Sparse Shifts for Univariate Polynomials

Lakshman Y. N.[1]

Department of Mathematics and Computer Science
Drexel University
Philadelphia, PA 19104

B. David Saunders[2]

Department of Computer and Information Sciences
University of Delaware
Newark, DE 19716
e-mail: `lakshman@mcs.drexel.edu, saunders@udel.edu`

Abstract

Let $f(x)$ be a polynomial of degree d with rational coefficients and let t be a positive integer $\leq \deg(f)$. We consider the problem of finding a t-sparse shift for $f(x)$. The problem is to find an α, if one exists (in some algebraic extension of the rationals), such that in the representation of $f(x)$ in the basis $1, x-\alpha, (x-\alpha)^2, \ldots$, i.e., $f(x) = \sum_{i=0}^{d} \mathtt{F}_i (x-\alpha)^i$, *at most* t of the coefficients $\mathtt{F}_i$ are non-zero. We derive explicit conditions for the uniqueness and rationality of a t-sparse shift for $f(x)$ and provide an efficient algorithm for computing a sparse shift when one exists. We also point out an application of our result to the problem of constructing sparse decompositions of univariate polynomials.

Keywords

Sparse polynomials, BCH decoding, Real roots, Pochhammer basis, Chebyshev basis, Descartes' rule of signs.

[1]Work supported by NSF grant CCR–9203062
[2]Work supported by NSF grant CCR-9123666.

Introduction

In this paper, we consider the problem of computing t-sparse shifts for univariate polynomials. Given a polynomial $f(x) \in \mathcal{F}[x]$ of degree d (where $\mathcal{F}$ is a field of characteristic 0), consider the representation of $f(x)$ in the basis $1, x-\alpha, (x-\alpha)^2, \ldots$ for some $\alpha \in \mathcal{K}$, an extension of $\mathcal{F}$, i.e.,

$$f(x) = \sum_{i=0}^{d} \mathrm{F}_i (x-\alpha)^i.$$

Let t be a positive integer $\leq d$. We say that α is a t-sparse shift for $f(x)$ (or, $f(x)$ is t-sparse in the shifted basis $1, x-\alpha, (x-\alpha)^2, \ldots$) if *at most* t of the F_i in the above representation are non-zero.

The main problem that we address is: given an $f(x)$ and t as above, can we efficiently compute a t-sparse shift for $f(x)$ if one exists? We construct an efficient algorithm for solving this problem and answer several related questions of interest, such as: When is a shift unique? When is a shift rational, meaning when does α belong to $\mathcal{F}$? How many evaluations does one need to distinguish 2 polynomials that are shifted t-sparse (with respect to possibly different shifts)? As an application of our algorithm, we construct a polynomial time algorithm for computing sparse decompositions of univariate polynomials. The problem is the following. Given a polynomial $f(x) \in \mathcal{F}[x]$, and a positive integer t, find polynomials $g_1, g_2, \ldots, g_s \in \mathcal{F}[x]$ each of which is t-sparse (meaning, has no more than t terms), such that $f(x) = g_1(g_2(\ldots(g_s(x))))$, if they exist.

Recently, there has been much interest in the design of efficient algorithms for computing sparse representations for various classes of functions such as polynomials, rational functions, and algebraic functions [1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 14, 18, 21]. Sparse interpolation is well recognized as a very useful tool for controlling intermediate expression swell in computer algebra [15, 24]. Sparse polynomials and rational functions can be evaluated quickly and that makes them attractive to several applications and an interesting line of research is to try to infer properties of sparse polynomials (such as divisibility, existence of non-trivial gcd, the existence of real roots etc.) from their values at a small number points (see [12]). Traditionally, "sparse polynomial (or rational function)" is taken to mean a polynomial (or rational function) with a "few terms" where the "terms" are power products of the variables involved. One can reasonably ask for sparse representations for polynomials in other bases such as the Chebyshev polynomials or the shifted power basis $1, x-\alpha, (x-\alpha)^2, \ldots$ Sparse univariate polynomial interpolation in the Chebyshev basis, the Pochhammer basis are treated in [18] and a generalized sparse interpolation algorithm that works for a class of bases (which includes the Pochhammer basis) can be found in [10].

Computing t-sparse shifts and the representation of $f(x)$ in the basis $1, x-\alpha, (x-\alpha)^2, \ldots$ was raised as an open problem by Borodin and Tiwari in [2]. Subsequently, the problem was taken up in a recent paper by Grigoriev and Karpinski [11] where they provide an elegant algebraic criterion to be satisfied by any t-sparse shift for $f(x)$. We use the above paper of

Grigoriev and Karpinski as our starting point for further refinement. The main departure is in our assumption that $f(x)$ is given *explicitly* (what we really need are the values of the first $2t$ derivatives of $f(x)$ at *two* distinct points, different from α). This is indeed a much stronger requirement than the requirement of Grigoriev and Karpinski [11] that only a black box for $f(x)$ be given. However, the algorithm of Grigoriev and Karpinski, in many cases, requires one to do more work than if one were to first interpolate $f(x)$ (dense) from its black box and then use the algorithm developed in this paper to compute a sparse shift. Furthermore, we are able to write down explicit conditions for the existence of a t-sparse shift as the vanishing of a certain number of derivatives of $f(x)$. This might provide some hints on attacking more general problems of this type. If $f(x)$ is given by a straight-line program, then the first $2t$ derivatives can be obtained by a polynomial sized blow-up in the length of the straight-line program for computing $f(x)$ [13] and the algorithm developed in this paper can be used to compute a sparse shift for $f(x)$, provided one exists. The main contributions of this paper are:

- a condition on the uniqueness and the rationality of the t-sparse shift;
- an efficient, randomized algorithm for computing a t-sparse shift α and the representation of $f(x)$ in the basis $1, x-\alpha, (x-\alpha)^2, \ldots$; Our algorithm performs $O(t^5 + t^2 d)$ field operations when the shift is unique. The algorithm can be made deterministic at the cost of an additional multiplicative factor of $O(td)$.
- a criterion for distinguishing 2 polynomials, each sparse with respect to a shift, given a bound for the shifts.
- a polynomial time algorithm for computing sparse decompositions of univariate polynomials over the rationals (in general, over fields of characteristic 0).

In section 1 we develop the condition on the uniqueness and the rationality of the t-sparse shift; in section 2, an efficient algorithm for computing a sparse shift α is given, and in section 3 we provide a criterion for distinguishing 2 polynomials which are sparse with respect to bounded shifts. Section 4 contains a description of the sparse decomposition algorithm. The proof of an important technical lemma in section 1 is relegated to an appendix in order to maintain the focus of the exposition on the uniqueness condition. A prelimminary version of this appeared as [19].

1. Uniqueness of Sparse Shifts

Consider the Taylor expansion of $f(x)$ about α, namely,

$$f(x) = f(\alpha) + f^{(1)}(\alpha)(x-\alpha) + f^{(2)}(\alpha)(x-\alpha)^2/2! + \ldots + f^{(d)}(\alpha)(x-\alpha)^d/d!$$

where $f^{(i)}(\alpha)$ denotes $\frac{d^i}{dx^i} f(x)\Big|_{x=\alpha}$. If α is a t-sparse shift for $f(x)$, i.e., there exist $\delta_1 = d > \delta_2 > \ldots > d_t \in \mathcal{Z}_+$ and $\mathtt{F}_1, \mathtt{F}_2, \ldots, \mathtt{F}_t \in \mathcal{K}$ such that $f(x) = \mathtt{F}_1(x-\alpha)^{\delta_1} + \mathtt{F}_2(x-\alpha)^{\delta_2} + \ldots + \mathtt{F}_t(x-\alpha)^{\delta_t}$, then, comparing it with the above Taylor expansion, we notice that at least $d+1-t$ of the derivatives of $f(x)$ at $x = \alpha$ must be zero. The vanishing of a derivative at a point is a linear constraint on the coefficients of $f(x)$ and there can be at most $d+1$ linearly independent constraints on the coefficients of $f(x)$. In what follows, we show that if $t \leq (d+1)/2$, then asking $f(x)$ to be t-sparse in 2 distinct shifted bases means imposing more constraints than can be satisfied on the coefficients of $f(x)$ and therefore $f(x)$ can be t-sparse in at most one shifted basis. We prove this intuitive observation in theorem 1. As a corollary, we conclude that such a t-sparse shift α must be rational. First some requisite lemmas.

Lemma 1 *Let $t \leq m < d$ be positive integers such that $t+m \leq d$. Let $\mathtt{L}$ be a non-increasing function mapping the interval $(0, \ldots, d)$ to $(1, \ldots, t)$. Finally, let $i_1 < i_2 < \ldots < i_m$ be a subsequence of $(0, \ldots, d)$. Then there is an s such that $\mathtt{L}(i_s) = m + 1 - s$.*

Proof: Consider the last k such that $\mathtt{L}(i_k) \leq m+1-k$. Such a k exists, since of necessity $\mathtt{L}(i_1) \leq t \leq m = (m+1)-1$ and $\mathtt{L}(i_m) \geq 1 = (m+1)-m$. For $k = m$ the result follows immediately. If $k < m$ then we have $\mathtt{L}(i_k) \leq (m+1)-k$ and $\mathtt{L}(i_{k+1}) > (m+1)-(k+1) = m-k$. But $\mathtt{L}$ is non-increasing, so we must have $\mathtt{L}(i_k) = (m+1)-k = \mathtt{L}(i_{k+1})$. □

Second, we'll need the non-singularity of a particular matrix. For positive integer i, let $j^{\underline{i}}$ denote $j(j-1)(j-2)\ldots(j-i+1)$.

Lemma 2 *Let integer sequences (p_l) and (q_l) be given such that $p_1 > p_2 > \ldots > p_r > 0$, $q_1 > q_2 > \ldots > q_r = 0$, and $p_l > q_l$ for $l = 1, \ldots, r$. Then $\det(A)$ is non-zero, where A is the $r \times r$ matrix*

$$A = \begin{pmatrix} p_1^{\underline{q_1}} & p_1^{\underline{q_2}} & \ldots & p_1^{\underline{q_r}} \\ p_2^{\underline{q_1}} & p_2^{\underline{q_2}} & \ldots & p_2^{\underline{q_r}} \\ \vdots & \vdots & \ddots & \vdots \\ p_r^{\underline{q_1}} & p_r^{\underline{q_2}} & \ldots & p_r^{\underline{q_r}} \end{pmatrix}.$$

Proof: See the appendix for the proof of this lemma. □

Theorem 1 *Let $f(x) \in \mathcal{F}[x]$ be of degree d and let $t \leq (d+1)/2$. If there exists an α in some algebraic extension $\mathcal{K}$ of $\mathcal{F}$ such that $f(x)$ is t-sparse in the shifted power basis $1, (x-\alpha), (x-\alpha)^2, \ldots$, then the shift α is unique.*

Proof: Suppose, for the sake of contradiction, that $f(x)$ is t-sparse with respect to two distinct shifts α, β in an extension $\mathcal{K}$. This implies that at least some $d+1-t$ of the derivatives of $f(x)$ vanish at $x = \alpha$ and some $d+1-t$ of the derivatives of $f(x)$ vanish at $x = \beta$.

Let us express the derivatives of $f(x)$ at $x = \alpha$ in terms of the derivatives of $f(x)$ at $x = \beta$ using Taylor expansions:

$$\begin{array}{rcccccccl} f^{(0)}(\alpha) & = & f^{(0)}(\beta) & + & f^{(1)}(\beta)(\alpha-\beta) & + & \cdots & + & f^{(d)}(\beta)(\alpha-\beta)^d/(d)! \\ f^{(1)}(\alpha) & = & f^{(1)}(\beta) & + & f^{(2)}(\beta)(\alpha-\beta) & + & \cdots & + & f^{(d)}(\beta)(\alpha-\beta)^{d-1}/(d-1)! \\ \vdots & & & & & & & & \\ f^{(d-1)}(\alpha) & = & & & & & f^{(d-1)}(\beta) & + & f^{(d)}(\beta)(\alpha-\beta) \\ f^{(d)}(\alpha) & = & & & & & & & f^{(d)}(\beta). \end{array} \tag{1}$$

It will be convenient to convert this triangular system to an equivalent Vandermonde-like system via diagonal scalings. Multiply each row i by $(d-i)!(\beta-\alpha)^i$; then factor $(\beta-\alpha)^j$ from column j, obtaining

$$(d-i)!(\alpha-\beta)^i f^{(i)}(\alpha) = \sum_{l=0}^{d}((d-i)^{\underline{(d-l)}}) \times (\alpha-\beta)^l f^{(l)}(\beta).$$

Note that $(d-i)! = (d-i)^{\underline{(d-l)}}(d-l)!$ Next we restrict to only the vanishing derivatives of $f(x)$ at $x = \alpha$. With the aid of lemma 1, we can extract a square system of equations in which the rows correspond to the *vanishing derivatives at* α and the columns correspond to all *the non-vanishing derivatives at* β on which they depend. Let $j_1, \ldots, j_t$ be the sequence of indices such that $f^{(j_k)}(\beta) \neq 0$, and let $\textsc{length}(f^{(i)}(\alpha))$ be the number of corresponding non-zero terms in the Taylor expansion of $f^{(i)}(\alpha)$ in (1). Apply lemma 1 to the non-increasing function $l(k) = \textsc{length}(f^{(i_k)}(\alpha))$ and the sequence $i_1 < \ldots < i_m$ of indices such that $f^{(i_k)}(\alpha) = 0$. Choose the largest s such that $\textsc{length}(f^{(i_s)}(\alpha)) = m+1-s$. Let $r = m+1-s$. The r equations corresponding to $f^{(i_s)}(\alpha)$ through $f^{(i_m)}(\alpha)$ then form the following square system

$$\begin{pmatrix} p_1^{q_1} & p_1^{q_2} & \cdots & p_1^{q_r} \\ p_2^{q_1} & p_2^{q_2} & \cdots & p_2^{q_r} \\ \vdots & \vdots & \ddots & \vdots \\ p_r^{q_1} & p_r^{q_2} & \cdots & p_r^{q_r} \end{pmatrix} \begin{pmatrix} (\alpha-\beta)^{d-q_1} f^{(d-q_1)}(\beta) \\ (\alpha-\beta)^{d-q_2} f^{(d-q_2)}(\beta) \\ \vdots \\ (\alpha-\beta)^{d-q_r} f^{(d-q_r)}(\beta) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where $p_k = i_{s+r-k}$ and $d - q_k = j_{(t-r+k)}$, for $k = 1, \ldots, r$. By lemma 2, this $r \times r$ matrix is non-singular so that the vector must be zero. The vector is zero only if $\alpha = \beta$, in other words, only if the shifts are identical. $\square$

Corollary 1 *Let $f(x) \in \mathcal{F}[x]$ be of degree d and let $t \leq (d+1)/2$. If α (in any extension $\mathcal{K}$ of $\mathcal{F}$) is a t-sparse shift (hence the unique t-sparse shift) for $f(x)$, then $\alpha \in \mathcal{F}$.*

Proof: Since α is a root of a number of polynomials (derivatives of $f(x)$) in $\mathcal{F}[x]$, α is algebraic over $\mathcal{F}$. Let $h(x) \in \mathcal{F}[x]$ be the minimal polynomial of α over $\mathcal{F}$. Since α is a t-sparse shift for $f(x)$, it is a common root of at least $d+1-t$ derivatives of $f(x)$. This means that $h(x)$ divides all those derivatives and therefore, all the conjugates of α are also common zeros of those derivatives. This implies that every conjugate of α is also a t-sparse shift for $f(x)$. Therefore, by the uniqueness of the shift α, $h(x)$ has degree 1, i.e., $\alpha \in \mathcal{F}$. □

2. An Algorithm to Determine Sparse Shifts

Let $f(x) = \mathrm{F}_1(x-\alpha)^{\delta_1} + \mathrm{F}_2(x-\alpha)^{\delta_2} + \ldots + \mathrm{F}_t(x-\alpha)^{\delta_t}$ where $\delta_1 > \delta_2 > \ldots > \delta_t$, and let $\Phi(z) = z^{\underline{t}} + \phi_{t-1} z^{\underline{t-1}} + \ldots + \phi_1 z^{\underline{1}} + \phi_0 z^{\underline{0}}$ be the polynomial of degree t whose roots are precisely $\delta_1, \delta_2, \ldots, \delta_t$, i.e., $\Phi(\delta_i) = 0$ for $i = 1, 2, \ldots, t$. Express $\Phi(z)$ as a monic polynomial in the *falling Pochhammer basis* for reasons that will be explained shortly ($z^{\underline{n}}$ denotes $z(z-1)(z-2)\ldots(z-n+1)$ for $n \geq 1$ with $z^{\underline{0}}$ denoting 1). Consider the following set of equalities:

$$\begin{array}{rcl} \phi_0 f(x) & = & \phi_0 \mathrm{F}_1(x-\alpha)^{\delta_1} + \phi_0 \mathrm{F}_2(x-\alpha)^{\delta_2} + \ldots + \phi_0 \mathrm{F}_t(x-\alpha)^{\delta_t} \\ \phi_1 f^{(1)}(x)(x-\alpha) & = & \phi_1 \delta_1 \mathrm{F}_1(x-\alpha)^{\delta_1} + \phi_1 \delta_2 \mathrm{F}_2(x-\alpha)^{\delta_2} + \ldots + \phi_1 \delta_t \mathrm{F}_t(x-\alpha)^{\delta_t} \\ \vdots & = & \vdots \\ f^{(t)}(x)(x-\alpha)^t & = & \delta_1^{\underline{t}} \mathrm{F}_1(x-\alpha)^{\delta_1} + \delta_2^{\underline{t}} \mathrm{F}_2(x-\alpha)^{\delta_2} + \ldots + \delta_t^{\underline{t}} \mathrm{F}_t(x-\alpha)^{\delta_t} \end{array} \tag{2}$$

Adding the left and right sides of the above set of equalities, we get

$$\begin{array}{rcl} \sum_{i=0}^{t} \phi_i f^{(i)}(x)(x-\alpha)^i & = & \Phi(\delta_1)\mathrm{F}_1(x-\alpha)^{\delta_1} + \Phi(\delta_2)\mathrm{F}_2(x-\alpha)^{\delta_2} + \ldots + \Phi(\delta_t)\mathrm{F}_t(x-\alpha)^{\delta_t} \\ & = & 0 \end{array} \tag{3}$$

Since the $t+1$ polynomials $(x-\alpha)^i f^{(i)}(x), i = 0, \ldots, t$ satisfy a linear relation, their their Wronskian vanishes (this is classical; see, for example, [22], page 664). The converse is also true.

Theorem 2 *For any $\alpha \in \mathcal{K}$, $f(x)$ is t-sparse in the α-shifted power basis for $\mathcal{K}[x]$ if* and only if *the Wronskian $\mathcal{W}_{t+1}(x, \alpha) = 0$, where $\mathcal{W}_k(x, z) = \det(W_k(x, z))$ and $W_k(x, z)$ is the $k \times k$ matrix*

$$\begin{pmatrix} g_0(x,z) & g_1(x,z) & \ldots & g_{k-1}(x,z) \\ (d/dx)(g_0(x,z)) & (d/dx)(g_1(x,z)) & \ldots & (d/dx)(g_{k-1}(x,z)) \\ \vdots & \vdots & \ddots & \vdots \\ (d^{k-1}/dx^{k-1})(g_0(x,z)) & (d^{k-1}/dx^{k-1})(g_1(x,z)) & \ldots & (d^{k-1}/dx^{k-1})(g_{k-1}(x,z)) \end{pmatrix},$$

for $g_i(x, z) = (x-z)^i f^{(i)}(x)$.

Proof: ($\Longrightarrow$) : Well-known (see [22], page 664).

($\Longleftarrow$) : Since $g_0(x,\alpha), g_1(x,\alpha), \ldots, g_t(x,\alpha) \in \mathcal{K}[x]$, $\mathcal{W}_{t+1}(x,\alpha) = 0$ implies that $g_0(x,\alpha)$, $g_1(x,\alpha)$, $\ldots$, $g_t(x,\alpha)$ are $\mathcal{K}$-linearly dependent [16]. In other words, $\exists c_0, c_1, \ldots, c_t \in \mathcal{K}$ s.t. $\sum_{i=0}^{t} c_i g_i(x,\alpha) = 0$. Let us look at the Taylor expansions of $g_0(x,\alpha)$, $g_1(x,\alpha)$, $\ldots, g_t(x,\alpha)$ about $x = \alpha$:

$$\begin{aligned}
g_0(x,\alpha) &= \frac{f^{(d)}(\alpha)}{d!}(x-\alpha)^d + \frac{f^{(d-1)}(\alpha)}{(d-1)!}(x-\alpha)^{d-1} + \ldots + f^{(1)}(\alpha)(x-\alpha) + f^{(0)}(\alpha) \\
g_1(x,\alpha) &= \frac{f^{(d)}(\alpha)}{(d-1)!}(x-\alpha)^d + \frac{f^{(d-1)}(\alpha)}{(d-2)!}(x-\alpha)^{d-1} + \ldots + f^{(1)}(\alpha)(x-\alpha) \\
&\vdots \\
g_t(x,\alpha) &= \frac{f^{(d)}(\alpha)}{(d-t)!}(x-\alpha)^d + \frac{f^{(d-1)}(\alpha)}{(d-1-t)!}(x-\alpha)^{d-1} + \ldots + f^{(t)}(\alpha)(x-\alpha)^t
\end{aligned} \tag{4}$$

If $f(x)$ is not t-sparse in the α-shifted power basis, at least $t+1$ derivatives of $f(x)$ are non-zero at α. Consider the $(t+1) \times (t+1)$ matrix

$$V_{t+1}(\alpha) = \begin{pmatrix} f^{(\delta_1)}(\alpha)/\delta_1! & f^{(\delta_2)}(\alpha)/\delta_2! & \ldots & f^{(\delta_{t+1})}(\alpha)/\delta_{t+1}! \\ f^{(\delta_1)}(\alpha)/(\delta_1-1)! & f^{(\delta_2)}(\alpha)/(\delta_2-1)! & \ldots & f^{(\delta_{t+1})}(\alpha)/(\delta_{t+1}-1)! \\ \vdots & \vdots & \ddots & \vdots \\ f^{(\delta_1)}(\alpha)/(\delta_1-t)! & f^{(\delta_2)}(\alpha)/(\delta_2-t)! & \ldots & f^{(\delta_{t+1})}(\alpha)/(\delta_{t+1}-t)! \end{pmatrix}$$

where the columns correspond to some $t+1$ non-vanishing derivatives of $f(x)$ at $x = \alpha$. The linear relation $\sum_{i=0}^{t} c_i g_i(x,\alpha) = 0$ implies that $V_{t+1}(\alpha)$ is singular. But it is easy to see that after factoring out $f^{(\delta_j)}(\alpha)/\delta_j!$ from the j-th column, we are left with the matrix $(\delta_j^{\underline{i-1}})$, which can be row-reduced to a Vandermonde matrix. Since $\delta_j \neq \delta_k$, $V_{t+1}(\alpha)$ cannot be singular. Therefore, it cannot happen that at least $t+1$ derivatives of $f(x)$ at α are non-zero, i.e., $f(x)$ must be t-sparse in the α-shifted power basis. $\square$

Corollary 2 *If α is the* unique *t-sparse shift for $f(x)$, and the representation of $f(x)$ in the α-shifted power basis has exactly t terms, then, the system of linear equations*

$$W_t(b,\alpha)\begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} g_t(b,\alpha) \\ (dg_t/dx)(b,\alpha) \\ \vdots \\ (d^{t-1}g_t/dx^{t-1})(b,\alpha) \end{pmatrix} \tag{5}$$

has $(\phi_0, \phi_1, \ldots, \phi_{t-1})^{Tr}$ (the vector of coefficients of $\Phi(z)$) as its unique *solution for any $b \neq \alpha$.*

Proof: By (3), $\sum_{i=0}^{t} \phi_i f^{(i)}(x)(x-\alpha)^i = 0$. Differentiating this equation $t-1$ times, we get $t-1$ new relations. If $y_0, \ldots, y_{t-1}$ is a solution to these equations, then $\sum_{i=0}^{t} y_i f^{(i)}(x)(x-\alpha)^i = 0$, where $y_t = 1$. Then, by the uniqueness of the representation of $f(x)$ in the α-shifted basis, each column in the set of equations (2) must be zero, so that $\sum_{i=0}^{t} y_i \delta_j^i = 0$. Thus the polynomial whose coefficients are the y_i must be $\Phi(z)$. □

Lemma 3 *For generic choices of* $b_1, b_2 \in \mathcal{F}$, $\gcd(\mathcal{W}_{t+1}(b_1, z), \mathcal{W}_{t+1}(b_2, z)) = c(z-\alpha)^\sigma$.

Proof: It follows from theorem 2 that $\mathcal{W}_{t+1}(x, z) = c(z-\alpha)^\sigma \mathcal{U}(x, z)$ where $\mathcal{U}(x, z)$ is primitive (as an element of $\mathcal{F}[z][x]$). Choose a $b_1 \in \mathcal{F}$ such that $\deg(\mathcal{U}_z(x, z)) = \deg(\mathcal{U}_z(b_1, z))$ and let $b_2 \in \mathcal{F}$ be such that it is *not* a root of $\mathrm{Res}_z(\mathcal{U}(x, z), \mathcal{U}(b_1, z))$ where $\mathrm{Res}_z(\mathcal{U}(x, z), \mathcal{U}(b_1, z))$ is the Sylvester resultant of $\mathcal{U}(x, z), \mathcal{U}(b_1, z) \in \mathcal{F}[x][z]$. Clearly, for such a choice of b_1, b_2, $\gcd(\mathcal{U}(b_1, z), \mathcal{U}(b_2, z)) = 1$ and the lemma follows. □

For a different condition relating the vanishing of a Wronskian to sparse representation, see [8, 11].

Algorithm Shifted-Sparse-Interpolation

Input: A polynomial $f(x) \in \mathcal{F}[x]$ of degree d and an integer $t \leq (d+1)/2$.

Output: $\alpha \in \mathcal{F}$, and a list of pairs $(\mathtt{F}_i, \delta_i), i = 1, \ldots, t$ such that $f(x) = \sum_{i=1}^{t} \mathtt{F}_i (x-\alpha)^{\delta_i}$, or, the message "$f(x)$ does not have a t-sparse shift for the given t."

1. *Compute* $\gcd(\mathcal{W}_{t+1}(b_1, z), \mathcal{W}_{t+1}(b_2, z))$ *for generic* $b_1, b_2 \in \mathcal{F}$.
   - *if the* gcd $\in \mathcal{F}$, *a t-sparse shift does not exist for $f(x)$.*
   - *if the square-free part of the* gcd $= c(z-\alpha)$, *α is the unique t-sparse shift; goto step 2.*
   - *if the square-free part of the* gcd *has degree* > 1, *$f(x)$ has a t-sparse shift; however, the number of terms in the shifted basis is strictly less than t. Find the smallest principle minor $W_{k+1}(x, z)$ of $W_{t+1}(x, z)$ such that the square-free part of the gcd of* $\det(W_{k+1}(b_1, z))$ *and* $\det(W_{k+1}(b_2, z))$ *is linear in z, set t equal to k and goto step 2.*
2. *Solve the system of linear equations* (5) *to get the auxiliary polynomial* $\Phi(z) = z^t + \phi_{t-1} z^{t-1} + \ldots + \phi_1 z + \phi_0$.
3. *Find the integer roots of* $\Phi(z)$. *The roots are* $\delta_1, \delta_2, \ldots, \delta_t$.
4. *Solve the linear system of equations* $\mathcal{V}_t(b_1, \alpha)\vec{\mathsf{f}} = \vec{\mathsf{a}}$ *to obtain the coefficients* $\mathtt{F}_1, \ldots, \mathtt{F}_t$ *of f. Output the list of pairs* $[(\mathtt{F}_1, \delta_1), \ldots, (\mathtt{F}_t, \delta_t)]$.

Remark: This algorithm can be used even when $t > (d+1)/2$. In such cases, the shift might not be unique and might be algebraic over $\mathcal{F}$. The third option in step 1 (the square-free part of the gcd has *degree* > 1) must be modified to take care of multiple shifts.

Analysis of the running time of the algorithm: We count the number of rational arithmetic operations performed by the above algorithm assuming that *classical algorithms* are used for doing polynomial arithmetic (for multiplication, gcd etc). In step 1, setting up the matrix $W_{t+1}(b_1, z)$ requires evaluating the first $2t$ derivatives of $f(x)$ at $x = b_1$ and $O(t^2 d)$ additional work. The $2t$ derivatives can be evaluated in $O(td)$ operations. The polynomial $\mathcal{W}_{t+1}(b_1, z)$ is of degree at most $\binom{t+1}{2}$ in z (this is the degree of the anti-diagonal term in $W_{t+1}(b_1, z)$). To compute this by interpolation, one needs its values at $\binom{t+1}{2}+1$ distinct values of z and each value is to be obtained by evaluating a $(t+1) \times (t+1)$ determinant. The total work involved in this is $O(t^5)$ operations (and a similar number for computing $\mathcal{W}_{t+1}(b_2, z)$). Since they are of degree $\binom{t+1}{2}$ each, their gcd can be computed using $O(\binom{t+1}{2}^2) = O(t^4)$ operations. The square-free part of the gcd can be computed within the same bounds on the number of operations. Step 2 can be performed in $O(t^3)$ operations, step 3 in $O(t^2 \log d)$ operations [20]. The linear system to be solved in Step 4 is a very special, Vandermonde-like system and can be solved in $O(t^2)$ steps. Clearly, step 1 dominates for an overall count of $O(t^2 d + t^5)$. The polynomial $f(x)$ and its first $2t$ derivatives need to be evaluated at 2 points b_1, b_2 for a total of $4t + 2$ evaluations. In contrast, the algorithm of Grigoriev and Karpinski needs $d^{O(1)}$ evaluations and its running time is dominated by $O(td^4)$. We can summarize the complexity of the algorithm as follows:

Theorem 3 *The shifted sparse interpolation algorithm computes the shifted t-sparse representation, if one exists, for a univariate polynomial f with coefficients in a field $\mathcal{F}$ (characteristic 0), and degree d using $O(t^2 d + t^5)$ $\mathcal{F}$-operations. The algorithm uses randomization with controllably small probability of error.* □

3. On Distinguishing Two Shifted Sparse Polynomials

If we are given black boxes for two polynomials that are known to be t-sparse with respect to some bounded, but unknown and possibly different shifts, the test developed in this section can be used to distinguish the two polynomials by evaluating them at no more than $3t$ sufficiently large points. The test applies to polynomials with real coefficients and it is based on the number of sign variations in the representation of such polynomials in the standard power basis.

We address the question of how often two shifted sparse polynomials (with possibly different sparse shifts) can have the same value at real points away from their sparse shifts. Let $H(x) = \sum_{i=1}^{t} h_i x^i$ be t-sparse in the standard power basis. Let $G(x) = H(x + a) = \sum_{i=0}^{d} g_i x^i$ for some positive real a with d being the degree of $H(x)$. We have

Lemma 4 *The number of sign variations, var(G), in the coefficient sequence of $G(x)$ is at most $t-1$.*

Proof: From Descartes' rule of signs, we know that $\mathrm{var}(H) - \mathrm{var}(G) = r + 2q$ where r is the number of real roots of $H(x)$ in the interval $(0, a]$ and q is a non-negative integer [3]. This implies that $\mathrm{var}(H) \geq \mathrm{var}(G)$. Since $H(x)$ is t-sparse, $\mathrm{var}(H) \leq t-1$, and hence, $\mathrm{var}(G) \leq t-1$. □

Let $f(x)$ be t-sparse in the α shifted power basis and $g(x)$ be t-sparse in the β-shifted power basis. Without loss of generality, assume that $\beta \leq \alpha$ and let $\gamma = \alpha - \beta$. With this understanding, we have,

Lemma 5 *The number of real roots of $g(x) - f(x)$ in the interval (α, ∞) is at most 3t-1.*

Proof: Let $f(x) = \sum_{i=1}^{t} f_i(x-\alpha)^{\delta_i}$ and $g(x) = \sum_{i=1}^{t} g_i(x-\beta)^{\epsilon_i}$. Denote by $\widehat{f}$ the polynomial $\sum_{i=1}^{t} f_i x^{\delta_i}$ and by $\widehat{g}$ the polynomial $\sum_{i=1}^{t} g_i(x+\gamma)^{\epsilon_i} = \sum_{j=0}^{\epsilon} \widehat{g}_i x^i$ where $\epsilon = \deg(g(x))$ so that $f(x) = \widehat{f}(x-\alpha)$ and $g(x) = \widehat{g}(x-\alpha)$. By lemma 4, the number of sign variations in the coefficients of $\widehat{g}(x)$, $\mathrm{var}(\widehat{g}(x))$ is less than t. Therefore, the number of sign variations in $\widehat{g}(x) - \widehat{f}(x)$ is less than $3t$ ($\widehat{f}(x)$ has at most t non-zero coefficients – when we "splice in" $\widehat{f}(x)$ into $\widehat{g}(x)$, each coefficient of $\widehat{f}(x)$ can produce two new sign variations in the difference $\widehat{g}(x) - \widehat{f}(x)$). Now, invoke Descartes' rule of signs on the difference $\widehat{g}(x) - \widehat{f}(x)$, to conclude that $\widehat{g}(x) - \widehat{f}(x)$ has fewer than $3t$ positive real roots, or, $g(x) - f(x)$ has fewer than $3t$ roots $> \alpha$. □

Hence, we have

Theorem 4 *Let $f(x), g(x)$ be polynomials with real coefficients each of which is t-sparse with respect to a real shift whose magnitude is bounded by some B. Let $\mathcal{A} = \{a_i\}$ be a set of $3t$ distinct real numbers greater than B. If $f(a_i) = g(a_i)$ for $a_i \in \mathcal{A}$, then, $f(x) = g(x)$.* □

4. An algorithm for finding sparse decompositions

Let us recall some facts about univariate polynomial decomposition from [17]. A decomposition of $f(x) \in \mathcal{F}[x]$ is a sequence of polynomials $f_1, f_2, \ldots, f_k \in \mathcal{F}[x]$ such that $f(x) = f_1 \circ f_2 \circ \ldots \circ f_k(x)$. If there is no such decomposition with *at least* two of the f_i being of degree 2 or more, then, $f(x)$ is said to be indecomposable. A complete decomposition of $f(x)$ is a decomposition in which each component $f_i(x)$ has degree at least 2 and is indecomposable. A theorem of Ritt guarantees that a complete decomposition is "essentially"

[3] We thank our colleague Prof. Jeremy Johnson for this observation.

unique for polynomials over fields of characteristic zero. The following types of ambiguities can occur in a complete decomposition:

$$f \circ g(x) = f(x+a) \circ (g(x) - a),\ x^n \circ x^m = x^m \circ x^n,\ T_n(x) \circ T_m(x) = T_m(x) \circ T_n(x),$$

where $T_j(x)$ is the j-th Chebyshev polynomial of the first kind. However, these are the only possible ambiguities [3,23]. In [4], an efficient algorithm for computing a complete decomposition of a polynomial $f(x) \in \mathcal{F}[x]$ is described (the authors treat coefficient fields of characteristic ≥ 0).

We call a decomposition such as $f(x) = g_1 \circ g_2 \circ \ldots \circ g_r(x)$ t-sparse if each component g_i is t-sparse. Note that such a decomposition need not be a complete decomposition and some of the g_i are allowed to have degree one. The sparse decomposition problem is to find a t-sparse decomposition for a given $f(x)$ and a t, if such a decomposition exists. The algorithm of von zur Gathen et al guarantees that if $f(x) \in \mathcal{F}[x]$ is decomposable over some algebraic extension of $\mathcal{F}$, then it is decomposable over $\mathcal{F}$. This, however, is not the case with sparse decomposition – for a given t, there may exist a t-sparse decomposition of $f(x)$ over an algebraic extension of $\mathcal{F}$, but not over $\mathcal{F}$.

Our algorithm involves repeated use of the algorithm of section 2, scanning for blocks in the complete decomposition of f that may be shifted t-sparse, taking advantage of the fact that composition in general does not commute. The ambiguity arising from the commutativity of composition of Chebyshev polynomials is dealt with separately (for example, consider $T_2(T_3(h))$ which is the same as $T_3(T_2(h))$; it may happen that $T_2(h)$ is t-sparse while $T_3(h)$ is not).

Consider a complete decomposition of f such as

$$f(x) = f_1 \circ f_2 \circ \ldots \circ f_m(x). \tag{6}$$

Identify those f_i which might be Chebyshev polynomials (this can be done in several simple ways, for instance, using the fact that the only polynomial solutions of the differential equation $(1-x^2)(dy/dx)^2 = m^2(1-y^2)$ are $\pm T_m(x)$).

- If $f_u, f_{u+1}, \ldots, f_v$ are all Chebyshev polynomials for some $u < v$ in (6) above, and f_{u-1}, f_{v+1} are not, we call $f_u, f_{u+1}, \ldots, f_v$ a Chebyshev block and denote it by $C_{u,v}$.
- By $\mathcal{T}_{u,v}$, we denote the power set of $\{u, u+1, \ldots, v\}$. For any index set $J \in \mathcal{T}_{u,v}$, let T_J denote the Chebyshev polynomial $f_{u_1} \circ f_{u_2} \circ \ldots \circ f_{u_l}$ where $J = \{u_1, u_2, \ldots, u_l\}$. Note that the degree of T_J is $d_J = \prod_{u_j \in I} \deg(f_{u_j})$.
- By $h_{i,j}$, we denote $f_i \circ f_{i+1} \circ \ldots \circ f_j(x)$ for $i \leq j$ if f_i, f_j are not Chebyshev polynomials.
- If f_i isn't Chebyshev, and f_j belongs to the Chebyshev block $C_{j,k}$ for some $k \geq j > i$, then, for any $J \in \mathcal{T}_{j,k}$, let $h_{i,J}$ denote $f_i \circ f_{i+1} \circ \ldots \circ f_{j-1} \circ T_J(x)$; if f_i belongs to the

Chebyshev block $C_{l,i}$ for some $l \leq i$, and f_j isn't Chebyshev, then, for any $I \in \mathcal{T}_{l,i}$, let $h_{I,j}$ denote $T_I \circ f_{i+1} \circ \ldots \circ f_{j-1} \circ f_j(x)$.

- If f_i, f_j are both Chebyshev but are in different Chebyshev blocks $C_{l,i}, C_{j,k,}$, then for any $I \in \mathcal{T}_{l,i}$ and $J \in \mathcal{T}_{j,k}$, let $h_{I,J}$ denote $T_I \circ f_{i+1} \circ \ldots \circ f_{j-1} \circ T_J(x)$.

We look for blocks in the complete decomposition of f that have shifted sparse representations.

For each $h_{i,j}$, if there is a shift $\alpha \in \mathcal{K}$ such that $h_{i,j}$ (respectively $h_{i,J}$, $h_{I,j}$, $h_{I,J}$) is t-sparse in the α-shifted basis, then, denote the t-sparse representation of $h_{i,j}$ in the α shifted basis by $g_{i,j}$ (respectively $g_{i,J}$, $g_{I,j}$, $g_{I,J}$).

If no t-sparse shift exists for $h_{i,j}$, check if there is a shift $\alpha \in \mathcal{K}$ such that $h_{i,j}$ is $t+1$-sparse in the α-shifted basis *with a non-zero constant term b*. If such a shift exists, then, set $g_{i,j}$ to $(x+b) \circ G_{i,j}$ where $G_{i,j}$ is the t-sparse representation of $h_{i,j} - b$ in the α shifted basis. If no such shift exists, set $g_{i,j}$ to zero (similarly for $g_{i,J}$, $g_{I,j}$, $g_{I,J}$).

The $g_{i,j}, g_{i,J}, g_{I,j}, g_{I,J}$ are the possible sparse components of f. In order to find out if f indeed has a t-sparse decomposition, we construct a directed acyclic graph $G = (V, E)$. It has the following vertices:

- a vertex labeled i for each f_i in the complete decomposition of f that is not in any Chebyshev block; a vertex labeled $m+1$;
- for each Chebyshev block $C_{u,v}$, vertices labeled $V_{u,v}(I), U_{u,v}(I)$ for each $I \in \mathcal{T}_{u,v}$.

The graph has the following labeled edges:

- an edge directed from i to j if $g_{i,j-1} \neq 0$, labeled $g_{i,j-1}$.
- if f_{j-1} is not Chebyshev but $f_j \in C_{j,k}$ for some k, and $g_{i,j-1} \neq 0$, an edge directed from i to $V_{j,k}(\phi)$ labeled $g_{i,j-1}$; if $f_k \in C_{j,k}$ for some j,k but f_{k+1} is not Chebyshev, an edge directed from $U_{j,k}(\phi)$ to $k+1$ with *no* label.
- if $g_{i,J} \neq 0$ for $J \in \mathcal{T}_{j,k}$ for some j,k, an edge directed from i to $V_{j,k}(J)$ labeled $g_{i,J}$; if $g_{I,j} \neq 0$ for $I \in \mathcal{T}_{k,i}$ for some k,i, an edge directed from $U_{k,i}(J)$ to j labeled $g_{I,j}$.
- if $g_{I,J} \neq 0$ for some $I \in \mathcal{T}_{l,i}$ and $J \in \mathcal{T}_{j,k}$, an edge directed from $U_{l,i}(I)$ to $V_{j,k}(J)$ labeled $g_{I,J}$.
- for every $(I,J) \in \mathcal{T}_{i,j} \times \mathcal{T}_{i,j}$, an edge directed from $V_{i,j}(I)$ to $U_{i,j}(J)$ if $I \cap J = \phi$ and for each $k \in \mathcal{T}_{i,j} \setminus (I \cup J)$, $g_{k,k} \neq 0$. The edge is labeled $g_{k_1,k_1} \circ \ldots \circ g_{k_r,k_r}$ (in any order) where $\{k_1, \ldots, k_r\} = \mathcal{T}_{i,j} \setminus (I \cup J)$.

We look for a simple path from the vertex labeled 1 to $m+1$ if f_1 is not in any Chebyshev block (if f_1 is in the Chebyshev block $C_{1,j}$, we look for a simple path from $V_{1,j}(\phi)$ to $m+1$). The labels of successive edges in such a path give a t-sparse decomposition of f.

Algorithm sparse decomposition:

1. *Find a complete decomposition of $f(x)$ into $f(x) = f_1 \circ f_2 \circ \ldots \circ f_m(x)$ with $f_i \in \mathcal{F}[x]$ using the algorithm of von zur Gathen et al. Identify the various Chebyshev blocks $C_{u,v}$ in the complete decomposition.*

2. *For each i, j $1 \leq i \leq j \leq m$, compute $h_{i,j}$, or all possible $h_{i,J}$, or $h_{I,j}$, or $h_{I,J}$ as the case may be (purge any duplicates in the case of $h_{i,J}, h_{I,j}$, and $h_{I,J}$).*

3. *For each $h_{i,j}, h_{i,J}, h_{I,j}$, and $h_{I,J}$, compute the corresponding $g_{i,j}, g_{i,J}, g_{I,j}$, and $g_{I,J}$.*

4. *Construct the directed graph $G = (V, E)$.*

5. *Find a simple path from the vertex labeled 1 to $m+1$ if f_1 is not in any Chebyshev block; if f_1 is in the Chebyshev block $C_{1,j}$, find a simple path from $V_{1,j}(\phi)$ to $m+1$. If such a path exists and the successive edges are labeled as $L_1, \ldots, L_r$, then return $L_1 \circ \ldots \circ L_r$ as a t-sparse decomposition of f. If no such path exists, f has no t-sparse decomposition.*

Suppose f has a t-sparse decomposition $f = g_1 \circ g_2 \circ \ldots \circ g_s$. Consider the complete decomposition of each g_i. Because of the uniqueness of the complete decomposition of f, the polynomials in the complete decomposition of each g_i must correspond to a block of indecomposable polynomials $f_l, f_{l+1}, \ldots, f_j$ in (6) (upto possible reordering of Chebyshev polynomials if they occur at the ends of the block and shifts such as $(x+a) \circ (x-a)$ introduced in the middle). The algorithm checks for all possible ways to split (6) into blocks and existence of t-sparse shifts for each possible split of (6). Hence, it will find a t-sparse decomposition of f if there is one.

Analysis of the running time of the algorithm: The simple observation required for showing that the algorithm runs in polynomial time (in the degree d of f) is that the number of composition factors in the complete decomposition of f (m above) is $\leq \log_2 d$. In step 1, we use the algorithm of von zur Gathen et al to find the complete decomposition of f and for each component f_i, we check whether it is a Chebyshev polynomial. Using classical polynomial arithmetic, both these tasks can be achieved in $O(d^2)$ rational arithmetic operations. Since $\prod_{i=1}^{m} \deg(f_i) = d$ and $\deg(f_i) > 1$, $m \leq \log_2 d$. Therefore, the union of all the power-sets $\mathcal{T}_{u,v}$ corresponding to the various Chebyshev blocks contains no more than d elements. Finding t-sparse representations for each $h_{i,j}$ or $h_{i,j} - b$ in steps 2 and 3 costs $O(d(t^2 d + t^5))$ "field" operations in all, where the "field" may be an algebraic extension of $\mathcal{F}$ of degree $O(d)$ (we are putting no constraints on t relative to d and as a result, t-sparse shifts need not be rational).

Again, using classical polynomial arithmetic, the number of $\mathcal{F}$-operations for steps 2 and 3 can be bounded by $O(t^2 d + t^5)d^3$. The graph G has at most $O(d)$ vertices and $O(d^2)$ edges. Constructing and traversing it costs $O(d^2)$ basic operations. The total cost is clearly dominated by steps 2 and 3 for an overall operation count of $O(t^2 d^3 (d + t^3))$. In summary, we have

Theorem 5 *Given a polynomial f with coefficients in a field $\mathcal{F}$ (characteristic 0), and degree d and a positive integer t, the sparse decomposition algorithm computes a t-sparse decomposition of f, if one exists. The algorithm uses randomization and performs $O(t^2 d^3 (d + t^3))$ $\mathcal{F}$-operations.*

Appendix

In this appendix, we prove lemma 2, i.e., that $\det(A)$ is non-zero where

$$A = \begin{pmatrix} p_1^{\underline{q_1}} & p_1^{\underline{q_2}} & \dots & p_1^{\underline{q_r}} \\ p_2^{\underline{q_1}} & p_2^{\underline{q_2}} & \dots & p_2^{\underline{q_r}} \\ \vdots & \vdots & \ddots & \vdots \\ p_r^{\underline{q_1}} & p_r^{\underline{q_2}} & \dots & p_r^{\underline{q_r}} \end{pmatrix}$$

with

$$p_1 > p_2 > \ldots > p_r > 0, \; q_1 > q_2 > \ldots > q_r = 0, \; p_l > q_l \text{ for } l = 1, \ldots, r.$$

If A is singular, then there is some $\widehat{f} = (\widehat{f}_1, \widehat{f}_2, \ldots, \widehat{f}_r)^{Tr}$, and an $s \leq r$ such that $\widehat{f}_s \neq 0, \widehat{f}_{s+1} = \ldots = \widehat{f}_r = 0$ and $A_s(\widehat{f}_1, \ldots, \widehat{f}_s)^{Tr} = 0$ where A_s is the $s \times s$ submatrix made from the first s rows and columns of A. Looking at it differently, this means that the polynomial $\zeta(z) = \widehat{f}_1 z^{\underline{q_1}} + \widehat{f}_2 z^{\underline{q_2}} + \ldots + \widehat{f}_s z^{\underline{q_s}}$ has $p_1, p_2, \ldots, p_s$ as its roots. However, as the polynomial $\zeta(z)$ can have at most $s - 1$ sign variations in the above representation, it cannot have s roots satisfying conditions above. This follows from an analogue of Descartes' rule of signs for polynomials expressed in the falling Pochhammer basis.

Lemma 6 *Let $h(z) = h_n z^{\underline{n}} + h_{n-1} z^{\underline{n-1}} + \ldots + h_1 z^{\underline{1}} + h_0$ be a polynomial with real coefficients. Let $a > n$ and let $g(z) = (z - a)h(z) = g_{n+1} z^{\underline{n+1}} + g_n z^{\underline{n}} + \ldots + g_1 z^{\underline{1}} + g_0$. Then, the sequence of coefficients $\{g_{n+1}, g_n, \ldots, g_1, g_0\}$ has* at least *one more sign variation than the sequence of coefficients $\{h_n, h_{n-1}, \ldots, h_1, h_0\}$.*

Proof: The proof is similar to the classical elementary proof of Descartes' rule of signs. Consider the product

$$\begin{aligned}(z + 1 - a - 1)h(z) = h_n (z+1)^{\underline{n+1}} &+ h_{n-1}(z+1)^{\underline{n}} + \ldots h_1 (z+1)^{\underline{2}} + h_0 (z+1)^{\underline{1}} \\ &- (a+1) h_n z^{\underline{n}} - \ldots - (a+1) h_1 z^{\underline{1}} - (a+1) h_0.\end{aligned}$$

Using the identity $(z+1)^{\underline{n}} = z^{\underline{n}} + nz^{\underline{n-1}}$, rewrite the above product as

$$\begin{aligned} h_n z^{\underline{n+1}} &+ (h_{n-1} + (n-a)h_n)z^{\underline{n}} + (h_{n-2} + (n-1-a)h_{n-1})z^{\underline{n-1}} \\ &+ \ldots + (h_0 + (1-a)h_1)z^{\underline{1}} + (-a)h_0. \end{aligned}$$

If h_i, h_{i-1} have different signs, then, $h_{i-1} + (i-a)h_i$ has the sign of h_{i-1} (since $a \geq n$, $i - a \leq 0$ for $i = 0, \ldots, n$). The next sign variation in the coefficients of $h(z)$ will create a coefficient of a different sign than that of $h_{i-1} + (i-a)h_i$ in the product. In other words, the sequence

$$h_n,\ h_{n-1} + (n-a)h_n,\ \ldots,\ h_0 + (1-a)h_1 \tag{7}$$

has at least as many variations of sign as the sequence $h_n, h_{n-1}, \ldots, h_1, h_0$. However, the last change of sign in $h_n, h_{n-1}, \ldots, h_1, h_0$ introduces at least one more sign change into the sequence (7), ensuring at least one extra sign variation in the coefficients of $g(z)$. □

Lemma 7 *Let $g(z) = g_n z^{\underline{n}} + g_{n-1} z^{\underline{n-1}} + \ldots + g_1 z^{\underline{1}} + g_0$, $g_0 \neq 0$ be a polynomial with real coefficients with integer roots $a_1, a_2, \ldots, a_k$ such that $a_1 \geq 1, a_2 \geq a_1 + 1, \ldots, a_k \geq a_{k-1} + 1$. Then, the sequence of coefficients $\{g_n, g_{n-1}, \ldots, g_1, g_0\}$ has* at least *k sign variations.*

Proof: Let $g(z) = (z - a_k)(z - a_{k-1}) \ldots (z - a_2)(z - a_1)h(z)$.

Case 1 $(a_1 > n - k)$: Apply lemma 6 repeatedly to

$$h(z), (z - a_1)h(z), \ldots, (z - a_{k-1}) \ldots (z - a_2)(z - a_1)h(z).$$

Case 2 $(a_1 \leq n - k)$: Let $(z - a_1)h(z) = H(z)$ where $H(z) = H_{n-k+1} z^{\underline{n-k+1}} + H_{n-k} z^{\underline{n-k}} + \ldots + H_1 z^{\underline{1}} + H_0$. Let $a_1 = l$. Therefore, $a_1^{\underline{l}} \neq 0$ and $a_1^{\underline{l+i}} = 0$ for any $i > 0$. This implies that a_1 is a root of the polynomial $H_l z^{\underline{l}} + H_{l-1} z^{\underline{l-1}} + \ldots + H_1 z^{\underline{1}} + H_0$ (this is not identically zero as at least $H_0 \neq 0$ since $g_0 \neq 0$) By lemma 6, the sequence of coefficients of this polynomial has at least one sign variation which carries over to $H(z)$, i.e., $H(z)$ has at least one sign variation.

Repeated applications of the above 2 cases with the assumption that $a_{i+1} \geq a_i + 1$, lead us to conclude the lemma. □

Now, we apply the last lemma to prove the non-singularity of A. If it is singular, then pick an element $b \neq 0$ in the kernel of A; Let b_j be the last non-zero entry in b (i.e., $b_{j+1} = \ldots = b_r = 0$). The projection $(b_1, ..., b_j)^{Tr}$ is in the kernel of the $j \times j$ principle submatrix of A. From this, it follows that $p_1 - q_j, p_2 - q_j, \ldots, p_j - q_j$ are all roots of the polynomial $\psi(z) = b_1 z^{\underline{q_1 - q_j}} + b_2 z^{\underline{q_2 - q_j}} + \ldots + b_j$. By lemma 7, the polynomial $\psi(z)$ must have at least j sign variations. But $\psi(z)$ has at most j terms and hence, can have at most $j - 1$ sign variations – a contradiction; therefore, the kernel of A is trivial, i.e., A is non-singular.

References

1. Ben Or, M. and Tiwari, P. (1988), "A deterministic algorithm for sparse multivariate polynomial interpolation," Proc. 20th Symp. Theory of Computing, ACM Press, pp. 301–309.
2. Borodin, A. and Tiwari, P. (1990), "On the decidability of sparse univariate polynomial interpolation," Proc. 22nd Symp. Theory of Computing, ACM Press, pp. 535–545.
3. Fried, M.D., and MacRae, R.E., (1969), "On the invariance of chains of fields," Ill. Jour. Math., **13**, 165–171.
4. von zur Gathen, J., Kozen, D., and Landau, S. (1987), "Functional decomposition of polynomials," Proc. 28th IEEE Symp. Found. Comp. Sci., Nov 1987, pp. 127–131.
5. Clausen, M., Dress, A., Grabmeier, J., Karpinski, M. (1988), "On zero testing and interpolation of k-sparse multivariate polynomials over finite fields," TR 88.06.006, IBM Germany, Heidelberg Scientific Center, June 1988.
6. Grigoriev, D. Yu. and Karpinski, M. (1987), "The matching problem for bipartite graphs with polynomially bounded permanents is in NC," Proc. 28th IEEE Symp. Foundations Comp. Sci., pp. 166–172.
7. Grigoriev, D., Karpinski, M., and Singer, M. (1990), "Fast parallel algorithms for sparse multivariate polynomial interpolation over finite fields," SIAM J.Comp., Vol. 19, pp. 1059–1063.
8. Grigoriev, D., Karpinski, M., and Singer, M., "Computational complexity of sparse rational interpolation," SIAM J.Comp., Vol. 23, pp. 1–11.
9. Grigoriev, D., Karpinski, M., and Singer, M., "Computational complexity of sparse real algebraic function interpolation," Proc. MEGA '92, Progress in Mathematics, Birkhauser, Vol. 109, pp. 91–104.
10. Grigoriev, D., Karpinski, M., and Singer, M., "The interpolation problem for k-sparse sums of eigenfunctions of operators," Advances in Appl Math, Vol 12, pp. 76-81.
11. Grigoriev, D., and Karpinski, M., "A zero-test and an interpolation algorithm for the shifted sparse polynomials," Proc. AAECC–93, Lect. Notes in Comp. Sci., Vol. 673, pp. 162–169.
12. Grigoriev, D., Karpinski, M., and Odlyzko, A.M.(1992a), "Existence of short proofs of non-divisibility of sparse polynomials under the extended Riemann hypothesis," Proc. ISSAC 92, ACM Press, pp. 117-122.
13. Kaltofen, E. (1987), "Single-factor Hensel lifting and its application to the straight-line complexity of certain polynomials," Proc. 19th Symp. Theory of Computing, ACM Press, pp. 443–452.
14. Kaltofen, E. and Lakshman, Y.N. (1988), "Improved sparse multivariate polynomial interpolation algorithms," Proc. ISSAC 1988, Rome, Italy, Springer-Verlag LNCS **358**, pp. 467–474.
15. Kaltofen, E., and Trager, B. (1990), "Computing with polynomials given by black boxes for their evaluations: Greatest common divisors, factorization, separation of numerators and denominators," J. Symb. Comp. **9** (1990), pp. 301-320.
16. Kaplanski, I., "An introduction to differential algebra," Hermann, Paris.
17. Kozen, D., and Landau, S., "Polynomial decomposition algorithms," JSC, Vol. 7, No. 5, pp. 445–456.
18. Lakshman, Y.N., and Saunders, B.D., "Sparse polynomial interpolation in non-standard bases," SIAM J. Comp., Vol. 24, No. 2, pp. 387-397.
19. Lakshman, Y.N., and Saunders, B.D., "On computing sparse shifts for univariate polynomials," Proc. ISSAC 1994, Oxford, OK, ACM Press.
20. Loos, R., "Computing rational zeros of integral polynomials by p-adic expansion," SIAM J.Comp., Vol. 12, pp. 286–293.
21. Mansour, Y., "Randomized interpolation and approximation of sparse polynomials ," SIAM J. Comp., Vol. 24, No. 2, pp. 357–368.
22. Muir, T. (1960, enlarged by Metzler, H.), "A treatise on the theory of determinants,", Dover Publishing Inc., New York.
23. Ritt, J.F., (1922), "Prime and composite polynomials," Trans. Amer. Math. Soc. **23**, pp. 51–66.
24. Zippel, R. (1990), "Interpolating polynomials from their values," Jour. Symb. Comp., Vol. 9, No. 3, pp. 375–403.