Moment inversion problem for piecewise *D*-finite functions

Moment inversion problem for piecewise *D*-finite functions

Dmitry Batenkov

Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel

E-mail: dima.batenkov@weizmann.ac.il

Abstract

We consider the problem of exact reconstruction of univariate functions with jump discontinuities at unknown positions from their moments. These functions are assumed to satisfy an *a priori unknown* linear homogeneous differential equation with polynomial coefficients on each continuity interval. Therefore, they may be specified by a finite amount of information. This reconstruction problem has practical importance in signal processing and other applications.

It is somewhat of a 'folklore' that the sequence of the moments of such 'piecewise *D*-finite' functions satisfies a linear recurrence relation of bounded order and degree. We derive this recurrence relation explicitly. It turns out that the coefficients of the differential operator which annihilates every piece of the function, as well as the locations of the discontinuities, appear in this recurrence in a precisely controlled manner. This leads to the formulation of a generic algorithm for reconstructing a piecewise *D*-finite function from its moments. We investigate the conditions for solvability of resulting linear systems in the general case, as well as analyse a few particular examples. We provide results of numerical simulations for several types of signals, which test the sensitivity of the proposed algorithm to noise.

1. Introduction

Consider the problem of reconstructing an unknown function $g : [a, b] \to \mathbb{R}$ from some finite number of its power moments

$$m_k(g) = \int_a^b x^k g(x) \, \mathrm{d}x, \qquad k = 0, 1, \ldots, M. \qquad (1.1)$$

This formulation is a 'prototype' for various problems in signal processing, statistics, computer tomography and other areas (see [1, 17, 27] and references therein). In all practical applications, it is assumed that g belongs to some *a priori known* class, and it can be faithfully

specified by a *finite* number of parameters in that class. For example, smooth signals may be represented as elements of some finite-dimensional Hilbert space and then the reconstruction problem is analysed in the classical framework of linear approximation—see [1, 32] for thorough expositions.

In recent years, novel algebraic techniques for moment inversion have been developed. One reason for their appearance is the unsatisfactory performance of the classical approximation methods when applied to irregular data. The famous 'Gibbs effect' due to a jump discontinuity probably provides the best-known example of such undesired behaviour. In [14], it is shown that only nonlinear methods have a chance to achieve the same order of approximation for such discontinuous signals as conventional (linear) methods do for smooth signals. The various nonlinear methods developed recently include the framework of signals with a finite rate of innovation [5, 9, 33], Padé-based methods [3, 10, 25], and other algebraic schemes [4, 11, 22]. Methods for reconstructing planar shapes from complex moments are developed in [17, 18]. In [21], an approach based on Cauchy-type integrals is presented. In [27], several of the above methods are reviewed in detail, while emphasizing the mathematical similarity between the corresponding inversion systems.

The present paper develops a general method for explicit inversion of the moment transform for functions which are piecewise solutions of linear ODEs with polynomial coefficients (precise definitions follow). In particular, the method allows us to locate the discontinuities of the function by purely algebraic means. The algebra involved is in the spirit of holonomic combinatorics [35]. We believe that the tools developed in this paper may shed a new light on the similar structure of the algebraic equations which appear in [11, 18, 25, 27, 33].

1.1. Overview of main results

Let $\mathfrak{D}$ denote an arbitrary linear differential operator with polynomial coefficients

$$\mathfrak{D} = \sum_{j=0}^{N} p_j(x)\partial^j, \tag{1.2a}$$

$$p_j(x) = \sum_{i=0}^{k_j} a_{i,j} x^i, \qquad a_{i,j} \in \mathbb{R}, \tag{1.2b}$$

where ∂ is the differentiation operator with respect to x: $\partial^i f \stackrel{\text{def}}{=} \frac{\mathrm{d}^i}{\mathrm{d}x^i} f = f^{(n)}(x)$ and $\partial^0 = \mathrm{I}$, the identity operator. If some function g satisfies $\mathfrak{D}g \equiv 0$, we say that $\mathfrak{D}$ *annihilates* g and also write $g \in \mathcal{N}_{\mathfrak{D}} \stackrel{\text{def}}{=} \{f : \mathfrak{D}f \equiv 0\}$. These functions are called '*D*-finite' ('differentiably finite' [30]).

The class $A_{\mathfrak{D}}$ of 'piecewise *D*-finite' functions is defined in the following way: for $n = 0, 1, \ldots, \mathcal{K}$, let $\Delta_n \stackrel{\text{def}}{=} [\xi_n, \xi_{n+1}]$ be a partition of $[a, b]$ such that $-\infty < a = \xi_0 < \xi_1 < \cdots < \xi_{\mathcal{K}+1} = b < +\infty$ (the case $\mathcal{K} = 0$ corresponds to functions consisting of a single 'piece'). We say that $g \in A_{\mathfrak{D}}$ if there exist operators $\mathfrak{D}_n$ such that on each 'continuity interval' Δ_n, the nth piece of g equals, say, $g_n(x)$, such that $\mathfrak{D}_n g_n \equiv 0$. In the most general setting, the annihilating operators $\mathfrak{D}_n$ may be pairwise different. In this paper, we explicitly treat functions for which $\mathfrak{D}_n \equiv \mathfrak{D}$ for all $0 \leqslant n \leqslant \mathcal{K}$. We write $g \in A_{\mathfrak{D}}^*$ in this particular case.

We shall always assume that the leading coefficient of $\mathfrak{D}_n$ does not vanish at any point of Δ_n.

We are interested in solving the following problem.

Operator-Based Moment Reconstruction Problem. Let us give the sequence of the moments (1.1) of an unknown function $g \in A_{\mathfrak{D}}$ and the constants $\mathcal{K}$, N, $\{k_j\}$, a, b.

(a) Find the operator $\mathfrak{D}$ (in the general case, the operators $\mathfrak{D}_n$), that is, determine the coefficients $a_{i,j}$ (respectively $a_{i,j,n}$) as in (1.2) such that $\mathfrak{D}_n g_n \equiv 0$ for all $0 \leqslant n \leqslant \mathcal{K}$.
(b) Find the 'jump points' $\{\xi_n\}_{n=1}^{\mathcal{K}}$.
(c) Let $\{u_i\}_{i=1}^{N}$ be a basis for the linear space $\mathcal{N}_{\mathfrak{D}}$. Find constants $\alpha_{i,n}$ such that $g_n(x) = \sum_{i=1}^{N} \alpha_{i,n} u_i(x)$. That is, determine each concrete solution of $\mathfrak{D}_n g_n = 0$ on Δ_n.

Example 1.1. Let $g(x) = \alpha \, \mathrm{e}^{\beta x}$ be our unknown function on $[0, 1]$. The parameters of the problem are $\mathcal{K} = 0$, $\mathfrak{D} = \partial - \beta \mathrm{I}$, $[a, b] = [0, 1]$, and the unknowns are α, β. Direct calculation of the moments yields

$$\begin{cases} \alpha(\mathrm{e}^{\beta} - 1) = \beta m_0 \\ \quad\;\; \alpha \, \mathrm{e}^{\beta} = \beta m_1 + m_0. \end{cases} \tag{1.3}$$

Denoting $\frac{m_1}{m_0} \stackrel{\text{def}}{=} A$, (1.3) amounts to

$$\begin{cases} \dfrac{1}{1 - \mathrm{e}^{-\beta}} - \dfrac{1}{\beta} = A \\ \quad\;\; \dfrac{\beta m_0}{\mathrm{e}^{\beta} - 1} = \alpha. \end{cases} \tag{1.4}$$

This system does not have an explicit analytic solution. It can be solved numerically (for example, using Newton's method), and a unique solution exists provided $A \in (0, 1)$ (the graph of the function $y = \frac{1}{1-\mathrm{e}^{-x}} - \frac{1}{x}$ is monotone in whole $\mathbb{R}$ and $0 < y(x) < 1$).

The above solution is unsatisfactory for several reasons. First, it is not general enough. Second, the solution is available only as an approximation and not in a closed form. However, there is an important positive feature: the minimal possible number of measurements are used. Using our method, example 1.1 will be solved later in a more convenient and general way—see examples 2.2 and 3.1.

Our method is based on the following results which we prove below (sections 2.1, 2.2 and 3.1). These results establish explicit relations between the known and the unknown parameters of the reconstruction problem.

Theorem [2.9, 3.1]. *Let $\mathcal{K} = 0$ and $\mathfrak{D}g \equiv 0$. Then the moment sequence of g satisfies a linear recurrence relation with coefficients linear in $a_{i,j}$. Consequently, the vector $\mathbf{a} = (a_{i,j})$ satisfies a linear homogeneous system $H\mathbf{a} = 0$, where the entries of H are certain linear combinations of the moments of g whose coefficients depend only on N and the endpoints a, b.*

Theorem [2.12]. *Let $\mathcal{K} > 0$ and let $\mathfrak{D}$ annihilate every piece of $g \in A_{\mathfrak{D}}^*$. Then the operator given by $\widehat{\mathfrak{D}} = \left(\prod_{i=1}^{\mathcal{K}} (x - \xi_i)^N \mathrm{I} \right) \cdot \mathfrak{D}$ annihilates g as a distribution. Consequently, conclusions of theorems 2.9 and 3.1 are true with $\mathfrak{D}$ replaced by $\widehat{\mathfrak{D}}$.*

Proposition [3.3]. *Let $g \in A_{\mathfrak{D}}^*$ with operator $\mathfrak{D}$ annihilating every piece g_n. Let $\{u_i\}_{i=1}^{N}$ be a basis for the linear space $\mathcal{N}_{\mathfrak{D}}$ and $g_n(x) = \sum_{i=1}^{N} \alpha_{i,n} u_i(x)$. Then the vector α of the coefficients $\alpha_{i,n}$ satisfies a linear system $C\alpha = \mathbf{m}$, where the matrix C contains the moments of u_i and the vector $\mathbf{m}$ contains the moments of g.*

Based on the above results, the proposed solution to the reconstruction problem is as follows (section 3.2).

(1) If $\mathcal{K} > 0$, replace $\mathfrak{D}$ with $\widehat{\mathfrak{D}} = \left(\prod_{i=1}^{\mathcal{K}} (x - \xi_i)^N \mathrm{I} \right) \cdot \mathfrak{D}$.

(2) Build the matrix H and solve the system $H\mathbf{x} = 0$, where $\mathbf{x}$ is the vector of unknown coefficients of $\mathfrak{D}$ or $\widehat{\mathfrak{D}}$ according to the previous step. Obtain a solution $\mathbf{a}$ and build the differential operator $\mathfrak{D}^* = \mathfrak{D}_{\mathbf{a}}$ which annihilates g in its entirety (as a distribution in the case $\mathcal{K} > 0$ or in the usual sense otherwise).

(3) If $\mathcal{K} > 0$, recover $\{\xi_i\}$ and the operator $\mathfrak{D}^\dagger$ (which annihilates every piece of g) from $\mathfrak{D}^*$.

(4) Compute the basis for $\mathcal{N}_{\mathfrak{D}^\dagger}$ and solve the system $C\alpha = \mathbf{m}$.

Conditions for the solvability of the above systems are discussed and few initial results in this direction are obtained in section 3.3. Results of numerical simulations, presented in section 4, suggest that the method is practically applicable to several different signal models, including piecewise constant functions (and piecewise polynomials), rational functions and piecewise sinusoids.

2. Recurrence relation for moments of piecewise *D*-finite functions

2.1. Single continuity interval

We start with the case of a *D*-finite function g on a single continuity interval $[a, b]$. The main tools used in the subsequent derivation are the discrete difference calculus and the Lagrange identity for a differential operator and its adjoint. Let us briefly introduce these tools.

Let $s : \mathbb{N} \to \mathbb{R}$ be a discrete sequence. The *discrete shift operator* E is defined by $\mathrm{E}\, s(n) = s(n + 1)$, and the forward difference operator by $\Delta \overset{\text{def}}{=} \mathrm{E} - \mathrm{I}$. We shall express recurrence relations in terms of polynomials in E or Δ. For example, the Fibonacci sequence F_k satisfies $\mathrm{E}^2 F_k = \mathrm{E} F_k + F_k$, so the operator $P(\mathrm{E}) \overset{\text{def}}{=} \mathrm{E}^2 - \mathrm{E} - \mathrm{I}$ is an annihilating difference operator for F_k. Likewise, the operator Δ annihilates every constant sequence.

Lemma 2.1 [13]**.** *Let $p(\mathrm{E})$ be a polynomial in the shift operator* E *and let $g(n)$ be any discrete function. Then*

$$p(\mathrm{E})(b^n g(n)) = b^n p(b\mathrm{E})g(n).$$

Lemma 2.2 [13]**.** *For any polynomial $p(n)$ of degree k and for $i \geqslant 1$,*

$$\Delta^{k+i} p(n) = 0.$$

Definition 2.3. *Let $x \in \mathbb{R}$ and $k \in \mathbb{N}$. The kth falling factorial[1] of x is*

$$(x)_k \overset{\text{def}}{=} x(x - 1) \cdots (x - k + 1).$$

The following well-known properties of the falling factorial are immediately derived from definition 2.3.

Proposition 2.4. *Let $x \in \mathbb{R}, k \in \mathbb{N}$. Then,*

(1) $(x)_k$ is a polynomial in x of degree k (thus it is also called the factorial polynomial).

(2) If $x = n \in \mathbb{N}$ and $n \geqslant k$, then $(n)_k = \frac{n!}{(n-k)!}$.

(3) If $x = n \in \mathbb{N} \cup \{0\}$ and $n < k$, then $(n)_k = 0$.

---

[1] The Pochhammer symbol $(x)_n$ is also used in the theory of special functions. There it usually represents the rising factorial $(x)_n = x \cdot (x + 1) \cdots (x + n - 1)$.

The formal adjoint of a differential operator $\mathfrak{D} = \sum_{j=0}^{N} p_j(x)\partial^j$ is given by

$$\mathfrak{D}^*\{\cdot\} \overset{\text{def}}{=} \sum_{j=0}^{N} (-1)^j \partial^j \{p_j(x)\cdot\}. \tag{2.1}$$

The operator and its adjoint are connected by the *Lagrange identity* [19]: for every $u, v \in C^N$

$$v\mathfrak{D}(u) - u\mathfrak{D}^*(v) = \frac{\mathrm{d}}{\mathrm{d}x} P_{\mathfrak{D}}(u, v) \tag{2.2}$$

where $P_{\mathfrak{D}}(u, v)$ is the *bilinear concomitant*—a homogeneous bilinear form which may be written explicitly as [19, p 211]

$$\begin{aligned}
P_{\mathfrak{D}}(u, v) &= u\{p_1 v - \partial(p_2 v) + \cdots + (-1)^{N-1}\partial^{N-1}(p_N v)\} \\
&+ u'\{p_2 v - \partial(p_3 v) + \cdots + (-1)^{N-2}\partial^{N-2}(p_N v)\} \\
&+ \cdots \\
&+ u^{(N-1)} p_N v. \tag{2.3}
\end{aligned}$$

If (2.2) is integrated between a and b, *Green's formula* is obtained:

$$\langle \mathfrak{D}u, v \rangle - \langle u, \mathfrak{D}^*v \rangle = [P_{\mathfrak{D}}(u, v)]_a^b, \tag{2.4}$$

where the inner product is defined by $\langle u, v \rangle \overset{\text{def}}{=} \int_a^b uv \, \mathrm{d}x$.

Let $\mathfrak{D}$ and g be arbitrary. Consider the 'differential moments' associated with $\mathfrak{D}$:

$$m_k^{\mathfrak{D}}(g) \overset{\text{def}}{=} m_k(\mathfrak{D}g).$$

By (2.4), we have

$$m_k(\mathfrak{D}g) = \langle \mathfrak{D}g, x^k \rangle = \langle g, \mathfrak{D}^*(x^k) \rangle + [P_{\mathfrak{D}}(g, x^k)]_a^b.$$

Let us define the following two sequences, indexed by k:

$$\mu_k = \mu_k(\mathfrak{D}, g) \overset{\text{def}}{=} \langle g, \mathfrak{D}^*(x^k) \rangle,$$

$$\varepsilon_k = \varepsilon_k(\mathfrak{D}, g) \overset{\text{def}}{=} [P_{\mathfrak{D}}(g, x^k)]_a^b.$$

The sequence of the differential moments is therefore the sum of the two sequences:

$$m_k(\mathfrak{D}g) = \mu_k(\mathfrak{D}, g) + \varepsilon_k(\mathfrak{D}, g). \tag{2.5}$$

Using (1.2) and (2.1) we have

$$\begin{aligned}
\mu_k = \langle g, \mathfrak{D}^*(x^k) \rangle &= \left\langle g, \sum_{j=0}^{N} (-1)^j \frac{\mathrm{d}^j}{\mathrm{d}x^j}(p_j(x)x^k) \right\rangle \\
&= \sum_{i=0}^{k_j} \sum_{j=0}^{N} a_{i,j}(-1)^j (i+k)_j m_{i+k-j}(g) \overset{\text{def}}{=} \Theta_{\mathfrak{D}}(k, \mathrm{E}) m_k, \tag{2.6}
\end{aligned}$$

where

$$\Theta = \Theta_{\mathfrak{D}}(k, \mathrm{E}) \overset{\text{def}}{=} \sum_{i,j} a_{i,j} \Pi^{(i,j)}(k, \mathrm{E}), \qquad \Pi^{(i,j)}(k, \mathrm{E}) \overset{\text{def}}{=} (-1)^j (i+k)_j \mathrm{E}^{i-j}. \tag{2.7}$$

Now let us return to our main problem. Recall that g is D-finite, so let $\mathfrak{D}g \equiv 0$. (2.5) combined with (2.6) gives $\Theta m_k + \varepsilon_k = 0$. As we demonstrate below, there exists a discrete difference operator $\mathcal{E} = \mathcal{E}(\mathrm{E})$ such that $\mathcal{E}\varepsilon_k \equiv 0$. Multiplying the last equation by this $\mathcal{E}$ from the left (multiplication being composition of difference operators) gives us the desired recurrence relation: $\mathcal{E} \cdot \Theta m_k = 0$.

The sequence ε_k is related to the behaviour of g at the endpoints of the interval $[a, b]$. The following lemma unravels its structure.

Lemma 2.5. *Let $\mathfrak{D}$ be of degree N as in (1.2). Then there exist polynomials $q_a(k)$ and $q_b(k)$ of degree at most $N - 1$ such that*

$$\varepsilon_k(\mathfrak{D}, g) = b^k q_b(k) - a^k q_a(k), \qquad k = 0, 1, \ldots \tag{2.8}$$

Proof. Write $\mathfrak{D} = \sum_{j=0}^{N} \mathfrak{D}_j$, where $\mathfrak{D}_j = p_j(x)\partial^j$. Denote $\varepsilon_{k,j} \stackrel{\text{def}}{=} \varepsilon_k(\mathfrak{D}_j, g)$. By (2.3) we have $\varepsilon_k = \sum_{j=0}^{N} \varepsilon_{k,j}$, where

$$\varepsilon_{k,j} = P_{\mathfrak{D}_j}(g, x^k) = \left\{ g^{(j-1)} x^k p_j - g^{(j-2)} \partial(x^k p_j) + \cdots + (-1)^{j-1} g \partial^{j-1}(x^k p_j) \right\}\Big|_{x=a}^{b}.$$

By the Leibniz rule, we have

$$\partial^i(x^k p_j(x)) = \sum_{l=0}^{i} \binom{i}{l} (x^k)^{(l)} p_j^{(i-l)}(x) = \sum_{l=0}^{i} \binom{i}{l} p_j^{(i-l)}(x)(k)_l x^{k-l} = x^k \sum_{l=0}^{i} (k)_l r_{i,j,l}(x),$$

where $r_{i,j,l}(x) = x^{-l} \binom{i}{l} p_j^{(i-l)}(x)$ is a rational function. Now

$$\varepsilon_{k,j} = \left\{ x^k \sum_{i=0}^{j-1} (-1)^i g^{(j-1-i)}(x) \sum_{l=0}^{i} (k)_l r_{i,j,l}(x) \right\}\Bigg|_{x=a}^{b} = \left\{ x^k \sum_{i=0}^{j-1} \sum_{l=0}^{i} (k)_l s_{i,j,l}(x) \right\}\Bigg|_{x=a}^{b}$$

$$= b^k q_{b,j}(k) - a^k q_{a,j}(k),$$

where $s_{i,j,l}(x) = (-1)^i g^{(j-1-i)}(x) r_{i,j,l}(x)$ and $q_{\alpha,j}(k) = \sum_{i=0}^{j-1} \sum_{l=0}^{i} (k)_l s_{i,j,l}(\alpha)$ for $\alpha \in \{a, b\}$. $q_{\alpha,j}(k)$ is a polynomial in k of degree at most $j - 1$ (see definition 2.3). Now

$$\varepsilon_k = \sum_{j=0}^{N} \varepsilon_{k,j} = b^k \sum_{j=0}^{N} q_{b,j}(k) - a^k \sum_{j=0}^{N} q_{a,j}(k).$$

Take $q_a(k) \stackrel{\text{def}}{=} \sum_{j=0}^{N} q_{a,j}(k)$ and $q_b(k) \stackrel{\text{def}}{=} \sum_{j=0}^{N} q_{b,j}(k)$. Since $\deg q_{a,j}, q_{b,j} < j$, we have $\deg q_a, q_b < N$ and this completes the proof. $\square$

As a side remark, we have the following simple condition for the sequence $\{\varepsilon_k\}$ to be a nonzero sequence.

Theorem 2.6. *Assume $\mathfrak{D} g \equiv 0$ and $p_N(x) \neq 0$ on $[a, b]$. Then $\varepsilon_k(\mathfrak{D}, g) \equiv 0$ if and only if $g \equiv 0$.*

Proof. From the proof of lemma 2.5 we have, for $\alpha \in \{a, b\}$,

$$q_\alpha(k) = \sum_{j=0}^{N} \sum_{i=0}^{j-1} (-1)^i g^{(j-1-i)}(\alpha) \sum_{l=0}^{i} (k)_l \alpha^{-l} \binom{i}{l} p_j^{(i-l)}(\alpha). \tag{2.9}$$

- In one direction, we have $g(\alpha) = \cdots = g^{(N-1)}(\alpha) = 0$ for $\alpha = a, b$. By direct substitution we obtain $q_\alpha(k) \equiv 0$.
- To prove the other direction, assume $b^k q_b(k) \equiv a^k q_a(k) = C$. Consider the following cases:
  (1) $C \neq 0$, i.e. $a, b \neq 0$ and $q_a(k), q_b(k) \not\equiv 0$. Then $\left(\frac{b}{a}\right)^k = \frac{q_a(k)}{q_b(k)}$ for all $k \in \mathbb{N}$. But this is impossible because the left-hand side is of exponential growth and the right-hand side is of at most polynomial growth.

(2) $C = 0$. Then at least one of $q_a(k), q_b(k)$ must be identically zero. So let $q_a(k) \equiv 0$ and $a \neq 0$. The coefficient of the highest-order term of q_a corresponds to $j = N, i = l = N - 1$ and equals

$$(-1)^{N-1} g(a) a^{-(N-1)} p_N(a).$$

This is zero only if $g(a) = 0$. So we can lower the limit of the second summation in (2.9) to $n - 2$. Repeating this argument with $j = N, i = l = N - 2$ shows that $g'(a) = 0$. Finally, we obtain that $g(a) = \cdots = g^{(N-1)}(a) = 0$, and by the uniqueness theorem for linear ODEs we conclude $g \equiv 0$. □

An immediate corollary is that generically we have $\mathcal{E} \neq \mathrm{I}$ (*generically* here means $g \not\equiv 0$).

With the structure of the sequence $\{\varepsilon_k\}$ at hand, we can now explicitly construct an annihilating difference operator for it.

Definition 2.7. *Given j, a and b, let $\mathcal{E}_{a,b}^j(\mathrm{E}) \overset{\text{def}}{=} (\mathrm{E} - a\mathrm{I})^j (\mathrm{E} - b\mathrm{I})^j$.*

Theorem 2.8. *Let $\mathfrak{D} = \sum_{j=0}^N p_j(x) \partial^j$. Then the difference operator $\mathcal{E} = \mathcal{E}_{a,b}^N$ annihilates the sequence $\varepsilon_k(\mathfrak{D}, g)$.*

Proof. By lemma 2.5, $\varepsilon_k = b^k q_b(k) - a^k q_a(k)$, where $\deg q_a, q_b \leqslant N - 1$. The factors $(\mathrm{E} - a\mathrm{I})$ and $(\mathrm{E} - b\mathrm{I})$ commute, so it would be sufficient to show that $(\mathrm{E} - a\mathrm{I})^N \{a^k q_a(k)\} \equiv 0$.
Let $p(\mathrm{E}) = (\mathrm{E} - a\mathrm{I})^N$, then $p(a\mathrm{E}) = a^N \Delta^N$. By lemma 2.1,

$$p(\mathrm{E})\{a^k q_a(k)\} = a^{k+N} \Delta^N q_a(k).$$

Since $\deg q_a \leqslant N - 1$, by lemma 2.2, we have $\Delta^N q_a(k) = 0$. □

Remark 2.1. Theorem 2.8 defines a connection between the moments of the two functions g and $\mathfrak{D}g$: $\mathcal{E}m_k(\mathfrak{D}g) = \mathcal{E}\Theta m_k(g)$. This implies a connection between their moment-generating functions as well (see further section 3.1 and also [21]).

Remark 2.2. The polynomial nature of the coefficients $p_j(x)$ of $\mathfrak{D}$ has not been used in the proof of theorem 2.8. Therefore it is true for every linear operator $\mathfrak{D}$ with sufficiently smooth coefficients $p_j(x)$.

Now we have all the necessary information in order to prove the main result of this section.

Theorem 2.9. *Assume that $\mathfrak{D}g \equiv 0$ on $[a, b]$. Then the sequence $\{m_k(g)\}$ satisfies the following recurrence relation:*

$$\mathfrak{S}m_k \overset{\text{def}}{=} \left(\mathcal{E}_{a,b}^N \cdot \Theta_{\mathfrak{D}} \right) m_k = \left((\mathrm{E} - a\mathrm{I})^N (\mathrm{E} - b\mathrm{I})^N \sum_{j=0}^N \sum_{i=0}^{k_j} a_{i,j} (-1)^j (i+k)_j \mathrm{E}^{i-j} \right) m_k = 0.$$

(2.10)

Proof. We have $m_k(\mathfrak{D}g) \equiv 0$. The proof is completed using (2.5), (2.6) and theorem 2.8. □

Remark 2.3. With respect to $\{m_k\}$, the length of the recurrence relation (2.10) is at most $3N + \max k_j + 1$. Its coefficients are linear in $a_{i,j}$ and polynomial in k.

Remark 2.4. The recurrence relation (2.10) is not trivial, i.e. $\mathfrak{S} \neq 0$. To see this, let $\mathfrak{D}g \equiv 0$. Now if $g \not\equiv 0$, then by theorem 2.6, $\varepsilon_k \not\equiv 0$ and thus $\mu_k = \Theta m_k \not\equiv 0$. It follows that Θ cannot be the identical zero operator, and therefore $\mathfrak{S} = \mathcal{E}\Theta \neq 0$.

We conclude this section with some examples which demonstrate the usefulness of theorem 2.9.

Example 2.1. In [6], the authors provide explicit recurrence relations satisfied by the moments of the powers of the *modified Bessel function* $f(x) = K_0(x)$. The method used to obtain these recurrences is integration by parts. However, an additional condition is imposed—namely, it is required that the integrals $\int_\Gamma x^k (f(x)^n)^{(j)} \, dx$ converge and the limits of the integrands coincide at the endpoints of Γ. In our setting, this is equivalent to putting $\mathcal{E} = I$.

The function $g = K_0^2(x)$ is annihilated by the operator $\mathfrak{D} = x^2 \partial^3 + 3x \partial^2 + (1 - 4x^2)\partial - 4xI$ [6, example 2]. The only nonzero coefficients are therefore

$a_{i,j}$	$i = 0$	$i = 1$	$i = 2$
$j = 0$		-4	
$j = 1$	1		-4
$j = 2$		3	
$j = 3$			1

By (2.10), we have

$$-4m_{k+1} - km_{k-1} + 4(k+2)m_{k+1} + 3k(k+1)m_{k-1} - (k+2)(k+1)km_{k-1} = 0,$$

which is just $4(k+1)m_{k+1} = k^3 m_{k-1}$. This result agrees with [6, example 3].

Example 2.2 (example 1.1 continued). $f = \alpha \, e^{\beta x}$ is annihilated by the operator $\mathfrak{D} = \partial - \beta I$. Thus $a_{0,0} = -\beta$ and $a_{1,0} = 1$. By (2.10) the sequence $\{m_k\}$ on $[0, 1]$ satisfies

$$\mathfrak{S} = E(E - I)(-\beta I - kE^{-1})m_k = -\beta m_{k+2} + (\beta - (k+2))m_{k+1} + (k+1)m_k = 0.$$

This recurrence relation with polynomial coefficients may be solved explicitly using computer algebra tools (see [34] for an overview of the existing algorithms). Using the maxima computer algebra system [28], the following explicit formula was obtained:

$$m_k = \alpha \int_0^1 x^k \, e^{\beta x} \, dx = \frac{(-1)^k k!}{\beta^k} \left\{ c_1 + c_2 \sum_{j=0}^{k-1} \frac{\beta^j (-1)^{j-1}}{(j+1)!} \right\}$$

$$c_1 = m_0$$

$$c_2 = \beta m_1 + m_0$$

This example is further continued in example 3.1.

2.2. Piecewise case

Until now we have been considering a function g which satisfies $\mathfrak{D}g \equiv 0$ on a *single interval* $[a, b]$. In particular, we have seen that the sequence of the moments of g satisfies a linear recurrence relation whose coefficients linearly depend on the coefficients of $\mathfrak{D}$. Now we are going to consider the piecewise case (depicted in figure 1) where g is assumed to consist of several 'pieces' $g_0, g_1, \ldots, g_{\mathcal{K}}$. On each continuity interval $\Delta_n = [\xi_n, \xi_{n+1}]$, the nth piece of g satisfies $\mathfrak{D}_n g_n \equiv 0$, where $\mathfrak{D}_n = \sum_{j=0}^{N_n} \sum_{i=0}^{k_n} (a_{i,j,n} x^i) \partial^j$. Denote $m_{k,n} \stackrel{\text{def}}{=} \int_{\xi_n}^{\xi_{n+1}} x^k g_n(x) \, dx$. Our goal is to find a recurrence relation satisfied by the sequence $m_k = \sum_{n=0}^{\mathcal{K}} m_{k,n}$.

As we shall see below, this recurrence relation is particularly easy to find explicitly in the case $g \in A_{\mathfrak{D}}^*$, i.e. $\mathfrak{D}_n = \mathfrak{D}$ for $n = 0, \ldots, \mathcal{K}$. We shall also discuss the general case where $\mathfrak{D}_m \neq \mathfrak{D}_n$ briefly.
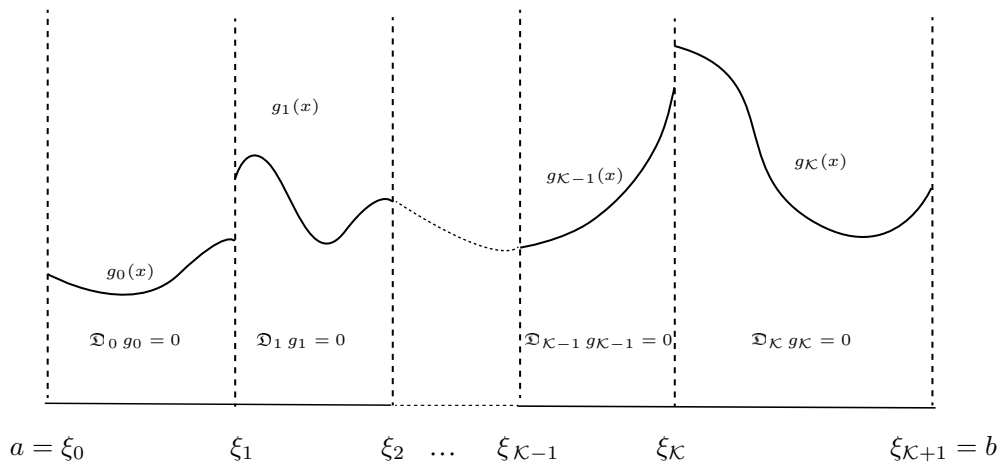
Figure 1. A piecewise *D*-finite function.

2.2.1. The same operator on each interval. We present two methods for computing the desired recurrence relation. Then we show that these methods indeed produce the same result.

Theorem 2.10 (Method I). *Let $g \in A_{\mathfrak{D}}^*$ with $\mathfrak{D}_m = \mathfrak{D}$ for $0 \leqslant m \leqslant \mathcal{K}$. Then $\{m_k(g)\}$ satisfies*

$$\left(\prod_{n=0}^{\mathcal{K}+1}(E - \xi_n I)^N \Theta_{\mathfrak{D}}(k, E)\right) m_k = 0. \tag{2.11}$$

Proof. By theorem 2.9, for every $n = 0, \ldots, \mathcal{K}$ the sequence $\{m_{k,n}\}$ satisfies

$$\left(\mathcal{E}_{\xi_n,\xi_{n+1}}^N(E) \cdot \Theta_{\mathfrak{D}}(k, E)\right) m_{k,n} = 0. \tag{2.12}$$

(2.11) follows by using the fact that the linear factors $(E - \xi_i I)$ and $(E - \xi_j I)$ commute. $\square$

It turns out that formula (2.11) may be obtained by considering the *piecewise* function $g \in A_{\mathfrak{D}}^*$ being annihilated *as a distribution on $[a, b]$* by some operator $\widehat{\mathfrak{D}} = \widehat{\mathfrak{D}}(\mathfrak{D}, \{\xi_n\})$. We now derive the explicit expression for $\widehat{\mathfrak{D}}$.

The general theory of distributions (generalized functions) may be found in [16]. By a *test function* we shall mean any $f \in C^{N-1}([a, b])$ (in fact, for our purposes it is sufficient to consider just the monomials $f = x^k$).

We shall identify the discontinuous function g with the following distribution (here $g_{-1} \equiv 0$ by definition):

$$g(x) = \widetilde{g}_0 + \sum_{n=1}^{\mathcal{K}} \widetilde{g}_n(x)\mathcal{H}(x - \xi_n) \tag{2.13a}$$

$$\widetilde{g}_n \stackrel{\text{def}}{=} g_n - g_{n-1} \tag{2.13b}$$

$$\mathcal{H}(x) \stackrel{\text{def}}{=} \begin{cases} 0 & x < 0 \\ 1 & x \geqslant 0. \end{cases} \tag{2.13c}$$

The functions $\widetilde{g}_n(x)$ belong to the solution space of $\mathfrak{D}$ on Δ_n, and thus $\widetilde{g}_n(x) \in C^{N-1}(\Delta_n)$. Since $p_n(x) \neq 0$ on $[a, b]$, we even have $\widetilde{g}_n(x) \in C^{N-1}([a, b])$.

The derivative of $\mathcal{H}$ is the Dirac δ. The distribution $\delta(x)$ has the following properties:

$$\langle \delta(x - t), f(x) \rangle = f(t),$$
$$\langle \delta^{(j)}(x - t), f(x) \rangle = (-1)^j f^{(j)}(t).$$

The second equality is valid provided $f(x)$ is j-times differentiable at $x = t$.

Now let $g(x) = c \cdot \delta(x - \xi)$ for $c, \xi \in \mathbb{R}$. This distribution is 'annihilated' by the operator $\mathfrak{D} = (x - \xi)\mathrm{I}$ in the sense that $\langle \mathfrak{D}g, f \rangle = \langle c \cdot (x - \xi)\delta(x - \xi), f \rangle = 0$ for every test function f. By linearity, the operator $\mathfrak{D} = \prod_{i=1}^{\mathcal{K}}(x - \xi_i)\mathrm{I}$ annihilates every g of the form $g = \sum_{i=1}^{\mathcal{K}} c_i \delta(x - \xi_i)$.

Lemma 2.11. *Let* $g = \sum_{i=1}^{\mathcal{K}} \sum_{j=0}^{S-1} c_{ij}(x)\delta^{(j)}(x - \xi_i)$ *such that* $c_{ij}(x)$ *is* j-*times differentiable at* ξ_i. *Then* g *is annihilated by* $\mathfrak{D} = p(x)\mathrm{I}$, *where* $p(x) = \prod_{i=1}^{\mathcal{K}}(x - \xi_i)^S$.

Proof. For every test function $f(x)$ we have

$$\langle f, \mathfrak{D}g \rangle = \langle p(x)f(x), g \rangle = \sum_{i=1}^{\mathcal{K}} \sum_{j=0}^{S-1} \langle c_{ij}(x)p(x)f(x), \delta^{(j)}(x - \xi_i) \rangle.$$

The function $r(x) = c_{ij}(x)p(x)f(x)$ has a zero of order S at $x = \xi_i$. Therefore, all its derivatives up to j vanish at ξ_i and so $\langle f, \mathfrak{D}g \rangle = 0$. $\square$

Theorem 2.12. *Assume* $g \in A_{\mathfrak{D}}^*$ *as in (2.13), with* $\mathfrak{D}$ *of degree* N *annihilating every piece of* g. *Then the entire* g *is annihilated as a distribution by the operator* $\widehat{\mathfrak{D}} = \prod_{i=1}^{\mathcal{K}}(x - \xi_i)^N \mathfrak{D}$.

Proof. By (2.13) and the fact that $\mathfrak{D}\widetilde{g}_n \equiv 0$ we obtain

$$\mathfrak{D}g = \mathfrak{D}\widetilde{g}_0 + \sum_{j=0}^{N} p_j(x)\partial^j \left\{ \sum_{i=1}^{\mathcal{K}} \widetilde{g}_i(x)\mathcal{H}(x - \xi_i) \right\}$$

$$= 0 + \sum_{i=1}^{\mathcal{K}} \sum_{j=0}^{N} p_j(x) \sum_{k=0}^{j} \binom{j}{k} \partial^{j-k}\{\widetilde{g}_i(x)\}\partial^k\{\mathcal{H}(x - \xi_i)\}$$

$$= \sum_{i=1}^{\mathcal{K}} \sum_{j=0}^{N} \sum_{k=1}^{j} p_j(x)\binom{j}{k}\partial^{j-k}\{\widetilde{g}_i(x)\}\partial^k\{\mathcal{H}(x - \xi_i)\} + \sum_{i=1}^{\mathcal{K}} \mathcal{H}(x - \xi_i) \sum_{j=0}^{N} p_j(x)\partial^j \widetilde{g}_i(x)$$

$$= \sum_{i=1}^{\mathcal{K}} \sum_{k=0}^{N-1} h_{ik}(x)\delta^{(k)}(x - \xi_i) + \sum_{i=1}^{\mathcal{K}} \mathcal{H}(x - \xi_i)\mathfrak{D}\widetilde{g}_i = \sum_{i=1}^{\mathcal{K}} \sum_{k=0}^{N-1} h_{ik}(x)\delta^{(k)}(x - \xi_i),$$

where $h_{ik}(x)$ is k-times differentiable at ξ_i. Then apply lemma 2.11. $\square$

Theorem 2.12 will later serve as the basis for recovering the locations of the discontinuities of $g \in A_{\mathfrak{D}}^*$. The factor $\prod(x - \xi_i)^N$ effectively 'encodes' the positions of the jumps into the operator itself. The 'decoding' will then simply be to find these 'extra' roots, once the 'enlarged' operator is reconstructed.

The second form of the recurrence relation now immediately follows from theorem 2.12.

Theorem 2.13 (Method II)**.** *The sequence of the moments of* $g \in A_{\mathfrak{D}}^*$ *satisfies the recurrence relation*

$$\left(\mathcal{E}_{a,b}^N(\mathrm{E}) \cdot \Theta_{\widehat{\mathfrak{D}}}(k, \mathrm{E}) \right) m_k = 0, \tag{2.14}$$

where $\widehat{\mathfrak{D}} = \prod_{n=1}^{\mathcal{K}}(x - \xi_n)^N \mathfrak{D}$.

Proof. Theorem 2.9 combined with theorem 2.12.   □

Lemma 2.14 (Equivalence of methods I and II). *Let $q(x)$ be an arbitrary polynomial. Then*

$$\Theta_{q(x)\mathfrak{D}}(k, \mathrm{E}) = q(\mathrm{E})\Theta_{\mathfrak{D}}(k, \mathrm{E}).$$

Equivalence of (2.11) and (2.14) then follows from lemma 2.14 by putting $q(x) = \prod_{i=1}^{\mathcal{K}}(x - \xi_i)^N$.

To prove lemma 2.14 we need the following result.

Proposition 2.15. *Let $p(x) = \sum_{i=0}^{\alpha} p_i x^i$, $q(x) = \sum_{i=0}^{\beta} q_i x^i$ and $r(x) = p(x)q(x) = \sum_{i=0}^{\alpha+\beta} r_i x^i$. Let E be the shift operator in k and let $\widehat{j}$ be fixed. Then*

$$\sum_{i=0}^{\alpha+\beta} r_i(i + k)_{\widehat{j}}\mathrm{E}^{i-\widehat{j}} = p(\mathrm{E})\sum_{i=0}^{\beta} q_i(i + k)_{\widehat{j}}\mathrm{E}^{i-\widehat{j}}$$

Proof. We extend the sequences of the coefficients $\{p_i\}$ and $\{q_i\}$ by zeros as necessary. By the rule of polynomial multiplication, $r_i = \sum_{j=0}^{\alpha} p_j q_{i-j}$. Now

$$\sum_{i=0}^{\alpha+\beta} r_i(i + k)_{\widehat{j}}\mathrm{E}^{i-\widehat{j}} = \sum_{i=0}^{\alpha+\beta}\sum_{j=0}^{\alpha} p_j q_{i-j}(i - j + k + j)_{\widehat{j}}\mathrm{E}^{i-j+\widehat{j}+j}$$

$$= \sum_{j=0}^{\alpha} p_j\mathrm{E}^j \sum_{i=0}^{\alpha+\beta} q_{i-j}((i - j) + k)_{\widehat{j}}\mathrm{E}^{(i-j)+\widehat{j}}$$

$$[i - j \to i] = \sum_{j=0}^{\alpha} p_j\mathrm{E}^j \sum_{i=-j}^{\alpha+\beta-j} q_i(i + k)_{\widehat{j}}\mathrm{E}^{i-\widehat{j}} = p(\mathrm{E})\sum_{i=0}^{\beta} q_i(i + k)_{\widehat{j}}\mathrm{E}^{i-\widehat{j}}$$
                                                                                                                                  □

Proof of Lemma 2.14. Recall that $\Theta_{\mathfrak{D}} = \sum_{i,j} a_{i,j}\Pi^{(i,j)}(k, \mathrm{E})$ where $\Pi^{(i,j)}(k, \mathrm{E}) = (-1)^j(i + k)_j\mathrm{E}^{i-j}$, $\mathfrak{D} = p_j(x)\partial^j$ and $p_j(x) = \sum_{i=0}^{k_j} a_{i,j}x^i$. Now let $\widehat{p}_j(x) = q(x)p_j(x) = \sum \widehat{a}_{i,j}x^i$, then

$$\Theta_{q(x)\mathfrak{D}}(k, \mathrm{E}) = \sum_{j=0}^{N}(-1)^j \sum_{i=0}^{k_j+\deg q} \widehat{a}_{i,j}(i + k)_j\mathrm{E}^{i-j}$$

$$[\textit{Proposition 2.15}] = \sum_{j=0}^{N}(-1)^j q(\mathrm{E}) \sum_{i=0}^{k_j} a_{i,j}(i + k)_j\mathrm{E}^{i-j}$$

$$= q(\mathrm{E})\Theta_{\mathfrak{D}}(k, \mathrm{E}).$$
                                                                                                                                  □

2.2.2. Different operators on each interval. Recall that we want to find a recurrence relation for the sequence $m_k = \sum m_{k,n}$, where each subsequence is annihilated by the difference operator

$$\mathfrak{S}_n = \mathcal{E}_{\xi_n,\xi_{n+1}}^{N_n} \cdot \Theta_{\mathfrak{D}_n}(k, \mathrm{E}).$$

There exist at least two approaches, both of which involve techniques from the theory of non-commutative polynomials—the so-called Ore polynomial rings (see [26]). Both differential and difference operators with polynomial coefficients are members of the appropriate Ore algebra. The *least common left multiple* (LCLM) of two polynomials p, q is the unique

polynomial r of minimal degree such that both p and q are right-hand factors of r, i.e. $r = p'p = q'q$ for some polynomials p', q'. The LCLM may be explicitly found by the non-commutative version of the polynomial division algorithm. The complete theory may be found in [26].

Approach I Given the operators $\mathfrak{S}_n$ such that $\mathfrak{S}_n m_{k,n} \equiv 0$, the operator $\mathfrak{S}^{\flat}$ which annihilates the sum $\sum_n m_{k,n}$ is given by the least common left multiple of $\{\mathfrak{S}_n\}$.

Approach II Given the operators $\mathfrak{D}_n$ which annihilate the pieces g_n separately, the operator $\mathfrak{D}^{\dagger}$ which annihilates every piece simultaneously is the least common left multiple of $\{\mathfrak{D}_n\}$. Then the annihilating operator for $\{m_k\}$ is $\mathfrak{S}^{\dagger} = \mathcal{E}_{a,b}^{N^{\dagger}} \Theta_{\mathfrak{D}^{\dagger}}(k, \mathrm{E})$.

We are not aware of any general procedure by which the coefficients of $\mathfrak{S}^{\dagger}$ or $\mathfrak{S}^{\flat}$ may be related to the coefficients of each $\mathfrak{D}_n$ in some tractable manner, unless the operators $\mathfrak{D}_n$ commute.

Example 2.3 (Piecewise sinusoids).   Let g consist of two sinusoid pieces: $g_1(x) = c_1 \sin(\omega_1 x + \phi_1)$ and $g_2(x) = c_2 \sin(\omega_2 x + \phi_2)$ with break point ξ. The annihilating operators for g_1 and g_2 are, respectively, $\mathfrak{D}_1 = \partial^2 + \omega_1^2 \mathrm{I}$ and $\mathfrak{D}_2 = \partial^2 + \omega_2^2 \mathrm{I}$. The operator $\mathfrak{D}^{\dagger} = \left(\partial^2 + \omega_1^2 \mathrm{I}\right) \cdot \left(\partial^2 + \omega_2^2 \mathrm{I}\right)$ annihilates both pieces simultaneously, therefore $(x - \xi)^4 \mathfrak{D}^{\dagger}$ annihilates the entire g.

3. Moment inversion

We now present our method for moment inversion for piecewise D-finite functions. Recall that our purpose is to reconstruct the parameters

$$\mathcal{P} \stackrel{\mathrm{def}}{=} \{\{a_{i,j}\}, \{\xi_i\}, \{\alpha_{i,n}\}\} \tag{3.1}$$

from the input

$$\mathcal{I} \stackrel{\mathrm{def}}{=} \{\{m_k\}, N, \{k_j\}, \mathcal{K}, a, b\}. \tag{3.2}$$

First, we establish explicit connections between $\mathcal{P}$ and $\mathcal{I}$—the 'forward mapping' $\mathcal{M} : \mathcal{P} \to \mathcal{I}$. Then we derive the inverse mapping $\mathcal{N} = \mathcal{M}^{-1}$ and provide simple conditions for the solvability of the resulting inverse systems.

3.1. Forward equations

Recall the polynomials $\Pi^{(i,j)}(k, \mathrm{E})$ which were defined in (2.7) during the derivation of the recurrence relation (2.10). Given a multi-index (i, j) we define the 'shifted' moment sequence

$$v_k^{(i,j)} \stackrel{\mathrm{def}}{=} \left(\mathcal{E}_{a,b}^N(\mathrm{E}) \cdot \Pi^{(i,j)}(k, \mathrm{E})\right) m_k. \tag{3.3}$$

For each $j = 0, \ldots, N$, let $h_j(z)$ be the formal power series

$$h_j(z) \stackrel{\mathrm{def}}{=} \sum_{k=0}^{\infty} v_k^{(0,j)} z^k.$$

Finally, for any $g(x)$ let

$$\mathfrak{g}_j^g(x) \stackrel{\mathrm{def}}{=} \mathcal{E}_{a,b}^N(x) \frac{\mathrm{d}^j}{\mathrm{d}x^j} g(x)$$

$$I_g(z) \stackrel{\mathrm{def}}{=} \sum_{k=0}^{\infty} m_k(g) z^k.$$

$I_g(z)$ is called the *moment-generating function* of g.

Theorem 3.1. *Let $g \in A_{\mathfrak{D}}^*$ be annihilated by $\mathfrak{D}$ (either in the usual sense if $\mathcal{K} = 0$ or as a distribution if $\mathcal{K} > 0$, in which case $\mathfrak{D} = \prod_{i=1}^{\mathcal{K}} (x - \xi_i)^N \mathfrak{D}^\dagger$ with $\mathfrak{D}^\dagger$ annihilating every piece), where $\mathfrak{D} = \sum_{j=0}^{N} p_j(x) \partial^j$ and $p_j(x) = \sum_{i=0}^{k_j} a_{i,j} x^i$. Then,*

(A) The vector $\mathbf{a} = (a_{i,j})$ satisfies a linear homogeneous system

$$
H\mathbf{a} = \begin{pmatrix} v_0^{(0,0)} & v_0^{(1,0)} & \cdots & v_0^{(k_N,N)} \\ v_1^{(0,0)} & v_1^{(1,0)} & \cdots & v_1^{(k_N,N)} \\ \vdots & \vdots & \vdots & \vdots \\ v_{\widehat{M}}^{(0,0)} & v_{\widehat{M}}^{(1,0)} & \cdots & v_{\widehat{M}}^{(k_N,N)} \end{pmatrix} \begin{pmatrix} a_{0,0} \\ a_{1,0} \\ \vdots \\ a_{k_N,N} \end{pmatrix} = 0 \tag{3.4}
$$

for all $\widehat{M} \in \mathbb{N}$.

(B) $v_k^{(i,j)} = m_{i+k}(\mathfrak{g}_j^g(x))$ for all $0 \leqslant j \leqslant N, 0 \leqslant i \leqslant k_j$ and $k \in \mathbb{N}$ (the moments are taken in $[a, b]$). Consequently, $h_j(z)$ is the moment-generating function of $\mathfrak{g}_j^g(x)$.

(C) The functions $\{1, h_0(z), \ldots, h_N(z)\}$ are polynomially dependent:

$$
q(z) + \sum_{j=0}^{N} h_j(z) p_j^*(z) = 0,
$$

where $p_j^*(z) = z^{\max k_j} p_j(z^{-1})$ and $q(z)$ is a polynomial with $\deg q < \max k_j$.

Proof of A. By theorem 2.13 and (3.3) we have

$$
\sum_{j=0}^{N} \sum_{i=0}^{k_j} a_{i,j} v_k^{(i,j)} = 0, \qquad k = 0, 1, \ldots.
$$

This is exactly (3.4). $\qquad\qquad\square$

Proposition 3.2. *Let $p(x)$ be a polynomial in x. Then for every $f(x)$*

$$
m_k(p(x)f(x)) = p(\mathrm{E})m_k(f(x)).
$$

Proof. Let $p(x) = x^r$. Then

$$
m_k(x^r f(x)) = \int_a^b x^k x^r f(x)\, \mathrm{d}x = \int_a^b x^{k+r} f(x)\, \mathrm{d}x = m_{k+r}(f(x)) = \mathrm{E}^r m_k(f(x)).
$$

The proof for an arbitrary polynomial $p(x)$ follows by linearity. $\qquad\qquad\square$

Proof of B. Fix $j \leqslant N, i \leqslant \max k_j$ and define $\mathfrak{D}_{ij} = x^i \partial^j$. By (2.5) and (2.6) we have

$$
m_k(\mathfrak{D}_{ij} g) = \Theta_{\mathfrak{D}_{ij}}(k, \mathrm{E}) m_k(g) + \varepsilon_k(\mathfrak{D}_{ij}, g). \tag{3.5}
$$

By (2.7), we have $\Theta_{\mathfrak{D}_{ij}}(k, \mathrm{E}) = \Pi^{(i,j)}(k, \mathrm{E})$. By theorem 2.8, we have $\mathcal{E}_{a,b}^N(\mathrm{E})\varepsilon_k(\mathfrak{D}_{ij}, g) = 0$. Finally,

$$
\begin{aligned}
v_k^{(i,j)} &= \left(\mathcal{E}_{a,b}^N(\mathrm{E}) \cdot \Pi^{(i,j)}(k, \mathrm{E})\right) m_k(g) = \mathcal{E}_{a,b}^N(\mathrm{E}) \Theta_{\mathfrak{D}_{ij}}(k, \mathrm{E}) m_k \\
[3.5] &= \mathcal{E}_{a,b}^N(\mathrm{E}) m_k(\mathfrak{D}_{ij} g) - \mathcal{E}_{a,b}^N(\mathrm{E}) \varepsilon_k(\mathfrak{D}_{ij}, g) = \mathcal{E}_{a,b}^N(\mathrm{E}) m_k(\mathfrak{D}_{ij} g) \\
[\text{\textit{Proposition} } 3.2] &= m_k(\mathcal{E}(x) x^i \partial^j g) = m_{i+k}(\mathfrak{g}_j^g(x)). \qquad\qquad\square
\end{aligned}
$$

Proof of C. Let $k^* = \max k_j$. We have $p_j^*(z) = \sum_{i=0}^{k_j} a_{i,j} z^{k^*-i}$. Denote the power series

$$
q(z) = -\sum_{j=0}^{N} h_j(z) p_j^*(z) = \sum_{k=0}^{\infty} q_k z^k.
$$

An immediate consequence of B is that $v_{k'}^{(i',j)} = v_k^{(i,j)}$ for $i' + k' = i + k$ and all j. Then for all $m \geqslant k^*$

$$-q_m = \sum_{j=0}^{N} \sum_{i=0}^{k_j} a_{i,j} v_{m-k^*+i}^{(0,j)} = \sum_{j=0}^{N} \sum_{i=0}^{k_j} a_{i,j} v_{m-k^*}^{(i,j)} = 0.$$

So $q(z)$ is a polynomial of degree at most $k^* - 1$. This completes the proof of theorem 3.1. $\qquad\square$

Remark 3.1. H has the structure of the *Hankel-striped* matrix $H = [V_0 \ldots V_N]$ where each 'stripe' is a Hankel matrix

$$V_j = \begin{bmatrix} v_0^{(0,j)} & v_0^{(1,j)} & \cdots & v_0^{(k_j,j)} \\ v_1^{(0,j)} & v_1^{(1,j)} & \cdots & v_1^{(k_j,j)} \\ \vdots & \vdots & \vdots & \vdots \\ v_{\widehat{M}}^{(0,j)} & v_{\widehat{M}}^{(1,j)} & \cdots & v_{\widehat{M}}^{(k_j,j)} \end{bmatrix} \tag{3.6}$$

Hankel-striped matrices appear as central objects in contexts such as the *Hermite–Padé* approximation (the standard Padé approximation being its special case), minimal realization problem in control theory and Reed–Solomon codes [7, 20, 23]. In fact, the system of polynomials $\{p_j^*(z)\}$ is called the Padé–Hermite form for $\Phi = \{1, h_0(z), \ldots, h_N(z)\}$.

Remark 3.2. Moment-generating functions are a powerful tool for the investigation of the properties of the sequence m_k. For instance, the asymptotic behaviour of the general term may be derived from the analytic properties of $I_g(z)$ [15].

Now suppose the operator $\mathfrak{D}$ annihilating every piece g_n of $g \in A_{\mathfrak{D}}^*$, as well as the jump points $\{\xi_n\}$, are known. Let $\{u_i\}_{i=1}^N$ be a basis for the space $\mathcal{N}_{\mathfrak{D}}$. Then $g_n(x) = \sum_{i=1}^N \alpha_{in} u_i(x)$. Applying the moment transform to both sides of the last equation and summing over $n = 0, \ldots, \mathcal{K}$ gives

Proposition 3.3. *Denote* $c_{i,k}^n = \int_{\xi_n}^{\xi_{n+1}} x^k u_i(x)$ *for* $n = 0, \ldots, \mathcal{K}$. *Then* $\forall \widetilde{M} \in \mathbb{N}$:

$$\begin{pmatrix} c_{1,0}^0 & \cdots & c_{N,0}^0 & \cdots & c_{N,0}^{\mathcal{K}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{1,\widetilde{M}}^0 & \cdots & c_{N,\widetilde{M}}^0 & \cdots & c_{N,\widetilde{M}}^{\mathcal{K}} \end{pmatrix} \begin{pmatrix} \alpha_{1,0} \\ \vdots \\ \alpha_{N,0} \\ \vdots \\ \alpha_{N,\mathcal{K}} \end{pmatrix} = \begin{pmatrix} m_0 \\ m_1 \\ \vdots \\ m_{\widetilde{M}} \end{pmatrix} \tag{3.7}$$

3.2. The inversion algorithm

The reconstruction algorithm which is based on the results of the previous section is depicted schematically in figure 2. The solvability of the corresponding systems is discussed in section 3.3. Note the following.

(a) At the initial stage, the 'encoding' of the (yet unknown) jump points takes place. In practice, this means the 'enlargement' of $\mathfrak{D}$ to $\widehat{\mathfrak{D}} \stackrel{\text{def}}{=} \prod_{i=1}^{\mathcal{K}} (x - \xi_i)^N \mathfrak{D}$. The parameters of the problem therefore change as follows: N remains the same while $k_j \leftarrow k_j + N\mathcal{K}$—see (2.14).
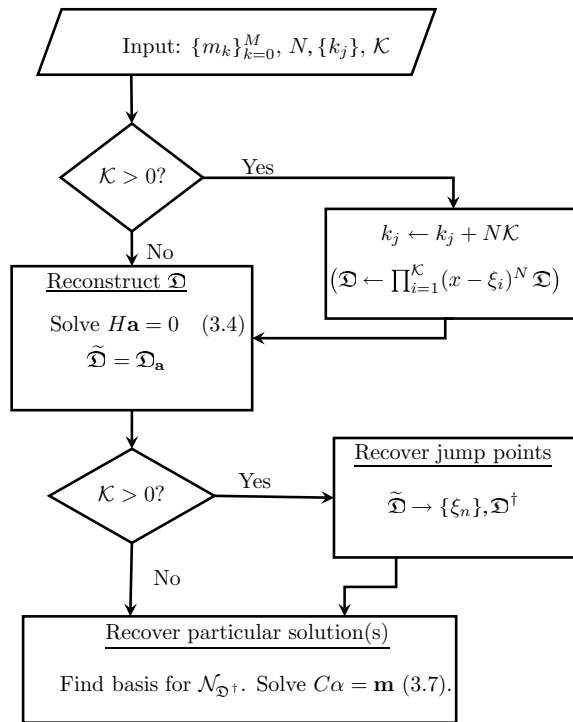
Figure 2. The reconstruction algorithm

(b) By theorem 2.12, $\{\xi_n\}$ are the $\mathcal{K}$ distinct common roots of the polynomials which are the coefficients of $\widetilde{\mathfrak{D}}$ of multiplicity N. The remaining part of the coefficients define the operator $\mathfrak{D}^\dagger$ which annihilates every piece of g.

Example 3.1 (Examples 1.1 and 2.2 continued)**.** $g(x) = \alpha\, e^{\beta x}$ on $[0, 1]$ is annihilated by $\mathfrak{D} = \partial - \beta \mathrm{I}$.

Writing down (3.4) with $\widehat{M} = 0$ yields

$$[m_2 - m_1 \quad -(2m_1 - m_0)] \begin{bmatrix} -\beta \\ 1 \end{bmatrix} = 0,$$

which has the solution

$$\beta = \frac{2m_1 - m_0}{m_1 - m_2}.$$

The constant α is then recovered by

$$\alpha = \frac{m_0(g)}{m_0(e^{\beta x})} = \frac{\beta m_0}{e^\beta - 1}.$$

Note that this solution requires the first three moments instead of two as in (1.4).

Additional examples of complete inversion procedures are elaborated in the appendix.

3.3. Solvability of inverse equations

The constants $\widehat{M}$ and $\widetilde{M}$ determine the minimal size of the corresponding linear systems (3.4) and (3.7) in order for all the solutions of these systems to be also solutions of the original problem.

Theorem 3.4. *If* $\mathbf{b} \in \mathcal{N}_H$, *then*

$$m_k(\mathcal{E}(x)(\mathfrak{D}_\mathbf{b} g)(x)) = 0, \qquad k = 0, 1, \ldots, \widehat{M}.$$

Proof. Denote $\mathbf{b} = (b_{ij})$ and let $k \geqslant 0$. By theorem 3.1, part B, and proposition 3.2, the product of the $(k + 1)$st row of H with $\mathbf{b}$ is

$$
\begin{aligned}
0 &= \sum_{j=0}^{N} \sum_{i=0}^{k_j} b_{ij} v_k^{(i,j)} = \sum_{j=0}^{N} \sum_{i=0}^{k_j} b_{ij} m_{i+k}(\mathfrak{g}_j^g(x)) \\
&= \sum_{j=0}^{N} \sum_{i=0}^{k_j} m_k(b_{ij} x^i \mathfrak{g}_j^g(x)) = m_k \left(\sum_{j=0}^{N} \left(\sum_{i=0}^{k_j} b_{ij} x^i \right) \mathfrak{g}_j^g(x) \right) \\
&= m_k(\mathcal{E}(x)(\mathfrak{D}_\mathbf{b} g)(x)).
\end{aligned}
$$

$\square$

If it is possible to estimate how many moments of $F = \mathcal{E}(x)(\mathfrak{D}_\mathbf{b} g)(x)$ should vanish in order to guarantee the identical vanishing of F and therefore also of $\mathfrak{D}_\mathbf{b} g$ for *all possible differential operators* of the prescribed complexity (i.e. of given order and given degrees of its coefficients), then $\widehat{M}$ may be taken to be this number. Then, every solution of (3.4) will correspond to some annihilating operator. For any specific g and $\mathfrak{D}_\mathbf{b}$, such a finite number exists because any nonzero piecewise-continuous integrable function has at least some nonzero moments. However, this number may be arbitrarily large as shown by the next example.

Example 3.2 (Legendre orthogonal polynomials)**.** It is known that every family of orthogonal polynomials satisfies a differential equation of order $N = 2$ of the following type:

$$\mathfrak{D} = q(x)\partial^2 + p(x)\partial + \lambda_n \mathrm{I}, \tag{3.8}$$

where q, p are fixed polynomials with $\deg q \leqslant 2$ and $\deg p \leqslant 1$, and λ_n is a scalar which is different for each member of the family.

Consider $\{L_n(x)\}$—the family of Legendre orthogonal polynomials. The interval of orthogonality is $[a, b] = [-1, 1]$ and $\mathcal{E} = (\mathrm{E}^2 - \mathrm{I})^2$. L_n is annihilated by $\mathfrak{D}_n = (1 - x^2)\partial^2 - 2x\partial + \lambda_n \mathrm{I}$. Furthermore, $\langle L_n(x), x^k \rangle = 0$ for $k \leqslant n - 1$. The 'reconstruction problem' for $L_n(x)$ (assume that it is normalized) is to find the constant λ_n from the moments (it is well known that $\lambda_n = n(n + 1)$).

Take an arbitrary vector $\mathbf{b} = [b_{00} \ b_{11} \ b_{02} \ b_{22}]^T$ so $\mathfrak{D}_\mathbf{b} = b_{00}\mathrm{I} + b_{11}x\partial + (b_{02} + b_{22}x^2)\partial^2$. The function $\mathcal{E}(x)\mathfrak{D}_\mathbf{b} L_n$ is a polynomial of degree $n + 4$, so it is uniquely determined by its $n + 5$ moments [27]. By theorem 3.4, this would be the minimal size of the system (3.4). The entries of the kth row of H are

$$
\begin{aligned}
v_k^{(0,0)} &= m_{k+4} - 2m_{k+2} + m_k & i = 0, j = 0 \\
v_k^{(1,1)} &= -((k + 5)m_{k+4} - 2(k + 3)m_{k+2} + (k + 1)m_k) & i = 1, j = 1 \\
v_k^{(0,2)} &= (k + 4)(k + 3)m_{k+2} - 2(k + 2)(k + 1)m_k + k(k - 1)m_{k-2} & i = 0, j = 2 \\
v_k^{(2,2)} &= (k + 6)(k + 5)m_{k+4} - 2(k + 4)(k + 3)m_{k+2} + (k + 2)(k + 1)m_k & i = 2, j = 2.
\end{aligned}
$$

The first $n-5$ rows are identically zero because the maximum moment of L_n involved is m_{n-1}. We are looking for a solution of the form $\mathbf{b} = [\lambda_n \quad -2 \quad 1 \quad -1]^T$. The $(n-4)$th row is

$$[m_n \quad -(n+1)m_n \quad 0 \quad (n+2)(n+1)m_n]$$

and that is just enough to reconstruct the operator: $m_n(\lambda_n + 2(n+1) - (n+2)(n+1)) = 0$ and thus $\lambda_n = n(n+1)$, as expected. The reason for the original estimate $n+5$ not being sharp is that some of the coefficients of $\mathfrak{D}$ (in fact, all but one) were known *a priori*.

Similar argument may be used in order to estimate $\widetilde{M}$. Straightforward computation leads to

Proposition 3.5. *If $\alpha = (\alpha_{in})$ is a solution of (3.7), then*

$$m_k(\mathfrak{G}_\alpha(x)) = 0, \qquad k = 0, 1, \ldots, \widetilde{M}, \tag{3.9}$$

where

$$\mathfrak{G}_\alpha(x) \overset{\text{def}}{=} g(x) - \sum_{n=0}^{\mathcal{K}} \sum_{i=1}^{N} \alpha_{in} u_i(x). \tag{3.10}$$

Every function $\mathfrak{G}_\alpha(x)$ is a (piecewise) solution of $\mathfrak{D}f \equiv 0$. By theorem 2.13, the moments of $\mathfrak{G}_\alpha(x)$ satisfy a linear recurrence relation (2.14). Therefore, the maximal number of moments of $\mathfrak{G}_\alpha(x) \not\equiv 0$ which are allowed to vanish is explicitly bounded by (2.14). For instance, $\widetilde{M}$ may be taken as the length of the recurrence plus the value of the largest positive integer zero of its leading coefficient.

Example 3.3 (Piecewise-constant functions on $[0, 1]$). Each piece of g is a constant $g_i(x) \equiv c_i$. The operator $\mathfrak{D} = \prod_{i=1}^{\mathcal{K}}(x - \xi_i)\partial$ annihilates the entire g and $\mathcal{E}(E) = E(E - I)$.

- First we determine the minimal size of the system (3.4). For every polynomial $q(x)$ of degree $\mathcal{K}$, the moments of the function $f_q = (\mathcal{E}(x)q(x)\partial)\, g$ are

$$m_k(f_q) = \sum_{i=1}^{\mathcal{K}} q_i \xi_i^k, \qquad q_i = \xi_i(\xi_i - 1)q(\xi_i).$$

Assume that the ξ_i's are pairwise distinct and do not coincide with the endpoints. We claim that $\widehat{M} = \mathcal{K} - 1$ is sufficient. Indeed, assume that $m_k(f_q) = 0$ for $k = 0, 1, \ldots, \mathcal{K} - 1$. Then

$$X\mathbf{q} = \mathbf{0}: \quad X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \xi_1 & \xi_2 & \cdots & \xi_{\mathcal{K}} \\ \xi_1^2 & \xi_2^2 & \cdots & \xi_{\mathcal{K}}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \xi_1^{\mathcal{K}-1} & \xi_2^{\mathcal{K}-1} & \cdots & \xi_{\mathcal{K}}^{\mathcal{K}-1} \end{bmatrix}, \qquad \mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_{\mathcal{K}} \end{bmatrix}$$

The matrix X is a Vandermonde matrix and it is nonsingular because $\xi_i \neq \xi_j$ for $i \neq j$. It follows that $\mathbf{q} = \mathbf{0}$ and therefore $q(\xi_i) = 0$ for all i. Therefore every solution of (3.4) is a multiple of $\prod_{i=1}^{\mathcal{K}}(x - \xi_i)$.

- Now we determine the minimal size of (3.7). The space $\mathcal{N}_{\mathfrak{D}}$ is spanned by piecewise-constant functions with the jump points ξ_i. Let α be a solution to (3.7) and let the function $\mathfrak{G}_\alpha(x)$ be as in (3.10). This $\mathfrak{G}_\alpha(x)$ is again a piecewise-constant function with $\mathcal{K}$ jump points ξ_i. Therefore, the moments of $\mathfrak{G}_\alpha(x)$ satisfy the recurrence relation (2.14). This recurrence has nonzero leading coefficient and its length is $\mathcal{K} + 2$. Therefore, vanishing of the first $\mathcal{K} + 1$ moments of $\mathfrak{G}_\alpha(x)$ implies $\mathfrak{G}_\alpha(x) \equiv 0$ and consequently $\widetilde{M} = \mathcal{K}$ will be sufficient. The constants c_i are precisely the solution of (3.7).

17

4. Stability of inversion in a noisy setting

The presented inversion scheme assumes that the unknown signal $g \in A_{\mathfrak{D}}^*$ is clean from noise and that the moment sequence $\{m_k\}$ is computed with infinite precision. In many applications this assumption is unrealistic. Therefore, it is practically important to analyse the sensitivity of the inversion to the noise in the data, both theoretically and numerically. In section 4.1, we give a theoretical stability estimate for one special case of linear combinations of Dirac δ-functions. Then we present results of numerical simulations for several test cases (section 4.2).

4.1. Theoretical stability analysis

The general case of an arbitrary piecewise D-finite function appears to be difficult to analyse directly. Here, we provide stability estimates for the reconstruction of the model

$$g(x) = \sum_{i=1}^{\mathcal{K}} a_i \delta(x - \xi_i). \tag{4.1}$$

We argue that it is crucial to understand the behaviour of the reconstruction in this special case. Consider a generic $g \in A_{\mathfrak{D}}^*$. Then $\mathfrak{D}g$ is a linear combination of δ-functions and their derivatives (see the proof of theorem 2.12). Thus the model (4.1) may be considered a 'prototype' which captures the discontinuous nature of g.

One can prove the following.

Theorem 4.1. *Let g be given by (4.1). Assume that the moments $m_k(g)$ are known with accuracy ϵ. There exists a constant C_1 such that the parameters may in principle be recovered with the following accuracy:*

$$|\Delta \xi_i| \leqslant C_1 \epsilon a_i^{-1}$$
$$|\Delta a_i| \leqslant C_1 \epsilon,$$

where C_1 depends only on the geometry of $\xi_1, \ldots, \xi_{\mathcal{K}}$:

$$C_1 \sim \prod_{i \neq j} |\xi_i - \xi_j|^{-1}.$$

Outline of the proof. Write the Jacobian matrix of $\mathcal{M}$ explicitly and use the inverse function theorem to get the Jacobian of $\mathcal{N}$. This matrix may be factorized as a product of a diagonal matrix diag$\{a_1, \ldots, a_{\mathcal{K}}\}$ and a Vandermonde matrix V on the grid $\xi_1, \ldots, \xi_{\mathcal{K}}$. Then C_1 may be taken to be the norm of V^{-1}. $\qquad\square$

The above result can be generalized to the case of the model

$$g(x) = \sum_{i=1}^{\mathcal{K}} \sum_{j=0}^{k_j} a_{ij} \delta^{(j)}(x - \xi_i).$$

We plan to present these results in detail separately.

The estimate of theorem 4.1 reflects only the *problem conditioning* (i.e. the sensitivity of the map $\mathcal{N}$ defined in section 3 with respect to perturbations). Moreover, this estimate is sharp in the worst-case scenario, since it is based on directly evaluating the norm of the Jacobian.

On the other hand, an accurate analysis of the *algorithm itself* should involve estimates for the condition numbers of the matrices H and C (see (3.4) and (3.7), respectively) as well as the sensitivity of the root finding step (figure 2). Let us briefly discuss these.
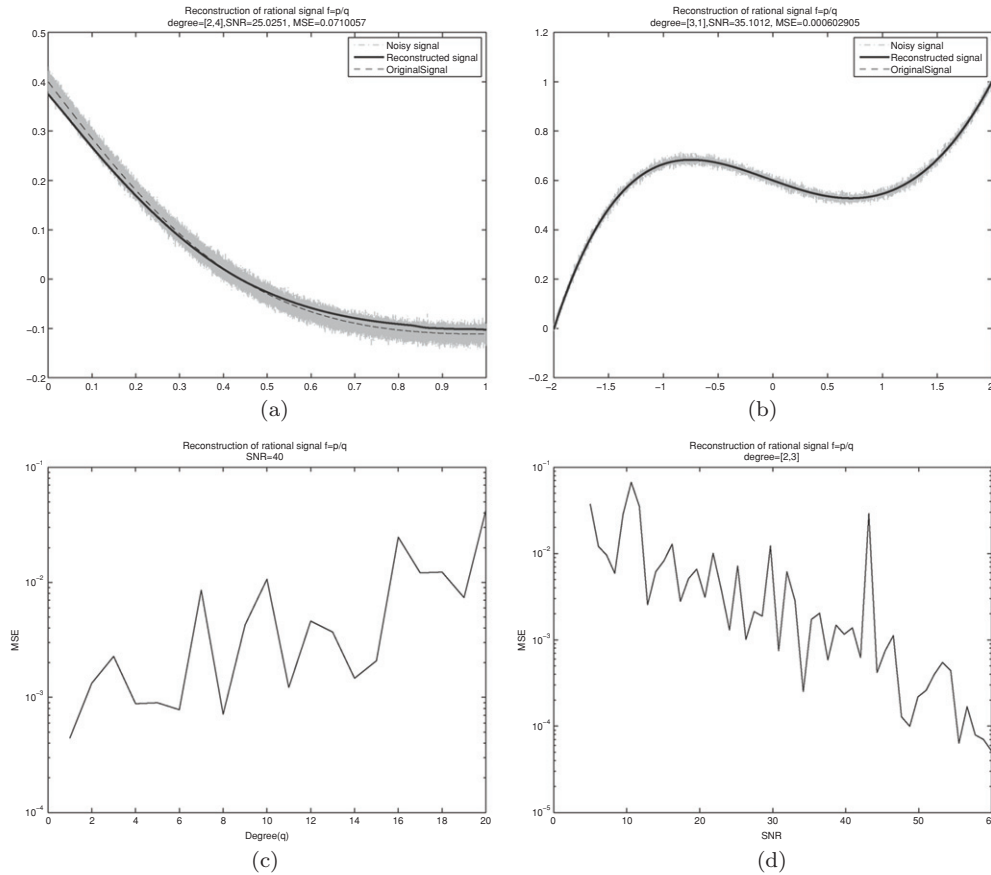
Figure 3. Reconstruction of rational signals $f = p/q$. (a) The signal with $\deg p = 2$, $\deg q = 4$ corrupted by noise with SNR $= 25$ dB. Reconstruction MSE is 0.07.   (b) The signal with $\deg p = 3$, $\deg q = 1$ corrupted by noise with SNR $= 35$ dB. Reconstruction MSE is 0.0006. The reconstructed signal is visually indistinguishable from the original. (c) Dependence of the MSE on the degree of the denominator with SNR $= 40$ dB. (d) Dependence of the MSE on the SNR with $\deg p = 2$, $\deg q = 3$.

(1) By theorem 3.1 and remark 3.1, solving (3.4) is equivalent to calculating a Hermite–Padé approximant to $\Phi = \{1, h_0(z), \dots, h_N(z)\}$. Estimates for the condition number of H in terms of Φ are available [8]. These estimates may hopefully be understood in terms of the differential operator $\mathfrak{D}$ and the function g itself, using the connection provided by theorem 3.1.

(2) It is known (e.g. [31]) that if the coefficients of a (monic) polynomial are known up to precision ϵ, then the error in recovering a root of multiplicity m may be as large as $\epsilon^{\frac{1}{m}}$.

(3) There is some freedom in the choice of the basis $\{u_i\}_{i=1}^{N}$ for the null space of $\mathfrak{D}$ (proposition 3.3). It should be investigated how this choice affects the condition number of C.

Considering the above, we hope that a stable reconstruction is possible at least for signals with relatively few and sufficiently separated points of discontinuity, as well as some mild conditions on the other model parameters.
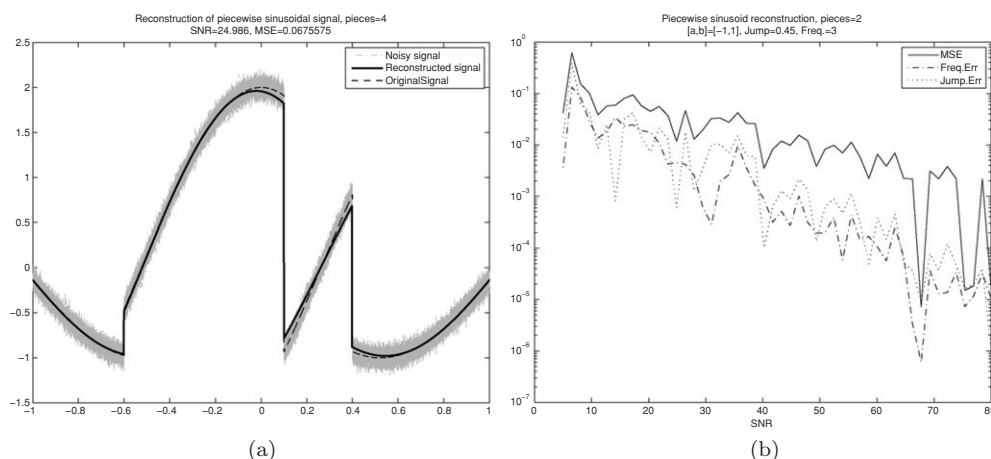
Figure 4. Reconstruction of piecewise sinusoids. (a) A sinusoid consisting of four pieces, corrupted by noise with SNR = 25 dB. Reconstruction MSE = 0.068. (b) Dependence of the MSE on the SNR in the reconstruction of a two-piece sinusoid. Also plotted are the relative errors in the estimated location of the jump point and the estimated frequency.

4.2. Numerical results

In order to further justify the hope expressed in the previous section, we have performed several numerical simulations using a straightforward implementation of the reconstruction algorithm from section 3 (the implementation details are provided in the appendix).

We have chosen three different models for the unknown signal: a rational function (no discontinuities), a piecewise sinusoid and a piecewise-polynomial function. In every simulation, the signal was first sampled on a sufficiently dense grid, then a white Gaussian noise of specified signal-to-noise ratio (SNR) was added to the samples, and finally the moments of the noised signal were calculated by trapezoidal numerical integration. To measure the success of the reconstruction, we calculated both the mean squared error (MSE) and the error in the relevant model parameters. The results are presented in figures 3, 4 and 5. Note that the SNR is measured in decibels (dB).

Several test signals were successfully recovered after a moderate amount of noise was added. Furthermore, an increase in SNR generally leads to improvement in accuracy. Nevertheless, some degree of instability is present, which is evident from the peaks in figures 3(d), 4(b) and 5(d).

While for some models, the increase in complexity leads to severe performance degradation, for other models it is not the case. Compare the growth in the degree of the rational function on one hand (figure 3(c)), and the increase in the number of jump points for the piecewise-constant reconstruction on the other hand (figure 5(c)).

5. Discussion

The piecewise *D*-finite moment inversion problem appears to be far from completely solved. In the theoretical direction, the solvability conditions need to be further refined. In particular, we hope that minimality results ('what is the minimal number of measurements which is sufficient to recover the model uniquely?') in the spirit of examples 3.2 and 3.3 concerning the
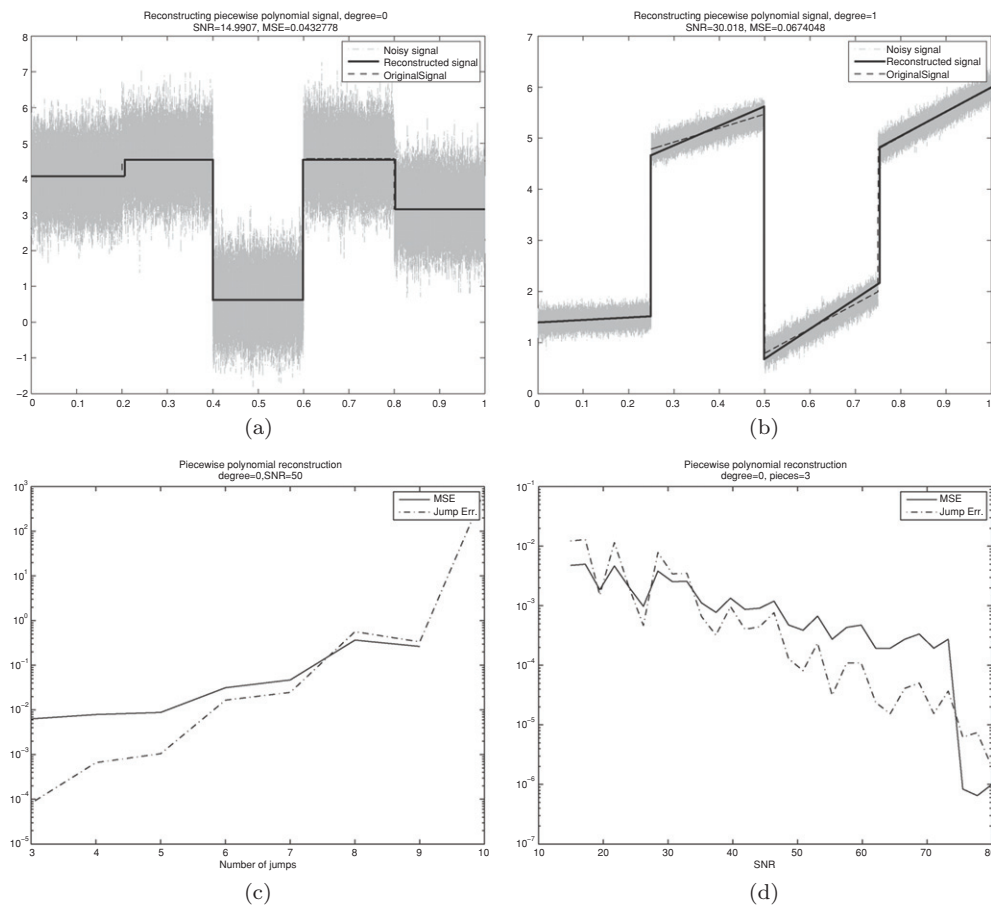
Figure 5. Reconstruction of piecewise-polynomials. (a) Piecewise-constant signal with five jumps corrupted by noise with SNR = 15 dB. Reconstruction MSE is 0.043. The reconstructed signal is visually indistinguishable from the original. (b) Piecewise-linear signal with three jumps, corrupted by noise with SNR = 30 dB. Reconstruction MSE is 0.067. (c) Dependence of the MSE on the number of jumps ($\mathcal{K}$) for a piecewise-constant signal, with SNR = 50 dB. The reconstruction failed for $\mathcal{K} = 10$. (d) Dependence of the MSE on the SNR for a piecewise-constant signal with two jumps. (c)+(d) Also plotted is the relative error in the recovered jump location.

reconstruction of additional classes of functions may be obtained using the methods presented in this paper (section 3.3). The importance of this question is discussed e.g. in [27], where estimates on the finite moment determinacy of piecewise-algebraic functions are given. In this context, the role of the moment-generating function is not yet fully understood. Another open question is the analysis of the case in which the function satisfies different operators on every continuity interval.

In the numerical direction, the results of section 4 suggest that a 'naive' implementation of the presented algorithm is relatively accurate for simple enough signals corrupted by low noise levels. We believe that attempts to improve the robustness of the algorithm should proceed in at least the following directions.

(1) A similar question regarding stability of reconstruction of signals with the finite rate of innovation is addressed in [24]. We propose to investigate the applicability of that method to our reconstruction problem.

(2) The connection of the system (3.4) to Hermite–Padé approximation may hopefully be exploited to build some kind of sequence of approximants which converge to the true solution, as more elements of the moment sequence are known. In fact, a similar approach (with standard Padé approximants) has been used in the 'noisy trigonometric moment problem' [3].

Furthermore, we believe that it is important to understand how the various model parameters influence the stability of the algorithm. The most general answer in our context would be to give stability estimates in terms of the differential operator $\mathfrak{D}$ and the geometry of the jump points.

On the other hand, the algebraic moments (1.1) are known to be a non-optimal choice for measurements (see [32]) due to their strong non-orthogonality. As stated in the introduction, the moment inversion is a 'prototype' for some real-world problems such as reconstruction from Fourier measurements. In fact, it is known that the Fourier coefficients (as well as coefficients with respect to other orthogonal systems) of many 'finite-parametric' signals satisfy various relations, and this fact has been utilized in numerous schemes for nonlinear Fourier inversion [2, 4, 5, 9, 12, 22]. The point of view presented in this paper, namely, the differential operator approach, can hopefully be generalized to include these types of measurements as well. In this regard, we expect that the methods of holonomic combinatorics [35] may provide useful insights.

Acknowledgments

Appendix. The inversion algorithm—implementation details

We have implemented the inversion algorithm of section 3 in the MATLAB environment [29]. For each of the three types of signals we built a specific inversion routine. All the three routines have a common base as follows.

- The matrix H is chosen to be square. The solution to the system $H\mathbf{a} = \mathbf{0}$ is obtained by taking the highest coefficient of $\mathbf{a}$ to be 1 and then performing standard Gaussian elimination with partial pivoting on the reduced system.
- The root finding step is performed with the `pejroot` routine of the `MULTROOT` package [36, 37], which takes into account the multiplicity structure of the polynomial.
- The calculation of the moments in the matrix C is done via numerical quadrature.

The difference between the routines lies in the construction of H and the computation of the basis for the null space of $\mathfrak{D}$. These are determined by the signal type, as described below.

(1) *Rational functions.* A rational function $f(x) = \frac{p(x)}{q(x)}$ is annihilated by the first-order operator

$$\mathfrak{D} = (-pq)\partial + (p'q - pq')\mathrm{I}.$$

Therefore, the matrix H consists of two 'stripes' V_0 and V_1 (the structure of the matrices V_i is given by (3.6)). If $\deg p(x) = r$ and $\deg q(x) = s$, then the degrees of the coefficients of $\mathfrak{D}$ are $k_0 = r + s - 1$ and $k_1 = r + s$. Subsequently, V_0 is $(k_0 + k_1 + 2) \times (k_0 + 1)$ and V_1 is $(k_0 + k_1 + 2) \times (k_1 + 1)$. After $\mathfrak{D}$ is reconstructed (without explicitly recovering p and q), a solution u to $\mathfrak{D}u \equiv 0$ is obtained by numerically solving (using the `ode45` solver)

the initial value problem $\{\mathfrak{D}u(x) \equiv 0, u(a) = 1, x = a \ldots b\}$. Then the final solution is obtained as

$$\widetilde{f}(x) = \frac{\int_a^b u(x)\,\mathrm{d}x}{m_0} u(x).$$

(2) *Piecewise sinusoids.* Let $g(x)$ consist of $\mathcal{K} + 1$ pieces of the form

$$g_n(x) = A_n \sin(\omega x + \varphi_n).$$

The annihilating operator for each piece is $\mathfrak{D} = \partial^2 + \omega^2 \mathrm{I}$; therefore, the entire g is annihilated by

$$\widehat{\mathfrak{D}} = \{(x - \xi_1)^2 \cdot \ldots \cdot (x - \xi_{\mathcal{K}})^2\}(\partial^2 + \omega^2 \mathrm{I}).$$

So $H = [V_0 \;\; V_2]$, where both V_0 and V_2 are $2(2\mathcal{K} + 1) \times (2\mathcal{K} + 1)$. From the operator reconstruction step we obtain $\widetilde{\mathfrak{D}} = p_1(x)\partial^2 + p_0(x)\mathrm{I}$. The jump points are taken to be the arithmetic means of the corresponding roots of p_0 and p_1. Because of the normalization performed in the operator recovery step, the frequency ω' may be taken to be the square root of the highest coefficient of p_0. The reconstruction is considered to be unsuccessful if at least one root fails to lie inside the interval $[a, b]$.

The basis for the null space of $\mathfrak{D}$ is chosen to be $u_1 = \sin(\omega' x)$ and $u_2 = \cos(\omega' x)$.

(3) *Piecewise polynomials.* Let g consist of $\mathcal{K} + 1$ polynomial pieces g_n, where $\deg g_n = d$. The annihilating operator for every piece is $\mathfrak{D} = \partial^{d+1}$, and therefore g is annihilated by

$$\widehat{\mathfrak{D}} = \{(x - \xi_1)^{d+1} \cdot \ldots \cdot (x - \xi_{\mathcal{K}})^{d+1}\}\partial^{d+1}.$$

Consequently, $H = V_{d+1}$ is a $\mathcal{K}(d + 1) \times \mathcal{K}(d + 1)$ Hankel matrix. Operator reconstruction step gives $\widetilde{\mathfrak{D}} = p(x)\partial^{d+1}$, and the jump locations are the roots of $p(x)$, each with multiplicity $d + 1$. The reconstruction is considered to be unsuccessful if at least one root fails to lie inside the interval $[a, b]$.

The basis for the null space of $\mathfrak{D}$ is always chosen to be $\{1, x, \ldots, x^d\}$.

References

[1] Ang D D 2002 *Moment Theory and Some Inverse Problems in Potential Theory and Heat Conduction* (Berlin: Springer)
[2] Banerjee N S and Geer J F 1997 Exponential approximations using Fourier series partial sums *Technical Report* TR-97-56 (ICASE)
[3] Barone P and March R 2000 Reconstruction of a piecewise constant function from noisy Fourier coefficients by Padé method *SIAM J. Appl. Math.* **60** 1137–56
[4] Beckermann B, Matos A C and Wielonsky F 2008 Reduction of the Gibbs phenomenon for smooth functions with jumps by the ε-algorithm *J. Comput. Appl. Math.* **219** 329–49
[5] Berent J and Dragotti P L 2006 Perfect reconstruction schemes for sampling piecewise sinusoidal signals *IEEE Int. Conf. Acoustics, Speech, Signal Processing (Toulouse, France, May 2006)* vol 3 (Piscataway, NJ: IEEE) pp III, 377–80
[6] Borwein J M and Salvy B 2008 A proof of a recursion for Bessel moments *Exp. Math.* **17** 223–30
[7] Byrnes C I and Lindquist A 1983 On the partial realization problem *Linear Algebra Appl.* **50** 277–319
[8] Cabay S, Jones A R and Labahn G 1996 Computation of numerical Pade-Hermite and simultaneous Pade systems I: near inversion of generalized Sylvester matrices *SIAM J. Matrix Anal. Appl.* **17** 248–67
[9] Dragotti P L, Vetterli M and Blu T 2007 Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-fix *IEEE Trans. Signal Process.* **55** 1741
[10] Driscoll T A and Fornberg B 2001 A Padé-based algorithm for overcoming the Gibbs phenomenon *Numer. Algorithms* **26** 77–92
[11] Eckhoff K S 1993 Accurate and efficient reconstruction of discontinuous functions from truncated series expansions *Math. Comput.* **61** 745–63
[12] Eckhoff K S 1995 Accurate reconstructions of functions of finite regularity from truncated Fourier series expansions *Math. Comput.* **64** 671–90

[13] Elaydi S 2005 *An Introduction to Difference Equations* (Berlin: Springer)
[14] Ettinger B, Sarig N and Yomdin Y 2008 Linear versus non-linear acquisition of step-functions *J. Geom. Anal.* **18** 369–99
[15] Flajolet P and Sedgewick R 2008 *Analytic Combinatorics* (UK: Cambridge University Press)
[16] Gelfand I M and Shilov G E 1958 *Generalized Functions and Operations on Them (Russian)* (Moscow: Fizmatgiz)
[17] Golub G H, Milanfar P and Varah J 2000 A stable numerical method for inverting shape from moments *SIAM J. Sci. Comput.* **21** 1222–43
[18] Gustafsson B, He C, Milanfar P and Putinar M 2000 Reconstructing planar domains from their moments *Inverse Problems* **16** 1053–70
[19] Ince E L 1956 *Ordinary Differential Equations* (New York: Dover)
[20] Kalman R E and DeClaris N (ed) 1971 On minimal partial realizations of a linear input/output map *Aspects of Network and System Theory* (New York: Holt, Rinehart & Winston) pp 385–407
[21] Kisunko V 2008 Cauchy type integrals and a D-moment problem *Math. Rep. Acad. Sci. R. Soc. Can.* **29** 115–22
[22] Kvernadze G 2004 Approximating the jump discontinuities of a function by its Fourier–Jacobi coefficients *Math. Comput.* **73** 731–52
[23] Labahn G 1992 Inversion components of block Hankel-like matrices *Linear Algebra Appl.* **177** 7–48
[24] Maravic I and Vetterli M 2005 Sampling and reconstruction of signals with finite rate of innovation in the presence of noise *IEEE Trans. Signal Process.* **53** 2788–805
[25] March R and Barone P 1998 Application of the Pade method to solve the noisy trigonometric moment problem: some initial results *SIAM J. Appl. Math.* **58** 324–43
[26] Ore O 1933 Theory of non-commutative polynomials *Ann. Math.* **34** 480–508
[27] Sarig N and Yomdin Y 2008 Signal acquisition from measurements via non-linear models *Math. Rep. Acad. Sci. R. Soc. Can.* **29** 97–114
[28] Schelter W *et al* 2001 *Maxima Manual* http://maxima.sourceforge.net/docs/manual/en/maxima.html
[29] Sigmon K and The MathWorks Staff 1998 *MATLAB Primer* 5th edn (Boca Raton, FL: CRC Press)
[30] Stanley R P 1980 Differentiably finite power series *Eur. J. Combin.* **1** 175–88
[31] Stewart G W 1980 The behavior of a multiplicity independent root-finding scheme in the presence of error *BIT Numer. Math.* **20** 526–8
[32] Talenti G 1987 Recovering a function from a finite number of moments *Inverse Problems* **3** 501–17
[33] Vetterli M, Marziliano P and Blu T 2002 Sampling signals with finite rate of innovation *IEEE Trans. Signal Process.* **50** 1417–28
[34] Weixlbaumer C 2001 *Solutions of Difference Equations with Polynomial Coefficients* (Australia: RISC Linz Johannes Kepler Universitat)
[35] Zeilberger D 1990 A holonomic systems approach to special functions identities *J. Comput. Appl. Math.* **32** 321–68
[36] Zeng Z 2004 Algorithm 835: MultRoot—a Matlab package for computing polynomial roots and multiplicities *ACM Trans. Math. Softw. (TOMS)* **30** 218–36
[37] Zeng Z 2005 Computing multiple roots of inexact polynomials *Math. Comput.* **74** 869–904