The Generalized Operator Based Prony Method

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
“Doctor rerum naturalium”
der Georg-August-Universität zu Göttingen

im Promotionsprogramm
“PhD School of Mathematical Sciences (SMS)”
der Georg-August University School of Science (GAUSS)

vorgelegt von
Kilian Stampfer
aus Kelheim

Göttingen, 2018

Betreuungsausschuss:

Prof. Dr. Stephan Huckemann
Institut für Mathematische Stochastik, Universität Göttingen

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

Mitglieder der Prüfungskommission:

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

Korreferent:
Prof. Dr. Stefan Kunis
Institut für Mathematik, Universität Osnabrück

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Stephan Huckemann
Institut für Mathematische Stochastik, Universität Göttingen

Prof. Dr. Gert Lube
Institut für Numerische und Angewandte Mathematik, Universität Göttingen

Prof. Dr. Russel Luke
Institut für Numerische und Angewandte Mathematik, Universität Göttingen

Prof. Dr. Dominic Schuhmacher
Institut für Mathematische Stochastik, Universität Göttingen

Tag der mündlichen Prüfung: 17.01.2019

Preface

In the last three years I had the opportunity to work at the Institute for Mathematical Stochastics at the University of Göttingen, an excellent place for research and learning. I have found colleagues, always ready for discussion about new ideas, which also became friends and learned much about the variety of mathematical research in stochastics as well as in numerics.
First of all, I would like to thank the German Research Foundation, that financed my position in the Research Training Group 2088. A special thank goes to my advisors Prof. Dr. Plonka-Hoch and Prof. Dr. Huckemann. Prof. Plonka-Hoch was always there for vivid discussion of new ideas and beyond and Prof. Huckemann gave me the freedom of doing things my own way.
Last but not least, I would like to thank my wife Maria, who convinced me to go this way, reinforced me when I was struggling and finally gave me serenity in the last few weeks. She was, is and will always be my strength.

Contents

List of Symbols

$\mathbb{N}$	set of positive integers
$\mathbb{N}_0$	set of non-negative integers
$\mathbb{Z}$	set of all integers
$\mathbb{R}$	set of real numbers
$\mathbb{R}_{\geq 0}$	set of non-negative real numbers
$\mathbb{R}_{\neq 0}$	set of positive real numbers without zero
$\overline{\mathbb{R}}$	all real numbers and $-\infty$ and ∞
$\mathbb{C}$	complex numbers
$\mathbb{C}^-$	complex numbers without the negative half axe
$\mathbb{C}_{\neq 0}$	complex numbers without zero
$\operatorname{rank}(X)$	rank of a matrix X
$\operatorname{span}(S)$	linear span of a set of column vectors S
$\lvert S \rvert$	cardinality of a set S
$x^\top$	transposed of a vector or matrix x
S^c	complement of a set S
I	identity operator
A	generator: Definition 2.2.2
Φ_φ	iteration operator: Definition 2.2.4
$\hat{\sigma}_P(A)$	subset of the point spectrum of A such that $\lambda \to v_\lambda$ is injective: Definition 2.2.2
$\hat{\sigma}_P^\varphi(A)$	subset of $\hat{\sigma}_P(A)$ such that $\lambda \to \varphi(\lambda)$ is injective: Definition 2.2.4
$P_\Lambda^\varphi(z)$	Prony polynomial: Definition 2.2.6
$P_\Lambda^\varphi(\Phi_\varphi)$	annihilation operator: Definition 2.2.6
$\mathcal{M}(A)$	signal space: Definition 2.2.2
$\mathcal{M}^\varphi(A)$	restricted signal space: Definition 2.2.4
$\mathcal{M}_M^\varphi(A)$	restricted signal model of all linear combination of length M
F_m	evaluation scheme: Definition 2.2.7
$S_{m,\ell}$	sampling scheme: Definition 2.2.11
$\mathcal{X}_{N,M}$	sampling matrix: Definition 2.2.12

CHAPTER 1

Introduction

> A problem well-stated is
> half-solved.
>
> Charles F. Kettering

In many fields, for example, classical mechanics and high-energy physics [30, 14], biological signal processing [28], as well as medical [20], electrical [25, 49], and mechanical engineering [41] , it is often necessary to decompose a complicated continuous time signal into a linear combination of finitely many pieces. Those pieces, often called atoms, do not necessarily possess basis or frame properties, but need to have a clear interpretation in the considered setting in the first place. Thereby and because in many cases only the parametric structures of the superimposed atoms are known, the problem at hand is severely non-linear. Over the years, several approaches were employed to solve this challenging task. The applied methods range from classical non-linear least squares over greedy algorithms [7] up to the very recent atomic norm minimization approach [6, 3].

The maybe most typical type of atoms used are complex exponential ones, because of their interpretation as pure frequency components. A classical approach to solve this non-linear reconstruction problem is the Prony method [13]. It is well suited for the exact reconstruction of a linear combination of complex exponentials, where the linear factors as well as the exponential frequencies are allowed to be complex; in particular, damped sinusoids are possible. For this and several other reasons, during the last years the method has experienced a notable renaissance in the signal processing community and beyond. Nonetheless, while most efforts during this time were made to stabilize the numerical algorithms in the presence of noise and to modify them for multivariate settings, only a few publications dealt with generalizations towards other types of atoms besides complex exponentials.

This chapter first wants to introduce the classical Prony method in the usual form and give a short history of further improvements. Afterwards it is shown how Peter & Plonka [32] reached a first generalization for more generic atomic expansions and how their work motivated further research and has finally led to the results in this thesis.

Although the major purpose of this work lies in the development of a new operator based approach to Prony's method, the classical derivation is chosen as a starting point in order to emphasize the differences and motivate the successive generalizations step by step.

1.1 The Classical Prony Method

The classical Prony method [13] was originally developed to reconstruct finite linear combinations of complex exponentials $f : \mathbb{R} \to \mathbb{C}$ such that

$$f(x) := \sum_{j=1}^{M} c_j \exp(\lambda_j x), \tag{1.1}$$

where it is assumed that $\lambda_j \in \mathbb{R} + i[0, 2\pi)$ and $c_j \in \mathbb{C}_{\neq 0}$. Associated with this parametric signal model is the so-called Prony polynomial

$$P(z) := \prod_{j=1}^{M} (z - z_j) = \sum_{k=0}^{M} p_k z^k, \tag{1.2}$$

where $z_j := \exp(\lambda_j)$ and $p_M = 1$ by definition of the polynomial. If the parametric function in (1.1) is seen as a solution of a difference equation, the Prony polynomial is identical with the characteristic polynomial of the corresponding equation. This can be easily seen by the following calculation.

$$\begin{aligned} \forall m \in \mathbb{N}_0 : \sum_{k=0}^{M} p_k f(m+k) &= \sum_{k=0}^{M} p_k \sum_{j=1}^{M} c_j \exp(\lambda_j (m+k)) = \sum_{j=1}^{M} \sum_{k=0}^{M} p_k c_j z_j^k z_j^m \\ &= \sum_{j=1}^{M} c_j z_j^m \sum_{k=0}^{M} p_k z_j^k = \sum_{j=1}^{M} c_j z_j^m P(z_j) = 0 \end{aligned} \tag{1.3}$$

Hence, for f holds that

$$\forall m \in \mathbb{N}_0 : \sum_{k=0}^{M-1} p_k f(m+k) = -f(m+M).$$

The key property in this derivation is the reproduction of the Prony polynomial inside the sum and its specific definition, since it vanishes on all z_j and only there. Based on this set of equations, the minimal linear system

$$\check{X} p := \begin{bmatrix} f(0) & f(1) & \dots & f(M-1) \\ f(1) & f(2) & \dots & f(M) \\ \vdots & \vdots & & \vdots \\ f(M-1) & f(M) & \dots & f(2M-2) \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{M-1} \end{bmatrix} = - \begin{bmatrix} f(M) \\ f(M+1) \\ \vdots \\ f(2M-1) \end{bmatrix} \tag{1.4}$$

can be constructed to solve for the coefficients $p := [p_k]_{k=0}^{M-1}$, and $p_M := 1$. Indeed, this system has a unique solution, since the matrix $\check{\mathcal{X}}$ can be diagonalized by a Vandermonde matrix.

$$\check{\mathcal{X}} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_M \\ z_1^2 & z_2^2 & \dots & z_M^2 \\ \vdots & \vdots & \dots & \vdots \\ z_1^{M-1} & z_2^{M-1} & \dots & z_M^{M-1} \end{bmatrix} \begin{bmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & c_{M-1} & 0 \\ 0 & 0 & \dots & 0 & c_M \end{bmatrix} \begin{bmatrix} 1 & z_1 & z_1^2 & \dots & z_1^{M-1} \\ 1 & z_2 & z_2^2 & \dots & z_2^{M-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & z_{M-1} & z_{M-1}^2 & \dots & z_{M-1}^{M-1} \\ 1 & z_M & z_M^2 & \dots & z_M^{M-1} \end{bmatrix} \tag{1.5}$$

This implies that the matrix $\check{\mathcal{X}}$ itself has full rank because $z_j \neq z_i$ for all $j \neq i$, which is sufficient for non-singularity of both Vandermonde matrices, and all $c_j \neq 0$ by definition.

Once the coefficients of the Prony polynomial are known, the zeros z_j can be calculated by any root finding algorithm. Since we assumed that $\mathrm{Im}\left(\lambda_j\right) \in [0, 2\pi)$, the actual parameters λ_j are uniquely defined by the zeros $z_j = \exp(\lambda_j)$ and their complex logarithm $\log(z_j) = \lambda_j$. In consequence, the active exponential atoms in (1.1) are no longer unknown and the linear system

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_M \\ \vdots & \vdots & & \vdots \\ z_1^M & z_2^M & \dots & z_M^M \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_M \end{bmatrix} = \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(M) \end{bmatrix} \tag{1.6}$$

can be constructed by a subset of the samples used in (1.4). Here the only unknowns are the linear coefficients c_j. Eventually, having solved this equation for the linear coefficients c_j, all parameters of the signal in (1.1) are exactly reconstructed from only $2M$ function values.

This is the classical Prony method as developed from Gaspard Clair Francois Marie Riche de Prony [13] in the 18$^{\text{th}}$ century. Despite the rich theory, the method itself was slightly abandoned due to the often involved large linear systems until the dawn of the computer age in the 1960s and 70s.

Prony's method is also known under the term annihilation filter method, in particular in the field of electrical engineering. Still, in this case the formulation is sometimes slightly different and based on the following equation system

$$\check{\mathcal{X}} p := \begin{bmatrix} f(-2M+1) & f(-2M+2) & \dots & f(-M) & f(-M+1) \\ f(-2M+2) & f(-2M+3) & \dots & f(-M+1) & f(-M+2) \\ \vdots & \vdots & & \vdots & \vdots \\ f(-M) & f(-M+1) & \dots & f(-1) & f(0) \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_M \end{bmatrix} = 0, \tag{1.7}$$

where the polynomial coefficients are interpreted as filter coefficients and the whole sampling scheme is shifted in time by $-2M+1$. It is obviously equivalent to (1.4) if we choose $p_M := 1$ and rename the sampling accordingly. With basically the same arguments it is shown that this system has a kernel of dimension one and therefore the polynomial coefficients are unique up to

some multiplicative constant; for more details the interested reader is referred to [45]. Instead of the normalization $p_M = 1$, more often $\|p\|_2 = 1$ is used to fix a single vector in the kernel.
Beyond that, it was long known that the Prony method in its very basic version performs quite poorly in the presence of noise, basically due to the instability of the involved root finding algorithms and the potentially badly conditioned Vandermonde matrix. The first one who proposed a more sophisticated error model to cope with the noise issues was Pisarenko [34] in 1973. He used the well-known covariance structure of the signal model to establish a set of noise reduced samples; more details on covariance based refinements of the Prony method can be found in [17]. Shortly after Pisarenko's results, but with a more heuristical approach, the famous MUSIC algorithm [42] followed, developed by Schmidt in 1986. In the same year a first maximum likelihood method was proposed by Bresler & Macovski [5]. Once started, a variety of new methods was developed, also a more geometrical approach by Roy and Kailath, called the ESPRIT algorithm [40] in 1990, followed by the matrix pencil method [19] by Hua & Sarkar one year later. In 1995 Osborne & Smyth [31] again used a more statistical point of view to establish a theoretically justified error model. They incorporated generalized least squares techniques to come up with an iterative maximum likelihood algorithm.
Although Prony-like methods were long established in signal processing since the first steps in the 18^{th} century, a real renaissance was induced by the invention of the so-called finite-rate-of-innovation (FRI) principle in 2002 by Vetterli et al. [47]. The concepts and the derived algorithms break the band limitations of the classical Whittaker-Kotelnikov-Shannon-sampling theorem [44, 48] by using a locally parameterized signal model instead of the usual Hilbert space approach. Basically, if the parameterization of the signal has only a finite number of free parameters per unit of time, it is said to have a finite rate of innovation. The associated algorithm is applicable if and only if the signal at hand can be locally reduced to the parametric form (1.1) in a finite number of preprocessing steps and in turn be essentially reconstructed by means of the classical Prony method or one of its offsprings.
Nevertheless, this new field of applications amplified the interest in the Prony method significantly; one of the more recent results is the approximate Prony method [37] by Potts & Tasche from 2010. Despite these vast improvements of the classical method in the presence of noise and even some multivariate extensions in recent years as for example by Potts et al., Kunis et al., and Cuyt et al. in [38, 23, 9], no systematic generalization for different types of atoms took place. A first approach in this direction was the Generalized Prony Method [32] by Peter & Plonka in 2013.

1.2 The Generalized Prony Method

Based on the classical Prony method and the knowledge of certain special applications of it, for example, sparse expansions of Chebyshev polynomials of arbitrary degree [39], Peter & Plonka developed the so-called Generalized Prony Method [32]. They directly reformulated the classical approach as seen in Section 1.1 in operator terms and applied the results to several groups of linear operators including scaling operators and those of Sturm-Liouville type. With

this approach they basically solved the problem of reconstructing more general expansions of the form

$$f(x) = \sum_{j=1}^{M} c_j v_j(x), \tag{1.8}$$

where v_j are eigenfunctions of certain linear operators and $c_j \in \mathbb{C}_{\neq 0}$ are constant complex coefficients, both unknown. They were able to show that the classical Prony method as well as the reconstruction algorithm for sparse expansions into Chebychev polynomials in [39] can be derived by this generalization and further examples were given.
The first step to the Generalized Prony Method as derived in [32] is the observation that the exponential functions $v_j(x) := \exp(\lambda_j x)$ are eigenfunctions of the shift operator S_1, i. e.,

$$S_1 v_j(x) := S_1 \exp(\lambda_j x) = \exp(\lambda_j(x+1)) = \exp(\lambda_j)\exp(\lambda_j x) = z_j v_j(x),$$

with eigenvalues $z_j := \exp(\lambda_j)$. Now we take equally spaced samples based on the point evaluation functional $Ff := f(0)$. We know from the classical Prony method that $2M$ samples of the form

$$F\left(S_1^k f\right) = F\left(f(\cdot + k)\right) = f(k)\,, \qquad k \in \{0, \ldots, 2M-1\}$$

are sufficient for the reconstruction of any finite linear combination of complex exponentials if the length M of the expansion (1.1) into exponential atoms $\exp(\lambda_j x)$ is known. This observation provokes the question whether this also works for more general expansions as in (1.8) and corresponding samples

$$F\left(A^k f\right),$$

where $k \in \{0, .., 2M-1\}$ and A is a suitable linear operator. As shown in [32], this is actually possible in the following way. Let

$$A : V \to V$$

be a linear operator on a vector space V with non-empty point spectrum, and in turn a non-empty family of eigenfunctions exists such that

$$Av_j = z_j v_j,$$

where z_j are the pairwise different eigenvalues uniquely identifying the eigenfunctions v_j. In other words, the mapping $\lambda_j \to v_j$ has to be injective up to some scaling factor $z \in \mathbb{C}_{\neq 0}$, i. e., we have simple eigenvalues. This directly suggests that we define the corresponding Prony polynomial similarly as before, but with the difference that we use the eigenvalues of the operator A as zeros, that are active in the expansion (1.8).

$$P(z) := \prod_{j=1}^{M}(z - z_j) =: \sum_{\ell=0}^{M} p_k z^\ell.$$

Following these definitions, basically the same calculation as in (1.3) can be done for functions of the form (1.8), with the major difference that we additionally use an almost arbitrary linear functional F for the evaluation,

$$\begin{aligned}\forall m \in \mathbb{N}_0 : \sum_{\ell=0}^{M} p_\ell F(A^{m+\ell} f) &= \sum_{\ell=0}^{M} p_\ell F\left(A^{m+\ell} \sum_{j=1}^{M} c_j v_j\right) = \sum_{j=1}^{M} \sum_{k=0}^{M} p_\ell c_j z_j^m z_j^\ell F(v_j) \\ &= \sum_{j=1}^{M} c_j z_j^m F(v_j) \sum_{\ell=0}^{M} p_\ell z_j^\ell = \sum_{j=1}^{M} \tilde{c}_j z_j^m P(z_j) = 0,\end{aligned} \tag{1.9}$$

where $\tilde{c}_j := c_j F(v_j)$. This again gives us a finite linear system to calculate the polynomial coefficients p_k, where $p_M = 1$ already holds by definition of the polynomial. We obtain the system

$$\check{\mathcal{X}} p := \begin{bmatrix} F\left(A^0 f\right) & F\left(Af\right) & \dots & F\left(A^{M-1} f\right) \\ F\left(A^1 f\right) & F\left(A^2 f\right) & \dots & F\left(A^M f\right) \\ \vdots & \vdots & & \vdots \\ F\left(A^{M-1} f\right) & F\left(A^M f\right) & \dots & F\left(A^{2M-2} f\right) \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{M-1} \end{bmatrix} = - \begin{bmatrix} F\left(A^M f\right) \\ F\left(A^{M+1} f\right) \\ \vdots \\ F\left(A^{2M-1} f\right) \end{bmatrix}, \tag{1.10}$$

that is solved to get $p := [p_k]_{k=0}^{M-1}$. Again, the reduced sampling matrix $\check{\mathcal{X}}$ can be factorized into two Vandermonde-type matrices and a diagonal matrix with entries $\tilde{c}_j := c_j F(v_j)$.

$$\check{\mathcal{X}} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_M \\ z_1^2 & z_2^2 & \dots & z_M^2 \\ \vdots & \vdots & \dots & \vdots \\ z_1^{M-1} & z_2^{M-1} & \dots & z_M^{M-1} \end{bmatrix} \begin{bmatrix} \tilde{c}_1 & 0 & 0 & \dots & 0 \\ 0 & \tilde{c}_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \tilde{c}_{M-1} & 0 \\ 0 & 0 & \dots & 0 & \tilde{c}_M \end{bmatrix} \left[\begin{array}{cccc:c} 1 & z_1 & \dots & z_1^{M-1} & \\ 1 & z_2 & \dots & z_2^{M-1} & \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & z_{M-1} & \dots & z_{M-1}^{M-1} & \\ 1 & z_M & \dots & z_M^{M-1} & \end{array}\right] \tag{1.11}$$

It is clear by the same arguments as before that the Vandermonde matrices have full rank and the linear coefficients c_j are non-zero by assumption. If we now additionally assume that the functional $F : V \to \mathbb{C}$ does not vanish on any eigenfunction of A, i. e.,

$$\forall v_j \text{ s. t. } Av_j = \lambda_j v_j : \ F(v_j) \neq 0,$$

the sampling matrix has full rank and in turn (1.10) has a unique solution.
Once the polynomial coefficients are known and the zeros are extracted, we are able to construct the further Vandermonde system

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_M \\ \vdots & \vdots & & \vdots \\ z_1^M & z_2^M & \dots & z_M^M \end{bmatrix} \begin{bmatrix} \tilde{c}_1 \\ \tilde{c}_2 \\ \vdots \\ \tilde{c}_M \end{bmatrix} = \begin{bmatrix} F\left(A^0 f\right) \\ F\left(A^1 f\right) \\ \vdots \\ F\left(A^M f\right) \end{bmatrix}, \tag{1.12}$$

that is also invertible by the same arguments as before. Solving this linear system, the vector $\left[\tilde{c}_j\right]_{j=1}^{M}$ provides us in turn with the true parameters

$$c_j = \frac{\tilde{c}_j}{F(v_j)}.$$

Although this generalized approach was quite promising to open new fields of applications, the majority of new reconstruction schemes was based on the sampling of very high signal derivatives. For example, a Legendre polynomial expansion of length ten [33] demands the sampling of each signal derivative up to order 19 at several points in time. Therefore, it became clear that a deeper insight into the method itself is necessary in order to find a systematic way to derive sampling schemes that can be realized more easily.

Contribution of this Thesis

This leads to the need for more research, for example, in the field of operator theory; in particular, on the theory of operator semi-groups as it can be found for instance in [15]. A completely different starting point and a fully operator based reformulation of the basic Prony method are given in Section 2 to reveal the fundamental ingredients which are necessary for the generalization. Based on these, the Generalized Operator based Prony Method is introduced in Section 2.2 as one of the most general forms of this approach. First examples are sparse cosine expansions and finite linear combinations of Stieltjes-Wigert polynomials, where the first one demonstrates the embedding of an existing example and the second one gives a new one, that is fully derived by means of the Generalized Operator based Prony Method.
In Chapter 3 a whole class of expansions is introduced which already covers almost all former applications, namely, the generalized shifts in Section 3.2.1. Furthermore, by using the former results, we derive a class of signals in Section 3.3 which generalizes the cosine expansion and embeds the case of finite linear combinations of Chebychev polynomials into the Generalized Operator based Prony Method approach.
Chapter 4 discusses the data acquisation process in more details. In particular, a method to circumvent the problem of sampling derivatives as mentioned above is given in the general form of a dual sampling scheme.
It is emphasized that the major contribution of this dissertation is a theoretical framework to systematically derive new reconstruction schemes for sparse expansions into eigenfunctions of special linear operators.

CHAPTER 2

Generalized Operator based Prony Method

> Always begin with the simplest examples.
>
> David Hilbert

In this section we derive the major result of this thesis, by starting again at the basic Prony method in order to extract its most important features which are necessary for further generalizations. This time, we first regard the exponential atoms as eigenfunctions of the differential operator $\frac{d}{dx}$, and give the connection to the shift operator approach. Second, the Generalized Operator based Prony Method (GOProM) is systematically introduced with all necessary definitions and summarized in the central theorem of this dissertation.

2.1 Prony's Method: The Operator Way

The basic Prony method as already mentioned is a way to reconstruct the parameters c_j and λ_j of finite weighted sums of complex exponentials, i. e., as in (1.1)

$$f(x) := \sum_{j=1}^{M} c_j \exp(\lambda_j x).$$

First of all, we define for a fixed $M \in \mathbb{N}$

$$\mathcal{M}_M := \left\{ x \mapsto \sum_{j=1}^{M} c_j \, e^{\lambda_j x} \;\middle|\; c_j \in \mathbb{C}_{\neq 0}, \lambda_j \in \mathbb{R} + i[0, 2\pi), \forall j \neq i : \lambda_j \neq \lambda_i \right\}$$

the set of all finite linear combinations of complex exponentials of known length M. The first simple observation is that all atoms $\exp(\lambda_j x)$ are eigenfunctions of the ordinary differential operator $\frac{d}{dx}$, i. e.,

$$\frac{d}{dx} \exp(\lambda_j x) = \lambda_j \exp(\lambda_j x) \iff \left(\lambda_j I - \frac{d}{dx} \right) \exp(\lambda_j x) = 0, \tag{2.1}$$

where the equation on the right side is used to define the polynomial

$$P_\Lambda(z) := \prod_{j=1}^{M} \left(\lambda_j - z\right) = \sum_{\ell=0}^{M} p_\ell z^\ell, \tag{2.2}$$

corresponding to the expansion in (1.1) with respect to $\Lambda := \{\lambda_1, \ldots, \lambda_M\}$. Due to the fact that

$$\forall\, i, j \in \mathbb{N} \,:\, \left(\lambda_j I - \frac{d}{dx}\right)\left(\lambda_i I - \frac{d}{dx}\right) = \left(\lambda_i I - \frac{d}{dx}\right)\left(\lambda_j I - \frac{d}{dx}\right),$$

i. e., the operators in (2.1) mutually commute, and each one annihilates a specific atom in the expansion (1.1), we have for all $f \in \mathcal{M}_M$ with active eigenvalues $\lambda_j \in \Lambda$,

$$P_\Lambda\left(\frac{d}{dx}\right) f := \prod_{j=1}^{M} \left(\lambda_j I - \frac{d}{dx}\right) f = \left(\sum_{\ell=0}^{M} p_\ell \frac{d^\ell}{dx^\ell}\right) f = \sum_{\ell=0}^{M} p_\ell\, f^{(\ell)} = 0. \tag{2.3}$$

Furthermore, since this new operator $P_\Lambda\left(\frac{d}{dx}\right)$ already annihilates these functions, this relation holds also true for all higher derivatives applied additionally from the left,

$$\forall m \in \mathbb{N}_0 \,:\, \frac{d^m}{dx^m} P_\Lambda\left(\frac{d}{dx}\right) f = 0. \tag{2.4}$$

Thereby, a possibly infinite homogeneous system of differential equations can be generated. In particular, a finite system of linear equations is achieved by a projection to the complex numbers using a suitable linear functional F,

$$F : \mathcal{M}_M \to \mathbb{C} \text{ such that } \forall \lambda_j \in \mathbb{C} :\; F(\exp(\lambda_j \cdot)) \neq 0.$$

For the moment, it will be sufficient to assume that F does not vanish on the eigenfunctions of $\frac{d}{dx}$, otherwise the functional F projects to a smaller and therefore false signal model. Eventually, we get

$$\forall\, m \in \mathbb{N}_0 \;:\; F\left(\frac{d^m}{dx^m} P_\Lambda\left(\frac{d}{dx}\right) f\right) = \sum_{\ell=0}^{M} p_\ell\, F\left(f^{(m+\ell)}\right) = 0. \tag{2.5}$$

It is noted that (2.4) are equations in $\mathcal{M}_M$, the signal model, and (2.5) are scalar linear equations over $\mathbb{C}$. The idea is now to identify the unknown atoms by finding the parameters λ_j, which are exactly the zeros of the polynomial in (2.2), and subsequently solve a Vandermonde system to identify the linear coefficients c_j in (1.1). Let for convenience the functional $F : \mathcal{M}_M \to \mathbb{C}$ be the point evaluation functional in zero, i. e., the samples are of the form

$$F\left(\frac{d^{m+\ell}}{dx^{m+\ell}} f\right) := f^{(m+\ell)}(0),$$

but keep in mind that any functional non-vanishing on the eigenfunctions could be used. Obviously, this point evaluation F does not vanish on any exponential because $F\left(\exp(\lambda_j \cdot)\right) = 1$.

Based on (2.5), the Hankel system

$$\mathcal{X}p = \begin{bmatrix} f^{(0)}(0) & f^{(1)}(0) & \dots & f^{(M)}(0) \\ f^{(1)}(0) & f^{(2)}(0) & \dots & f^{(M+1)}(0) \\ \vdots & \vdots & & \vdots \\ f^{(M-1)}(0) & f^{(M)}(0) & \dots & f^{(2M-1)}(0) \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_M \end{bmatrix} = 0, \tag{2.6}$$

where $p := [p_\ell]_{\ell=0}^{M}$, can be constructed from $2M-1$ derivates evaluated at zero, where the index m in (2.5) runs along the rows and ℓ along the columns. At this point it becomes clear that the annihilation formulation, as mentioned before in the basic Prony method, is slightly more general as the formulation in (1.4), since one can choose any proper norm to select a fixed coefficient vector p by the side condition $\|p\| = 1$.

The only important property for unique identification is that the kernel of the sampling matrix $\mathcal{X}$ has dimension one. This holds because it is possible (analogous to (1.11)) to factorize $\mathcal{X}$ into two Vandermonde-type matrices and a diagonal matrix,

$$\mathcal{X} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_M \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_M^2 \\ \vdots & \vdots & \dots & \vdots \\ \lambda_1^{M-1} & \lambda_2^{M-1} & \dots & \lambda_M^{M-1} \end{bmatrix} \begin{bmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & c_{M-1} & 0 \\ 0 & 0 & \dots & 0 & c_M \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^M \\ 1 & \lambda_2 & \dots & \lambda_2^M \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \lambda_{M-1} & \dots & \lambda_{M-1}^M \\ 1 & \lambda_M & \dots & \lambda_M^M \end{bmatrix}. \tag{2.7}$$

The quadratic Vandermonde-type matrix on the left has full rank since all λ_j are mutually distinct. The diagonal matrix is non-singular due to $c_j \neq 0$ by definition of the signal model and the rectangular matrix on the right has exactly rank M since it is only the transposed version of the left matrix if we delete the last column. This implies that the kernel of $\mathcal{X}$ has exactly dimension one and a closer look reveals that the multiplication with p reproduces the characteristic polynomial at the values λ_j and annihilates in turn the whole system. Therefore, the polynomial coefficients are uniquely determined up to some scaling factor by a non-zero vector from the kernel of $\mathcal{X}$. In case we choose the normalization $p_M = 1$, the linear systems (2.6) and (1.4) are equivalent concerning the set of solutions. Alternatively, we can also use the restriction $\|p\|_2 = 1$ to get a unique coefficient vector up to some unitary factor, that of course does not change the zeros of the corresponding polynomial.

If we now solve (2.6) with respect to some norm constraint, for example, $\|p\|_2 = 1$, we get a coefficient vector, that generates a polynomial with the same zeros as P_Λ in (2.2). These polynomials are simply multiples of P_Λ and therefore also denoted by P_Λ.

Afterwards, any root finding algorithm can be used to determine the active eigenvalues in f

which in turn uniquely identify the atoms. Thus, the Vandermonde system

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1^1 & \lambda_2^1 & \dots & \lambda_M^1 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{M-1} & \lambda_2^M & \dots & \lambda_M^{2M-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_M \end{bmatrix} = \begin{bmatrix} f^{(0)}(0) \\ f^{(1)}(0) \\ \vdots \\ f^{(M)}(0) \end{bmatrix} \tag{2.8}$$

can be constructed and solved to find the linear coefficients c_j. The algorithm derived so far can be summarized as follows:

Input: $f^{(\ell)}(0)$, $\ell = 0, \dots, 2M-1$

1. Solve (2.6) w. r. t. $\|p\|_2 = 1$.
2. Compute the zeros λ_j of $P_\Lambda(z) = \sum_{\ell=0}^{M} p_\ell z^\ell$.
3. Compute c_j by solving (2.8).

Output: parameters λ_j and c_j, $j = 1, \dots, M$

It is emphasized that in contrast to Prony's method, that is based on the shift operator, the current algorithm needs no restriction on the imaginary part of the frequencies $\lambda_j \in \mathbb{C}$ because the parameters are directly given as the zeros of the polynomial P_Λ in (2.2).
Despite this advantage, the algorithm comes with the issue that we need to know derivatives of the signal up to order $2M-1$, but in practice usually only direct samples of the original signal are available. Any derivative has to be estimated by some difference scheme and the numerical stability of this estimation is highly dependent on the smoothness and oscillatory behavior of the functions in $\mathcal{M}_M$. In particular, the estimation usually becomes very unstable with increasing derivative order. Therefore, we want to circumvent this problem by using the fact that the differential operator $\frac{d}{dx}$ and the shift operator $S_\tau f := f(\cdot + \tau)$ with $\tau \in \mathbb{R}_{\neq 0}$ share the same eigenfamily corresponding to different spectra, in particular

$$\frac{d}{dx}\exp(\lambda x) = \lambda \exp(\lambda x) \;\Rightarrow\; S_\tau \exp(\lambda x) = \exp(\lambda\tau)\exp(\lambda x). \tag{2.9}$$

Implication (2.9) can be understood by the following considerations. First of all, a short calculation gives us a representation of the shift operator on the monomials with respect to the exponential power series, namely

$$\begin{aligned} \exp\left(\tau\frac{d}{dx}\right)x^m &= \sum_{k=0}^{\infty} \frac{\tau^k}{k!}\frac{d^k}{dx^k}x^m = \sum_{k=0}^{m} \frac{\tau^k}{k!}\frac{m!}{(m-k)!}x^{m-k} = \sum_{k=0}^{m}\binom{m}{k}\tau^k x^{m-k} \\ &= (x+\tau)^m = S_\tau x^m. \end{aligned}$$

Since this exponential operator is linear and the equation holds for any monomial, it is also true for every analytic function. In particular, it holds for all $f \in \mathcal{M}_M$ because the signal model is a

subset of the analytic functions on $\mathbb{R}$, i. e.,

$$\exp\left(\tau\frac{d}{dx}\right)f(x) = f(\tau + x). \tag{2.10}$$

Equation (2.10), that was first introduced by Lagrange, implies that

$$S_\tau \exp(\lambda x) = \exp\left(\tau\frac{d}{dx}\right)\exp(\lambda x) = \sum_{k=0}^{\infty}\frac{\tau^k}{k!}\frac{d^k}{dx^k}\exp(\lambda x) = \sum_{k=0}^{\infty}\frac{\tau^k}{k!}\lambda^k\exp(\lambda x) = \exp(\lambda\tau)\exp(\lambda x)$$

which directly delivers implication (2.9). Since we are only interested in the parameters of the expansion and want to choose the simplest operator which has the exponential functions as eigenfunctions, a new equation system can be derived by using a slightly changed polynomial, that we now explicitly call *Prony polynomial*. Plugging in the one-parametric shift operator S_τ yields

$$P_\Lambda^{\exp(\tau\cdot)}(S_\tau) := \prod_{j=1}^{M}\left(\exp(\lambda_j\tau)I - S_\tau\right) = \sum_{\ell=0}^{M} p_\ell S_\tau^\ell = \sum_{\ell=0}^{M} p_\ell S_{\tau\ell}. \tag{2.11}$$

The operator $P_\Lambda^{\exp(\tau\cdot)}(S_\tau)$ annihilates the same functions as $P_\Lambda\left(\frac{d}{dx}\right)$ before, since it corresponds to the same set of active eigenfunctions to different eigenvalues. Based on this knowledge we derive a modified linear system by analogously employing further iterations of the shift operator. For all f as in (1.1) with respect to a fixed Λ it holds that

$$\forall m \in \mathbb{N}_0 : F\left(S_\tau^m P_\Lambda^{\exp(\tau\cdot)}(S_\tau)\, f\right) = F\left(S_{\tau m} P_\Lambda^{\exp(\tau\cdot)}(S_\tau)\, f\right) = \sum_{\ell=0}^{M} p_\ell f((m+\ell)\tau) = 0 \in \mathbb{C}, \tag{2.12}$$

where $F(f) := f(0)$ again. From this set of equations, a Hankel system different from (2.6) can be constructed using only time discrete point evaluations of the original signal f, namely

$$\begin{bmatrix} f(0) & f(\tau) & \dots & f(M\tau) \\ f(\tau) & f(2\tau) & \dots & f((M+1)\tau) \\ \vdots & \vdots & & \vdots \\ f((M-1)\tau) & f(M\tau) & \dots & f((2M-1)\tau) \end{bmatrix}\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_M \end{bmatrix} = 0. \tag{2.13}$$

Although this system is well-known from the introduction as part of the basic Prony method [13], the essential difference is an additional sampling parameter $\tau \in \mathbb{R}_{\neq 0}$. We already know that 2.13 has kernel-dimension one, the additional parameter does not change this property, and we get a unique polynomial coefficient vector p by solving 2.13 with respect to some normalization constraint. Given these polynomial coefficients, the zeros of the Prony polynomial can

be calculated and the Vandermonde system

$$\begin{bmatrix} \exp(\lambda_1\tau)^0 & \exp(\lambda_2\tau)^0 & \dots & \exp(\lambda_M\tau)^0 \\ \exp(\lambda_1\tau)^1 & \exp(\lambda_2\tau)^1 & \dots & \exp(\lambda_M\tau)^1 \\ \vdots & \vdots & & \vdots \\ \exp(\lambda_1\tau)^{M-1} & \exp(\lambda_2\tau)^M & \dots & \exp(\lambda_M\tau)^{2M-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_M \end{bmatrix} = \begin{bmatrix} f(0) \\ f(\tau) \\ \vdots \\ f(M\tau) \end{bmatrix} \tag{2.14}$$

can be constructed to recover the linear coefficients c_j. Therefore, the Prony method can be summarized as follows: First we identify the span of fixed length M of the active eigenfunctions and afterwards the linear coefficients are determined.
Since we want to reconstruct the true exponential parameters λ_j, but only have exponentials $\exp(\lambda_j)$ of the active eigenvalues as zeros, it is necessary to restrict the parameter space of $\mathcal{M}_M$ in order to achieve uniqueness. In particular, we have to assume that $\operatorname{Im}\left(\lambda_j\right) \in [0, 2\pi/\tau)$, otherwise we cannot recover λ_j uniquely.
To put it all in a nutshell, this delivers a similar algorithm as for the differential operator, but with a sampling scheme that is much easier to realize:

Input: $f(\ell\tau)$, $\ell = 0, \dots, 2M-1$

1. Solve (2.13) w. r. t. $\|p\|_2 = 1$.
2. Solve $P_\Lambda^{\exp(\tau\cdot),f}(z) = 0$ and find the zeros $\exp(\lambda_j\tau)$.
3. Extract $\{\lambda_j\}_{j=1}^M$ from $\{\exp(\lambda_j)\}_{j=1}^M$.
4. Solve (2.14) to find c_j.

Output: (λ_j, c_j), $j = 1, \dots, M$

This algorithm is of course the well-known Prony method as seen before, with a slight generalization, namely the introduction of a sampling parameter $\tau \in \mathbb{R}_{\neq 0}$. Although this modification is not new and often used in different other works, in this approach it comes very naturally and without any additional effort as also for the Generalized Prony Method [32].
The second generalization is the use of a different evaluation functional. This is actually not a real generalization from the operator point of view, because for convenience we have always artificially restricted the sampling scheme to a point evaluation. It is emphasized that the choice of the functional is quite arbitrary since we basically only need linearity. We could have used, for example, another point evaluation functional at any point beside zero, or even integrals over suitable compact intervals. For the evaluation at an arbitrary point $x_0 \in \mathbb{R}$, the linear system 2.13 would be almost the same with the only difference of using $f(x_0 + \ell\tau)$ as samples instead of $f(\ell\tau)$. In this special case the changed starting point x_0 can be seen as a simple additional shift applied after the actual annihilation already took place, i. e.,

$$F\left(S_{x_0} S_{m\tau} P_\Lambda^{\exp(\tau\cdot)}\left(S_\tau\right) f\right) = 0.$$

Later on we will see that it is quite natural in the generalized Prony context to interpret the additional shift S_{x_0} as part of the functional and not of the annihilating operator $P_\Lambda^{\exp(\tau\cdot)}$. Also the Vandermonde system (2.14) is constructed slightly different by

$$\begin{bmatrix} \exp(\lambda_1 x_0) & \dots & \exp(\lambda_M x_0) \\ \exp(\lambda_1(x_0+\tau)) & \dots & \exp(\lambda_M(x_0+\tau)) \\ \vdots & \vdots & \vdots \\ \exp(\lambda_1(x_0+(M-1)\tau)) & \dots & \exp(\lambda_M(x_0+(2M-1)\tau)) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_M \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_0+1) \\ \vdots \\ f(x_0+M) \end{bmatrix}. \tag{2.15}$$

Moreover, concerning the construction of the sampling scheme, we are not even restricted to the shift or derivative operator. Any operator which has the exponentials as eigenfunctions can be used and every functional which does not vanish there is suitable. For example, every convolution operator in combination with a point-evaluation in zero works as well, if we adjust the algorithm accordingly.

Remark 2.1.1. The idea of using the product of suitable linear polynomial factors $(\lambda_j - z)$ in combination with the differential operator to annihilate the signal at hand can be found in [43], where this variation of Prony's method is called Operator based Annihilation. The difference to the present thesis is that in [43] this idea is only used for exponential eigenfunctions.
Although the examples concerning the differential and the difference operators with arbitrary starting points are also developed in [43], the connection between these two schemes was not given. Moreover, in [43] only point evaluation functionals were taken into account and the results therein were already implied by the Generalized Prony Method of Peter & Plonka [32].

Before we employ the ideas above for a systematic generalization, we summarize the essential properties which were used so far to reach a more abstract derivation of the basic Prony method using an operator based approach.

1. linearity of the differential operator $\frac{d}{dx}$ and the shift operator S_τ
2. The operators $\frac{d}{dx}$ and S_τ share the same eigenfunctions to different eigenvalues.
3. The mapping $\lambda \to \exp(\lambda\cdot)$ is injective on at least a restricted subset of the complex plane.
4. The signal is a finite linear combination of eigenfunctions $\exp(\lambda\cdot)$ corresponding to distinct eigenvalues.
5. We can use any functional which does not vanish on this restricted set of eigenvalues to get a finite sampling matrix.

In conclusion, it should be possible to reconstruct any atomic expansion defined by a linear operator with non-empty point spectrum by the same ideas as above. Indeed, once the operator formulation of the classical Prony method is understood, we can derive the Generalized Operator-based Prony Method within a few additional steps.

2.2 Generalized Operator based Prony Method

So far, the presented results were already known albeit we derived them in a different because operator based way. In this section we start by gradually generalizing the equation systems in (2.5) and (2.12), by the generalization of the corresponding sampling schemes

$$F\left(\frac{d^m}{dx^m}P_\Lambda\left(\frac{d^m}{dx^m}\right)\right) \quad \text{and} \quad F\left(S_\tau^m P_\Lambda^{\exp(\tau\cdot)}(S_\tau)\right),$$

in terms of the operator approach and we extract their common structure. Afterwards a strict definition of the involved terms and the main result of this thesis follows.

2.2.1 Informal preparation for the Generalization

First of all, as it was seen before, the sampling scheme is the central object which the annihilation system (2.13) and the Vandermonde system (2.14) are based on and in turn the key to all Prony-like algorithms which can be derived. Those samples are, as we know, sufficient to reconstruct functions $f : \mathbb{R} \to \mathbb{C}$ such that

$$f(x) = \sum_{\lambda\in\Lambda} c_\lambda \exp(\lambda x),$$

where $\Lambda := \{\lambda_1, \ldots, \lambda_M\} \subset \mathbb{C}$ is the finite set of the M active eigenvalues. Since the relation between eigenvalue λ and eigenfunction v_λ is one to one by assumption, we slightly modified the notation and indexed directly over the eigenvalues. Henceforth, we want to recover the parameter sets Λ and $\{c_\lambda\}_{\lambda\in\Lambda}$ corresponding to more general expansions

$$f = \sum_{\lambda\in\Lambda} c_\lambda v_\lambda.$$

For this purpose, we successively modify the sampling scheme similar as Peter & Plonka in [32]. First, the atoms v_λ must belong to an eigenfamily of a linear operator A which has a non-empty point spectrum and it has to hold that the eigenvalues uniquely identify the eigenfunctions. This means for the Prony method that $A = \frac{d}{dx}$ and $v_\lambda = \exp(\lambda\cdot)$. If we now define the Prony polynomial as

$$P_\Lambda(z) := \prod_{\lambda\in\Lambda}(\lambda - z) = \sum_{\ell=0}^{M} p_\ell z^\ell,$$

where $M = |\Lambda| < \infty$, this is the polynomial (2.2). Since the only property of A actually used in the calculations later on is linearity, we think of it from now on as a general linear operator $A : V \to V$ with non-empty point spectrum, where V is a vector space which contains the set of all finite linear combinations of eigenfunctions of A, denoted by $\mathcal{M}(A)$.

Furthermore, the operator A basically determines the model $\mathcal{M}(A)$; thus, we call A the *generator*. As mentioned for the differential sampling scheme, the action of the generator can be difficult to realize, but we are only interested in the identification of the active eigenfunctions,

so we use a different operator $\Phi_\varphi : \mathcal{M}(A) \to \mathcal{M}(A)$, defined on the eigenfunctions v_λ of A by a function φ such that

$$\Phi_\varphi v_\lambda = \varphi(\lambda) v_\lambda.$$

The use of Φ_φ instead of A is a possible way to simplify the sampling scheme. In the operator based derivation of the Prony method this was nothing else but the transition from $A = \frac{d}{dx}$ to $\Phi_\varphi = S_\tau$, where $\varphi := \exp(\tau\cdot)$. Furthermore, we had to restrict the point spectrum of A such that the exponential φ is invertible. It is stressed that the sampling parameter $\tau \in \mathbb{R}_{\neq 0}$, which is independent of the generator A, is a part of the function φ and not of the generator.
The operator Φ_φ is called *iteration operator*, because it will be iterated over and over again to recover the structure of the more general *Prony polynomial* $P_\Lambda^\varphi : \mathbb{C} \to \mathbb{C}$ which is defined as

$$P_\Lambda^\varphi(z) := \prod_{\lambda\in\Lambda} (\varphi(\lambda) - z) = \sum_{\ell=0}^{M} p_\ell z^\ell.$$

This definition directly corresponds to the modification of P_Λ in definition (2.2) to the Prony polynomial $P_\Lambda^{\exp(\tau\cdot)}$ in Definition (2.11). Thereby, we get the first general annihilation equation for all $f = \sum_{\lambda\in\Lambda} c_\lambda v_\lambda$ such that $c_\lambda \in \mathbb{C}_{\neq 0}$,

$$0 = P_\Lambda^\varphi(\Phi_\varphi) f = \prod_{\lambda\in\Lambda} (\varphi(\lambda)\,\mathrm{I} - \Phi_\varphi) f = \sum_{\ell=0}^{M} p_\ell \Phi_\varphi^\ell f,$$

where the annihilation takes place because the factors again mutually commute and

$$(\varphi(\lambda)\,\mathrm{I} - \Phi_\varphi) v_\lambda = 0$$

by assumption. The last step in a full generalization is the projection of this equation to a scalar sampling, which is used to build the necessary linear systems for the parameter identification. In (2.6) and (2.13) we used a sampling scheme based on the point evaluation functional $F(f) = f(0)$ and iterated $A = \frac{d}{dx}$ or $\Phi_\varphi = S_\tau$ to achieve the corresponding annihilation equations to generate a linear system of equations. Although this is quite straightforward, no one prevents us from using any sequence $F_m : V \to V$ of functionals, as long as the sampling matrix still has full rank. Therefore, the general sampling scheme is

$$\forall\, m \in \mathbb{N}_0 \;:\; S_{m,\ell} := F_m \circ \Phi_\varphi^\ell,$$

which results in the sampling matrix

$$\mathcal{X}_{N,M} := \left[S_{m,\ell}\right]_{m,\ell=0}^{N,M},$$

where $N \in \mathbb{N}$, $N \geq M - 1$, and $M \in \mathbb{N}$ are chosen accordingly.

Remark 2.2.1. It is emphasized that the major innovations compared to the generalized Prony method in [32] are

1. the discrimination between the generator A that defines the atoms and the iteration operator Φ_φ that is actually used for the sampling scheme,
2. a strict separation of the operator based annihilation by $P_\Lambda^\varphi(\Phi_\varphi)$ and the functional evaluation of the signal by the sequence F_m,
3. and in turn more freedom in choosing the sampling scheme $S_{m,\ell}$.

2.2.2 Strict Generalization of Prony's Method

After having clarified the basic ideas of the generalization by the rather informal discussion above, we now start the strict derivation by defining the signal model.

Definition 2.2.2 (Sparse Signal Model)**.** Let $A : V \to V$ be a linear operator on a vector space V with non-empty point spectrum $\sigma_P(A)$. Let further $\{v_\lambda\}_{\lambda\in\sigma_P(A)}$ be the family of eigenfunctions of A, i. e., $Av_\lambda = \lambda v_\lambda$. Given a finite $M \in \mathbb{N}$ the *sparse signal model* $\mathcal{M}_M(A) \subset V$ is defined as

$$\mathcal{M}_M(A) := \left\{ \sum_{\lambda\in\Lambda} c_\lambda v_\lambda \;\middle|\; \Lambda \subset \hat{\sigma}_P(A) \wedge |\Lambda| = M \wedge c_\lambda \in \mathbb{C}_{\neq 0} \right\},$$

where $\hat{\sigma}_P(A) \subseteq \sigma_P(A)$ is a subset of the point spectrum such that $\lambda \mapsto v_\lambda$ is injective. The operator A is called *generator* of the model. The signal model is called sparse because Λ will always be a finite subset of the possibly uncountable point spectrum.

Although for the general algorithm below and all of its applications we will always focus on one specific sparse signal model corresponding to a fixed and known order M, for theoretical considerations we define the surrounding general vector space of all finite linear combinations of eigenfunctions.

Definition 2.2.3 (Sparse Signal Space)**.** Given a linear operator $A : V \to V$ as in Definition 2.2.2, we define

$$\mathcal{M}(A) := \left\{ \sum_{\lambda\in\Lambda} c_\lambda v_\lambda \;\middle|\; \Lambda \subset \hat{\sigma}_P(A) \wedge |\Lambda| < \infty \wedge c_\lambda \in \mathbb{C}_{\neq 0} \right\},$$

the vector space spanned by all finite linear combinations of eigenfunctions of A corresponding to $\lambda \in \hat{\sigma}_P(A)$. The sparsity refers again to the fact that we only have finitely many active eigenvalues.

The observed major simplification of the sampling scheme in (2.13) was triggered by the transition from the differential operator to the shift operator. This transform is formalized in the next definition.

Definition 2.2.4 (Iteration Operator)**.** Let $A : V \to V$ be a generator corresponding to the sparse signal space $\mathcal{M}(A)$. Let further $\varphi : \hat{\sigma}_P^{\varphi}(A) \to \mathbb{C}$ be an injective function, where $\hat{\sigma}_P^{\varphi}(A) \subseteq \hat{\sigma}_P(A)$ is called the *restricted* point spectrum. We define the operator

$$\Phi_\varphi : \mathcal{M}(A) \to \mathcal{M}(A)$$

such that it fulfills the *spectral mapping property*,

$$\forall\, \lambda \in \hat{\sigma}_P^{\varphi}(A) \;:\; \Phi_\varphi v_\lambda = \varphi(\lambda) v_\lambda.$$

The operator Φ_φ is called the *iteration operator* for the new signal space

$$\mathcal{M}^{\varphi}(A) := \left\{ f = \sum_{\lambda\in\Lambda} c_\lambda v_\lambda \;\middle|\; \Lambda \subset \hat{\sigma}_P^{\varphi}(A) \wedge |\Lambda| < \infty \wedge c_\lambda \in \mathbb{C}_{\neq 0} \right\}.$$

As in Definition 2.2.2 we denote the *restricted sparse signal model* by $\mathcal{M}_M^{\varphi}(A)$.

This definition formalizes the fact that we always use the unique correspondence between eigenvalues and eigenfunctions, as we have seen for the classical Prony method. Even if the operator is transformed from A to Φ_φ the unique correspondence of eigenvalues and eigenfunctions has to be preserved. This may imply some further restrictions on the point spectrum to ensure the identifiability of the true parameters after switching from A to Φ_φ.
Moreover, the iteration operator is well-defined for all $f \in \mathcal{M}(A)$ as a direct consequence of the spectral mapping property

$$\Phi_\varphi f = \sum_{\lambda\in\Lambda} c_\lambda \Phi_\varphi v_\lambda = \sum_{\lambda\in\Lambda} c_\lambda \varphi(\lambda) v_\lambda \in \mathcal{M}(A).$$

In particular, the space $\mathcal{M}(A)$ is invariant under the actions of A and Φ_φ.

Remark 2.2.5. A special example for an iteration operator is the inverse $\Phi_\varphi := A^{-1}$. If the inverse A^{-1} for A as in Definition 2.2.2 exists, it holds that

$$Av_\lambda = \lambda v_\lambda \Leftrightarrow \Phi_\varphi v_\lambda = \varphi(\lambda) v_\lambda, \tag{2.16}$$

where $\varphi(z) := \frac{1}{z}$ is injective on $\mathbb{C}_{\neq 0}$.

Based on the definition of the signal model $\mathcal{M}_M(A)$ for a known expansion length $M \in \mathbb{N}$ and a given iteration operator Φ_φ, the generalized Prony polynomial and the corresponding annihilation operator are defined as follows.

Definition 2.2.6 (Generalized Prony Polynomial and Annihilation Operator)**.** Given a sparse signal model $\mathcal{M}_M^{\varphi}(A)$ with iteration operator Φ_φ, the polynomial $P_\Lambda^{\varphi} : \mathbb{C} \to \mathbb{C}$ is defined for a finite $\Lambda \subset \hat{\sigma}_P^{\varphi}(A)$ as

$$P_\Lambda^{\varphi}(z) := \prod_{\lambda\in\Lambda} (\varphi(\lambda) - z) := \sum_{l=0}^{M} p_l z^l,$$

where $M := |\Lambda| < \infty$, and it is called *Prony polynomial* with respect to φ and Λ. The corresponding operator

$$\begin{aligned} P_\Lambda^{\varphi}(\Phi_\varphi) &: \mathcal{M}(A) \to \mathcal{M}(A) \\ P_\Lambda^{\varphi}(\Phi_\varphi) &:= \prod_{\lambda\in\Lambda} \left(\varphi(\lambda) I - \Phi_\varphi\right) := \sum_{l=0}^{M} p_l \Phi_\varphi^{\ell} \end{aligned}$$

is called *annihilation operator* for all $f \in \mathcal{M}_M(A)$ corresponding to a fixed Λ.

The Prony polynomial depends on the iteration operator as well as on Λ, the set of all active eigenvalues of the signal $f \in \mathcal{M}_M(A)$, which has to be reconstructed. The annihilation operator has by definition the important property that it annihilates only the functions $f = \sum_{\lambda\in\Lambda} c_\lambda v_\lambda$, where $c_\lambda \in \mathbb{C}_{\neq 0}$ and $\Lambda \subset \hat{\sigma}_P(A)$ are fixed,

$$\begin{aligned} P_\Lambda^{\varphi}(\Phi_\varphi) f &= \sum_{l=0}^{M} p_l \Phi_\varphi^{\ell} f = \sum_{\lambda\in\Lambda} \sum_{l=0}^{M} p_l \Phi_\varphi^{\ell} c_\lambda v_\lambda \\ &= \sum_{\lambda\in\Lambda} c_\lambda v_\lambda \sum_{l=0}^{M} p_l \left(\varphi(\lambda)\right)^{\ell} = \sum_{\lambda\in\Lambda} c_\lambda v_\lambda P_\Lambda^{\varphi}(\varphi(\lambda)) = 0. \end{aligned}$$

Since we want to end up with finite linear systems, we need a certain sequence of functionals F_m as it was motivated in Section 2.2.1.

Definition 2.2.7 (Evaluation Scheme)**.** We call a sequence of linear functionals

$$F_m : \mathcal{M}(A) \to \mathbb{C} \quad \text{with} \quad m \in \mathbb{N}_0$$

an *evaluation scheme*. Moreover, let $E_F := [F_m(v_\lambda)]_{m=0,\lambda\in\Lambda}^{N}$, then the evaluation scheme is called *admissible* if and only if for a given signal space $\mathcal{M}(A)$ holds

$$\forall N \geq |\Lambda| - 1 \; \forall \; \Lambda \subset \hat{\sigma}_P^{\varphi}(A) \text{ s.t. } |\Lambda| < \infty : \operatorname{rank}(E_F) = |\Lambda|. \tag{2.17}$$

A special case of such a scheme is given by a sequence

$$F_m := F \circ B_m,$$

with respect to suitable linear operators $B_m : \mathcal{M}(A) \to \mathcal{M}(A)$. In many cases this operator sequence can be generated by powers of a single operator, i. e., $B_m := B^m$. This particular evaluation scheme plays a central role for more specific admissibility conditions. A simple example is given for the basic Prony method, where $F_m := F \circ S_\tau^m = F \circ S_{\tau m}$.

Definition 2.2.8 (Generated Evaluation Scheme)**.** Let $F_m : \mathcal{M}(A) \to \mathbb{C}$ with $m \in \mathbb{N}_0$ be an evaluation scheme and $B : \mathcal{M}(A) \to \mathcal{M}(A)$ a linear operator. Let further $F : \mathcal{M}(A) \to \mathbb{C}$ be a linear functional such that $F(v_\lambda) \neq 0$ for all $\lambda \in \hat{\sigma}_P(A)$. Then we call

$$F_m := F \circ B^m$$

the evaluation scheme generated by B. If B is an iteration operator we call it a *canonical evaluation scheme*.

Remark 2.2.9. Although it is sufficient to define the evaluation scheme only on the signal model, i. e., $F_m : \mathcal{M}(A) \to \mathbb{C}$, in most cases we can assume that the sequence F_m is defined for all elements of V.

After these definitions we can give the first theorem about the admissibility of canonical evaluation schemes.

Theorem 2.2.10. *A canonical evaluation scheme is always admissible.*

Proof. Let $\Psi_\psi : \mathcal{M}(A) \to \mathcal{M}(A)$ be an iteration operator, i. e., $\Psi_\psi v_\lambda = \psi(\lambda) v_\lambda$. Let further $F_m : \mathcal{M}(A) \to \mathbb{C}$ be an evaluation scheme generated by Ψ_ψ, i. e., $F_m := F \circ \Psi_\psi^m$, then $E_F := [F_m(v_\lambda)]_{m=0,\lambda\in\Lambda}^N$ factorizes into

$$E_F = [F_m(v_\lambda)]_{m=0,\lambda\in\Lambda}^N = [\psi^m(\lambda) F(v_\lambda)]_{m=0,\lambda\in\Lambda}^N = [\psi^m(\lambda)]_{m=0,\lambda\in\Lambda}^N \operatorname{diag}\left([F(v_\lambda)]_{\lambda\in\Lambda}\right).$$

The left factor matrix is a Vandermonde matrix with distinct entries because ψ is injective on the restricted point spectrum; thus, the Vandermonde matrix has full rank. The diagonal matrix on the right is invertible since $F(v_\lambda) \neq 0$ by definition of a generated evaluation scheme. Thus, the matrix E_F has full rank for all finite choices of $\Lambda \subset \hat{\sigma}_P(A)$ and $N \in \mathbb{N}$. □

Eventually, we come to the central object of all further modifications of the Generalized Operator based Prony Method.

Definition 2.2.11 (Sampling Scheme and Realizability)**.** Given an evaluation scheme F_m as in Definition 2.2.7 with $m \in \mathbb{N}_0$ and an iteration operator Φ_φ as in Definition 2.2.4 with respect to a signal space $\mathcal{M}(A)$, the composition

$$S_{m,\ell} := F_m \circ \Phi_\varphi^\ell : \mathcal{M}(A) \to \mathbb{C}$$

is called a *sampling scheme* for $\mathcal{M}(A)$, where $m, \ell \in \mathbb{N}_0$. A sampling scheme is called *admissible* if the corresponding evaluation scheme F_m is admissible. Additionally, a sampling scheme is called *realizable* if it can be written as a sequence of linear functionals $\tilde{F}_{m,\ell} : \mathcal{M}(A) \to \mathbb{C}$ applied to the signal itself without using signal derivatives. The result of applying $S_{m,\ell}$ to a signal is called *sampling*.

Although the definition of realizability sounds a bit strict here, in Chapter 4 it is actually shown that under certain circumstances almost every sampling scheme can be realized in the sense

above, even those based on differential operators, but for this purpose additional assumptions on the evaluation scheme are necessary. In particular, the iteration operator and the evaluation scheme do no longer only depend on the signal space, but also on each other.
Based on the definition of the sampling scheme we define the central object of every Prony-like algorithm.

Definition 2.2.12 (Sampling Matrix)**.** Given a sampling scheme $S_{m,\ell}$ as above, $f \in \mathcal{M}_M(A)$, and $N, M \in \mathbb{N}$ fixed and known such that $N \geq M-1$, the matrix valued mapping

$$\mathcal{X}_{N,M} : \mathcal{M}_M(A) \to \mathbb{C}^{N+1\times M+1}$$

is called *sampling matrix*, where

$$\mathcal{X}_{N,M}(f) := [S_{m,\ell}(f)]_{m=0,\ell=0}^{N,M} = \left[F_m\left(\Phi_\varphi^\ell f\right)\right]_{m=0,\ell=0}^{N,M}.$$

The following theorem summarizes the central statement of this thesis and the corresponding proof subsequently gives rise to an algorithm.

Theorem 2.2.13 (Generalized Operator based Prony Method (GOProM))**.** *Let $S_{m,\ell} = F_m \circ \Phi_\varphi^\ell$ be an admissible sampling scheme and*

$$\mathcal{M}_M^\varphi(A) := \left\{ \sum_{\lambda\in\Lambda} c_\lambda v_\lambda \;\middle|\; \Lambda \subset \hat{\sigma}_P^\varphi(A) \wedge |\Lambda| = M \wedge c_\lambda \in \mathbb{C}_{\neq 0} \right\},$$

where $A : V \to V$ is a generator of this model. Every $f \in \mathcal{M}_M^\varphi(A)$ is uniquely determined by the sampling matrix

$$\mathcal{X}_{N,M}(f) = \left[F_m\left(\Phi_\varphi^\ell f\right)\right]_{m,l=0}^{N,M},$$

where $M := |\Lambda|$ and $N \geq M-1$ are fixed and known natural numbers.

Proof. Let $\{v_\lambda \mid \lambda \in \hat{\sigma}_P^\varphi(A)\}$ be a fixed family of atoms such that $Av_\lambda = \lambda v_\lambda$. Let Φ_φ be an iteration operator. In particular, it holds for all $f = \sum_{\lambda\in\Lambda} c_\lambda v_\Lambda$, $c_\lambda \in \mathbb{C}_{\neq 0}$, by construction that

$$P_\Lambda^\varphi(\Phi_\varphi) f = \prod_{\lambda\in\Lambda} (\varphi(\lambda) I - \Phi_\varphi) f = 0.$$

This is a direct consequence of $\lambda \in \Lambda$ and $(\varphi(\lambda)I - \Phi_\varphi)v_\lambda = 0$ because all factors mutually commute. This implies for all evaluation schemes F_m that

$$0 = F_m\left(P_\Lambda^\varphi(A) f\right) = \sum_{\ell=0}^{M} p_\ell F_m\left(\Phi_\varphi^\ell f\right);\ m \in \mathbb{N}.$$

The involved samplings $F_m\left(\Phi_\varphi^\ell f\right)$ give rise to the sampling matrix

$$\mathcal{X}_{N,M}(f) := \left[F_m\left(\Phi_\varphi^\ell f\right)\right]_{m,\ell=0}^{N,M},$$

where $N \geq M-1$. Using the sampling matrix, we get the following linear system for the polynomial coefficients $p := [p_k]_{k=0}^{M}$:

$$\mathcal{X}_{N,M}(f)p = \left[F_m\left(\Phi_\varphi^\ell f\right)\right]_{m,\ell=0}^{N,M}[p_\ell]_{\ell=0}^{M} = 0. \tag{2.18}$$

As a necessary and sufficient condition for a reconstruction of the polynom coefficients p_ℓ and in turn of the unique active eigenvalues, the kernel of $\mathcal{X}_{N,M}(f)$ has to be one-dimensional and $p \in \ker[\mathcal{X}_{N,M}(f)] \setminus \{0\}$. To see this, we decompose the sampling matrix $\mathcal{X}_{N,M}(f)$ into

$$\mathcal{X}_{N,M}(f) = E_F D_c V_\varphi, \tag{2.19}$$

where

$$E_F := [F_m(v_\lambda)]_{m=0,\lambda\in\Lambda}^{N}\,,\; D_c := \operatorname{diag}\left([c_\lambda]_{\lambda\in\Lambda}\right),\text{ and } V_\varphi := \left[\varphi^\ell(\lambda)\right]_{\lambda\in\Lambda,\ell=0}^{M}.$$

Now, the reduced matrix

$$\check{V}_\varphi := \left[\varphi^\ell(\lambda)\right]_{\lambda\in\Lambda,\ell=0}^{M-1} \in \mathbb{C}^{M\times M}$$

is invertible since it is a Vandermonde matrix for $\varphi(\Lambda) = \{\varphi(\lambda) \mid \lambda \in \Lambda\}$, all λ are distinct, and $\varphi : \hat{\sigma}_P^\varphi \to \mathbb{C}$ is injective. Thus, the extended matrix V_φ has a kernel of exact dimension one. Since $V_\varphi p = 0$ by construction of the Prony polynomial, it follows that $p \in \ker V_\varphi \setminus \{0\}$ is unique up to a complex factor, which of course does not change the zeros of the polynomial. Since we actually need this property, it remains to ensure that both D_c and E_F have full rank. First, D_c has full rank because $c_\lambda \in \mathbb{C}_{\neq 0}$, and second E_F has full rank because we assumed the evaluation scheme F_m to be admissible. Therefore, it holds that

$$\ker\left(\mathcal{X}_{N,M}(f)\right) = \ker\left(V_\varphi\right) = \{\alpha p : \alpha \in \mathbb{C}\}.$$

Once the polynomial coefficients p_ℓ are known, we can calculate the zeros $\varphi(\lambda)$ of the Prony polynomial P_Λ^φ. Afterwards, it is possible to reconstruct the spectral parameters uniquely from the zeros $\varphi(\lambda)$, because of the injectivity of φ on the restricted point spectrum $\hat{\sigma}_P^\varphi(A)$.
For the rest of the proof it is assumed that we know the active eigenvalues $\lambda \in \Lambda$ and in turn V_φ as well as E_F. The remaining parameters c_λ can be recovered by constructing a suitable linear system. For this purpose it is sufficient to take M entries of the first column of the sampling matrix $\mathcal{X}_{N,M}(f)$ denoted by

$$\left[F_{m_j}(f)\right]_{j=1}^{M};$$

thus, we only take those sampling values into account where $\varphi^0(A)f = f$. With these selected sampling values at hand we can construct the following linear system,

$$\left[F_{m_j}(f)\right]_{j=1}^{M} = \left[\sum_{\lambda\in\Lambda} c_\lambda F_{m_j}(v_\lambda)\right]_{j=1}^{M} = \left[F_{m_j}(v_\lambda)\right]_{j=1,\lambda\in\Lambda}^{M} \cdot [c_\lambda]_{\lambda\in\Lambda}\,, \tag{2.20}$$

where $m_j \in \{0, 1, \ldots, N\}$. The left side of equation (2.20) is simply a subvector of the given

sampling matrix and the matrix on the right side is a submatrix of E_F. Therefore, the only unknowns are the linear coefficients c_λ. Since E_F has full rank it is always possible to find a selection of indices $m_j \in \{0, 1, \dots, N\}$, $j \in \{1, \dots, M\}$, such that the system has a unique solution. □

Although it is sufficient for the proof to take only a suitable part of the actual sampling matrix into account to recover the linear coefficients, the following lemma delivers an alternative which is capable of using all the available information. In particular, this can be helpful in the case of noisy samples because we use linear combinations of the rows of the sampling matrix.

Lemma 2.2.14. *Under the same assumptions as in Theorem 2.2.13, and Λ be given, the Vandermonde system*

$$\mathcal{X}_{N,M-1}^{\top}(f)\omega = \left[\varphi^{\ell}(\lambda)\right]_{\ell=0,\lambda\in\Lambda}^{M-1} \Gamma\,[c_\lambda]_{\lambda\in\Lambda}\,, \tag{2.21}$$

uniquely determines the linear coefficients c_λ of $f \in \mathcal{M}_M(A)$ for all choices of $\omega = (\omega_m)_{m=0}^{N} \in \mathbb{C}_{\neq 0}^{N+1}$ such that $\Gamma := \operatorname{diag}\left(\left[\sum_{m=0}^{N} \omega_m F_m(v_\lambda)\right]_{\lambda\in\Lambda}\right)$ is invertible and $\mathcal{X}_{N,M-1}(f) := [S_{m,\ell}]_{m,\ell=0}^{N,M-1}$.

Proof. Let be everything as in the proof before, we start with the knowledge of Λ and therefore we also know $E_F = [F_m(v_\lambda)]_{m=0,\lambda\in\Lambda}^{N}$ and $V_\Phi := \left[\varphi^{\ell}(\lambda)\right]_{\lambda\in\Lambda,\ell=0}^{M-1} \in \mathbb{C}^{M\times M}$. The sampling matrix

$$\mathcal{X}_{N,M}(f) = [F_m(v_\lambda)]_{m=0,\lambda\in\Lambda}^{N} \cdot \operatorname{diag}\left([c_\lambda]_{\lambda\in\Lambda}\right) \cdot \left[\varphi^{\ell}(\lambda)\right]_{\lambda\in\Lambda,\ell=0}^{M-1}$$

has full rank because E_F, D_c, and V_φ have full rank as before. If we now choose an arbitrary $\omega \in \mathbb{C}_{\neq 0}^{N+1}$ such that for all $\lambda \in \Lambda$ holds that $\gamma(\lambda) := \sum_{m=0}^{N} \omega_m F_m(v_\lambda) \neq 0$, which is always possible, we get a new linear system

$$\begin{aligned}\mathcal{X}_{N,M-1}^{\top}(f)\,\omega &= \left(\left[F_m(\Phi_\varphi^{\ell} f)\right]_{m,\ell=0}^{N,M-1}\right)^{\top} \omega = \left[\sum_{\lambda\in\Lambda}\sum_{m=0}^{N} c_\lambda \omega_m F_m(v_\lambda)\varphi^{\ell}(\lambda)\right]_{\ell=0}^{M-1} \\ &= \left[\sum_{\lambda\in\Lambda} c_\lambda\,\gamma(\lambda)\,\varphi^{\ell}(\lambda)\right]_{\ell=0}^{M-1}.\end{aligned}$$

This system can equivalently be written as generalized Vandermonde system

$$\mathcal{X}_{N,M-1}^{\top}(f)\omega = \left[\varphi^{\ell}(\lambda)\right]_{\ell=0,\lambda\in\Lambda}^{M-1} \Gamma\,[c_\lambda]_{\lambda\in\Lambda}\,,$$

where $\Gamma := \operatorname{diag}\left([\gamma(\lambda)]_{\lambda\in\Lambda}\right)$. The diagonal matrix Γ is invertible because $\gamma(\lambda) \neq 0$ for all $\lambda \in \Lambda$ by the choice of ω. The linear system (2.21) has in turn a unique solution, because the quadratic Vandermonde matrix $\left[\varphi^{\ell}(\lambda)\right]_{\ell=0,\lambda\in\Lambda}^{M-1}$ on the right side has full rank since all $\varphi(\lambda)$ are distinct by the construction of the restricted signal model $\mathcal{M}_M^{\varphi}(A)$ and the definition of φ. □

Remark 2.2.15. This Lemma gives us the possibility to incorporate all available information in the sampling matrix into the reconstruction of the linear coefficients c_λ. Depending on the error model one can try to find a weight vector ω to reduce the influence of noise.

The proof of Theorem 2.2.13 leads in summary to an algorithm based on the linear systems (2.18) and (2.21). The GOProM can be summarized as follows:

Input: $F_m\left(\Phi_\varphi^\ell f\right)$, $\ell = 0, \dots, M$ and $m = 0, \dots, N \geq M-1$

- Solve (2.18) with respect to a suitable normalization, e. g. $\|p\|_2 = 1$.
- Identify the active eigenvalues $\lambda \in \Lambda \subset \hat{\sigma}_P^\varphi(A)$ from the zeros $\varphi(\lambda)$ of the Prony polynomial
$$P_\Lambda^\varphi(z) = \sum_{\ell=0}^{M} p_\ell z^\ell.$$
- Recover the linear coeffients c_λ, $\lambda \in \Lambda$, by solving (2.20) or (2.21).

Output: parameter sets Λ and $\{c_\lambda \in \mathbb{C}_{\neq 0} \mid \lambda \in \Lambda\}$

2.3 Analytic Sampling Schemes

It was shown in Section 2.1 how to employ the differential operator $\frac{d}{dx}$ and the exponential representation of the shift $S_\tau = \exp\left(\tau \frac{d}{dx}\right)$ to derive the classical Prony method [13] from an operator point of view. The key property was that we preserved the eigenfunctions and only changed the eigenvalues. This corresponds to the change from the generator A to the iteration operator Φ_φ.
In this section a class of iteration operators is discussed, that includes the exponential case of Section 2.1. First, we remember the definition of an analytic function.

Definition 2.3.1 (Analytic Function)**.** A function $\varphi : \mathcal{I} \to \mathbb{C}$ on an open set $\mathcal{I} \subseteq \mathbb{C}$ is *analytic*, if for all $z \in \mathcal{I}$ holds that
$$\varphi(z) = \sum_{n=0}^{\infty} \phi_n z^n,$$
where $\{\phi_n\}_{n=0}^{N_\Psi} \subset \mathbb{C}$ such that the sum converges absolutely for all $z \in \mathcal{I}$.

Remark 2.3.2. Although for GOProM it is sufficient to assume that φ has a power series representation on the restricted point spectrum $\hat{\sigma}_P^\varphi(A)$, in all examples given in this thesis we will use analytic functions. It is noted that also polynomials fall into this category.

The following lemma delivers a class of iteration operators defined by analytic functions. In particular, the spectral mapping property has a very intuitive form.

Lemma 2.3.3 (Analytic Iteration Operator)**.** *Let $A : V \to V$ be a generator of $\mathcal{M}(A)$. Furthermore, let $\Phi_\varphi : \mathcal{M}(A) \to \mathcal{M}(A)$ be a linear operator with injective and analytic $\varphi : \mathcal{I} \to \mathbb{C}$ such that $\hat{\sigma}_P^\varphi \subset \mathcal{I}$, where $\mathcal{I} \subseteq \mathbb{C}$ is open, and $\varphi(z) := \sum_{k=0}^{\infty} \phi_k z^k$. Then*
$$\Phi_\varphi := \varphi(A) = \sum_{k=0}^{\infty} \phi_k A^k$$

is a well defined iteration operator called analytic iteration operator*; in particular, the spectral mapping property*

$$\Phi_\varphi v_\lambda = \varphi(\lambda) v_\lambda$$

holds for all $\lambda \in \hat{\sigma}_P^\varphi$.

Proof. Since $A^n v_\lambda = \lambda^n v_\lambda$ for all $n \in \mathbb{N}_0$, and all $\lambda \in \hat{\sigma}_P^\varphi(A) \subset \mathcal{I}$ are in the region of convergence of φ, it holds directly that

$$\Phi_\varphi v_\lambda = \sum_{k=0}^{\infty} \phi_k A^k v_\lambda = \sum_{k=0}^{\infty} \phi_k \lambda^k v_\lambda = \varphi(\lambda) v_\lambda.$$

Furthermore, because φ is assumed to be injective on the restricted point spectrum $\hat{\sigma}_P^\varphi$, it is an iteration operator. □

In Section 2.1 we have seen that such analytic iteration operators, namely the exponential representation of the shift, can have simpler actions as the corresponding generator. If such a simple action is given, it is reasonable to use this analytic iteration operator to generate the evaluation scheme.

Corollary 2.3.4 (Analytic Evaluation Scheme)**.** *Let* $\Psi_\psi : \mathcal{M}(A) \to \mathcal{M}(A)$ *be an analytic iteration operator . Let* $F : \mathcal{M}(A) \to \mathbb{C}$ *be a linear functional such that* $F(v_\lambda) \neq 0$ *for all* $\lambda \in \hat{\sigma}_P^\psi(A) \subset \mathcal{I}$*. Then the generated evaluation scheme*

$$F_m := F \circ \Psi_\psi^m$$

is admissible. In other words, every analytically generated sampling scheme is canonical.

Proof. This is a consequence of Lemma 2.3.3 and Theorem 2.2.10. □

Example 2.3.5. A simple example for a sampling scheme based on an analytic evaluation scheme is a mixed sampling scheme of shift and differential operator for the reconstruction of signals $f(x) = \sum_{\lambda \in \Lambda} c_\lambda \exp(\lambda x)$. Let therefore $\Phi_\varphi := A = \frac{d}{dx}$ be the iteration operator, where $\varphi(z) = z$. Let further $\Psi_\psi := \exp(\tau A) = S_\tau$, where $\psi(z) = \exp(\tau z)$, be another analytic iteration operator, that generates the evaluation scheme. Using the point evaluation $F(f) = f(0)$ we get the following canonical and therefore admissible sampling scheme

$$F_m := F \circ \Psi_\psi^m = F \circ S_{\tau m}.$$

Then

$$S_{m,\ell} := F_m \circ \Phi_\varphi^\ell = F \circ S_{\tau m} \frac{d^\ell}{dx^\ell}$$

is an admissible sampling scheme with sampling matrix

$$\mathcal{X}_{N,M}(f) = \left[f^{(\ell)}(\tau m) \right]_{m,\ell=0}^{N,M}$$

for all $f \in \mathcal{M}_M^{\varphi}(A)$. This matrix is sufficient for the reconstruction of f by Theorem 2.2.13. The advantage is that we have no further restrictions on the point spectrum $\hat{\sigma}_P^{\varphi}(A)$, since $\varphi(z) = z$. The disadvantage is that we have to sample derivatives and therefore the sampling scheme is not realizable, but we have reduced the order of the highest derivative to the minimum, namely the length M of the signals in $\mathcal{M}_M^{\varphi}(A)$.

A particular case of Corollary 2.3.4 is provided by the Generalized Prony Method [32], which corresponds to $\Psi_{\psi} := \Phi_{\varphi} := I$. The following corollary summarizes the Generalized Prony Method [32] in the new terms of this dissertation.

Corollary 2.3.6 (Generalized Prony Method)**.** *Let $A : V \to V$ be a generator of the model $\mathcal{M}_M(A)$, $M \in \mathbb{N}$, and $F : \mathcal{M}(A) \to \mathbb{C}$ a linear functional such that $|F(v_\lambda)| > 0$ for all $\lambda \in \hat{\sigma}_P(A)$. Then*

$$S_{m,\ell} := F \circ A^{m+\ell} : \mathcal{M}(A) \to \mathbb{C}$$

is an admissible sampling scheme and all $f \in \mathcal{M}_M(A)$ can be reconstructed from the sampling matrix

$$\mathcal{X}_{N,M}(f) = \left[F\left(A^{m+\ell} f\right)\right]_{m,l=0}^{N,M},$$

$N \geq M - 1$, which is made of only $2M$ different samples and has Hankel structure.

Proof. The sampling scheme in this corollary is admissible by Corollary 2.3.4 because it is generated be the generator A itself, i. e., $F_m := F \circ A^m$. This corresponds to $\varphi(z) = z$. The rest is a direct consequence of Theorem 2.2.13. □

The Generalized Prony Method [32] is applicable if the given sampling scheme is already realizable using the generator A as iteration operator; examples besides the classical exponential expansion are the monomials, Stieltjes-Wigert polynomials and expansions into cosines, which are given later in this section. Of course, if a sampling scheme as in Corollary 2.3.6 is already realizable and known, the GOProM approach seems a bit cumbersome. In contrast, the direct generalization of the basic form of Prony's method [32], that already gives the result in Corollary 2.3.6, left some structures hidden, which makes it more complicated to generate new sampling schemes for given atomic decompositions.
Moreover, GOProM gives a clear constructive way from the defining operator equation of the atoms to a realizable sampling scheme. Although GOProM is not a guarantee to find such sampling schemes, it clearifies by its very detailed structure where we have to work to find new realizable examples. In this sense, the GOProM approach also constitutes a search program for the applications of the Prony principle. Especially in Chapter 4, this structure is used to introduce a, for the Prony method, new way of data acquisation that makes almost every practically relevant sampling scheme realizable.
The numerically useful minimal sampling property of the Generalized Prony Method in [32] comes and goes with the Hankel and Toeplitz structure, which can be expressed as follows.

Definition 2.3.7 (Hankel and Toeplitz structure)**.** Let the sampling scheme be given as

$$S_{m,\ell} := F_m \circ \Phi_\varphi^\ell$$

and the sampling matrix $\mathcal{X}_{N,M} := [S_{m,\ell}]_{m,\ell=0}^{N,M}$, $N \geq M-1$. Then

(i) $S_{m,\ell}$ only depends on $m+\ell \Longleftrightarrow \mathcal{X}_{N,M}$ has Hankel structure,

(ii) $S_{m,\ell}$ only depends on $|m-\ell| \Longleftrightarrow \mathcal{X}_{N,M}$ has Toeplitz structure.

The following lemma summarizes simple conditions for theses structures.

Lemma 2.3.8 (Conditions for Hankel and Toeplitz structure)**.** *Let the sampling scheme $S_{m,\ell} := F_m \circ \Phi_\varphi^\ell$ be given as and the sampling matrix $\mathcal{X} := [S_{m,\ell}]_{m=0,\ell=0}^{N,M}$, $N \geq M-1$.*

(i) If $S_{m,\ell} = F \circ \Psi_\psi^m \Phi_\varphi^\ell$ s. t. $\Psi_\psi = \Phi_\varphi \Rightarrow \mathcal{X}_{N,M}$ has Hankel structure.

(ii) If $\Phi_\varphi = \Psi_\psi$ and

$$\forall f \in \mathcal{M}_M(A) : \Phi_\varphi^{\ell-m} f = \Phi_\varphi^{m-\ell} f$$

holds, then $\mathcal{X}$ has Toeplitz-structure.

Actually, $2M$ samples is the minimal number of needed samples. It is always achieved if the matrix is a permutation of either a Hankel or a Toeplitz matrix. In special cases it is also possible to employ symmetries to reduce the number of necessary samples to a minimum.
The next two sections are dedicated to the first examples for the application of the GOProM.

2.3.1 Sparse Cosine Expansions

In the last few sections we have seen that the central objects to reconstruct signals $f \in \mathcal{M}_M^{\varphi}(A)$ are admissible sampling schemes and the corresponding sampling matrices.
In this section we demonstrate the details of GOProM on explicit examples of sparse atomic expansions of several types; the first example is also considered in [35].
Let f be a linear combination of cosines of different frequencies $\alpha \in [0, \pi)$ and linear coefficients $c_j \in \mathbb{C}$ that both have to be recovered, i. e.,

$$f(x) := \sum_{j=0}^{M} c_j \cos(\alpha_j x).$$

Furthermore, let the generator $A := \frac{d^2}{dx^2}$. It can be easily seen that

$$A\cos(\alpha x) = -\alpha^2 \cos(\alpha x),$$

which means that $\cos(\alpha\cdot)$ is an eigenfunction of A to the eigenvalue $-\alpha^2$.
The problem is that a sampling scheme directly based on the generator A is not realizable. Thus, we have to find a suitable Φ_φ to circumvent this issue. Therefore we define for a fixed $\tau \neq 0$

$$\varphi(z) := \sum_{k=0}^{\infty} \frac{\tau^{2k}}{(2k)!} z^k,$$

which is an analytic function on $\mathbb{C}$. Moreover, the iteration operator Φ_φ is given by

$$\Phi_\varphi = \varphi(A) = \sum_{k=0}^{\infty} \frac{\tau^{2k}}{(2k)!} A^k = \sum_{k=0}^{\infty} \frac{\tau^{2k}}{(2k)!} \left(\frac{d^2}{dx^2}\right)^k = \frac{1}{2}\left(\exp\left(\tau\frac{d}{dx}\right) + \exp\left(-\tau\frac{d}{dx}\right)\right) = \frac{1}{2}\left(S_\tau + S_{-\tau}\right),$$

where S_τ is the ordinary shift with parameter $\tau \in \mathbb{R}_{\neq 0}$. By Theorem 2.3.3 we directly know that

$$\Phi_\varphi \cos(\alpha x) = \varphi(-\alpha^2)\cos(\alpha x) = \sum_{k=0}^{\infty} \frac{(-1)^k \tau^{2k}}{(2k)!} \alpha^{2k} \cos(\alpha x) = \cos(\alpha\tau)\cos(\alpha x).$$

Thus, we have found an analytical mapping φ such that the eigenfamily is preserved and the mapping is injective for all $\alpha \in [0, \pi/\tau]$; thus, it defines an iteration operator. In turn, GOProM can be applied; consequently, we have to evaluate the iteration of Φ_φ explicitely. This is possible by the standard binomial law, since an operator obviously commutes with its inverse and therefore

$$\Phi_\varphi^\ell = \left[\frac{1}{2}\left(S_\tau + S_{-\tau}\right)\right]^\ell = \frac{1}{2^\ell} \sum_{k=0}^{\ell} \binom{\ell}{k} S_\tau^{(\ell-k)} S_{-\tau}^{k} = \frac{1}{2^\ell} \sum_{k=0}^{\ell} \binom{\ell}{k} S_{(\ell-2k)\tau}.$$

This result can be used to formulate the following corollary for sparse cosine expansions.

Corollary 2.3.9 (Sparse Cosine Expansions)**.** *Let $f \in \mathcal{M}_M^{\varphi}\left(\frac{d^2}{dx^2}\right)$, $M \in \mathbb{N}$, be a linear combination of cosine atoms with respect to frequencies $\alpha \in [0, \pi K)$ and $\tau \in (0, 1/K] \subset \mathbb{R}$, $K \in \mathbb{N}$, i. e.,*

$$f(x) = \sum_{j=1}^{M} c_j \cos(\alpha_j x).$$

Then the following sampling scheme is realizable,

$$S_{m,\ell} = \frac{1}{2^\ell} \sum_{k=0}^{\ell} \binom{\ell}{k} F_m \circ S_{(\ell-2k)\tau},\ m, \ell \in \mathbb{N},$$

where F_m is an evaluation scheme. Let $F(f) = f(x_0)$, $x_0 \in \mathbb{R}$, be a linear functional such that $|F(\cos(\alpha x_0))| > 0$ for all $\alpha \in [0, \pi K)$. The sampling matrix with respect to the canonical evaluation scheme $F_m := F_{x_0} \circ S_\tau^m$, $F_{x_0}(f) := f(x_0)$, is given by

$$\mathcal{X}_{N,M}(f) := [S_{m,\ell}(f)]_{m,\ell=0}^{N,M} = \left[\frac{1}{2^{m+\ell}} \sum_{k=0}^{m+\ell} \binom{m+\ell}{k} f(x_0 + \tau(m + \ell - 2k)) \right]_{m,\ell=0}^{N,M}$$

and uniquely determines all cosine expansions $f \in \mathcal{M}_M^{\varphi}\left(\frac{d^2}{dx^2}\right)$. Actually, the matrix is made of $3M + N + 1$ samples of the form $F\left(f(x_0 + \tau n)\right)$, $n \in \{-2M, \dots, M + N\}$.

Proof. The corollary is a consequence of Theorem 2.2.13, Corollary 2.3.4, and the calculations above. The number of samples is

$$\max(m + \ell - 2k) - \min(m + l - 2k) + 1 = N + M - (-2M) + 1 = 3M + N + 1.$$

In case $N = M - 1$, the minimal size of N, we have $4M$ samples to use. □

Remark 2.3.10. Since cosine is an even function, a clever evaluation can reduce the number of necessary samples. For example, if we use the point evaluation at zero, $F_0(f) = f(0)$, the samples with respect to k and $-k$ fulfill the equation $f(\tau k) = f(-\tau k)$ and therefore we only need to know half of the samples, i. e., $2M$ samples instead of $4M$, in case $N = M - 1$.

Remark 2.3.11. Examples concerning trigonometric functions were already derived in [39, 32] using the special properties of those expansions more directly. The advantage of the GOProM is the constructive way in which the examples are generated, starting only with the defining linear differential equation and ending up with a realizable sampling scheme. The operator approach has also revealed the connection of several examples sharing a common structure that will be discussed in Section 3.3.

2.3.2 Sparse Polynomial Expansions

In this section we start with an example that is already known from the Generalized Prony Method [32], namely sparse monomial expansions, and proceed with a new family of atoms,

the so-called Stieltjes-Wigert polynomials.

Sparse Monomial Expansions

One of the simplest non-trivial examples for the Generalized Prony Method [32] are sparse expansions $f : \mathbb{R} \to \mathbb{R}$ into arbitrary monomials, i. e.,

$$f(x) := \sum_{j=1}^{M} c_j x^{n_j},$$

where $n_j \in \mathbb{N}$ and $c_j \in \mathbb{C}_{\neq 0}$. The first step to identify a suitable GOProM is to find a generator for the atoms. In this case, it is simply the dilation operator $D_q : C(\mathbb{R}) \to C(\mathbb{R})$ with $D_q f := f(q\,\cdot)$, $q \in \mathbb{R}^+ \setminus \{0, 1\}$. By

$$\left(D_q(\cdot)^{n_j}\right)(x) = (qx)^{n_j} = q^{n_j} x^{n_j}$$

we see that the monomials are eigenfunctions of this operator with q^{n_j} as eigenvalues. Thus, with the generator $A := D_q$ the restricted point spectrum is $\hat{\sigma}_P(A) = \{q^n \,|\, n \in \mathbb{N}\}$. The family of atoms is given by $v_{\lambda_n} = v_{q^n} = x^n$ for $n \in \mathbb{N}$. The next step in the GOProM approach is to verify if the action of A and all of its powers A^ℓ are already realizable. Otherwise, it has to be modified by a suitable mapping φ. Fortunately, for the dilation operator the iterated action is simply

$$D_q^\ell f = f(q^\ell \cdot).$$

This is clearly realizable and therefore we focus on the simplest version of possible sampling schemes, namely the canonical sampling scheme

$$S_{m,\ell} = F_m \circ D_q^\ell := F \circ D_q^{m+\ell},$$

where F is a linear functional such that $F(x^n) \neq 0$ for all $n \in \mathbb{N}$, e. g., $F(v_{q^n}) = 1^n \neq 0$. Based on the considerations above, the following corollary is a direct application of Corollary 2.3.6.

Corollary 2.3.12 (Sampling Theorem for Sparse Monomial Expansions)**.** *Let $f \in \mathcal{M}_M\left(D_q\right)$ be a linear combination of arbitrary monomials, i. e., $f(x) = \sum_{j=1}^{M} c_j x^{n_j}$ and F be a functional such that $|F(x^n)| > 0$ for all $n \leq N \in \mathbb{N}$, $N \geq M - 1$. Then*

$$S_{m,\ell} = F \circ D_{q^{m+\ell}}$$

is a realizable and admissible sampling scheme with $m + \ell \in \{0, 1, \ldots, N + M\}$. The sampling matrix corresponding to the point evaluation functional $F(f) = f(1)$ is given by

$$\mathcal{X}_{N,M}(f) := \left[f\left(q^{m+\ell}\right)\right]_{m,\ell=0}^{N,M}$$

and uniquely determines all sparse monomial expansions $f \in \mathcal{M}_M\left(D_q\right)$. Actually, the matrix is made of only $2M$ samples of f for $N = M - 1$.

Proof. It is a consequence of Corollary 2.3.6. □

Other possible choices for the functional F would be point evaluations in different $x_0 \in \mathbb{R}_{\neq 0}$. In contrast, non-admissible choices for F are for example all symmetric integrals $F(f) = \int_{-a}^{a} f(x)dx$, because those vanish on all odd monomials.

Stieltjes-Wigert Polynomials

Besides the monomials, in applications it is often more interesting to use orthogonal polynomials as expansion atoms. Although later on general algorithms for sparse Tschebychev expansions and other Sturm-Liouville type eigenfamilies are given, as a first example of a set of orthogonal atoms we discuss the so-called Stieltjes-Wigert polynomials. This example demonstrates that the complicated part in finding realizable sampling schemes is the operator iteration Φ_{φ}^{ℓ}. Afterwards, the evaluation scheme F_m can be chosen quite arbitrarily and it should not be a problem to realize it.
Following the definition of the Stieltjes-Wigert polynomials in [24], a new example for the application of the Generalized Prony Method in Corollary 2.3.6 can be given.
We first need some notation, let therefore $k \in \mathbb{R}_{\neq 0}^{+}$ and

$$q := \exp\left(-\frac{1}{2k^2}\right).$$

Further we remember the definition of the $q-$shifted factorial for $q \in (0,1)$, namely

$$(a,q)_0 := 1 \text{ and, } (a,q)_n := \prod_{j=0}^{n-1}(1-aq^j),\ n \in \mathbb{N}_0.$$

With this notation, the Stieltjes-Wigert polynomials for a fixed $q \in (0,1)$ can be written as

$$p_n(x) := \sum_{j=0}^{n} \frac{q^{j^2}}{(q,q)_j (q,q)_{n-j}} (-x)^j.$$

These polynomials are orthogonal to each other with respect to the inner product

$$\langle f,g\rangle_{\omega} := \int_0^{\infty} (fg)(x)\omega(x)dx \text{ with } \omega(x) = \frac{k}{\sqrt{\pi}} \exp\left(-k^2 \log^2(x)\right).$$

Actually, there are other choices of inner products which preserve the orthogonality of these polynomials.
Following Christiansen & Koelink [8], the Stieltjes-Wigert polynomials $p_n : \mathbb{R}_{\times}^{+} \to \mathbb{R}$ can also be defined as eigenfunctions of the linear operator

$$L_q = D_q - \frac{1}{x}\left(I - D_q^{-1}\right),$$

where $D_q f(x) = f(qx)$ is again the dilation operator with respect to a fixed q as above. Thus, we define p_n by

$$L_q p_n := \lambda_n p_n,$$

where $\lambda_n = q^n$ and hence we choose the restricted point spectrum $\hat{\sigma}_P(A) = \{q^n \,|\, n \in \mathbb{N}\}$.
Since the simple action of this operator is directly realizable without any modification, we choose $\varphi(z) := z$ and in turn $\Phi_\varphi = L_q$. Now we want to reconstruct sparse expansions into Stieltjes-Wigert polynomials by means of GOProM, i. e., we want to recover the parameters of functions of the following form

$$f(x) := \sum_{j=1}^{M} c_j p_{n_j}(x).$$

From Corollary 2.3.6 it is known that we first have to find the explicit form of the sampling scheme

$$S_{m,\ell} = F_m \circ L_q^\ell = F \circ L_q^m L_q^\ell = F \circ L_q^{m+\ell},$$

$(m + l) \in \{0, 1, \ldots, M + N\}$, that can be used to construct the concrete sampling matrices. Although the simple action of the generator is already realizable in terms of the signal itself, the iterated action L_q^n, $n \in \mathbb{N}_0$, is non-trivial because of the non-commutativity

$$D_q \frac{1}{x}(I - D_q^{-1}) = \frac{1}{xp}(D_q - I) \neq \frac{1}{x}(I - D_q^{-1})D_q = \frac{1}{x}(D_q - I),\ q \neq 1.$$

Thus, the binomial formula cannot be employed to evaluate L_q^n and a direct expansion of such non-commuting multinomials becomes intractable very quickly. In a first step to cope with this problem, we calculate the commutator of both operators.

Definition 2.3.13 (Commutator)**.** If X and Y are linear operators, their *commutator* is defined as

$$[X, Y] := XY - YX.$$

They commute if and only if $[X, Y] = 0$.

For the example of Stieltjes-Wigert polynomials we choose $X := D_q$ and $Y := -\frac{1}{x}\left(I - D_q^{-1}\right)$; thus, a short calculation delivers

$$[X, Y] = \left(\frac{1}{q} - 1\right) YX$$

and in turn $qXY = YX$. Since we want to reduce the action of $L_q^n = (X + Y)^n$ to the iterated actions of X and Y, we first prove a binomial law for this commutation relation.

Lemma 2.3.14 (q-Binomial Law)**.** *If X and Y are linear operators and $qXY = YX$, then*

$$(X+Y)^n = \sum_{k=0}^{n} \alpha_{n,k}(q) X^{n-k} Y^k,$$

where the coefficient polynomials $\alpha_{n,k}$ in $q \in \mathbb{R}$ are recursively defined by

$$\alpha_{n,k}(q) = q^{n-k}\alpha_{n-1,k-1}(q) + \alpha_{n-1,k}(q),$$

with $\alpha_{n,k}(q) = 0$ for all $k \in \mathbb{N} \setminus \{0, 1, \ldots, n\}$ and $\alpha_{n,k} = 1$ if $k \in \{0, n\}$.

Proof. By using the commutation relation $qXY = YX$, the first few coefficient polynomials can be calculated.

$$(X+Y)^0 = 1I \qquad (X+Y)^1 = X+Y \qquad (X+Y)^2 = X^2 + (1+q)XY + Y^2\,;$$

thus,

$$\begin{aligned} &\alpha_{0,0} = 1 \\ &\alpha_{1,0} = 1 \quad \alpha_{1,1} = 1 \\ &\alpha_{2,0} = 1 \quad \alpha_{2,1} = 1+q \quad \alpha_{2,2} = 1. \end{aligned}$$

This gives the result for the base case. Furthermore, for all $k < 0$ or $k > n$ the coefficient polynomials have to be zero. The induction step is then

$$\begin{aligned} (X+Y)^{n+1} &= (X+Y)\sum_{k=0}^{n} \alpha_{n,k}(q) X^{n-k}Y^k = \sum_{k=0}^{n} \alpha_{n,k}(q) X^{n+1-k}Y^k + \sum_{k=0}^{n} \alpha_{n,k}(q) q^{n-k} X^{n-k}Y^{k+1} \\ &= \sum_{k=0}^{n} \alpha_{n,k}(q) X^{n+1-k}Y^k + \sum_{k=1}^{n+1} \alpha_{n,k-1}(q) q^{n+1-k} X^{n+1-k}Y^k \\ &= \alpha_{n,0}(q)X^{n+1} + \sum_{k=1}^{n} \left(\alpha_{n,k}(q) + \alpha_{n,k-1}(q)q^{n+1-k}\right) X^{n+1-k}Y^k + \alpha_{n,n}(q)Y^{n+1} \\ &= X^{n+1} + \sum_{k=1}^{n} \left(\alpha_{n,k}(q) + \alpha_{n,k-1}(q)q^{n+1-k}\right) X^{n+1-k}Y^k + Y^{n+1} \\ &= \sum_{k=0}^{n+1} \alpha_{n+1,k}(q) X^{n+1-k}Y^k. \end{aligned}$$

□

By using this lemma, the iterated action L_q^n can be reduced to the simpler actions of X and Y. In particular, the iteration of $X = D_q$ is given by $X^{n-k} = D_{q^{n-k}}$ and therefore obviously realizable. In contrast, the action of

$$Y = -\frac{1}{x}\left(I - D_q^{-1}\right) = \frac{1}{x}D_q^{-1} - \frac{1}{x}I$$

is a sum of two operators $X' := \frac{1}{x}D_q^{-1}$ and $Y' := -\frac{1}{x}I$. Following this notation, Y can be expressed as

$$Y = X' + Y'.$$

Fortunately, X' and Y' commute also up to a factor by $\frac{1}{q}X'Y' = Y'X'$. Therefore, Lemma 2.3.14 can be applied again and we have

$$Y^k = (X' + Y')^k = \sum_{s=0}^{k} \alpha_{k,s}\left(\frac{1}{q}\right) X'^{k-s} Y'^s.$$

The iteration of Y' is given by

$$Y'^s = \left(-\frac{1}{x}I\right)^s = (-1)^s \frac{1}{x^s} I$$

and is therefore realizable. Moreover, the iteration of X' can be seen to be

$$X'^{k-s} = \left(\frac{1}{x}D_q^{-1}\right)^{k-s} = q^{\frac{1}{2}(k-s-1)(k-s)} \frac{1}{x^{k-s}} D_{q^{-(k-s)}}$$

by a short calculation. Since every action is now reduced to simply realizable scaling operators, we get the final result for this type of polynomials.

Corollary 2.3.15 (Sampling Theorem for Stieltjes-Wigert Polynomials)**.** *Let $f \in \mathcal{M}_M\left(L_q\right)$, $q \in (0,1)$, be a finite linear combination of arbitrary Stieltjes-Wigert Polynomials p_{n_j}, i. e.,*

$$f(x) := \sum_{j=1}^{M} c_j p_{n_j}(x).$$

Let further F be a linear functional such that $0 < |F(p_n)| < \infty$ for all $n \leq N \in \mathbb{N}$, $N \geq M-1$. Then an admissible sampling scheme is given by

$$S_{m,\ell} := F \circ L_q^n = \sum_{k=0}^{n} \sum_{s=0}^{k} \omega_{n,k,s}(q) \frac{1}{x^{k+s-1}} F \circ D_{q^{n-2k+s}},$$

where $n := m + \ell \in \{0, \ldots, M+N\}$, and

$$\omega_{n,k,s}(q) := \alpha_{n,k}(q)\alpha_{k,s}\left(\frac{1}{q}\right)(-1)^s q^{\kappa_{n,k,s}},$$

with $\kappa_{n,k,s} := \frac{1}{2}(k-s-1)(k-s) + s(2k-s-n) - (k-1)(n-k)$. The sampling matrix corresponding to the point evaluation functional $F(f) := F_1(f) = f(1)$ is given by

$$\mathcal{X}_{N,M}(f) := \left[\sum_{k=0}^{n} \sum_{s=0}^{k} \omega_{n,k,s}(q) \frac{1}{x^{k+s-1}} f\left(q^{n-2k+s}\right)\right]_{m,\ell=0}^{N,M}$$

and uniquely determines all sparse Stieltjes-Wigert expansions $f \in \mathcal{M}_M\left(L_q\right)$. Actually, the matrix is made of only $2M$ samples of f if $N = M-1$.

Proof. This is a direct consequence of Corollary 2.3.6 and a backwards substitution of the results above. The number of samples is determined by $0 \leq n-2k+s \leq n$, where $k \in \{0, \ldots, n\}$

and $s \in \{0, \ldots, k\}$. It is clear that $2M$ samples of the form $F(D_{q^\ell} f)$ with $\ell \in \{0, 1, \ldots, 2M-1\}$ are enough for reconstruction, which is achieved for $N = M - 1$. □

It is emphasized that this is only one possible sampling theorem, since we have quite a bit of freedom in choosing the evaluation scheme. In particular, the monomial and Stieltjes-Wigert example demonstrate again that any expansion generated by a finite linear combination of operators with realizable actions can be directly handled with the Generalized Prony Method. In principle, it is always possible to calculate an explicit expression for the iteration of superimposed operators X_k, $k \in \mathbb{N}$, i. e.,

$$A^n := \left(\sum_{k=1} X_k \right)^n .$$

The problem is that these expressions become very quickly intractable in length and complexity if n gets bigger. The sampling schemes based on a very high number of different operator combinations are likely to be difficult to handle in practice, which makes it complicated to realize the sampling scheme. In terms of implementability this is not desirable and we should always try to find the simplest expression available.
The commutation relation in the Stieltjes-Wigert example is only a special case of a more general commutation by an arbitrary monomial power of a known parameter, i. e., $q^s XY = YX$. In this case we always have a binomial law type expansion in terms of the so called q-binomial coefficient or Gaussian polynomials $\binom{n}{k}_q$. Their iterated actions can be even explicitly given for all powers in the following form

$$(X+Y)^n = \sum_{k=0}^{n} \binom{n}{k}_q X^k Y^{n-k}.$$

A detailed definition and discussion of this and other useful expansions of non commuting operators can be found in [16, 21]. Many other possible atoms defined by difference equations are discussed in [22].
Once those samples are realized, the linear systems, which are the heart of GOProM, can be constructed and the algorithm can be applied. Therefore, the main problem is always how to get those samples.

In summary, the crucial point which distinguishes GOProM from all former approaches is the clear separation of the following steps.

1. Define a signal space $\mathcal{M}(A)$ based on the eigenfunctions of the generator A.

2. Determine whether this generator already has a realizable action.

3. If it has no such simple action, try to find a modification Φ_φ which has one, and determine the restricted signal space $\mathcal{M}^\varphi(A)$.

4. Find a law for the iterated action Φ_φ^ℓ.

5. Choose a realizable and admissible evaluation scheme F_m to get an admissible sampling scheme $S_{m,\ell} = F_m \circ \Phi_\varphi^\ell$.

6. Determine the order M of the signal model and

7. construct the sampling matrix $\mathcal{X}_{N,M}$ according to the sampling scheme $S_{m,\ell}$.

8. Apply GOProM to recover $f \in \mathcal{M}_M^\varphi(A)$.

After having seen that the most complicated steps are finding the explicit action of the iteration operator and the iteration law, the next section introduces a wide class of atoms which has suprisingly simple actions and easily tractable iteration laws. This class already covers all realizable examples which have been known so far.

CHAPTER 3

Generalized Exponential Prony Method

> All truths are easy to understand once they are discovered; the point is to discover them.
>
> Galileo Galilei

This section presents a class of applications for GOProM which is a natural generalization of Prony's method. Although examples for this class are already known, e. g., finite linear combinations of monomials, in the Prony context they were not considered as part of the class of exponential atoms before. With the theory of Chapter 2 at hand, the corresponding signal model can be easily derived and the algorithms are particularly simple to implement.
As a motivation we start again with the basic Prony method. The signal model is the set of all finite linear combinations of complex exponentials, i. e., functions of the form

$$f(x) = \sum_{j=1}^{M} c_j \exp(\lambda_j x) = \sum_{\lambda\in\Lambda} c_\lambda v_\lambda(x),$$

where $\Lambda = \{\lambda_1, \dots, \lambda_M\}$. The generator $A = \frac{d}{dx}$ is the ordinary derivative, the analytical mapping $\varphi = \exp(\tau\cdot)$, where $\tau \in \mathbb{R}_{\neq 0}$ is a sampling parameter, and an admissible evaluation scheme $F_m = F \circ \Phi_\varphi^m$ is simply generated by the iteration operator $\Phi_\varphi = \exp\left(\tau\frac{d}{dx}\right) = S_\tau$ with respect to the point evaluation functional $F := F_{x_0}$, $x_0 \in \mathbb{R}$. This leads to the canonical sampling scheme

$$S_{m,\ell} = F_m \circ \Phi_\varphi^\ell = F \circ \Phi_\varphi^m \Phi_\varphi^\ell = F_{x_0} \circ S_{(k+\ell)\tau},$$

as seen in former sections. Besides this concrete example, we can ask for more general linear operators and their exponential actions. As a first step in this direction, we make the following formal calculation,

$$\exp(\tau A) f = \exp(\tau A) \sum_{\lambda\in\Lambda} c_\lambda v_\lambda = \sum_{\lambda\in\Lambda} c_\lambda \exp(\tau A) v_\lambda = \sum_{\lambda\in\Lambda} c_\lambda \exp(\lambda\tau) v_\lambda.$$

This looks quite promising, since it reproduces the classical Prony form up to the functions v_λ. Nevertheless, the major issue in the examples given so far was the action of the iteration operator for an arbitrary integer power $\ell \in \mathbb{N}$, namely Φ_φ^ℓ. Especially in the Stieltjes-Wigert

example, the iteration law was non-trivial and one can easily imagine that this gets even more complicated for other commutators besides the relatively simple case of commutation up to a constant. Fortunately, for exponential operators this iteration is as simple as for exponential functions, i. e.,

$$\Phi_\varphi^\ell = (\exp(\tau A))^\ell = \exp(\ell\tau A).$$

Therefore, if we use a suitable evaluation scheme F_m, for example $F_m = F \circ \Phi_\varphi^m$, we get the following equation as a consequence of Lemma 2.3.3,

$$\begin{aligned} S_{m,\ell}(f) &= F\left(\Phi_\varphi^m \Phi_\varphi^\ell f\right) = F\left(\exp(\tau A)^m \exp(\tau A)^\ell f\right) = F\left(\exp(\tau m A)\exp(\tau \ell A) f\right) \\ &= F\left(\exp(\tau(m+\ell)A) f\right) = \sum_{\lambda\in\Lambda} c_\lambda F(v_\lambda)\exp(\lambda\tau(m+\ell)) = \sum_{\lambda\in\Lambda} \tilde{c}_\lambda \exp(\lambda\tau(m+\ell)), \end{aligned}$$

where $\tilde{c}_\lambda := c_\lambda F(v_\lambda)$. In this sense, the classical Prony problem is restored by using the exponential of the generator. As long as the action of this exponential iteration operator is known, we can reconstruct Λ and $\{c_\lambda\}_{\lambda\in\Lambda}$ and thereby identify the function $f = \sum\limits_{\lambda\in\Lambda} c_\lambda v_\lambda$. Moreover, let $t \in \mathbb{R}$ be interpreted as a variable instead of a fixed sampling parameter like τ before. Then it also holds that

$$\rho_t(f) := F\left(\exp(tA) f\right) = \sum_{\lambda\in\Lambda} \tilde{c}_\lambda \exp(\lambda t), \tag{3.1}$$

where the coefficients $\tilde{c}_\lambda$ are defined as above. This can be seen as a parametric transform

$$\begin{aligned} \rho_t &: \mathcal{M}(A) \to \mathcal{M}\left(\frac{d}{dt}\right) \\ \rho_t &\left(\sum_{\lambda\in\Lambda} c_\lambda v_\lambda\right) = \sum_{\lambda\in\Lambda} \tilde{c}_\lambda \exp(\lambda t), \end{aligned}$$

where we map a general sparse signal space $\mathcal{M}(A)$, parameterized by $t \in \mathbb{R}$, to a basic exponential signal space with t as the variable.

The identifiability is guaranteed under the same restrictions as for the classical Prony method and the assumption that the functional F exists and does not vanish on the eigenfunctions of A for all $\lambda \in \hat{\sigma}_P^{\exp}(A)$.

Actually, this kind of transform is implicitly used by Mourrain in [29] to derive a Prony-like algorithm. There the mapping ρ_t is called the generating sequence

$$\rho_n(f) = F(\exp(nA) f) \text{ for } n \in \mathbb{N}$$

of f if it is evaluated at the integers. Mourrain developed a multivariate Prony method for the reconstruction of sparse atomic signals including also generalized eigenfunctions. The approach is intrinsically connected with Hankel operators and their corresponding matrices. Therefore, it is not clear how to generalize it directly to sampling schemes where the Hankel structure is missing, for example, the cosine or the new Stieltjes-Wigert expansions.

3.1 Exponential Operators and Operator Semi-Groups

Before introducing the *Generalized Exponential Prony Method* (GEProM) formally, we discuss some issues concerning exponential operators in general, following [15]. In particular, this section shows that parametric iteration operators as above always have an exponential form and are well behaved.
The major obstacle on the way to a realizable sampling scheme, even for trivial actions Φ_φ, is the iteration law for Φ_φ^ℓ. Despite this issue for other examples, we have learned from the basic Prony method that an exponential iteration operator can yield very simple iterative actions. In particular, the motivating calculation before has shown that the key property which makes it possible to simplify the iteration in many cases can be formulated in the following terms:

$$\forall n \in \mathbb{N}_0 : \Phi_\varphi^n = \varphi(nA). \tag{3.2}$$

This means that instead of finding a decomposition of the operator Φ_φ into a sum of realizable operators, just to get a possibly untractable multinomial law for the iterations Φ_φ^n afterwards, we only need to know the parametric action of $\varphi(nA)$. Of course, this may be also challenging, but in some cases we are thereby able to simplify the iteration tremendously. Although we know that $\Phi_\varphi = \exp(\tau A)$ fulfills this property, it is not obvious whether the special form of the iteration law in (3.2) already determines that Φ_φ has an exponential representation.
Therefore, at first let Φ_φ be an iteration operator with parametric iteration law as in (3.2). If we also want to rescale with a sampling parameter $\tau \in \mathbb{R}$ arbitrarily, we have to assume that

$$\forall h \in \mathbb{R}^+ : \Phi(hA)^n = \varphi^{nh}(A) = \varphi(nhA),$$

which is implied by

$$\forall t \in \mathbb{R}^+ : \Phi_\varphi^t = \varphi(tA). \tag{3.3}$$

Substituting $t := x + y \in \mathbb{R}^+$, this can also be written as

$$\forall x, y \in \mathbb{R}^+ : \Phi_\varphi^x \Phi_\varphi^y = \varphi^x(A)\varphi^y(A) = \varphi^{x+y}(A) = \varphi((x+y)A).$$

If we now substitute $T_t := \Phi_\varphi^t$ we get

$$\forall x, y \in \mathbb{R}^+ : T_{x+y} = T_x T_y = T_y T_x, \tag{3.4}$$

which is one of the defining properties of a semi-group of operators. Moreover, from a sampling point of view it seems quite natural to define $T_0 := I$. This leads to the well-known concept of a semi-group of operators.

Definition 3.1.1 (One-parametric Semi-Group of Operators)**.** A family of linear operators $(T_t)_{t\geq 0}$ on a Banach space is called *one-parametric semi-group*, if

$$T_{x+y} = T_x T_y \; ; \forall \; x, y \in R^+ \text{ and } T(0) := I.$$

It is called *one-parametric group* if the factorization even holds for all $x, y \in \mathbb{R}$.

Since we derived this definition from a property of the exponential iteration operator, the two following questions are natural.

1. For which generators are exponential iteration operators well-defined?

2. Can we use other φ besides the exponential power series to yield the same parametric sampling structure as in (3.2)?

To answer these questions we have to assume from now on that V is at least a Banach space equipped with

$$||A|| := \sup_{||f||=1} ||Af||,$$

the operator norm.

Definition 3.1.2 (Bounded Operators)**.** An operator $A : V \to V$ is called *bounded* if and only if

$$||A|| := \sup_{||f||=1} (||Af||) < \infty.$$

Furthermore, for all bounded operators we can define uniform continuity as follows.

Definition 3.1.3 (Uniform Continuity)**.** A one-parametric semi-group of operators T_t, $t \in \mathbb{R}^+$, is *uniformly continuous* if

$$\lim_{h\to 0} ||T_{t+h} - T_t|| = 0,$$

with respect to the operator norm.

With these definitions at hand, both questions above can be answered.

Theorem 3.1.4. *Let V be a Banach space, then the following holds.*

1. *If $A : V \to V$ is bounded, $(\exp(tA))_{t\geq 0}$ is a uniformly continuous semi-group.*

2. *If $(T_t)_{t\geq 0}$ is a uniformly continuous semi-group, a linear bounded operator A exists such that*

$$T_t = \exp(tA).$$

Proof. We only prove the first assertion because the second one would need theory which is not in the scope of this thesis. The second assertion is a consequence of Theorem 3.5 and 3.7

in [15].
First we show that $\exp(tA) = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$ is well-defined which is true since A is bounded:

$$\|\exp(tA)\| = \|\sum_{k=0}^{\infty} \frac{t^k}{k!} A^k\| \leq \sum_{k=0}^{\infty} \frac{t^k}{k!} \|A\|^k = \exp(t\|A\|) < \infty.$$

Therefore,

$$\|(\exp((t+h)A) - \exp(tA))\| = \|\exp(tA)(\exp(hA) - I)\| \leq \exp\left(t \cdot \|A\|\right) \|(\exp(hA) - I)\|,$$

and in turn it is sufficient to show that $\|\exp(hA) - I\| \to 0$ for $h \to 0$, which is true since

$$\|\exp(hA) - I\| = \left\| \sum_{k=1}^{\infty} \frac{h^k}{k!} A^k \right\| \leq \exp(|h| \cdot \|A\|) - 1 \overset{h \to 0}{\longrightarrow} 0.$$

The semi-group property can be assured by using the Cauchy product:

$$\exp(tA)\exp(hA) = \sum_{k=0}^{\infty} \sum_{s=0}^{k} \frac{t^{k-s} A^{k-s}}{(k-s)!} \frac{h^k A^k}{k!} = \sum_{k=0}^{\infty} \frac{(t+h)^k A^k}{k!} = \exp((t+h)A).$$

□

This actually answers two questions. First, if we have a bounded generator, the exponential iteration operator is defined for all non-negative real powers and we can easily construct a sampling matrix with a Hankel structure by using the canonical sampling scheme. The Hankel structure of this specific sampling matrix is a consequence of Lemma 2.3.8. Second, if we already have a realizable iteration operator which is uniformly continuous, it possesses exponential structure and we can find a corresponding bounded generator. Moreover, the theorem shows that we even have a group structure, since the exponential operator is actually defined for all real powers $t \in \mathbb{R}$. Although a bounded operator A can be singular, the exponential iteration operator is always invertible due to

$$\exp(-tA)\exp(tA) = \exp(tA)\exp(-tA) = \exp(0A) = I.$$

Remark 3.1.5. In the theory of operator semi-groups and dynamical systems, the iteration operator $\Phi_\varphi = \exp(A)$ is called *evolution operator*, since it determines the time-evolution of a linear system by formally solving a generalized Cauchy boundary value problem. In particular, exponential operators play an important role in quantum mechanics where they are used to investigate solutions of generalized difference equations, as for example done in [10, 12, 11]. Particularly, the special exponential operators discussed in the next section are of interest in this field.

3.2 Generalized Exponential Prony Method

Following the discussion in Section 2.2, the Generalized Exponential Prony Method (GEProM) can be easily derived by setting $\Phi_\varphi := \exp(\tau\cdot)$ for a fixed $\tau \in \mathbb{R}_{\neq 0}$, and subsequently choosing the other terms accordingly. So far, the discussion of this special case revealed a few advantages:

1. a smooth structure of the sampling scheme
2. parametric iteration laws
3. invertibility of the iteration operator and therefore group structure
4. reduction of an arbitrary model to the classical exponential problem
5. possibility of using all existing Prony-like algorithms for the reconstruction

Since all results are consequences of Theorem 2.2.13 or even explicitly discussed in the introduction, we summarize the GEProM in the following theorem.

Theorem 3.2.1 (Generalized Exponential Prony Method)**.** *Let $A : V \to V$ be a generator of the signal model $\mathcal{M}_M(A)$ with non-empty point spectrum $\hat{\sigma}_P(A) \cap \mathbb{R} \times i\,[0, 2\pi) \neq \emptyset$, $\Phi_\varphi := \exp(\tau A)$ be a realizable iteration operator with sampling parameter $\tau \in \mathbb{R}_{\neq 0}$. Furthermore, the restricted point spectrum is $\hat{\sigma}_P^{\exp}(A) \subseteq \mathbb{R} \times i\left[0, \frac{2\pi}{\tau}\right)$. Then the sampling scheme*

$$S_{m,\ell} := F_m \circ \exp(\ell\tau A)$$

is admissible and realizable, if F_m is an admissible evaluation scheme. If we choose the canonical evaluation scheme $F_t := F \circ \exp(tA)$ such that $|F(v_\lambda)| > 0$ and $t \in \mathbb{R}_{\neq 0}$, we get the transform

$$\rho_t : \mathcal{M}(A) \to \mathcal{M}\left(\frac{d}{dt}\right)$$
$$\rho_t\left(\sum_{\lambda\in\Lambda} c_\lambda v_\lambda\right) = \sum_{\lambda\in\Lambda} \tilde{c}_\lambda \exp(\lambda t),$$

where $\tilde{c}_\lambda = c_\lambda F(v_\lambda)$. The sampling matrix for the canonical sampling scheme is given by

$$\mathcal{X}_{N,M}(f) = \left[F\left(\exp(\tau(m+\ell)A)\, f\right)\right]_{m,\ell=0}^{N,M}$$

which uniquely determines all sparse expansions $f \in \mathcal{M}_M^{\exp}(A)$. The matrix size is determined by the signal order $M \in \mathbb{N}$ and the number of rows $N \in \mathbb{N}$, $N \geq M-1$. If we choose $N = M-1$, we only need $2M$ samples.

Proof. Although all assertions are implied by Theorem 2.2.13, we shortly discuss the assumptions in the theorem above. The first one is

$$\hat{\sigma}_P(A) \cap \mathbb{R} \times i\,[0, 2\pi) \neq \emptyset.$$

This property guarantees that we always find a non-empty subset of eigenvalues and injectively corresponding eigenfunctions such that

$$\forall \lambda \in \hat{\sigma}_P^{\exp}(A) \subseteq \hat{\sigma}_P(A) \cap \mathbb{R} \times i\,[0, 2\pi) : \lambda \to \varphi(\lambda) \text{ is injective .}$$

In turn, we can uniquely identify the active λ by the zeros $\varphi(\lambda) = \exp(\tau\lambda)$ of the generalized Prony polynomial. The admissibility of $F_m := F \circ \exp(\tau m A)$ is guaranteed by Theorem 2.3.4 under the assumption that $0 < |F(v_\lambda)|$. The transform $\rho_{\tau m+l}$ is the same as in equation (3.1), where $t := \tau m + l$. □

In the following examples of this section, we focus on analytical schemes. Although this sounds very straightforward, the difficult part is actually to find the explicit expressions of the iteration operator and its iterated action. The rest of the chapter is devoted to examples where these actions are realizable and it is shown that most of the classical applications of the Prony method are already covered by these terms.

3.2.1 Generalized Shift Operators

Section 3.1 gave a general result for bounded operators which shows the importance of exponential operators concerning natural sampling schemes. Still, many of the interesting cases we will see are based on unbounded generators. In general, the theory for semi-groups becomes slightly more complicated in this case. Fortunately, we are only interested in functions $f \in \mathcal{M}(A)$, and as discussed before, A and all possible iteration operators Φ_φ are well-defined on this set.
In this section we want to use a specific analytic sampling scheme to reconstruct sparse expansions $f : [x_0, x_1] \to \mathbb{C}$, with $x_0, x_1 \in \overline{\mathbb{R}}$, such that

$$f(x) := \sum_{\lambda \in \Lambda} c_\lambda \frac{\exp(\lambda G(x))}{H(x)}, \tag{3.5}$$

where $c_\lambda \in \mathbb{C}$ and $\lambda \in \mathbb{R} \times i[0, 2\pi)$. This signal model generalizes the classical Prony method, where it held that $G(x) = x$ and $H(x) \equiv 1$. We assume that the function $G : [x_0, x_1] \to \mathbb{R}$ is of the form

$$G(x) := \int_{x_0}^{x} \frac{1}{g(y)} dy, \tag{3.6}$$

where $g : [x_0, x_1] \to R$ is continuous and non-zero. Thus, the so-called *phase function* G is differentiable and monotone. The function $1/g$ can be seen as the change in phase and therefore it is called *instantaneous phase function*. In addition, we assume that $H : [x_0, x_1] \to \mathbb{R}$ is defined by

$$H(x) := \exp\left(\int_{x_0}^{x} \frac{a(y)}{g(y)} dy\right), \tag{3.7}$$

where g is the same function as in the definition of G and $a : [x_0, x_1] \to \mathbb{R}$ is assumed to be integrable and chosen such that H is non-zero. Actually, by these assumptions the atoms $v_\lambda(x) :=$

$\frac{\exp \lambda G(x)}{H(x)}$ form a one-parametric eigenfamily of the generator $A : C^1[x_0, x_1] \to C[x_0, x_1]$

$$A := \left(g(x)\frac{d}{dx} + a(x)I \right),$$

since

$$\begin{aligned} Av_\lambda(x) &= \left(g(x)\frac{d}{dx} + a(x)I \right) \frac{\exp(\lambda G(x))}{H(x)} = \left(g(x)\frac{d}{dx} + a(x)I \right) \exp\left(\lambda G(x) - \int_{x_0}^{x} \frac{a(y)}{g(y)} dy \right) \\ &= g(x)\left(\lambda \frac{1}{g(x)} - \frac{a(x)}{g(x)} \right) \exp\left(\lambda G(x) - \int_{x_0}^{x} \frac{a(y)}{g(y)} dy \right) + a(x) \exp\left(\lambda G(x) - \int_{x_0}^{x} \frac{a(y)}{g(y)} dy \right) \\ &= (\lambda - a(x) + a(x)) \frac{\exp(\lambda G(x))}{H(x)} = \lambda v_\lambda(x). \end{aligned}$$

By using $\varphi := \exp(\tau \cdot)$, we get an analytic iteration operator, which is well-defined since exp is an entire function. In contrast to the case of a bounded operator A, the iteration operator

$$\Phi_\varphi := \varphi(A) = \exp(\tau A),$$

where $\tau \in \mathbb{R}_{\neq 0}$, does not have to be uniformly continuous anymore. It is not even clear that $\exp(\tau A)$ exists on the whole domain of A. Therefore, we restrict ourselves to the signal space $\mathcal{M}(A)$, because there the iteration operator is always well-defined.
Eventually, it remains to evaluate the parametric action of the semi-group $\exp(\tau A)$, that is given for all $f \in \mathcal{M}(A)$ by the following lemma.

Lemma 3.2.2 (Generalized Shift Operators)**.** *Let the generator* $A := \left(g\frac{d}{dx} + aI \right)$ *on* $V := C^1[x_0, x_1]$ *with* $x_0 < x_1 \in \overline{\mathbb{R}}$ *and* $a, g : [x_0, x_1] \to \mathbb{R}$ *be given as in* (3.6) *and* (3.7)*. The action of* $\exp(tA)$ *on all* $f \in \mathcal{M}(A) \subset V$ *is given by*

$$\exp(tA) f(x) = \frac{H\left(G^{-1}(t + G(x))\right)}{H(x)} f\left(G^{-1}(t + G(x))\right)$$

for all $t \in \mathbb{R}$.

Proof. It is sufficient to prove the action for an arbitrary element $f \in \mathcal{M}(A)$. Let us first consider the exponential action of the simpler operator

$$\check{A} := g(x)\frac{d}{dx},$$

where $a \equiv 0$, i. e., for the eigenfunctions $\check{v}_\lambda(x) = H(x)\, v_\lambda(x) = \exp(\lambda G(x))$. By Lemma 2.3.3 for analytic functions, we know that

$$\exp(\tau \check{A})\, \check{v}_\lambda = \exp(\lambda \tau)\, \check{v}_\lambda,$$

and thus

$$\begin{aligned}\exp(\tau\check{A})\,\check{v}_\lambda(x) &= \exp(\lambda\tau)\exp(\lambda G(x)) = \exp(\lambda(\tau+G(x)))\\ &= \exp\left(\lambda G\left(G^{-1}\left(\tau+G(x)\right)\right)\right) = \check{v}_\lambda\left(G^{-1}\left(\tau+G(x)\right)\right).\end{aligned}$$

The same calculation can now be done for the generator A

$$\begin{aligned}\exp(\tau A)\,v_\lambda(x) &= \exp(\lambda\tau)\,v_\lambda(x) = \exp(\lambda\tau)\frac{\exp(\lambda G(x))}{H(x)} = \frac{1}{H(x)}\exp(\lambda(\tau+G(x)))\\ &= \frac{1}{H(x)}\exp\left(\lambda G\left(G^{-1}\left(\tau+G(x)\right)\right)\right) = \frac{1}{H(x)}\exp(\tau\check{A})(H(x)\cdot v_\lambda(x))\\ &= \left(\frac{\exp\left(\tau\check{A}\right)H(x)}{H(x)}\right)\exp(\tau\check{A})v_\lambda(x),\end{aligned}$$

which proves the lemma for all v_λ with $\lambda\in\mathbb{C}$ and in turn for all $f\in\mathcal{M}(A)$. □

Remark 3.2.3. If there is a dense subfamily of v_λ in $C[x_0,x_1]$ the action is actually the same for all $f\in C[x_0,x_1]$.

The action of generalized shift operators is obviously realizable and can directly be used to define an admissible sampling scheme $S_{m,\ell} := F\circ\exp\left((m+\ell)\tau A\right)$, where F is a linear functional such that $0<|F(v_\lambda)|<\infty$. Since the action of the iteration operator is quite simple, it is reasonable to use a canonical evaluation scheme $F_m := F\circ\exp(\tau mA)$. The results are summarized in the following theorem.

Theorem 3.2.4 (Sparse General Exponential Expansions)**.** *Let*

$$g:[x_0,x_1]\to\mathbb{R}\ \text{ and }\ a[x_0,x_1]\to\mathbb{R}$$

be functions as in (3.6) *and* (3.7)*, respectively. Furthermore, let* $A := g\frac{d}{dx} + aI$ *be a generator of*

$$\mathcal{M}_M^{\exp}(A) := \left\{\sum_{\lambda\in\Lambda} c_\lambda\frac{\exp(\lambda G(x))}{H(x)}\;\middle|\; c_\lambda\in\mathbb{C}\wedge\Lambda\subset\hat{\sigma}_P^{\exp}(A)\wedge|\Lambda|=M\in\mathbb{N}\right\},$$

with $H(x)=\exp\left(\int_{x_0}^x\frac{a(y)}{g(y)}dy\right)$ *and* $G(x)=\int_{x_0}^x\frac{1}{g(y)}dy$*, then* $\Phi_\varphi := \exp(\tau A)$*, with* $\tau\in\mathbb{R}_{\neq0}$*, is an iteration operator with the restricted point spectrum* $\hat{\sigma}_P^{\exp}(A)=\mathbb{R}\times i[0,2\pi/\tau]$*. Given a canonical evaluation scheme* $F_m := F\circ\exp(\tau mA)$ *such that* $0<|F(v_\lambda)|<\infty$ *for all* $\lambda\in\hat{\sigma}_P(A)$*, then*

$$S_{m,\ell} = F\circ\left[\frac{H\left(G^{-1}(\tau(m+\ell)+G(\cdot))\right)}{H(\cdot)}f\left(G^{-1}(\tau(m+\ell)+G(\cdot))\right)\right]$$

is an admissible sampling scheme. Using the point evaluation in x_0*, it holds for the sampling matrix*

$$\mathcal{X}_{N,M}(f) := \left[F\left(H\left(G^{-1}(\tau(m+\ell))\right)f\left(G^{-1}(\tau(m+\ell))\right)\right)\right]_{m,\ell=0}^{N,M},$$

where $N\geq M-1$*, is sufficient for the reconstruction of* $f\in\mathcal{M}_M^{\exp}(A)$ *by GEProM.*

Proof. The existence of the inverse G^{-1} is ensured because G is strictly monotone and continuously differentiable by definition of g. Lemma 3.2.2 and Theorem 2.2.13 provide the rest. The sampling matrix is achieved by using the point evaluation of $\exp(\tau(m+\ell)A)f$, $f \in \mathcal{M}_M^{\exp}(A)$ and A as in (3.2.4), in the lower bound x_0 of the definition of G in (3.6), i. e., $F_{x_0} f = f(x_0)$. $|x_0| < \infty$, of the integral $G(x) = \int_{x_0}^{x} \frac{1}{g(y)} dy$, because $G(x_0) = 0$ holds by definition for any choice of g.
The great advantage of the point evaluation in the lower bound is that it guarantees admissibility due to

$$F(v_\lambda) = F(\exp(\lambda G(x_0))) = 1 \neq 0.$$

□

In particular, $\mathcal{X}_{N,M}(f)$ has Hankel structure and therefore it is possible to get a minimal sampling matrix, with only $2M$ distinct entries if we set $N = M - 1$.
Table 3.1 summarizes some examples for several functions G. For convenience, the function H is set constant to one, i. e., $a(x) = 0$ for all x. The evaluation functional $F(f) = f(\tilde{x})$ was chosen to be the point evaluation at $\tilde{x} \in [x_0, x_1] \subseteq \mathbb{R}$.

Inst. Phase: g	Phase Fct.: G	Atoms: v_λ	Sampling Scheme: $S_{m,\ell}$
$-1/x$	$-\frac{1}{2}x^2$	$\exp\left(-\frac{\lambda}{2}x^2\right)$	$f\left(\sqrt{\tilde{x}^2 - \tau(m+\ell)}\right)$
1	x	$\exp(\lambda x)$	$f(\tau(m+\ell) + \tilde{x})$
x	$\log(x)$	x^λ	$f\left(e^{\tau(m+\ell)}\tilde{x}\right)$
$-\sqrt{1-x^2}$	$\arccos x$	$\exp(\lambda \arccos x)$	$f(\cos(\tau(m+\ell) + \arccos(\tilde{x})))$
$\sqrt{1-x^2}$	$\arcsin x$	$\exp(\lambda \arcsin x)$	$f(\sin(\tau(m+\ell) + \arcsin(\tilde{x})))$
$\frac{1}{\cos(x)}$	$\sin x$	$\exp(\lambda \sin x)$	$f(\arcsin(\tau(m+\ell) + \sin(\tilde{x})))$
$-\frac{1}{\sin(x)}$	$\cos x$	$\exp(\lambda \cos x)$	$f(\arccos(\tau(m+\ell) + \cos(\tilde{x})))$

Table 3.1: Examples concerning the GEProM with sampling parameter h

The function H can be seen as an almost arbitrary amplitude function, that is known and can therefore be removed before the actual group action takes place. In the basic Prony method the evaluation functional is simply the point evaluation in zero, which yields the simplest sampling scheme. Such a simplification can often be derived by using a suitable point evaluation functional. A reasonable example is given in the proof of Theorem 3.2.4.
The rest of the section is devoted to several examples, including cosine expansions from a more general point of view and linear combinations of arbitrary Chebychev polynomials.

Sparse Exponential Cosine Expansions

Atomic decompositions into linear combinations of functions $\exp(\lambda \cos(x))$ as in Table 3.1 play an important role in the field of electromagnetic engineering, e. g., [26]. Therefore, they are worth a closer investigation by means of GEProM in Theorem 3.2.4. In this example it is shown

how to reconstruct functions $f : \mathbb{R} \to \mathbb{C}$ of the type

$$f(x) = \sum_{\lambda \in \Lambda} c_\lambda \exp(\lambda \cos(x))$$

in terms of generalized shift operators. The corresponding signal model is

$$\mathcal{M}_M^{\exp}(A) := \left\{ \sum_{\lambda \in \Lambda} c_\lambda \exp(\lambda \cos(x)) \;\middle|\; c_\lambda \in \mathbb{C} \setminus \{0\} \wedge \Lambda \subset \hat{\sigma}_P^{\exp}(A) \wedge |\Lambda| = M \in \mathbb{N} \right\},$$

where $\tau \in \mathbb{R}_{\neq 0}$ is a sampling parameter and $\hat{\sigma}_P^{\exp}(A) = \mathbb{R} \times i[0, 2\pi/\tau)$. The atoms are defined as

$$v_\lambda(x) := \exp(\lambda \cos(x)) : \mathbb{R} \to \mathbb{C}.$$

The first derivative of v_λ using the chain rule delivers the corresponding generator

$$A := -\frac{1}{\sin(x)} \frac{d}{dx} : C^1\left[\frac{\pi}{2}, \pi\right) \to C\left[\frac{\pi}{2}, \pi\right),$$

that is the generator of a generalized shift, as in Lemma 3.2.2, with $1/g(y) = -\sin(y)$ and $a \equiv 0$. Thus,

$$G(x) = \int_{x_0}^{x} \frac{1}{g(y)} dy = \int_{\pi/2}^{x} -\sin(y) dy = \cos(x).$$

Furthermore, since $a \equiv 0$ we have that $H(x) = \exp\left(\int_{\pi/2}^{x} \frac{0}{g(y)} dy\right) = 1$. The inverse of $G : [0, \pi] \to [-1, 1]$ is given by

$$G^{-1} : [-1, 1] \to [0, \pi]$$
$$G^{-1}(x) = \arccos(x)$$

From Theorem 3.2.4 it is known that the following sampling matrix is sufficient for reconstruction of $f \in \mathcal{M}_M^{\exp}(A)$,

$$\mathcal{X}_{N,M}(f) := \left[F\left(\frac{H\left(G^{-1}(\tau(m+\ell))\right)}{H(\cdot)} f\left(G^{-1}(\tau(m+\ell))\right) \right) \right]_{m,\ell=0}^{N,M} = \left[f\left(\arccos(\tau(m+l))\right) \right]_{m,\ell=0}^{N,M}.$$

Here we used the point evaluation $F_{\pi/2} f(x) = f(\pi/2)$ as evaluation functional and the evaluation scheme $F_m := F_{\pi/2} \circ \exp(\tau A)$. This evaluation scheme is admissible since

$$0 < F_{\pi/2}(v_\lambda) = 1 < \infty.$$

Eventually, we know that $\tau(m + \ell) \in [-1, 1]$ has to be fulfilled and therefore $\tau \in [-1/\hat{N}, 1/\hat{N}]$, $\hat{N} = M + N - 1$, but τ is assumed to be positive; thus $\tau \in (0, 1/\hat{N}]$. This implies that we can reconstruct any $\lambda \in [0, 2\pi\hat{N}]$ and in turn GEProM can reconstruct any $f \in \mathcal{M}_M^{\exp}\left(-\frac{1}{\sin(x)} \frac{d}{dx}\right)$ by adjusting the sampling parameter τ.

Piecewise Equidistant Sampling

If we forget the theory behind the generalized shift operators for a moment, we can think of more general examples for G and H than only differentiable functions. The only properties which have to be imposed on them to achieve the eigenstructure necessary for the reconstruction of

$$f(x) = \sum_{\lambda\in\Lambda} c_\lambda \frac{\exp(\lambda G(x))}{H(x)} : I \to \mathbb{C}$$

are continuity and monotonicity, that are sufficient for the existence of G^{-1}. The function H only has to be continuous and unequal to zero over the whole interval $I \subseteq \mathbb{R}$ and in turn either strictly smaller or greater than zero. We define the generalized shift operator for all $t \in \mathbb{R}$ directly by

$$T_{t,G,H} f(x) := \frac{H\left(G^{-1}(t + G(x))\right)}{H(x)} f\left(G^{-1}(t + G(x))\right).$$

Thus, it is clear by

$$\begin{aligned} T_{t,G,H}\, v_\lambda(x) = T_{\tau,G,H} \frac{\exp(\lambda G(x))}{H(x)} &= \frac{H\left(G^{-1}(\tau + G(x))\right)}{H(x)} \frac{\exp\left(\lambda G(G^{-1}(\tau + G(x)))\right)}{H\left(G^{-1}(\tau + G(x))\right)} \\ &= \frac{\exp(\lambda G(x))}{H(x)} \exp(\lambda\tau) = \exp(\lambda\tau) v_\lambda(x) \end{aligned}$$

that the atoms in the expansion above are eigenfunctions of $T_{\tau,G,H}$ and the action is realizable. This enables for instance the use of piecewise defined polynomials, for example,

$$G(x) := \begin{cases} x & x \in (-\infty, 0] \\ 0.8x & x \in (0, 1] \\ 2x - 1.2 & x \in (1, 2] \\ 0.1x + 2.6 & \end{cases} \qquad G^{-1}(x) := \begin{cases} x & x \in (-\infty, 0] \\ \frac{5}{4}x & x \in (0, 1] \\ \frac{1}{2}(x + 1.2) & x \in (1, 2] \\ 10(x - 2.6) & x \in (2, \infty)\,. \end{cases}$$

This corresponds to a change in the sampling frequency during the measurement process which has to be compensated afterwards.

Remark 3.2.5 (Mixed Generalized Shift Generators)**.** If we assume that for all $j \in \{1, 2, \ldots, d\}$ the operators

$$g_j \frac{d}{dx} : C^1[x_0, x_1] \to C[x_0, x_1],$$

$x_0, x_1 \in \overline{\mathbb{R}}$, are the generators of generalized shifts, $a_j \in C[x_0, x_1]$, and the linear combination

$$\tilde{A} := \sum_{j=1}^{d} \alpha_j \exp\left(g_j \frac{d}{dx}\right)$$

is well-defined with known eigenfamily v_λ, then we always get a realizable sampling scheme by

$$S_{m,\ell} = F_m \circ \tilde{A}^\ell,$$

since it is made of operators with realizable actions. Operators of the form of $\tilde{A}$ include, for instance, all difference operators with constant coefficients. Another example of this form are the Stieltjes-Wigert polynomials from Corollary 2.3.15, which have the generator

$$L_q := D_q + \frac{1}{x} D_q^{-1} - \frac{1}{x} I = \exp\left(\tau x \frac{d}{dx}\right) + \frac{1}{x} \exp\left(-\tau x \frac{d}{dx}\right) - \frac{1}{x} \exp\left(0 x \frac{d}{dx}\right),$$

where $q := \exp(\tau)$. The connection of the dilation operator and its exponential representation can be seen in Table 3.1.

In the last examples we always already knew the generators to the desired atoms. In the following section we give an example with a non-trivial and in some sense non-intuitive generator, that has to be found first.

3.2.2 Sparse Expansions into Gaussian Chirps

A signal model which is often encountered in signal processing and other fields of electrical engineering are sparse sums of so called Gaussian chirps, i. e., atoms $v_{\mu,\omega} : \mathbb{R} \to \mathbb{C}$ of the form

$$v_{\mu,\omega}(x) := \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) \exp\left(i\omega(x-\mu)\right),$$

where $\sigma^2 \in \mathbb{R}_\times^+$ is a known scale parameter, $\mu \in \mathbb{R}$ is an unknown time-shift, and $\omega \in [0, 2\pi)$ an unknown frequency parameter determining the imaginary and real oscillation pattern of the signal. In this section we want to reconstruct sparse expansions into those atoms, namely

$$f(x) := \sum_{k=1}^{M} c_k \exp\left(-\frac{1}{2\sigma^2}(x-\mu_k)^2\right) \exp\left(i\omega_k(x-\mu_k)\right),$$

where c_k, μ_k and ω_k are the unknown parameters. They are used to model signals of damped complex oscillations, for example, to describe natural speech signals [27, 46].
First of all, we define $\lambda := \mu + i\sigma^2\omega$ to write these atoms in a more compact form,

$$\begin{aligned}
v_{\mu,\omega}(x) :&= \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2 + i\omega(x-\mu)\right) = \exp\left(-\frac{1}{2\sigma^2}\left(x^2 - 2\mu x + \mu^2\right) + i\omega(x-\mu)\right) \\
&= \exp\left(-\frac{1}{2\sigma^2}\left(x^2 - 2\mu x + \mu^2 - 2i\sigma^2\omega x + 2i\sigma^2\omega\mu\right)\right) \\
&= \exp\left(-\frac{1}{2\sigma^2}\left(x^2 - 2\mu x + \mu^2 - 2i\sigma^2\omega x + 2i\sigma^2\omega\mu - \sigma^4\omega^2 + \sigma^4\omega^2\right)\right) \\
&= \exp\left(-\frac{(\sigma\omega)^2}{2}\right) \exp\left(-\frac{1}{\sigma^2}\left(x^2 - 2(\mu + i\sigma^2\omega)x + (\mu + i\sigma^2\omega)^2\right)\right) \\
&= \exp\left(-\frac{(\sigma\omega)^2}{2}\right) \exp\left(-\frac{1}{\sigma^2}\left(x^2 - 2\lambda x + \lambda^2\right)\right) \\
&= \exp\left(-\frac{(\sigma\omega)^2}{2}\right) \exp\left(-\frac{1}{2\sigma^2}(x-\lambda)^2\right).
\end{aligned}$$

Thus, we can rewrite the Gaussian chirps more compactly as

$$v_{\mu,\omega} = \exp\left(-\frac{(\sigma\omega)^2}{2}\right) v_\lambda,$$

where $v_\lambda(x) := \exp\left(-\frac{1}{2\sigma^2}(x-\lambda)^2\right)$. Since we are only interested in the reconstruction of linear combinations we can rewrite $f(x)$ as

$$f(x) := \sum_{k=1}^{M} c_k \exp\left(-\frac{1}{2\sigma^2}(x-\mu_k)^2\right) \exp\left(i\omega_k(x-\mu_k)\right) = \sum_{k=1}^{M} \tilde{c}_k \exp\left(-\frac{1}{2\sigma^2}(x-\lambda_k)^2\right),$$

where $\tilde{c}_k := c_k \exp\left(-\frac{(\sigma\omega)^2}{2}\right) \in \mathbb{C}$. In contrast to the cosine and piecewise polynomial example, it is not obvious which generator corresponds to these atoms. Therefore we start by deriving the generator. For this reason, the first derivative of v_λ is calculated, that is

$$\frac{d}{dx}\exp\left(-\frac{1}{2\sigma^2}(x-\lambda)^2\right) = -\frac{1}{\sigma^2}(x-\lambda)\exp\left(-\frac{1}{2\sigma^2}(x-\lambda)^2\right)$$
$$-\sigma^2\frac{d}{dx}v_\lambda = xv_\lambda - \lambda v_\lambda,$$

and a rearrangement yields the equation

$$\left(\sigma^2\frac{d}{dx} + x\,\mathrm{I}\right)v_\lambda = \lambda v_\lambda.$$

This means that the operator $A := \left(\sigma^2\frac{d}{dx} + xI\right)$ has the atoms v_λ as eigenfunctions. It is apparently the generator of a generalized shift operator. Therefore, Theorem 3.2.4 can be applied to get the following corollary.

Corollary 3.2.6 (Sampling theorem for sparse Expansion of Gaussian Chirps)**.** *Let* $A := \sigma^2\frac{d}{dx} + xI$, $\Phi_\varphi := \exp(\tau A)$, *and*

$$f(x) := \sum_{\lambda\in\Lambda}^{M} c_\lambda \exp\left(-\frac{1}{2\sigma^2}(x-\lambda)^2\right),$$

where $\sigma^2 \in \mathbb{R}^+_{\neq 0}$ *is a known scale parameter. The parameters* $c_\lambda \in \mathbb{C}$ *and* $\lambda \in \mathbb{R} \times i[0, 2\pi/\tau)$ *are unknown. Let* $k(x) := \exp\left(x\tau\ell + \frac{1}{2}\sigma^2\tau^2\ell^2\right)$, *the corresponding sampling scheme is given by*

$$S_{m,\ell} = F_m \circ \Phi_\varphi^\ell = F_m \circ \left(k \cdot \exp\left(\sigma^2\tau\ell\frac{d}{dx}\right)\right).$$

The corresponding sampling matrix with $F_m := F \circ \exp(\tau m A)$ *as evaluation scheme, where* $F(f) := f(0)$, *is given by*

$$\mathcal{X}_{N,M}(f) = \left[\exp\left(\frac{1}{2}\sigma^2\tau^2(m+l)^2\right) f\left(\sigma^2\tau(m+\ell)\right)\right]_{m,\ell=0}^{N,M},$$

The sampling matrix $\mathcal{X}_{N,M}(f)$ *is sufficient for the reconstruction of* f *by GEProM.*

Proof. Since $A := \left(\sigma^2 \frac{d}{dx} + xI\right)$ is the generator of a generalized shift as in Lemma 3.2.2, we first calculate G and H. Due to $g(x) = \sigma^2$ and $a(x) = x$ we get

$$G(x) := \int_0^x \frac{1}{g(y)} dy = \frac{1}{\sigma^2} x \Longleftrightarrow G^{-1}(x) = \sigma^2 x$$

and

$$H(x) = \exp\left(\int_0^x \frac{a(y)}{g(y)} dy\right) = \exp\left(\frac{1}{2\sigma^2} x^2\right).$$

With these integrals we already know the actions of the iteration operator $\Phi_\varphi := \exp(\tau A)$ and all of its powers $\Phi_\varphi^\ell = \exp(\tau \ell A)$ from Theorem 3.2.2, which is

$$\Phi_\varphi^\ell f(x) = \exp(\tau \ell A) f(x) = \frac{H\left(G^{-1}(\tau\ell + G(x))\right)}{H(x)} f\left(G^{-1}(\tau\ell + G(x))\right),$$

where $\tau \in \mathbb{R}_\times^+$ is an arbitrary sampling parameter. With $G^{-1}(\tau\ell + G(x)) = x + \sigma^2\tau\ell$ it follows that

$$\Phi_\varphi^\ell f(x) = \exp(\tau \ell A) f(x) = \frac{H\left(x + \sigma^2\tau\ell\right)}{H(x)} f\left(x + \sigma^2\tau\ell\right).$$

Furthermore,

$$\frac{H\left(x + \sigma^2\tau\ell\right)}{H(x)} = \exp\left(\frac{1}{2\sigma^2}\left[(x + \sigma^2\tau\ell)^2 - x^2\right]\right) = \exp\left(x\tau\ell + \frac{1}{2}\sigma^2\tau^2\ell^2\right),$$

which gives with k as above the sampling scheme

$$S_{m,\ell} = F_m \circ \Phi_\varphi^\ell = F_m \circ \exp(\tau\ell A) = F_m \circ \left(k \cdot \exp\left(\sigma^2\tau\ell \frac{d}{dx}\right)\right).$$

If we now use the evaluation scheme $F_m := F \circ \Phi_\varphi^m$, with $F(f) = 0$, we get as a consequence the sampling matrix

$$\mathcal{X}_{N,M}(f) = [S_{m,\ell}(f)]_{m,\ell=0}^{N,M} = \left[\exp\left(\frac{1}{2}\sigma^2\tau^2(m+l)^2\right) f\left(\sigma^2\tau(m+\ell)\right)\right]_{m,\ell=0}^{N,M}.$$

This matrix has full rank and is therefore sufficient for the reconstruction of f, because the sampling scheme $S_{m,\ell}$ is admissible by Theorem 2.3.4 due to

$$\forall\ \lambda \in \mathbb{C}\ :\ 0 < |F(v_\lambda)| = \left|\exp\left(-\frac{1}{2\sigma^2}\lambda^2\right)\right| < \infty.$$

□

If we use GEProM to reconstruct the parameters λ and c_λ of f in Corollary 3.2.6, we can extract the complete information about the time shift and frequency parameter in the Gaussian chirp

model, since

$$\operatorname{Re}(\lambda) = \mu \qquad \frac{1}{\sigma^2}\operatorname{Im}(\lambda) = \omega.$$

We can also think of a slight generalization of the signals in Corollary 3.2.6. Let therefore be

$$v_\lambda(x) := \exp\left(\frac{1}{2\sigma^2}(k(x) - \lambda)^2\right),$$

where $k : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable strictly monotone function. A generator for those atoms is

$$A_k = \left(\frac{\sigma^2}{k'}\frac{d}{dx} + k\,\mathrm{I}\right).$$

Then it holds for G and H analogously to the proof above that

$$G(x) = \frac{1}{\sigma^2}k(x) \Longleftrightarrow G^{-1}(x) = k^{-1}(\sigma^2 x) \text{ and } H(x) = \exp\left(\frac{1}{2\sigma^2}k^2(x)\right).$$

A short calculation delivers the corresponding sampling scheme; first, we easily see that

$$\frac{H\left(G^{-1}(t + G(x)\right)}{H(x)} = \exp\left(\frac{1}{2}\sigma^2 t^2 + tk(x)\right),$$

for all $t \in \mathbb{R}$. By using the point evaluation $Ff(x) = f(x_0)$, $x_0 \in \mathbb{R}$, and based on the evaluation scheme $F_m(f) = F\left(\exp^m(A)f\right)$, we get for $t = \tau(m + \ell)$ the following sampling matrix of a signal $f \in \mathcal{M}_M^{\exp}(A_k)$,

$$S_{m,\ell}(f) = \exp\left(\tau(m + l)k(x_0) + \frac{\sigma^2}{2}\tau^2(m + \ell)^2\right) f\left(k^{-1}\left(\sigma^2\tau(m + l) + k(x_0)\right)\right).$$

Depending on the function k it can be necessary to restrict the point spectrum even further to ensure its invertibility.

Numerical examples for Gaussian Chirps

Now, we give a small numerical example to demonstrate that the approach works. Therefore, we want to reconstruct a linear combination of Gaussian chirps of order five, i. e.,

$$f(x) = \sum_{\lambda\in\Lambda} c_\lambda \exp\left(-\frac{1}{2\sigma^2}(x - \lambda)^2\right),$$

where $|\Lambda| = 5$ and the eigenvalues λ and the linear coefficients are drawn randomly.
For a concrete example we draw five realizations of $c_\lambda \sim U[-10, 10] + iU[-10, 10]$, where U denotes the uniform distribution. Furthermore, we draw the complex shift parameters λ uniformly from $[1, 15] + i\,[-\pi, \pi)$. The used realizations are summarized in Table 3.2. If we

$\mathrm{Re}(\lambda)$	12.011	12.549	8.112	3.787	2.675
$\mathrm{Im}(\lambda)$	1.101	2.913	−1.984	1.410	0.560
$\mathrm{Re}(c_\lambda)$	−2.444	−9.723	−3.223	−6.578	9.195
$\mathrm{Im}(c_\lambda)$	1.773	−4.966	−4.111	9.068	1.878

Table 3.2: Shift parameters λ and the corresponding linear coefficients

now sample with the canonical sampling scheme

$$S_{m,\ell}(f) = \exp\left(\frac{1}{2}\sigma^2\tau^2(m+l)^2\right) f\left(\sigma^2\tau(m+\ell)\right),$$

where $\sigma^2 = 2$, $x_0 = 0$, and use GEProM we can exactly recover the function f, which is depicted in Figure 3.1. We sample equidistantly over $[0, 7]$, thus $\tau = 0.7$. The approximated mean squared error over the interval $[-5, 15]$ is 3.319×10^{-7}. Since we only used noiseless samples, we set $N = M = 5$, i. e., we take only 10 samples $S_{m,\ell}(f)$ into account, with $m + \ell \in \{0, \dots, 2M - 1\}$. Thus, the sampling matrix has the form

$$\mathcal{X}_{5,5} = \begin{bmatrix} S_{0,0}(f) & S_{0,1}(f) & S_{0,2}(f) & S_{0,3}(f) \\ S_{1,0}(f) & S_{1,1}(f) & S_{1,2}(f) & S_{1,3}(f) \\ S_{2,0}(f) & S_{2,1}(f) & S_{2,2}(f) & S_{2,3}(f) \end{bmatrix}.$$

This kind of example was also given in [35], the above sampling scheme for Gaussian chirps was derived therein in a slightly different way, but with the same final sampling matrix.

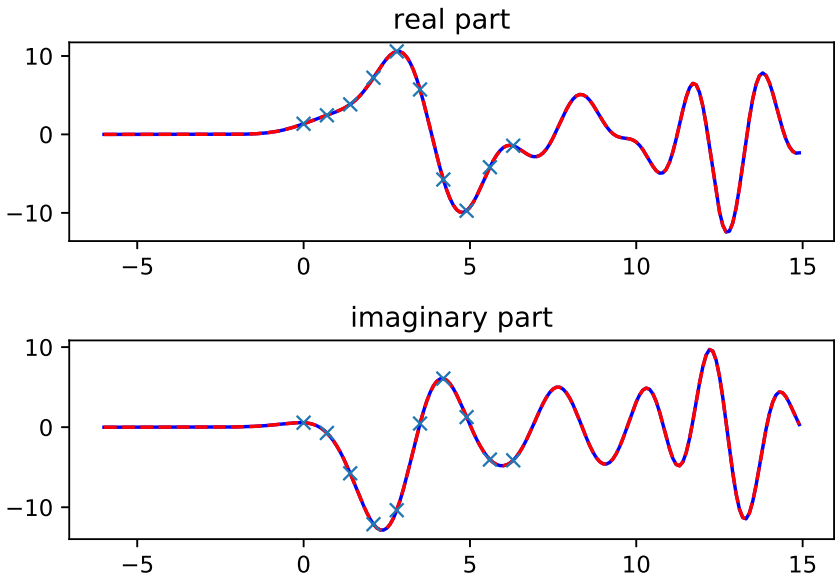

Figure 3.1: ground truth in blue, reconstruction red, and samples as blue crosses

3.3 Squared Operators as Generators

A further interesting example for the GEProM are generators which can be written as squares of simpler operators.
In Section 2.3.1 it was demonstrated how to reduce a certain generator of the cosine model to a linear combination of ordinary shifts. In this section the ideas of the cosine example are generalized to generators that can be written as squares of operators with a realizable exponential action. Principally, there are two possible approaches, the first one is a direct generalization of the cosine example and the second one includes a certain convolution operator.

3.3.1 Symmetric Generalized Shift Approach

Before we start with the direct generalization of the idea in Section 2.3.1, the following theorem gives a more general result concerning the representation of squared operators.

Lemma 3.3.1 (Symmetric Generalized Shift Operators)**.** *Let $B : V \to V$ be a linear operator on a vector space V such that the point spectrum $\sigma_P(B) \neq \emptyset$ and $\exp(B) : \mathcal{M}(A) \to \mathcal{M}(A)$ is well-defined. If $A := B^2$, then the point spectrum of A is not empty. Furthermore, let $\varphi(z) := \sum_{n=0}^{\infty} \frac{\tau^{2n}}{(2n)!} z^n$, $z \in \mathbb{C}$, and $\Phi_\varphi := \varphi(A)$. Then*

$$\Phi_\varphi = \frac{1}{2}\left(\exp(\tau B) + \exp(-\tau B)\right)$$

and for all $\ell \in \mathbb{N}$ it holds that

$$\Phi_\varphi^\ell = \frac{1}{2^\ell}\left(\exp(\tau B) + \exp(-\tau B)\right)^\ell = \frac{1}{2^\ell}\sum_{k=0}^{\ell}\binom{\ell}{k}\exp\left(\tau(\ell-2k)B\right).$$

Furthermore, for all $\omega \in \sigma_P(A)$ it holds that

$$\Phi_\varphi v_\omega^A = \varphi(\omega) v_\omega^A,$$

where v_ω^A is the eigenfunction of A to the eigenvalue ω.

Proof. Let $A := B^2$ and further $\varphi(z)$ as above, than it holds that

$$\Phi_\varphi = \sum_{n=0}^{\infty}\frac{\tau^{2n}}{(2n)!}A^n = \sum_{n=0}^{\infty}\frac{\tau^{2n}}{(2n)!}(B^2)^n = \sum_{n=0}^{\infty}\frac{\tau^{2n}}{(2n)!}B^{2n} = \frac{1}{2}\left(\exp(\tau B) + \exp(-\tau B)\right).$$

Assume that the point spectrum $\sigma_P(B)$ is non-empty. Since

$$A v_\lambda^B = B^2 v_\lambda^B = \lambda^2 v_\lambda^B,$$

where v_λ^B is the eigenfunction of B to the eigenvalue λ, it follows that

$$\left\{\lambda^2 \,\middle|\, \lambda \in \sigma_P(B)\right\} \subseteq \sigma_P(A)$$

and therefore the point spectrum of A is not empty either. In turn, we get by Lemma 2.3.3

$$\Phi_\varphi v_\omega = \varphi(\omega) v_\omega.$$

The iteration law for Φ_φ^ℓ is simply the binomial law since $\exp(\tau B)$ and $\exp(-\tau B)$ commute. □

This lemma enables us to reduce the action of a generator A to the possibly simpler exponential action of the operator B. In particular, if B has a realizable exponential action, Lemma 3.3.1 delivers a suitable iteration operator to construct a realizable sampling scheme. The following theorem gives a concrete example, that includes the cosine example of Section 2.3.1.

Theorem 3.3.2 (Squared Generalized Shifts I)**.** *Let the operator $A : C^2[x_0, x_1] \to C[x_0, x_1]$ be defined as*

$$A := \left(g\frac{d}{dx}\right)^2 = g^2\frac{d^2}{dx^2} + gg'\frac{d}{dx}, \tag{3.8}$$

where $g \in C^1[x_0, x_1]$ as in (3.6)*, then A has the point spectrum $\sigma_P(A) = \mathbb{C}$. Furthermore, let be $\varphi(z) := \sum_{n=0}^{\infty} \frac{\tau^{2n}}{(2n)!} z^n$ for all $z \in \sigma_P(A)$ as in Lemma 3.3.1, then the sampling scheme*

$$S_{m,\ell} = F_m \circ \Phi_\varphi^\ell = \frac{1}{2^\ell}\sum_{k=0}^{\ell}\binom{\ell}{k} F_m \circ \exp\left((\ell - 2k)\tau g\frac{d}{dx}\right)$$

is realizable. It is admissible if we choose the canonical evaluation scheme $F_m := F \circ \Phi^m$, such that $0 < |F(v_\lambda)| < \infty$ for all $\lambda \in \hat{\sigma}_P(A)$.
Thus, the sampling matrix

$$\mathcal{X}_{N,M} = \left[\frac{1}{2^{m+\ell}}\sum_{k=0}^{m+\ell}\binom{m+\ell}{k} f\left(G^{-1}((m+\ell-2k)\tau + G(x_0))\right)\right]_{m,\ell=0}^{N,M},$$

where $N \geq M - 1 \in \mathbb{N}$ and $Ff(x) := f(x_0)$ such that $\infty > |v_\lambda(x_0)| > 0$, is sufficient for the reconstruction of all $f \in \mathcal{M}_M^\varphi(A)$.

Proof. We first prove that the point spectrum of A is equal to $\mathbb{C}$. We remember that $B := g\frac{d}{dx}$ is actually a generator of a generalized shift as in Theorem 3.2.4. Thus, it has $v_\lambda^B(x) = \exp(\lambda G(x)))$ as eigenfunctions for all $\lambda \in \mathbb{C}$ as eigenvalues. Since $A := B^2$ and in turn $Av_\lambda^B = \lambda^2 v_\lambda^B$, it holds that

$$\mathbb{C} = \sigma_P(B) = \left\{\lambda^2 \;\middle|\; \lambda \in \sigma_P(B)\right\} \subseteq \sigma_P(A) \subseteq \mathbb{C}.$$

Furthermore, we know from Lemma 3.3.1 that

$$\Phi_\varphi^\ell = \frac{1}{2}\left(\exp(\tau B) + \exp(-\tau B)\right) = \frac{1}{2^\ell}\sum_{k=0}^{m+\ell}\binom{\ell}{k}\exp\left(\tau(\ell - 2k)B\right).$$

Applying an evaluation scheme $F_m : C^2[x_0, x_1] \to \mathbb{C}$ yields the sampling scheme in Theorem

3.3.2 above,

$$S_{m,\ell} = F_m \circ \Phi_\varphi^\ell = \frac{1}{2^\ell} \sum_{k=0}^{\ell} \binom{\ell}{k} F_m \circ \exp\left((\ell - 2k)\tau g \frac{d}{dx}\right).$$

The sampling matrix is found by choosing the canonical sampling scheme $S_{m,\ell} := F \circ \Phi_\varphi^{m+\ell}$, i. e.,

$$F \circ \Phi_\varphi^{m+\ell} = \frac{1}{2^\ell} \sum_{k=0}^{m+\ell} \binom{m+\ell}{k} F \circ \exp\left(\tau(m + \ell - 2k)B\right),$$

where $0 < |F(v_\lambda)| < \infty$. This condition is fulfilled by any point evaluation $Ff(x) = f(x_0)$ such that $0 < |v_\lambda(x_0)| < \infty$. Using the point evaluation and applying this sampling scheme to a function f for all $m \in \{0, \dots, N\}$ and $\ell \in \{0, \dots, M\}$, $N \geq M - 1$, we get

$$\mathcal{X}_{N,M} = \left[\frac{1}{2^{m+\ell}} \sum_{k=0}^{m+\ell} \binom{m+\ell}{k} f\left(G^{-1}((m + \ell - 2k)\tau + G(x_0))\right) \right]_{m,\ell=0}^{N,M},$$

where we used the known action of the generalized shift operator. This sampling matrix is sufficient for the reconstruction of f, since the sampling scheme is admissible by Lemma 2.3.4 and the point spectrum is, in dependence on v_λ, restricted such that

$$\lambda \to v_\lambda \quad \text{and} \quad \varphi(z) : \hat{\sigma}_P^\varphi(A) \to \mathbb{C}$$

are injective, because we only work with functions $f \in \mathcal{M}_M^\varphi(A)$. □

The restriction of the point spectrum $\hat{\sigma}_P(A)$ in Theorem 3.3.2 depends on the concrete choice of eigenfunctions v_λ which are used as atoms. In the next section we demonstrate how to employ the results on squared generalized shift operators to reconstruct sparse expansions into Chebychev polynomials of arbitrary degree. In particular, it is clarified in which way the restriction of the point spectrum corresponds to the choice of a certain family of atoms.

Sparse Expansions of Chebychev Polynomials

Besides the cosine expansions, that were the motivation for the generalized symmetric shift approach, we can also reconstruct sparse functions $f : [-1, 1] \to \mathbb{C}$ of the form

$$f(x) := \sum_{\lambda \in \Lambda} c_\lambda T_\lambda(x), \tag{3.9}$$

where $T_\lambda : [-1, 1] \to \mathbb{R}$ is defined by $T_\lambda(x) := \cos(\lambda \arccos(x))$, i. e., the Chebychev polynomials for $\lambda \in \mathbb{N}$. The linear coefficients $c_\lambda \in \mathbb{C}$ are arbitrary complex numbers. We focus in this example only on sparse expansions into Chebychev polynomials. The generator for these atoms is given by $A : C^2[-1, 1] \to C[-1, 1]$,

$$Af(x) := \left(1 - x^2\right) \frac{d^2}{dx^2} f(x) - x \frac{d}{dx} f(x),$$

which is of the form of a squared generalized shift as in equation (3.8). This can be seen if we define $B := -\sqrt{1-(\cdot)^2}\frac{d}{dx} : C^1[-1,1] \to C[-1,1]$; thus,

$$B^2 f(x) = \sqrt{1-x^2}\frac{d}{dx}\sqrt{1-x^2}\frac{d}{dx}f(x) = \left(1-x^2\right)\frac{d^2}{dx^2}f(x) - x\frac{d}{dx}f(x) = Af(x).$$

Therefore, we can apply Theorem 3.3.2.
The point spectrum of B is $\sigma_P(B) = \mathbb{C}$, because of Theorem 3.3.2. If we define

$$v_n(x) := \exp(in\arccos(x)) = \cos(n\arccos(x)) + i\sin(n\arccos(x)) = T_n(x) + i\sqrt{1-T_n^2(x)},$$

we observe that

$$Bv_n(x) = -in\, v_n(x),$$

i. e., v_n are eigenfunctions of B to the eigenvalue $-in$. This implies that

$$Av_n(x) = B^2 v_n(x) = -n^2 v_n(x).$$

In particular, if v_n is a solution to the differential equation $Af = -n^2 f$, the real part $\mathrm{Re}\,(v_n) = \cos(n\arccos(x))) = T_n(x)$ is also a solution, i. e., the Chebychev polynomial of degree n over the interval $[-1,1]$. Thus, we know that

$$AT_n = -n^2 T_n$$

and therefore A is a generator for the model of all finite linear combinations of Chebychev polynomials, because $\lambda := -n^2 \to T_n$ is obviously injective if we restrict the point spectrum to $\hat{\sigma}_P(A) := \left\{-n^2 \,\middle|\, n \in \mathbb{N}\right\}$.
Moreover, we want to apply GOProM for the reconstruction of functions f as in equation (3.9) to circumvent the problem of iterating the differential operator A directly. Therefore, we define

$$\varphi(z) := \sum_{k=0}^{\infty} \frac{\tau^{2k}}{(2k)!} z^k,$$

as in Lemma 3.3.1. Since $A = B^2$ we get also by Lemma 3.3.1 that the iteration operator $\Phi_\varphi = \frac{1}{2}\left(\exp(\tau B) + \exp(-\tau B)\right)$. This implies

$$\Phi_\varphi T_n = \sum_{k=0}^{\infty} \frac{\tau^{2k}}{(2k)!} A^k T_n = \sum_{k=0}^{\infty} \frac{\tau^{2k}}{(2k)!}(-n^2)^k T_n = \Phi(-n^2)T_n = \cos(\tau n)T_n,$$

for all $\lambda \in \hat{\sigma}_P(A)$, where $\tau \in \mathbb{R}_\times^+$ is a sampling parameter. Since we only have injectivity of the mapping

$$n \mapsto \Phi(-n^2) = \cos(\tau n)$$

for $n \in [0, \pi/\tau]$, the sampling parameter τ has to be bounded from above by $0 < \tau \leq \pi/\hat{n}$, where $\hat{n}$ is the maximum degree of the potentially active Chebychev polynomials. In other words, by using this approach, we are only able to reconstruct signals based on finite alphabets made of Chebychev polynomials of at most degree $\hat{n}$. Nevertheless, the following corollary provides a first reconstruction procedure for finite sparse Chebychev expansions.

Corollary 3.3.3 (Sparse Chebychev Expansions I)**.** *Let*

$$A := \left(-\sqrt{1-x^2}\frac{d}{dx}\right)^2 = (1-x^2)\frac{d^2}{dx^2} - x\frac{d}{dx}$$

be as above. With the restricted point spectrum $\hat{\sigma}_P(A) := \left\{-n^2 \,\middle|\, n \leq \hat{n} \in \mathbb{N}\right\}$ *it is a generator for all signals of the form*

$$f = \sum_{\lambda \in \Lambda} c_\lambda v_\lambda,$$

where $v_\lambda = v_{-n^2} = T_n$ *is the Chebychev polynomial* T_n *of degree n. Furthermore, let the iteration operator be defined as* $\Phi_\varphi := \sum_{k=0}^{\infty} \frac{\tau^{2k}}{(2k)!} A^k$*, where* $\tau \in (0, \pi/\hat{n}]$*, then a realizable sampling scheme is given by*

$$S_{m,\ell} = \frac{1}{2^\ell} \sum_{k=0}^{\ell} \binom{\ell}{k} F_m \circ \exp\left((\ell - 2k)\tau \sqrt{1-x^2}\frac{d}{dx}\right).$$

If we choose the canonical evaluation scheme $F_m = F \circ \Phi_\varphi^m$*, such that* $0 < |F(v_\lambda)| < \infty$*, the sampling scheme* $S_{m,\ell}$ *is also admissible. Using the point evaluation* $Ff(x) = f(1)$*, we get an admissible sampling matrix*

$$\mathcal{X}_{N,M} := \left[\frac{1}{2^{m+\ell}} \sum_{k=0}^{m+\ell} \binom{m+\ell}{k} f\left(\cos((m+\ell-2k)\tau)\right)\right]_{m,\ell=0}^{N,M},$$

$N \geq M - 1 \in \mathbb{N}$*, which is sufficient for the reconstruction of* f *by GOProM.*

Proof. From the discussion above it is clear that $A = B^2 = \left(-\sqrt{1-x^2}\frac{d}{dx}\right)^2$ is a generator for sparse Chebychev expansions and it explains why the restriction of the point spectrum to $\left\{-n^2 \,\middle|\, n \leq \hat{n} \in \mathbb{N}\right\}$ is necessary. The sampling scheme $S_{m,\ell}$ is derived by using Theorem 3.3.2 and the known exponential action of B, which is for all linear combinations of Chebychev polynomials and all $t \in \mathbb{R}$ given by Lemma 3.2.2

$$\exp\left(-t\sqrt{1-x^2}\frac{d}{dx}\right) f(x) = f\left(\cos(t + \arccos(x))\right),$$

where $g(x) = -\sqrt{1-x^2}$. Based on this sampling scheme and in connection with the canonical evaluation scheme $F_m := F \circ \Phi^m(A)$, where $F(f) = f(1)$, we get the sampling matrix $\mathcal{X}_{N,M}$ as given above. It is admissble, since for all $v_\lambda(x) = \cos(\lambda \arccos(x))$ it holds that

$$0 < |F(v_\lambda)| = 1 < \infty,$$

which is sufficient for the admissibility of the canonical sampling scheme by Corollary 2.3.4. □

This example for the application of the Prony method was first introduced in [39] and reformulated in the context of the Generalized Prony Method in [32]. In particular, in [32] the idea of the so-called symmetric shift operator

$$\frac{1}{2}\left(S_{\tau n} + S_{-\tau n}\right)$$

was introduced, where $S_{\tau n} f(x) = f(\cos(\arccos(x) + \tau n))$, with $\tau \in \mathbb{R}$ and $n \in \mathbb{N}$. It was shown that this operator has the Chebychev polynomials as eigenfunctions and afterwards the Generalized Prony Method was applied.
The symmetric shift operator corresponds to the decomposition of the iteration operator

$$\Phi_\varphi = \frac{1}{2}\left(\exp(\tau B) + \exp(-\tau B)\right),$$

as in Theorem 3.3.2 for the special case of $B = -\sqrt{1-x^2}\frac{d}{dx}$. In contrast to these former approaches, the derivation in this section started directly at the defining differential equation for the Chebychev polynomials and gave a constructive way of finding a suitable sampling scheme.

3.3.2 Integral Operator Approach

In the last section it was demonstrated that generators with a certain structure can be decomposed by a suitable mapping into operators with simpler exponential actions. In contrast, in this section we show how the same generator structure can be used to directly evaluate the exponential action.
We start with a general result concerning the exponential action of squared operators.

Lemma 3.3.4 (One-parametric Exponential Squared Operators)**.** *Let $A : V \to V$ be a linear operator on a vector space V over $\mathbb{R}$, such that it can be written as $A = B^2$, where $B : V \to V$ is also linear and has a non-empty point spectrum. Then A also has a non-empty point spectrum. Furthermore, let*

$$f(x) := \sum_{\lambda \in \Lambda} c_\lambda v_\lambda(x),$$

where v_λ is the eigenfunction of A to the eigenvalue $\lambda \in \Lambda$. Then the following equation holds

$$\exp(\tau A) f(x) = \frac{1}{2\sqrt{\pi\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\tau}\right) g(t,x)\,dt, \tag{3.10}$$

where $g(t,x) := \exp(-tB) f(x)$.

Proof. We first show that

$$\exp\left(x^2\right) = \frac{1}{\sqrt{4\pi}} \int_{\mathbb{R}} \exp\left(-\frac{\xi^2}{4}\right) \exp(-\xi x)\,d\xi$$

for all $x \in \mathbb{R}$. By using the exponential power series and the linearity of the integral it holds that

$$\begin{aligned}\frac{1}{\sqrt{4\pi}}\int_{\mathbb{R}} \exp\left(-\frac{\xi^2}{4}\right)\exp(-\xi x)d\xi &= \sum_{k=0}^{\infty}\frac{(-1)^k x^k}{k!}\frac{1}{\sqrt{4\pi}}\int_{\mathbb{R}}\exp\left(-\frac{\xi^2}{4}\right)\xi^k d\xi \\ &= \sum_{k=0}^{\infty}\frac{(-1)^k x^k}{k!}\mu_k(0,2)\,,\end{aligned}$$

where

$$\mu_{\tilde{k}}(0,2) = \begin{cases} 0\ ;\ \tilde{k} \in 2\mathbb{N}-1, \\ 2^{\tilde{k}/2}(\tilde{k}-1)!!\ ;\ \tilde{k} \in 2\mathbb{N}_0 \end{cases}$$

is the well known k$^{\text{th}}$ central moment of the normal distribution with variance $\sigma^2 = 2$, expectation value 0, and !! denotes the double faculty, i. e., in this case the product of all odd numbers smaller than $k-1$. By substituting $2k = \tilde{k} \in 2\mathbb{N}_0$ we get that

$$\mu_{2k}(0,2) = 2^k(2k-1)!! = \frac{(2k)!}{k!}$$

and since $\mu_{\tilde{k}} = 0$ for all odd indices, it follows for all even indices

$$\begin{aligned}\frac{1}{\sqrt{4\pi}}\int_{\mathbb{R}} \exp\left(-\frac{\xi^2}{4}\right)\exp(-\xi x)d\xi &= \sum_{k=0}^{\infty}\frac{(-1)^k x^k}{k!}\mu_k(0,2) = \sum_{k=0}^{\infty}\frac{(-1)^{2k}x^{2k}}{(2k)!}\mu_{2k}(0,2) \\ &= \sum_{k=0}^{\infty}\frac{x^{2k}}{(2k)!}\frac{(2k)!}{k!} = \sum_{k=0}^{\infty}\frac{\left(x^2\right)^k}{k!} = \exp\left(x^2\right).\end{aligned}$$

By defining $x := \sqrt{\tau}\,z$ and substituting $t := \xi\sqrt{\tau}$ we get

$$\exp\left(x^2\right) = \exp\left(\tau z^2\right) = \frac{1}{\sqrt{4\pi}}\int_{\mathbb{R}}\exp\left(-\frac{\xi^2}{4}\right)\exp\left(-\xi\sqrt{\tau}x\right)d\xi = \frac{1}{\sqrt{4\pi\tau}}\int_{\mathbb{R}}\exp\left(-\frac{t^2}{4\tau}\right)\exp(-tz)dt.$$

If we now formally substitute B for z, which is possible since we are in $\mathcal{M}(A)$, we get

$$\exp(\tau A)f = \exp(\tau B^2)f = \frac{1}{2\sqrt{\pi\tau}}\int_{\mathbb{R}}\exp\left(-\frac{t^2}{4\tau}\right)[\exp(-tB)f]\,dt$$

for all f as in the Lemma. This substitution is justified by the fact that z is constant with respect to the integral over all $t \in \mathbb{R}$. □

Remark 3.3.5. The result on the representation of exponentiated square operators was also used in [12] for the solution of Cauchy problems concerning certain exponential operators but it was surely known long before.

The parameter $4\tau = 2\sigma^2$, where σ is the standard deviation of the normal distribution, describes the width of the Gaussian kernel. Thus, we gather information about the signal f over a range of different scales.

It must be emphasized that one has to be very careful with this kind of equation as often in the case of exponential operators. Intrinsically we always think of $\exp(\tau A)$ as the formal exponential power series

$$\exp(\tau A) := \sum_{k=0}^{\infty} \frac{\tau}{k!} A^k,$$

where convergence can only be guaranteed if A is bounded. In contrast, if A is unbounded $\exp(\tau A)$ is only defined for certain functions. Even worse, there exist functions such that the integral on the right side of equation (3.10) is defined, but the formal power series on the left side does not converge for any τ except $\tau = 0$.
Fortunately, we have already seen that the generator and the iteration operator are always well-defined for all $f \in \mathcal{M}(A)$. In particular, it is sufficient that the restricted point spectrum $\hat{\sigma}(A)$ is in the region of convergence of φ. This is always the case, because the exponential power series converges on the whole complex plane. In turn, we have no essential restriction concerning the convergence of both sides of equation (3.10) and therefore the proof above is complete.
Using Lemma 3.3.4 and the results on generalized shift operators, we can extend the set of applications of GEProM to the class of generators known from Theorem 3.3.2, but with a different sampling scheme.

Corollary 3.3.6 (Squared Generalized Shifts II)**.** *Let the operator $A : C^2[x_0, x_1] \to C[x_0, x_1]$, $x_0 < x_1 \in \overline{\mathbb{R}}$, be defined as*

$$A := \left(g\frac{d}{dx}\right)^2 = g^2\frac{d^2}{dx^2} + gg'\frac{d}{dx},$$

where $g \in C^1[x_0, x_1]$, then A has the point spectrum $\sigma_P(A) = \mathbb{C}$. A realizable exponential sampling scheme $S_{m,\ell} := F_m \circ \exp(\tau\ell A)$, for $\tau \in \mathbb{R}_{\neq 0}$, is given by

$$S_{m,\ell}(\,\cdot\,) := \frac{1}{2\sqrt{\pi\ell\tau}} F_m\left(\int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\ell\tau}\right)\exp\left(-tg\frac{d}{dx}\right)(\,\cdot\,)dt\right),$$

with respect to an admissible evaluation scheme $F_m : C^2[x_0, x_1] \to \mathbb{C}$. Using the canonical evaluation scheme $F_m := F \circ \exp(\tau m A)$ and the point evaluation $F(f) = f(x_0)$, $x_0 \in \mathbb{R}$ such that $0 < |v_\lambda(x_0)| < \infty$ for all $\lambda \in \hat{\sigma}_P(A)$,

$$\mathcal{X}_{N,M}(f) := \left[\frac{1}{2\sqrt{\pi(m+\ell)h}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4(m+\ell)h}\right) f\left(G^{-1}\left(-t + G(x_0)\right)\right) dt\right]_{m,\ell=0}^{N,M}$$

is admissible, and sufficient for the reconstruction of all $f \in \mathcal{M}_M(A)$.

Proof. The argumentation for the spectrum of A is the same as in Theorem 3.3.2. The sampling scheme $S_{m,\ell}$ is a direct consequence of Lemma 3.3.4 and the composition with an admissible

sampling scheme F_m. If we choose $F_m := F \circ \exp(\tau m A)$ and $F(f) := f(x_0)$ we get the samplings

$$\begin{aligned} S_{m,\ell}(f) &= F\left[\exp(\tau(m+l)A)f\right] \\ &= \frac{1}{2\sqrt{\pi(m+\ell)\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4(m+\ell)\tau}\right) F\left[\exp\left(-\tau(m+\ell)g\frac{d}{dx}\right) f\right] dt \\ &= \frac{1}{2\sqrt{\pi(m+\ell)\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4(m+\ell)\tau}\right) f\left(G^{-1}(-t+G(x_0)\right) dt. \end{aligned}$$

This leads to an admissible sampling matrix if we choose x_0 such that $0 < |v_\lambda(x_0)| < \infty$, for all λ. □

Sparse Cosine Expansions

One of the first examples of this thesis was the embedding of the sparse cosine expansions into the GOProM framework. In the last section we have seen that this type of expansion is actually only a special case of GOProM, based on linear combinations of generalized shift operators. In a second approach it was shown that we can also use certain types of integral operators to reconstruct sparse expansions into eigenfunctions of squared generalized shift operators. In this section we revisit the cosine expansions under this new integral operator approach.
Therefore, let

$$A := \frac{d^2}{dx^2} : C^2[x_0, x_1] \to C[x_0, x_1]$$

be the generator for the space of all finite linear combinations of cosine as well as sine functions. In this example we focus on sparse cosine expansions, i. e., functions $f : \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) := \sum_{\lambda \in \Lambda} c_\lambda v_\lambda(x),$$

where the atoms are defined by $v_\lambda(x) := \cos\left(\sqrt{\lambda}x\right) : \mathbb{R} \to \mathbb{R}$ with $\omega := \sqrt{\lambda} \in (0, 2\pi T]$ and $T \in \mathbb{R}^+$. Using Lemma 3.3.4 we know the action of $\exp(\tau A)$, which is

$$\exp(tA)f = \frac{1}{2\sqrt{\pi\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\tau}\right) \exp\left(-t\frac{d}{dx}\right) f dt = \frac{1}{2\sqrt{\pi\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\tau}\right) f(\cdot - t) dt, \quad (3.11)$$

and thus it is realizable. In this special case it is obviously the convolution of f with a Gaussian kernel. This kind of transformation is called Weierstrass transform [4] or Gaussian filter and denoted by

$$\mathcal{W}_\tau f := \frac{1}{2\sqrt{\pi\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\tau}\right) f(\cdot - t) dt.$$

Although Corollary 3.3.9 can directly be derived from Theorem 3.3.6, we first give an additional lemma on the eigenfunctions of $\mathcal{W}_\tau$.

Lemma 3.3.7 (eigenfunctions of the Weierstrass Transform)**.** *Let be* $v_\lambda \in \{\cos(\omega x), \sin(\omega x), \exp(\omega x)\}$, $\lambda \in \{-\omega^2, \omega^2\}$. *For all* $\tau \in \mathbb{R}_\times^+$ *it holds that*

$$\forall\ \omega \in \mathbb{C}\ :\ \mathcal{W}_\tau \exp(\omega x) = \exp(\omega^2\tau)\exp(\omega x), \tag{3.12}$$

$$\forall\ \omega \in \mathbb{R}\ :\ \mathcal{W}_\tau \cos(\omega x) = \exp(-\omega^2\tau)\cos(\omega x), \tag{3.13}$$

$$\forall\ \omega \in \mathbb{R}\ :\ \mathcal{W}_\tau \sin(\omega x) = \exp(-\omega^2\tau)\sin(\omega x). \tag{3.14}$$

Proof. We first prove the eigenfunction property for $v_\lambda(x) := \exp(\omega x)$: $A = \frac{d^2}{dx^2}$, and $\lambda := \omega^2$,

$$\begin{aligned}
\exp(\tau A)v_\lambda(x) &= \exp\left(\tau\frac{d^2}{dx^2}\right)\exp(\omega x) = \frac{1}{2\sqrt{\pi\tau}}\int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\tau}\right)\exp\left(-t\frac{d}{dx}\right)\exp(\omega x)dt \\
&= \frac{1}{2\sqrt{\pi\tau}}\int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\tau}\right)\exp(\omega(x-t))dt = \frac{1}{2\sqrt{\pi\tau}}\int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\tau}\right)\exp(-\omega t)dt\ v_\lambda(x) \\
&= \frac{1}{2\sqrt{\pi\tau}}\int_{\mathbb{R}} \exp\left(-\left[\left(\frac{t}{2\sqrt{\tau}}\right)^2 + 2\frac{t}{2\sqrt{\tau}}\omega\sqrt{\tau} + \omega^2\tau\right]\right)dt\exp\left(\omega^2\tau\right)v_\lambda(x) \\
&= \frac{1}{2\sqrt{\pi\tau}}\int_{\mathbb{R}} \exp\left(-\left[\frac{t}{2\sqrt{\tau}} + \omega\sqrt{\tau}\right]^2\right)dt\exp\left(\omega^2\tau\right)v_\lambda(x) \\
&= \frac{1}{2\sqrt{\pi\tau}}\int_{\mathbb{R}} \exp\left(-\frac{1}{2(2\tau)}\left[t + 2\omega\tau\right]^2\right)dt\exp\left(\omega^2\tau\right)v_\lambda(x)
\end{aligned}$$

The last equation implies that the integral, seen as integral over the density of a normal distribution, has variance $\sigma^2 := 2\tau$ which implies that $\tau = \sigma^2/2$. The constant in front of the integral is

$$\frac{1}{2\sqrt{\pi\tau}} = \frac{1}{\sqrt{4\pi\sigma^2/2}} = \frac{1}{\sqrt{2\pi\sigma^2}},$$

which is exactly the normalization constant of the normal distribution, and thus the integral and the constant multiply to one, i. e., $\exp(\tau A)v_\lambda = \exp\left(\omega^2\tau\right)v_\lambda(x)$.
The cosine and sine case can be proven by Euler's formula, which holds for all $\omega \in \mathbb{R}$, and the linearity of the Weierstrass transform

$$\begin{aligned}
\mathcal{W}_\tau \cos(\omega x) &= \frac{1}{2}\left(\mathcal{W}_\tau \exp(i\omega x) + \mathcal{W}_\tau \exp(-i\omega x)\right) \\
&= \frac{1}{2}\left(\exp(-\omega^2\tau)\exp(i\omega x) + \exp(-\omega^2\tau)\exp(-i\omega x)\right) = \exp(-\omega^2\tau)\cos(\omega x).
\end{aligned}$$

The eigenfunction property for the sinus can be derived analogously. □

Remark 3.3.8 (Sparse Hermitian Expansions)**.** Another interesting property of the Weierstrass transform is that it maps Hermite polynomials H_n defined by the differential equation

$$\frac{d^2}{dx^2}H_n(x) - 2x\frac{d}{dx}He_n(x) = -2n\,H_n(x)$$

to monomials of degree n [4], i. e., $\mathcal{W}_{0.5}\left(H_n(x)\right) = x^n$. Thus, we can employ Theorem 3.2.4 with $g(x) = x$ and $a(x) = 0$ to reconstruct sparse expansions into Hermite polynomials. For this

reason we first apply the Weierstrass transform as above to get

$$\mathcal{W}_{0.5}\left(\sum_{j=0}^{M} c_{n_j} H_{n_j}(x)\right) = \sum_{j=0}^{M} c_{n_j} x^{n_j},$$

i. e., an expansion of monomials of arbitrary degree and in turn we use Theorem 3.2.4 with $g(x) = x$ and $a(x) = 0$ to reconstruct it.

Having the Lemma 3.3.7 at hand there are actually two ways to derive realizable and admissible sampling matrices for the reconstruction of sinus, cosine or exponential expansions.
Either, we can simply use Theorem 3.3.6 choosing $g(x) \equiv 1$ and restricting the point spectrum accordingly, or we use Lemma 3.3.7 in connection with the Generalized Prony Method from Peter & Plonka described in Corollary 2.3.6 to get an admissible and realizable sampling scheme. We choose the first approach for the cosine example yielding the following Corollary.

Corollary 3.3.9 (Sparse Cosine Expansions II)**.** *Let the operator $A : C^2(\mathbb{R}) \to C(\mathbb{R})$ be defined as $A := \frac{d^2}{dx^2}$, then A has the point spectrum $\sigma_P(A) = \mathbb{C}$. We define the restricted point spectrum as $\hat{\sigma}_P(A) := (-\infty, 0]$. Let the signals $f : \mathbb{R} \to \mathbb{C}$ be of the form*

$$f(x) = \sum_{\omega \in \Omega_f} c_\omega \cos(\omega x),$$

where $\Omega_f \subset \mathbb{R}^+$ is finite and the eigenvalues $\lambda_\omega = -\omega^2 \in \hat{\sigma}_P(A)$. A realizable exponential sampling scheme $S_{m,\ell} := F_m \circ \exp(\tau \ell A)$ to recover f is given by

$$S_{m,\ell} := F_m \circ \mathcal{W}_{\tau\ell},$$

with respect to an evaluation scheme $F_m : C^2(\mathbb{R}) \to \mathbb{C}$ and the Weierstrass transform $\mathcal{W}_{\tau\ell}$. Using the canonical evaluation scheme $F_m := F \circ \exp(\tau m A)$ and the point evaluation $F(f) = f(0)$, the sampling matrix

$$\mathcal{X}_{N,M}(f) := \left[\frac{1}{2\sqrt{\pi(m+\ell)\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4(m+\ell)\tau}\right) f(t)\, dt\right]_{m,\ell=0}^{N,M}$$

is admissible, and sufficient for the reconstruction of all f.

Proof. By equation (3.11) we know that the ℓ^{th} power of the exponential iteration operator $\Phi_\varphi := \exp(\tau A)$, of the generator $A = \frac{d^2}{dx^2}$, is given by

$$\Phi_\varphi^\ell = \exp(\tau \ell A) = \mathcal{W}_{\tau\ell},$$

$\ell \in \mathbb{N}$, and $\tau \in \mathbb{R}_{\neq 0}$. Thus, using an evaluation scheme $F_m : C^2(\mathbb{R}) \to \mathbb{C}$ we get a sampling scheme $S_{m,\ell} := F_m \circ \mathcal{W}_{\tau\ell}$. Furthermore, since we have

$$\forall\ \omega \in \mathbb{R} \ :\ \mathcal{W}_\tau \cos(\omega x) = \exp(-\omega^2 \tau) \cos(\omega x),$$

the relation $\lambda \mapsto \omega$, with $\lambda := -\omega^2$, where λ is the eigenvalue to $v_\lambda(x) = \cos(\omega x)$, is injective for all $\omega \in \mathbb{R}^+$. This is a consequence of the bijectivity of the exponential function on $\mathbb{R}$ and $\omega = \sqrt{-\lambda}$. If we now choose the canonical sampling scheme $F_m := F \circ \Phi_\varphi^m$, with $F(f) = f(0)$, we get the sampling matrix $\mathcal{X}_{N,M} = [S_{m,\ell}]_{m,\ell=0}^{N,M}$ by

$$S_{m,\ell} = F_m \circ \Phi_\varphi^\ell := F \circ \Phi_\varphi^{m+\ell} = F \circ \mathcal{W}_{\tau(m+\ell)}$$

applied to a signal f and evaluated at zero, which gives by equation (3.11) the samples

$$S_{m,\ell}(f) = \frac{1}{2\sqrt{\pi(m+\ell)\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4(m+\ell)\tau}\right) f(-t)dt,$$

but $f(-x) = f(x)$, since the cosine is an even function. Thus, we yield the sampling matrix as in the corollary above. This sampling matrix is admissible, because

$$0 < |F(\cos(\omega \cdot))| = 1 < \infty.$$

□

The major difference to the former results on the reconstruction of cosine expansions is that we have no longer a band limitation. This advantage is accompanied with slightly more difficult sampling values. We have to evaluate the Weierstrass transform of the signal over the whole real line, which of course can only be done approximatively in real applications; but, since the Gaussian term in the Weierstrass transform decays exponentially and we integrate bounded signals, this can be done with high accurracy even for compact sampling intervals.

Sparse Chebychev Expansions

As we have already seen, also the Chebychev polynomials can be written as eigenfunctions of the squared generator of a certain generalized shift. In this section we apply the integral operator approach to reconstruct sparse expansions of those atoms.

Corollary 3.3.10 (Sparse Chebychev Expansions II)**.** *Let*

$$A := \left(-\sqrt{1-x^2}\frac{d}{dx}\right)^2 = (1-x^2)\frac{d^2}{dx^2} - x\frac{d}{dx}$$

be given as in Corollary 3.3.3. With the restricted point spectrum $\hat{\sigma}_P(A) := \left\{-n^2 \,\middle|\, n \in \mathbb{N}\right\}$ *the operator A is a generator for all signals of the form*

$$f = \sum_{\lambda_n \in \Lambda} c_{\lambda_n} v_{\lambda_n},$$

where $v_{\lambda_n} = v_{-n^2} = T_n$ *is the Chebychev polynomial* $T_n : [-1,1] \to \mathbb{R}$ *of degree n. A sampling scheme is given by*

$$S_{m,\ell} := F_m \circ \exp(\tau\ell A) = \frac{1}{\sqrt{4\pi\tau\ell}} F_m\left(\int_{\mathbb{R}} \exp\left(-\frac{1}{4\tau\ell}t^2\right) f(\cos(-t + \arccos(\cdot)))dt\right),$$

where $F_m : C^2[-1,1] \to \mathbb{C}$ *is an evaluation scheme. Using the canonical evaluation scheme* $F_m = F \circ \exp(\tau m A)$*, where* $F(f) = f(1)$*, we get the admissible sampling matrix*

$$\mathcal{X}_{N,M}(f) := \left[\frac{1}{2\sqrt{\pi(m+\ell)\tau}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{4(m+\ell)\tau}\right) f\left(\cos(t)\right) dt\right]_{m,\ell=0}^{N,M}.$$

Proof. We first prove that the Chebychev polynomials are eigenfunctions of $\exp(yA)$, for all $y \in \mathbb{R}$,

$$\begin{aligned}
\exp(yA)\, T_n(x) &= \frac{1}{\sqrt{4\pi y}} \int_{\mathbb{R}} \exp\left(-\frac{1}{4y}t^2\right) \exp\left(-t\left(-\sqrt{1-x^2}\right)\frac{d}{dx}\right) T_n(x) dt \\
&= \frac{1}{\sqrt{4\pi y}} \int_{\mathbb{R}} \exp\left(-\frac{1}{4y}t^2\right) T_n\left(\cos(-t + \arccos(x))\right) dt \\
&= \frac{1}{\sqrt{4\pi y}} \int_{\mathbb{R}} \exp\left(-\frac{1}{4y}t^2\right) \cos(-nt + n\arccos(x)) dt \\
&= \frac{1}{\sqrt{4\pi y}} \int_{\mathbb{R}} \exp\left(-\frac{1}{4y}t^2\right) \left(\cos(nt)\cos(n\arccos(x)) + \sin(nt)\sin(n\arccos(x))\right) dt \\
&= \frac{1}{\sqrt{4\pi y}} \int_{\mathbb{R}} \exp\left(-\frac{1}{4y}t^2\right) \cos(nt) dt \;\; \cos(n\arccos(x)) \\
&= \mathcal{W}_y(\cos(n\,\cdot))T_n(x) = \exp(-n^2 y) T_n(x).
\end{aligned}$$

The first line is a consequence of Lemma 3.3.4 and $A = \left(-\sqrt{1-x^2}\frac{d}{dx}\right)^2$. The second line is given by the action of the generalized shift operator $\exp\left(-t\left(-\sqrt{1-x^2}\right)\frac{d}{dx}\right)$, which is known from Lemma 3.2.2. Afterwards, we use the identity of the Chebychev polynomials $T_n(\cos(x)) = \cos(nx)$, that holds for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$, and apply the addition theorems to decompose the remaining sum under the cosine. Since the integral over the sine term is zero, we finally get the Weierstrass transform of the cosine times the Chebychev polynomial T_n. The last equation is a consequence of Lemma 3.3.7.
This calculation shows that the Chebychev polynomials are actually eigenfunctions of the iteration operator $\exp(yA)$ and in turn a sampling scheme is given by

$$S_{m,\ell} := F_m \circ \exp(\tau \ell A),$$

where $\tau \in \mathbb{R}^+_{\neq 0}$ and $\ell \in \mathbb{N}_0$. We explicitly allow $\ell = 0$ because it is well-defined on the eigenfunctions T_n as the point evaluation at zero. This is a consequence of the fact that the Gaussian term approximates the delta distribution for $y \to 0$. The sampling scheme and the sampling matrix are achieved by using Theorem 3.3.6 with $g = -\sqrt{1-x^2}$ and $x_0 = 1$. This sampling matrix is admissible since $0 < |T_n(1)| = 1 < \infty$. □

Since this is a GEProM approach with canonical sampling scheme, the sampling matrix has already Hankel structure. Moreover, the exponential function is bijective on the whole real line and in turn no restriction on the point spectrum has to be imposed. This implies that we can even use infinite dictionaries of Chebychev polynomials. Of course, as in the cosine case, it may be more difficult to realize the sampling scheme since we have no longer compact sampling support, but with the same arguments as above it is possible to approximate the integral over a compact interval for an arbitrary given accuracy. The last theorem of this section concerns a combination of the results on exponential operators as presented before.

Theorem 3.3.11 (Quadratic Polynomials of Generators)**.** *Let $A : V \to V$ be a generator defined as*

$$A := \alpha_2 C^2 + \alpha_1 C + \alpha_0 I$$

with respect to a second operator $C : V \to V$ such that the exponential iteration operator $\exp(\tau C)$, $\tau \in \mathbb{R}^+_\times$, is well-defined and realizable for $\alpha_i \in \mathbb{C}$. The sampling scheme

$$S_{m,\ell} := F_m \circ \exp(\tau \ell A) = \exp(\alpha_0)\, F_m \circ \exp\left(\tau \alpha_2 C^2\right) \exp\left(\tau \alpha_1 C\right)$$

is realizable and also admissible as long as $F_m : V \to \mathbb{C}$ is an admissible evaluation scheme. For all $f \in \mathcal{M}(A)$ and the canonical sampling scheme $F_m := F \circ \exp(\tau m A)$

$$S_{m,\ell} := \frac{e^{\alpha_0}}{2\sqrt{\pi \alpha_2 (m+l)\tau}} F\left(\int_{\mathbb{R}} \exp\left(-\frac{t^2}{4\alpha_2(m+l)h}\right) \exp\left(-t(\alpha_1+\alpha_2)(m+l)\tau C\right) f dt\right).$$

is admissible if $0 < |F(v_\lambda)| < \infty$.

Proof. Assume that we are interested in the reconstruction of $f \in \mathcal{M}_M(A)$, i. e., linear com-

binations of atoms defined as eigenfunctions of A. We have seen how the exponential action of every single operator in this sum can be iterated. The only remaining part is to show that $\exp(\tau A)$ factorizes so that we can apply the single iterative action successively. But this is trivially given, since all $\alpha_i \in \mathbb{C}$ are constant and therefore the summation terms commute, which implies that we can factorize the exponential iteration operator. Since it is assumed that we know the action of $\exp(\tau\alpha_1 C)$ from the discussion above, we also know the action of $\exp\left(\tau\alpha_2 C^2\right)$. The only action we have to find is $\exp(\alpha_0 \mathrm{I})$, which is

$$\exp(\alpha_0 \mathrm{I}) f = \sum_{k=0}^{\infty} \frac{\alpha_0}{k!} I^k f = \left(\sum_{k=0}^{\infty} \frac{\alpha_0}{k!} \right) f = \exp(\alpha_0) f,$$

i. e., only the multiplication operator. □

At this point we have finished the central part of this thesis. We derived the currently most general form of the Prony method in Theorem 2.2.13 and gave several examples for its application. In particular, we demonstrated that there is a whole subclass of algorithms, namely GEProM, which includes already almost all existing examples. Those not included were shown to be of GOProM type.

The central mathematical object of every GOProM approach is the sampling matrix $\mathcal{X}_{N,M}$. The sampling scheme $S_{m,\ell}$ was only a mean to construct these matrices and most of the time we simply assumed the admissibility of $S_{m,\ell}$ or it was given because we used the canonical evaluation scheme in connection with a suitable point evaluation.

The next chapter is therefore devoted to a discussion of several aspects of the realizability and admissibility of the sampling scheme. In particular, we give conditions under which almost every differential generator can be used to generate a realizable sampling scheme.

CHAPTER 4

Dual Sampling Schemes

> What I cannot create, I do not understand.
>
> Richard Feynman

So far, GOProM has been theoretically established and first examples have demonstrated the large freedom in constructing old and new sampling schemes and matrices for several atomic expansions. Thereby, it has become clear that the key element to new applicable examples is always a realizable sampling scheme; the mere theoretical construction of the same is often quite simple. Given a certain atomic expansion, the reconstruction problem can only be considered to be completely solved if we have found such a scheme, which gives us the necessary linear measurements to construct the sampling matrix.
In this section different ways of the data acquisition are presented, concerning the difficulties arising from the realizability of the general sampling scheme as in Definition 2.2.11,

$$S_{m,\ell} := F_m\left(\Phi_\varphi^\ell f\right),$$

which is the central object of this chapter. Choosing a suitable mapping Φ to get a realizable sampling scheme often includes a restriction of the point spectrum that we are able to recover. Examples of such restrictions are finite a priori known alphabets like in the case of the Chebychev polynomials. Therefore, in some instances it would be more advantageous to use the generator A directly as the iteratation operator, for example, as in the Generalized Prony Method of Peter & Plonka [33]. The basic problem of this approach in applications where the generator A is a differential operator is that one has to sample derivatives of very high orders. Therefore, in this section we demonstrate how the evaluation scheme F_m can be employed to simplify the sampling process by using dual sampling schemes. This approach enables us to cope, at least theoretically, with actually all linear differential operators which are used as generators. Moreover, we can combine it with GEProM to construct new realizable variants of this method; for instance, if the direct evaluation of the sampling scheme $S_{m,\ell} = \exp(\tau(m+\ell)A)$ is not possible, the dual approach can help to find realizable versions of this scheme. This approach combines the advantages of GEProM, above all the reproduction of the classical Prony problem as seen in (3.1), with the flexibility of the dual approach developed in this section.

4.1 Dual Sampling

The ideas for the dual data acquisition are explained best by examples, but before we start with the derivation of dual sampling schemes for linear differential operators, we need some notation concerning dual spaces. Throughout this chapter we assume that the considered vector spaces are normed.

Definition 4.1.1 (Dual Space)**.** Let V be a vector space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The space of all linear functionals

$$V^* := \{F : V \to \mathbb{K} \mid \ F \text{ linear } \}$$

is called the *dual space*. The natural pairing of an element $f \in V$ and $\phi \in V^*$ is denoted by

$$\langle \cdot, \cdot \rangle : V \times V^* \to \mathbb{K},$$

which is a linear form.

This implies that we can write the evaluation scheme as

$$F_m := \langle \cdot, \phi_m \rangle : V \to \mathbb{C}$$

with respect to a sequence $\phi_m \in V^*$. We emphasize that we do not need to know the complete dual space for the purpose of this work. We only need suitable elements of the dual space that fulfill some conditions explained later on, so we can use them as representations of the evaluation functionals. Moreover, in the concrete instances given below we always use integrals with suitable kernels as evaluation functionals. Nevertheless, the only properties we need to define the general dual sampling scheme is the linearity of the evaluation scheme, i. e., it has to be a linear form with respect to some elements in the dual, and the existence of an explicitely constructable adjoint operator.

Definition 4.1.2 ($\mathcal{M}(A)$-Adjoint Operator)**.** Given a linear operator $A : V \to V$ on a linear space V with dual V^*, a second operator $B : V^* \to V^*$ is called an *adjoint operator* for A with respect to $\phi \in V^*$ if for all $f \in V$ and all $\phi \in V^*$ holds that

$$\langle Af, \phi \rangle = \langle f, B\phi \rangle ,$$

where $\langle \cdot, \cdot \rangle$ is the natural pairing on V and V^*. We denote the adjoint B by A^*.

Although an adjoint of A is usually only denoted by A^* if it is unique, in a slight abuse of notation we always use this symbol for the adjoint operators. This is justified because for every sampling scheme we will choose a certain evaluation scheme in advance and calculate an explicit adjoint for the iteration operator.
With these two definitions at hand we are able to define a dual sampling scheme that will be used to derive sampling schemes for linear differential operators.

Theorem 4.1.3 (Dual Sampling Scheme)**.** *Let $A : V \to V$ be a generator of the signal model $\mathcal{M}_M(A)$ and V a linear space with dual V^*. Further, let $\Phi_\varphi : \mathcal{M}(A) \to \mathcal{M}(A)$ be a corresponding iteration operator with an adjoint*

$$\Phi_\varphi^* : (\mathcal{M}(A))^* \to (\mathcal{M}(A))^*,$$

and $F_m : \mathcal{M}(A) \to \mathbb{C}$ an admissible evaluation scheme induced by $\phi_m \in (\mathcal{M}(A))^*$ such that

$$\left\langle \Phi_\varphi^\ell \cdot, \phi_m \right\rangle = \left\langle \cdot, (\Phi_\varphi^*)^\ell \phi_m \right\rangle$$

for all $\ell \in \mathbb{N}$, $\ell \leq M \in \mathbb{N}$, and $m \in \mathbb{N}$, $m \leq N \in \mathbb{N}$, $N \geq M - 1$. Then

$$S_{m,\ell} = F_m \circ \Phi_\varphi^\ell = \left\langle \Phi_\varphi^\ell \cdot, \phi_m \right\rangle = \left\langle \cdot, (\Phi_\varphi^*)^\ell \phi_m \right\rangle$$

is an admissible and realizable sampling scheme. We call $\left\langle \cdot, (\Phi_\varphi^*)^\ell \phi_m \right\rangle$ a dual sampling scheme*. The corresponding sampling matrix for all $f \in \mathcal{M}_M(A)$ is*

$$\mathcal{X}_{N,M}(f) = \left[\left\langle f, (\Phi_\varphi^*)^\ell \phi_m \right\rangle \right]_{m,\ell}^{N,M},$$

where $N \geq M - 1$, $N \in \mathbb{N}$.

Proof. The admissibility of the dual sampling scheme is directly inherited from the admissibility of the evaluation scheme F_m, since

$$S_{m,\ell} = F_m \circ \Phi_\varphi^\ell = \left\langle \Phi_\varphi^\ell \cdot, \phi_m \right\rangle = \left\langle \cdot, (\Phi_\varphi^*)^\ell \phi_m \right\rangle$$

coincides with the dual scheme for all $f \in \mathcal{M}_M(A)$. The dual sampling scheme is realizable because $\left\langle \cdot, (\Phi_\varphi^*)^\ell \phi_m \right\rangle$ is a given sequence of kernels and therefore $F_{m,\ell}(f) := \left\langle f, (\Phi_\varphi^*)^\ell \phi_m \right\rangle$ are linear measurements of the signal itself. □

Since we are only able to calculate an explicit adjoint for the iteration operator Φ_φ with respect to a certain evaluation scheme F_m, they are no longer as independent as before. For the sampling schemes in GOProM and GEProM the only connection between evaluation scheme and iteration operator was the admissibility condition based on the eigenfunctions of A as given in Definition 2.2.7. In the dual setting they are connected more directly, since we can only define an adjoint iteration operator for GOProM with respect to a certain evaluation scheme or in certain cases for classes of these schemes.

Corollary 4.1.4. *Let $A : V \to V$ be a generator of the signal model $\mathcal{M}_M(A)$ and V a linear space with dual V^*. Further, let $\Phi_\varphi : \mathcal{M}(A) \to \mathcal{M}(A)$ be a corresponding iteration operator with an adjoint*

$$\Phi_\varphi^* : (\mathcal{M}(A))^* \to (\mathcal{M}(A))^* .$$

Moreover, let $F_m : V \to \mathbb{C}$ be an evaluation scheme induced by $\phi \in (\mathcal{M}(A))^*$ and

$$\forall\, \lambda \in \hat{\sigma}_P^\varphi(A) \;:\; 0 < |F(v_\lambda)| = |\langle v_\lambda, \phi\rangle| < \infty$$

such that

$$\left\langle \Phi_\varphi^{m+\ell} \cdot , \phi \right\rangle = \left\langle \cdot , (\Phi_\varphi^*)^{m+\ell} \phi \right\rangle$$

for all $m+\ell \leq (N+M-1) \in \mathbb{N}$, $N \geq M-1$. Then, the following sampling scheme is admissible

$$S_{m,\ell} := F \circ \Phi_\varphi^{m+\ell} = \left\langle \cdot , (\Phi_\varphi^*)^{m+\ell} \phi_0 \right\rangle = \langle \cdot , \phi_{m,\ell} \rangle ,$$

where $\phi_{m,\ell} := (\Phi_\varphi^*)^{m+\ell} \phi_0$.

Proof. This is a reformulation of the results in Theorem 2.3.4 using the admissibility of the canonical sampling scheme in terms of the adjoint iteration operator. It is sufficient to assume the admissibility condition for $\lambda \in \hat{\sigma}_P^\varphi(A)$, i. e., the restricted point spectrum, such that the mappings $\lambda \to v_\lambda$ and $\lambda \to \varphi(\lambda)$ are both injective. Thus, we can uniquely identify the active eigenvalues and in turn the active eigenfunctions by GOProM. □

Although we basically only need the linearity of the evaluation scheme, the assumption that V is a Hilbert space, i. e., the dual notation of the evaluation functional can be chosen to be an inner product, delivers a geometrical interpretation of the admissibility condition

$$|F_m(v_\lambda)| = |\langle f, \phi_m \rangle| > 0.$$

The admissible functional kernels $\phi_m \in (\mathcal{M}(A))^*$ necessarily satisfy that they are not orthogonal to any atom $v_n \in \mathcal{M}(A)$. In particular, if $\{v_n\}_{n=0}^\infty := \{v_{\lambda_n}\}_{n=0}^\infty$, with $\lambda_n \in \hat{\sigma}_P^\varphi(A)$, is an orthonormal basis for the Hilbert space V, we can always construct an admissible kernel $\phi : \mathcal{M}(A) \to \mathbb{C}$ for the canonical dual sampling scheme as in Corollary 4.1.4 by

$$\phi(x) := \sum_{k=0}^{\infty} \beta_k v_k(x),$$

where $|\beta_n| > 0$ for all $n \in \mathbb{N}$ such that $\sum\limits_{k=0}^{\infty} |\beta_k|^2 < \infty$, because

$$\forall\, \lambda_n \in \hat{\sigma}_P^\varphi(A) :\; |F(v_n)| = |\langle v_n, \phi\rangle| = \left| \sum_{k=0}^{\infty} \beta_k \langle v_n, v_k \rangle \right| = |\beta_n| > 0.$$

As an example we can always think of spaces spanned by orthonormal polynomials, for instance the space of all square integrable functions over $[-1, 1]$ with respect to a certain inner

product such that there is an orthonormal polynomial basis.
In general, the convergence of ϕ demands that β_k has to be a null sequence and although it does not identically vanish by construction on any atom v_n, there is an infinite number of atoms such that the corresponding coefficients β_k are smaller than any a priori given bound. This is a clear issue with this type of kernels and it seems that we can only circumvent it by employing some information about the possible active atoms. In other words, we have to assume a finite alphabet again. If a finite alphabet $\mathcal{A}$ with cardinality M is given, we can simply take the null sequence

$$\varphi_k := \begin{cases} \alpha_k \neq 0 & v_k \in \mathcal{A} \\ \alpha_k = 0 & v_k \notin \mathcal{A}, \end{cases}$$

with arbitrary large coefficients α_k. This means we can choose

$$\phi(x) = \sum_{k=1}^{M} \alpha_k v_k(x),$$

i. e., a finite linear combination of the possibly active eigenfunctions with linear coefficients α_k bounded away from zero.
Furthermore, in the case of finite alphabets of orthonormal functions it is even possible to construct more general admissible evaluation schemes F_m, beyond the canonical one. This can be seen by the following calculation. Given a signal model $\mathcal{M}_M(A)$, such that the eigenfunctions v_λ are orthonormal for all $\lambda \in \hat{\sigma}_P(A)$, the set of all eigenvalues to eigenfunctions in $\mathcal{A}$, we have for

$$F_m = \langle \cdot, \phi_m \rangle := \left\langle \cdot, \sum_{k=1}^{M} \alpha_k^m v_k \right\rangle$$

and all $f \in \mathcal{M}_M(A)$, restricted to the finite alphabet $\mathcal{A}$, that

$$E_F := [F_m(v_n)]_{m,n=0}^{N,M} = [\langle v_n, \phi_m \rangle]_{m,n=0}^{N,M} = \left[\sum_{k=1}^{M} \alpha_k^m \langle v_n, v_k \rangle \right]_{m,n=0}^{N,M} = [\alpha_n^m]_{m,n=0}^{N,M},$$

where v_n are the currently active eigenfunctions, i. e., a subset of the alphabet $\mathcal{A}$. At this point we remember the exact definition for admissibility in Definition 2.2.7, namely, for a given signal model $\mathcal{M}_M(A)$ it must hold that

$$\forall N \in \mathbb{N} \, \forall \, \Lambda \subseteq \hat{\sigma}_P^{\varphi}(A) \text{ s.t. } |\Lambda| = \tilde{M} : \operatorname{rank}(E_F) = \tilde{M}. \tag{4.1}$$

Compared to the calculation above for finite alphabets this means that

$$\forall N \in \mathbb{N} \, \forall \, \Lambda \subseteq \hat{\sigma}_P(A) \text{ s.t. } |\Lambda| = \tilde{M} : \operatorname{rank}\left([\alpha_n^m]_{m,n=0}^{N,M} \right) = \tilde{M},$$

where $\hat{\sigma}_P(A)$ is as above the set of finitely many eigenvalues corresponding to the alphabet $\mathcal{A}$. If we assume that all eigenfunctions in $\mathcal{A}$ are active, the matrix $E_F = \hat{E}_F$, the biggest matrix possible, with respect to $M = |\mathcal{A}|$ has to have full rank. If we only take a finite subset of

$\mathcal{A}$ as active eigenfunctions, this corresponds to a submatrix of $\hat{E}_F$ made of the corresponding columns and all rows. This means, for admissibility of an evaluation scheme as above it is sufficient to find a matrix $\hat{E}_F$ such that every selection of columns is linear independent.

Remark 4.1.5. An interesting observation is the formal connection of the dual GOProM approach and the field of dynamical sampling [1, 2]. Dynamical sampling is a recently developed framework for the reconstruction of signals f that evolve over time by a given law, i. e., we can only observe samples of the form

$$\left[\left\langle B^\ell f, \phi_m \right\rangle\right]_{m,\ell\in\mathbb{N}},$$

where B is an operator that governs the evolution of f and ϕ_m is a possibly countable set of sampling kernels. If we now assume that the function f, which is to be reconstructed, is a sparse linear combination of eigenfunctions of the evolution operator B and ϕ_m constitutes an admissible sampling scheme, GOProM can be interpreted as a dynamical sampling method to reconstruct sparse initial states of the observed signals.

4.1.1 Linear Differential Operators

In the examples given so far we have often used the iteration operator Φ_φ together with the generalized Prony polynomials to establish the annihilation equations as well as to generate the evaluation scheme and called the resulting sampling scheme canonical. This was possible since this sampling scheme was already realizable, i. e., it was possible to write it down as linear functionals applied to the signal itself without using signal derivatives.
In this section we discuss the example of linear differential operators that are used as generators of sparse expansions into their eigenfunctions. In the previous chapters we already considered examples of such differential generators, for instance, the ordinary derivative $\frac{d}{dx}$ as generator of finite expansions into complex exponentials in Section 2.1 or $\left(\sqrt{1-x^2}\frac{d}{dx}\right)^2$ as generator of sparse linear combinations into Chebychev polynomials in Section 3.3. In contrast to these special cases, it is in general not always obvious how to find suitable mappings Φ to construct a simple iteration operator for differential generators. Thus, we want to apply the dual approach to get sampling schemes for all linear differential operators, that are defined as follows.

Definition 4.1.6 (Linear Differential Operator)**.** Given real-valued functions $\alpha_n \in C[x_0, x_1]$ we define $D_d : C^d[x_0, x_1] \to C[x_0, x_1]$ with

$$D_d f(x) := \sum_{n=0}^{d} \alpha_n(x) \frac{d^n}{dx^n} f(x),$$

for $f \in C^d[x_0, x_1]$ and call it a *linear differential operator* of order d, where $x_0, x_1 \in \mathbb{R}$.

The eigenfunctions of these differential operators, if the operator has eigenfunctions, are infinitely often differentiable and can be used to form new families of atoms. Vice versa, a possible first approach to derive a generator $A := D_d$ for a given smooth eigenfamily is to try

to deduce a defining differential operator, for example, as it was done in Section 3.2.2 for the Gaussian chirp model.
A direct use of such a differential operator as annihilation operator or for the generation of the evaluation scheme leads to non-realizable sampling schemes based on high signal derivatives. Thus, we have to find suitable elements in the dual space of $C^d[x_0, x_1]$ with respect to an admissible evaluation scheme

$$F_m : C^d[x_0, x_1] \to \mathbb{C},$$

to construct an admissible and realizable sampling scheme $S_{m,\ell}$. For all practical reasons and in order to avoid unnecessary technicalities, in this section we define the evaluation scheme F_m always by using a sufficiently smooth kernel function ϕ.
We first focus on canonical evaluation schemes, i. e., generated by the iteration operator itself, $\Phi_\varphi := A = D_d$. Therefore,

$$F_m := F \circ \Phi_\varphi^m = \langle D_d^m(\cdot), \phi \rangle.$$

The complexity of linear differential operators D_d makes it in general very difficult to find a mapping φ to get a realizable iteration operator $\varphi(D_d)$; remember, the choice $\Phi_\varphi := D_d$, where $\varphi(z) = z$, is not realizable in the sense of this thesis because we have to use signal derivatives. To circumvent this issue, we use the following classical result for differential operators to construct a very general adjoint operator concerning a certain evaluation functional F.

Lemma 4.1.7 (Formal Adjoint)**.** *Let $D_d : C^d[x_0, x_1] \to C[x_0, x_1]$, $x_0, x_1 \in \bar{R}$, be a linear differential operator as in Definition 4.1.6 with $\alpha_n \in C^d[x_0, x_1]$ real-valued for all integers $n \leq d \in \mathbb{N}_0$ and let $\phi \in C^d[x_0, x_1]$ such that*

$$\forall k \leq d-1 : \lim_{x \to x_0} \frac{d^k}{dx^k}\phi(x) = \lim_{x \to x_1} \frac{d^k}{dx^k}\phi(x) = 0,$$

then the formal adjoint *of D_d with respect to the linear functional $\langle \cdot, \phi \rangle := \int_{x_0}^{x_1} (\cdot)\phi(x)dx$ is*

$$D_d^* = \sum_{n=0}^{d} (-1)^n \frac{d^n}{dx^n}\alpha_n I = \sum_{n=0}^{d} (-1)^n \sum_{k=0}^{n} \binom{n}{k} \alpha_n^{(n-k)} \frac{d^k}{dx^k}.$$

Proof. First we show that $A^* := \sum_{n=0}^{d} B_n^*$ is an adjoint for $A = \sum_{n=0}^{d} B_n$, given that B_n^* is an adjoint for B_n. Let ϕ be an arbitrary element of $C^d[x_0, x_1]$, then

$$\langle Af, \phi \rangle = \left\langle \sum_{n=0}^{d} B_n f, \phi \right\rangle = \sum_{n=0}^{d} \langle B_n f, \phi \rangle = \sum_{n=0}^{d} \langle f, B_n^* \phi \rangle = \left\langle f, \sum_{n=0}^{d} B_n^* \phi \right\rangle = \langle f, A^* \phi \rangle .$$

This shows that A^* is a dual operator for A. We now define the differential generator as

$$D_d := \sum_{n=0}^{d} B_n \quad \text{with} \quad B_n := \alpha_n \frac{d^n}{dx^n}.$$

Since the adjunction is linear, we only need to calculate the dual of B_n. This can be done by the extended chain rule

$$\langle B_n f, \phi \rangle = \int_{x_0}^{x_1} (B_n f)(x)\phi(x)dx = \int_{x_0}^{x_1} \alpha_n(x) f^{(n)}(x)\phi(x)dx = \int_{x_0}^{x_1} f^{(n)}(x)u_n(x)dx,$$

where we substituted $u_n := \alpha_n \cdot \phi$. Integration by parts delivers

$$\langle B_n f, \phi \rangle = \int_{x_0}^{x_1} f^{(n)}(x)u_n(x)dx = \left[\sum_{k=0}^{n-1}(-1)^k f^{(n-1-k)}(x)u_n^{(k)}(x)\right]_{x_0}^{x_1} + (-1)^n \int_{x_0}^{x_1} f(x)u_n^{(n)}(x)dx.$$

Since we assumed that f and α_n are $d-$times continuously differentiable and the kernel and all its derivatives up to order d vanish at the boundaries, it follows by the Leibniz rule that

$$\begin{aligned}
\lim_{x\to x_0} \sum_{k=0}^{n-1}(-1)^k f^{(n-1-k)}(x)u_n^{(k)}(x) &= \lim_{x\to x_0} \sum_{k=0}^{n-1}(-1)^k f^{(n-1-k)}(x)u_n^{(k)}(x) \\
&= \lim_{x\to x_0} \sum_{k=0}^{n-1}(-1)^k f^{(n-1-k)}(x)(\alpha_n \cdot \phi)^{(k)}(x) \\
&= \lim_{x\to x_0} \sum_{k=0}^{n-1}(-1)^k f^{(n-1-k)}(x) \sum_{s=0}^{k}\binom{k}{s}\alpha_n^{(k-s)}(x)\phi^{(k)}(x) \\
&= \sum_{k=0}^{n-1}(-1)^k \sum_{s=0}^{k}\binom{k}{s} \lim_{x\to x_0} f^{(n-1-k)}(x)\alpha_n^{(k-s)}(x)\phi^{(k)}(x) = 0
\end{aligned}$$

and analogously it works for $\lim_{x\to x_1}$. This directly implies

$$\langle B_n f, \phi \rangle = \int_{x_0}^{x_1} f^{(n)}(x)u_n(x)dx = (-1)^n \int_{x_0}^{x_1} f(x)u_n^{(n)}(x)dx = \left\langle f, B_n^*\phi \right\rangle,$$

where

$$B_n^*\phi(x) := (-1)^n \frac{d^n}{dx^n}(\alpha_n\phi)(x) = (-1)^n \sum_{k=0}^{n}\binom{n}{k}\alpha_n^{(n-k)}(x)\phi^{(k)}(x),$$

which yields the assertion, since $D_d^* = \sum_{n=0}^{d} B_n^*$ as shown before. □

Furthermore, it is emphasized that we have got a new restriction to the set of suitable evaluation functionals $F := \langle \cdot, \phi \rangle$, namely

$$\phi \in \left\{ \phi \in C^{d-1}[x_0, x_1] \;\middle|\; \forall k \le d-1 : \lim_{x\to x_0} \phi^{(k)}(x) = \lim_{x\to x_1} \phi^{(k)}(x) = 0 \right\}.$$

This is a necessary condition to guarantee that the operator D_d^* in Lemma 4.1.7 is an adjoint of D_d, but for GOProM we also need powers $(\Phi_\varphi^*)^\ell$ of this operator, which is a simple consequence

of the definition of the canonical sampling scheme in this section, namely

$$S_{m,\ell} = \left\langle D_d^{m+\ell} \cdot, \phi \right\rangle = \left\langle \cdot, (D_d^*)^{m+\ell} \phi \right\rangle.$$

Thus, the kernel ϕ that defines the evaluation functional F for signals $f \in \mathcal{M}_M(D_d)$ has to be even smoother,

$$\phi \in \left\{ \phi \in C^{d(N+M)}[x_0, x_1] \;\middle|\; \forall k \leq d(N+M) : \lim_{x \to x_0} \phi^{(k)}(x) = \lim_{x \to x_1} \phi^{(k)}(x) = 0 \right\},$$

where $N + 1 \geq M$ is the number of columns of the sampling matrix $\mathcal{X}_{N,M}$.

The following lemmata present four types of suitable kernels inducing admissible dual sampling schemes. The first type of kernels has the advantage that it is infinitely often differentiable on the real line and all derivatives are compactly supported.

Lemma 4.1.8 (Bump kernels for arbitrary intervals)**.** *Let $D_d : C^d[x_0, x_1] \to C[x_0, x_1]$ be a linear differential operator as in Definition 4.1.6 with $\alpha_n \in C^\infty[x_0, x_1]$. Let the restricted point spectrum of D_d be denoted by $\hat{\sigma}_P(D_d) \neq \emptyset$ with corresponding eigenfunctions v_λ. Moreover, let $I := [x_0, x_1] \subset \mathbb{R}$ and $\kappa \in \mathbb{R}_\times^+$ be a scale parameter. Then the kernel*

$$\phi(x) := \begin{cases} \exp\left(-\frac{\kappa}{(x-x_0)^2}\right) \exp\left(-\frac{\kappa}{(x-x_1)^2}\right) & x \in I \\ 0 & x \notin I \end{cases}$$

fulfills the vanishing property of Lemma 4.1.7

$$\lim_{x \to x_0} \frac{d^k}{dx^k} \phi(x) = \lim_{x \to x_1} \frac{d^k}{dx^k} \phi(x) = 0,$$

for all $k \in \mathbb{N}_0$ and is infinitely often differentiable. Moreover, it is assumed for all $\lambda \in \hat{\sigma}_P(D_d)$ that $0 < |\langle v_\lambda, \phi \rangle| < \infty$ holds, i. e., ϕ is the kernel of an admissible evaluation scheme. Then the sampling scheme based on the generator D_d and the evaluation scheme $F_m := \left\langle \cdot, (D_d^*)^m \phi \right\rangle$, where D_d^* is as in Lemma 4.1.7, is given by $S_{m,\ell} : \mathcal{M}_\infty(D_d) \to \mathbb{C}$,

$$S_{m,\ell} := \left\langle \cdot, D_d^{*m+\ell} \phi \right\rangle.$$

This sampling scheme is admissible and realizable for all $f \in \mathcal{M}_M(D_d)$.

Proof. Since the support of the derivative of a function is a subset of the support of the function itself, it is sufficient to prove that the derivatives of ϕ are all well-defined and vanish at the boundaries x_0 and x_1. The first derivative of ϕ in I is

$$\frac{d}{dx} \exp\left(-\frac{\alpha}{(x-x_0)^2}\right) \exp\left(-\frac{\alpha}{(x-x_1)^2}\right) = \left(\frac{2\alpha}{(x-x_0)^3} + \frac{2\alpha}{(x-x_1)^3}\right) \phi(x).$$

Therefore, any further derivative is given by the product rule and the chain rule and is of the

form

$$\frac{d^k}{dx^k}\exp\left(-\frac{\alpha}{(x-x_0)^2}\right)\exp\left(-\frac{\alpha}{(x-x_1)^2}\right) = R_k(x)\exp\left(-\frac{\alpha}{(x-x_0)^2}\right)\exp\left(-\frac{\alpha}{(x-x_1)^2}\right),$$

where R_k is a rational function with poles of finite order at x_0 and x_1. Since the exponential function dominates every polynomial of finite degree and the product $\phi(x)$ vanishes for $x \to x_0$ and $x \to x_1$, also $\phi^{(k)}(x_0)$ and $\phi^{(k)}(x_1)$ vanish for all k and $\phi^{(k)}$ is bounded on I. Moreover, the condition $0 < |\langle v_\lambda, \phi\rangle| < \infty$ implies that the evaluation functional $F := \langle v_\lambda, \phi\rangle$ does not vanish for any $\lambda \in \hat{\sigma}_P(A)$ and therefore it is an admissible evaluation functional. The dual scheme inherits the admissibility directly. The realizability is obvious, since $S_{m,\ell}$ is a linear function on $\mathcal{M}_M(D_d)$. □

We can often choose a symmetric interval $I := [-x_0, x_0]$. Although the kernel in Lemma 4.1.8 is also suitable for this situation, a slightly different kernel can be used, too.

Lemma 4.1.9 (Bump Kernels for symmetric Intervals)**.** *Let $D_d : C^d[-x_0, x_0] \to C[-x_0, x_0]$ be a linear differential operator as in Definition 4.1.6 with $\alpha_n \in C^\infty[-x_0, x_0]$. Let the restricted point spectrum of D_d be denoted by $\hat{\sigma}_P(D_d) \neq \emptyset$ with corresponding eigenfunctions v_λ. Moreover, let $I := [-x_0, x_0] \subset \mathbb{R}$ and $\kappa \in \mathbb{R}^+_\times$ a scale parameter. Then the kernel*

$$\phi(x) := \begin{cases} \exp\left(-\frac{\kappa}{x_0 - x^2}\right) & x \in I \\ 0 & x \notin I \end{cases}$$

fulfills the vanishing property of Lemma 4.1.7

$$\lim_{x\to x_0}\frac{d^k}{dx^k}\phi(x) = \lim_{x\to x_1}\frac{d^k}{dx^k}\phi(x) = 0,$$

for all $k \in \mathbb{N}$ and is infinitely often differentiable. Moreover, it is assumed for all $\lambda \in \hat{\sigma}_P(D_d)$ that $0 < |\langle v_\lambda, \phi\rangle| < \infty$ holds, i. e., ϕ is the kernel of an admissible evaluation scheme. Then the sampling scheme based on the generator D_d and the evaluation scheme $F_m := \left\langle \cdot, (D_d^*)^m\phi\right\rangle$, where D_d^* is as in Lemma 4.1.7, is given by $S_{m,\ell} : \mathcal{M}_\infty(D_d) \to \mathbb{C}$,

$$S_{m,\ell} := \left\langle \cdot, D_d^{*m+\ell}\phi\right\rangle.$$

This sampling scheme is admissible and realizable for all $f \in \mathcal{M}_M(D_d)$.

Proof. The same arguments as in the proof of Lemma 4.1.8 can be used. □

Although these kernels are infinitely often continously differentiable and have in turn the advantage that they need no a priori knowledge of the order M of the model $\mathcal{M}_M(D_d)$ or a predefined sampling size, they behave very badly for high derivatives. The higher derivatives show a severe oscillation pattern with very rapid sign changes. Thus, for theoretical purposes these kernels are ideal, but impractical if we use them in connection with adjoint differential operators. The next family of kernels provides a compromise in a sense. They also vanish outside

a compact interval but they show a much better oscillation pattern. This is achieved by losing some generality and incorporating the knowledge about the model order M. Before concrete examples will be given, we formulate the basic properties of these kernels in the following lemma.

Lemma 4.1.10 (Gaussian Modulated Polynomial Kernels)**.** *Let $D_d : C^d[x_0, x_1] \to C[x_0, x_1]$ be a linear differential operator as in Definition 4.1.6 with $\alpha_n \in C^\infty[x_0, x_1]$. Let the restricted point spectrum of D_d be denoted by $\hat{\sigma}_P(D_d) \neq \emptyset$ with corresponding eigenfunctions v_λ. Moreover, let $I := [x_0, x_1] \subset \mathbb{R}$ and $\alpha, \beta \in \mathbb{R}_\times^+$ be scale parameters such that $\beta x_i \notin I$ for $i \in \{0, 1\}$, where $K \geq d(N + M)$. Then the kernel*

$$\phi_P(x) := \begin{cases} (x - x_0)^K (x - x_1)^K \exp\left(-\alpha(x - \beta x_0)^2 (x - \beta x_1)^2\right) & x \in I \\ 0 & x \notin I \end{cases}$$

is $K-$times continously differentiable and fulfills the vanishing property of Lemma 4.1.7

$$\lim_{x \to x_0} \frac{d^k}{dx^k} \phi_P(x) = \lim_{x \to x_1} \frac{d^k}{dx^k} \phi_P(x) = 0,$$

for all $k \leq K - 1$. Moreover, it is assumed for all $\lambda \in \hat{\sigma}_P(A)$ that $0 < |\langle v_\lambda, \phi_G \rangle| < \infty$ holds, i. e., ϕ is the kernel of an admissible evaluation scheme. Then the sampling scheme based on the generator D_d and the evaluation scheme $F_m := \left\langle \cdot, (D_d^*)^m \phi_G \right\rangle$, where D_d^* is as in Lemma 4.1.7, is given by $S_{m,\ell} : \mathcal{M}_\infty(D_d) \to \mathbb{C}$,

$$S_{m,\ell} := \left\langle \cdot, D_d^{*m+\ell} \phi \right\rangle.$$

This sampling scheme is admissible and realizable for all $f \in \mathcal{M}_M(D_d)$.

Proof. The admissibility and realizability are inherited from the canonical sampling scheme as already seen in the proof of Lemma 4.1.8. Therefore, we only show the differentiability and the vanishing property in detail. We start by calculating the k^{th}-derivative of $p(x) := (x - x_0)^K (x - x_1)^K$ over I by using the Leibniz rule,

$$p^{(k)}(x) = \frac{d^k}{dx^k} (x - x_0)^K (x - x_1)^K = \sum_{s=0}^{k} \binom{k}{s} \frac{K!^2}{(K - k + s)!(K - s)!} (x - x_0)^{K-k+s} (x - x_1)^{K-s}.$$

The continous derivatives exist for all $k \leq K - 1$ and they obviously vanish outside $\mathbb{R} \setminus (x_0, x_1)$. This implies that also $p^{(k)}(x) \mathbb{1}_I(x)$ is k-times continously differentiable. Furthermore, the exponential term is infinitely often differentiable and therefore also the product $p(x) \exp\left(-\alpha(x - \beta x_0)^2 (x - \beta x_1)^2\right)$. In turn, the function $\phi_{\alpha,\beta}$ is $(K - 1)$-times continously differentiable. □

A fourth type of kernels, that are also encountered in other applications, are cardinal $B-$splines. A cardinal $B-$spline is a piecewise polynomial which can be calculated by the following iteration. Let $\phi_0 := \mathbb{1}_{[0,1)}$ be per definition the B-spline of degree zero, i. e. the characteristic

function of the interval $[0, 1)$. The n^{th} cardinal B-spline is defined by successive convolutions of ϕ_0 with itself, i. e.,

$$\phi_n(x) := (\phi_{n-1} * \phi_0)(x) = \int_{\mathbb{R}} \phi_{n-1}(x - s)\phi_0(s)ds.$$

The first two B-splines are

$$\phi_1(x) = \begin{cases} x, & x \in [0, 1) \\ 2 - x & x \in [1, 2) \\ 0, & otherwise \end{cases} \quad \text{and} \quad \phi_2(x) = \begin{cases} \frac{1}{2}x^2, & x \in [0, 1) \\ \frac{1}{2} + (x - 1)(2 - x) & x \in [1, 2] \\ \frac{1}{2} + \frac{1}{2}(x - 2)(x - 4) & x \in [2, 3) \\ 0, & otherwise\ . \end{cases}$$

The most interesting property of those kernels from the GOProM perspective is the fact that they give rise to non-canonically generated evaluation schemes. In order to see this let a convulotion operator B with respect to ϕ_0 for all $g \in C(\mathbb{R})$ be defined as

$$Bf := \phi * g = \int_{\mathbb{R}} g(\cdot - t)\phi(t)dt.$$

With this operator at hand, we define the following evaluation scheme for all $f \in C(\mathbb{R})$ by

$$F_m := \left\langle \cdot, B^{m-1}\phi_0 \right\rangle = \langle \cdot, \phi_m \rangle,$$

where ϕ_m is the m^{th} cardinal B-spline as defined above, since $\phi_m = (\phi_{m-1} * \phi_0) = B^{m-1}\phi_0$. This means that F_m is an evaluation scheme, generated by the successive convolution of the characteristic function of the interval $[0, 1)$ with itself.

Lemma 4.1.11 (B-Spline kernels)**.** *Let $D_d : C^d[x_0, x_1] \to C[x_0, x_1]$ be a linear differential operator as in Definition 4.1.6 with $\alpha_n \in C^\infty[x_0, x_1]$. Let the restricted point spectrum of D_d be denoted by $\hat{\sigma}_P(A) \neq \emptyset$ with corresponding eigenfunctions v_λ. Moreover,*

$$Bg := \int_{\mathbb{R}} g(t)\phi_0(\cdot - t)dt,$$

where $\phi_0(x) := \mathbb{1}_{[x_0,x_1)}(x)$ and $g \in C[x_0, x_1]$. Then the sampling scheme $S_{m,\ell} : \mathcal{M}_\infty(D_d) \to \mathbb{C}$,

$$S_{m,\ell} := \left\langle \cdot, D_d^{*\ell} B^{m-1}\phi_{dM} \right\rangle = \int_{\mathbb{R}} (\cdot)(x)(D_d^*)^\ell \phi_{dM+m}(x)dx = \left\langle \cdot, (D_d^*)^\ell \phi_{dM+m} \right\rangle,$$

where ϕ_n is the n^{th} cardinal B-spline over $[x_0, x_1)$, is admissible if $E_F = [\langle v_\lambda, \phi_m \rangle]_{m=0,\lambda\in\Lambda}^N$ has full rank for all choices of $\Lambda \in \hat{\sigma}_P(A)$, $N \geq M - 1$, is realizable, and the formal adjoint D_d^* is well-defined since all derivatives of ϕ_m vanish outside the interval I.

Proof. It holds that

$$S_{m,\ell} f = \left\langle D_d^\ell f, B^{m-1}\phi_0 \right\rangle = \left\langle f, (D_d^*)^\ell B^{m-1}\phi_0 \right\rangle = \left\langle f, (D_d^*)^\ell \phi_m \right\rangle,$$

where ϕ_m is by definition the m^{th} B-spline supported on the interval $[x_0, m+1+x_1)$. The result follows by substituting ϕ_0 by ϕ_{dM}, to have sufficiently smooth kernels; they have to be at least $d\ell$-times differentiable, which is ensured by the choice ϕ_{dM}. The admissibility is just a reformulated version of the usual one, because $F_m(v_\lambda) = \langle v_\lambda, \phi_m \rangle$. □

Of course, we could have used analogously to the polynomial example any B-spline with sufficiently high degree K and use the sampling scheme

$$S_{m,\ell} = \left\langle \cdot, (D_d^*)^{m+\ell} \phi_K \right\rangle.$$

This sampling scheme is admissible under the same conditions as for the polynomial kernel. The advantage of the splines over this scheme is that we keep the coefficients of the resulting kernels smaller. Instead, if we iterate a differential operator on a polynomial of high degree, the coefficients of the resulting polynomial quickly become very large.

Remark 4.1.12. Besides the fact that the convolution operator $B^* f := \int_{\mathbb{R}} f(\cdot - t)\phi(t)dt$ and the correlation operator $Bf := \int_{\mathbb{R}} f(t)\phi(\cdot + t)dt$ can be used to generate certain evaluation schemes, they also have the exponentials as eigenfunctions. This holds analogously for both by

$$B\exp(\lambda x) = \int_{\mathbb{R}} \exp(\lambda(x-t))\phi_0(t)dt = \int_{\mathbb{R}} \exp(-\lambda t)\phi_0(t)dt \, \exp(\lambda x) = \varphi(\lambda)\exp(\lambda x),$$

as long as $\varphi(\lambda) := \int_{\mathbb{R}} \exp(-\lambda t)\phi_0(t)dt < \infty$. This implies that we could use them directly as iteration operators for $f \in \mathcal{M}_\infty\left(\frac{d}{dx}\right)$ if $\lambda \to \varphi(\lambda)$ is one to one. In particular, the convolution operator commutes with the shift operator; but not only that, if an operator commutes with the shift, it is already a convolution. This implies that we can switch between the shift and the convolution operator to get an iteration operator or to generate the evaluation scheme.

4.2 Application of the Dual Sampling Approach

4.2.1 Exponential Expansions

We again consider as a first example for the dual sampling approach sparse exponential sums, i. e., $f : \mathbb{R} \to \mathbb{C}$ such that

$$f(x) = \sum_{\lambda \in \Lambda} c_\lambda v_\lambda(x),$$

where $v_\lambda(x) := \exp(\lambda x)$ are the eigenfunctions of the generator $A = \frac{d}{dx}$ and Λ is a finite subset of the restricted point spectrum $\hat{\sigma}_P(A) = \mathbb{R} \times i[0, 2\pi)$, with $|\Lambda| = M$. A possible sampling scheme for these expansions based on some evaluation functional $F : C(\mathbb{R}) \to \mathbb{C}$ is

$$S_{m,\ell} = F_m \circ A^\ell = F \circ \frac{d^{m+\ell}}{dx^{m+\ell}}, \; m + \ell \in \{0, 1, \ldots, N + M - 1\},$$

which is admissible if $0 < |F(v_\lambda)| < \infty$ for all $\lambda \in \hat{\sigma}_P(A)$. In the introduction the connection to the shift operator was used to get a realizable sampling scheme. Now we use the dual approach;

therefore, we need the formal adjoint of the ordinary differential operator provided by Lemma 4.1.7, i. e.,

$$(A^*)^n = \left(\frac{d^n}{dx^n}\right)^* = (-1)^n \frac{d^n}{dx^n} \implies S_{m,\ell} = \left\langle \cdot, (-1)^{m+\ell} \frac{d^{m+\ell}}{dx^{m+\ell}} \phi \right\rangle,$$

where ϕ is a kernel as in Lemma 4.1.7 and $A := \frac{d}{dx} : C^1(\mathbb{R}) \to C(\mathbb{R})$. From Corollary 4.1.4 we also know that this sampling scheme is admissible as long as $0 < |\langle v_\lambda, \phi \rangle| < \infty$ for all $\lambda \in \hat{\sigma}_P(A)$.

A kernel which satisfies these properties is by Lemma 4.1.10 the polynomial kernel for $\alpha = 0$ defined by $\phi(x) := x^K \left(\frac{1}{2} - x\right)^K \mathbb{1}_{\left[0,\frac{1}{2}\right)}(x)$, where $M := |\Lambda|$, and $K := N + M - 1$, because it holds that

$$F(v_\lambda) := \int_0^{\frac{1}{2}} \exp(\lambda x) x^K \left(\frac{1}{2} - x\right)^K dx \neq 0$$

for all $k \in \mathbb{N}_0$ and $\lambda \in \hat{\sigma}_P(A)$. This can be seen by

$$F(v_\lambda) := \int_0^{\frac{1}{2}} \exp(u_\lambda x) \cos(w_\lambda x) x^K \left(\frac{1}{2} - x\right)^K dx + i \int_0^{\frac{1}{2}} \exp(u_\lambda x) x^K \sin(w_\lambda x) \left(\frac{1}{2} - x\right)^K dx,$$

where $\lambda := u_\lambda + i w_\lambda$ with $u_\lambda \in \mathbb{R}$ and $w_\lambda \in [0, 2\pi)$. Due to the fact that the imaginary part of $F(v_\lambda)$ is an integral over a strictly positive function over $(0, 0.5)$ for all $\lambda \in \hat{\sigma}_P(A)$, the functional F does not vanish for any λ, i. e. $|F(v_\lambda)| > 0$.

The sampling scheme applied to a function $f \in \mathcal{M}_M(A)$ is then given by Lemma 4.1.10 and the Leibniz rule. We substitute for convenience $k := m + \ell$ and get

$$\begin{aligned} S_{m,\ell}(f) &= \left\langle f, (-1)^{m+\ell} \frac{d^{m+\ell}}{dx^{m+\ell}} \phi \right\rangle = (-1)^k \int_0^{\frac{1}{2}} f(x) \left[x^{N+M-1} \left(\frac{1}{2} - x\right)^{N+M-1} \right]^{(k)} dx \\ &= \sum_{s=0}^{k} \sum_{j=0}^{K-s} \alpha_{K,k,s,j} \int_0^{\frac{1}{2}} f(x) x^{K-k+s+j} dx, \end{aligned}$$

where $\alpha_{K,k,s,j} := (-1)^{k+j} \left(\frac{1}{2}\right)^{K-s-j} \binom{k}{s} \binom{K}{k-s} \binom{K}{s} \binom{K-s}{j} s!(k-s)!$ and $K = N + M$. This means we essentially have to sample at most $N + M - 1$ moments of f, $N \geq M - 1$. In the case of $N = M$ we need $2M - 1$ moments of f, where we start with the moment to the power of zero and calculate all moments up to the power of $2M - 1$. Eventually, this yields the sampling matrix

$$\mathcal{X}_{N,M}(f) = \sum_{s=0}^{m+\ell} \sum_{j=0}^{K-s} \left[\alpha_{K,m+\ell,s,j} \int_0^{\frac{1}{2}} f(x) x^{K-m-\ell+s+j} dx \right]_{m,\ell=0}^{N,M},$$

where $N \geq M - 1 \in \mathbb{N}$.

4.2.2 Sparse Sturm-Liouville Expansions

In this section we discuss expansions into eigenfunctions of Sturm-Liouville-type differential operators. We follow the notation in [18] with the boundary conditions

$$R_1 f := \alpha_1 f(a) + \alpha_2 f^{(1)}(a) = 0 \text{ and } R_2 f = \beta_1 f(b) + \beta_2 f^{(1)}(b) = 0,$$

defined by operators $R_1 := \alpha_1 \mathrm{I} + \alpha_2 \frac{d}{dx}$ and $R_2 := \beta_1 \mathrm{I} + \beta_2 \frac{d}{dx}$ and the assumptions that for the real functions p, q, and r it holds that

1. $p \in C^1[a,b]$ and $q, r \in C[a,b]$
2. $p(t) > 0$ and $r(t) > 0$ for all $t \in [a,b]$
3. $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}_{\neq 0}$.

The corresponding Sturm-Liouville operator is defined as $\mathcal{L} : C^2[a,b] \to C[a,b]$,

$$\mathcal{L} := \frac{1}{r}\left(\frac{d}{dx} p \frac{d}{dx} + q\,\mathrm{I}\right).$$

Using this notation, the following eigenvalue problem can be formulated

$$\mathcal{L} v_{\lambda_n} = \lambda_n v_{\lambda_n}.$$

From the theory of Sturm-Liouville problems it is known that this equation has countably many orthonormal solutions for certain $\lambda_n \in \mathbb{N}$ for $n \in \mathbb{N}$. These solutions form a basis for the Hilbert space $L^2([a,b], rdx)$ equipped with the inner product

$$\langle f, g \rangle_r := \int_a^b (fg)(x) \cdot r(x) dx.$$

In particular, the eigenvalues uniquely identify the eigenfunctions and therefore GOProM can be applied. Unfortunately, a direct iteration leads to non-realizable sampling schemes and an exponential iteration is not explicitly realizable since the terms in $\mathcal{L}$ generally do not commute in a simple way. Thus, we employ the dual approach to come up with realizable sampling schemes. A useful property of the Sturm-Liouville operator is that it is self-adjoint with respect to kernels $\phi \in L^2([a,b], rdx) \cap C^1[a,b]$, where $\phi : [a,b] \to \mathbb{R}$ such that $(R_1\phi)(a) = (R_2\phi)(b) = 0$, i. e.,

$$\langle \mathcal{L} f, \phi \rangle_r = \langle f, \mathcal{L}\phi \rangle_r$$

for all $f \in L^2([a,b], r(x)dx)$. This is a consequence of

$$\begin{aligned} \phi r \mathcal{L} f - f r \mathcal{L}\phi &= \phi \frac{d}{dx} p f^{(1)} + \phi q f - f \frac{d}{dx} p \phi^{(1)} - f q \phi = \phi \frac{d}{dx} p f^{(1)} - f \frac{d}{dx} p \phi^{(1)} \\ &= \frac{d}{dx}\left[p\left(\phi f^{(1)} - \phi^{(1)} f\right)\right]. \end{aligned}$$

Taking integrals on both sides gives

$$\int_a^b r(x)(\mathcal{L}f)(x)\phi(x) - r(x)f(x)(\mathcal{L}\phi)(x)dx = 0,$$

and because of $R_1\phi = R_2\phi = 0$ it holds that

$$\begin{aligned}\int_a^b \frac{d}{dx}\left[p\left(\phi f^{(1)} - \phi^{(1)} f\right)\right] dx &= p(b)\left(f^{(1)}(b)\phi(b) - f(b)\phi^{(1)}(b)\right) - p(a)\left(f^{(1)}(a) - f(a)\phi^{(1)}(a)\right) \\ &= p(b)f^{(1)}(b)\left(\phi(b) + \frac{\beta_2}{\beta_1}\phi^{(1)}(b)\right) - p(a)f^{(1)}(a)\left(\Phi_\varphi + \frac{\alpha_2}{\alpha_1}\phi^{(1)}(a)\right) \\ &= p(b)f^{(1)}(b)R_2\phi - p(a)f^{(1)}(a)R_1\phi = 0.\end{aligned}$$

Thus, we have

$$\langle \mathcal{L}f, \phi\rangle_r = \int_a^b (\mathcal{L}f)(x)\phi(x)r(x)dx = \int_a^b f(x)(\mathcal{L}\phi)(x)r(x)dx = \langle f, \mathcal{L}\phi\rangle_r .$$

In contrast, this well-known result for Sturm-Liouville operators imposes a different condition on suitable evaluation kernels ϕ than in Lemma 4.1.7. Instead of a certain number of vanishing derivatives, it is sufficient to assume that the boundary conditions R_1 and R_2 are fulfilled for all $\mathcal{L}^{m+\ell}\phi$, $m + \ell = 0, \ldots, N + M - 1$. In principle, this is a weaker condition as in Lemma 4.1.7 in the sense that all kernels therein fulfill it automatically. In general, this is a non-trivial assumption, since R_1 and R_2 do not commute with $\mathcal{L}$ and therefore we have to impose this boundary condition on every iteration $\mathcal{L}^{m+\ell}\phi$.
Fortunately, in the special case of kernels

$$\phi(x) := \sum_{n=0}^{\infty} \alpha_n v_{\lambda_n}(x),$$

where $\{v_{\lambda_n} \mid n \in \mathbb{N}_0\}$ is an orthonormal basis for $L^2([a,b], r(x)dx)$ which solves the Sturm-Liouville problem, it holds that

$$\left(R_1\mathcal{L}^{m+\ell}\phi\right)(a) = \left(R_2\mathcal{L}^{m+\ell}\phi\right)(b) = 0.$$

This can be seen by the following calculation, using the linearity of R_1 and R_2. For $i \in \{1, 2\}$ it holds that

$$R_i\mathcal{L}^{m+\ell}\phi = R_i\mathcal{L}^{m+\ell}\sum_{n=0}^{\infty}\alpha_n v_n = R_i\sum_{n=0}^{\infty}\alpha_n\lambda_n^{m+\ell}v_n = \sum_{n=0}^{\infty}\alpha_n\lambda_n^{m+\ell}R_i v_n\,;$$

therefore, they vanish on the boundary points

$$(R_1\mathcal{L}^{m+\ell}\phi)(a) = (R_2\mathcal{L}^{m+\ell}\phi)(b) = 0$$

for all $(m + l) \in \mathbb{N}$. Thus, for these kernels the Sturm-Liouville operator and all its iterations are self-adjoint.

Although this is a nice result, the issues of this type of kernels discussed shortly after Corollary 4.1.4 are evident. Either we assume a finite alphabet or we accept the fact that $F := \langle \cdot, \phi \rangle$ gets arbitrarily close to zero. Since we want to have no restriction on the alphabet in the upcoming example, we use the symmetric polynomial kernel from Lemma 4.1.10.

This kernel fulfills the boundary conditions R_1 and R_2 since it simply vanishes for all derivatives in a and b. Therefore, the assumptions in Lemma 4.1.7 are given and the following calculation shows that $\mathcal{L}$ is also self-adjoint with respect to this type of kernels.

$$\begin{aligned}\left(\frac{d}{dx}p\frac{d}{dx} + q\,\mathrm{I}\right)^* &= \left(p\frac{d^2}{dx^2} + p^{(1)}\frac{d}{dx} + q\,\mathrm{I}\right)^* = \frac{d^2}{dx^2}p - \frac{d}{dx}p^{(1)} + q\,\mathrm{I} \\ &= \frac{d}{dx}\left(p^{(1)}\,\mathrm{I} + p\frac{d}{dx}\right) - p^{(2)}\,\mathrm{I} - p^{(1)}\frac{d}{dx} + q\,\mathrm{I} \\ &= p^{(2)}\,\mathrm{I} + p^{(1)}\frac{d}{dx} + \frac{d}{dx}p\frac{d}{dx} - p^{(2)}\,\mathrm{I} - p^{(1)}\frac{d}{dx} + q\,\mathrm{I} \\ &= \frac{d}{dx}p\frac{d}{dx} + q\,\mathrm{I}\end{aligned}$$

Thus, we proceed using the sufficiently differentiable and compactly supported sampling kernels of Lemma 4.1.10 to demonstrate the dual approach for Sturm-Liouville type expansions, namely, finite linear combinations of Legendre polynomials.

Remark 4.2.1. A special type of Sturm-Liouville operators and polynomials was already discussed from different perspectives, namely the Chebychev polynomials, which are defined as eigenfunctions of

$$(1-x^2)\frac{d^2}{dx^2} - x\frac{d}{dx} = (1-x^2)\frac{d^2}{dx^2} - 2x\frac{d}{dx} + x\frac{d}{dx} = \frac{d}{dx}(1-x^2)\frac{d}{dx} + x\frac{d}{dx}.$$

Sparse Legendre-Expansions

In this section we construct a sampling scheme for sparse expansions into Legendre polynomials of arbitrary degree. This problem was already considered in [33]. The great advantage of the new dual approach is that the sampling scheme can be formulated without any use of signal derivatives.

First, we note that the Legendre polynomial L_n can be defined by using the following operator as generator,

$$A := (1-x^2)\frac{d^2}{dx^2} - 2x\frac{d}{dx} = \frac{d}{dx}(1-x^2)\frac{d}{dx}.$$

This operator is obviously of Sturm-Liouville type and has the Legendre polynomials L_n as eigenfunctions, i. e.,

$$AL_n = -n(n+1)L_n,$$

where $\lambda_n = -n(n+1)$. Therefore, the restricted point spectrum in this case is $\hat{\sigma}_P(A) = \{-n(n+1) \,|\, n \in \mathbb{N}_0\}$. The goal is to reconstruct sparse linear combinations of Legendre polynomials, i. e.,

$$f(x) = \sum_{j=1}^{M} c_{n_j} L_{n_j}(x).$$

The operator A is by the considerations above formally self-adjoint with respect to a suitable vanishing kernel. A possible choice is delivered by Lemma 4.1.10.

$$\phi_P(x) := \begin{cases} (x-x_0)^K (x-x_1)^K \exp\left(-\alpha(x-\beta_0)^2(x-\beta_1)^2\right) & x \in I \\ 0 & x \notin I \end{cases}$$

Here,we choose the interval $I := [-0.5, 0.75]$, $\beta_0 = \beta_1$, and $K := d(N+M)$, where M is the a priori known model order, $N+1$ the number of rows of the sampling matrix, and $d = 2$ the order of the differential operator A. Let

$$F(f) := \int_{-0.5}^{0.75} f(x)\phi_P(x)dx \neq 0$$

be the evaluation functional based on the kernel above. We will use the canonical sampling scheme $S_{m,\ell} := F \circ A^{m+\ell}$; thus, we have to check the admissibility condition. Since the kernel is an asymmetric function concerning the y-axis and no pure polynomial term, the integral does not vanish on any Legendre polynomial L_n, i. e.,

$$\forall\, n \in \mathbb{N} \;:\; F(L_n) = \int_{-0.5}^{0.75} L_n(x)\phi_P(x)dx \neq 0.$$

In turn, the dual sampling scheme based on the canonical one is also admissible.
If we now assume that the given signal is

$$f(x) := \sum_{j=1}^{3} c_{n_j} L_{n_j}(x),$$

i. e., a three sparse linear combination of arbitrary Legendre polynomial, we need five iterations of $A^n\phi_P$, $n \in \{0,1,2,3,4,5\}$, to generate the sampling matrix based on the canonical sampling scheme, where $n := m + \ell$. Furthermore, we fix $N := M$, $\alpha = 0.1$, and $-\beta_0 = \beta_1 = 2$. As a result we have a sampling kernel of the form

$$\phi_P(x) := \begin{cases} (x+0.5)^{12}(x-0.75)^{12} \exp\left(-\frac{1}{10}(x^2-4)^2\right) & x \in I \\ 0 & x \notin I \end{cases}.$$

The parameters of the sampling kernel are chosen such that the amplitude of $A^n\phi_P$ stays in a tractable range for all $n \in \{0, 1, 2, 3, 4, 5\}$. Since the kernels are not really demonstrative, we depict them in Figure 4.1 by their graphs on $[-1, 1]$. They correspond, starting at the top left, to the kernels $A^0\phi_P$, $A^1\phi_P$, $A^2\phi_P$ (upper row), $A^3\phi_P$, $A^4\phi_P$, and $A^5\phi_P$ (lower row).

Actually, this type of sampling kernels reminds of wavelets, and in a sense they also gather

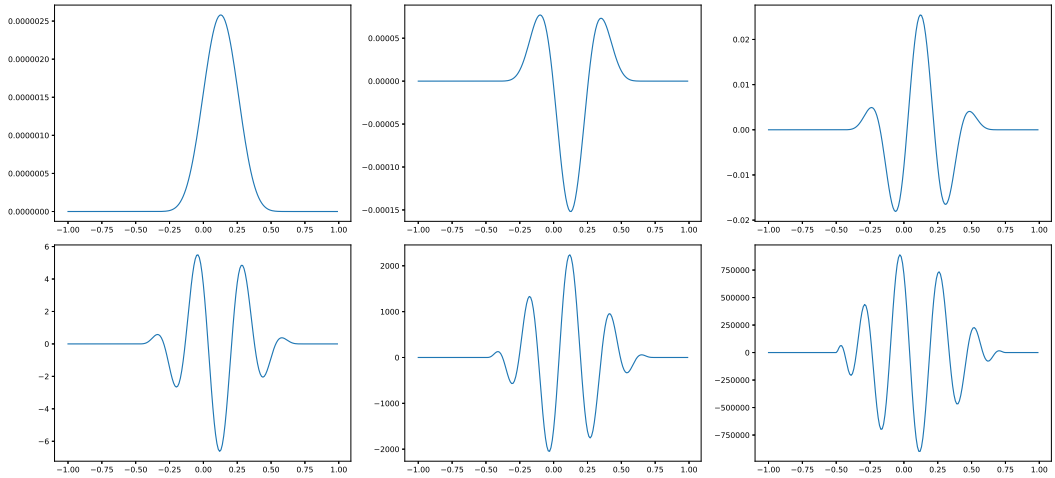

Figure 4.1: Sampling kernels for a Legendre Expansion

information over multiple scales of the signal at hand. Of course, there are fundamental differences between the sampling kernels here and wavelets of any kind, but it gives a hint that Prony's method could be interpreted as a sort of multiscale analysis in its own right, at least in the dual approach for sparse signals.

These kernels can now be used for any kind of 3−sparse linear combination of arbitrary Legendre polynomials. For a concrete example we choose as maximum degree $\hat{n} = 20$, but we could have taken any other higher natural number. Furthermore, we draw without replacement three distinct numbers $n \in \{0, 1, \ldots, \hat{n}\}$, that define the active atoms of the exemplarily signal. Moreover, the linear coefficients are drawn uniformely from the interval $[1, 10]$. The concrete realizations in this example are shown in Table 4.1. The signal with these parameters is

n_j	1	4	9
c_{n_j}	1.703	3.193	3.710

Table 4.1: Polynomial degrees n and the corresponding linear coefficients c_n

depicted in Figure 4.2. For the reconstruction we take samples

$$S_{m,\ell} = \left\langle f\,,\, A^{m+\ell}\phi_P \right\rangle.$$

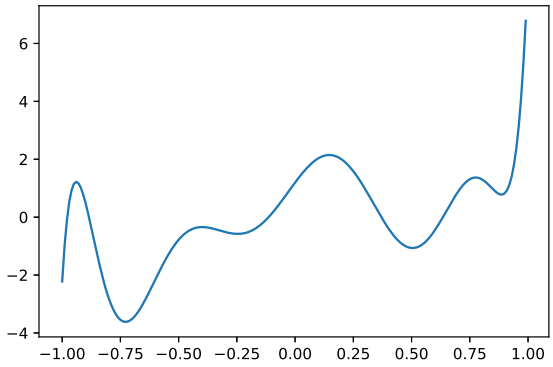

Figure 4.2: 3-sparse Legendre Expansion

and apply the GOProM approach with the sampling matrix

$$\mathcal{X}_{3,3} = \begin{bmatrix} S_{0,0}(f) & S_{0,1}(f) & S_{0,2}(f) & S_{0,3}(f) \\ S_{1,0}(f) & S_{1,1}(f) & S_{1,2}(f) & S_{1,3}(f) \\ S_{2,0}(f) & S_{2,1}(f) & S_{2,2}(f) & S_{2,3}(f) \end{bmatrix}.$$

The reconstructed parameters can be seen in Table 4.2. Due to rounding errors the polynomial

n	1.00008823	4.00001099	9.00000026
c_n	1.703	3.193	3.710

Table 4.2: Polynomial degrees n and the corresponding linear coefficients c_n

degrees are not exactly recovered. Therefore, we round to the closest natural number and get the exact values. Although this is a quite heuristic estimation procedure, in the noiseless scenario it works quite well because of the small rounding errors. In contrast, in the presence of noise, where we cannot expect that we are always closer to the true parameter than to any other integer, more work has to be done to achieve a realiable reconstruction.
Despite the fact that we usually use a Vandermonde system based on the same samples as for the reconstruction of the active eigenvalues, we recover the linear parameters in a different way. Instead of solving the Vandermonde system, we use the orthonormality of the Legendre polynomials with respect to $\langle g, h\rangle = \int_{-1}^{1} g(x)h(x)dx$, for all functions g and h square integrable over $[-1, 1]$ and get

$$c_\lambda = (n + 0.5)\langle f, L_n\rangle.$$

This leads to a perfect reconstruction as seen above.
An issue with this approach is the very high amplitude of the last iterations $A^{m+\ell}\phi$, $0 \ll m + \ell$, compared to the amplitude of the kernel ϕ. A clever choice of the parameters of ϕ can help to control these amplitudes, but a systematic way of finding suitable sampling kernels with small amplitude ranges over several iteration is an open problem.

4.3 Conclusion and ongoing Research

In this thesis, an operator based formulation of Prony's method was introduced and systematically generalized to the GOProM framework. Thereby, we defined the concepts of a generator A, iteration operator Φ_φ, evaluation scheme F_m and sampling scheme $S_{m,\ell}$, and the sampling matrix $\mathcal{X}_{N,M}$ as a preparation for GOProM. Based on these terms, an identifiable signal model $\mathcal{M}_M(A)$ was defined and embedded in the signal space $\mathcal{M}(A)$ of all finite linear combinations of eigenfunctions. In a second step, we further restricted this model with respect to the iteration operator Φ_φ to the final signal model $\mathcal{M}_M^\varphi(A)$, in order to preserve the identifiability also for the modified sampling schemes.
Afterwards, based on the sampling scheme $S_{m,\ell}$ and matrix $\mathcal{X}_{N,M}$, it was shown how the existing applications of the Prony method fit into this framework. Explicit examples were given, for instance, sparse expansions into cosine functions or Stieltjes-Wigert polynomials, where the last one was a completely new application of the generalized Prony method.
Later on, a whole class of examples was discussed in more detail, namely sparse expansions into eigenfunctions of exponential operators, in particular the generalized shifts, which already cover almost all former applications. It was shown how to employ the GEProM approach to reconstruct a variety of exponential atoms and even get a way to recover the parameters of signals based on squared operators. Two particular cases were the cosine and Chebychev expansions, that were treated from two different points of view, namely the symmetric shift approach as an example of the GOProM and more specifically the integral approach, that is an example of GEProM.
In the last section we discussed a new point of view of data acquisitation in the context of generalized Prony methods. By the decomposition of iteration operator and evaluation scheme we have seen that a suitable choice of adjoint operators can be used to formally derive sampling schemes for all linear differential operators. These schemes are no longer realized by sampling derivatives, but by linear measurements and are therefore in our terms realizable. Examples for this approach were the classical problem of reconstructing sparse expansions of complex exponentials and linear combinations of Legendre polynomials. It is emphasized that the used techniques are the same for all linear differential equations. First, we fix an admissible evaluation scheme F_m, deriving the adjoint of the iteration operator Φ_φ, in this case often simply A, and evaluate the orbit of $(\Phi_\varphi^*)^\ell \phi_m$ dependent on the model order M and the intended total number of sampels. Of course, we could use this approach also for GEProM or in general for every GOProM. It has to be decided in the individual cases if something is gained by changing the sampling from the direct to the dual.
The topics discussed in this thesis do not cover all possibilities opened by GOProM. Fields of active research are approximative sampling schemes, which make it possible to include the work of Wischerhoff & Plonka [36] and some cases of the FRI-approach similar to it. It seems also possible to include the FRI approach in total, but this is an open question as well. Another possible research direction is the already mentioned field of dynamical sampling and the question about deeper connection besides a mere reinterpretation of GOProM. Last but not least, as

often for Prony methods, the stabilization of the envolved algorithms in the presence of noise has to be adressed in further work. Due to current results, it seems possible to generalize the autocovariance approach in [17] for orthogonal eigenfunctions and the Maximum-Likelihood method in [31] for certain error models based on the GOProM signal space.

Bibliography

[1] Akram Aldroubi, Carlos Cabrelli, Ursula Molter, and Sui Tang. Dynamical sampling. *Applied and Computational Harmonic Analysis*, 42:378–401, 2017.

[2] Akram Aldroubi and Armenak Petrosyan. Dynamical Sampling and Systems from Iterative Actions of Operators. In Isaac Pesenson, Quoc Thong Le Gia, Azita Mayeli, Hrushikesh Mhaskar, and Ding-Xuan Zhou, editors, *Frames and Other Bases in Abstract and Function Spaces*, pages 15–26. Birkhäuser, Basel, 2017.

[3] Badri Narayan Bhaskar, Gongguo Tang, and Benjamin Recht. Atomic Norm Denoising With Applications to Line Spectral Estimation. *IEEE Transactions on Signal Processing*, 61(23):5987–5999, 2013.

[4] Gerald Gustave Bilodeau. The Weierstrass transform and Hermite polynomials. *Duke Mathematical Journal*, 29(2):293–308, 1962.

[5] Yoram Bresler and Albert Macovski. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(5):1081–1089, 1986.

[6] Venkat Chandrasekaran, Benjamin Recht, Pablo A. Parrilo, and Alan S. Willsky. The Convex Geometry of Linear Inverse Problems. *Foundations of Computational Mathematics*, 12(6):805–849, 2012.

[7] Scott Shaobing Chen, David L. Donoho, and Michael L. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[8] Jacob S. Christiansen and Erik Koelink. Self-adjoint difference operators and classical solutions to the Stieltjes–Wigert moment problem. *Journal of Approximation Theory*, 140(1):1–26, 2006.

[9] Annie Cuyt and Wen-shin Lee. Multivariate exponential analysis from the minimal number of samples. *Advances in Computational Mathematics*, 44(4):987–1002, 2018.

[10] Giuseppe Dattoli and Decio Levi. Exponential operators and generalized dierence equations. *Il Nuovo Cimento B*, 115(6):653–662, 2000.

[11] Giuseppe Dattoli, Paolo Emilio Ricci, and Dario Sacchetti. Generalized shift operators and pseudo-polynomials of fractional order. *Applied Mathematics and Computation*, 141(1):215–224, 2003.

[12] Giuseppe Dattoli, Amalia Torre, P. L. Ottaviani, and L. Vázquez. Evolution operator equations: Integration with algebraic and finitedifference methods. Applications to physical problems in classical and quantum mechanics and quantum field theory. *La Rivista del Nuovo Cimento*, 20(2):3–133, 1997.

[13] Gaspard Riche de Prony. Essai experimental et analytique sur les lois de la delatabilite des fluides et sur celles de la force expansive del la vapeur de l'eau el de l'alkool differentes temperatures. *l'Ecole Polytech.*, 1(2):24–76, 1795.

[14] Federico Delfino, Renato Procopio, Mansueto Rossi, and Farhad Rachidi. Prony Series Representation for the Lightning Channel Base Current. *IEEE Transactions on Electromagnetic Compatibility*, 54(2):308–315, 2012.

[15] Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.

[16] Thomas Ernst. *A Comprehensive Treatment of q-Calculus*. Birkhäuser, Basel, 2012.

[17] Frank Filbir, Hrushikesh Mhaskar, and Jürgen Prestin. On the Problem of Parameter Estimation in Exponential Sums. *Constructive Approximation*, 35(3):323–343, 2012.

[18] Harro Heuser. *Funktionalanalysis: Theorie und Anwendung*. Mathematische Leitfäden. Vieweg+Teubner Verlag, Wiesbaden, 2006.

[19] Yingbo Hua and Tapan Kumar Sarkar. On SVD for estimating generalized eigenvalues of singular matrix pencil in noise. *IEEE Transactions on Signal Processing*, 39(4):892–900, 1991.

[20] Jae-Wook Jung, Jung-Wuk Hong, Hyoung-Ki Lee, and Kiwan Choi. Estimation of viscoelastic parameters in Prony series from shear wave propagation. *Journal of Applied Physics*, 119:id. 234701, 2016.

[21] Victor Kac and Pokman Cheung. *Quantum Calculus*. Universitext. Springer-Verlag, New York, 2002.

[22] Roelof Koekoek, Peter A. Lesky, and René F. Swarttouw. *Hypergeometric Orthogonal Polynomials and Their q-Analogues*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, Heidelberg, 2010.

[23] Stefan Kunis, Thomas Peter, Tim Römer, and Ulrich von der Ohe. A multivariate generalization of Prony's method. *Linear Algebra and its Applications*, 490:31–47, 2016.

[24] Yutian Li and Wong, Richard, R. Global Asymptotics of Stieltjes-Wigert Polynomials. *Analysis and Applications*, 11(5):id. 1350028, 2013.

[25] Tadeusz Lobos, Jacek Rezmer, and Peter Schegner. Parameter Estimation of Distorted Signals Using Prony Method. In IEEE, editor, *2003 IEEE Bologna Power Tech Conference Proceedings*, pages 692–696, 2003.

[26] Edmund Miller. Using Prony's Method To Synthesize Discrete Arrays for Prescribed Source Distributions and Exponentiated Patterns. *IEEE Antennas and Propagation Magazine*, 57(1):147–163, 2015.

[27] Bishwarup Mondal and T. V. Sreenivas. Mixture Gaussian envelope chirp model for speech and audio. In IEEE, editor, *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings*, pages 857–860, 2001.

[28] José de Jesús Moreno Vázquez, Aldo Rafael Sartorius Castellanos, Raúl Antonio Ortiz, Marcia Lorena Hernández Nieto, and Antonia Zamudio Radilla. Prony's method implementation for slow wave identification of electroenterogram signals. *Revista Facultad de Ingeniería*, 76:114–122, 2015.

[29] Bernard Mourrain. Polynomial–Exponential Decomposition From Moments. *Foundations of Computational Mathematics*, 18(6):1435–1492, 2018.

[30] Raoul R. Nigmatullin, Sergey I. Osokin, Jan Awrejcewicz, and Grzegorz Kudra. Application of the generalized Prony spectrum for extraction of information hidden in chaotic trajectories of triple pendulum. *Central European Journal of Physics*, 12(8):565–577, 2014.

[31] Michael R. Osborne and Gordon K. Smyth. A modified Prony algorithm for exponential function fitting. *SIAM Journal on Scientific Computing*, 16(1):119–138, 1995.

[32] Thomas Peter and Gerlind Plonka. A Generalized Prony Method for Reconstruction of Sparse Sums of Eigenfunctions of Linear Operators. *Inverse Problems*, 29(2):id. 025001, 2013.

[33] Thomas Peter, Gerlind Plonka, and Daniela Daniela Roşca. Representation of sparse Legendre expansions. *Journal of Symbolic Computation*, 50:159–169, 2013.

[34] V. F. Pisarenko. The Retrieval of Harmonics from a Covariance Function. *Geophysical Journal International*, 33(3):347–366, 1973.

[35] Gerlind Plonka, Kilian Stampfer, and Ingeborg Keller. Reconstruction of Stationary and Non-stationary Signals by the Generalized Prony Method. *Analysis and Applications*, 141:1–32, 2018.

[36] Gerlind Plonka and Marius Wischerhoff. How many Fourier samples are needed for real function reconstruction? *Journal of Applied Mathematics and Computing*, 42(1-2):117–137, 2013.

[37] Daniel Potts and Manfred Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Processing*, 90(5):1631–1642, 2010.

[38] Daniel Potts and Manfred Tasche. Parameter estimation for multivariate exponential sums. *Electronic Transactions on Numerical Analysis*, 40, 2013.

[39] Daniel Potts and Manfred Tasche. Sparse polynomial interpolation in Chebyshev bases. *Linear Algebra and its Applications*, 441:61–87, 2014.

[40] Richard Roy and Thomas Kailath. ESPRIT - Estimation of Signal Parameters via Rotational Invariance Techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, 1989.

[41] Mohamed Sahraoui, Antonio J. Marques Cardoso, Khaled Yahia, and Adel Ghoggal. The use of the modified Prony's method for rotor speed estimation in squirrel cage induction motors. In IEEE, editor, *2015 IEEE 10th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, pages 455–460, 2015.

[42] Ralph O. Schmidt. Multiple Emitter Location and Signal Parameter - - Estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

[43] Chandra Sekhar Seelamantula. OperA: Operator-Based Annihilation for Finite-Rate-of-Innovation Signal Sampling. In Götz Pfander, editor, *Sampling Theory, a Renaissance*, Applied and Numerical Harmonic Analysis, pages 461–484. Birkhäuser, Basel, 2015.

[44] Claude E. Shannon. Communication in the Presence of Noise. *Proceedings of the IRE*, 86(1):10–21, 1949.

[45] Petre Stoica and Randolph Moses. *Spectral Analysis of Signals*. Pearson/Prentice Hall, Upper Saddle River, 2005.

[46] Bob L. Sturm and Jerry D. Gibson. Matching Pursuit Decompositions of Non-Noisy Speech Signals Using Several Dictionaries. In IEEE, editor, *2006 IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 3, pages 456–459, 2006.

[47] Martin Vetterli, Pina Marziliano, and Thierry Blu. Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002.

[48] John M. Whittaker. The "Fourier" Theory of the Cardinal Function. *Proceedings of the Edinburgh Mathematical Society*, 1(3):169–176, 1928.

[49] Shuang Zhao and Kenneth Loparo. Forward and Backward Extended Prony (FBEP) Method for Power System Small-Signal Stability Analysis. *IEEE Transactions on Power Systems*, 32:3618–3626, 2017.

Curriculum Vitae

Kilian Stampfer

Born May 8th 1988 in Kelheim, Germany.

Education

since 2015 **PhD student**, *Georg-August-Universität Göttingen.*
(Supervisor: Prof. Dr. Gerlind Plonka, Prof. Dr. Stephan Huckemann)

2015 **M. Sc.in Mathematics**, *Katholische Universität Eichstätt.*
Thesis: *Der Informationstheoretische Entropiebegriff und die Kapazität ausgewählter Kanalmodelle*
(Supervisor: Prof. Dr. Götz Pfander)

2013 **B. Sc. in Mathematics**, *Katholische Universität Eichstätt.*
Thesis: *Betrachtung des Graphenisomorphieproblems unter der Verwendung von Invarianten und Ähnlichkeitsbegriffen*
(Supervisor: Dipl.-Inf. Dr. Andreas Harrer)

2010 **Allgemeine Hochschulreife**, *Staatliche Berufsoberschule*, Kelheim.

Academic funding

PhD-Position in the Research Training Group 2088, *Discovering structure in complex data: Statistics meets Optimization and Inverse Problems*, Member.

Publications & Preprints

2018 Gerlind Plonka, **Kilian Stampfer** and Ingeborg Keller, *Reconstruction of Stationary and Non-stationary Signals by the Generalized Prony Method*, [35]

Talks & Workshops

2018 **9th International Conference on Curves and Surfaces (talk)**, Arcachon.

2018 **7th International Conference on Comp. Harmonic Analysis (talk)**, Nashville.

2017 **HeKKSaGOn Workshop (talk)**, Karlsruhe.

2016 **Workshop on Manifold-valued Image Processing** , Kaiserslautern.