j$, and $\widetilde{c}_j$, i.e.

$$\widetilde{a}_0 = \widetilde{a}_N, \quad \widetilde{a}_{N+1} = \widetilde{a}_1, \qquad \widetilde{b}_0 = \widetilde{b}_N, \quad \widetilde{b}_{N+1} = \widetilde{b}_1, \qquad \widetilde{c}_0 = \widetilde{c}_N, \quad \widetilde{c}_{N+1} = \widetilde{c}_1.$$

We start with a vertex of the convex hull of $\widetilde{G}$. Without loss of generality, let this be $\widetilde{\boldsymbol{v}}_1$. Then initialize the set G of ordered vertices of D to

$$G = \{\boldsymbol{v}_1\} \quad \text{with} \quad \boldsymbol{v}_1 := \widetilde{\boldsymbol{v}}_1.$$

Further, we update the set $\widetilde{G}$ of the not yet ordered vertices, that is to say, we set

$$\widetilde{G} := \widetilde{G} \setminus \{\boldsymbol{v}_1\} = \{\widetilde{\boldsymbol{v}}_2, \dots, \widetilde{\boldsymbol{v}}_N\}.$$

Using (6.40), we compute the slope m_1 of the line which contains the line segment that connects the vertex $\boldsymbol{v}_1$ with its successor $\boldsymbol{v}_2$ by

$$m_1 = \frac{\widetilde{a}_1}{2\widetilde{c}_1 \sin(\vartheta\pi)\cos(\vartheta\pi)} - \frac{\widetilde{a}_1}{2\widetilde{b}_1}\tan(\vartheta\pi) + \frac{\widetilde{a}_1}{2} - \cot(\vartheta\pi).$$

The line containing the edge between $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is then given by the equation

$$x_2 = m_1 \cdot (x_1 - v_{1,1}) + v_{1,2}.$$

Thus, we determine all points $\widetilde{\boldsymbol{v}}_j = (\widetilde{v}_{j,1}, \widetilde{v}_{j,2})^{\mathrm{T}}$ in G which fulfil

$$\widetilde{v}_{j,2} = m_1 \cdot (\widetilde{v}_{j,1} - v_{1,1}) + v_{1,2} \tag{6.42}$$

and hence obtain the set

$$S_1 := \{\widetilde{\boldsymbol{v}}_j \in \widetilde{G} \mid \widetilde{\boldsymbol{v}}_j \text{ fulfils (6.42)}\}$$

of possible successors of $\boldsymbol{v}_1$. Now we have to establish which point in S_1 is actually the successor $\boldsymbol{v}_2$. For this purpose, we have to consider the following cases:

(1a) $|S_1| = 1$: If S_1 contains only one element, i.e.

$$S_1 = \{\widetilde{\boldsymbol{v}}_{j_1}\} \quad \text{for some } j_1 \in \{2, \dots, N\},$$

then the successor $\boldsymbol{v}_2$ is given by $\widetilde{\boldsymbol{v}}_{j_1}$, and we set

$$G = \{\boldsymbol{v}_1, \boldsymbol{v}_2\} \quad \text{with} \quad \boldsymbol{v}_2 := \widetilde{\boldsymbol{v}}_{j_1}$$

and

$$\widetilde{G} := \widetilde{G} \setminus \{\boldsymbol{v}_2\}.$$

(1b) $|S_1| \geq 2$: Remember that $\boldsymbol{v}_1$ is a vertex of the convex hull of $\widetilde{G}$. Thus, if S_1 contains at least two elements, all these elements can only lie in the same direction from $\boldsymbol{v}_1$, and only the point in S_1 which is nearest to $\boldsymbol{v}_1$ is a possible choice for $\boldsymbol{v}_2$, see Figure 6.4. Otherwise, we would have a degenerate case, but such a case is excluded by assumption. Having given the only possible choice for the successor $\boldsymbol{v}_2$, we proceed as described in the case (1a).
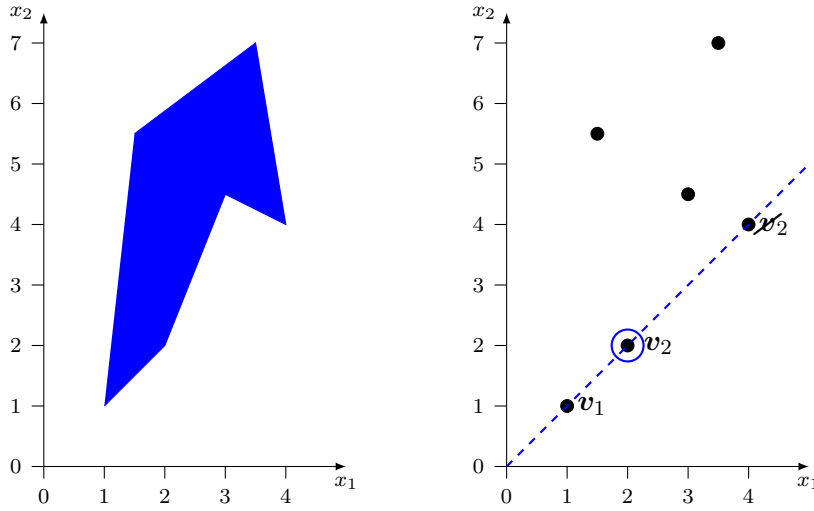
Figure 6.4. Left: *Original unit-height polygon.* Right: *Determination of the order of the computed vertices. Start with* $\boldsymbol{v}_1 := (1,1)^{\mathrm{T}}$. *Computation of* m_1 *accordingly to* (6.40) *leads to* $(2,2)^{\mathrm{T}}$ *and* $(4,4)^{\mathrm{T}}$ *as candidates for* $\boldsymbol{v}_2$. *Choice of* $(4,4)^{\mathrm{T}}$ *would lead to a degenerate case. Hence,* $\boldsymbol{v}_2 = (2,2)^{\mathrm{T}}$.

Now we repeat the approach explained above in order to determine the successor $\boldsymbol{v}_3$ of the vertex $\boldsymbol{v}_2$. Then the line which contains the edge connecting $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$ is described by

$$x_2 = m_2 \cdot \left(x_1 - v_{2,1}\right) + v_{2,2} \tag{6.43}$$

where the slope m_2 is computed using (6.40). The set of possible successors of $\boldsymbol{v}_2$ is given by

$$S_2 := \left\{\widetilde{\boldsymbol{v}}_j \in \widetilde{G} \mid \widetilde{\boldsymbol{v}}_j \text{ fulfils (6.43)}\right\}.$$

The first case which we have to consider now is similar to the case (1a).

(2a) $|S_2| = 1$: If S_2 contains only one element, i.e.

$$S_2 = \left\{\widetilde{\boldsymbol{v}}_{j_2}\right\} \quad \text{for some } j_2 \in \{2, \ldots, N\} \setminus \{j_1\},$$

then $\widetilde{\boldsymbol{v}}_{j_2}$ is the successor of $\boldsymbol{v}_2$, and we set

$$G = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\} \quad \text{with} \quad \boldsymbol{v}_3 := \widetilde{\boldsymbol{v}}_{j_2}$$

and

$$\widetilde{G} := \widetilde{G} \setminus \{\boldsymbol{v}_3\}.$$

The case $|S_2| \geq 2$ is a bit different from the case (1b).

(2b) $|S_2| \geq 2$: Note that we only know the line containing the edge between $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$, but we do not know in which direction we have to go on this line, starting at $\boldsymbol{v}_2$, in order to reach the successor $\boldsymbol{v}_3$. For each of the both directions, only the point in S_2 which is nearest to $\boldsymbol{v}_2$ is a possible choice for $\boldsymbol{v}_3$. Otherwise, we would have a degenerate case. Thus, we obtain a new set $\widetilde{S}_2$ which contains at most two points that fulfil (6.43) and hence are possible choices for the successor $\boldsymbol{v}_3$. If $|\widetilde{S}_2| = 1$, then we follow the lines in the case (2a).

Consider now the case $|\widetilde{S}_2| = 2$, i.e.

$$\widetilde{S}_2 = \left\{ \widetilde{\boldsymbol{v}}_{j_2}, \widetilde{\boldsymbol{v}}_{j_3} \right\} \quad \text{for some } j_2, j_3 \in \{2, \ldots, N\} \setminus \{j_1\}.$$

For each *candidate point* $\boldsymbol{v}_{3,c,j}$ in $\widetilde{S}_2$ ($j = j_2, j_3$), we can compute the slope $m_{3-1,c,j}$ of the line containing the edge between this point and its predecessor by using formula (6.39). The successor $\boldsymbol{v}_3$ is now given by the point $\boldsymbol{v}_{3,c,j}$ for which we have

$$m_2 = m_{3-1,c,j}, \tag{6.44}$$

see also Figure 6.5, and we proceed as described in the case (2a).
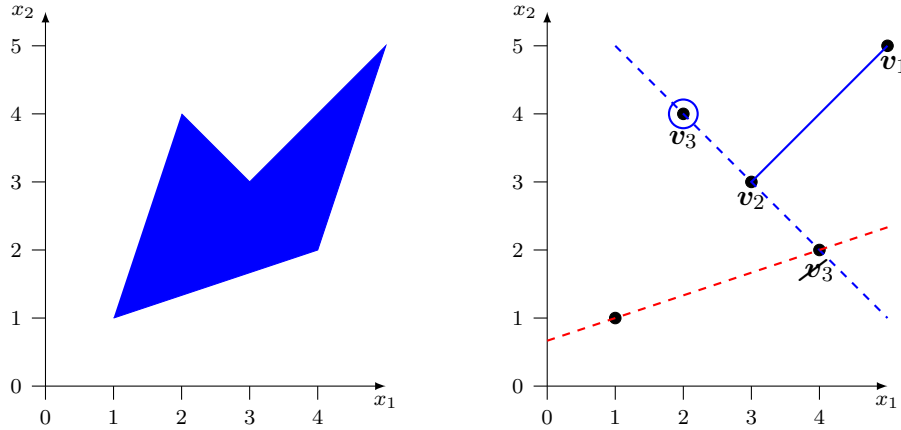
Figure 6.5. Left: *Original unit-height polygon*. Right: *Determination of the order of the computed vertices. Starting with* $\boldsymbol{v}_1 := (5,5)^{\mathrm{T}}$ *leads to* $\boldsymbol{v}_2 = (3,3)^{\mathrm{T}}$. *Computation of* m_2 *accordingly to* (6.40) *leads to* $(2,4)^{\mathrm{T}}$ *and* $(4,2)^{\mathrm{T}}$ *as candidates for* $\boldsymbol{v}_3$ *(blue, dashed line). Equation* (6.39) *yields that the predecessor of* $(2,4)^{\mathrm{T}}$ *lies on the blue, dashed line, and that the predeccessor of* $(4,2)^{\mathrm{T}}$ *lies on the red, dashed line. Thus,* $\boldsymbol{v}_3 = (2,4)^{\mathrm{T}}$.

We use this approach iteratively in order to determine the order of the remaining vertices until we have computed a successor for every element in the set given in (6.41) such that the successor of the last considered vertex is the vertex $\boldsymbol{v}_1$, with which we have started.

If (6.44) holds for both candidate points at some stage of the iteration, we have to choose arbitrarily one of those points as the next successor vertex. If this choice is the wrong one, the algorithm will also terminate when a computed successor is equal to the first considered vertex $\boldsymbol{v}_1$. But in this case we will not have determined a successor for all elements in the set $\widetilde{G}$ in (6.41) such that we have to turn back to the stage of the iteration where we had to choose a successor arbitrarily and continue the iteration by choosing the other candidate point. See also Example 6.3 for more details.

In this manner, we will uniquely reconstruct the polygonal domain D. This concludes the proof. $\qquad\square$
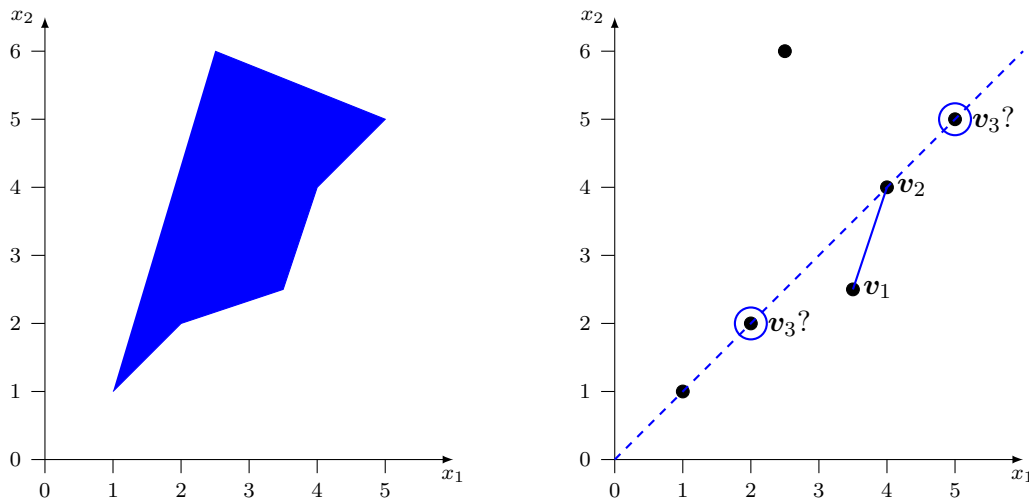

Figure 6.6. Left: *Original unit-height polygon.* Right: *Example for a situation where an arbitrary choice between two candidates for a successor vertex is necessary. Here, $\boldsymbol{v}_3$ has to be chosen.*

Example 6.3 (Necessity for an arbitrary choice of a successor vertex).
Consider the unit-height polygon given in Figure 6.6, and assume that the vertices, which we have computed as described in the proof of Theorem 6.2, are numbered as follows:

$$\widetilde{\boldsymbol{v}}_1 := \begin{pmatrix} 3.5 \\ 2.5 \end{pmatrix}, \quad \widetilde{\boldsymbol{v}}_2 := \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \quad \widetilde{\boldsymbol{v}}_3 := \begin{pmatrix} 5 \\ 5 \end{pmatrix}, \quad \widetilde{\boldsymbol{v}}_4 := \begin{pmatrix} 2.5 \\ 6 \end{pmatrix}, \quad \widetilde{\boldsymbol{v}}_5 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \widetilde{\boldsymbol{v}}_6 := \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

By starting with $\boldsymbol{v}_1 := \widetilde{\boldsymbol{v}}_1$, we find that the successor of the vertex $\boldsymbol{v}_1$ is the point $\widetilde{\boldsymbol{v}}_2$ such that $\boldsymbol{v}_2 = \widetilde{\boldsymbol{v}}_2$.

The computation of the slope of the line containing the edge between $\boldsymbol{v}_2$ and its successor $\boldsymbol{v}_3$ yields that $\boldsymbol{v}_3$ has to lie on the line $x_2 = x_1$. Hence, $\widetilde{\boldsymbol{v}}_3$ and $\widetilde{\boldsymbol{v}}_6$ are candidates for the vertex $\boldsymbol{v}_3$.

We compute the slopes of the lines containing the edges between $\widetilde{\boldsymbol{v}}_j$ and its predecessor for $j = 3$ and $j = 6$ in order to check if both points are still possible candidates. Since this is the case, that is, also the predecessors of $\widetilde{\boldsymbol{v}}_3$ and $\widetilde{\boldsymbol{v}}_6$ lie on the line $x_2 = x_1$, we cannot finally decide which candidate point is actually the successor vertex $\boldsymbol{v}_3$ such that we have to make an arbitrary choice, see Figure 6.6.

Suppose that we choose $\boldsymbol{v}_3 := \widetilde{\boldsymbol{v}}_6$. Determination of the successor of $\boldsymbol{v}_3$ then leads to $\boldsymbol{v}_4 = \boldsymbol{v}_1$. Hence, the algorithm terminates, but we do not have computed successors for all vertices $\widetilde{\boldsymbol{v}}_j$, $j = 1, \ldots, 6$. Especially, the vertex $\widetilde{\boldsymbol{v}}_3$, which we have not chosen as $\boldsymbol{v}_3$, has not been used in the terminated algorithm, see Figure 6.7. This tells us that the choice of $\boldsymbol{v}_3$ has been the wrong one such that we have to take $\boldsymbol{v}_3 := \widetilde{\boldsymbol{v}}_3$ and erase the edge between $\boldsymbol{v}_2$ and $\widetilde{\boldsymbol{v}}_6$. The other information which we have already acquired, i.e. the information that $\widetilde{\boldsymbol{v}}_6$ and $\boldsymbol{v}_1$ are connected by an edge, can still be used. In the next step, we consider the vertex $\boldsymbol{v}_3$ and determine its successor.

We apply this procedure to each next successor until we have considered every vertex $\widetilde{\boldsymbol{v}}_j$. Then we have established the right order of them, see also Figure 6.7.
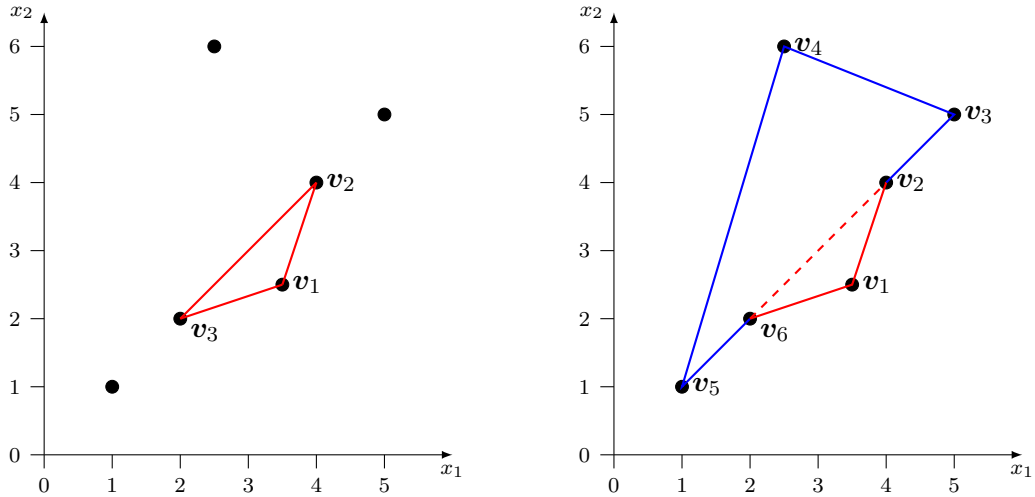
Figure 6.7. Left: *Wrong choice of successor $\boldsymbol{v}_3$ (cf. Figure 6.6) leads to early termination of the algorithm for construction of all edges of the polygon. Some vertices have no successor or predecessor; i.e., they are not connected with other vertices by line segments.* Right: *Edge between $(4,4)^{\mathrm{T}}$ and $(2,2)^{\mathrm{T}}$ has to be erased, and $(5,5)^{\mathrm{T}}$ has to be taken as successor $\boldsymbol{v}_3$.*

Remarks 6.4.

1. It may happen that an incorrect, arbitrary choice as illustrated with the example given above is not discovered instantly as being wrong, but it is possible that other arbitrary choices have to be made until the algorithm terminates without

determining successors for all vertices. In this case, one has to go back to the stages of the iteration where an arbitrary choice has been made, and where the at that time not considered candidate has not been used as a successor afterwards. The edges constructed there have to be erased, and one has to proceed with the unused candidates.

2. In order to avoid the problems mentioned in the first remark, one can determine sets of possible successors and sets of possible predecessors for each vertex $\widetilde{\boldsymbol{v}}_j$ in the set $\widetilde{G}$ given in (6.41). Starting with sets which contain only one possible successor or predecessor, one can construct all edges of the polygon step by step. For this purpose, one has to update the sets of possible successors and predecessors every time when an edge is constructed.

   Note that at least each of the vertices of the convex hull of $\widetilde{G}$ has only one possible successor and one possible predecessor, see the case (1b) in the proof of Theorem 6.2.

3. The scheme for the reconstruction of polygonal domains from sparse Fourier data as proposed in Theorem 6.2 can also be used if the number N of the vertices of the polygon D is not known. In that case, we need an upper bound $M \geq N$ and $3M$ sampling values of the Fourier transform of $\mathbf{1}_D$, compare Remarks 2.3, 3.

We summarize the proposed reconstruction scheme in the following algorithm:

Algorithm 6.5 (Reconstruction of polygonal domains).

- **Input:**
  - Step size $h > 0$ with $h\|\boldsymbol{v}_j\|_2 < \pi$ for $j = 1, \ldots, N$;
  - Fourier samples $\widehat{f}(\boldsymbol{\xi})$ for

    $$\boldsymbol{\xi}^{\mathrm{T}} \in \big\{ \big(\ell h, 0\big), \big(0, \ell h\big), \big(\cos(\vartheta \pi)\ell h, \sin(\vartheta \pi)\ell h\big) \big\}, \quad \ell = 1, \ldots, N,$$

    where $\vartheta \in (0, 1) \setminus \{\frac{1}{2}\}$ needs to be chosen suitably, see Step 6 of the computation.

- **Computation:**
  1. Compute
     $$g(\ell h, 0) = |\ell h|^2 \widehat{f}(\ell h, 0) \quad \text{for } \ell = 1, \ldots, N$$
     and set $g(0, 0) = 0$.
  2. Use Algorithm 2.5 in order to compute the parameters $v_{j,1}$ and a_j of g in (6.18).

3. Compute
$$g(0, \ell h) = |\ell h|^2 \widehat{f}(0, \ell h) \quad \text{for } \ell = 1, \ldots, N,$$
and set $g(0, 0) = 0$.

4. Use Algorithm 2.5 in order to compute the parameters $v_{j,2}$ and b_j of g in (6.21).

5. Compute the Cartesian product of $\{v_{1,1} \ldots, v_{N,1}\}$ and $\{v_{1,2}, \ldots, v_{N,2}\}$ as the set of possible candidates for the true vertices, i.e.
$$K := \big\{ (v_{k,1}, v_{\ell,2})^{\mathrm{T}} : \ k = 1, \ldots, N, \ \ell = 1, \ldots, N \big\}.$$

6. Choose a parameter $\vartheta \in (0, 1) \setminus \{\frac{1}{2}\}$ such that the orthogonal projections of all candidate points in K onto the line $x_2 = \tan(\vartheta\pi)x_1$ are pairwise different.

7. Acquire the N Fourier samples $\widehat{f}\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big)$, $\ell = 1, \ldots, N$. Then compute
$$g\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big) = |\ell h|^2 \widehat{f}\big(\cos(\vartheta\pi)\ell h, \sin(\vartheta\pi)\ell h\big), \quad \ell = 1, \ldots, N,$$
and set $g(0, 0) = 0$.

8. Use Algorithm 2.5 in order to compute the parameters of g in (6.24), namely $\gamma_j = \big(\cos(\vartheta\pi)v_{j,1} + \sin(\vartheta\pi)v_{j,2}\big)$ and c_j.

9. Determine the true vertices $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ by comparison of K with the set $\{\gamma_1, \ldots, \gamma_N\}$; that is, determine all points $(\alpha_k, \beta_\ell)^{\mathrm{T}}$ in the set K for which there exist indices $j \in \{1, \ldots, N\}$ such that
$$\cos(\vartheta\pi)\alpha_k + \sin(\vartheta\pi)\beta_\ell = \gamma_j.$$

   This results in the set $\widetilde{G}$ of all true vertices, but here they are ordered arbitrarily.

10. Establish the right order of the elements in $\widetilde{G}$ by computing successors and predecessors for each vertex. For this purpose, consider the explanations following (6.41), Example 6.3, and Remarks 6.4, 1. and 2.

- **Output:**

   Set $G = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N\}$ of ordered vertices; that is, the boundary ∂D of the polygon D is determined by the closed polygonal chain
   $$[\boldsymbol{v}_1 — \boldsymbol{v}_2 — \boldsymbol{v}_3 — \ldots — \boldsymbol{v}_{N-1} — \boldsymbol{v}_N — \boldsymbol{v}_1].$$

Remark 6.6.

In the proof of Theorem 6.2 and in the above algorithm, we rely upon the assumption that the coordinates of the vertices of the polygons to be reconstructed are pairwise different in both dimensions.

Our reconstruction method uses the orthogonal projections of the vertices onto the x_1- and the x_2-axis in the time domain, which are computed by using Fourier samples from corresponding sampling lines in the frequency domain, i.e. the ω_1- and the ω_2-axis.

The assumption that the vertex coordinates $v_{j,1}$, $j = 1, \ldots, N$, as well as the coordinates $v_{j,2}$, $j = 1, \ldots, N$, are pairwise different is needed since the computed orthogonal projections onto the x_1- and the x_2-axis have to be distinct such that we can use the formula (6.14) for the Fourier transform of the unit-height polygons.

Unless we have a formula for the Fourier transform of unit-height polygons where this restriction is not necessary, we can reconstruct general polygons, i.e. polygons without any restrictions on the vertex coordinates, only if we use much more Fourier samples.

By using Fourier data sampled on distinct lines $\widetilde{l}_k := \{\lambda \boldsymbol{u}_k \,|\, \lambda \in \mathbb{R}\}$, $k = 1, \ldots, r$, in the frequency domain where $\boldsymbol{u}_k$ are unit vectors, we obtain the orthogonal projections of the vertices onto the distinct lines $l_k := \{\lambda \boldsymbol{u}_k \,|\, \lambda \in \mathbb{R}\}$, $k = 1, \ldots, r$, in the time domain. For all $k \in \{1, \ldots, r\}$, these orthogonal projections are given by[4]

$$\langle \boldsymbol{u}_k, \boldsymbol{v}_j \rangle \boldsymbol{u}_k, \quad j = 1, \ldots, N.$$

Thus, we have to ensure that we have always two distinct lines on which the projections are pairwise different. For this purpose, we need $\frac{N(N-1)}{2} + 2$ pairwise different lines l_k and $\widetilde{l}_k$ in the time domain and the frequency domain respectively, see [12, §3], where Buhmann and Pinkus use this approach for a similar problem.

We take the usual coordinate axes and consider the vertical and horizontal axes of $\frac{N(N-1)}{2}$ rotated versions of the usual Cartesian coordinate system as further sampling lines. Let us assume that the orthogonal projections of the vertices are pairwise different onto the lines l_1 and l_2. Then we can use the Fourier samples from the lines $\widetilde{l}_1$ and $\widetilde{l}_2$ as initializations for Algorithm 6.5. Observe that we have to use an appropriate coordinate transformation, corresponding to the coordinate system spanned by l_1 and l_2, in order to continue with Algorithm 6.5.

Altogether, we need Fourier samples from $\frac{N(N-1)}{2} + 3$ sampling lines. But we will only use the Fourier data from three lines. Therefore, it would be of high interest to have an approach similar to our proposed reconstruction method where only three sampling lines are needed. This means that another representation for the Fourier transform of a unit-height polygon is required.

---

[4]See [39, p. 118, (1.88)], for instance.

6.2. Numerical results

This section concludes the chapter about the reconstruction of polygonal shapes with a few numerical examples. We use simulated Fourier data on the x_1-axis, on the x_2-axis, and on a third, adaptively chosen sampling line, which is determined accordingly to the explanations in Section 4.3, that is, we solve a max-min problem similar to (4.45) in order to decide on the third sampling line.

Observe in the following examples that some vertices have nearly the same first or second coordinate. But nevertheless we are able to recover the original vertices and the maximal reconstruction error for the vertex coordinates has an order of magnitude equal to -7.

Further, note that concave polygons are considered in the second, third, and fourth example such that the step of determining the order of the computed vertices is a very important step in order to recover the original shape.
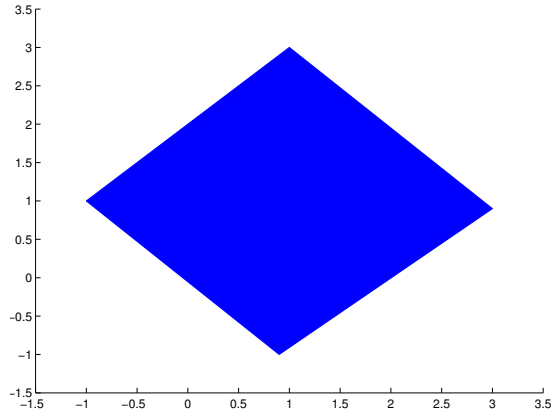
Figure 6.8. *Unit-height polygon determined by $(\boldsymbol{v}_j)_{j=1}^4$ given in* Table 6.1.

j	$v_{j,1}$	$|v_{j,1} - v_{j,1}^*|$	$v_{j,2}$	$|v_{j,2} - v_{j,2}^*|$
1	-1	$8.882 \cdot 10^{-16}$	1	$7.649 \cdot 10^{-14}$
2	1	$1.91 \cdot 10^{-14}$	3	$4.441 \cdot 10^{-16}$
3	3	$2.665 \cdot 10^{-15}$	0.9	$8.737 \cdot 10^{-14}$
4	0.9	$8.626 \cdot 10^{-14}$	-1	$2.22 \cdot 10^{-16}$

Table 6.1. *Vertices of the unit-height polygon displayed in* Figure 6.8 *and approximate reconstruction errors. The sampling step size is $h = 0.7$, and the measure of the angle between the first and the third sampling line is $64°$.*

Figure 6.9. *Unit-height polygon determined by $(\boldsymbol{v}_j)_{j=1}^4$ given in* Table 6.2.

j	$v_{j,1}$	$|v_{j,1} - v_{j,1}^*|$	$v_{j,2}$	$|v_{j,2} - v_{j,2}^*|$
1	0.05	$2.732 \cdot 10^{-12}$	0	$9.992 \cdot 10^{-16}$
2	0	$2.273 \cdot 10^{-12}$	4	$3.109 \cdot 10^{-15}$
3	0.5	$2.224 \cdot 10^{-12}$	3	$2.665 \cdot 10^{-15}$
4	2	$1.776 \cdot 10^{-15}$	2.4	0

Table 6.2. *Vertices of the unit-height polygon displayed in* Figure 6.9 *and approximate reconstruction errors. The sampling step size is $h = 0.7$, and the measure of the angle between the first and the third sampling line is $10°$.*
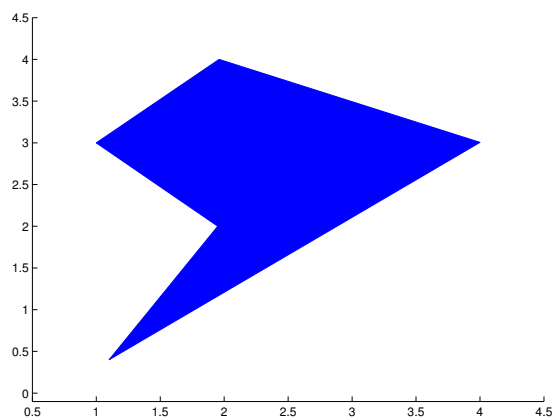
Figure 6.10. *Unit-height polygon determined by $(\boldsymbol{v}_j)_{j=1}^5$ given in* Table 6.3.

j	$v_{j,1}$	$\lvert v_{j,1} - v_{j,1}^* \rvert$	$v_{j,2}$	$\lvert v_{j,2} - v_{j,2}^* \rvert$
1	1	$1.601 \cdot 10^{-9}$	3	$2.792 \cdot 10^{-7}$
2	1.95	$2.677 \cdot 10^{-7}$	2	$5.788 \cdot 10^{-12}$
3	1.1	$5.524 \cdot 10^{-9}$	0.4	$1.197 \cdot 10^{-13}$
4	4	$1.403 \cdot 10^{-13}$	3.005	$1.68 \cdot 10^{-7}$
5	1.96	$4.96 \cdot 10^{-7}$	4	$7.994 \cdot 10^{-13}$

Table 6.3. *Vertices of the unit-height polygon displayed in* Figure 6.10 *and approximate reconstruction errors. The sampling step size is $h = 0.4$, and the measure of the angle between the first and the third sampling line is $45°$.*
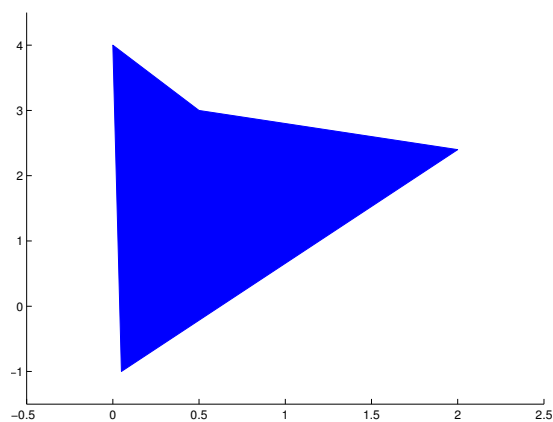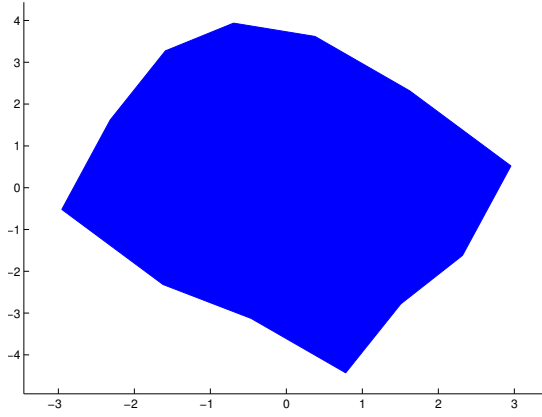

Figure 6.11. *Unit-height polygon with 12 vertices. The approximate reconstruction errors are given in* Table 6.4.

j	$\lvert v_{j,1} - v_{j,1}^* \rvert$	$\lvert v_{j,2} - v_{j,2}^* \rvert$	j	$\lvert v_{j,1} - v_{j,1}^* \rvert$	$\lvert v_{j,2} - v_{j,2}^* \rvert$
1	$1.014 \cdot 10^{-11}$	$1.11 \cdot 10^{-14}$	7	$1.94 \cdot 10^{-12}$	$2.665 \cdot 10^{-15}$
2	$1.399 \cdot 10^{-10}$	$2.132 \cdot 10^{-14}$	8	$3.758 \cdot 10^{-11}$	$9.77 \cdot 10^{-15}$
3	$8.304 \cdot 10^{-14}$	$2.887 \cdot 10^{-15}$	9	$1.052 \cdot 10^{-13}$	0
4	$4.441 \cdot 10^{-16}$	$9.992 \cdot 10^{-16}$	10	$8.882 \cdot 10^{-16}$	$1.11 \cdot 10^{-16}$
5	$3.161 \cdot 10^{-10}$	$8.882 \cdot 10^{-16}$	11	$8.695 \cdot 10^{-11}$	$4.885 \cdot 10^{-15}$
6	$3.238 \cdot 10^{-11}$	$9.326 \cdot 10^{-16}$	12	$6.564 \cdot 10^{-12}$	$3.197 \cdot 10^{-14}$

Table 6.4. *Vertices of the unit-height polygon displayed in* Figure 6.11 *and reconstruction errors. The sampling step size is $h = 0.7$, and the measure of the angle between the first and the third sampling line is $56.75°$.*

7. Conclusion and Outlook

In this dissertation, we have asked how to reconstruct structured functions by means of a smallest possible set of Fourier data. Answering this question, we have derived novel algorithms for the unique reconstruction of different classes of structured functions, based upon the Prony method.

In particular, we have derived reconstruction methods for step functions, linear combinations of non-uniform B-splines, and linear combinations of non-uniform translates of low-pass filter functions in the one-dimensional case using sparse Fourier samples, and we have generalized these results to the tensor-product case in two dimensions.

Furthermore, we have considered the reconstruction of linear combinations of non-uniform translates of multivariate functions where the variables are not separable. We have developed reconstruction methods using adaptive sampling and have shown that we are able to uniquely recover such functions from only a small number of Fourier samples taken on some lines through the origin in the frequency domain. The adaptive sampling scheme is essential in order to obtain unique reconstructions by means of a smallest possible set of Fourier data. In the two-dimensional case, we have shown that one needs only $\mathcal{O}(N)$ Fourier samples to uniquely recover linear combinations of N non-uniform translates if one uses the adaptive sampling scheme. Contrarily, one needs $\mathcal{O}(N^2)$ sampling values if a deterministic sampling approach is applied.

We have also shown that polygonal shapes in the real plane can be uniquely reconstructed by using a similar approach of adaptive sampling such that $\mathcal{O}(N)$ Fourier samples suffice to recover polygons with N vertices where we have emphasized that the polygons do not have to be convex but can also be concave.

In several cases, we have illustrated our proposed reconstruction methods with numerical experiments where we use exact (simulated) Fourier data. However, in the case of noisy measurements, the performance of the reconstruction can be greatly improved if a larger number of Fourier data is available, see [23, 51, 55]. In particular, for small data sets we recommend the preprocessing step of data filtering presented in [23].

Recently, in [13], Candès and Fernandez-Granda have proposed the reconstruction of functions of the form (2.1) using a total variation minimization formulation. In order to tackle this minimization problem, a semi-definite program is applied to solve the dual problem in a first step. The obtained result is used to define a special polynomial whose zeros on the unit circle are related to the wanted parameters T_j. The exact

connections between that minimization approach in the context of super-resolution and the direct algorithms for the Prony method are not yet established.

It would also be of great interest to generalize, and maybe combine, the approaches for the reconstruction of non-uniform translates in Section 4.2 and the reconstruction of polygonal shapes in Chapter 6 such that more general functions and shapes can be considered. For example, the *shape from moments problem*, see [21,27,47], is extended from polygons to algebraic curves in [29].

In this context, the question arises if the approach from Chapter 5 for a d-dimensional setting with $d > 2$ can be combined with the reconstruction of polygonal shapes in Chapter 6 such that we can transfer it to the reconstruction of polytopes. This is, for example, treated in [28], where integral moments of a d-dimensional polytope are used, together with the Prony method, in order to reconstruct the polytope. But there are considered only convex polytopes. Thus, it would be interesting to extend the theory developed in this dissertation.

A. Basic properties of the Fourier transform

The thesis on hand deals with the reconstruction of functions from sparse Fourier data. Therefore, the Fourier transform is an essential mathematical instrument. We give here an overview of the definition and basic properties of the Fourier transform which we need throughout this work. The following definitions and properties are collected from [11, pp. 103–109], [30, pp. 38–41], [52, pp. 89–97], and [63, pp. 1–5].

Definition and Proposition A.1 (*Fourier transform on L^1*).
Let $f \in L^1(\mathbb{R}^d)$. The *Fourier transform* $\widehat{f} : \mathbb{R}^d \to \mathbb{C}$ is defined by

$$\mathcal{F}\{f\}(\boldsymbol{\omega}) := \widehat{f}(\boldsymbol{\omega}) := \int_{\mathbb{R}^d} f(\boldsymbol{x}) \, \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\omega},\boldsymbol{x}\rangle} \, \mathrm{d}\boldsymbol{x}, \quad \boldsymbol{\omega} \in \mathbb{R}^d.$$

The mapping $\mathcal{F} : f \mapsto \widehat{f}$ is a linear transformation from $L^1(\mathbb{R}^d)$ into $C_0(\mathbb{R}^d)$. This transformation is called *Fourier transform*. The domain of the original function f is regarded as the *spatial* domain (or *time* domain), and the domain of the transformed function $\widehat{f}$ is considered as the *frequency* domain (or *Fourier* domain).

In the following propositions, we assume $f \in L^1(\mathbb{R}^d)$.

Proposition A.2 (Conjugate symmetry).
If the function f is real-valued, then we have the relation

$$\widehat{f}(-\boldsymbol{\omega}) = \overline{\widehat{f}(\boldsymbol{\omega})} \quad \text{for all } \boldsymbol{\omega} \in \mathbb{R}^d.$$

Proposition A.3 (Translation in the spatial domain).
The Fourier transform of the translated function $f(\cdot - \boldsymbol{a})$ for $\boldsymbol{a} \in \mathbb{R}^d$ is given by

$$\mathcal{F}\{f(\cdot - \boldsymbol{a})\}(\boldsymbol{\omega}) = \mathrm{e}^{-\mathrm{i}\langle\boldsymbol{\omega},\boldsymbol{a}\rangle} \, \mathcal{F}\{f\}(\boldsymbol{\omega}), \quad \boldsymbol{\omega} \in \mathbb{R}^d.$$

Proposition A.4 (Differentiation in the spatial domain).
If the mixed partial derivative $\mathrm{D}^{\boldsymbol{\alpha}} f$ exists and $\mathrm{D}^{\boldsymbol{\alpha}} f \in L^1(\mathbb{R}^d)$ holds, then the Fourier transform of this derivative is given by

$$\mathcal{F}\{\mathrm{D}^{\boldsymbol{\alpha}} f\}(\boldsymbol{\omega}) = \mathrm{i}^{|\boldsymbol{\alpha}|}\, \boldsymbol{\omega}^{\boldsymbol{\alpha}} \mathcal{F}\{f\}(\boldsymbol{\omega}), \quad \boldsymbol{\omega} \in \mathbb{R}^d.$$

Here, $\boldsymbol{\alpha}$ denotes a multi-index, i.e. a d-tuple $\boldsymbol{\alpha} := (\alpha_1, \alpha_2, \ldots, \alpha_d)$ of non-negative integers, and $\mathrm{D}^{\boldsymbol{\alpha}}$ is the corresponding differential operator. The derivative $\mathrm{D}^{\boldsymbol{\alpha}} f$ is then given by

$$\mathrm{D}^{\boldsymbol{\alpha}} f(\boldsymbol{\omega}) := \frac{\partial^{\alpha_1 + \alpha_2 + \ldots + \alpha_d}}{\partial \omega_1^{\alpha_1} \partial \omega_2^{\alpha_2} \cdots \partial \omega_d^{\alpha_d}} f(\boldsymbol{\omega}) \quad \text{for } \boldsymbol{\omega} := (\omega_1, \omega_2, \ldots, \omega_d)^{\mathrm{T}} \in \mathbb{R}^d.$$

Further, we have

$$|\boldsymbol{\alpha}| := \sum_{j=1}^{d} \alpha_j \qquad \text{and} \qquad \boldsymbol{\omega}^{\boldsymbol{\alpha}} := \prod_{j=1}^{d} \omega_j^{\alpha_j}.$$

Proposition A.5 (Differentiation in the frequency domain).
If the function $\mathbb{R}^d \ni \boldsymbol{x} \mapsto g(\boldsymbol{x}) := \boldsymbol{x}^{\boldsymbol{\alpha}} f(\boldsymbol{x})$ is an element of $L^1(\mathbb{R}^d)$, then we have

$$\mathrm{D}^{\boldsymbol{\alpha}}\big(\mathcal{F}\{f\}\big)(\boldsymbol{\omega}) = (-\mathrm{i})^{|\boldsymbol{\alpha}|} \mathcal{F}\{g\}(\boldsymbol{\omega}) \quad \text{for all } \boldsymbol{\omega} \in \mathbb{R}^d.$$

The Fourier transform can also be defined on the Schwartz space $\mathcal{S}(\mathbb{R}^d)$.

Definition and Proposition A.6 (Fourier transform on $\mathcal{S}(\mathbb{R}^d)$).
Consider $f \in \mathcal{S}(\mathbb{R}^d)$. Then the Fourier transform $\mathcal{F} : f \mapsto \widehat{f}$ is a linear transformation from $\mathcal{S}(\mathbb{R}^d)$ onto $\mathcal{S}(\mathbb{R}^d)$.

Remark A.7.
The statements from Propositions A.2–A.5 are also valid if we replace the space $L^1(\mathbb{R}^d)$ by the Schwartz space $\mathcal{S}(\mathbb{R}^d)$. Note that the conditions in Propositions A.4 and A.5 are always fulfilled for functions $f \in \mathcal{S}(\mathbb{R}^d)$, cf. [66, pp. 211-212].

In the special case that $\varphi \in \mathcal{S} = \mathcal{S}(\mathbb{R})$, the m-th derivative of the Fourier transform is computed as follows.

Proposition A.8.
Let $\varphi \in \mathcal{S}$. Then we have

$$\frac{\partial^m}{\partial \omega^m} \mathcal{F}\{\varphi\}(\omega) = \mathcal{F}\{(-\mathrm{i}\,\cdot)^m \varphi(\cdot)\}(\omega) \quad \text{for all } \omega \in \mathbb{R}.$$

Proof. We have

$$\frac{\partial^m}{\partial\omega^m}\mathcal{F}\{\varphi\}(\omega) = \frac{\partial^m}{\partial\omega^m}\widehat{\varphi}(\omega) = \frac{\partial^m}{\partial\omega^m}\int\limits_{-\infty}^{\infty}\varphi(x)\,\mathrm{e}^{-\mathrm{i}\,\omega x}\,\mathrm{d}x. \tag{A.1}$$

For $k = 0,\ldots,m-1$, define the function $g_k : \mathbb{R}\times\mathbb{R}\to\mathbb{C}$ by

$$g_k(\omega,x) := \varphi(x)(-\mathrm{i}\,x)^k\,\mathrm{e}^{-\mathrm{i}\,\omega x}.$$

Since the function $\Phi_k : \mathbb{R}\ni x\mapsto x^k\varphi(x)\in\mathbb{C}$ is absolutely integrable by Lemma C.1, g_k fulfils the following conditions:

- $g_k(\omega,\cdot)$ is absolutely integrable for all $\omega\in\mathbb{R}$.

- g_k is differentiable with respect to the first argument, that is, $\frac{\partial}{\partial\omega}g_k(\omega_0,x)$ exists for all $x\in\mathbb{R}$ for a given ω_0.

- For $k = 0,\ldots,m-1$, and for all $\omega\in\mathbb{R}$, we have

$$\left|\frac{\partial}{\partial\omega}g_k(\omega,x)\right| = \left|(-\mathrm{i}\,x)^{k+1}\varphi(x)\,\mathrm{e}^{-\mathrm{i}\,\omega x}\right| = |x^{k+1}\varphi(x)| = |\Phi_{k+1}|$$

  where Φ_{k+1} is absolutely integrable.

Hence, we can apply the theorem of differentiation under the integral sign[1] in (A.1) m times. This yields

$$\begin{aligned}
\frac{\partial^m}{\partial\omega^m}\widehat{\varphi}(\omega) &= \frac{\partial^m}{\partial\omega^m}\int\limits_{-\infty}^{\infty}\varphi(x)\,\mathrm{e}^{-\mathrm{i}\,\omega x}\,\mathrm{d}x \\
&= \int\limits_{-\infty}^{\infty}\frac{\partial^m}{\partial\omega^m}\varphi(x)\,\mathrm{e}^{-\mathrm{i}\,\omega x}\,\mathrm{d}x \\
&= \int\limits_{-\infty}^{\infty}(-\mathrm{i}\,x)^m\varphi(x)\,\mathrm{e}^{-\mathrm{i}\,\omega x}\,\mathrm{d}x,
\end{aligned}$$

which proves the claim. $\qquad\square$

Remark A.9 (Fourier transform of distributions).

Since the concept of distributions in the univariate case and in the bivariate case is introduced in Appendix B and D respectively, the reader is referred to Sections B.4 and D.3 for the definition of the Fourier transform of distributions.

---

[1]See [22, pp. 147 f., 5.7].

B. Proofs for non-uniform spline representations

In Section 3.2, we compute the derivatives of a non-uniform spline function f of the form (3.10). This works with the concept of ordinary functions as long as the order of the B-splines is greater than or equal to three. B-splines of order two and one are piecewise linear functions and step functions respectively. These functions are not differentiable at their knots. In order to be able to compute derivatives in these cases, we have to work with distributions and generalized functions.

B.1. Distributions and generalized functions

Let $g : \mathbb{R} \to \mathbb{C}$ be a locally integrable function. Then a regular distribution[1] $\mu_{(g)}$ will be defined by the functional

$$C_c^\infty \ni \varphi \mapsto \langle \mu_{(g)}, \varphi \rangle := \int\limits_{-\infty}^{+\infty} \varphi(x) g(x)\, \mathrm{d}x, \tag{B.1}$$

see [35, p. 11]. The distribution $\mu_{(g)}$ can be identified with the function g, which will then be called generalized function.

The derivative of the distribution $\mu_{(g)}$ is the distribution $\mathrm{D}\,\mu_{(g)}$ which is defined by

$$\langle \mathrm{D}\,\mu_{(g)}, \varphi \rangle = \langle \mu_{(g)}, -\varphi' \rangle \tag{B.2}$$

for all $\varphi \in C_c^\infty$, see [35, p. 37]. Since $\varphi \in C_c^\infty$, this means that $\mu_{(g)}$ is infinitely differentiable. The kth derivative (in this distributional sense) is denoted by $\mathrm{D}^k\,\mu_{(g)}$.

If the generalized function g is continuously differentiable, the distributional derivative $\mathrm{D}\,\mu_{(g)}$ is the distribution defined by g' and (B.1). For all $\varphi \in C_c^\infty$, we have

$$\langle \mathrm{D}\,\mu_{(g)}, \varphi \rangle = \langle \mu_{(g)}, -\varphi' \rangle$$
$$= -\int\limits_{-\infty}^{\infty} \varphi'(x) g(x)\, \mathrm{d}x = -\left\{ \left[\varphi(x) g(x)\right]_{-\infty}^{\infty} - \int\limits_{-\infty}^{\infty} \varphi(x) g'(x)\, \mathrm{d}x \right\}, \tag{B.3}$$

---

[1] This work deals with the concept of distributions and generalized functions only as detailed as necessary. For more information about distributions, the reader is referred to books on the subject, e.g. [26, 30, 35].

see [35, p. 37]. The function φ has compact support. Further, g is locally integrable, which yields $\lim_{|x| \to \infty} |g(x)| < \infty$. Thus, the first term on the right side of (B.3) vanishes, and we get

$$\left\langle \mathrm{D}\, \mu_{(g)}, \varphi \right\rangle = \int\limits_{-\infty}^{\infty} \varphi(x) g'(x)\, \mathrm{d}x = \left\langle \mu_{(g')}, \varphi \right\rangle. \tag{B.4}$$

B.2. Proof of Lemma 3.5

Note that the function f from (3.10) is locally integrable. If we apply (B.4) k times to f and the distribution $\mu_{(f)}$, defined accordingly to (B.1), we get the following for $m \geq 3$ and $k = 1, \ldots, m-2$:

$$\left\langle \mathrm{D}^k\, \mu_{(f)}, \varphi \right\rangle \overset{\text{(B.4)}}{=} \left\langle \mu_{(f^{(k)})}, \varphi \right\rangle \overset{\text{Lemma}}{\underset{3.4}{=}} \int\limits_{-\infty}^{\infty} \left(\sum_{j=1}^{N+k} c_j^k B_j^{m-k}(x) \right) \varphi(x)\, \mathrm{d}x \tag{B.5}$$

for all $\varphi \in C_c^\infty$ where the coefficients c_j^k, $j = 1, \ldots, N+k$, are defined in Lemma 3.4.

In the case $k = m-2$, the generalized function $f^{(m-2)}$ is a linear combination of B-splines of order two. These B-splines are piecewise linear. Previously, when we examined ordinary functions, we could not differentiate the function $f^{(m-2)}$. Since we are now working with distributions, we can also compute the derivative $\mathrm{D}^{m-1}\, \mu_{(f)}$ of the distribution $\mu_{(f)}$ in the case $m \geq 2$. For all $\varphi \in C_c^\infty$, we have

$$\begin{aligned}
\left\langle \mathrm{D}^{m-1}\, \mu_{(f)}, \varphi \right\rangle &= \left\langle \mathrm{D} \left(\mathrm{D}^{m-2}\, \mu_{(f)} \right), \varphi \right\rangle = \left\langle \mathrm{D}^{m-2}\, \mu_{(f)}, -\varphi' \right\rangle && \text{by (B.2)} \\
&= \left\langle \mu_{(f^{(m-2)})}, -\varphi' \right\rangle && \text{by (B.5)} \\
&= -\int\limits_{-\infty}^{\infty} \left(\sum_{j=1}^{N+m-2} c_j^{m-2} B_j^2(x) \right) \varphi'(x)\, \mathrm{d}x && \text{by Lemma 3.4} \\
&= \sum_{j=1}^{N+m-2} c_j^{m-2} \cdot (-1) \int\limits_{-\infty}^{\infty} B_j^2(x) \varphi'(x)\, \mathrm{d}x. && \tag{B.6}
\end{aligned}$$

We denote the integral in (B.6) by I, and we use integration by parts to compute it. The function B_j^2 does not have a derivative in the classical sense everywhere, that is to say, it is not differentiable at the knots T_j, T_{j+1}, and T_{j+2}. Further, B_j^2 and φ are compactly supported. Therefore, we get

$$-I = (-1) \int\limits_{-\infty}^{\infty} B_j^2(x) \varphi'(x)\, \mathrm{d}x = -\left\{ \left[B_j^2(x)\varphi(x) \right]_{-\infty}^{\infty} - \int\limits_{-\infty}^{\infty} (B_j^2)'(x)\varphi(x)\, \mathrm{d}x \right\}$$

$$= \lim_{a \searrow T_j} \lim_{b \nearrow T_{j+1}} \int_a^b \left(\frac{x - T_j}{T_{j+1} - T_j} \right)' \varphi(x) \, dx$$

$$+ \lim_{c \searrow T_{j+1}} \lim_{d \nearrow T_{j+2}} \int_c^d \left(\frac{T_{j+2} - x}{T_{j+2} - T_{j+1}} \right)' \varphi(x) \, dx \qquad \text{by (3.11), (3.12)}$$

$$= \lim_{a \searrow T_j} \lim_{b \nearrow T_{j+1}} \int_a^b \frac{1}{T_{j+1} - T_j} \varphi(x) \, dx + \lim_{c \searrow T_{j+1}} \lim_{d \nearrow T_{j+2}} \int_c^d \frac{-1}{T_{j+2} - T_{j+1}} \varphi(x) \, dx$$

$$= \int_{-\infty}^{\infty} \left(\frac{1}{T_{j+1} - T_j} \mathbf{1}_{[T_j, T_{j+1})}(x) - \frac{1}{T_{j+2} - T_{j+1}} \mathbf{1}_{[T_{j+1}, T_{j+2})}(x) \right) \varphi(x) \, dx. \qquad \text{(B.7)}$$

Plugging this into (B.6) yields for all $\varphi \in C_c^\infty$

$$\left\langle \mathrm{D}^{m-1} \mu_{(f)}, \varphi \right\rangle = \sum_{j=1}^{N+m-2} c_j^{m-2} \int_{-\infty}^{\infty} \left(\frac{1}{T_{j+1} - T_j} \mathbf{1}_{[T_j, T_{j+1})}(x) \right.$$

$$\left. - \frac{1}{T_{j+2} - T_{j+1}} \mathbf{1}_{[T_{j+1}, T_{j+2})}(x) \right) \varphi(x) \, dx$$

$$= \int_{-\infty}^{\infty} \left(\sum_{j=1}^{N+m-2} \frac{1}{T_{j+1} - T_j} c_j^{m-2} \mathbf{1}_{[T_j, T_{j+1})}(x) \right.$$

$$\left. - \sum_{j=1}^{N+m-2} \frac{1}{T_{j+2} - T_{j+1}} c_j^{m-2} \mathbf{1}_{[T_{j+1}, T_{j+2})}(x) \right) \varphi(x) \, dx$$

$$= \int_{-\infty}^{\infty} \left(\sum_{j=1}^{N+m-2} \frac{1}{T_{j+1} - T_j} c_j^{m-2} \mathbf{1}_{[T_j, T_{j+1})}(x) \right.$$

$$\left. - \sum_{j=2}^{N+m-1} \frac{1}{T_{j+1} - T_j} c_{j-1}^{m-2} \mathbf{1}_{[T_j, T_{j+1})}(x) \right) \varphi(x) \, dx$$

$$= \int_{-\infty}^{\infty} \sum_{j=1}^{N+m-1} c_j^{m-1} B_j^1(x) \varphi(x) \, dx \qquad \text{(B.8)}$$

with

$$c_j^{m-1} := \frac{1}{T_{j+1} - T_j} \cdot \left(c_j^{m-2} - c_{j-1}^{m-2} \right), \qquad j = 1, \ldots, N + m - 1,$$

and the convention $c_0^{m-2} = c_{N+m-1}^{m-2} = 0$.

Equation (B.8) means that the derivative $\mathrm{D}^{m-1}\mu_{(f)}$ $(m \geq 2)$ is the distribution $\mu_{(f^{(m-1)})}$ defined by (B.1) and the generalized function

$$f^{(m-1)} : \mathbb{R} \ni x \mapsto \sum_{j=1}^{N+m-1} c_j^{m-1} B_j^1(x) \in \mathbb{R}. \tag{B.9}$$

This proves Lemma 3.5. Observe that denoting this generalized function by $f^{(m-1)}$ means an abuse of notation since $f^{(m-1)}$ does not denote a derivative in the classical sense. $\qquad\square$

B.3. Proof of Lemma 3.6

In the next step, we want to compute the m-th distributional derivative $\mathrm{D}^m\mu_{(f)}$ for $m \geq 1$. Note that $f^{(m-1)}$ given in (B.9) is locally integrable. For $\varphi \in C_c^\infty$, we have

$$
\begin{aligned}
\left\langle \mathrm{D}^m\mu_{(f)}, \varphi \right\rangle &= \left\langle \mathrm{D}\left(\mathrm{D}^{m-1}\mu_{(f)}\right), \varphi \right\rangle \\
&= \left\langle \mathrm{D}^{m-1}\mu_{(f)}, -\varphi' \right\rangle && \text{by (B.2)} \\
&= -\int_{-\infty}^{\infty} f^{(m-1)}(x)\varphi'(x)\,\mathrm{d}x && \text{by (B.8), (B.9)} \\
&= -\sum_{j=1}^{N+m-1} c_j^{m-1} \int_{-\infty}^{\infty} B_j^1(x)\varphi'(x)\,\mathrm{d}x \\
&= -\sum_{j=1}^{N+m-1} c_j^{m-1} \int_{T_j}^{T_{j+1}} \varphi'(x)\,\mathrm{d}x \\
&= \sum_{j=1}^{N+m-1} c_j^{m-1}\Big(\varphi(T_j) - \varphi(T_{j+1})\Big) && \text{(B.10)} \\
&= \sum_{j=1}^{N+m-1} c_j^{m-1}\Big((\tau_{-T_j}\varphi)(0) - (\tau_{-T_{j+1}}\varphi)(0)\Big)
\end{aligned}
$$

where τ_a is the translation operator which is defined for $a \in \mathbb{R}$ and an ordinary function φ by

$$\tau_a\varphi(x) := \varphi(x - a). \tag{B.11}$$

The Dirac delta distribution[2] δ is defined by the functional

$$C_c^\infty \ni \varphi \mapsto \langle \delta, \varphi \rangle := \varphi(0). \tag{B.12}$$

---

[2]The Dirac delta distribution is not a regular distribution, that is, it cannot be defined by a locally integrable function as in (B.1).

Using this definition, we get the following for $\varphi \in C_c^\infty$:

$$
\begin{aligned}
\left\langle \mathrm{D}^m \, \mu_{(f)}, \varphi \right\rangle &= \sum_{j=1}^{N+m-1} c_j^{m-1} \left(\left\langle \delta, \tau_{-T_j} \varphi \right\rangle - \left\langle \delta, \tau_{-T_{j+1}} \varphi \right\rangle \right) \\
&= \sum_{j=1}^{N+m-1} c_j^{m-1} \left(\left\langle \tau_{T_j} \delta, \varphi \right\rangle - \left\langle \tau_{T_{j+1}} \delta, \varphi \right\rangle \right).
\end{aligned}
\tag{B.13}
$$

Here, $\tau_a \mu$ denotes the translate of a distribution μ and is defined for $a \in \mathbb{R}$ by

$$
\left\langle \tau_a \mu, \varphi \right\rangle = \left\langle \mu, \tau_{-a} \varphi \right\rangle \quad \text{for all } \varphi \in C_c^\infty,
\tag{B.14}
$$

see [35, p. 37]. From (B.13) we compute

$$
\begin{aligned}
\left\langle \mathrm{D}^m \, \mu_{(f)}, \varphi \right\rangle &= \sum_{j=1}^{N+m-1} c_j^{m-1} \left\langle \tau_{T_j} \delta, \varphi \right\rangle - \sum_{j=1}^{N+m-1} c_j^{m-1} \left\langle \tau_{T_{j+1}} \delta, \varphi \right\rangle \\
&= \sum_{j=1}^{N+m-1} c_j^{m-1} \left\langle \tau_{T_j} \delta, \varphi \right\rangle - \sum_{j=2}^{N+m} c_{j-1}^{m-1} \left\langle \tau_{T_j} \delta, \varphi \right\rangle \\
&= \sum_{j=1}^{N+m} c_j^m \left\langle \tau_{T_j} \delta, \varphi \right\rangle
\end{aligned}
$$

where the coefficients c_j^m for $j = 1, \ldots, N + m$ are defined by

$$
c_j^m := c_j^{m-1} - c_{j-1}^{m-1}
\tag{B.15}
$$

with the usual convention $c_0^{m-1} = c_{N+m}^{m-1} = 0$. For addition and scalar multiplication of distributions, the same rules as for general functionals apply, see [35, p. 36]. Hence,

$$
\left\langle \mathrm{D}^m \, \mu_{(f)}, \varphi \right\rangle = \left\langle \sum_{j=1}^{N+m} c_j^m \left(\tau_{T_j} \delta \right), \varphi \right\rangle \quad \text{for all } \varphi \in C_c^\infty,
$$

and the m-th distributional derivative of the generalized function f is a linear combination of translated Dirac delta distributions. We will denote this distribution by $f^{(m)}$, i.e.

$$
f^{(m)} := \mathrm{D}^m \, \mu_{(f)} = \sum_{j=1}^{N+m} c_j^m \left(\tau_{T_j} \delta \right).
\tag{B.16}
$$

Thus, the proof of Lemma 3.6 is completed. Observe that $f^{(m)}$ does not denote a derivative in the classical sense. $\qquad \square$

B.4. Proof of Lemma 3.7

As a last step, we have to introduce the definition of the Fourier transform of a distribution in order to get a representation of the Fourier transform of f in (3.10).

Let μ be one of the examined distributions above. Then the Fourier transform $\mathcal{F}\{\mu\}$ is defined on the linear space $\mathcal{S}$ as the linear functional which fulfils[3]

$$\big\langle \mathcal{F}\{\mu\}, \varphi \big\rangle = \big\langle \mu, \widehat{\varphi} \big\rangle \quad \text{for all } \varphi \in \mathcal{S}, \tag{B.17}$$

see [35, pp. 108-110]. Here, $\widehat{\varphi}$ denotes the ordinary Fourier transform of φ. The space $\mathcal{S} := \mathcal{S}(\mathbb{R})$ consists of all complex-valued functions defined on the real axis which are infinitely differentiable and, together with all their derivatives, rapidly decreasing as the absolute value of the argument tends to infinity[4]. Note that $\mathcal{S} \subset L^p(\mathbb{R})$ for all $p \geq 1$, see [66, pp. 211-212]. The elements of $\mathcal{S}$ are called Schwartz functions.

Now we have all needed concepts to compute the Fourier transform of the distribution $f^{(m)}$. For all $\varphi \in \mathcal{S}$, we get

$$
\begin{aligned}
\big\langle \mathcal{F}\{f^{(m)}\}, \varphi \big\rangle &\overset{\text{(B.16)}}{=} \sum_{j=1}^{N+m} c_j^m \big\langle \mathcal{F}\{(\tau_{T_j}\delta)\}, \varphi \big\rangle \overset{\text{(B.17)}}{=} \sum_{j=1}^{N+m} c_j^m \big\langle \tau_{T_j}\delta, \widehat{\varphi} \big\rangle \\
&\overset{\text{(B.14)}}{=} \sum_{j=1}^{N+m} c_j^m \big\langle \delta, \tau_{-T_j}\widehat{\varphi} \big\rangle \\
&\overset{\text{(B.11)}}{\underset{\text{(B.12)}}{=}} \sum_{j=1}^{N+m} c_j^m \widehat{\varphi}(T_j) \\
&= \sum_{j=1}^{N+m} c_j^m \int_{-\infty}^{\infty} \varphi(\omega)\, e^{-\,i\,\omega T_j}\, d\omega \\
&= \int_{-\infty}^{\infty} \left(\sum_{j=1}^{N+m} c_j^m\, e^{-\,i\,\omega T_j} \right) \varphi(\omega)\, d\omega \\
&= \big\langle \mu_{(F)}, \varphi \big\rangle
\end{aligned}
\tag{B.18}
$$

---

[3] Actually, one has to introduce the concept of tempered distributions in order to define the Fourier transform of a distribution, but in this work only the necessary concepts are introduced. For the examined distributions, the Fourier transform is well defined as one can see in [35, p. 108-109].

[4] A function g of rapid decrease as the absolute value of the argument tends to infinity is defined by the property

$$\lim_{|x| \to \infty} |x^n g(x)| = 0 \quad \text{for all } n \in \mathbb{N}_0,$$

see [35, p. 102].

with the generalized function

$$F : \mathbb{R} \ni \omega \mapsto \sum_{j=1}^{N+m} c_j^m \, \mathrm{e}^{-\mathrm{i}\,\omega T_j} \in \mathbb{C}. \tag{B.19}$$

Note that F is locally integrable. Hence, the distribution $\mu_{(F)}$ is well-defined by (B.1).

Since $f^{(m)} = \mathrm{D}^m \mu_{(f)}$, we also have the following for all $\varphi \in \mathcal{S}$ by applying (B.2) m times:

$$\langle \mathcal{F}\{f^{(m)}\}, \varphi \rangle = \langle \mathcal{F}\{\mathrm{D}^m \mu_{(f)}\}, \varphi \rangle = \langle \mathrm{D}^m \mu_{(f)}, \widehat{\varphi} \rangle = \langle \mu_{(f)}, (-1)^m \widehat{\varphi}^{(m)} \rangle. \tag{B.20}$$

The m-th derivative of the Fourier transform $\widehat{\varphi}$ can be computed as follows:

$$(-1)^m \widehat{\varphi}^{(m)}(\omega) = (-1)^m \int_{-\infty}^{\infty} (-\mathrm{i}\,x)^m \varphi(x) \, \mathrm{e}^{-\mathrm{i}\,\omega x} \, \mathrm{d}x \qquad \text{by Proposition A.8}$$

$$= \int_{-\infty}^{\infty} (\mathrm{i}\,x)^m \varphi(x) \, \mathrm{e}^{-\mathrm{i}\,\omega x} \, \mathrm{d}x.$$

Inserting this into (B.20), we obtain

$$\langle \mathcal{F}\{f^{(m)}\}, \varphi \rangle = \langle \mu_{(f)}, \mathcal{F}\{(\mathrm{i}\,\cdot)^m \varphi(\cdot)\} \rangle.$$

Further, we have

$$
\begin{aligned}
\langle \mathcal{F}\{f^{(m)}\}, \varphi \rangle &= \langle \mu_{(f)}, \mathcal{F}\{(\mathrm{i}\,\cdot)^m \varphi(\cdot)\} \rangle \\
&= \int_{-\infty}^{\infty} f(t) \left(\int_{-\infty}^{\infty} (\mathrm{i}\,x)^m \varphi(x) \, \mathrm{e}^{-\mathrm{i}\,t x} \, \mathrm{d}x \right) \mathrm{d}t \\
&= \int_{-\infty}^{\infty} (\mathrm{i}\,x)^m \varphi(x) \left(\int_{-\infty}^{\infty} f(t) \, \mathrm{e}^{-\mathrm{i}\,t x} \, \mathrm{d}t \right) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} (\mathrm{i}\,x)^m \widehat{f}(x) \varphi(x) \, \mathrm{d}x = \left\langle \mu_{\left((\mathrm{i}\,\cdot)^m \widehat{f}\right)}, \varphi \right\rangle
\end{aligned}
\tag{B.21}
$$

for all $\varphi \in \mathcal{S}$. Exchanging the order of integration is possible due to Lemma C.2.

From (B.18) and (B.21) we conclude

$$\langle \mu_{(F)}, \varphi \rangle = \left\langle \mu_{\left((\mathrm{i}\,\cdot)^m \widehat{f}\right)}, \varphi \right\rangle \quad \text{for all } \varphi \in \mathcal{S}$$

with the generalized function F from (B.19). Therefore, we can use the following representation of the Fourier transform of the function f in (3.10):

$$(\mathrm{i}\,\omega)^m \widehat{f}(\omega) = \sum_{j=1}^{N+m} c_j^m \, \mathrm{e}^{-\mathrm{i}\,\omega T_j}$$

for $\omega \in \mathbb{R}$ where the coefficients c_j^m are defined by (B.15).

This proves Lemma 3.7. Observe that the equations and definitions from the previous sections which we have used here are also valid for $\varphi \in \mathcal{S}$. $\qquad\square$

C. Lemmas for integration

In the previous chapter of the Appendix, we have given a short introduction to the theory of distributions, including derivatives and the Fourier transform, in order to explain the computation of the Fourier transform of a non-uniform spline function in Section 3.2. In the process of discussing the concept of distributions, we have used properties of Schwartz functions which we will now prove in the following lemmas.

Lemma C.1.
Let $\varphi \in \mathcal{S}$. The function Φ_m defined by (C.1) is absolutely integrable for all $m \in \mathbb{N}_0$.

$$\Phi_m : \mathbb{R} \ni x \mapsto x^m \varphi(x) \in \mathbb{C} \tag{C.1}$$

Proof. By definition of $\mathcal{S}$, the function φ is of rapid decrease[1]. Before we prove the claim, we need the following equivalent statements for the property of rapid decrease[2]:

φ is of rapid decrease

$$\begin{aligned}
\Leftrightarrow \quad & \sup_{x \in \mathbb{R}} |x^n \varphi(x)| < \infty \text{ for all } n \in \mathbb{N}_0 \\
\Leftrightarrow \quad & \forall\, (n \in \mathbb{N}_0)\, \exists\, (C \geq 0)\, \forall\, (x \in \mathbb{R}) : |\varphi(x)| \leq \frac{C}{1 + x^{2n}}.
\end{aligned} \tag{C.2}$$

By the definition of the property of rapid decrease, also the function Φ_m is rapidly decreasing for all $m \in \mathbb{N}_0$. Since φ is infinitely differentiable, Φ_m is infinitely differentiable too. This means that $\Phi_m \in \mathcal{S}$. Therefore, we can apply (C.2) and find a constant $C \geq 0$ with

$$|\Phi_m(x)| = |x^m \varphi(x)| \leq \frac{C}{1 + x^2} \quad \text{for all } x \in \mathbb{R}.$$

This yields $\Phi_m \in L^1(\mathbb{R})$ for all $m \in \mathbb{N}_0$ by the direct comparison test[3] since

$$\int_{\mathbb{R}} \frac{1}{1 + x^2}\, \mathrm{d}x = \lim_{R \to \infty} \big(\arctan(R) - \arctan(-R) \big) = \pi < \infty.$$

$\square$

---

[1]The property of rapid decrease is defined in Section B.4, footnote 4 on p. 148.
[2]These equivalences are taken from [38].
[3]See [22, p. 130, 3.3a),e)].

Lemma C.2.
Let f be defined as in (3.10) and $\varphi \in \mathcal{S}$. Then we have

$$\int\limits_{-\infty}^{\infty} f(t) \left(\int\limits_{-\infty}^{\infty} (\mathrm{i}\,x)^m \varphi(x)\, \mathrm{e}^{-\mathrm{i}\,tx}\, \mathrm{d}x \right) \mathrm{d}t = \int\limits_{-\infty}^{\infty} (\mathrm{i}\,x)^m \varphi(x) \left(\int\limits_{-\infty}^{\infty} f(t)\, \mathrm{e}^{-\mathrm{i}\,tx}\, \mathrm{d}t \right) \mathrm{d}x.$$

Proof. Let us denote the integral on the left side of the equation to prove by I. If we take the absolute value of the integrand in I, we get the following:

$$\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \left| f(t)(\mathrm{i}\,x)^m \varphi(x)\, \mathrm{e}^{-\mathrm{i}\,tx} \right| \mathrm{d}t\, \mathrm{d}x = \int\limits_{-\infty}^{\infty} |x^m \varphi(x)| \left(\int\limits_{-\infty}^{\infty} |f(t)|\, \mathrm{d}t \right) \mathrm{d}x$$

$$= \int\limits_{-\infty}^{\infty} |x^m \varphi(x)| \cdot C \, \mathrm{d}x$$

with a constant $C < \infty$ since $f \in L^1(\mathbb{R})$ by definition of f as an almost everywhere continuous function with compact support. Further, we have

$$\int\limits_{-\infty}^{\infty} |x^m \varphi(x)|\, \mathrm{d}x < \infty$$

by Lemma C.1. Hence,

$$\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \left| f(t)\, \mathrm{i}\, x^m \varphi(x)\, \mathrm{e}^{-\mathrm{i}\,tx} \right| \mathrm{d}t\, \mathrm{d}x < \infty,$$

and the conditions for Fubini's theorem[4] are fulfilled. Therefore, we can exchange the order of integration in I, which yields the claim. $\qquad\square$

---

[4]See [22, pp. 175–176, 2.1c)].

D. Proofs for non-uniform tensor-product spline representations

In Subection 4.1.1, we have given results for the reconstruction of non-uniform tensor-product spline representations from sparse Fourier samples. Those results are due to the structure of the Fourier transform of tensor-products f of the form (4.1). In order to get the representation of the Fourier transform $\widehat{f}$, we have to compute the derivatives of f, where we have to work with distributions and generalized functions. This approach is similar to the univariate case, see Section B.1. We still need to prove the statements from Lemmas 4.1–4.3, which will be done in the following sections.

D.1. Proof of Lemma 4.1

In Section B.1, we have introduced the concept of distributions and generalized functions. There, we have only considered the univariate case. That concept easily transfers to the multidimensional setting, cf. [30, Chapter 2].

Let us consider a locally integrable function $g : \mathbb{R}^2 \to \mathbb{C}$. Then the functional

$$C_c^\infty(\mathbb{R}^2) \ni \varphi \mapsto \left\langle \mu_{(g)}, \varphi \right\rangle := \int_{\mathbb{R}^2} \varphi(\boldsymbol{x}) g(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \tag{D.1}$$

defines a regular distribution $\mu_{(g)}$, see [30, p. 30]. We identify this distribution with the so-called generalized function g.

Note that the function f of the form (4.1) is locally integrable. Further, observe that the partial derivatives of f with respect to the first and second variable are continuously differentiable up to order $m_1 - 3$ and $m_2 - 3$ respectively (in the case $m_1, m_2 \geq 3$). Thus, we have

$$\mathrm{D}^{(m_1-2,m_2-2)} \mu_{(f)} = \mu_{\left(\frac{\partial^{m_2-2}}{\partial \cdot_2^{m_2-2}} \frac{\partial^{m_1-2}}{\partial \cdot_1^{m_1-2}} f(\cdot_1, \cdot_2) \right)},$$

cf. [30, pp. 32–33], where the right-hand side of the equation given above is defined accordingly to (D.1). Here, $(m_1 - 2, m_2 - 2)$ denotes a multi-index.

At the beginning of Subsection 4.1.1, we have seen that the following holds for the case $m_1, m_2 \geq 3$:

$$\frac{\partial^{m_2-2}}{\partial x_2^{m_2-2}} \left(\frac{\partial^{m_1-2}}{\partial x_1^{m_1-2}} f(x_1, x_2) \right) = \sum_{j=1}^{N_1+m_1-2} \sum_{k=1}^{N_2+m_2-2} c_{j,k}^{m_1-2,m_2-2} B_j^2(x_1) B_k^2(x_2)$$

where the coefficients $c_{j,k}^{m_1-2,m_2-2}$ are defined in (4.2)–(4.5). Therefore, the derivative $\mathrm{D}^{(m_1-2,m_2-2)} \mu_{(f)}$ is the distribution $\mu_{(f^{(m_1-2,m_2-2)})}$ defined by (D.1) and the generalized function

$$f^{(m_1-2,m_2-2)} : \mathbb{R}^2 \ni (x_1, x_2)^{\mathrm{T}} \mapsto \sum_{j=1}^{N_1+m_1-2} \sum_{k=1}^{N_2+m_2-2} c_{j,k}^{m_1-2,m_2-2} B_j^2(x_1) B_k^2(x_2) \in \mathbb{R}.$$

In the following part, we consider the case $m_1, m_2 \geq 2$. For all $\varphi \in C_c^\infty(\mathbb{R}^2)$, we have[1]

$$\begin{aligned}
\langle \mathrm{D}^{(m_1-1,m_2-1)} \mu_{(f)}, \varphi \rangle \\
&= \langle \mathrm{D}^{(1,1)} \big(\mathrm{D}^{(m_1-2,m_2-2)} \mu_{(f)} \big), \varphi \rangle \\
&= \left\langle \big(\mathrm{D}^{(m_1-2,m_2-2)} \mu_{(f)} \big), (-1)^{|(1,1)|} \frac{\partial}{\partial \cdot_1} \frac{\partial}{\partial \cdot_2} \varphi(\cdot_1, \cdot_2) \right\rangle \\
&= \left\langle \mu_{(f^{(m_1-2,m_2-2)})}, \frac{\partial}{\partial \cdot_1} \frac{\partial}{\partial \cdot_2} \varphi(\cdot_1, \cdot_2) \right\rangle \\
&= \int_{\mathbb{R}^2} \sum_{j=1}^{N_1+m_1-2} \sum_{k=1}^{N_2+m_2-2} c_{j,k}^{m_1-2,m_2-2} B_j^2(x_1) B_k^2(x_2) \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} \varphi(x_1, x_2) \, \mathrm{d}\boldsymbol{x}.
\end{aligned}$$

Since B_j^2 and B_k^2 are continuous functions with compact support and $\varphi \in C_c^\infty(\mathbb{R}^2)$, the conditions for Fubini's theorem[2] are fulfilled. Thus, we get

$$\langle \mathrm{D}^{(m_1-1,m_2-1)} \mu_{(f)}, \varphi \rangle = \int_{-\infty}^{\infty} \sum_{j=1}^{N_1+m_1-2} \sum_{k=1}^{N_2+m_2-2} c_{j,k}^{m_1-2,m_2-2} B_k^2(x_2) I_1(x_2) \, \mathrm{d}x_2 \qquad \text{(D.2)}$$

where we use the notation

$$I_1(x_2) := \int_{-\infty}^{\infty} B_j^2(x_1) \frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_2} \varphi(x_1, x_2) \right) \, \mathrm{d}x_1.$$

---

[1]See [30, pp. 32-33] for the first two equalities.
[2]See [22, pp. 175 f., 2.1c)].

Using the same steps as in (B.7), we find

$$
I_1(x_2) = -\int\limits_{-\infty}^{\infty} \left(\frac{1}{T_{j+1} - T_j} \mathbf{1}_{[T_j,T_{j+1})}(x_1) - \frac{1}{T_{j+2} - T_{j+1}} \mathbf{1}_{[T_{j+1},T_{j+2})}(x_1) \right) \frac{\partial}{\partial x_2} \varphi(x_1,x_2) \, dx_1.
$$

Next, we insert this into (D.2) and obtain

$$
\left\langle D^{(m_1-1,m_2-1)} \mu_{(f)}, \varphi \right\rangle
$$

$$
= -\int\limits_{-\infty}^{\infty} \sum_{k=1}^{N_2+m_2-2} B_k^2(x_2) \int\limits_{-\infty}^{\infty} \left(\sum_{j=1}^{N_1+m_1-2} \frac{1}{T_{j+1}-T_j} c_{j,k}^{m_1-2,m_2-2} \mathbf{1}_{[T_j,T_{j+1})}(x_1) \right.
$$

$$
\left. - \sum_{j=1}^{N_1+m_1-2} \frac{1}{T_{j+2}-T_{j+1}} c_{j,k}^{m_1-2,m_2-2} \mathbf{1}_{[T_{j+1},T_{j+2})}(x_1) \right) \frac{\partial}{\partial x_2} \varphi(x_1,x_2) \, dx_1 \, dx_2
$$

$$
= -\int\limits_{-\infty}^{\infty} \sum_{k=1}^{N_2+m_2-2} B_k^2(x_2) \int\limits_{-\infty}^{\infty} \left(\sum_{j=1}^{N_1+m_1-2} \frac{1}{T_{j+1}-T_j} c_{j,k}^{m_1-2,m_2-2} \mathbf{1}_{[T_j,T_{j+1})}(x_1) \right.
$$

$$
\left. - \sum_{j=2}^{N_1+m_1-1} \frac{1}{T_{j+1}-T_j} c_{j-1,k}^{m_1-2,m_2-2} \mathbf{1}_{[T_j,T_{j+1})}(x_1) \right) \frac{\partial}{\partial x_2} \varphi(x_1,x_2) \, dx_1 \, dx_2
$$

$$
= -\int\limits_{-\infty}^{\infty} \sum_{k=1}^{N_2+m_2-2} B_k^2(x_2) \int\limits_{-\infty}^{\infty} \sum_{j=1}^{N_1+m_1-1} c_{j,k}^{m_1-1,m_2-2} B_j^1(x_1) \frac{\partial}{\partial x_2} \varphi(x_1,x_2) \, dx_1 \, dx_2
$$

where

$$
c_{j,k}^{m_1-1,m_2-2} := \frac{c_{j,k}^{m_1-2,m_2-2} - c_{j-1,k}^{m_1-2,m_2-2}}{T_{j+1} - T_j}, \qquad \begin{aligned} &j = 1, \ldots, N_1 + m_1 - 1, \\ &k = 1, \ldots, N_2 + m_2 - 2, \end{aligned}
$$

and

$$
c_{0,k}^{m_1-2,m_2-2} = c_{N_1+m_1-1,k}^{m_1-2,m_2-2} = 0, \qquad\qquad k = 1, \ldots, N_2 + m_2 - 2.
$$

Since we can exchange the order of integration in the last equation by Fubini's theorem, we get

$$
\left\langle D^{(m_1-1,m_2-1)} \mu_{(f)}, \varphi \right\rangle = -\int\limits_{-\infty}^{\infty} \sum_{j=1}^{N_1+m_1-1} B_j^1(x_1) \sum_{k=1}^{N_2+m_2-2} c_{j,k}^{m_1-1,m_2-2} I_2(x_1) \, dx_1.
$$

where we define $I_2(x_1)$ by

$$
I_2(x_1) := \int\limits_{-\infty}^{\infty} B_k^2(x_2) \frac{\partial}{\partial x_2} \varphi(x_1,x_2) \, dx_2.
$$

In order to simplify the integral $I_2(x_1)$, we use the same approach as in (B.7). Hence, we obtain

$$I_2(x_1) = -\int\limits_{-\infty}^{\infty} \left(\frac{1}{S_{k+1} - S_k} \mathbf{1}_{[S_k, S_{k+1})}(x_2) - \frac{1}{S_{k+2} - S_{k+1}} \mathbf{1}_{[S_{k+1}, S_{k+2})}(x_2) \right) \varphi(x_1, x_2) \, dx_2.$$

Putting the last two equations together, we get

$$\left\langle D^{(m_1-1, m_2-1)} \mu_{(f)}, \varphi \right\rangle$$

$$= \int\limits_{-\infty}^{\infty} \sum_{j=1}^{N_1+m_1-1} B_j^1(x_1) \int\limits_{-\infty}^{\infty} \left(\sum_{k=1}^{N_2+m_2-2} \frac{1}{S_{k+1} - S_k} c_{j,k}^{m_1-1, m_2-2} \mathbf{1}_{[S_k, S_{k+1})}(x_2) \right.$$

$$\left. - \sum_{k=1}^{N_2+m_2-2} \frac{1}{S_{k+2} - S_{k+1}} c_{j,k}^{m_1-1, m_2-2} \mathbf{1}_{[S_{k+1}, S_{k+2})}(x_2) \right) \varphi(x_1, x_2) \, dx_2 \, dx_1$$

$$= \int\limits_{-\infty}^{\infty} \sum_{j=1}^{N_1+m_1-1} B_j^1(x_1) \int\limits_{-\infty}^{\infty} \left(\sum_{k=1}^{N_2+m_2-2} \frac{1}{S_{k+1} - S_k} c_{j,k}^{m_1-1, m_2-2} \mathbf{1}_{[S_k, S_{k+1})}(x_2) \right.$$

$$\left. - \sum_{k=2}^{N_2+m_2-1} \frac{1}{S_{k+1} - S_k} c_{j,k-1}^{m_1-1, m_2-2} \mathbf{1}_{[S_k, S_{k+1})}(x_2) \right) \varphi(x_1, x_2) \, dx_2 \, dx_1$$

$$= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \sum_{j=1}^{N_1+m_1-1} \sum_{k=1}^{N_2+m_2-1} c_{j,k}^{m_1-1, m_2-1} B_j^1(x_1) B_k^1(x_2) \varphi(x_1, x_2) \, dx_2 \, dx_1$$

where the coefficients are defined by

$$c_{j,k}^{m_1-1, m_2-1} := \frac{c_{j,k}^{m_1-1, m_2-2} - c_{j,k-1}^{m_1-1, m_2-2}}{S_{k+1} - S_k}, \qquad \begin{array}{l} k = 1, \ldots, N_2 + m_2 - 1, \\ j = 1, \ldots, N_1 + m_1 - 1, \end{array}$$

with the convention

$$c_{j,0}^{m_1-1, m_2-2} = c_{j, N_2+m_2-1}^{m_1-1, m_2-2} = 0, \qquad j = 1, \ldots, N_1 + m_1 - 1.$$

Again, we use Fubini's theorem in the last double integral (B_j^1 and B_k^1 are continuous functions with compact support, and $\varphi \in C_c^{\infty}(\mathbb{R}^2)$), which finally results in

$$\left\langle D^{(m_1-1, m_2-1)} \mu_{(f)}, \varphi \right\rangle = \int\limits_{\mathbb{R}^2} \left(\sum_{j=1}^{N_1+m_1-1} \sum_{k=1}^{N_2+m_2-1} c_{j,k}^{m_1-1, m_2-1} B_j^1(x_1) B_k^2(x_2) \right) \varphi(x_1, x_2) \, d\boldsymbol{x}.$$

According to (D.1), this means that $\mathrm{D}^{(m_1-1,m_2-1)} \mu_{(f)}$ for $m_1, m_2 \geq 2$ is the distribution generated by the generalized function

$$f^{(m_1-1,m_2-1)} : \mathbb{R}^2 \ni (x_1, x_2)^{\mathrm{T}} \mapsto \sum_{j=1}^{N_1+m_1-1} \sum_{k=1}^{N_2+m_2-1} c_{j,k}^{m_1-1,m_2-1} B_j^1(x_1) B_k^1(x_2) \in \mathbb{R}. \quad \text{(D.3)}$$

This proves Lemma 4.1. □

D.2. Proof of Lemma 4.2

In order to prove Lemma 4.2, we have to compute the derivative $\mathrm{D}^{(1,1)} \mu_{(f^{(m_1-1,m_2-1)})}$ in the case $m_1, m_2 \geq 1$ where $f^{(m_1-1,m_2-1)}$ is the generalized function defined in (D.3). Observe that $f^{(m_1-1,m_2-1)}$ is locally integrable. Therefore, we have the following for $\varphi \in C_c^\infty(\mathbb{R}^2)$:

$$\begin{aligned}
\langle \mathrm{D}^{(m_1,m_2)} \mu_{(f)}, \varphi \rangle &= \langle \mathrm{D}^{(1,1)} \big(\mathrm{D}^{(m_1-1,m_2-1)} \mu_{(f)} \big), \varphi \rangle \\
&= \Big\langle \big(\mathrm{D}^{(m_1-1,m_2-1)} \mu_{(f)} \big), (-1)^{|(1,1)|} \frac{\partial}{\partial \bullet_1} \frac{\partial}{\partial \bullet_2} \varphi(\bullet_1, \bullet_2) \Big\rangle \\
&= \Big\langle \mu_{(f^{(m_1-1,m_2-1)})}, \frac{\partial}{\partial \bullet_1} \frac{\partial}{\partial \bullet_2} \varphi(\bullet_1, \bullet_2) \Big\rangle \\
&= \int_{\mathbb{R}^2} \sum_{j=1}^{N_1+m_1-1} \sum_{k=1}^{N_2+m_2-1} c_{j,k}^{m_1-1,m_2-1} B_j^1(x_1) B_k^1(x_2) \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} \varphi(x_1, x_2) \, \mathrm{d}\boldsymbol{x}, \quad \text{(D.4)}
\end{aligned}$$

where we have applied (D.3). In the first two equalities, we have used propoerties of derivatives of distributions, cf. [30, pp. 32-33]. We can apply Fubini's theorem to the double integral in (D.4) since B_j^1 and B_k^1 are almost everywhere continuous functions with compact support and $\varphi \in C_c^\infty(\mathbb{R}^2)$. Then we have

$$\langle \mathrm{D}^{(m_1,m_2)} \mu_{(f)}, \varphi \rangle = \int_{-\infty}^{\infty} \sum_{j=1}^{N_1+m_1-1} \sum_{k=1}^{N_2+m_2-1} c_{j,k}^{m_1-1,m_2-1} B_k^1(x_2) L_1(x_2) \, \mathrm{d}x_2 \quad \text{(D.5)}$$

with the integral

$$L_1(x_2) := \int_{-\infty}^{\infty} B_j^1(x_1) \frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_2} \varphi(x_1, x_2) \right) \mathrm{d}x_1.$$

Using the same steps as in (B.10) for the computation of $L_1(x_2)$, we obtain

$$L_1(x_2) = - \left(\frac{\partial}{\partial x_2} \varphi(T_j, x_2) - \frac{\partial}{\partial x_2} \varphi(T_{j+1}, x_2) \right). \quad \text{(D.6)}$$

Combination of (D.5) and (D.6) produces the result

$$
\begin{aligned}
\big\langle &\mathrm{D}^{(m_1,m_2)}\mu_{(f)},\varphi\big\rangle \\
&= -\int_{-\infty}^{\infty}\sum_{k=1}^{N_2+m_2-1} B_k^1(x_2)\Bigg(\sum_{j=1}^{N_1+m_1-1} c_{j,k}^{m_1-1,m_2-1}\frac{\partial}{\partial x_2}\varphi(T_j,x_2) \\
&\qquad\qquad\qquad - \sum_{j=1}^{N_1+m_1-1} c_{j,k}^{m_1-1,m_2-1}\frac{\partial}{\partial x_2}\varphi(T_j,x_2)\Bigg)\,\mathrm{d}x_2 \\
&= -\int_{-\infty}^{\infty}\sum_{k=1}^{N_2+m_2-1} B_k^1(x_2)\Bigg(\sum_{j=1}^{N_1+m_1-1} c_{j,k}^{m_1-1,m_2-1}\frac{\partial}{\partial x_2}\varphi(T_j,x_2) \\
&\qquad\qquad\qquad - \sum_{j=2}^{N_1+m_1} c_{j-1,k}^{m_1-1,m_2-1}\frac{\partial}{\partial x_2}\varphi(T_j,x_2)\Bigg)\,\mathrm{d}x_2 \\
&= -\sum_{k=1}^{N_2+m_2-1}\int_{-\infty}^{\infty} B_k^1(x_2)\sum_{j=1}^{N_1+m_1} c_{j,k}^{m_1,m_2-1}\frac{\partial}{\partial x_2}\varphi(T_j,x_2)\,\mathrm{d}x_2 \qquad\text{(D.7)}
\end{aligned}
$$

with coefficients $c_{j,k}^{m_1,m_2-1}$ defined by

$$
c_{j,k}^{m_1,m_2-1} = c_{j,k}^{m_1-1,m_2-1} - c_{j-1,k}^{m_1-1,m_2-1}, \qquad
\begin{aligned}
&j = 1,\ldots,N_1+m_1, \\
&k = 1,\ldots,N_2+m_2-1,
\end{aligned}
\qquad\text{(D.8)}
$$

where we use the usual convention

$$
c_{0,k}^{m_1-1,m_2-1} = c_{N_1+m_1,k}^{m_1-1,m_2-1} = 0, \qquad\qquad k = 1,\ldots,N_2+m_2-1. \qquad\text{(D.9)}
$$

Due to

$$
\int_{-\infty}^{\infty} B_k^1(x_2)\frac{\partial}{\partial x_2}\varphi(T_j,x_2)\,\mathrm{d}x_2 = -\big[\varphi(T_j,S_k) - \varphi(T_j,S_{k+1})\big]
$$

(cf. the computational steps in (B.10)), (D.7) yields

$$
\begin{aligned}
\big\langle \mathrm{D}^{(m_1,m_2)}\mu_{(f)},\varphi\big\rangle &= \sum_{j=1}^{N_1+m_1}\sum_{k=1}^{N_2+m_2-1} c_{j,k}^{m_1,m_2-1}\big[\varphi(T_j,S_k) - \varphi(T_j,S_{k+1})\big] \\
&= \sum_{j=1}^{N_1+m_1}\Bigg(\sum_{k=1}^{N_2+m_2-1} c_{j,k}^{m_1,m_2-1}\varphi(T_j,S_k) - \sum_{k=2}^{N_2+m_2} c_{j,k-1}^{m_1,m_2-1}\varphi(T_j,S_k)\Bigg) \\
&= \sum_{j=1}^{N_1+m_1}\sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2}\varphi(T_j,S_k). \qquad\text{(D.10)}
\end{aligned}
$$

Here, we use the definition

$$c_{j,k}^{m_1,m_2} := c_{j,k}^{m_1,m_2-1} - c_{j,k-1}^{m_1,m_2-1}, \qquad \begin{aligned} &k = 1, \ldots, N_2 + m_2, \\ &j = 1, \ldots, N_1 + m_1, \end{aligned} \tag{D.11}$$

with the convention

$$c_{j,0}^{m_1,m_2-1} = c_{j,N_2+m_2}^{m_1,m_2-1} = 0, \qquad j = 1, \ldots, N_1 + m_1. \tag{D.12}$$

In order to express the right-hand side in (D.10) in terms of distributions, we have to use the definitions of the translation operator and the Dirac delta distribution, which we have only introduced for the univariate case so far. In the case of ordinary functions, the translation operator $\tau_{a,b}$ for $(a,b)^{\mathrm{T}} \in \mathbb{R}^2$ is given by

$$\tau_{a,b}\varphi(x_1, x_2) := \varphi(x_1 - a, x_2 - b). \tag{D.13}$$

The translate $\tau_{a,b}\mu$ of a distribution μ is defined for $(a,b)^{\mathrm{T}} \in \mathbb{R}^2$ by

$$\langle \tau_{a,b}\mu, \varphi \rangle = \langle \mu, \tau_{-a,-b}\varphi \rangle \quad \text{for all } \varphi \in C_c^{\infty}(\mathbb{R}^2), \tag{D.14}$$

cf. [26, p. 9]. Further, the rule

$$C_c^{\infty}(\mathbb{R}^2) \ni \varphi \mapsto \langle \delta, \varphi \rangle := \varphi(0,0) \tag{D.15}$$

defines the Dirac delta distribution δ, see [30, p. 31].

Incorporating these definitions, we get the following from (D.10):

$$\begin{aligned}
\left\langle \mathrm{D}^{(m_1,m_2)}\mu_{(f)}, \varphi \right\rangle &= \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \varphi(T_j, S_k) \\
&= \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \tau_{-T_j,-S_k}\varphi(0,0) \\
&= \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \left\langle \delta, \tau_{-T_j,-S_k}\varphi \right\rangle \\
&= \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \left\langle \tau_{T_j,S_k}\delta, \varphi \right\rangle \\
&= \left\langle \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \left(\tau_{T_j,S_k}\delta \right), \varphi \right\rangle
\end{aligned}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^2)$. Hence, the distributional derivative $\mathrm{D}^{(m_1,m_2)} \mu_{(f)}$ of the generalized function f in the case $m_1, m_2 \geq 1$ is a linear combination of translated Dirac delta distributions, which we denote by $f^{(m_1,m_2)}$, i.e.

$$f^{(m_1,m_2)} := \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \left(\tau_{T_j, S_k} \delta \right) \tag{D.16}$$

where the coefficients $c_{j,k}^{m_1,m_2}$ are defined by (D.8), (D.9), (D.11), and (D.12).

This completes the proof of Lemma 4.2. $\qquad\square$

D.3. Proof of Lemma 4.3

The statement from Lemma 4.3 provides a representation of the Fourier transform $\widehat{f}$ of the function f which is considered in (4.1). That representation is used in Theorem 4.4 in order to derive a theory for the reconstruction of such functions f from sparse Fourier data.

In the section on hand, we want to prove the claim of Lemma 4.3. This will be done by computing the Fourier transform of the generalized function $f^{(m_1,m_2)}$ which we have determined in the previous section, see (D.16). Therefore, we need the definition of the Fourier transform of a distribution. Let μ be one of the examined distributions in the previous sections. The Fourier transform $\mathcal{F}\{\mu\}$ is defined analogously as in the univariate case (see p. 148). It is the linear functional on the space $\mathcal{S}(\mathbb{R}^2)$ which fulfils

$$\langle \mathcal{F}\{\mu\}, \varphi \rangle = \langle \mu, \widehat{\varphi} \rangle \qquad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^2), \tag{D.17}$$

cf. [30, p. 47]. The notation $\widehat{\varphi}$ indicates the ordinary Fourier transform of φ. Further, $\mathcal{S}(\mathbb{R}^2)$ denotes the Schwartz space of rapidly decreasing, infinitely differentiable functions in $\mathbb{R}^2$.

Consider $\varphi \in \mathcal{S}(\mathbb{R}^2)$. Then we have

$$\langle \mathcal{F}\{f^{(m_1,m_2)}\}, \varphi \rangle \stackrel{(D.16)}{=} \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \langle \mathcal{F}\{(\tau_{T_j, S_k}\delta)\}, \varphi \rangle$$

$$\stackrel{(D.17)}{=} \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \langle \tau_{T_j, S_k}\delta, \widehat{\varphi} \rangle$$

$$\stackrel{(D.14)}{=} \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \langle \delta, \tau_{-T_j, -S_k}\widehat{\varphi} \rangle$$

$$\stackrel{(D.15)}{=} \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \tau_{-T_j, -S_k}\widehat{\varphi}(0,0)$$

$$\stackrel{\text{(D.13)}}{=} \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} \widehat{\varphi}(T_j, S_k)$$

$$= \int_{\mathbb{R}^2} \left(\sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} e^{-i \cdot (\omega_1 T_j + \omega_2 S_k)} \right) \varphi(\omega_1, \omega_2) \, d\boldsymbol{\omega}$$

$$= \left\langle \mu_{(F)}, \varphi \right\rangle \tag{D.18}$$

where the generalized function F is given by

$$F : \mathbb{R}^2 \ni (\omega_1, \omega_2)^{\mathrm{T}} \mapsto \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2} e^{-i \cdot (\omega_1 T_j + \omega_2 S_k)} \in \mathbb{C}. \tag{D.19}$$

Since F is locally integrable, $\mu_{(F)}$ is well-defined by (D.1). This means that the Fourier transform $\mathcal{F}\{f^{(m_1,m_2)}\}$ of the distribution $f^{(m_1,m_2)}$ is identified with the generalized function F.

We still need a representation of the Fourier transform $\widehat{f}$. Using the relation $f^{(m_1,m_2)} = D^{(m_1,m_2)}\mu_{(f)}$, we also have the following for $\varphi \in \mathcal{S}(\mathbb{R}^2)$:

$$\left\langle \mathcal{F}\{f^{(m_1,m_2)}\}, \varphi \right\rangle \stackrel{\text{(D.17)}}{=} \left\langle D^{(m_1,m_2)}\mu_{(f)}, \widehat{\varphi} \right\rangle = \left\langle \mu_{(f)}, (-1)^{|(m_1,m_2)|} \frac{\partial^{m_1}}{\partial \bullet_1^{m_1}} \frac{\partial^{m_2}}{\partial \bullet_2^{m_2}} \widehat{\varphi} \right\rangle. \tag{D.20}$$

The last equality is due to properties of derivatives of distributions, cf. [30, p. 46]. Moreover, the mixed partial derivative of $\widehat{\varphi}$ in (D.20) is given by

$$(-1)^{|(m_1,m_2)|} \frac{\partial^{m_1}}{\partial \omega_1^{m_1}} \frac{\partial^{m_2}}{\partial \omega_2^{m_2}} \widehat{\varphi}(\omega_1, \omega_2) = \int_{\mathbb{R}^2} (i\,x_1)^{m_1} (i\,x_2)^{m_2} \varphi(x_1, x_2) e^{-i \cdot (\omega_1 x_1 + \omega_2 x_2)} \, d\boldsymbol{x}$$

$$= \mathcal{F}\{(i\bullet_1)^{m_1}(i\bullet_2)^{m_2}\varphi(\bullet_1,\bullet_2)\}(\omega_1,\omega_2),$$

see Proposition A.5. Thus, we have

$$\left\langle \mathcal{F}\{f^{(m_1,m_2)}\}, \varphi \right\rangle = \left\langle \mu_{(f)}, \mathcal{F}\{(i\bullet_1)^{m_1}(i\bullet_2)^{m_2}\varphi(\bullet_1,\bullet_2)\} \right\rangle$$

$$= \int_{\mathbb{R}^2} f(t_1, t_2) \left(\int_{\mathbb{R}^2} (i\,x_1)^{m_1}(i\,x_2)^{m_2}\varphi(x_1, x_2) e^{-i \cdot (t_1 x_1 + t_2 x_2)} \, d\boldsymbol{x} \right) d\boldsymbol{t}.$$

Observe that $\mathbb{R}^2 \ni (x_1, x_2)^{\mathrm{T}} \mapsto x_1^{m_1} x_2^{m_2} \varphi(x_1, x_2) \in \mathbb{R}$ is also a Schwartz function by definition of $\mathcal{S}(\mathbb{R}^2)$, cf. [66, p. 211], and recall that $\mathcal{S}(\mathbb{R}^2) \subset L^p(\mathbb{R})$ for all $p \geq 1$, see [66, pp. 211-212]. Further, we have $f \in L^1(\mathbb{R}^2)$ since f is an almost everywhere continuous function with compact support. These properties ensure that the conditions

for Fubini's theorem are fulfilled such that we can exchange the order of integration in the multiple integral above. Then we get

$$
\begin{aligned}
\left\langle \mathcal{F}\{f^{(m_1,m_2)}\}, \varphi \right\rangle &= \int_{\mathbb{R}^2} f(t_1,t_2) \left(\int_{\mathbb{R}^2} (\mathrm{i}\,x_1)^{m_1} (\mathrm{i}\,x_2)^{m_2} \varphi(x_1,x_2)\, \mathrm{e}^{-\mathrm{i}\cdot(t_1 x_1 + t_2 x_2)}\, \mathrm{d}\boldsymbol{x} \right) \mathrm{d}\boldsymbol{t} \\
&= \int_{\mathbb{R}^2} (\mathrm{i}\,x_1)^{m_1} (\mathrm{i}\,x_2)^{m_2} \varphi(x_1,x_2) \left(\int_{\mathbb{R}^2} f(t_1,t_2)\, \mathrm{e}^{-\mathrm{i}\cdot(t_1 x_1 + t_2 x_2)}\, \mathrm{d}\boldsymbol{t} \right) \mathrm{d}\boldsymbol{x} \\
&= \int_{\mathbb{R}^2} (\mathrm{i}\,x_1)^{m_1} (\mathrm{i}\,x_2)^{m_2} \widehat{f}(x_1,x_2) \varphi(x_1,x_2)\, \mathrm{d}\boldsymbol{x} \\
&= \left\langle \mu_{((\mathrm{i}\cdot_1)^{m_1}(\mathrm{i}\cdot_2)^{m_2}\widehat{f})}, \varphi \right\rangle.
\end{aligned}
\tag{D.21}
$$

From (D.18) and (D.21) we deduce

$$
\left\langle \mu_{(F)}, \varphi \right\rangle = \left\langle \mu_{((\mathrm{i}\cdot_1)^{m_1}(\mathrm{i}\cdot_2)^{m_2}\widehat{f})}, \varphi \right\rangle \quad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^2),
$$

where the generalized function F is defined in (D.19). Hence, the Fourier transform $\widehat{f}$ of the function f considered in (4.1) can be represented by

$$
(\mathrm{i}\,\omega_1)^{m_1} (\mathrm{i}\,\omega_2)^{m_2} \widehat{f}(\omega_1,\omega_2) = \sum_{j=1}^{N_1+m_1} \sum_{k=1}^{N_2+m_2} c_{j,k}^{m_1,m_2}\, \mathrm{e}^{-\mathrm{i}\cdot(\omega_1 T_j + \omega_2 S_k)}, \quad (\omega_1,\omega_2)^{\mathrm{T}} \in \mathbb{R}^2.
$$

The coefficients $c_{j,k}^{m_1,m_2}$ are defined accordingly to (D.8), (D.9), (D.11), and (D.12).

Thus, the proof of Lemma 4.3 is completed. In the computations above, we have used some definitions and equations from the previous sections. Note that they also hold for $\varphi \in \mathcal{S}(\mathbb{R}^2)$. $\qquad\square$

Bibliography

[1] M. Aigner. *Combinatorial Theory*. Number 234 in Grundlehren der mathematischen Wissenschaften. Springer, New York, 1979.

[2] R. Badeau, B. David, and G. Richard. High-Resolution Spectral Analysis of Mixtures of Complex Exponentials Modulated by Polynomials. *IEEE Transactions on Signal Processing*, 54(4):1341–1350, 2006.

[3] D. Batenkov. Decimated generalized Prony systems. arXiv preprint, `http://arxiv.org/abs/1308.0753v1`, August 2013.

[4] D. Batenkov. Complete Algebraic Reconstruction of Piecewise-Smooth Functions from Fourier Data. arXiv preprint, `http://arxiv.org/abs/1211.0680v2`, March 2014. *Mathematics of Computation* (to appear).

[5] D. Batenkov, N. Sarig, and Y. Yomdin. Accuracy of Algebraic Fourier Reconstruction for Shifts of Several Signals. arXiv preprint, `http://arxiv.org/abs/1311.3468v2`, April 2014. *Special Issue of STSIP, 2014* (to appear).

[6] D. Batenkov and Y. Yomdin. Algebraic Fourier reconstruction of piecewise-smooth functions. *Mathematics of Computation*, 81(277):277–318, 2012.

[7] D. Batenkov and Y. Yomdin. On the accuracy of solving confluent Prony systems. *SIAM Journal on Applied Mathematics*, 73(1):134–154, 2013.

[8] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. *Applied and Computational Harmonic Analysis*, 19:17–48, 2005.

[9] R. N. Bracewell. *The Fourier Transform and its Applications*. McGraw-Hill, New York, 3rd edition, 2000.

[10] D. Braess. *Nonlinear approximation theory*. Springer, Berlin, 1986.

[11] K. Bredies and D. Lorenz. *Mathematische Bildverarbeitung*. Vieweg+Teubner, Wiesbaden, 2011.

[12] M. Buhmann and A. Pinkus. On a Recovery Problem. *Annals of Numerical Mathematics*, 4:129–142, 1997.

[13] E. J. Candès and C. Fernandez-Granda. Towards a Mathematical Theory of Super-Resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.

[14] E. J. Candès, J. Romberg, and T. Tao. Robust Uncertainty Principle: Exact Signal Reconstruction From Highly Incomplete Frequency Information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[15] I. Daubechies. *Ten Lectures on Wavelets*. Number 61 in CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, 4th printing, 1995.

[16] C. de Boor. *A Practical Guide to Splines*. Springer, New York, Revised edition, 2001.

[17] G. C. F. M. R. de Prony. Essai expérimental et analytique sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures. *Journal de l'École Polytechnique*, 1:24–76, 1795. Quoted from [48].

[18] D. L. Donoho. Compressed Sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[19] P. L. Dragotti, M. Vetterli, and T. Blu. Sampling Moments and Reconstructing Signals of Finite Rate of Innovation: Shannon Meets Strang-Fix. *IEEE Transactions on Signal Processing*, 55(5):1741–1757, 2007.

[20] S. Durocher. Graph-Theoretic and Geometric Algorithms Associated with Moment-based Polygon Reconstruction. Master's thesis, Department of Computer Science, University of British Columbia, 1999.

[21] M. Elad, P. Milanfar, and G. H. Golub. Shape from Moments — An Estimation Theory Perspective. *IEEE Transactions on Signal Processing*, 52(7):1814–1829, 2004.

[22] J. Elstrodt. *Maß- und Integrationstheorie*. Springer, Berlin, 5th edition, 2007.

[23] F. Filbir, H. N. Mhaskar, and J. Prestin. On the Problem of Parameter Estimation in Exponential Sums. *Constructive Approximation*, 35(3):323–343, 2012.

[24] G. Fischer. *Lernbuch Lineare Algebra und Analytische Geometrie*. Springer Spektrum, Wiesbaden, 2nd edition, 2012.

[25] O. Forster. *Analysis 3*. Springer Spektrum, Wiesbaden, 7th edition, 2012.

[26] I. M. Gel'fand and G. E. Shilov. *Generalized functions – Volume I – Properties and Operations*. Academic Press, New York, 2nd printing, 1966.

[27] G. H. Golub, P. Milanfar, and J. Varah. A stable numerical method for inverting shape from moments. *SIAM Journal on Scientific Computing*, 21(4):1222–1243, 1999.

[28] N. Gravin, J. Lasserre, D. V. Pasechnik, and S. Robins. The Inverse Moment Problem for Convex Polytopes. *Discrete & Computational Geometry*, 48(3):596–621, 2012.

[29] B. Gustafsson, C. He, P. Milanfar, and M. Putinar. Reconstructing planar domains from their moments. *Inverse Problems*, 16(4):1053–1070, 2000.

[30] D. D. Haroske and H. Triebel. *Distributions, Sobolev Spaces, Elliptic Equations.* European Mathematical Society, Zürich, 2008.

[31] H. Heuser. *Lehrbuch der Analysis – Teil 1.* B. G. Teubner, Stuttgart, 15th edition, 2003.

[32] H. Heuser. *Lehrbuch der Analysis – Teil 2.* B. G. Teubner, Wiesbaden, 13th edition, 2004.

[33] F. B. Hildebrand. *Introduction to numerical analysis.* Dover Publications, New York, 2nd edition, 1987. Unabridged, slightly corrected republication.

[34] R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, 1996.

[35] R. F. Hoskins and J. S. Pinto. *Distributions, ultradistributions and other generalized functions.* Ellis Horwood, New York, 1994.

[36] Y. Hua and T. K. Sarkar. Matrix Pencil Method for Estimatin Parameter of Exponentially Damped/Undamped Sinusoids in Noise. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(5):814–824, 1990.

[37] A. Iske. *Charakterisierung bedingt positiv definiter Funktionen für multivariate Interpolationsmethoden mit radialen Basisfunktionen.* PhD thesis, University of Göttingen, 1994.

[38] H. Klein. Mathematik für Physiker IV, §6 Fourierreihen und Fouriertransformation, pp. 15–16. Online, http://www.math.uni-kiel.de/geometrie/klein/physs13/mo1006.pdf, lecture notes, summer semester 2013. Accessed on 2 April 2014.

[39] P. Knabner and W. Barth. *Lineare Algebra: Grundlagen und Anwendungen.* Springer Spektrum, Berlin, 2013.

[40] D. E. Knuth. *The Art of Computer Programming / Volume 1 / Fundamental Algorithms.* Addison-Wesley, Reading, 3rd edition, 1997.

[41] J. Komrska. Simple derivation of formulas for Fraunhofer diffraction at polygonal apertures. *J. Opt. Soc. Am.*, 72(10):1382–1384, 1982.

[42] C. Lanczos. *Applied Analysis.* Sir Isaac Pitman & Sons, London, 4th printing, 1967.

[43] J. Liesen and V. Mehrmann. *Lineare Algebra.* Vieweg+Teubner, Wiesbaden, 2011.

[44] I. Maravić and M. Vetterli. Exact Sampling Results for Some Classes of Parametric Nonbandlimited 2-D Signals. *IEEE Transactions on Signal Processing*, 52(1):175–189, 2004.

[45] H. N. Mhaskar and J. Prestin. On the detection of singularities of a periodic function. *Advances in Computational Mathematics*, 12(2–3):95–131, 2000.

[46] H. N. Mhaskar and J. Prestin. On Local Smoothness Classes of Periodic Functions. *Journal of Fourier Analysis and Applications*, 11(3):353–373, 2005.

[47] P. Milanfar, G. C. Verghese, W. C. Karl, and A. S. Willsky. Reconstructing Polygons from Moments with Connections to Array Processing. *IEEE Transactions on Signal Processing*, 43(2):432–443, 1995.

[48] T. Peter. *Generalized Prony Method.* Der Andere Verlag, Uelvesbüll, 2014.

[49] T. Peter and G. Plonka. A Generalized Prony Method for Reconstruction of Sparse Sums of Eigenfunctions of Linear Operators. *Inverse Problems*, 29(2):(025001), 2013.

[50] T. Peter, G. Plonka, and D. Roşca. Representation of Sparse Legendre Expansions. *Journal of Symbolic Computation*, 50:159–169, 2013.

[51] T. Peter, D. Potts, and M. Tasche. Nonlinear approximation by sums of exponentials and translates. *SIAM Journal on Scientific Computing*, 33(4):1920–1947, 2011.

[52] M. A. Pinsky. *Introduction to Fourier Analysis and Wavelets.* Number 102 in Graduate Studies in Mathematics. American Mathematical Society, Providence, 2009.

[53] G. Plonka and M. Tasche. Prony Methods for Recovery of Structured Functions. *GAMM-Mitteilungen*, 37(2):239–258, 2014.

[54] G. Plonka and M. Wischerhoff. How many fourier samples are needed for real function reconstruction? *Journal of Applied Mathematics and Computing*, 42(1–2):117–137, 2013.

[55] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Processing*, 90(5):1631–1642, 2010.

[56] D. Potts and M. Tasche. Nonlinear approximation by sums of nonincreasing exponentials. *Appl. Anal.*, 90:609–626, 2011.

[57] D. Potts and M. Tasche. Parameter estimation for multivariate exponential sums. *Electronic Transactions on Numerical Analysis*, 40:204–224, 2013.

[58] D. Potts and M. Tasche. Parameter estimation for nonincreasing exponential sums by prony-like methods. *Linear Algebra and its Applications*, 439(4):1024–1039, 2013.

[59] H. Rauhut. Random sampling of sparse trigonometric polynomials. *Applied and Computational Harmonic Analysis*, 22(1):16–42, 2007.

[60] A. Rényi. On projections of probability distributions. *Acta Mathematica Academiae Scientiarum Hungarica*, 3(3):131–142, 1952.

[61] R. Roy and T. Kailath. ESPRIT — Estimation of Signal Parameters Via Rotational Invariance Techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(7):984–995, 1989.

[62] R. O. Schmidt. Multiple Emitter Location and Signal Parameter Estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

[63] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, Princeton, 1971.

[64] F. Vanpoucke, M. Moonen, and Y. Berthoumieu. An efficient subspace algorithm for 2-D harmonic retrieval. *Proceedings of the 1994 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 4:461–464, 1994.

[65] M. Vetterli, P. Marziliano, and T. Blu. Sampling Signals With Finite Rate of Innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002.

[66] D. Werner. *Funktionalanalysis*. Springer, Berlin, 6th edition, 2007.

[67] M. Wischerhoff. Real function reconstruction from sparse fourier samples. *Proceedings in Applied Mathematics and Mechanics*, 13(1):491–492, 2013.

[68] E. Zeidler, editor. *Springer-Taschenbuch der Mathematik*. Springer Spektrum, Wiesbaden, 3rd edition, 2013.

Curriculum Vitae

DIPL.-MATH. MARIUS WISCHERHOFF

Address	Institute for Numerical and Applied Mathematics
	University of Göttingen
	Lotzestr. 16-18
	37083 Göttingen, Germany
Telephone	+495513920075
Email	m.wischerhoff@math.uni-goettingen.de
Homepage	http://na.math.uni-goettingen.de/index.php?section=wischerhoff

Personal Details

Gender	Male
Date of birth	3 April 1985
Place of birth	Duisburg, Germany

Academic Education

10/2010–01/2015	Doctoral student of mathematics at the University of Göttingen. Advisor: Prof. Dr. G. Plonka-Hoch.
10/2005–07/2010	Diploma student of mathematics at the University of Duisburg-Essen, with business studies as minor subject.

Internship

11/2009–01/2010	Internship in the area of value-risk-management at AXA Service AG, branch Life Controlling in Cologne (Germany).

Alternative civilian service

08/2004–05/2005	Employed at the 'Duisburger Werkstatt für Menschen mit Behinderung gGmbH' in Duisburg.

Education

2001–2004	Secondary school 'Albert-Einstein-Gymnasium' in Duisburg.
1995–2001	Secondary school 'Realschule Rheinhausen' in Duisburg.

Research Experience

11/2011–12/2013	Associated member of the 'Research Training Group 1023: Identification in Mathematical Models' at the Institute for Numerical and Applied Mathematics, University of Göttingen.
Since 10/2010	Scientific assistant and doctoral student at the Institute for Numerical and Applied Mathematics, University of Göttingen.

Teaching Experience

2014/2015	*Image and Geometry Processing I*, winter term, lecture assistance and tutorial.
2014	*Applied Mathematics for Teacher Trainees*, summer term, lecture assistance and tutorial.
2013/2014	*Convex Analysis and Applications*, winter term, lecture assistance and tutorial.
2013	*Mathematics for Computer Science students II*, summer term, lecture assistance and tutorial.
2012/2013	*Mathematics for Computer Science students I*, winter term, lecture assistance and tutorial.
2012	*Mathematics for Computer Science students II*, summer term, lecture assistance and tutorial.
2011/2012	*Mathematics for Computer Science students I*, winter term, lecture assistance and tutorial.
2011	*Applied Mathematics for Teacher Trainees*, summer term, lecture assistance and tutorial.
2010/2011	*Introduction to Signal and Image Processing: application-oriented Fourier analysis*, winter term, lecture assistance and tutorial.

Publications

[1] M. Wischerhoff. Behandlung von Poisson- und Gamma-Störungen in digitalen Bildern mit Hilfe von Splitting-Methoden. Diploma thesis. Faculty of Mathematics, University of Duisburg-Essen, July 2010.

[2] G. Plonka and M. Wischerhoff. How many Fourier samples are needed for real function reconstruction? *Journal of Applied Mathematics and Computing*, 42(1–2):117–137, 2013.

[3] M. Wischerhoff. Real function reconstruction from sparse Fourier samples. *Proceedings in Applied Mathematics and Mechanics*, 13(1):491—492, 2013.

Participation in Conferences and Workshops

2014	Workshop: Mathematical Signal Processing and Phase Retrieval, 1–3 September in Gleichen-Reinhausen.
	5th International Conference on Computational Harmonic Analysis (talk), 19–23 May in Nashville, Tennessee (USA).
	24th Rhein-Ruhr-Workshop, 31 January to 1 February in Bestwig.
	First French-German Mathematical Image Analysis Conference, 13–15 January in Paris (France).
2013	Final Colloquium of the Research Training Group 1023 (poster), 29–30 November in Göttingen.
	Workshop on Statistical Issues in Compressive Sensing, 11–13 November in Göttingen.
	Conference: 84th Annual Meeting of the International Association of Applied Mathematics and Mechanics (talk), 18–22 March in Novi Sad (Serbia).
	23rd Rhein-Ruhr-Workshop (poster), 1–2 February in Bestwig.
2012	8th Annual Workshop of the Research Training Group 1023 (talk), 17–19 October in Goslar.
	Workshop on Advances in Mathematical Image Processing (talk and poster), 4–6 September in Gleichen-Reinhausen.
	Summer School: Image Reconstruction: Mathematics & Applications (talk), 23–27 July in Munich.
	22nd Rhein-Ruhr-Workshop (talk), 3–4 February in Bestwig.
2011	21st Rhein-Ruhr-Workshop, 4–5 February in Königswinter.