Journal *of* Singularities
Volume 10 (2014), 1-25

Proc. of 12th International Workshop on Singularities, São Carlos, 2012
DOI: 10.5427/jsing.2014.10a

GEOMETRY AND SINGULARITIES OF THE PRONY MAPPING

DMITRY BATENKOV AND YOSEF YOMDIN

Abstract. The Prony mapping provides the global solution of the Prony system of equations

$$\Sigma_{i=1}^n A_i x_i^k = m_k,\ k = 0, 1, \ldots, 2n-1.$$

This system appears in numerous theoretical and applied problems arising in Signal Reconstruction. The simplest example is the problem of reconstruction of linear combination of δ-functions of the form $g(x) = \sum_{i=1}^n a_i \delta(x - x_i)$, with the unknown parameters $a_i,\ x_i,\ i = 1, \ldots, n$, from the "moment measurements" $m_k = \int x^k g(x) dx$.

The global solution of the Prony system, i.e., the inversion of the Prony mapping, encounters several types of singularities. One of the most important ones is a collision of some of the points x_i. The investigation of this type of singularities has been started in [21] where the role of finite differences was demonstrated.

In the present paper we study this and other types of singularities of the Prony mapping, and describe its global geometry. We show, in particular, close connections of the Prony mapping with the "Vieta mapping" expressing the coefficients of a polynomial through its roots, and with hyperbolic polynomials and "Vandermonde mapping" studied by V. Arnold.

1. Introduction

Prony system appears as we try to solve a very simple "algebraic signal reconstruction" problem of the following form: assume that the signal $F(x)$ is known to be a linear combination of shifted δ-functions:

$$F(x) = \sum_{j=1}^d a_j \delta(x - x_j). \tag{1.1}$$

We shall use as measurements the polynomial moments:

$$m_k = m_k(F) = \int x^k F(x)\, \mathrm{d}\, x. \tag{1.2}$$

After substituting F into the integral defining m_k we get

$$m_k(F) = \int x^k \sum_{j=1}^d a_j \delta(x - x_j)\, \mathrm{d}\, x = \sum_{j=1}^d a_j x_j^k.$$

Considering a_j and x_j as unknowns, we obtain equations

$$m_k(F) = \sum_{j=1}^d a_j x_j^k,\ k = 0, 1, \ldots. \tag{1.3}$$

2000 *Mathematics Subject Classification.* 94A12 62J02, 14P10, 42C99.

Key words and phrases. Singularities, Signal acquisition, Non-linear models, Moments inversion.

This research is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, ISF grant 264/09 and the Minerva Foundation.

This infinite set of equations (or its part, for $k = 0, 1, \ldots, 2d-1$), is called Prony system. It can be traced at least to R. de Prony (1795, [19]) and it is used in a wide variety of theoretical and applied fields. See [2] for an extensive bibligoraphy on the Prony method.

In writing Prony system (1.3) we have assumed that all the nodes $x_1, \ldots, x_d$ are pairwise different. However, as the left-hand side $\mu = (m_0, \ldots, m_{2d-1})$ of (1.3) is provided by the actual measurements of the signal F, we cannot guarantee a priori, that this condition is satisfied for the solution. Moreover, we shall see below that multiple nodes may naturally appear in the solution process. In order to incorporate possible collisions of the nodes, we consider "confluent Prony systems".

Assume that the signal $F(x)$ is a linear combination of shifted δ-functions and their derivatives:

$$F(x) = \sum_{j=1}^{s} \sum_{\ell=0}^{d_j-1} a_{j,\ell} \delta^{(\ell)} (x - x_j). \tag{1.4}$$

Definition 1.1. For $F(x)$ as above, the vector $D(F) \stackrel{\text{def}}{=} (d_1, \ldots, d_s)$ is *the multiplicity vector* of F, $s = s(F)$ is the size of its support, $T(F) \stackrel{\text{def}}{=} (x_1, \ldots, x_s)$, and $\operatorname{rank}(F) \stackrel{\text{def}}{=} \sum_{j=1}^{s} d_j$ is its rank. For avoiding ambiguity in these definitions, it is always understood that $a_{j,d_j-1} \neq 0$ for all $j = 1, \ldots, s$ (i.e. d_j is the maximal index for which $a_{j,d_j-1} \neq 0$).

For the moments $m_k = m_k(F) = \int x^k F(x) \,\mathrm{d}\, x$ we now get

$$m_k = \sum_{j=1}^{s} \sum_{\ell=0}^{d_j-1} a_{j,\ell} \frac{k!}{(k-\ell)!} x_j^{k-\ell}.$$

Considering x_i and $a_{j,\ell}$ as unknowns, we obtain a system of equations

$$\sum_{j=1}^{s} \sum_{\ell=0}^{d_j-1} \frac{k!}{(k-\ell)!} a_{j,\ell} x_j^{k-\ell} = m_k, \quad k = 0, 1, \ldots, 2d-1, \tag{1.5}$$

which is called a confluent Prony system of order d with the multiplicity vector $D = (d_1, \ldots, d_s)$. The original Prony system (1.3) is a special case of the confluent one, with D being the vector $(1, \ldots, 1)$ of length d.

The system (1.5) arises also in the problem of reconstructing a planar polygon P (or even an arbitrary semi-analytic *quadrature domain*) from its moments

$$m_k(\chi_P) = \iint_{\mathbb{R}^2} z^k \chi_P \,\mathrm{d}\, x \,\mathrm{d}\, y, \; z = x + \imath y,$$

where χ_P is the characteristic function of the domain $P \subset \mathbb{R}^2$. This problem is important in many areas of science and engineering [11]. The above yields the confluent Prony system

$$m_k = \sum_{j=1}^{s} \sum_{i=0}^{d_j-1} c_{i,j} k(k-1) \cdots (k-i+1) z_j^{k-i}, \qquad c_{i,j} \in \mathbb{C}, \; z_j \in \mathbb{C} \setminus \{0\}.$$

Definition 1.2. For a given multiplicity vector $D = (d_1, \ldots, d_s)$, its *order* is $\sum_{j=1}^{s} d_j$.

As we shall see below, if we start with the measurements $\mu(F) = \mu = (m_0, \ldots, m_{2d-1})$, then a natural setting of the problem of solving the Prony system is the following:

Problem 1.3 (Prony problem of order d)**.** *Given the measurements*

$$\mu = (m_0, \ldots, m_{2d-1}) \in \mathbb{C}^{2d}$$

in the right hand side of (1.5)*, find the multiplicity vector* $D = (d_1, \dots, d_s)$ *of order*

$$r = \sum_{j=1}^{s} d_j \leq d,$$

and find the unknowns x_j *and* $a_{j,\ell}$*, which solve the corresponding confluent Prony system* (1.5) *with the multiplicity vector* D *(hence, with solution of rank* r*).*

It is extremely important in practice to have *a stable method of inversion*. Many research efforts are devoted to this task (see e.g. [3, 7, 10, 17, 18, 20] and references therein). A basic question here is the following.

Problem 1.4 (Noisy Prony problem)**.** Given the *noisy* measurements

$$\tilde{\mu} = (\tilde{m_0}, \dots, \tilde{m}_{2d-1}) \in \mathbb{C}^{2d}$$

and an estimate of the error $|\tilde{m}_k - m_k| \leq \varepsilon_k$*, solve Problem 1.3 so as to minimize the reconstruction error.*

In this paper we study the global setting of the Prony problem, stressing its algebraic structure. In Section 2 the space where the solution is to be found (Prony space) is described. It turns out to be a vector bundle over the space of the nodes $x_1, \dots, x_d$. We define also three mappings: "Prony", "Taylor", and "Stieltjes" ones, which capture the essential features of the Prony problem and of its solution process.

In Section 3 we investigate solvability conditions for the Prony problem. The answer leads naturally to a stratification of the space of the right-hand sides, according to the rank of the associated Hankel-type matrix and its minors. The behavior of the solutions near various strata turns out to be highly nontrivial, and we present some initial results in the description of the corresponding singularities.

In Section 4, we study the multiplicity-restricted Prony problem, fixing the collision pattern of the solution, and derive simple bounds for the stability of the solution via factorization of the Jacobian determinant of the corresponding Prony map.

In Section 5 we consider the rank-restricted Prony problem, effectively reducing the dimension to $2r$ instead of $2d$, where r is precisely the rank of the associated Hankel-type matrix. In this formulation, the Prony problem is solvable in a small neighborhood of the exact measurement vector.

In Section 6 we study one of the most important singularities in the Prony problem: collision of some of the points x_i. The investigation of this type of singularities has been started in [21] where the role of finite differences was demonstrated. In the present paper we introduce global bases of finite differences, study their properties, and prove that using such bases we can resolve in a robust way at least the linear part of the Prony problem at and near colliding configurations of the nodes.

In Section 7 we discuss close connections of the Prony problem with hyperbolic polynomials and "Vandermonde mapping" studied by V.I.Arnold in [1] and by V.P.Kostov in [13, 14, 15], and with "Vieta mapping" expressing the coefficients of a polynomial through its roots. We believe that questions arising in theoretical study of Prony problem and in its practical applications justify further investigation of these connections, as well as further applications of Singularity Theory.

Finally, in Appendix A we describe a solution method for the Prony system based on Padé approximation.

2. Prony, Stieltjes and Taylor Mappings

In this section we define "Prony", "Taylor", and "Stieltjes" mappings, which capture some essential features of the Prony problem and of its solution process. The main idea behind the spaces and mappings introduced in this section is the following: associate to the signal $F(x)=\sum_{i=1}^{d}a_i\delta(x-x_i)$ the rational function $R(z)=\sum_{i=1}^{d}\frac{a_i}{z-x_i}$. (In fact, R is the Stieltjes integral transform of F). The functions R obtained in this way can be written as $R(z)=\frac{P(z)}{Q(z)}$ with $\deg P\leq\deg Q-1$, and they satisfy $R(\infty)=0$. Write R as

$$R(z)=\sum_{i=1}^{d}\frac{a_i}{z(1-x_i/z)}.$$

Developing the summands into geometric progressions we conclude that $R(z)=\sum_{k=0}^{\infty}m_k(\frac{1}{z})^{k+1}$, with

$$m_k=\sum_{i=1}^{d}a_ix_i^k,$$

so the moment measurements m_k in the right hand side of the Prony system (1.3) are exactly the Taylor coefficients of $R(z)$. We shall see below that this correspondence reduces solution of the Prony system to an appropriate Padé approximation problem.

Definition 2.1. For each $w=(x_1,\dots,x_d)\in\mathbb{C}^d$, let $s=s(w)$ be the number of distinct coordinates τ_j, $j=1,\dots,s$, and denote $T(w)=(\tau_1,\dots,\tau_s)$. The multiplicity vector is

$$D=D(w)=(d_1,\dots,d_s),$$

where d_j is the number of times the value τ_j appears in $\{x_1,\dots,x_d\}$. The order of the values in $T(w)$ is defined by their order of appearance in w.

Example 2.2. For $w=(3,1,2,1,0,3,2)$, we have

$$s(w)=4,\qquad T(w)=(3,1,2,0)\,,\text{and}\qquad D(w)=(2,2,2,1)\,.$$

Remark 2.3. Note the slight abuse of notations between Definition 1.1 and Definition 2.1. Note also that the *order* of $D(w)$ equals to d ***for all*** $w\in\mathbb{C}^d$.

Definition 2.4. For each $w\in\mathbb{C}^d$, let $s=s(w)$, $T(w)=(\tau_1,\dots,\tau_s)$ and $D(w)=(d_1,\dots,d_s)$ be as in Definition 2.1.

(1) V_w is the vector space of dimension d containing the linear combinations

$$g=\sum_{j=1}^{s}\sum_{\ell=0}^{d_j-1}\gamma_{j,\ell}\delta^{(\ell)}(x-\tau_j)\tag{2.1}$$

of δ-functions and their derivatives at the points of $T(w)$. The "standard basis" of V_w is given by the distributions

$$\delta_{j,\ell}=\delta^{(\ell)}(x-\tau_j)\,,\qquad j=1,\dots,s(w)\,;\ \ell=0,\dots,d_j-1.\tag{2.2}$$

(2) W_w is the vector space of dimension d of all the rational functions with poles $T(w)$ and multiplicities $D(w)$, vanishing at ∞ :

$$R(z)=\frac{P(z)}{Q(z)},\qquad Q(z)=\prod_{j=1}^{s}(z-\tau_j)^{d_j}\,,\ \deg P(z)<\deg Q\leqslant d.$$

The "standard basis" of W_w is given by the elementary fractions

$$R_{j,\ell} = \frac{1}{(z-\tau_j)^{\ell}}, \qquad j=1,\dots,s;\ \ell = 1,\dots,d_j.$$

Now we are ready to formally define the Prony space $\mathcal{P}_d$ and the Stieltjes space $\mathcal{S}_d$ as certain (trivial) vector bundles over $\mathbb{C}^d$.

Definition 2.5. The Prony space $\mathcal{P}_d$ is the vector bundle over $\mathbb{C}^d$, consisting of all the pairs

$$(w,g): \quad w\in\mathbb{C}^d,\ g\in V_w.$$

The topology on $\mathcal{P}_d$ is induced by the natural embedding $\mathcal{P}_d \subset \mathbb{C}^d \times \mathcal{D}$, where $\mathcal{D}$ is the space of distributions on $\mathbb{C}$ with its standard topology.

Definition 2.6. The Stieltjes space $\mathcal{S}_d$ is the vector bundle over $\mathbb{C}^d$, consisting of all the pairs

$$(w,\gamma): \qquad w\in\mathbb{C}^d,\ \gamma\in W_w.$$

The topology on $\mathcal{S}_d$ is induced by the natural embedding $\mathcal{S}_d \subset \mathbb{C}^d \times \mathcal{R}$, where $\mathcal{R}$ is the space of complex rational functions with its standard topology.

Definition 2.7. The Stieltjes mapping $\mathcal{S}M : \mathcal{P}_d \to \mathcal{S}_d$ is defined by the Stieltjes integral transform: for $(w,g)\in\mathcal{P}_d$

$$\mathcal{S}M\left((w,g)\right) = (w,\gamma), \qquad \gamma(z) = \int_{-\infty}^{\infty} \frac{g(x)\,\mathrm{d}\,x}{z-x}.$$

Sometimes we abuse notation and write for short $\mathcal{S}M(g)=\gamma$, with the understanding that $\mathcal{S}M$ is also a map $\mathcal{S}M : V_w \to W_w$ for each $w\in\mathbb{C}^d$.

The following fact is immediate consequence of the above definitions.

Proposition 2.8. *$\mathcal{S}M$ is a linear isomorphism of the bundles $\mathcal{P}_d$ and $\mathcal{S}_d$ (for each $w\in\mathbb{C}^d$, $\mathcal{S}M$ is a linear isomorphism of the vector spaces V_w and W_w). In the standard bases of V_w and W_w, the map $\mathcal{S}M$ is diagonal, satisfying*

$$\mathcal{S}M\left(\delta_{j,\ell}\right) = (-1)^{\ell}\,\ell! R_{j,\ell}(z).$$

Furthermore, for any $(w,g)\in\mathcal{P}_d$

$$\mathcal{S}M(g) = \underbrace{\frac{P(z)}{Q(z)}}_{irreducible}, \qquad \deg P < \deg Q = \operatorname{rank}(g) \leqslant d. \tag{2.3}$$

Definition 2.9. The Taylor space $\mathcal{T}_d$ is the space of complex Taylor polynomials at infinity of degree $2d-1$ of the form $\sum_{k=0}^{2d-1} m_k (\frac{1}{z})^{k+1}$. We shall identify $\mathcal{T}_d$ with the complex space $\mathbb{C}^{2d}$ with the coordinates $m_0,\dots,m_{2d-1}$.

Definition 2.10. The Taylor mapping $\mathcal{T}M : \mathcal{S}_d \to \mathcal{T}_d$ is defined by the truncated Taylor development at infinity:

$$\mathcal{T}M\left((w,\gamma)\right) = \sum_{k=0}^{2d-1} \alpha_k \left(\frac{1}{z}\right)^{k+1}, \qquad \text{where } \gamma(z) = \sum_{k=0}^{\infty} \alpha_k \left(\frac{1}{z}\right)^{k+1}.$$

We identify $\mathcal{T}M((w,\gamma))$ as above with $(\alpha_0,\dots,\alpha_{2d-1})\in\mathbb{C}^{2d}$. Sometimes we write for short $\mathcal{T}M(\gamma)=(\alpha_0,\dots,\alpha_{2d-1})$.

Finally, we define the Prony mapping $\mathcal{P}M$ which encodes the Prony problem.

Definition 2.11. The Prony mapping $\mathcal{P}M:\mathcal{P}_d\to\mathbb{C}^{2d}$ for $(w,g)\in\mathcal{P}_d$ is defined as follows:

$$\mathcal{PM}((w,g))=(m_0,\dots,m_{2d-1})\in\mathbb{C}^{2d},\qquad m_k=m_k(g)=\int x^k g(x)\,\mathrm{d}\,x.$$

By the above definitions, we have

$$\mathcal{P}M=\mathcal{T}M\circ\mathcal{S}M. \tag{2.4}$$

Solving the Prony problem for a given right-hand side $(m_0,\dots,m_{2d-1})$ is therefore equivalent to inverting the Prony mapping $\mathcal{P}M$. As we shall elaborate in the subsequent section, the identity (2.4) allows us to split this problem into two parts: inversion of $\mathcal{T}M$, which is, essentially, the Padé approximation problem, and inversion of $\mathcal{S}M$, which is, essentially, the decomposition of a given rational function into the sum of elementary fractions.

3. Solvability of the Prony problem

3.1. **General condition for solvability.** In this section we provde a necessary and sufficient condition for the Prony problem to have a solution (which is unique, as it turns out by Proposition 3.2). As mentioned in the end of the previous section, our method is based on inverting (2.4) and thus relies on the solution of the corresponding (diagonal) *Padé approximation problem* [4].

Problem 3.1 (Diagonal Padé approximation problem)**.** *Given $\mu=(m_0,\dots,m_{2d-1})\in\mathbb{C}^{2d}$, find a rational function $R_d(z)=\frac{P(z)}{Q(z)}\in\mathcal{S}_d$ with $\deg P<\deg Q\leqslant d$, such that the first $2d$ Taylor coefficients at infinity of $R_d(z)$ are $\{m_k\}_{k=0}^{2d-1}$.*

Proposition 3.2. *If a solution to Problem 3.1 exists, it is unique.*

Proof. Writing $R(z)=\frac{P(z)}{Q(z)}$, $R_1(z)=\frac{P_1(z)}{Q_1(z)}$, with $\deg P<\deg Q\leqslant d$ and $\deg P_1<\deg Q_1\leqslant d$, we get

$$R-R_1=\frac{PQ_1-P_1Q}{QQ_1},$$

and this function, if nonzero, can have a zero of order at most $2d-1$ at infinity. □

Let us summarize the above discussion with the following statement.

Proposition 3.3. *The tuple*

$$\left\{s,\ D=(d_1,\dots,d_s),\ r=\sum_{j=1}^s d_j\le d,\ X=\{x_j\}_{j=1}^s,\ A=\{a_{j,\ell}\}_{j=1,\dots,s;\ \ell=0,\dots,d_j-1}\right\}$$

is a (unique, up to a permutation of the nodes $\{x_j\}$) solution to Problem 1.3 with right-hand side

$$\mu=(m_0,\dots,m_{2d-1})\in\mathbb{C}^{2d}$$

if and only if the rational function

$$R_{D,X,A}(z)=\sum_{j=1}^s\sum_{\ell=1}^{d_j}(-1)^{\ell-1}(\ell-1)!\frac{a_{j,\ell-1}}{(z-x_j)^\ell}=\sum_{k=0}^{2d-1}\frac{m_k}{z^{k+1}}+O\left(z^{-2d-1}\right)$$

is a (unique) solution to Problem 3.1 with input μ. In that case,

$$R_{D,X,A}(z) = \int_{-\infty}^{\infty} \frac{g(x)\,\mathrm{d}\,x}{z-x} \qquad \text{where } g(x) = \sum_{j=1}^{s} \sum_{\ell=0}^{d_j-1} a_{j,\ell}\delta^{(\ell)}(x-x_j),$$

i.e., $R_{D,X,A}(z)$ is the Stieltjes transform of $g(x)$.

Proof. This follows from the definitions of Section 2, (2.4), Proposition 3.2 and the fact that the problem of representing a given rational function as a sum of elementary fractions of the specified form (i.e., inverting $\mathcal{S}M$) is always uniquely solvable up to a permutation of the poles. □

The next result provides necessary and sufficient conditions for the solvability of Problem 3.1. It summarizes some well-known facts in the theory of Padé approximation, related to "normal indices" (see, for instance, [4]). However, these facts are not usually formulated in the literature on Padé approximation in the form we need in relation to the Prony problem. Consequently, we give a detailed proof of this result in Appendix A. This proof contains, in particular, some facts which are important for understanding the solvability issues of the Prony problem.

Definition 3.4. Given a vector $\mu = (m_0, \dots, m_{2d-1})$, let $\tilde{M}_d$ denote the $d \times (d+1)$ Hankel matrix

$$\tilde{M}_d = \begin{bmatrix} m_0 & m_1 & m_2 & \dots & m_d \\ m_1 & m_2 & m_3 & \dots & m_{d+1} \\ \iddots & \iddots & \iddots & \iddots & \iddots \\ m_{d-1} & m_d & m_{d+1} & \dots & m_{2d-1} \end{bmatrix}. \tag{3.1}$$

For each $e \leqslant d$, denote by $\tilde{M}_e$ the $e \times (e+1)$ submatrix of $\tilde{M}_d$ formed by the first e rows and $e+1$ columns, and let M_e denote the corresponding square matrix.

Theorem 3.5. *Let $\mu = (m_0, \dots, m_{2d-1})$ be given, and let $r \leqslant d$ be the rank of the Hankel matrix $\tilde{M}_d$ as in (3.1). Then Problem 3.1 is solvable for the input μ if and only if the upper left minor $|M_r|$ of $\tilde{M}_d$ is non-zero.*

As an immediate consequence of Theorem 3.5 and Proposition 3.3, we obtain the following result.

Theorem 3.6. *Let $\mu = (m_0, \dots, m_{2d-1})$ be given, and let $r \leqslant d$ be the rank of the Hankel matrix $\tilde{M}_d$ as in (3.1). Then Problem 1.3 with input μ is solvable if and only if the upper left minor $|M_r|$ of $\tilde{M}_d$ is non-zero. The solution, if it exists, is unique, up to a permutation of the nodes $\{x_j\}$. The multiplicity vector $D = (d_1, \dots, d_s)$, of order $\sum_{j=1}^{s} d_j = r$, of the resulting confluent Prony system of rank r is the multiplicity vector of the poles of the rational function $R_{D,X,A}(z)$, solving the corresponding Padé problem.*

As a corollary we get a complete description of the right-hand side data $\mu \in \mathbb{C}^{2d}$ for which the Prony problem is solvable (unsolvable). Define for $r = 1, \dots, d$ sets $\Sigma_r \subset \mathbb{C}^{2d}$ (respectively, $\Sigma'_r \subset \mathbb{C}^{2d}$) consisting of $\mu \in \mathbb{C}^{2d}$ for which the rank of $\tilde{M}_d = r$ and $|M_r| \neq 0$ (respectively, $|M_r| = 0$). The set Σ_r is a difference $\Sigma_r = \Sigma_r^1 \setminus \Sigma_r^2$ of two algebraic sets: Σ_r^1 is defined by vanishing of all the $s \times s$ minors of $\tilde{M}_d$, $r < s \leq d$, while Σ_r^2 is defined by vanishing of $|M_r|$. In turn, $\Sigma'_r = \Sigma_r'^1 \setminus \Sigma_r'^2$, with $\Sigma_r'^1 = \Sigma_r^1 \cap \Sigma_r^2$ and $\Sigma_r'^2$ defined by vanishing of all the $r \times r$ minors of $\tilde{M}_d$. The union $\Sigma_r \cup \Sigma'_r$ consists of all μ for which the rank of $\tilde{M}_d = r$, which is $\Sigma_r^1 \setminus \Sigma_r'^2$.

Corollary 3.7. *The set Σ (respectively, Σ') of $\mu \in \mathbb{C}^{2d}$ for which the Prony problem is solvable (respectively, unsolvable) is the union $\Sigma = \cup_{r=1}^{d} \Sigma_r$ (respectively, $\Sigma' = \cup_{r=1}^{d} \Sigma'_r$). In particular, $\Sigma' \subset \{\mu \in \mathbb{C}^{2d}, \det M_d = 0\}$.*

So for a generic right hand side μ we have $|M_d| \neq 0$, and the Prony problem is solvable. On the algebraic hypersurface of μ for which $|M_d| = 0$, the Prony problem is solvable if $|M_{d-1}| \neq 0$, etc.

Let us now consider some examples.

Example 3.8. Let us fix $d = 1, 2, \ldots$. Consider $\mu = (m_0, \ldots, m_{2d-1}) \in \mathbb{C}^{2d}$, the right hand sides of the Prony problem, to be of the form $\mu = \mu_\ell = (\delta_{k\ell}) = (0, \ldots, 0, \underbrace{1}_{\text{position } \ell+1}, 0, \ldots, 0)$, with all the $m_k = 0$ besides $m_\ell = 1,\ \ell = 0, \ldots, 2d-1$, and let $\tilde{M}_d^\ell$ be the corresponding matrix.

Proposition 3.9. *The rank of $\tilde{M}_d^\ell$ is equal to $\ell + 1$ for $\ell \leq d - 1$, and it is equal to $2d - \ell$ for $\ell \geq d$. The corresponding Prony problem is solvable for $\ell \leq d-1$, and it is unsolvable for $\ell \geq d$.*

Proof. For $d = 5$ and $\ell = 2, 4, 5, 9$, the corresponding matrices $\tilde{M}_\ell^d$ are as follows.

$$\tilde{M}_5^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \ \tilde{M}_5^4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \text{(solvable)}$$

$$\tilde{M}_5^5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \ \tilde{M}_5^9 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \qquad \text{(unsolvable)}$$

In general, the matrices $\tilde{M}_d^\ell$ have the same pattern as in the special cases above, so their rank is $\ell+1$ for $\ell \leqslant d-1$, and $2d - \ell$ for $\ell \geqslant d$, as stated above. Application of Theorem 3.6 completes the proof. ☐

In fact, μ_ℓ is a moment sequence of

$$F(x) = \frac{1}{\ell!} \delta^{(\ell)}(x),$$

and this signal belongs to $\mathcal{P}_d$ if and only if $\ell \leqslant d-1$. In notations of Corollary 3.7 we have

$$\begin{aligned} \mu_\ell \in \quad & \Sigma_{\ell+1}, \quad \ell \leqslant d-1, \\ \mu_\ell \in \quad & \Sigma'_{2d-\ell}, \qquad \ell \geqslant d. \end{aligned}$$

It is easy to provide various modifications of the above example. In particular, for

$$\mu = \tilde{\mu}_\ell = (0, \ldots, 0, 1, 1, \ldots, 1),$$

the result of Proposition 3.9 remains verbally true.

Example 3.10. Another example is provided by μ_{ℓ_1,ℓ_2}, with all the $m_k = 0$ besides

$$m_{\ell_1} = 1,\ m_{\ell_2} = 1,\ 0 \leq \ell_1 < d \leq \ell_2 \leq 2d-1.$$

For $\ell_1 < \ell_2 - d + 1$ the rank of the corresponding matrix $\tilde{M}_d$ is $r = 2d + \ell_1 - \ell_2 + 1$ while $|M_r| = 0$, so the Prony problem for such μ_{ℓ_1,ℓ_2} is unsolvable. For $d = 5$ and $\ell_1 = 2,\ \ell_2 = 8$ the matrix is as follows:

$$\tilde{M}_5^{(2,8)} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

3.2. **Near-singular inversion.** The behavior of the inversion of the Prony mapping near the unsolvability stratum Σ' and near the strata where the rank of $\tilde{M}_d$ drops, turns out to be pretty complicated. In particular, in the first case at least one of the nodes tends to infinity. In the second case, depending on the way the right-hand side μ approaches the lower rank strata, the nodes may remain bounded, or some of them may tend to infinity. In this section we provide one initial result in this direction, as well as some examples. We believe that a comprehensive description of the inversion of the Prony mapping near Σ' and near the lower rank strata is important both in theoretical study and in applications of Prony-like systems, and consider it to be an important direction for future research.

Theorem 3.11. *As the right-hand side $\mu \in \mathbb{C}^{2d} \setminus \Sigma'$ approaches a finite point $\mu_0 \in \Sigma'$, at least one of the nodes $x_1, \ldots, x_d$ in the solution tends to infinity.*

Proof. By assumptions, the components $m_0, \ldots, m_{2d-1}$ of the right-hand side

$$\mu = (m_0, \ldots, m_{2d-1}) \in \mathbb{C}^{2d}$$

remain bounded as $\mu \to \mu_0$. By Theorem 6.17, the finite differences coordinates of the solution $\mathcal{P}M^{-1}(\mu)$ remain bounded as well. Now, if all the nodes are also bounded, by compactness we conclude that $\mathcal{P}M^{-1}(\mu) \to \omega \in \mathcal{P}_d$. By continuity in the distribution space (Lemma 6.9) we have $\mathcal{P}M(\omega) = \mu_0$. Hence the Prony problem with the right-hand side μ_0 has a solution $\omega \in \mathcal{P}_d$, in contradiction with the assumption that $\mu_0 \in \Sigma'$. □

Example 3.12. Let us consider an example: $d = 2$ and $\mu_0 = (0, 0, 1, 0)$. Here the rank ℓ of $\tilde{M}_2$ is 2, and $|M_2| = 0$, so by Theorem 3.6 we have $\mu_0 \in \Sigma'_2 \subset \Sigma'$. Consider now a perturbation $\mu(\epsilon) = (0, \epsilon, 1, 0)$ of μ_0. For $\epsilon \neq 0$ we have $\mu(\epsilon) \in \Sigma_2 \subset \Sigma$, and the Prony system is solvable for μ_ϵ. Let us write an explicit solution: the coefficients c_0, c_1 of the polynomial $Q(z) = c_0 + c_1 z + z^2$ we find from the system (A.⋆⋆):

$$\begin{bmatrix} 0 & \epsilon \\ \epsilon & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

whose solution is $c_1 = -\frac{1}{\epsilon}$, $c_0 = \frac{1}{\epsilon^2}$. Hence the denominator $Q(z)$ of $R(z)$ is $Q(z) = \frac{1}{\epsilon^2} - \frac{1}{\epsilon} z + z^2$, and its roots are $x_1 = \frac{1+\imath\sqrt{3}}{2\epsilon}$, $x_2 = \frac{1-\imath\sqrt{3}}{2\epsilon}$. The coefficients b_0, b_1 of the numerator $P(z) = b_0 + b_1 z$ we find from (A.⋆):

$$\begin{bmatrix} 0 & 0 \\ 0 & \epsilon \end{bmatrix} \begin{bmatrix} -\frac{1}{\epsilon} \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_0 \end{bmatrix},$$

i.e., $b_1 = 0$, $b_0 = \epsilon$. Thus the solution of the associated Padé problem is

$$R(z) = \frac{P(z)}{Q(z)} = \frac{\epsilon}{(z - x_1)(z - x_2)} = \frac{\epsilon^2}{\imath\sqrt{3}} \frac{1}{(z - x_1)} - \frac{\epsilon^2}{\imath\sqrt{3}} \frac{1}{(z - x_2)}.$$

Finally, the (unique up to a permutation) solution of the Prony problem for μ_ϵ is

$$a_1 = \frac{\epsilon^2}{\imath\sqrt{3}},\ a_2 = -\frac{\epsilon^2}{\imath\sqrt{3}}, \quad x_1 = \frac{1+\imath\sqrt{3}}{2\epsilon},\ x_2 = \frac{1-\imath\sqrt{3}}{2\epsilon}.$$

As ϵ tends to zero, the nodes x_1, x_2 tend to infinity while the coefficients a_1, a_2 tend to zero.

As it was shown above, for a given $\mu \in \Sigma$ (say, with pairwise different nodes) the rank of the matrix $\tilde{M}_d$ is equal to the number of the nodes in the solution for which the corresponding δ-function enters with a non-zero coefficients. So μ approaches a certain μ_0 belonging to a stratum of a lower rank of $\tilde{M}_d$ if and only if some of the coefficients a_j in the solution tend to zero. We do not analyze all the possible scenarios of such a degeneration, noticing just that if $\mu_0 \in \Sigma'$, i.e., the Prony problem is unsolvable for μ_0, then Theorem 3.11 remains true, with essentially the same proof. So at least one of the nodes, say, x_j, escapes to infinity. Moreover, one can show that $a_j x_j^{2d-1}$ cannot tend to zero - otherwise the remaining linear combination of δ-functions would provide a solution for μ_0.

If $\mu_0 \in \Sigma$, i.e., the Prony problem is solvable for μ_0, all the nodes may remain bounded, or some x_j may escape to infinity, but in such a way that $a_j x_j^{2d-1}$ tends to zero.

4. Multiplicity-restricted Prony problem

Consider Problem 1.4 at some point $\mu_0 \in \Sigma$. By definition, $\mu_0 \in \Sigma_{r_0}$ for some $r_0 \le d$. Let $\mu_0 = \mathcal{P}M\left((w_0, g_0)\right)$ for some $(w_0, g_0) \in \mathcal{P}_d$. Assume for a moment that the multiplicity vector $D_0 = D(g_0) = (d_1, \ldots d_{s_0})$, $\sum_{j=1}^{s_0} d_j = r_0$, has a non-trivial collision pattern, i.e., $d_j > 1$ for at least one $j = 1, \ldots, s_0$. It means, in turn, that the function $R_{D_0,X,A}(z)$ has a pole of multiplicity d_j. Evidently, there exists an arbitrarily small perturbation $\tilde{\mu}$ of μ_0 for which this multiple pole becomes a cluster of single poles, thereby changing the multiplicity vector to some $D' \neq D_0$. While we address this problem in Section 6 via the bases of divided differences, in this section we consider a "multiplicity-restricted" Prony problem.

Definition 4.1. Let $\mathbf{x} = (x_1, \ldots, x_s) \in \mathbb{C}^s$ and $D = (d_1, \ldots, d_s)$ with $d = \sum_{j=1}^{s} d_j$ be given. The $d \times d$ *confluent Vandermonde* matrix is

$$V = V(\mathbf{x}, D) = V(x_1, d_1, \ldots, x_s, d_s) = \begin{bmatrix} \mathbf{v_{1,0}} & \mathbf{v_{2,0}} & \cdots & \mathbf{v_{s,0}} \\ \mathbf{v_{1,1}} & \mathbf{v_{2,1}} & \cdots & \mathbf{v_{s,1}} \\ & & \cdots & \\ \mathbf{v_{1,d-1}} & \mathbf{v_{2,d-1}} & \cdots & \mathbf{v_{s,d-1}} \end{bmatrix} \tag{4.1}$$

where the symbol $\mathbf{v_{j,k}}$ denotes the following $1 \times d_j$ row vector

$$\mathbf{v_{j,k}} \stackrel{\text{def}}{=} \begin{bmatrix} x_j^k, & k x_j^{k-1}, & \ldots & , k(k-1)\cdots(k-d_j)\, x_j^{k-d_j+1} \end{bmatrix}.$$

Proposition 4.2. *The matrix V defines the linear part of the confluent Prony system* (1.5) *in the standard basis for V_w, namely,*

$$V(x_1, d_1, \ldots, x_s, d_s) \begin{bmatrix} a_{1,0} \\ \vdots \\ a_{1,d_1-1} \\ \vdots \\ a_{s,d_s-1} \end{bmatrix} = \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ \\ m_{d-1} \end{bmatrix}. \tag{4.2}$$

Definition 4.3. Let $\mathcal{P}M(w_0, g_0) = \mu_0 \in \Sigma_{r_0}$ with $D(g_0) = D_0$ and $s(g_0) = s_0$. Let $\mathcal{P}_{D_0}$ denote the following subbundle of $\mathcal{P}_d$ of dimension $s_0 + r_0$:

$$\mathcal{P}_{D_0} = \left\{ (w, g) \in \mathcal{P}_d : \quad D(g) = D_0 \right\}.$$

The multiplicity-restricted Prony mapping $\mathcal{PM}_{D_0}^* : \mathcal{P}_{D_0} \to \mathbb{C}^{s_0+r_0}$ is the composition

$$\mathcal{PM}_{D_0}^* = \pi \circ \mathcal{P}M \restriction_{\mathcal{P}_{D_0}},$$

where $\pi : \mathbb{C}^{2d} \to \mathbb{C}^{s_0+r_0}$ is the projection map on the first $s_0 + r_0$ coordinates.

Inverting this $\mathcal{PM}^*_{D_0}$ represents the solution of the confluent Prony system (1.5) with fixed structure D_0 from the first $k = 0, 1, \ldots, s_0 + r_0 - 1$ measurements.

Theorem 4.4 ([7]). *Let $\mu_0^* = \mathcal{PM}^*_{D_0}((w_0, g_0)) \in \mathbb{C}^{s_0+r_0}$ with the unperturbed solution*

$$g_0 = \sum_{j=1}^{s_0} \sum_{\ell=0}^{d_j-1} a_{j,\ell} \delta^{(\ell)}(x - \tau_j).$$

In a small neighborhood of $(w_0, g_0) \in \mathcal{P}_{D_0}$, the map $\mathcal{PM}^*_{D_0}$ is invertible. Consequently, for small enough ε, the multiplicity-restricted Prony problem with input data $\tilde{\mu}^* \in \mathbb{C}^{r_0+s_0}$ satisfying $\|\tilde{\mu}^* - \mu_0^*\| \leq \varepsilon$ has a unique solution. The error in this solution satisfies

$$\begin{aligned}
|\Delta a_{j,\ell}| &\leq \frac{2}{\ell!}\left(\frac{2}{\delta}\right)^{s_0+r_0}\left(\frac{1}{2} + \frac{s_0+r_0}{\delta}\right)^{d_j-\ell}\left(1 + \frac{|a_{j,\ell-1}|}{|a_{j,d_j-1}|}\right)\varepsilon, \\
|\Delta \tau_j| &\leq \frac{2}{d_j!}\left(\frac{2}{\delta}\right)^{s_0+r_0} \frac{1}{|a_{j,d_j-1}|}\varepsilon,
\end{aligned}$$

where $\delta \stackrel{def}{=} \min_{i \neq j} |\tau_i - \tau_j|$ (for consistency we take $a_{j,-1} = 0$ in the above formula).

Proof outline. The Jacobian of $\mathcal{PM}^*_{D_0}$ can be easily computed, and it turns out to be equal to the product

$$\mathcal{J}_{\mathcal{PM}^*_{D_0}} = V(\tau_1, d_1 + 1, \ldots, \tau_{s_0}, d_{s_0} + 1) \operatorname{diag}\{E_j\}$$

where V is the *confluent Vandermonde matrix* (4.1) on the nodes $(\tau_1, \ldots, \tau_{s_0})$, with multiplicity vector

$$\tilde{D}_0 = (d_1 + 1, \ldots, d_{s_0} + 1),$$

while E is the $(d_j + 1) \times (d_j + 1)$ block

$$E_j = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & a_{j,0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{j,d_j-1} \end{bmatrix}.$$

Since $\mu_0 \in \Sigma_r$, the highest order coefficients a_{j,d_j-1} are nonzero. Furthermore, since all the τ_j are distinct, the matrix V is nonsingular. Local invertibility follows. To estimate the norm of the inverse, use bounds from [6]. □

Remark 4.5. Note that as two nodes collide ($\delta \to 0$), the inversion of the multiplicity-restricted Prony mapping $\mathcal{PM}^*_{D_0}$ becomes ill-conditioned proportionally to $\delta^{-(s_0+r_0)}$.

Let us stress that we are not aware of any general method of inverting $\mathcal{PM}^*_{D_0}$, i.e., solving the multiplicity-restricted confluent Prony problem with the smallest possible number of measurements. As we demonstrate in [5], such a method exists for a very special case of a single point, i.e., $s = 1$.

5. Rank-restricted Prony problem

Recall that the Prony problem consists in inverting the Prony mapping $\mathcal{PM} : \mathcal{P}_d \to \mathcal{T}_d$. So, given $\mu = (m_0, \ldots, m_{2d-1}) \in \mathcal{T}_d$ we are looking for $(w, g) \in \mathcal{P}_d$ such that

$$m_k(g) = \int x^k g(x) dx = m_k,$$

with $k = 0, 1, \ldots, 2d - 1$. If $\mu \in \Sigma_r$ with $r < d$, then in fact any neighborhood of μ will contain points from the non-solvability set Σ'. Indeed, consider the following example.

Example 5.1. Slightly modifying the construction of Example 3.10, consider $\mu_{\ell_1,\ell_2,\epsilon} \in \mathbb{C}^{2d}$ with all the $m_k = 0$ besides $m_{\ell_1} = 1$ and $m_{\ell_2} = \epsilon$, such that $\ell_2 > \ell_1 + d - 1$. For example, if $d = 5$ and $\ell_1 = 2,\ \ell_2 = 8$, the corresponding matrix is

$$\tilde{M}_5^{(2,8,\epsilon)} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \epsilon \\ 0 & 0 & 0 & 0 & \epsilon & 0 \end{bmatrix}.$$

For $\epsilon = 0$ the Prony problem is solvable, while for any small perturbation $\epsilon \neq 0$ it becomes unsolvable. However, if we restrict the whole problem just to $d = 3$, it remains solvable for any small perturbation of the input.

We therefore propose to consider the *rank-restricted Prony problem* analogous to the construction of Section 4, but instead of fixing the multiplicity $D(g)$ we now fix the rank r (recall Definition 1.1).

Definition 5.2. Denote by $\mathcal{P}_r$ the following vector bundle:

$$\mathcal{P}_r = \left\{(w, g) : \quad w \in \mathbb{C}^r,\ g \in V_w\right\},$$

where V_w is defined exactly as in Definition 2.4, replacing d with r.

Likewise, we define the Stieltjes bundle of order r as follows.

Definition 5.3. Denote by $\mathcal{S}_r$ the following vector bundle:

$$\mathcal{S}_r = \left\{(w, \gamma) : \qquad w \in \mathbb{C}^r,\ \gamma \in W_w\right\},$$

where W_w is defined exactly as in Definition 2.4, replacing d with r.

The Stieltjes mapping acts naturally as a map $\mathcal{S}M : \mathcal{P}_r \to \mathcal{S}_r$ with exactly the same definition as Definition 2.7.

The restricted Taylor mapping $\mathcal{T}M_r : \mathcal{S}_r \to \mathbb{C}^{2r}$ is, as before, given by the truncated development at infinity to the first $2r$ Taylor coefficients.

Definition 5.4. Let $\pi : \mathbb{C}^{2d} \to \mathbb{C}^{2r}$ denote the projection operator onto the first $2r$ coordinates. Denote $\Sigma_r^* \stackrel{\text{def}}{=} \pi(\Sigma_r)$. The rank-restricted Prony mapping $\mathcal{PM}_r^* : \mathcal{P}_r \to \Sigma_r^*$ is given by by

$$\mathcal{PM}_r^*((w, g)) = (m_0, \dots, m_{2r-1}), \qquad m_k = m_k(g) = \int x^k g(x)\,\mathrm{d}\,x.$$

Remark 5.5. $\mathcal{P}_r$ can be embedded in $\mathcal{P}_d$, for example by the map $\Xi_r : \mathcal{P}_r \to \mathcal{P}_d$

$$\Xi_r :\ (w, g) \in \mathcal{P}_r \longmapsto (w', g') \in \mathcal{P}_d : \qquad w' = \left(x_1, \dots, x_r, \underbrace{0, \dots 0}_{\times(d-r)} \right),\ g' = g.$$

With this definition, $\mathcal{PM}_r^*$ can be represented also as the composition

$$\mathcal{PM}_r^* = \pi \circ \mathcal{P}M \circ \Xi_r.$$

Proposition 5.6. *The rank-restricted Prony mapping satisfies*

$$\mathcal{PM}_r^* = \mathcal{T}M_r \circ \mathcal{S}M.$$

Inverting $\mathcal{PM}_r^*$ represents the solution of the rank-restricted Prony problem. Unlike in the multiplicity-restricted setting of Section 4, here we allow two or more nodes to collide (thereby changing the multiplicty vector $D(g)$ of the solution).

The basic fact which makes this formulation useful is the following result.

Theorem 5.7. *Let $\mu_0^* \in \Sigma_r^*$. Then in a small neighborhood of $\mu_0^* \in \mathbb{C}^{2r}$, the Taylor mapping $\mathcal{T}M_r$ is continuously invertible.*

Proof. This is a direct consequence of the solution method to the Padé approximation problem described in Appendix A. Indeed, if the rank of $\tilde{M}_r$ is full, then it remains so in a small neighborhood of the *entire space* $\mathbb{C}^{2r}$. Therefore, the system (A.⋆⋆) remains continuously invertible, producing the coefficients of the denominator $Q(z)$. Consequently, the right-hand side of (A.⋆) depends continuously on the moment vector $\mu^* = (m_0, \ldots, m_{2r-1}) \in \mathbb{C}^{2r}$. Again, since the rank always remains full, the polynomials $P(z)$ and $Q(z)$ cannot have common roots, and thereby the solution $R = \frac{P}{Q} = \mathcal{T}M_r^{-1}(\mu^*)$ depends continuously on μ^* (in the topology of the space of rational functions). □

In the next section, we consider the remaining problem: how to invert $\mathcal{S}M$ in this setting.

6. Collision singularities and bases of finite differences

6.1. **Introduction.** Collision singularities occur in Prony systems as some of the nodes x_i in the signal $F(x) = \sum_{i=1}^d a_i \delta(x - x_i)$ approach one another. This happens for μ near the discriminant stratum $\Delta \subset \mathbb{C}^{2d}$ consisting of those $(m_0, \ldots, m_{2d-1})$ for which some of the coordinates $\{x_j\}$ in the solution collide, i.e., the function $R_{D,X,A}(z)$ has multiple poles (or, nontrivial multiplicity vector D). As we shall see below, typically, as μ approaches $\mu_0 \in \Delta$, i.e. some of the nodes x_i collide, the corresponding coefficients a_i tend to infinity. Notice, that all the moments $m_k = m_k(F)$ remain bounded. This behavior creates serious difficulties in solving "near-colliding" Prony systems, both in theoretical and practical settings. Especially demanding problems arise in the presence of noise. The problem of improvement of resolution in reconstruction of colliding nodes from noisy measurements appears in a wide range of applications. It is usually called a "super-resolution problem" and a lot of recent publications are devoted to its investigation in various mathematical and applied settings. See [8] and references therein for a very partial sample.

Here we continue our study of collision singularities in Prony systems, started in [21]. Our approach uses bases of finite differences in the Prony space $\mathcal{P}_r$ in order to "resolve" the linear part of collision singularities. In these bases the coefficients do not blow up any more, even as some of the nodes collide.

Example 6.1. Let $r = 2$, and consider the signal $F = a_1 \delta(x - x_1) + a_2 \delta(x - x_2)$ with

$$\begin{aligned} x_1 &= t,\ x_2 = t + \epsilon, \\ a_1 &= -\epsilon^{-1},\ a_2 = \epsilon^{-1}. \end{aligned}$$

The corresponding Prony system is

$$\left(a_1 x_1^k + a_2 x_2^k =\right) m_k = k t^{k-1} + \underbrace{\sum_{j=2}^{k} \binom{k}{j} t^{k-j} \epsilon^{j-1}}_{\stackrel{\text{def}}{=} \rho_k(t,\epsilon)}, \qquad k = 0, 1, 2, 3.$$

As $\epsilon \to 0$, the Prony system as above becomes ill-conditioned and the coefficients $\{a_j\}$ blow up, while the measurements remain bounded. Note that

$$\tilde{M}_2 = \begin{bmatrix} 0 & 1 & 2t + \rho_2(t,\epsilon) \\ 1 & 2t + \rho_2(t,\epsilon) & 3t^2 + \rho_3(t,\epsilon) \end{bmatrix},$$

therefore $\operatorname{rank} \tilde{M}_2 = 2$ and $|M_2| = 1 \neq 0$, i.e. the Prony problem with input $(m_0, \ldots, m_3)$ remains solvable for all ϵ. However, the standard basis $\{\delta(x - x_1),\ \delta(x - x_2)\}$ degenerates, and in the limit it is no more a basis. If we represent the solution

$$F_{\epsilon}(x) = -\frac{1}{\epsilon}\delta(x-t) + \frac{1}{\epsilon}\delta(x-t-\epsilon)$$

in the basis

$$\begin{aligned} \Delta_1(x_1, x_2) &= \delta(x - x_1), \\ \Delta_2(x_1, x_2) &= \frac{1}{x_1 - x_2}\delta(x - x_1) + \frac{1}{x_2 - x_1}\delta(x - x_2), \end{aligned}$$

then we have

$$F_{\epsilon}(x) = 1 \cdot \Delta_2(t, t+\epsilon),$$

i.e., the coefficients in this new basis are just $\{b_1 = 0,\ b_1 = 1\}$. As $\epsilon \to 0$, in fact we have

$$\Delta_2(t, t+\epsilon) \to \delta'(x-t),$$

where the convergence is in the topology of the bundle $\mathcal{P}_r$.

Our goal in this section is to generalize the construction of Example 6.1 and [21] to handle the general case of colliding configurations.

6.2. **Divided finite differences.** For modern treatment of divided differences, see e.g. [9, 12, 16]. We follow [9] and adopt what has become by now the standard definition.

Definition 6.2. Let an arbitrary sequence of points $w = (x_1, x_2, \dots,)$ be given (repetitions are allowed). The *(n-1)*-st *divided difference* $\Delta^{n-1}(w) : \Pi \to \mathbb{C}$ is the linear functional on the space Π of polynomials in one variable x, associating to each $p \in \Pi$ its (uniquely defined) n-th coefficient in the Newton form

$$p(x) = \sum_{j=1}^{\infty} \left\{ \Delta^{j-1}(x_1, \dots, x_j)\, p \right\} \cdot q_{j-1,w}(x), \qquad q_{i,w}(x) \overset{def}{=} \prod_{k=1}^{i} (x - x_k). \tag{6.1}$$

Example 6.3. For $n = 1$, we have $\Delta^0(x_1)\, p = p(x_1)$, and the 0-th order Newton interpolation polynomial is the constant

$$P_1(x) = p(x_1) \cdot \underbrace{1}_{=q_{0,w}(x)}.$$

Example 6.4. For $n = 2$ consider two cases.

(1) If $x_1 \neq x_2$, we have $\Delta^1(x_1, x_2)\, p = \frac{p(x_2) - p(x_1)}{x_2 - x_1}$, and the first order Newton interpolation polynomial is

$$P_2(x) = p(x_1) \cdot \underbrace{1}_{=q_{o,w}(x)} + \frac{p(x_2) - p(x_1)}{x_2 - x_1} \cdot \underbrace{(x - x_1)}_{=q_{1,w}(x)}.$$

It can be readily verified that $P_2(x_k) = p(x_k)$ for $k = 1, 2$.

(2) If $x_1 = x_2$, then $\Delta^1(x_1, x_1)\, p = p'(x_1)$, and so

$$P_2(x) = p(x_1) + p'(x_1)(x - x_1).$$

It can be readily verified that $P_2(x_1) = p(x_1)$ and $P_2'(x_1) = p'(x_1)$.

It turns out that this definition can be extended to all sufficiently smooth functions for which the interpolation problem is well-defined.

Definition 6.5 ([9])**.** For any smooth enough function f, defined at least on $x_1,\dots,x_n$, the divided finite difference $\Delta^{n-1}(x_1,\dots,x_n)f$ is the n-th coefficient in the Newton form (6.1) of the Hermite interpolation polynomial P_n, which agrees with f and its derivatives of appropriate order on $x_1,\dots,x_n$:

$$f^{(\ell)}(x_j)=P_n^{(\ell)}(x_j): \qquad 1\leqslant j\leqslant n,\ 0\leqslant \ell<d_j \stackrel{\text{def}}{=} \#\{i:\quad x_i=x_j\}. \tag{6.2}$$

Example 6.6. Consider the rational function depending on a parameter $z\in\mathbb{C}$:

$$f_z(x)=\frac{1}{z-x}.$$

The 0th divided difference is $\Delta^0(x_1)f=f(x_1)=\frac{1}{z-x_1}$, and the Newton interpolation polynomial is

$$P_1(x)=\frac{1}{z-x_1}.$$

For $n=2$ and $x_1\neq x_2$, we have $\Delta^1(x_1,x_2)=\frac{1}{(z-x_1)(z-x_2)}$, and

$$P_2(x)=\frac{1}{z-x_1}+\frac{x-x_1}{(z-x_1)(z-x_2)},$$

thus $P_2(x_k)=f(x_k)$ for $k=1,2$. If $x_1=x_2$ then $\Delta^1(x_1,x_1)=f_z'(x_1)=\frac{1}{(z-x_1)^2}$, and so

$$P_2(x)=\frac{1}{z-x_1}+\frac{x-x_1}{(z-x_1)^2}.$$

Again, $P_2(x_1)=f_z(x_1)$ and $P_2'(x_1)=f_z'(x_1)$.

Therefore, each divided difference can be naturally associated with an element of the Prony space (see Item 5 in Proposition 6.7 and Definition 6.8 below for an accurate statement).

Let us now summarize relevant properties of the functional Δ which we shall use later on.

Proposition 6.7. *For $w=(x_1,\dots,x_n)\in\mathbb{C}^n$, let $s(w)$, $T(w)$ and $D(w)$ be defined according to Definition 2.1. Let $q_{n,w}(z)=\prod_{j=1}^{s}(z-\tau_j)^{d_j}$ be defined as in (6.1).*

1. *The functional $\Delta^{n-1}(x_1,\dots,x_n)$ is a symmetric function of its arguments, i.e., it depends only on the set $\{x_1,\dots,x_n\}$ but not on its ordering.*
2. *$\Delta^{n-1}(x_1,\dots,x_n)$ is a continuous function of the vector $(x_1,\dots,x_n)$. In particular, for any test function f*

$$\lim_{(x_1,\dots,x_n)\to(t_1,\dots,t_n)}\Delta^{n-1}(x_1,\dots,x_n)f=\Delta^{n-1}(t_1,\dots,t_n)f.$$

3. *Δ may be computed by the recursive rule*

$$\Delta^{n-1}(x_1,\dots,x_n)f=\begin{cases}\frac{\Delta^{n-2}(x_2,\dots,x_n)f-\Delta^{n-2}(x_1,\dots,x_{n-1})f}{x_n-x_1} & x_1\neq x_n,\\ \left\{\frac{\mathrm{d}}{\mathrm{d}\,\xi}\Delta^{n-2}(\xi,x_2,\dots,x_{n-1})f\right\}|_{\xi=x_n}, & x_1=x_n,\end{cases} \tag{6.3}$$

where $\Delta^0(x_1)f=f(x_1)$.

4. *(Generalization of Example 6.6) Let $f_z(x)=(z-x)^{-1}$. Then for all $z\notin\{x_1,\dots,x_n\}$*

$$\Delta^{n-1}(x_1,\dots,x_n)f_z=\frac{1}{q_{n,w}(z)}. \tag{6.4}$$

5. *By (6.2), $\Delta^{n-1}(x_1,\dots,x_n)$ is a linear combination of the functionals*

$$\delta^{(\ell)}(x-\tau_j),\qquad 1\leqslant j\leqslant s,\ 0\leqslant \ell<d_j.$$

In fact, using (6.4) *we obtain the Chakalov's expansion (see* [9]*)*

$$\mathbf{\Delta}^{n-1}(x_1,\dots,x_n)=\sum_{j=1}^{s}\sum_{\ell=0}^{d_j-1}a_{j,\ell}\delta^{(\ell)}(x-\tau_j), \tag{6.5}$$

where the coefficients $\{a_{j,\ell}\}$ *are defined by the partial fraction decomposition*[1]

$$\frac{1}{q_{n,w}(z)}=\sum_{j=1}^{s}\sum_{\ell=0}^{d_j-1}\frac{\ell!a_{j,\ell}}{(z-\tau_j)^{\ell+1}}. \tag{6.6}$$

(6) *By* (6.5) *and* (6.6)

$$\mathbf{\Delta}^{n-1}\Bigg(\underbrace{t,\dots,t}_{\times n}\Bigg)=\frac{1}{(n-1)!}\delta^{(n-1)}(x-t). \tag{6.7}$$

(7) *Popoviciu's refinement lemma* [9, Proposition 23]*: for every index subsequence*

$$1\leqslant\sigma(1)<\sigma(2)<\dots<\sigma(k)\leqslant n,$$

there exist coefficients $\alpha(j)$ *such that*

$$\mathbf{\Delta}^{k-1}\left(x_{\sigma(1)},\dots,x_{\sigma(k)}\right)=\sum_{j=\sigma(1)-1}^{\sigma(k)-k}\alpha(j)\,\mathbf{\Delta}^{k-1}(x_{j+1},x_{j+2},\dots,x_{j+k}). \tag{6.8}$$

Based on the above, we may now identify $\mathbf{\Delta}$ with elements of the bundle $\mathcal{P}_r$.

Definition 6.8. Let $w=(x_1,\dots,x_r)\in\mathbb{C}^r$, and $X=\{n_1,n_2,\dots,n_\alpha\}\subseteq\{1,2,\dots,r\}$ of size $|X|=\alpha$ be given. Let the elements of X be enumerated in increasing order, i.e.

$$1\leqslant n_1<n_2<\dots<n_\alpha\leqslant r.$$

Denote by w_X the vector

$$w_X\stackrel{\text{def}}{=}(x_{n_1},x_{n_2},\dots,x_{n_\alpha})\in\mathbb{C}^\alpha.$$

Then we denote

$$\Delta_X(w)\stackrel{\text{def}}{=}\mathbf{\Delta}^{\alpha-1}(w_X).$$

We immediately obtain the following result.

Lemma 6.9. *For all* $w\in\mathbb{C}^r$ *and* $X\subseteq\{1,2,\dots,r\}$*, we have* $\Delta_X(w)\in V_w$*. Moreover, letting* $\alpha=|X|$ *we have*

$$\mathcal{S}M(\Delta_X(w))=\mathbf{\Delta}^{\alpha-1}(w_X)\frac{1}{z-x}=\frac{1}{q_{\alpha,w_X}(z)}. \tag{6.9}$$

Finally, $(w,\Delta_X(w))$ *is a continuous section of* $\mathcal{P}_r$*.*

---

[1]The coefficients $\left\{a_{j,\ell}\right\}$ may be readily obtained by the Cauchy residue formula

$$a_{j,\ell}=\frac{1}{(d_j-1-\ell)!}\lim_{z\to\tau_j}\left(\frac{\mathrm{d}}{\mathrm{d}\,z}\right)^{d_j-1-\ell}\left\{\frac{(z-\tau_j)^{\ell+1}}{q_{n,w}(z)}\right\}.$$

6.3. **Constructing a basis.** The following result is well-known, see e.g. [9, Proposition 35].

Theorem 6.10. *Denote* $N_j = \{1, 2, \ldots, j\}$ *for* $j = 1, 2, \ldots, r$. *Then for every* $w \in \mathbb{C}^r$, *the collection*

$$\left\{\Delta_{N_j}(w)\right\}_{j=1}^{r}$$

is a basis for V_w.

There are various proofs of this statement. Below we show how to construct sets which do not necessarily remain basis for all $w \in \mathbb{C}^r$, but only for w in a small neighborhood of a given $w_0 \in \mathbb{C}^r$. Theorem 6.10 will then follow as a special case of this construction.

Informally, if two coordinates x_i and x_j can collide, then it is necessary to allow them to be glued by some element of the basis, i.e., we will need $\Delta_X(w)$ where $i, j \in X$ (in Theorem 6.10 all coordinates might be eventually glued into a single point because w is unrestricted.) In order to make this statement formal, let us introduce a notion of *configuration*, which is essentially a partition of the set of indices.

Definition 6.11. A *configuration* $\mathcal{C}$ is a partition of the set $N_r = \{1, 2, \ldots, r\}$ into $s = s(\mathcal{C})$ disjoint nonempty subsets

$$\sqcup_{i=1}^{s} X_i = N_r, \qquad |X_i| = d_i > 0.$$

The multiplicity vector of $\mathcal{C}$ is

$$T(\mathcal{C}) = (d_1, \ldots, d_s).$$

Every configuration defines a continuous family of divided differences as follows.

Definition 6.12. Let a configuration $\mathcal{C} = \{X_j\}_{j=1}^{s(\mathcal{C})}$. Enumerate each X_j in increasing order of its elements

$$X_j = \left\{n_1^j < n_2^j < \ldots n_{d_j}^j\right\}$$

and denote for every $m = 1, 2, \ldots, d_j$

$$X_{j,m} \stackrel{\text{def}}{=} \left\{n_k^j : \ k = 1, 2, \ldots, m\right\}.$$

For every $w \in \mathbb{C}^r$, the collection $\mathcal{B}_C(w) \subset V_w$ is defined as follows:

$$\mathcal{B}_C(w) \stackrel{\text{def}}{=} \left\{\Delta_{X_{j,m}}(w)\right\}_{j=1,\ldots,s(\mathcal{C})}^{m=1,\ldots,d_j}.$$

Now we formally define when a partition is "good" with respect to a point $w \in \mathbb{C}^r$.

Definition 6.13. The point $w = (x_1, \ldots, x_r) \in \mathbb{C}^r$ is *subordinated* to the configuration

$$\mathcal{C} = \{X_j\}_{j=1}^{s(\mathcal{C})}$$

if whenever $x_k = x_\ell$ for a pair of indices $k \neq \ell$, then necessarily $k, \ell \in X_j$ for some X_j.

Now we are ready to formulate the main result of this section.

Theorem 6.14. *For a given* $w_0 \in \mathbb{C}^r$ *and a configuration* $\mathcal{C}$, *the collection* $\mathcal{B}_C(w_0)$ *is a basis for* V_{w_0} *if and only if* w_0 *is subordinated to* $\mathcal{C}$. *In this case,* $\mathcal{B}_C(w)$ *is a continuous family of bases for* V_w *in a sufficiently small neighborhood of* w_0.

Let us first make a technical computation.

Lemma 6.15. *For a configuration* $\mathcal{C}$ *and a point* $w \in \mathbb{C}^r$, *consider for every fixed* $j = 1, \ldots, s(\mathcal{C})$ *the set*

$$S_j \stackrel{\text{def}}{=} \left\{\Delta_{X_{j,m}}(w)\right\}_{m=1}^{d_j}. \tag{6.10}$$

(1) *Define for any pair of indices $1 \leqslant k \leqslant \ell \leqslant d_j$ the index set*

$$X_{j,k:\ell} \stackrel{def}{=} \left\{n_k^j < n_{k+1}^j < \dots < n_\ell^j\right\} \subseteq X_j = X_{j,1:d_j} = X_{j,d_j}.$$

Then

$$\Delta_{X_{j,k:\ell}}(w) \in \operatorname{span} S_j.$$

(2) *For an arbitrary subset $Y \subseteq X_j$ (and not necessarily containing segments of consecutive indices), we also have*

$$\Delta_Y(w) \in \operatorname{span} S_j.$$

Proof. For clarity, we denote $y_i = x_{n_i^j}$ and $[k:\ell] = \Delta_{X_{j,k:\ell}}(w)$. By (6.3) we have in all cases (including repeated nodes)

$$(y_\ell - y_k)[k:\ell] = [k+1:\ell] - [k:\ell-1]. \tag{6.11}$$

The proof of the first statement is by backward induction on $n = \ell - k$. We start from $n = d_j$, and obviously $[1:d_j] \in S_j$. In addition, by definition of S_j we have $[1:m] \in S_j$ for all $m = 1, \dots, d_j$. Therefore, in order to obtain all $[k:\ell]$ with $\ell - k = n - 1$, we apply (6.11) several times as follows.

$$\begin{aligned}
[2:n] &= (y_n - y_1)[1:n] + [1:n-1] \\
[3:n+1] &= (y_{n+1} - y_2)\underset{\longleftrightarrow}{[2:n+1]} + \underline{[2:n]} \\
&\dots \\
[d_j - n + 2 : d_j] &= \left(y_{d_j} - y_{d_j - n+1}\right)\underset{\longleftrightarrow}{[d_j - n + 1 : d_j]} + \underline{[d_j - n + 1 : d_j - 1]}
\end{aligned}$$

Here the symbol $\underline{\cdots}$ under a term means that the term is taken directly from the previous line, while $\underset{\longleftrightarrow}{\cdots}$ indicates that the induction hypothesis is used. In the end, the left-hand side terms are shown to belong to $\operatorname{span} S_j$.

In order to prove the second statement, we employ the first statement, (6.8) and Proposition 6.7, Item 1. □

Proof of Theorem 6.14. In one direction, assume that $w_0 = (x_1, \dots, x_r)$ is subordinated to $\mathcal{C}$. It is sufficient to show that every element of the standard basis (2.2) belongs to $\operatorname{span}\{\mathcal{B}_C(w_0)\}$.

Let $\tau_j \in T(w_0)$, let d_j be the corresponding multiplicity, and let $Y_j \subseteq N_r$ denote the index set of size d_j

$$Y_j \stackrel{\text{def}}{=} \{i : \quad x_i = \tau_j\}.$$

By the definition of subordination, there exists an element in the partition of $\mathcal{C}$, say X_k, for which $Y_j \subseteq X_k$. By Lemma 6.15 we conclude that for all subsets $Z \subseteq Y_j$,

$$\Delta_Z(w_0) \in \operatorname{span}\left\{\Delta_{X_{k,m}}(w_0)\right\}_{m=1}^{|X_k|} \subseteq \operatorname{span}\{\mathcal{B}_C(w_0)\}.$$

By (6.7), $\Delta_Z(w_0)$ is nothing else but

$$\Delta_Z(w_0) = \Delta^{|Z|-1}\left(\underbrace{\tau_j, \dots, \tau_j}_{\times |Z|}\right) = \frac{1}{(|Z|-1)!}\delta^{(|Z|-1)}(x - \tau_j).$$

This completes the proof of the necessity. In the other direction, assume by contradiction that $x_k = x_\ell = \tau$ but nevertheless there exist two distinct elements of the partition $\mathcal{C}$, say X_α and

X_β such that $k \in X_\alpha$ and $\ell \in X_\beta$. Let the sets $\{S_j\}_{j=1}^{s(\mathcal{C})}$ be defined by (6.10). Again, by Lemma 6.15 and (6.7) we conclude that

$$\delta(x-\tau) \in \operatorname{span} S_\alpha \cap \operatorname{span} S_\beta.$$

But notice that $\mathcal{B}_C(w_0) = \bigcup_{j=1}^{s(\mathcal{C})} S_j$ and $\sum_{j=1}^{s} |S_j| = d$, therefore by counting dimensions we conclude that

$$\dim \operatorname{span}\{\mathcal{B}_C(w_0)\} < d,$$

in contradiction to the assumption that $\mathcal{B}_C(w_0)$ is a basis.

Finally, one can evidently choose a sufficiently small neighborhood $U \subset \mathbb{C}^r$ of w_0 such that for all $w \in U$, no new collisions are introduced, i.e., w is still subordinated to $\mathcal{C}$. The continuity argument (Lemma 6.9) finishes the proof. ☐

Remark 6.16. Another possible method of proof is to consider the algebra of elementary fractions in the Stieltjes space $\mathcal{S}_r$, and use the correspondence (6.9).

As we mentioned, Theorem 6.10 follows as a corollary of Theorem 6.14 for the configuration $\mathcal{C}$ consisting of a single partition set N_r.

6.4. **Resolution of collision singularities.** Let $\mu_0^* \in \Sigma_r^* \subset \mathbb{C}^{2r}$ be given, and let $(w_0, g_0) \in \mathcal{P}_r$ be a solution to the (rank-restricted) Prony problem. The point w_0 is uniquely defined up to a permutation of the coordinates, so we just fix a particular permutation. Let $T(w_0) = (\tau_1, \dots, \tau_s)$.

Our goal is to solve the rank-restricted Prony problem for every input $\mu^* \in \mathbb{C}^{2r}$ in a small neighborhood of μ_0^*. According to Theorem 5.7, this amounts to a continuous representation of the solution $R_{\mu^*}(z) = \frac{P_{\mu^*}(z)}{Q_{\mu^*}(z)} = \mathcal{T} M_r^{-1}(\mu^*)$ to the corresponding diagonal Padé approximation problem as an element of the bundle $\mathcal{P}_r$.

Define $\delta = \min_{i \neq j} |\tau_i - \tau_j|$ to be the "separation distance" between the clusters. Since the roots of Q_{μ^*} depend continuously on μ^* and the degree of Q_{μ^*} does not drop, we can choose some μ_1^* sufficiently close to μ_0^*, for which

(1) all the roots of $Q_{\mu_1^*}(z)$ are distinct, and
(2) these roots can be grouped into s clusters, such that each of the elements of the j-th cluster is at most $\delta/3$ away from τ_j.

Enumerate the roots of $Q_{\mu_1^*}$ within each cluster in an arbitrary manner. This choice enables us to define locally (in a neighborhood of μ_1^*) r algebraic functions $x_1(\mu^*), \dots, x_r(\mu^*)$, satisfying

$$Q_{\mu^*}(z) = \prod_{j=1}^{s} (z - x_j(\mu^*)).$$

Then we extend these functions by analytic continuation according to the above formula into the entire neighborhood of μ_0^*. Consequently,

$$w(\mu^*) \stackrel{\text{def}}{=} (x_1(\mu^*), \dots, x_r(\mu^*))$$

is a continuous (multivalued) algebraic function in a neighborhood of μ_0^*, satisfying

$$w(\mu_0^*) = w_0.$$

After this "pre-processing" step, we can solve the rank-restricted Prony problem in this neighborhood of μ_0^*, as follows.

Algorithm 1 Solving rank-restricted Prony problem with collisions.

Let $\mu_0^* \in \Sigma_r^* \subset \mathbb{C}^{2r}$ be given, and let $(w_0, g_0) \in \mathcal{P}_r$ be a solution to the (rank-restricted) Prony problem. Let w_0 be subordinated to some configuration $\mathcal{C}$.

The input to the problem is a measurement vector $\mu^* = (m_0, \dots, m_{2r-1}) \in \mathbb{C}^{2r}$, which is in a small neighborhood of μ_0^*.

(1) Construct the function $w = w(\mu^*)$ as described above.

(2) Build the basis $\mathcal{B}_C(w) = \left\{\Delta_{X_{j,\ell}}(w)\right\}_{j=1,\dots,s(\mathcal{C})}^{\ell=1,\dots,d_j}$ for V_w.

(3) Find the coefficients $\{\beta_{j,\ell}\}_{j=1,\dots,s(\mathcal{C})}^{\ell=1,\dots,d_j}$ such that

$$\mathcal{S}M\left(\sum_{j,\ell}\beta_{j,\ell}\Delta_{X_{j,\ell}}(w)\right) = R(z),$$

by solving the linear system

$$\underbrace{\sum_{j,\ell}\beta_{j,\ell}(w)\,\Delta_{X_{j,\ell}}(w)}_{=g(w)}\left(x^k\right) = m_k\left(=\int x^k g(w)(x)\,\mathrm{d}\,x\right), \qquad k = 0,1,\dots,2r-1. \tag{6.12}$$

Theorem 6.17. *The coordinates $\{\beta_{j,\ell}\}$ of the solution to the rank-restricted Prony problem, given by Algorithm 6.4, are (multivalued) algebraic functions, continuous in a neighborhood of the point μ_0^* .*

Proof. Since the divided differences $\Delta_{j,\ell}(w)$ are continuous in w, then clearly for each

$$k = 0,1,\dots,2r-1$$

the functions

$$\nu_{j,\ell,k}(w) = \Delta_{j,\ell}(w)\left(x^k\right) = \Delta^{\ell-1}\left(w_{X_{j,\ell}}\right)\left(x^k\right)$$

are continuous[2] in w, and hence continuous, as multivalued functions, in a neighborhood of μ_0^*. Since $\mathcal{B}_C(w(\mu^*))$ remains a basis in a (possibly smaller) neighborhood of μ_0^*, the system (6.12), taking the form

$$\sum_{j,\ell}\nu_{j,\ell,k}(w)\,\beta_{j,\ell}(w) = m_k, \qquad k = 0,1,\dots,2r-1,$$

remains non-degenerate in this neighborhood. We conclude that the coefficients $\{\beta_{j,\ell}(w(\mu^*))\}$ are multivalued algebraic functions, continuous in a neighborhood of μ_0^*. □

7. Real Prony space and hyperbolic polynomials

In this section we shall restrict ourselves to the real case. Notice that in many applications only real Prony systems are used. On the other hand, considering the Prony problem over the real numbers significantly simplifies some constructions. In particular, we can easily avoid topological problems, related with the choice of the ordering of the points $x_1, \dots, x_d \in \mathbb{C}$. So in a definition of the real Prony space $R\mathcal{P}_d$ we assume that the coordinates $x_1, \dots, x_d$ are taken with their natural ordering $x_1 \le x_2 \le \dots \le x_d$. Accordingly, the real Prony space $R\mathcal{P}_d$ is defined as the bundle (w, g), $w \in \prod_d \subset \mathbb{R}^d, g \in RV_w$. Here $\prod_d$ is the prism in $\mathbb{R}^d$ defined by the inequalities $x_1 \le x_2 \le \dots \le x_d$, and RV_w is the space of linear combinations with real coefficients of δ-functions and their derivatives with the support $\{x_1, \dots, x_d\}$, as in Definition

[2] In fact, $\nu_{j,\ell,k}(w)$ are symmetric polynomials in some of the coordinates of w.

2.4. The Prony, Stieltjes and Taylor maps are the restrictions to the real case of the complex maps defined above.

In this paper we just point out a remarkable connection of the real Prony space and mapping with hyperbolic polynomials, and Vieta and Vandermonde mappings studied in Singularity Theory (see [1, 13, 14, 15] and references therein).

Hyperbolic polynomials (in one variable) are real polynomials $Q(z) = z^d + \sum_{j=1}^{d} \lambda_j z^{d-j}$, with all d of their roots real. We denote by Γ_d the space of the coefficients $\Lambda = (\lambda_1, \ldots, \lambda_d) \subset \mathbb{R}^d$ of all the hyperbolic polynomials, and by $\hat{\Gamma}_d$ the set of $\Lambda \in \Gamma_d$ with $\lambda_1 = 0,\ |\lambda_2| \leq 1$. Recalling (2.3), it is evident that all hyperbolic polynomials appear as the denominators of the irreducible fractions in the image of $R\mathcal{P}_d$ by $\mathcal{S}M$. This shows, in particular, that the geometry of the boundary $\partial\Gamma$ of the hyperbolicity domain Γ is important in the study of the real Prony map $\mathcal{P}M$: it is mapped by $\mathcal{P}M$ to the boundary of the solvability domain of the real Prony problem. This geometry has been studied in a number of publications, from the middle of 1980s. In [13] V. P. Kostov has shown that $\hat{\Gamma}$ possesses the Whitney property: there is a constant C such that any two points $\lambda_1, \lambda_2 \in \hat{\Gamma}$ can be connected by a curve inside $\hat{\Gamma}$ of the length at most $C\|\lambda_2 - \lambda_1\|$. "Vieta mapping" which associates to the nodes $x_1 \leq x_2 \leq \cdots \leq x_d$ the coefficients of $Q(z)$ having these nodes as the roots, is also studied in [13]. In our notations, Vieta mapping is the composition of the Stieltjes mapping $\mathcal{S}M$ with the projection to the coefficients of the denominator.

In [1] V.I.Arnold introduced and studied the notion of maximal hyperbolic polynomial, relevant in description of $\hat{\Gamma}$. Furthermore, the Vandermonde mapping $\mathcal{V} : \mathbb{R}^d \to \mathbb{R}^d$ was defined there by

$$\begin{cases} y_1 = a_1 x_1 + \ldots + a_d x_d, \\ \ldots \\ y_d = a_1 x_1^d + \ldots + a_d x_d^d, \end{cases}$$

with $a_1, \ldots, a_d$ fixed. In our notations $\mathcal{V}$ is the restriction of the Prony mapping to the pairs $(w, g) \in R\mathcal{P}_d$ with the coefficients of g in the standard basis of RV_w fixed. It was shown in [1] that for $a_1, \ldots, a_d > 0$ $\mathcal{V}$ is a one-to-one mapping of $\prod_d$ to its image. In other words, the first d moments uniquely define the nodes $x_1 \leq x_2 \leq \cdots \leq x_d$. For $a_1, \ldots, a_d$ with varying signs, this is no longer true in general. This result is applied in [1] to the study of the colliding configurations.

Next, the "Vandermonde varieties" are studied in [1], which are defined by the equations

$$\begin{cases} a_1 x_1 + \ldots + a_d x_d & = \alpha_1, \\ & \ldots \\ a_1 x_1^\ell + \ldots + a_d x_d^\ell & = \alpha_\ell. \end{cases} \qquad \ell \leqslant d.$$

It is shown that for $a_1, \ldots, a_d > 0$ the intersections of such varieties with $\prod_d$ are either contractible or empty. Finally, the critical points of the next Vandermonde equation on the Vandermond variety are studied in detail, and on this base a new proof of Kostov's theorem is given.

We believe that the results of [1, 13] and their continuation in [14, 15] and other publications are important for the study of the Prony problem over the reals, and we plan to present some results in this direction separately.

Appendix A. Proof of Theorem 3.5

Recall that we are interested in finding conditions for which the Taylor mapping $\mathcal{T}M: \mathcal{S}_d \to \mathcal{T}_d$ is invertible. In other words, given

$$S(z) = \sum_{k=0}^{2d-1} m_k \left(\frac{1}{z}\right)^{k+1},$$

we are looking for a rational function $R(z) \in \mathcal{S}_d$ such that

$$S(z) - R(z) = \frac{d_1}{z^{2d+1}} + \frac{d_2}{z^{2d+2}} + \dots. \tag{A.1}$$

Write $R(z) = \frac{P(z)}{Q(z)}$ with $Q(z) = \sum_{j=0}^{d} c_j z^j$ and $P(z) = \sum_{i=0}^{d-1} b_i z^i$. Multiplying (A.1) by $Q(z)$, we obtain

$$Q(z) S(z) - P(z) = \frac{e_1}{z^{d+1}} + \frac{e_2}{z^{d+2}} + \dots. \tag{A.2}$$

Proposition A.1. *The identity* (A.2)*, considered as an equation on P and Q with*

$$\deg P < \deg Q \le d,$$

always has a solution.

Proof. Substituting the expressions for S, P and Q into (A.2) we get

$$\left(c_0 + c_1 z + \dots + c_d z^d\right)\left(\frac{m_0}{z} + \frac{m_1}{z^2} + \dots\right) - b_0 - \dots - b_{d-1} z^{d-1} = \frac{e_1}{z^{d+1}} + \dots. \tag{A.3}$$

The highest degree of z in the left hand side of (A.3) is $d-1$. So equating to zero the coefficients of z^s in (A.3) for $s = d-1, \dots, -d$ we get the following systems of equations:

$$\begin{bmatrix} 0 & 0 & 0 & m_0 \\ 0 & 0 & m_0 & m_1 \\ \iddots & \iddots & & \\ m_0 & m_1 & \dots & m_{d-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_d \end{bmatrix} = \begin{bmatrix} b_{d-1} \\ b_{d-2} \\ \vdots \\ b_0 \end{bmatrix}. \tag{A.$\star$}$$

From this point on, the equations become homogeneous:

$$\begin{bmatrix} m_0 & m_1 & \dots & m_d \\ m_1 & m_2 & \dots & m_{d+1} \\ \iddots & \iddots & & \\ m_{d-1} & m_d & \dots & m_{2d-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{A.$\star\star$}$$

The homogeneous system (A.$\star\star$) has the Hankel-type $d \times (d+1)$ matrix $\tilde{M}_d = (m_{i+j})$ with $0 \leqslant i \leqslant d-1$ and $0 \leqslant j \leqslant d$. This system has d equations and $d+1$ unknowns $c_0, \dots, c_d$. Consequently, it always has a nonzero solution $c_0, \dots, c_d$. Now substituting these coefficients $c_0, \dots, c_d$ of Q into the equations (A.$\star$) we find the coefficients $b_0, \dots, b_{d-1}$ of the polynomial P, satisfying (A.$\star$). Notice that if $c_j = 0$ for $j \geqslant \ell + 1$ then it follows from the structure of the equations (A.$\star$) that $b_j = 0$ for $j \ge \ell$. Hence these P, Q provide a solution of (A.2), satisfying $\deg P < \deg Q \le d$, and hence belonging to $\mathcal{S}_d$. □

However, in general (A.2) does not imply (A.1). This implication holds only if $\deg Q = d$. The following proposition describes a possible "loss of accuracy" as we return from (A.2) to (A.1) and $\deg Q < d$:

Proposition A.2. *Let (A.2) be satisfied with the highest nonzero coefficient of Q being c_ℓ, $\ell \le d$. Then*

$$S(z) - \frac{P(z)}{Q(z)} = \frac{d_1}{z^{d+\ell+1}} + \frac{d_2}{z^{d+\ell+2}} + \dots. \tag{A.4}$$

Proof. We notice that if the leading nonzero coefficient of Q is c_ℓ then we have

$$\frac{1}{Q} = \frac{1}{z^\ell}\left(\frac{1}{c_\ell + \frac{c_{\ell-1}}{z} + \dots}\right) = \frac{1}{z^\ell}(f_0 + f_1\frac{1}{z} + \dots).$$

So multiplying (A.2) by $\frac{1}{Q}$ we get (A.4). □

Proof of Theorem 3.5. Assume that the rank of $\tilde{M}_d$ is $r \le d$, and that $|M_r| \ne 0$. Let us find a polynomial $Q(z)$ of degree r of the form $Q(z) = z^r + \sum_{j=0}^{r-1} c_j z^j$, whose coefficients satisfy system (A.$\star\star$). Put $\mathbf{c}_r = (c_0, \dots, c_{r-1}, 1)^T$ and consider a linear system $\tilde{M}_r \mathbf{c}_r = 0$. Since by assumptions $|M_r| \ne 0$, this system has a unique solution. Extend this solution by zeroes, i.e., put $\mathbf{c}_d = (c_0, \dots, c_{r-1}, 1, 0, \dots, 0)^T$. We want $\mathbf{c}_d$ to satisfy (A.$\star\star$), which is $\tilde{M}_d \mathbf{c}_d = 0$. This fact is immediate for the first r rows of $\tilde{M}_d$. But since the rank of $\tilde{M}_d$ is r by the assumption, its other rows are linear combinations of the first r ones. Hence $\mathbf{c}_d$ satisfies (A.$\star\star$).

Now the equations (A.$\star$) produce a polynomial $P(z)$ of degree at most $r-1$. So we get a rational function $R(z) = \frac{P(z)}{Q(z)} \in \mathcal{S}_r \subseteq \mathcal{S}_d$ which solves the Padé problem (A.2), with $\deg Q(z) = r$. Write $R(z) = \sum_{k=0}^{\infty} \alpha_k (\frac{1}{z})^{k+1}$. By Proposition A.2 we have $m_k = \alpha_k$ till $k = d + r - 1$.

Now, the Taylor coefficients α_k of $R(z)$ satisfy a linear recurrence relation

$$m_k = -\sum_{s=1}^{r} c_s m_{k-s}, \qquad k = r, r+1, \dots. \tag{A.5}$$

Considering the rows of the system $\tilde{M}_d \mathbf{c}_d = 0$ we see that m_k satisfy the same recurrence relation (A.5) till $k = d + r - 1$ (we already know that $m_k = \alpha_k$ till $k = d + r - 1$). We shall show that in fact m_k satisfy (A.5) till $k = 2d - 1$.

Consider a $d \times r$ matrix $\bar{M}_d$ formed by the first r columns of M_d, and denote its row vectors by $\mathbf{v}_i = (m_{i,0}, \dots, m_{i,r-1}),\ i = 1, \dots, d-1$. The vectors $\mathbf{v}_i$ satisfy

$$\mathbf{v}_i = -\sum_{s=1}^{r} c_s \mathbf{v}_{i-s}, \quad i = r, \dots, d-1, \tag{A.6}$$

since their coordinates satisfy (A.5) till $k = d + r - 1$. Now $\mathbf{v}_0, \dots, \mathbf{v}_{r-1}$ are linearly independent, and hence each $\mathbf{v}_i,\ i = r, \dots, d-1$, can be expressed as

$$\mathbf{v}_i = \sum_{s=0}^{r-1} \gamma_{i,s} \mathbf{v}_s. \tag{A.7}$$

Denote by $\tilde{\mathbf{v}}_i = (m_{i,0}, \dots, m_{i,d}),\ i = 1, \dots, d-1$ the row vectors of $\tilde{M}_d$. Since by assumptions the rank of $\tilde{M}_d$ is r, the vectors $\tilde{\mathbf{v}}_i$ can be expressed through the first r of them exactly in the same form as $\mathbf{v}_i$:

$$\tilde{\mathbf{v}}_i = \sum_{s=0}^{r-1} \gamma_{i,s} \tilde{\mathbf{v}}_s, \quad i = r, \dots, d-1. \tag{A.8}$$

Now the property of a system of vectors to satisfy the linear recurrence relation (A.6) depends only on the coefficients $\gamma_{i,s}$ in their representation (A.7) or (A.8). Hence from (A.6) we conclude that the full rows $\tilde{\mathbf{v}}_i$ of $\tilde{M}_d$ satisfy the same recurrence relation. Coordinate-wise this implies that m_k satisfy (A.5) till $k=2d-1$, and hence $m_k=\alpha_k$ till $k=2d-1$. So $R(z)$ solves the original Problem 3.1.

In the opposite direction, assume that $R(z)$ solves Problem 3.1, and that the representation $R(z)=\frac{P(z)}{Q(z)}\in\mathcal{S}_r\subset\mathcal{S}_d$ is irreducible, i.e., $\deg Q=r$. Write $Q(z)=z^r+\sum_{j=0}^{r-1}c_jz^j$. Then m_k, being the Taylor coefficients of $R(z)$ till $k=2d-1$, satisfy a linear recurrence relation (A.5): $m_k=-\sum_{s=1}^{r}c_sm_{k-s},\ k=r,r+1,\dots,2d-1$. Applying this relation coordinate-wise to the rows of $\tilde{M}_d$ we conclude that all the rows can be linearly expressed through the first r ones. So the rank of $\tilde{M}_d$ is at most r.

It remains to show that the left upper minor $|M_r|$ is non-zero, and hence the rank of $\tilde{M}_d$ is exactly r.

By Proposition 3.3, if the decomposition of $R(z)$ in the standard basis is

$$R\left(z\right)=\sum_{j=1}^{s}\sum_{\ell=1}^{d_j}a_{j,\ell-1}\frac{\left(-1\right)^{\ell-1}\left(\ell-1\right)!}{\left(z-x_j\right)^{\ell}},$$

where $\sum_{j=1}^{s}d_j=r$ and $\{x_j\}$ are pairwise distinct, then the Taylor coefficients of $R(z)$ are given by (1.5). Clearly, we must have $a_{j,d_j-1}\neq 0$ for all $j=1,\dots,s$, otherwise $\deg Q<r$, a contradiction. Now consider the following well-known representation of M_r as a product of three matrices (see e.g. [7]):

$$M_r=V\left(x_1,d_1,\dots,x_s,d_s\right)\times\operatorname{diag}\left\{A_j\right\}_{j=1}^{s}\times V\left(x_1,d_1,\dots,x_s,d_s\right)^{T},\tag{A.9}$$

where $V(\dots)$ is the confluent Vandermonde matrix (4.1) and each A_j is the following $d_j\times d_j$ block:

$$A_j\stackrel{\text{def}}{=}\begin{bmatrix} a_{j,0} & a_{j,1} & \cdots & \cdots & a_{j,d_j-1}\\ a_{j,1} & & & \binom{d_j-1}{d_j-2}a_{j,d_j-1} & 0\\ \cdots & & & \cdots & 0\\ & \binom{d_j-1}{2}a_{j,d_j-1} & 0 & \cdots & 0\\ a_{j,d_j-1} & 0 & \cdots & \cdots & 0\end{bmatrix}.$$

The formula (A.9) can be checked by direct computation. Since $\{x_j\}$ are pairwise distinct and $a_{j,d_j-1}\neq 0$ for all $j=1,\dots,s$, we immediately conclude that $|M_r|\neq 0$.

This finishes the proof of Theorem 3.5. ☐

References

[1] V.I. Arnol'd. Hyperbolic polynomials and Vandermonde mappings. *Functional Analysis and Its Applications*, 20(2):125–127, 1986. DOI: 10.1007/BF01077267

[2] J.R. Auton. Investigation of Procedures for Automatic Resonance Extraction from Noisy Transient Electromagnetics Data. Volume III. Translation of Prony's Original Paper and Bibliography of Prony's Method. Technical report, Effects Technology Inc., Santa Barbara, CA, 1981.

[3] R. Badeau, B. David, and G. Richard. Performance of ESPRIT for estimating mixtures of complex exponentials modulated by polynomials. *IEEE Transactions on Signal Processing*, 56(2):492–504, 2008. DOI: 10.1109/TSP.2007.906744

[4] G.A. Baker and P. Graves-Morris. Pade approximants. Part 1: Basic theory. *Encyclopedia of Mathematics and its applications*, 1981.

[5] D. Batenkov. Complete Algebraic Reconstruction of Piecewise-Smooth Functions from Fourier Data. *To appear in Mathematics of Computation.*

[6] D. Batenkov. Numerical stability bounds for algebraic systems of Prony type and their accurate solution by decimation. September 2014. arXiv:1409.3137

[7] D. Batenkov and Y. Yomdin. On the accuracy of solving confluent Prony systems. *SIAM J. Appl. Math.*, 73(1):134–154, 2013. DOI: 10.1137/110836584

[8] Emmanuel J. Candès and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, June 2014. DOI: 10.1002/cpa.21455

[9] C. de Boor. Divided differences. *Surv. Approx. Theory*, 1:46–69, 2005.
[10] D.L. Donoho, M. Elad, and V.N. Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *Information Theory, IEEE Transactions on*, 52(1):6–18, 2006.
[11] B. Gustafsson, C. He, P. Milanfar, and M. Putinar. Reconstructing planar domains from their moments. *Inverse Problems*, 16(4):1053–1070, 2000. DOI: 10.1088/0266-5611/16/4/312
[12] W. Kahan and R.J. Fateman. Symbolic computation of divided differences. *ACM SIGSAM Bulletin*, 33(2):7–28, 1999. DOI: 10.1145/334714.334716
[13] V.P. Kostov. On the geometric properties of Vandermonde's mapping and on the problem of moments. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 112(3-4):203–211, 1989. DOI: 10.1017/S0308210500018679
[14] V.P. Kostov. Root arrangements of hyperbolic polynomial-like functions. *Revista matemática complutense*, 19(1):197–225, 2006.
[15] V.P. Kostov. On root arrangements for hyperbolic polynomial-like functions and their derivatives. *Bulletin des sciences mathematiques*, 131(5):477–492, 2007. DOI: 10.1016/j.bulsci.2006.12.004
[16] L.M. Milne-Thompson. *The calculus of finite differences*. Macmillan, 1933.
[17] T. Peter, D. Potts, and M. Tasche. Nonlinear approximation by sums of exponentials and translates. *SIAM Journal on Scientific Computing*, 33(4):1920, 2011. DOI: 10.1137/100790094
[18] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Processing*, 90(5):1631–1642, 2010. DOI: 10.1016/j.sigpro.2009.11.012
[19] R. Prony. Essai experimental et analytique. *J. Ec. Polytech.(Paris)*, 2:24–76, 1795.
[20] P. Stoica and N. Arye. MUSIC, maximum likelihood, and Cramer-Rao bound. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(5):720–741, 1989. DOI: 10.1109/29.17564
[21] Y. Yomdin. Singularities in Algebraic Data Acquisition. In M. Manoel, M.C.R. Fuster, and C.T.C. Wall, editors, *Real and Complex Singularities*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511731983.026

Department of Computer Science, The Technion - Israel Institute of Technology, Technion City, Haifa 32000, Israel
E-mail address: batenkov@cs.technion.ac.il
URL: http://dimabatenkov.info

Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel
E-mail address: yosef.yomdin@weizmann.ac.il
URL: http://www.wisdom.weizmann.ac.il/~yomdin