PRONY'S METHOD IN BANACH MODULES

ILYA A. KRISHTAL AND GÖTZ E. PFANDER

ABSTRACT. We show that the classical Prony's method for recovery of a sparse signal from its consecutive Fourier coefficients can be viewed as a spectral identification problem for an unknown restriction of a known linear operator. This presents a unified point of view on various existing and novel generalizations and applications of the method, some of which are discussed in this paper.

1. Introduction

In 1795, Prony demonstrated in his work [12] that it is possible to reconstruct any s-sparse d-dimensional vector from any 2s of its consecutive Fourier coefficients. Since then, Prony's ideas have been expanded in various directions, leading to advancements in several subfields of applied and computational mathematics (see [9, 11, 23, 30, 36] and references therein for a small sample of these developments).

In the paper [3], the authors take Prony's method a step further and present a novel perspective on it. They show that the core of Prony's method can be seen as a spectral identification problem, where one aims to identify an unknown restriction of a known linear operator in a Euclidean space. In this paper, we extend this viewpoint to a much more abstract level, considering Banach modules over group algebras instead of just Euclidean spaces.

By generalizing Prony's method in this way, we achieve a unification of various existing extensions of Prony's method [9, 30] and address other related problems [3, 23]. One of the problems discussed in the paper is the identification of operators represented as linear combinations of time-frequency shifts [18, 19, 21, 25]. These operators are commonly used to model time-varying communication channels [38]. Accurately identifying the channel is crucial for reliable communication, and it typically involves observing a limited number of input and output pairs of the operator [19, 26, 32, 20, 31]. See [39] for an overview and an history of the subjects, and additional references. In Section 4

Date: March 29, 2024.

 $[\]it Key\ words\ and\ phrases.$ Banach modules, Beurling spectrum, superresolution, annihilating filters.

of the paper, we present a novel approach to tackle this problem that is based on a special case of our abstract method.

To provide the necessary background, we introduce the relevant concepts in the beginning of Section 2. The rest of the section is the centerpiece of the paper where our abstract approach is presented. We proceed to demonstrate how Prony's method and some of its generalizations can be derived as special cases in Section 3. Finally, in the concluding remarks, we discuss further potential applications of the method.

2. Banach modules and spectral identification

We begin by introducing the notation and relevant notions of the spectral theory of Banach modules. We refer to [6, 7, 8] and references therein for more information on the subject.

The symbol \mathcal{X} will denote a complex Banach space and $B(\mathcal{X})$ will be the Banach algebra of all bounded linear operators on \mathcal{X} . By \mathbb{G} we will denote a locally compact Abelian group (LCA-group) and by $\widehat{\mathbb{G}}$ – its *Pontryagin dual*, i.e. the group of all continuous unitary characters of \mathbb{G} . We will write the operation additively on all LCA-groups, except when the group $\mathbb{T} = \{\theta \in \mathbb{C} : |\theta| = 1\}$ is used.

Definition 2.1. A (Beurling) weight is a (Haar) measurable locally bounded even function $w : \mathbb{G} \to [1, \infty)$ such that

$$w(g_1 + g_2) \le w(g_1)w(g_2)$$
, for all $g_1, g_2 \in \mathbb{G}$.

We denote by $L_w(\mathbb{G})$ the Banach algebra of (equivalence classes of) complex-valued functions w-integrable on \mathbb{G} with respect to the normalized Haar measure. The norm is given by

$$\|f\|_{w} = \int_{\mathbb{G}} |f(g)| w(g) dg < \infty,$$

and the role of multiplication in $L_w(\mathbb{G})$ is played by the convolution

$$(f*h)(g) = \int_{\mathbb{G}} f(t)h(g-t)dt.$$

These algebras are typically called *Beurling algebras* [15, 16, 22, 24, 27, 29, 33, 34, 37, and references therein]. If $w \equiv 1$, we obtain the standard group algebra, that is, $L_1(\mathbb{G}) \equiv L^1(\mathbb{G})$.

In this paper, we typically assume that the weight w is non-quazianalytic [13, 27, 28], i.e. it satisfies the Beurling-Domar condition

(2.1)
$$\sum_{n=0}^{\infty} \frac{\ln w(ng)}{1+n^2} < \infty \quad \text{for all } g \in \mathbb{G}, \quad ng = \underbrace{g+g+\ldots+g}_{n \text{ times}}.$$

This condition ensures that the algebra $L_w(\mathbb{G})$ is regular and its spectrum is isomorphic to the dual group $\widehat{\mathbb{G}}$ [13, 24].

We denote by $\hat{f}: \widehat{\mathbb{G}} \to \mathbb{C}$ the Fourier transform of $f \in L_w(\mathbb{G})$:

$$\hat{f}(\gamma) = \int_{\mathbb{G}} f(g)\gamma(-g)dg, \ \gamma \in \widehat{\mathbb{G}}.$$

The inverse Fourier transform of a function $h: \widehat{\mathbb{G}} \to \mathbb{C}$ is denoted by \check{h} or h^{\vee} .

For some results, we may also require that the weight w satisfies $Ditkin's\ condition\ [17]$, which states that given $\gamma\in\widehat{\mathbb{G}}$ and a function $f\in L_w(\mathbb{G})$ with $\widehat{f}(\gamma)=0$ there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of functions in $L_w(\mathbb{G})$ such that $\widehat{f}_n\equiv 0$ in a neighborhood of γ and $\lim f_n*f=f$. Ditkin's condition is satisfied [5] if, for example,

(2.2)
$$\lim_{n \to \infty} \frac{w(ng)}{n} = 0 \text{ for all } g \in \mathbb{G}.$$

In this paper, the space \mathcal{X} is endowed with a Banach $L_w(\mathbb{G})$ -module structure associated with some representation $\mathcal{T}: \mathbb{G} \to B(\mathcal{X})$. We shall occasionally use the notation $(\mathcal{X}, \mathcal{T})$ for such modules. If \mathcal{X} is an appropriate space of functions (or distributions) on \mathbb{G} , the following representations are used most often.

The translation representation $\mathcal{T} = T : \mathbb{G} \to B(\mathcal{X})$ is given by

(2.3)
$$T(t)x(s) = x(s+t), \ x \in \mathcal{X}, \ s, t \in \mathbb{G},$$

and the modulation representation $\mathcal{T} = M : \widehat{\mathbb{G}} \to B(\mathcal{X})$ is defined by

(2.4)
$$M(\xi)x(s) = \xi(s)x(s), \ x \in \mathcal{X}, \ \xi \in \widehat{\mathbb{G}}, \ s \in \mathbb{G}.$$

We will only consider representations $\mathcal{T}: \mathbb{G} \to B(\mathcal{X})$ that are strongly continuous; in other words, we require that the function $x_{\mathcal{T}}: \mathbb{G} \to \mathcal{X}$ given by $x_{\mathcal{T}}(g) = \mathcal{T}(g)x$ is continuous for every $x \in \mathcal{X}$. We also assume $\|\mathcal{T}(g)\| \leq w(g)$ for all $g \in \mathbb{G}$. The Banach $L_w(\mathbb{G})$ -module structure [10, 16, 35] on \mathcal{X} is then defined via

(2.5)
$$fx = \int_{\mathbb{G}} f(g) \mathcal{T}(-g) x dg, \quad f \in L_w(\mathbb{G}), \ x \in \mathcal{X}.$$

We will always assume that this structure is non-degenerate, that is, fx = 0 for all $f \in L_w(\mathbb{G})$ implies x = 0. We also note the following important property of the defined module structure:

(2.6)
$$\mathcal{T}(g)(fx) = (T(g)f)x = f(\mathcal{T}(g)x),$$

where T is the translation representation (2.3).

Definition 2.2. The Beurling spectrum of a set $N \subseteq (\mathcal{X}, \mathcal{T})$ is the subset $\Lambda(N) = \Lambda(N, \mathcal{T})$ of the dual group $\widehat{\mathbb{G}}$ given by

$$\{\gamma \in \widehat{\mathbb{G}} : \text{ given any } f \in L_w(\mathbb{G}) \text{ with } \widehat{f}(\gamma) \neq 0$$

there is $x \in N$ such that $fx \neq 0\}.$

When $N = \{x\}$ is a singleton, we shall write $\Lambda(x)$ instead of $\Lambda(\{x\})$.

The notion of the Beurling spectrum serves as a natural generalization of the notion of a support. Among many things the Beurling spectrum may coincide with, specific choices of Banach modules yield the support of a function, the support of the (distributional) Fourier transform, the set of indices of non-zero matrix diagonals, etc [7]. The following example serves as an illustration.

Example 2.1. Let $\mathcal{X} = L^2(\mathbb{G})$. If $\mathcal{T} = M$ is the modulation representation (2.4), then the $L^1(\widehat{\mathbb{G}})$ -module structure of (\mathcal{X}, M) is given by $(fx)(s) = \widehat{f}(s)x(s)$ so that $\Lambda(x) = \operatorname{supp} x$. If $\mathcal{T} = T$ is the translation representation (2.3), then the $L^1(\mathbb{G})$ -module structure of (\mathcal{X}, T) is given by the convolution (fx)(s) = (f*x)(s). As f*x = 0 if and only if $\widehat{f}\widehat{x} = 0$ we conclude that $\Lambda(x) = \operatorname{supp}\widehat{x}$.

Given a closed set $F \subset \widehat{\mathbb{G}}$ we shall denote by $\mathcal{X}(F)$ the (closed) spectral submodule of all vectors $x \in \mathcal{X}$ such that $\Lambda(x) \subseteq F$. If $F = \{\gamma\}$ is a singleton, we shall write \mathcal{X}_{γ} instead of $\mathcal{X}(\{\gamma\})$.

Remark 2.1. A submodule \mathcal{X}_{γ} , $\gamma \in \widehat{\mathbb{G}}$, contains all $x \in \mathcal{X}$ such that $\mathcal{T}(g)x = \gamma(g)x$ for all $g \in \mathbb{G}$. Moreover, if Ditkin's condition is satisfied, \mathcal{X}_{γ} does not contain any other vectors (see, e.g., [4]).

The key objects of this paper are the spectral submodules $\mathcal{X}(F)$ generated by a *finite* set F of a fixed cardinality æ.

Given some $x \in \mathcal{X}(F)$, the goal is to use certain "measurements" of x to identify initially the set F and ultimately the vector x itself. Only a finite number of measurements shall be used. Therefore, it is natural to assume that $F \subseteq \Omega$ for some known subset $\Omega \subseteq \widehat{\mathbb{G}}$, and each \mathcal{X}_{γ} , $\gamma \in \Omega$, has a finite dimension $m_{\gamma} \leq M \in \mathbb{N}$. The set $\{x_{\gamma}^m : m = 1, \ldots, m_{\gamma}\}$ will be a (known) basis in \mathcal{X}_{γ} , $\gamma \in \widehat{\mathbb{G}}$. Thus, our signal model is

(2.7)
$$x = \sum_{\gamma \in F} \sum_{m=1}^{m_{\gamma}} c_{\gamma m} x_{\gamma}^{m}, \ c_{\gamma m} \in \mathbb{C}.$$

The set F and the coefficients $c_{\gamma m}$ are unknown and to be recovered from certain "measurements" of x. It will be assumed that for each $\gamma \in F$ there exists $m \in \{1, \ldots, m_{\gamma}\}$ such that $c_{\gamma m} \neq 0$.

To compute the measurements, we shall choose two linear operators $A: \mathcal{X} \to \mathbb{C}^S$, $S \in \mathbb{N}$, and $B \in B(\mathcal{X})$. The measurements y_{ℓ} shall be of the form

(2.8)
$$y_{\ell} = AB^{\ell}x, \ \ell = 0, 1, \dots, L.$$

The choice of the operator A is predicated by two goals it needs to achieve. One of them is to allow for unique identification of the coefficients $c_{\gamma m}$ once the set F is recovered. In general, one can only ensure this by requiring that all matrices in a certain class have a left inverse. In most of the specific examples we present in this paper, however, we check that this condition is satisfied. The other goal in the choice of the operator A is to make sure that it does not obscure the recovery of F made possible via the design of the operator B. Sufficient conditions for achieving this goal are included in the statements of the general results.

The operator B shall have the form

(2.9)
$$B = \sum_{n=1}^{N} b_n \mathcal{T}(g_n), \ g_n \in \mathbb{G}, b_n \in \mathbb{C} \setminus \{0\}.$$

By B_F we shall denote the restriction of the operator B to $\mathcal{X}(F)$. By definition, all spectral submodules are invariant for B of the form (2.9), so that the restrictions are well defined. The following important result, which is a special case of the spectral mapping theorem, indicates why using B of the form (2.9) may lead to the recovery of F.

Theorem 2.1 ([6, Theorem 3.3.14]). The spectrum $\sigma(B_F)$ of the operator B_F satisfies

(2.10)
$$\sigma(B_F) = \overline{\left\{ \sum_{n=1}^N b_n \gamma(g_n) : \gamma \in F \right\}}.$$

The right-hand-side of (2.10) motivates us to introduce the function $h = h_B : \widehat{\mathbb{G}} \to \mathbb{C}$ given by

(2.11)
$$h(\gamma) = \sum_{n=1}^{N} b_n \gamma(g_n), \ \gamma \in \widehat{\mathbb{G}}.$$

With this notation, using the fact that the image of a compact set F under a continuous function h is compact, we can write (2.10) more succinctly:

(2.12)
$$\sigma(B_F) = h_B(F).$$

In view of Theorem 2.1, we would like to be able to choose the coefficients b_n in (2.9) so that $\sigma(B_F) = h(F)$ completely determines

F. If this is, indeed, possible, the method of [3] outlined below can be used to determine the "observable" part of F. For convenience, we always choose the coefficients b_n in such a way that $0 \notin \sigma(B_F)$ for any nonempty $F \subseteq \Omega$.

Observe that we have $\dim \mathcal{X}(F) \leq \mathfrak{E}M$. Thus, for $x \in \mathcal{X}(F)$, the system of vectors $\{x, Bx, \dots, B^{\mathfrak{E}M}x\}$ is always linearly dependent and we can find the coefficients $\alpha_0, \alpha_1, \dots, \alpha_{\mathfrak{E}M}$ such that $\alpha_{\mathfrak{E}M} = 1$ and

(2.13)
$$\sum_{\ell=0}^{\infty M} \alpha_{\ell} y_{\ell+k} = \sum_{\ell=0}^{\infty M} \alpha_{\ell} A B^{\ell+k} x = 0, \ k = 0, 1, \dots$$

We let

(2.14)
$$p_{\min}(z) = \sum_{\ell=0}^{\infty M} \alpha_{\ell} z^{\ell}, \quad z \in \mathbb{C},$$

where α_{ℓ} are such that (2.13) holds; if the choice is not unique, we pick $0 = \alpha_0 = \alpha_1 = \dots$ for as many coefficients as possible. Observe that p_{\min} can be found from the measurements (2.8) with $L = 2 \approx M - 1$.

The following result is now an immediate consequence of [3, Proposition 2.4].

Proposition 2.2. The set R_{\min} of all non-zero roots of the polynomial p_{\min} is a subset of $\sigma(B_F)$.

The above proposition can be considerably strengthened when all singleton spectral submodules are one-dimensional, i.e. M=1. In that case, we have $\mathcal{X}_{\gamma} = \{\alpha x_{\gamma} : \alpha \in \mathbb{C}\}$ and $\mathcal{T}(g)x_{\gamma} = \gamma(g)x_{\gamma}, g \in \mathbb{G}$.

Theorem 2.3. Assume that the restriction of h to Ω is one-to-one and $0 \notin h(\Omega)$. Assume also that M = 1 and for each $\gamma \in \Omega$ we have $Ax_{\gamma} \neq 0$. Then $R_{\min} = h(F)$ and F can be recovered from the measurements (2.8) with $L = 2 \otimes -1$.

Proof. From (2.13) and [3, Proposition 2.4], we have

(2.15)
$$\sum_{\ell=0}^{\infty} \alpha_{\ell} A B^{\ell+k} x = \sum_{\gamma \in F} c_{\gamma} h^{k}(\gamma) p_{\min}(h(\gamma)) A x_{\gamma} = 0$$

for all $k=0,1,\ldots$ Our assumptions on h ensure that the Vandermonde matrix generated by $\{h(\gamma): \gamma \in F\}$ is invertible. Since we have also assumed that $c_{\gamma}Ax_{\gamma} \neq 0$, (2.15) may only hold if $p_{\min}(h(\gamma)) = 0$ for all $\gamma \in F$. Hence, our choice of p_{\min} ensures that $R_{\min} = h(F)$. The set F can now be recovered since the restriction of h to Ω is one-to-one. \square

To obtain an analogous result for the case M>1 we will need the following definitions.

Definition 2.3. A spectral value $\gamma \in F$ is called *observable* if the restriction of h to Ω is one-to-one, $0 \notin h(\Omega)$, and $h(\gamma) \in R_{\min}$. The set F is called *observable* if all $\gamma \in F$ are observable.

Definition 2.4. A Krylov subspace of order r generated by an operator B and a vector $x \in \mathcal{X}$ is

$$\mathcal{K}_r(B, x) = \operatorname{span}\{x, Bx, ..., B_{r-1}x\}.$$

The maximal Krylov subspace will be denoted by $\mathcal{K}_{\infty}(B, x)$.

Definition 2.5. For $x \in \mathcal{X}$, the (A, B, x)-annihilator, denoted by $p_{A,x}^B$, is the monic polynomial of the smallest degree among all the polynomials p such that $Ap(B)\mathcal{K}_{\infty}(B, x) = \{0\}$.

We have the following sufficient criterion for observability.

Proposition 2.4. Assume that the restriction of h to Ω is one-to-one and $0 \notin h(\Omega)$. Assume also that $\gamma \in F$ is such that there is a (non-zero) eigenvector $x_{\gamma}^1 \in \mathcal{K}_{\infty}(B,x)$ satisfying $Bx_{\gamma}^1 = h(\gamma)x_{\gamma}^1$ and $Ax_{\gamma}^1 \neq 0$. Then γ is observable.

Proof. By construction, we have $p_{\min}(z) = z^k p_{A,x}^B(z)$ for some $k \geq 0$. Since $x_{\gamma}^1 \in \mathcal{K}_{\infty}(B,x)$, it follows that

$$0 = Ap_{\min}(B)x_{\gamma}^{1} = p_{\min}(h(\gamma))Ax_{\gamma}^{1} = (h(\gamma))^{k}p_{A,x}^{B}(h(\gamma))Ax_{\gamma}^{1}.$$

Since $Ax_{\gamma}^1 \neq 0$ and $0 \notin h(\Omega)$, we must have $p_{\min}(h(\gamma)) = 0$. It follows that $h(\gamma) \in R_{\min}$.

The following lemma is implied by standard linear algebraic facts. We provide a proof for completeness.

Lemma 2.5. For any $\gamma \in F$ there is $x_{\gamma}^1 \in \mathcal{K}_{\infty}(B,x) \setminus \{0\}$ such that $Bx_{\gamma}^1 = h(\gamma)x_{\gamma}^1$.

Proof. Pick $\gamma \in F$ and let $E = h(F) \setminus \{h(\gamma)\}$. We will use a polynomial q defined in the following way. If $E = \emptyset$, we let $q \equiv 1$. Otherwise, we let

$$q(z) = \prod_{\tau \in E} (z - \tau).$$

By construction, we have $q(B)x \in \mathcal{K}_{\infty}(B,x) \setminus \{0\}$. Let \mathcal{Y} be the submodule generated by q(B)x. From Theorem 2.1, we deduce that the restriction $B_{\mathcal{Y}}$ of B to \mathcal{Y} has the spectrum $\sigma(B_{\mathcal{Y}}) = \{h(\gamma)\}$. The conclusion of the lemma now follows from the fundamental theorem of algebra.

Combining Proposition 2.4 with Lemma 2.5 and Remark 2.1, we get

Theorem 2.6. Assume that the restriction of h to Ω is one-to-one and $0 \notin h(\Omega)$. Then $\gamma \in F$ is observable if and only if $Ax \neq 0$ for each $x \in \mathcal{X}$ such that $\mathcal{T}(g)x = \gamma(g)x$ for all $g \in \mathbb{G}$.

The results of this section yield the following Algorithm 1, which is essentially the same as the classical Prony algorithm.

Algorithm 1 Spectral recovery.

- 1: Goal: Find all elements $x \in \mathcal{X}$ of the form (2.7) that satisfy the given measurements (2.8).
- 2: **Input:** The measurements (2.8), function h satisfying conditions of Theorems 2.3 and 2.6, maximal cardinality x of the spectrum x, and maximal dimension x of the spectral submodules x.
- 3: Set up the linear system (2.13) and solve it with $0 = \alpha_0 = \alpha_1 = \dots$ for as many coefficients as possible.
- 4: Form the polynomial p_{\min} given by (2.14) and find the set R_{\min} of its non-zero roots.
- 5: Find $F = h^{-1}(R_{\min})$.
- 6: Find all possible solutions for $c_{\gamma m}$ by solving the linear system yielded by (2.8) when x with the recovered F is plugged in.
- 7: Output: All $x \in \mathcal{X}$ of the form (2.7) that satisfy (2.8).

Remark 2.2. Algorithm 1 fully determines the spectrum F of any $x \in \mathcal{X}$ of the form (2.7) satisfying (2.8). It is also clear, that the final linear system in the algorithm is guaranteed to have a solution. In general, however, there is no guarantee that the solution is unique, i.e. there may be infinitely many $x \in \mathcal{X}$ of the form (2.7) that satisfy the given measurements (2.8) (they will all have the same spectrum F). We show in the following section that sufficient conditions for the uniqueness of the solution may often be formulated in terms of an appropriate choice of the operator A.

3. Reduction to Prony's method.

In this section, we recover Prony's method [9] and develop its more general versions.

3.1. **Standard Prony's method.** (Maybe we should also address the on-grid case, for flow of reading. We could even make a short section, but I think that would not work.)

Here, we let $\mathbb{G} = \mathbb{R}$, $\widehat{\mathbb{G}} \simeq \mathbb{R}$, $\mathcal{X} = C_{ub}(\mathbb{R})$ – the space of all bounded uniformly continuous functions on \mathbb{R} , and $F \subset [0,1)$. The role of $\gamma(g)$,

 $g \in \mathbb{G}$, $\gamma \in \widehat{\mathbb{G}}$, in this case, is played by $e^{2\pi i g \gamma}$, $g, \gamma \in \mathbb{R}$. The signal model is then

(3.1)
$$x(t) = \sum_{\gamma \in F} c_{\gamma} e^{2\pi i t \gamma}, \ t \in \mathbb{R},$$

i.e. x is an almost periodic function with a finite spectrum. The representation \mathcal{T} is the translation $\mathcal{T}(t)y(s)=y(s+t), \ y\in\mathcal{X}$, so that the module structure is the convolution: $fy=f*y, \ f\in L^1(\mathbb{R})$. It is then easy too see from Remark 2.1 that letting $x_{\gamma}(t)=e^{2\pi it\gamma}$ yields $\mathcal{X}_{\gamma}=\{ax_{\gamma}: a\in\mathbb{C}\}$. Thus, (3.1) is indeed a special case of (2.7) and all spectral submodules \mathcal{X}_{γ} are one-dimensional.

We now let $B = \mathcal{T}(1)$ and $A : \mathcal{X} \to \mathbb{C}$ be given by Ay = y(0). This yields standard Prony's measurements

(3.2)
$$y_{\ell} = AB^{\ell}x = x(\ell), \ \ell = 0, 1, \dots$$

The function $h: \mathbb{R} \to \mathbb{C}$ in (2.11) becomes $h(\gamma) = e^{2\pi i \gamma}$. Clearly, h is one-to-one on $\Omega = [0, 1)$, and $0 \notin h(\Omega) = \mathbb{T}$. Thus, Theorem 2.3 applies ascertaining the validity of the standard off-grid Prony's method.

To summarize, the system (2.13) in this case becomes

$$\sum_{\ell=0}^{\infty} \alpha_{\ell} y(\ell+k) = 0, \ k = 0, 1, \dots, \infty - 1.$$

The solution with as many $0 = \alpha_0 = \alpha_1 = \dots$ as possible provides the coefficients of p_{\min} . Finding the non-zero roots of p_{\min} , one obtains the spectrum of x from $R_{\min} = h(F)$. To find x itself it remains to solve another linear system, which is guaranteed to have a unique solution since its matrix is Vandermonde.

Remark 3.1. One could make another choice of B above. Potentially this allows one to better distinguish the frequencies that are close by.

3.2. Exponential sums with polynomial coefficients. A more general form of the off-grid Prony's method arises if we let the signal x be a continuous function of the form

(3.3)
$$x(t) = \sum_{\gamma \in F} q_{\gamma}(t)e^{2\pi i \gamma t},$$

where each q_{γ} is a (non-zero) polynomial of degree $m_{\gamma} \leq M$. The space \mathcal{X} is then the closed linear span of such signals in

$$\mathcal{Y} = \left\{ y : \frac{y}{(1+|\cdot|)^M} \in C_{ub}(\mathbb{R}) \right\},\,$$

with $||y||_{\mathcal{Y}} = \sup_{t \in \mathbb{R}} |y(t)|(1+|t|)^{-M}$. We again have that the representation \mathcal{T} is the translation $\mathcal{T}(t)y(s) = y(s+t)$, $y \in \mathcal{X}$, and the module

structure is the convolution: fy = f * y. It is crucial, however, that this time the module is not over $L^1(\mathbb{R})$ but only over $L_w(\mathbb{R})$, where $w(t) = (1 + |t|)^M$. Thus, even though the functions $x_\gamma = x_\gamma^1$ given by $x_\gamma(t) = e^{2\pi i \gamma t}$ are still the eigenvectors of the representation \mathcal{T} and, therefore, belong to the spectral submodules \mathcal{X}_γ , they no longer span them. In this case, the submodules \mathcal{X}_γ are not one-dimensional. In fact, one can check that the functions x_γ^m , $m = 1, \dots M+1$, given by $x_\gamma^m(t) = t^{m-1}e^{2\pi i \gamma t}$ form a basis of \mathcal{X}_γ . In particular, to see that $x_\gamma^m \in \mathcal{X}_\gamma$, notice that integration by parts yields $(f*x_\gamma^m)(t) = e^{2\pi i \gamma t} \sum_{k=0}^m t^k \widehat{f}_k(\gamma)$, where supp $\widehat{f}_k = \text{supp } \widehat{f}$ for each $k = 0, \dots, m$. Choosing f with $\gamma \notin \text{supp } \widehat{f}$, therefore, yields $f*x_\gamma^m = 0$ resulting in $\Lambda(x_\gamma^m) = \{\gamma\}$. It follows that (3.3) is also a special case of (2.7).

We now choose A and B the same way as in the previous subsection. We have that all eigenspaces of B are of the form $\{\alpha x_{\gamma}^{1}: \alpha \in \mathbb{C}\}$. Moreover, $Ax_{\gamma}^{1}=1\neq 0$, and Theorem 2.6 applies yielding the following corollary.

Corollary 3.1. Let $x(t) = \sum_{\gamma \in F} q_{\gamma}(t)e^{2\pi i\gamma t}$, where $t, \gamma \in \mathbb{R}$, q_{γ} is a non-zero polynomial of degree at most M, and F is a finite subset of [0,1) of cardinality \mathfrak{A} . Then x can be completely recovered from its values $x(\ell)$, $\ell = 0, 1, \ldots 2\mathfrak{A}(M+1) - 1$.

Proof. From Theorem 2.6 we deduce that F is observable. It remains to prove that the polynomials q_{γ} can be recovered from the values of x once F is known. This is shown in the following Lemma 3.2.

Lemma 3.2. Let $x(t) = \sum_{n=1}^{N} q_n(t)(\theta_n)^t$, where $t \in \mathbb{R}$, $\theta_n \in \mathbb{T}$ are distinct, and q_n are unknown non-zero polynomials of degree at most M, n = 1, 2, ..., N. Then x can be completely recovered from its values $x(\ell)$, $\ell = 0, 1, ..., N(M+1) - 1$.

Proof. The values $x(\ell)$ allow one to write a linear system of equations for the unknown coefficients of the polynomials q_n . The lemma follows from the invertibility of the matrix of that system, which can be written as $(V_1 \ V_2 \dots V_N)$, where K = N(M+1) - 1 and

$$V_{n} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \theta_{n} & \theta_{n} & \theta_{n} & \dots & \theta_{n} \\ \theta_{n}^{2} & 2\theta_{n}^{2} & 2^{2}\theta_{n}^{2} & \dots & 2^{M}\theta_{n}^{2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta_{n}^{K} & K\theta_{n}^{K} & K^{2}\theta_{n}^{K} & \dots & K^{M}\theta_{n}^{K} \end{pmatrix}, \quad n = 1, \dots, N.$$

By Gaussian elimination in each V_n , n = 1, ..., N, the above matrix reduces to $(\widetilde{V}_1 \ \widetilde{V}_2 ... \widetilde{V}_N)$, where

$$\widetilde{V}_{n} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \theta_{n} & 1 & 0 & \dots & 0 \\ \theta_{n}^{2} & 2\theta_{n} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta_{n}^{M} & M\theta_{n}^{M-1} & {M \choose 2}\theta_{n}^{M-2} & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta_{n}^{K} & K\theta_{n}^{K-1} & {K \choose 2}\theta_{n}^{K-2} & \dots & {K \choose M}\theta_{n}^{K-M} \end{pmatrix}, \quad n = 1, \dots, N.$$

We show that this matrix is invertible by augmenting the standard elementary argument for computing the determinant of a Vandermonde matrix with Gaussian elimination in each \widetilde{V}_n , $n = 1, \ldots, N$.

The proof is by induction in M. In the base case M=0 the matrix is Vandermonde and is, therefore, invertible. The beginning of the inductive step is precisely the same as in the Vandermonde induction: the first column of \widetilde{V}_1 , i.e. the block corresponding to θ_1 , is subtracted from the first columns of all other blocks and common factors are eliminated in each modified column; then from each row except the first one we subtract the previous row multiplied by θ_1 . After this step, the M+1 columns corresponding to θ_1 become

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & \theta_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \theta_1^{M-1} & {\binom{M-1}{1}} \theta_1^{M-2} & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \theta_1^{K-1} & {\binom{K-1}{1}} \theta_1^{K-2} & \dots & {\binom{K-1}{M-1}} \theta_1^{K-M} \end{pmatrix}.$$

In \widetilde{V}_n , i.e. the block of M+1 columns corresponding to θ_n , $n=2,3,\ldots N$, the first column becomes $(0\ 1\ \theta_n\ \ldots\ \theta_n^{K-1})^T$. For the other columns we perform the additional step, which we used in the initial reduction, i.e. Gaussian elimination. In this case, we consecutively subtract the previous column and eliminate the common factors. Thus, the element $\binom{k}{\ell}\theta_n^{k-\ell}$ becomes

$$\frac{1}{\theta_n - \theta_1} \left(\binom{k}{\ell} \theta_n^{k-\ell} - \binom{k-1}{\ell} \theta_1 \theta_n^{k-1-\ell} - \binom{k-1}{\ell-1} \theta_n^{k-\ell} \right)$$
$$= \binom{k-1}{\ell} \theta_n^{k-1-\ell}.$$

Repeating the Vandermonde and Gaussian reductions for each n = 2, 3, ..., N completes the inductive step for M.

3.3. **Sparse Dynamical Sampling.** Here, we exhibit an example that generalizes on-grid Prony's method and provides a different point of view on some of the results in [3]. For simplicity of exposition, we present it in the case of a finite-dimensional module \mathcal{X} .

We consider an abstract initial value problem

(3.4)
$$\begin{cases} \dot{x}(t) = \mathcal{A}x(t) \\ x(0) = x_0, \end{cases} \quad t \in \mathbb{R}_+, \ x \in \mathcal{X},$$

where \mathcal{X} is a high-dimensional Euclidean space and $\mathcal{A} \in B(\mathcal{X})$ is an operator with the spectrum $\sigma(\mathcal{A})$ and a known basis of generalized eigenvectors $\mathscr{B}_{\mathcal{A}} = \{x_{\lambda}^m : \lambda \in \sigma(\mathcal{A}), m = 1, \dots, m_{\lambda}\}$. We consider the standard ordering of the basis so that $\mathcal{A}x_{\lambda}^1 = \lambda x_{\lambda}^1$ and $(\mathcal{A} - \lambda I)x_{\lambda}^m = x_{\lambda}^{m-1}, m = 2, \dots m_{\lambda}$.

We assume that the initial value x_0 is of the form (2.7), i.e.

$$x_0 = \sum_{\lambda \in F} \sum_{m=1}^{m_{\lambda}} c_{\lambda m} x_{\lambda}^m,$$

where F is a relatively small subset of $\sigma(A)$. We wish to identify x_0 from the dynamical samples

$$y_{\ell}(s) = \langle x(\beta \ell), e_s \rangle, \quad s \in \mathcal{I}, \beta > 0, \ell = 0, 1, \dots L,$$

where x(t) is the solution of the IVP (3.4) and $\{e_1, \ldots, e_d\}$ is an orthonormal basis of \mathcal{X} . The set $\mathcal{I} \subseteq \{1, \ldots d\}$ is assumed to be such that no eigenvector in $\mathcal{B}_{\mathcal{A}}$ is orthogonal to the span of $\{e_s : s \in \mathcal{I}\}$.

Remark 3.2. To recover the classical on-grid Prony problem from this setting, we simply let \mathcal{A} be the diagonal operator with the spectrum $\frac{2\pi i}{d} \cdot \{0, \ldots, d-1\}$, $\mathcal{B}_{\mathcal{A}}$ – the standard basis, $\{e_1, \ldots, e_d\}$ – the Fourier basis, and \mathcal{I} – a singleton.

The above problem fits in the framework of dynamical sampling introduced in [2] and further developed in e.g. [1]. Our goal here is to show that this problem can be solved by Algorithm 1. The goal is easily achieved by putting this problem in the notation of Section 2. To this end, we let $A: \mathcal{X} \to \mathbb{C}^{|\mathcal{I}|}$ be defined by $Ax = (\langle x, e_s \rangle)_{s \in \mathcal{I}}$, and $B \in B(\mathcal{X})$ – by $B = e^{\beta \mathcal{A}}$, so that $x(\beta \ell) = B^{\ell}x_0$. The function $h: \mathbb{R} \to \mathbb{C}$ in (2.11) is then $h(\lambda) = e^{\beta \lambda}$ and the representation $\mathcal{T}: \mathbb{R} \to B(\mathcal{X})$ is $\mathcal{T}(t) = e^{t\mathcal{A}}$. The non-orthogonality assumption on \mathcal{I} yields that the assumption of Theorem 2.6 holds and we conclude that Algorithm 1 can indeed be used to recover all x_0 fitting the dynamical

samples as long as β was chosen in such a way that h is one-to-one on $\sigma(A)$. The criteria for uniqueness of such x_0 can be derived from [1].

Remark 3.3. Strictly speaking we reached the above conclusion prematurely because the group \mathcal{T} will grow exponentially (unless $\sigma(\mathcal{A}) \subset i\mathbb{R}$). This violates the non-quasi-analyticity condition (2.1) and prevents us from using Theorem 2.1. For this case, however, we can use a much more basic spectral theorem instead, such as [14, Ch. I, Lemma 3.13]. We also note that a similar example can be considered in an infinite dimensional space but it will not provide new insights, beyond those in the previous subsection. Finally, we point out that in this example Algorithm 1 does not need to have access to the spectrum $\sigma(\mathcal{A})$ but only to the basis of generalized eigenvectors $\mathscr{B}_{\mathcal{A}}$ (provided that $\|\mathcal{A}\|$ is known and β was chosen to keep h one-to-one in $\{\lambda \in \mathbb{C} : |\lambda| \leq \|\mathcal{A}\|\}$).

4. Identification of time-varying channels

The identification of time-varying channels is a central task in communications engineering. In mobile communications, for example, the transmission channel is a superposition of various signal paths, each of which is characterized by a gain factor, a time-delay and a frequency shift, the latter being a consequence of the Doppler effect. Below, we give a application of the method developed above to identify a transmission channel as described above, that is, the parameters of each signal path. Clearly, the identification of a communications channel is of utmost importance in order to transmit information close to channel capacity. For background on the problem see [19, 20, 26, 31, 32, 38, 39].

In order to apply Prony's method for Banach modules in this setting, we give a formal definition of the channel as follows. First, we define the time-frequency shift operators $\pi_{\lambda} = \pi_{(t,\nu)} = M_{\nu}T_t$ acting on a signal $u : \mathbb{R} \to \mathbb{R}$ by $(\pi_{\lambda}u)(r) = e^{2\pi i r \nu}u(r+t)$ with $\lambda = (t,\nu) \in \mathbb{R}^2$. Observe that $(\pi_{\lambda})^{-1} = e^{2\pi i t \nu}\pi_{-\lambda}$ [18]. Second, a generic time-varying communications channel is expressed as the operator

$$u \mapsto \sum_{\gamma \in F} c_{\gamma} \pi(\gamma) u,$$

where the gain factor c_{γ} represents the energy transfer on the path (or the collection of paths) with time-frequency shift γ . The set S of all potentially appearing time-frequency shifts is a bounded subset of the time-frequency plane, the so called spreading domain. The assumption of boundedness is justified due to physical limitations to the maximal size of time and frequency shifts that a transmission signal may be subjected to. In the current context, the size of the spreading domain is irrelevant and we choose, without loss of generality, the set $S = [0,1)^2$. Channel identification in this setting, therefore, reduces to the identification of the operator $X = \sum_{\gamma \in F} c_{\gamma} \pi_{\gamma}$, where F is a finite subset

in $[0,1)^2$. A standard method applied to the identification problem is to transmit a well chosen signal u and to recover the unknown parameters of F from the observed channel output Xu.

To put this problem into our general framework, we choose $\mathbb{G} = \mathbb{R}^2$, and, hence, $\widehat{\mathbb{G}} \simeq \mathbb{R}^2$. We then treat $F \subset [0,1)^2$ as a subset of $\widehat{\mathbb{G}}$, and, in order to be compatible with the time-frequency theory, we use the following isomorphism of \mathbb{R}^2 and $\widehat{\mathbb{G}}$: $(t_{\gamma}, \nu_{\gamma}) \mapsto \gamma \in \widehat{\mathbb{G}}$, where

(4.1)
$$\gamma(g) = \gamma(x,\xi) = e^{2\pi i(x\nu_{\gamma} - t_{\gamma}\xi)}, \quad g = (x,\xi) \in \mathbb{R}^2.$$

In the following, \mathcal{H} is the Hilbert space $L^2(\mathbb{R})$ and $B(\mathcal{H})$ is the algebra of all bounded linear operators in \mathcal{H} . The space \mathcal{X} is then chosen to be the closure of span $\{\pi_{\lambda} : \lambda \in \mathbb{R}^2\}$ in $B(\mathcal{H})$. It is endowed with a Banach module structure by means of the representation $\mathcal{T} : \mathbb{R}^2 \to B(\mathcal{X})$ given by $\mathcal{T}(g)X = \pi_g X \pi_g^{-1}, g \in \mathbb{R}^2$.

Observe that for $X = \sum_{\gamma \in F} c_{\gamma} \pi_{\gamma}$ with $F \subset [0,1)^2$ finite, we have

$$\mathcal{T}(g)X = \sum_{\gamma \in F} c_{\gamma} \pi_{g} \pi_{\gamma}(\pi_{g})^{-1} = \sum_{\gamma \in F} c_{\gamma} e^{2\pi i (t_{\gamma} \nu_{g} - t_{g} \nu_{\gamma})} \pi_{\gamma} = \sum_{\gamma \in F} c_{\gamma} \gamma(g) \pi_{\gamma},$$

i.e., each π_{γ} is an eigenvector of $\mathcal{T}(g)$ that corresponds to the eigenvalue $\gamma(g)$. This justifies our choice of isomorphism in (4.1). It also follows that $\Lambda(\pi_{\gamma}, \mathcal{T}) = \{\gamma\}$. Moreover, the spectral submodules \mathcal{X}_{γ} are one-dimensional. Thus, we showed that time-varying communication channels are a special case of signals of the form (2.7).

To define the operator $A: \mathcal{X} \to \mathbb{C}^S$, we fix the function $u \in \mathcal{H}$ to be the Gaussian

$$u(t) = e^{-t^2}, \ t \in \mathbb{R},$$

and pick a finite subset $K \subset \mathbb{R}^2$ of cardinality S. We let

$$(AX)(s) = \langle X\pi_s u, \pi_s u \rangle, \ s \in K.$$

We note that similar types of measurements were used in [19], but the recovery algorithm developed there is different.

For B of the form (2.9), we have $h(\gamma) = \sum_{n=1}^{N} b_n \gamma(g_n)$, as before. We can still choose $B = \mathcal{T}(g_1)$ with $g_1 = (1, 1)$. In this case, however, the function $h(\gamma) = \gamma(g)$ is not one-to-one on $\Omega = [0, 1)^2$. Nevertheless,

the measurements

$$y_{\ell}(s) = (AB^{\ell}X)(s) = \sum_{n=1}^{N} \sum_{\gamma \in F} b_n c_{\gamma} \gamma^{\ell}(g_n) \langle \pi_{\gamma} \pi_s u, \pi_s u \rangle$$

still allow one to find the polynomial p_{\min} with the set R_{\min} of non-zero roots guaranteed to be a subset of h(F). Observe also that with this choice of A and B all measurements $y_{\ell}(s)$ are linear combinations of a finite collection of numbers of the form $\langle X\pi_r u, \pi_g u \rangle$, for some $r, q \in \mathbb{R}^2$.

Restricting the set Ω in such a way that h satisfies the assumptions of Theorem 2.6 yields $R_{\min} = h(F)$. We also remark that choosing several different B's one can get F after recovering the corresponding subsets of h(F).

Example 4.1. Choosing N=2, $b_1=1$, $b_2=i$, $g_1=(0,1/12)$, and $g_2=(1/12,0)$, we get the function

$$(4.2) \quad h(t,\nu) = e^{\frac{2\pi it}{12}} + ie^{-\frac{2\pi i\nu}{12}} = \cos\frac{\pi t}{6} + i\sin\frac{\pi t}{6} + i\cos\frac{\pi\nu}{6} + \sin\frac{\pi\nu}{6}$$

which is injective on $[0,1)^2$. To see that, assume that $t,t',\nu,\nu' \in [0,1)$ are such that $h(t,\nu) = h(t',\nu')$. Then we must have

$$\begin{cases} \cos(\pi t/6) + \sin(\pi \nu/6) = \cos(\pi t'/6) + \sin(\pi \nu'/6) \\ \sin(\pi t/6) + \cos(\pi \nu/6) = \sin(\pi t'/6) + \cos(\pi \nu'/6) \end{cases}$$

or, equivalently,

$$\begin{cases} \cos(\pi t/6) - \cos(\pi t'/6) = \sin(\pi \nu'/6) - \sin(\pi \nu/6) \\ \sin(\pi t/6) - \sin(\pi t'/6) = \cos(\pi \nu'/6) - \cos(\pi \nu/6). \end{cases}$$

Using the standard trig identities the above is equivalent to

$$\begin{cases} \sin\frac{\pi(t+t')}{12}\sin\frac{\pi(t-t')}{12} = \cos\frac{\pi(\nu+\nu')}{12}\sin\frac{\pi(\nu-\nu')}{12} \\ \cos\frac{\pi(t+t')}{12}\sin\frac{\pi(t-t')}{12} = \sin\frac{\pi(\nu+\nu')}{12}\sin\frac{\pi(\nu-\nu')}{12}. \end{cases}$$

Observe that for $t, t', \nu, \nu' \in [0, 1)$ both sides of the last equation can only equal 0 if t = t' and $\nu = \nu'$, which would not contradict the injectivity of h. Otherwise, we can divide the first equation by the second one and obtain

$$\tan \frac{\pi(t+t')}{12} = \cot \frac{\pi(\nu+\nu')}{12},$$

which would only hold if

$$\frac{\pi(t+t')}{12} = \frac{\pi}{2} - \frac{\pi(\nu+\nu')}{12} + \pi k$$

for some $k \in \mathbb{Z}$. But that would require $t + t' + \nu + \nu' = 6 + 12k$ which is impossible for $t, t', \nu, \nu' \in [0, 1)$. We conclude that the function h in (4.2) is, indeed, injective on $[0, 1)^2$. In fact, an elementary computation yields that the inverse function $h^{-1}: h([0, 1)^2) \subset \mathbb{C} \to [0, 1)^2$ is given by

(4.3)
$$h^{-1}(x+iy) = \frac{3}{\pi} \left(\arcsin \frac{x^2+y^2-2}{2} - \arcsin \frac{x^2-y^2}{x^2+y^2} \right).$$

To describe a more general model, we let \mathcal{H}_n be the *n*-th order Sobolev space and define the following two operators. We let $D: \mathcal{H}_1 \to \mathcal{H}$ be the derivative operator: $(Du)(t) = \dot{u}(t)$ and $\Delta: \mathcal{F}\mathcal{H}_1 \to \mathcal{H}$ be given by $\Delta = \mathcal{F}^{-1}D\mathcal{F}$, where $\mathcal{F} \in B(\mathcal{H})$ is the classical unitary Fourier transform. The space \mathcal{X} will then be a subspace of $L(\mathcal{H}_M \cap \mathcal{F}\mathcal{H}_M, \mathcal{H})$ spanned by $\Delta^j D^k \pi_\gamma$, $\gamma \in \mathbb{R}^2$, $j, k \in \{0, 1, ..., M\}$. The representation \mathcal{T} remains to be defined in the same way, albeit with a different range space. Since

$$\mathcal{T}(g)\Delta = \pi_g \mathcal{F}^{-1} D \mathcal{F}(\pi_g)^{-1} = 2\pi i t I + \Delta$$
 and
 $\mathcal{T}(g)D = \pi_g D(\pi_g)^{-1} = 2\pi i \nu I + D, \ g = (t, \nu),$

we have $\Lambda(D, \mathcal{T}) = \Lambda(\Delta, \mathcal{T}) = \{0\}$, and it follows [7, Corollary 7.8] that $\Lambda(\Delta^j D^k \pi_\gamma) = \{\gamma\}, \ \gamma \in \mathbb{R}^2, \ j, k \in \{0, 1, \dots, M\}.$

Thus, in this setting, one seeks to recover an operator of the form

$$X = \sum_{\gamma \in F} q_{\gamma}^{\Delta}(\Delta) q_{\gamma}^{D}(D) \pi_{\gamma},$$

where q_{γ}^{Δ} and q_{γ}^{D} are non-zero polynomials of degree at most M. The same choice of A and B as earlier in this section, once again, allows one to find a polynomial p_{\min} whose non-zero roots are guaranteed to be in h(F), where h is still given by (2.11). As before, this is a direct consequence of Proposition 2.2 and Theorem 2.1.

5. Concluding remarks

In this short note, we presented only a few instances of how our point of view on Prony's method can be applied. Many more remain to be developed. One problem that can be investigated is identifying rational functions using the fact that they are Laplace transforms of almost periodic functions with a finite spectrum. Another problem is identifying polynomial splines using the observation that their high order derivatives are linear combinations of Dirac masses. We invite the readers to pursue these and other related problems.

Acknowledgements

This paper is dedicated to Karlheinz Gröchenig. Charly has not only been the leader of the vast area of time-frequency analysis since the 1990s, but also significantly influenced the work of the authors of this paper. We have greatly appreciated his direct, honest, and unapologetic criticism as well as the substantial (albeit rare) praise for the handful of our papers that actually deserved to be published (in Charly's eyes). Above all, we value our ongoing friendship, with hopes that it will endure for many more years to come.

The first author acknowledges the support by the NSF grant DMS-2208031 and thanks the second author for his hospitality during the visits to KU Eichstätt-Ingolstadt. The second author acknowledges support by the DFG grants PF 450/11-1 and PF 450/9-1.

References

- [1] A. Aldroubi, C. Cabrelli, U. Molter, and S. Tang, *Dynamical sampling*, Applied and Computational Harmonic Analysis, 42 (2017), pp. 378–401. doi: 10.1016/j.acha.2015.08.014.
- [2] A. Aldroubi, J. Davis, and I. Krishtal, *Dynamical sampling: time-space trade-off*, Appl. Comput. Harmon. Anal., 34 (2013), pp. 495–503.
- [3] A. Aldroubi and I. Krishtal, Krylov subspace methods in dynamical sampling, Sampl. Theory Signal Image Process., 15 (2016), pp. 9–20.
- [4] A. G. Baskakov, Spectral tests for the almost periodicity of the solutions of functional equations, Mat. Zametki, 24 (1978), pp. 195–206, 301. English translation: Math. Notes 24 (1978), no. 1–2, pp. 606–612 (1979).
- [5] —, On Ditkin's condition in some Beurling algebras, Izv. Vyssh. Uchebn. Zaved. Mat., (1982), pp. 3–5.
- [6] —, Theory of representations of Banach algebras, and abelian groups and semigroups in the spectral analysis of linear operators, Sovrem. Mat. Fundam. Napravl., 9 (2004), pp. 3–151 (electronic). English translation: J. Math. Sci. (N. Y.) 137 (2006), no. 4, pp. 4885–5036.
- [7] A. G. BASKAKOV AND I. A. KRISHTAL, Harmonic analysis of causal operators and their spectral properties, Izv. Ross. Akad. Nauk Ser. Mat., 69 (2005), pp. 3–54. English translation: Izv. Math. 69 (2005), no. 3, pp. 439–486.
- [8] —, Memory estimation of inverse operators, J. Funct. Anal., 267 (2014), pp. 2551–2605.
- [9] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, Sparse sampling of signal innovations, Signal Processing Magazine, IEEE, 25 (2008), pp. 31–40.
- [10] O. Bratteli and D. W. Robinson, *Unbounded derivations of C*-algebras*, Comm. Math. Phys., 42 (1975), pp. 253–268.
- [11] A. CUYT AND W.-S. LEE, Multivariate exponential analysis from the minimal number of samples, Adv. Comput. Math., 44 (2018), pp. 987–1002.
- [12] G. C. F. M. R. DE Prony, Essai expérimental et analytique: sur les lois de la dilatabilité de fluides élastiques et sur les celles de la force expansive de la

- vapeur de l'eau et de la vapeur de l'alkool, à différentes températures., J. de l'École Polytechnique Floréal et Plairial, An. III (1795), pp. 24–76.
- [13] Y. Domar, Harmonic analysis based on certain commutative Banach algebras, Acta Math., 96 (1956), pp. 1–66.
- [14] K.-J. ENGEL AND R. NAGEL, One-parameter semigroups for linear evolution equations, vol. 194 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.
- [15] H. G. FEICHTINGER, Gewichtsfunktionen auf lokalkompakten Gruppen, Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II, 188 (1979), pp. 451–471.
- [16] —, Compactness in translation invariant Banach spaces of distributions and compact multipliers, J. Math. Anal. Appl., 102 (1984), pp. 289–327.
- [17] I. GELFAND, D. RAIKOV, AND G. SHILOV, Commutative normed rings, Translated from the Russian, with a supplementary chapter, Chelsea Publishing Co., New York, 1964.
- [18] K. GRÖCHENIG, Foundations of Time-Frequency Analysis, Applied and Numerical Harmonic Analysis, Birkhäuser, Boston, MA, 2001.
- [19] K. GRÖCHENIG AND E. PAUWELS, Uniqueness and reconstruction theorems for pseudodifferential operators with a bandlimited Kohn-Nirenberg symbol, Adv. Comput. Math., 40 (2014), pp. 49–63.
- [20] R. HECKEL AND H. BÖLCSKEI, *Identification of sparse linear operators*, IEEE Transactions on Information Theory, 59 (2013), pp. 7985–8000.
- [21] C. Heil, Integral Operators, Pseudodifferential Operators, and Gabor Frames, Birkhäuser Boston, Boston, MA, 2003, pp. 153–169.
- [22] S.-Z. Huang, J. van Neerven, and F. Räbiger, *Ditkin's condition for certain Beurling algebras*, Proc. Amer. Math. Soc., 126 (1998), pp. 1397–1407.
- [23] C. Josz, J. B. Lasserre, and B. Mourrain, Sparse polynomial interpolation: sparse recovery, super-resolution, or Prony?, Adv. Comput. Math., 45 (2019), pp. 1401–1437.
- [24] E. Kaniuth, A course in commutative Banach algebras, vol. 246 of Graduate Texts in Mathematics, Springer, New York, 2009.
- [25] W. KOZEK, Adaptation of Weyl-Heisenberg frames to underspread environments, Birkhäuser Boston, Boston, MA, 1998, pp. 323–352.
- [26] W. Kozek and G. Pfander, *Identification of operators with bandlimited symbols*, SIAM J. Math. Anal., 37 (2006), pp. 867–888.
- [27] Y. I. LYUBICH, Vvedenie v teoriyu banakhovykh predstavlenii grupp, "Vishcha Shkola", Khar'kov, 1985. English translation: Introduction to the theory of Banach representations of groups. Operator Theory: Advances and Applications, 30. Birkhäuser Verlag, Basel, 1988. x+223 pp.
- [28] Y. I. LYUBICH, V. I. MACAEV, AND G. M. FEL'DMAN, Representations with a separable spectrum, Funkcional. Anal. i Priložen., 7 (1973), pp. 52–61.
- [29] D. OLESEN, Inner*-automorphisms of simple C*-algebras, Comm. Math. Phys., 44 (1975), pp. 175–190.
- [30] T. Peter and G. Plonka, A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators, Inverse Problems, 29 (2013). 025001.

- [31] G. Pfander and P. Zheltov, *Identification of stochastic operators*, Applied and Computational Harmonic Analysis, 36 (2014), pp. 256 279.
- [32] G. E. PFANDER AND D. F. WALNUT, Measurement of time-variant linear channels, IEEE Trans. Inform. Theory, 52 (2006), pp. 4808–4820.
- [33] H. Reiter, Classical harmonic analysis and locally compact groups, Clarendon Press, Oxford, 1968.
- [34] H. REITER AND J. D. STEGEMAN, Classical harmonic analysis and locally compact groups, vol. 22 of London Mathematical Society Monographs. New Series, The Clarendon Press Oxford University Press, New York, second ed., 2000.
- [35] M. A. RIEFFEL, Induced Banach representations of Banach algebras and locally compact groups, J. Functional Analysis, 1 (1967), pp. 443–491.
- [36] T. SAUER, Prony's method in several variables, Numer. Math., 136 (2017), pp. 411–438.
- [37] J. D. Stegeman, Wiener-Ditkin sets for certain Beurling algebras, Monatsh. Math., 82 (1976), pp. 337–340.
- [38] H. VAN TREES, Detection, Estimation and Modulation Theory, Wiley, New York, 1971.
- [39] D. Walnut, G. E. Pfander, and T. Kailath, Cornerstones of sampling of operator theory, in Excursions in harmonic analysis. Vol. 4, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, Cham, 2015, pp. 291–332.

Department of Mathematical Sciences, Northern Illinois University, DeKalb, IL 60115, email: ikrishtal@niu.edu

LEHRSTUHL WISSENSCHAFTLICHES RECHNEN, MATHEMATISCH-GEOGRAPHISCHE FAKULTÄT, KARHOLISCHE UNIVERSITÄT EICHSTÄTT INGOLSTADT, GERMANY, EMAIL: GOETZ@PFANDER.ORG