Solving Algebraic Equations in Terms of $\mathcal{A}$-Hypergeometric Series

Bernd Sturmfels
Department of Mathematics
University of California
Berkeley, CA 94720, U.S.A.
bernd@math.berkeley.edu

Abstract

The roots of the general equation of degree n satisfy an $\mathcal{A}$-hypergeometric system of differential equations in the sense of Gel'fand, Kapranov and Zelevinsky. We construct the n distinct $\mathcal{A}$-hypergeometric series solutions for each of the 2^{n-1} triangulations of the Newton segment. This works over any field whose characteristic is relatively prime to the lengths of the segments in the triangulation.

1. Solving the Quintic

A classical problem in mathematics is to find a formula for the roots of the general equation of degree n in terms of its $n+1$ coefficients. While there are formulas in terms of radicals for $n \leq 4$, Galois theory teaches us that no such formula exists for the general quintic

$$a_5 x^5 \;+\; a_4 x^4 \;+\; a_3 x^3 \;+\; a_2 x^2 \;+\; a_1 x \;+\; a_0 \quad = \quad 0. \tag{1.1}$$

An alternative approach is to expand the roots into fractional power series (or Puiseux series). In 1757 Johann Lambert expressed the roots of the trinomial equation $x^p + x + r$ as a Gauss hypergeometric function in the parameter r. Series expansions of more general algebraic functions were subsequently given by Euler, Chebyshev and Eisenstein, among others. The poster "Solving the Quintic with Mathematica" [12] gives a nice introduction to these classical techniques and underlines their relevance for symbolic computation.

The state of the art in the first half of our century appears in works of Richard Birkeland [2] and Karl Mayr [10]. They proved that the roots are multivariate hypergeometric functions (in the sense of Horn) in all of the coefficients and they gave series expansions for the roots and their powers. The purpose of this note is to refine the these results.

Our point of departure is the fact that the roots satisfy the $\mathcal{A}$-hypergeometric differential equations introduced by Gel'fand, Kapranov and Zelevinsky [6],[7]. Here $\mathcal{A}$ denotes the configuration of $n+1$ equidistant points on the affine line. It follows from recent work of McDonald [11] that there are 2^{n-1} distinct complete sets of series solutions, one for each of the 2^{n-1} triangulations of $\mathcal{A}$. We shall construct these series solutions explicitly.

Let us illustrate our general construction for the example of the quintic (1.1). Here the set $\mathcal{A}$ has 16 distinct triangulations. The finest triangulation divides $\mathcal{A}$ into five segments of unit length. The coarsest triangulation of $\mathcal{A}$ is just a single segment of length 5.

For the finest triangulation we get the following expressions for the five roots of (1.1):

$$X_{1,-1} \;=\; -[\frac{a_0}{a_1}], \qquad X_{2,-1} \;=\; -[\frac{a_1}{a_2}] \,+\, [\frac{a_0}{a_1}], \qquad X_{3,-1} \;=\; -[\frac{a_2}{a_3}] \,+\, [\frac{a_1}{a_2}],$$
$$X_{4,-1} \;=\; -[\frac{a_3}{a_4}] \,+\, [\frac{a_2}{a_3}], \qquad X_{5,-1} \;=\; -[\frac{a_4}{a_5}] \,+\, [\frac{a_3}{a_4}].$$

Each bracket represents a power series having the monomial in the bracket as its first term:

$$\begin{aligned}
[\frac{a_0}{a_1}] &= \frac{a_0}{a_1} + \frac{a_0^2 a_2}{a_1^3} - \frac{a_0^3 a_3}{a_1^4} + 2\frac{a_0^3 a_2^2}{a_1^5} + \frac{a_0^4 a_4}{a_1^5} - 5\frac{a_0^4 a_2 a_3}{a_1^6} - \frac{a_0^5 a_5}{a_1^6} + \cdots \\
[\frac{a_1}{a_2}] &= \frac{a_1}{a_2} + \frac{a_1^2 a_3}{a_2^3} - \frac{a_1^3 a_4}{a_2^4} - 3\frac{a_0 a_1^2 a_5}{a_2^4} + 2\frac{a_1^3 a_3^3}{a_2^5} + \frac{a_1^4 a_5}{a_2^5} - 5\frac{a_1^4 a_3 a_4}{a_2^6} + \cdots \\
[\frac{a_2}{a_3}] &= \frac{a_2}{a_3} - \frac{a_0 a_5}{a_3^2} - \frac{a_1 a_4}{a_3^2} + 2\frac{a_1 a_2 a_5}{a_3^3} + \frac{a_2^2 a_4}{a_3^3} - \frac{a_2^3 a_5}{a_3^4} + 2\frac{a_2^3 a_4^2}{a_3^5} + \cdots \\
[\frac{a_3}{a_4}] &= \frac{a_3}{a_4} - \frac{a_2 a_5}{a_4^2} + \frac{a_3^2 a_5}{a_4^3} + \frac{a_1 a_5^2}{a_4^3} - 3\frac{a_2 a_3 a_5^2}{a_4^4} - \frac{a_0 a_5^3}{a_4^4} + 4\frac{a_1 a_3 a_5^3}{a_4^5} + \cdots \\
[\frac{a_4}{a_5}] &= \frac{a_4}{a_5}
\end{aligned}$$

Note that the last bracket is just a single Laurent monomial. The other four brackets $\left[\frac{a_{i-1}}{a_i}\right]$ can easily be written as an explicit sum over $\mathbf{N}^4$. For instance,

$$[\frac{a_0}{a_1}] \;=\; \sum_{i,j,k,l \geq 0} \frac{(-1)^{2i+3j+4k+5l}\,(2i+3j+4k+5l)!}{i\,!\,j\,!\,k\,!\,l\,!\,(i+2j+3k+4l+1)!} \cdot \frac{a_0^{i+2j+3k+4l+1} a_2^i a_3^j a_4^k a_5^l}{a_1^{2i+3j+4k+5l+1}}$$

Each coefficient appearing in one of these series is integral. Therefore our five series solutions of the general quintic are characteristic-free. They work over any base field.

The situation is different for the coarsest triangulation of $\mathcal{A}$. Here we must assume that the characteristic is different from 5. The five series solutions of (1.1) are

$$X_{5,\xi} \;=\; \xi \cdot [\frac{a_0^{1/5}}{a_5^{1/5}}] \,+\, \frac{1}{5} \cdot \left(\xi^2 [\frac{a_1}{a_0^{3/5} a_5^{2/5}}] \,+\, \xi^3 [\frac{a_2}{a_0^{2/5} a_5^{3/5}}] \,+\, \xi^4 [\frac{a_3}{a_0^{1/5} a_5^{4/5}}] \,-\, [\frac{a_4}{a_5}] \right)$$

where ξ runs over the five roots of the equation $\xi^5 = -1$. The brackets denote the series

$$\begin{aligned}
[\frac{a_0^{1/5}}{a_5^{1/5}}] &= \frac{a_0^{1/5}}{a_5^{1/5}} - \frac{1}{25}\frac{a_1a_4}{a_0^{4/5}a_5^{6/5}} - \frac{1}{25}\frac{a_2a_3}{a_0^{4/5}a_5^{6/5}} + \frac{2}{125}\frac{a_1^2a_3}{a_0^{9/5}a_5^{6/5}} + \frac{3}{125}\frac{a_2a_4^2}{a_0^{4/5}a_5^{11/5}} + \cdots \\
[\frac{a_1}{a_0^{3/5}a_5^{2/5}}] &= \frac{a_1}{a_0^{3/5}a_5^{2/5}} - \frac{1}{5}\frac{a_3^2}{a_0^{3/5}a_5^{7/5}} - \frac{2}{5}\frac{a_2a_4}{a_0^{3/5}a_5^{7/5}} + \frac{7}{25}\frac{a_3a_4^2}{a_0^{3/5}a_5^{12/5}} + \frac{6}{25}\frac{a_1a_2a_3}{a_0^{8/5}a_5^{7/5}} + \cdots \\
[\frac{a_2}{a_0^{2/5}a_5^{3/5}}] &= \frac{a_2}{a_0^{2/5}a_5^{3/5}} - \frac{1}{5}\frac{a_1^2}{a_0^{7/5}a_5^{3/5}} - \frac{3}{5}\frac{a_3a_4}{a_0^{2/5}a_5^{8/5}} + \frac{6}{25}\frac{a_1a_2a_4}{a_0^{7/5}a_5^{8/5}} + \frac{3}{25}\frac{a_1a_3^2}{a_0^{7/5}a_5^{8/5}} + \cdots \\
[\frac{a_3}{a_0^{1/5}a_5^{4/5}}] &= \frac{a_3}{a_0^{1/5}a_5^{4/5}} - \frac{1}{5}\frac{a_1a_2}{a_0^{6/5}a_5^{4/5}} - \frac{2}{5}\frac{a_4^2}{a_0^{1/5}a_5^{9/5}} + \frac{1}{25}\frac{a_1^3}{a_0^{11/5}a_5^{4/5}} + \frac{4}{25}\frac{a_1a_3a_4}{a_0^{6/5}a_5^{9/5}} + \cdots
\end{aligned}$$

Each of these four series can be expressed as an explicit sum over the lattice points in a 4-dimensional polyhedron. The general formula will be presented in Theorem 3.2 below.

2. The Roots are $\mathcal{A}$-hypergeometric

Our problem is to compute the roots of the general equation of degree n,

$$f(x) \quad = \quad a_0 \,+\, a_1x \,+\, a_2x^2 \,+\, \ldots \,+\, a_{n-1}x^{n-1} \,+\, a_nx^n. \tag{2.1}$$

Each root of $f(x)$ is an algebraic function in the indeterminate coefficients:

$$X \quad = \quad X(a_0, a_1, a_2, \ldots, a_{n-1}, a_n).$$

Proposition 2.1. (Karl Mayr [10, p. 284]) *The roots of the general equation of degree n satisfy the following system of linear partial differential equations:*

$$\frac{\partial^2 X}{\partial a_i \partial a_j} = \frac{\partial^2 X}{\partial a_k \partial a_l} \qquad \textit{whenever} \quad i+j = k+l, \tag{2.2}$$

$$\sum_{i=0}^{n} i a_i \frac{\partial X}{\partial a_i} \;=\; -X \qquad \textit{and} \qquad \sum_{i=0}^{n} a_i \frac{\partial X}{\partial a_i} \;=\; 0. \tag{2.3}$$

The system (2.2)-(2.3) is a special instance of the class of $\mathcal{A}$-hypergeometric differential equations introduced by Gel'fand, Kapranov and Zelevinsky [5],[7]. Namely, (2.2)-(2.3) is the $\mathcal{A}$-hypergeometric system with parameters $\binom{-1}{0}$ associated with the integer matrix

$$\mathcal{A} \quad := \quad \begin{pmatrix} 0 & 1 & 2 & 3 & \cdots & n-1 & n \\ 1 & 1 & 1 & 1 & \cdots & 1 & 1 \end{pmatrix}. \tag{2.4}$$

The column vectors of $\mathcal{A}$ are homogeneous coordinates of $n+1$ equidistant points on the line. Their convex hull is a line segment: it is the Newton polytope of $f(x)$.

The Euler-type equations (2.3) follow readily from the homogeneity relations

$$\begin{aligned} X(a_0, ta_1, t^2a_2, \ldots, t^{n-1}a_{n-1}, t^n a_n) &= \frac{1}{t} \cdot X(a_0, a_1, a_2, \ldots, a_{n-1}, a_n), \\ X(ta_0, ta_1, ta_2, \ldots, ta_{n-1}, ta_n) &= X(a_0, a_1, a_2, \ldots, a_{n-1}, a_n). \end{aligned}$$

The equations (2.2) appeared in Karl Mayr's 1937 paper [10, equation (2) on page 284]. I shall present two proofs different proofs. The first one was shown to me in the spring of 1992 by Jean-Luc Brylinski. See [3] for an appearance of (2.2) in differential geometry.

Brylinsky's proof of (2.2): It uses implicit differentiation and works over any base field. We consider the first derivative $f'(x) = \sum_{i=1}^n i a_i x^{i-1}$ and the second derivative $f''(x) = \sum_{i=2}^n i(i-1)a_i x^{i-2}$. Note that $f'(X) \neq 0$, since $a_0, \ldots, a_n$ are indeterminates. Differentiating the defining identity $\sum_{i=0}^n a_i X(a_0, a_1, \ldots, a_n)^i = 0$ with respect to a_j, we get

$$X^j + f'(X) \cdot \frac{\partial X}{\partial a_j} = 0. \tag{2.5}$$

We next differentiate $\partial X / \partial a_j$ with respect to the indeterminate a_i:

$$\frac{\partial^2 X}{\partial a_i \partial a_j} = \frac{\partial}{\partial a_i}\left(-\frac{X^j}{f'(X)}\right) = \frac{\partial f'(X)}{\partial a_i} X^j f'(X)^{-2} - jX^{j-1}\frac{\partial X}{\partial a_i} f'(X)^{-1}. \tag{2.6}$$

Using (2.5) and the identity $\frac{\partial f'(X)}{\partial a_i} = -\frac{f''(X)}{f'(X)} \cdot X^i + iX^{i-1}$, we can rewrite (2.6) as

$$\frac{\partial^2 X}{\partial a_i \partial a_j} = -f''(X)X^{i+j} f'(X)^{-3} + (i+j)X^{i+j-1} f'(X)^{-2}. \tag{2.7}$$

The expression (2.7) depends only on the sum of indices $i+j$. This proves (2.2). ∎

A complex analysis proof of (2.2): Suppose we are working over the field of complex numbers **C**. Consider the logarithmic derivative $[log(f(x))]' = f'(x)/f(x)$. We view it as a rational function in $ahabe_0, a_1, \ldots, a_n$ and differentiate with respect to these variables:

$$\frac{\partial^2}{\partial a_i \partial a_j}[log(f(x))]' = \left(\frac{-x^{i+j}}{f^2(x)}\right)'.$$

This shows that $[log(f(x))]'$ satisfies the quadratic $\mathcal{A}$-hypergeometric equations (2.2). Proposition 2.1 follows by differentiating under the integral sign in Cauchy's formula

$$X = \frac{1}{2\pi i}\int_\Gamma \frac{z f'(z)}{f(z)} dz, \tag{2.8}$$

where Γ is a sufficiently small loop in the complex plane. ∎

3. Series Expansions

For any rational number u and any integer v we abbreviate

$$\gamma(u,v) \;:=\; \begin{cases} 1 & \text{if } v=0, \\ u(u-1)(u-2)\cdots(u+v+1) & \text{if } v<0, \\ 0 & \text{if } u \text{ is a negative integer and } u \geq -v, \\ \frac{1}{(u+1)(u+2)\cdots(u+v)} & \text{otherwise.} \end{cases}$$

If u is not a negative integer then $\gamma(u,v-1) = (u+v)\cdot\gamma(u,v)$ and we have $\gamma(u,v) = \Gamma(u+1)/\Gamma(u+v+1)$, where Γ is the usual Gamma function.

Let $\mathcal{L}$ denote the integer kernel of $\mathcal{A}$. This is the $(n-1)$-dimensional sublattice of $\mathbf{Z}^{n+1}$ spanned by $\{e_{i-1}-2e_i+e_{i+1} \,:\, i=1,\ldots,n-1\}$. Consider any monomial $a_0^{u_0}a_1^{u_1}\cdots a_n^{u_n}$ with rational exponents in the coefficients of $f(x)$. We define the formal power series

$$\left[a_0^{u_0}a_1^{u_1}\cdots a_n^{u_n}\right] \quad := \quad \sum_{(v_0,\ldots,v_n)\in\mathcal{L}} \prod_{i=0}^{n} \left(\gamma(u_i,v_i)\, a_i^{u_i+v_i}\right). \tag{3.1}$$

This series satisfies the linear differential equations (2.3) with their right hand sides $-X$ and 0 replaced by $\gamma_1 X$ and $\gamma_2 X$, and it satisfies the quadratic equations (2.2) in the following two cases. To prove the second case it is best to first derive formula (4.2) below.

Lemma 3.1. ([6, Lemma 1]) *Let $(u_0,u_1,\ldots,u_n)$ be a rational vector which either has no negative integer coordinate or has the form $(0,\ldots,0,1,-1,0,\ldots,0)$. Then the series $\left[a_0^{u_0}\cdots a_n^{u_n}\right]$ is a formal solution of the $\mathcal{A}$-hypergeometric system with parameters $\binom{\gamma_1}{\gamma_2} = \mathcal{A}\cdot(u_0,u_1,\ldots,u_n)^T$.*

Gel'fand, Kapranov and Zelevinsky constructed a complete set of series solutions for each regular triangulations of the set $\mathcal{A}$. We shall adapt their general construction to our special case. Extra care must be taken, however, because our equations (2.3) do not satisfy the non-resonance hypothesis which is necessary for [6, Theorem 3] to hold.

We write $0,1,\ldots,n$ for the points in our configuration $\mathcal{A}$ in (2.4). It has 2^{n-1} triangulations, all of which are regular. Each triangulation is indexed by a subset I of $\{1,\ldots,n-1\}$. Writing the complementary subset as $\{0,1,\ldots,n\}\backslash I = \{0=i_0<i_1<i_2<\cdots<i_{r-1}<i_r=n\}$, the triangulation of $\mathcal{A}$ indexed by I consists of the r segments $[i_0,i_1],[i_1,i_2],\ldots,[i_{r-1},i_r]$. See [8, Sections 7.3.A and 12.2.A] for details. If $r=n$ (resp. $r=1$) then this is the finest (resp. coarsest) triangulation referred to in Section 1.

We are now prepared to present the main result of this note. In the remainder of Section 3 we shall be working over the field of complex numbers $\mathbf{C}$. Fix any of the 2^{n-1} triangulations of $\mathcal{A}$. For $j=1,\ldots,r$ we write $d_j := i_j - i_{j-1}$ for the length of the j-th segment in that triangulation. Clearly, $d_1+d_2+\cdots+d_r=n$. Let $\xi = (-1)^{1/d_j}$ be any of the d_j-th roots of -1. We define the $\mathcal{A}$-hypergeometric series

$$X_{j,\xi} \quad := \quad \xi \cdot \left[a_{i_{j-1}}^{\frac{1}{d_j}} a_{i_j}^{-\frac{1}{d_j}} \right] \quad + \quad \frac{1}{d_j} \cdot \sum_{k=2}^{d_j} \xi^k \cdot \left[a_{i_{j-1}+k-1}\, a_{i_{j-1}}^{\frac{k-d_j}{d_j}}\, a_{i_j}^{-\frac{k}{d_j}} \right] \quad + \quad \frac{1}{d_j} \cdot \left[\frac{a_{i_{j-1}-1}}{a_{i_{j-1}}} \right]$$

If $j = 1$ then the expression $\left[\frac{a_{-1}}{a_0}\right]$ appears in the rightmost summand. We define it to be zero. Note that by varying j and ξ we have defined n distinct series in total.

Theorem 3.2. *The n series $X_{j,\xi}$ are roots of the general equation of order n, that is, $f(X_{j,\xi}) = 0$. There exists a constant M such that all n series $X_{j,\xi}$ converge whenever*

$$|a_{i_{j-1}}|^{i_j - k} \cdot |a_{i_j}|^{k - i_{j-1}} \;\leq\; M \cdot |a_k|^{d_j} \quad \textit{for all } 1 \leq j \leq r \textit{ and } k \notin \{i_{j-1}, i_j\}. \tag{3.2}$$

Proof: Consider the open convex cone

$$C \;=\; \big\{ w \in \mathbf{R}^{n+1} \,:\, (i_j - k)\cdot w_{i_{j-1}} + (k - i_{j-1})\cdot w_{i_j} < d_j \cdot w_k \text{ for } j = 1, \ldots, r,\ k \notin \{i_{j-1}, i_j\} \big\}.$$

This is the outer normal cone to the secondary polytope $\Sigma(\mathcal{A})$ at the vertex corresponding to the triangulation I of $\mathcal{A}$ (cf. [8, Theorem 12.2.2]). Equivalently, the cone C consists of all vectors $w = (w_0, w_1, \ldots, w_n)$ which induce the triangulation in question.

Let $\mathcal{U}$ be the region in the coefficient space $\mathbf{C}^{n+1}$ defined by the inequalities (3.2) for $M \gg 0$. There exists a vector $V \in \mathbf{R}^{n+1}$ such that $(log(|a_0|), \ldots, log(|a_n|)) - V \in C$ for all $(a_0, \ldots, a_n)$ in $\mathcal{U}$. Let $\mathcal{H}$ be the space of all complex-valued functions on $\mathcal{U}$ which are $\mathcal{A}$-hypergeometric with parameters $\binom{-1}{0}$. The $\mathcal{A}$-hypergeometric system is holonomic of rank n and its singular locus, the discriminantal locus of f, is disjoint from $\mathcal{U}$ (see [6]). Hence $\mathcal{H}$ is a complex vector space of dimension at most n. We shall identify n linearly independent elements in $\mathcal{H}$, which will imply that $\mathcal{H}$ has dimension exactly n.

It follows from [6, Proposition 2] that the series $\left[a_0^{u_0} \cdots a_n^{u_n}\right]$ defined in (3.1) converges in $\mathcal{U}$ provided *at most two* of the exponents u_i are non-integers. Each of the summands in the definition of $X_{j,\xi}$ has this property. Therefore $X_{j,\xi}$ converges in $\mathcal{U}$. Using Lemma 3.1, we conclude that the n series $X_{j,\xi}$ lie in the vector space $\mathcal{H}$.

We fix an vector $w = (w_0, \ldots, w_n)$ in C such that all coordinates w_i are integers and such that no two of the lines spanned by pairs $\{(w_i, i), (w_j, j)\}$ in $\mathbf{R}^2$ are parallel. The weight of a monomial $a_0^{i_0} a_1^{i_1} \cdots a_n^{i_n}$ (with rational exponents) is defined to be $w_0 i_0 + w_1 i_1 + \cdots + w_n i_n$. We replace the input equation by its *toric deformation*

$$f_t(x) \quad = \quad a_0 t^{w_0} \,+\, a_1 t^{w_1} x \,+\, a_2 t^{w_2} x^2 \,+\, \ldots \,+\, a_{n-1} t^{w_{n-1}} x^{n-1} \,+\, a_n t^{w_n} x^n. \tag{3.3}$$

We shall study the n roots as an algebraic function of t. For t close to the origin the n roots split into r groups, one for each segment $[i_{j-1}, i_j]$, for $j = 1, \ldots, r$. (This is a special case of the multivariate construction in [9, §3].) The roots in the j-th group possess a Puiseux expansion of the form

$$\psi_{j,\xi}(t) \quad = \quad \xi \cdot a_{i_{j-1}}^{\frac{1}{d_j}} \cdot a_{i_j}^{-\frac{1}{d_j}} \cdot t^{\frac{1}{d_j}(w_{i_{j-1}} - w_{i_j})} + \text{higher terms in } t,$$

where ξ satisfies $\xi^5 = -1$. We shall prove that $\psi_{j,\xi}(1) = X_{j,\xi}$ for all $(a_0, \ldots, a_n) \in \mathcal{U}$. To this end we first determine the second lowest term in the series $\psi_{j,\xi}(t)$. After modifying the weight vector w by an affine transformation $w_i \mapsto \alpha w_i + \beta$, which does not alter the structure of the series, we may assume that $w_{i_j} = w_{i_{j-1}} = 0$, and there is a unique index r with $w_r = 1$, and $w_l > 1$ for $l \in \{0, 1, \ldots, n\} \backslash \{i_j, i_{j-1}, r\}$. An explicit calculation reveals

$$\begin{aligned} \psi_{j,\xi}(t) \quad &= \quad \xi a_{i_{j-1}}^{\frac{1}{d_j}} a_{i_j}^{-\frac{1}{d_j}} \\ &+ \ \frac{1}{d_j} \cdot \xi^{r+1-i_{j-1}} \cdot a_r \cdot a_{i_{j-1}}^{(r+1-i_j)/d_j} \cdot a_{i_j}^{(i_{j-1}-r-1)/d_j} \cdot t \ + \ \text{higher terms in } t. \end{aligned} \tag{3.4}$$

By varying w within the cone C we can arrange that the role of r is played by any index in the set

$$\big(\{0, 1, \ldots, n\} \backslash \{i_1, i_2, \ldots, i_r\}\big) \cup \{i_{j-1}, i_{j+2}\}. \tag{3.5}$$

Note that the cardinality of this set is at least $d_j - 1$. Consider the linear map that extracts from a series in $a_0, \ldots, a_n$ all those terms which are lowest or second lowest with respect to the grading defined by some $w \in C$. Call this map T. Consider the image of a root $\psi_{j,\xi}(1)$ under T. This is a polynomial (with fractional exponents) having at least d_j distinct terms, one for the lowest term in (3.4) and at least $d_j - 1$ for the distinct values of r coming from (3.5). The coefficients of these terms are distinct powers of ξ divided by d_j. These considerations imply that the images of the roots $\psi_{j,\xi}(1)$ under T are linearly independent. Therefore the roots themselves are linearly independent over **C**. Moreover, they all satisfy the $\mathcal{A}$-hypergeometric system with parameter $\binom{-1}{0}$, by Proposition 2.1. We conclude that the space $\mathcal{H}$ is n-dimensional and the roots $\psi_{j,\xi}(1)$ form a basis for $\mathcal{H}$.

We wish to prove $\psi_{j,\xi}(1) = X_{j,\xi}$. It suffices to show $T(\psi_{j,\xi}(1)) = T(X_{j,\xi})$ because the functional T defined above separates the space $\mathcal{H}$. Equivalently, given any generic w in C, we must show that, in the w-grading, $X_{j,\xi}$ has the same first two terms as $\psi_{j,\xi}(1)$. This is clear for the first term, so we only need to look at the second term $\frac{1}{d_j}\xi^{r+1-i_{j-1}} \cdot a_r \cdot a_{i_{j-1}}^{(r+1-i_j)/d_j} \cdot a_{i_j}^{(i_{j-1}-r-1)/d_j}$. Let a^v denote this monomial. Let k be the unique integer between 1 and d_j such that r is congruent to $i_{j-1} - 1 + k$ modulo d_j. If $k = 1$ then a^v equals the second lowest term (with respect to w) of the series $\xi \cdot [a_{i_{j-1}}^{1/d_j} a_{i_j}^{-1/d_j}]$. If $2 \le k \le d_j - 1$ then a^v equals the w-lowest term of the series indexed by k in the central sum in the definition of $X_{j,\xi}$. For $k = d_j$ there are two subcases: if $r > i_{j-1}$ then a^v equals the w-lowest term of the series indexed by $k = d_j$ in the center sum; if $r < i_{j-1}$ then a^v equals the w-lowest term of the rightmost series $\frac{1}{d_j}[a_{i_{j-1}-1}/a_{i_j-1}]$. In each of these four cases a^v is the second lowest term in $X_{j,\xi}$. This completes the proof of Theorem 3.2. ■

4. Integrality Issues

To prove the identity $f(X_{j,\xi}) = 0$ in Theorem 3.2 we used the complex numbers. But, a posteriori, there is no need to stick with an algebraically closed field of characteristic zero. If K is any integral domain such that the equation $\xi^{d_j} = -1$ has d_j distinct solutions and the coefficients of $X_{j,\xi}$ are defined in K, then $f(X_{j,\xi}) = 0$ is a valid identity in a suitable fractional power series ring over K. The following result shows that K can be an algebraically closed field of characteristic p, provided p is relatively prime to $d_1 d_2 \cdots d_r$. The proof of Theorem 4.1 using Hensel's Lemma was suggested to me by Hendrik Lenstra.

Theorem 4.1. *All coefficients of the series $X_{j,\xi}$ lie in the ring $\mathbf{Z}[\frac{1}{d_j}][\xi]$.*

Proof: The series $\psi_{j,\xi}(t)$ in (3.4) lies in the fractional power series ring $R_1[[t^{\frac{1}{d_j}}]]$, where

$$R_1 \quad := \quad \mathbf{Q}[\,\xi\,]\big[a_s : s \notin \{i_{j-1}, i_j\}\big]\,[a_{i_{j-1}}^{\pm\frac{1}{d_j}}, a_{i_j}^{\pm\frac{1}{d_j}}].$$

This holds because ω is integral and each term of $X_{j,\xi} = \psi_{j,\xi}(1)$ lies in R_1. To prove Theorem 4.1, it suffices to show that $\psi_{j,\xi}(t)$ is an element of the subring $R_2[[t^{\frac{1}{d_j}}]]$, where

$$R_2 \quad := \quad \mathbf{Z}[\frac{1}{d_j}][\,\xi\,]\big[a_s : s \notin \{i_{j-1}, i_j\}\big]\,[a_{i_{j-1}}^{\pm\frac{1}{d_j}}, a_{i_j}^{\pm\frac{1}{d_j}}].$$

We shall apply Hensel's Lemma [5, Theorem 7.3] to the integral domain $R_2[[t^{\frac{1}{d_j}}]]$. This domain is complete with respect to the principal ideal $\mathbf{m} := \langle t^{\frac{1}{d_j}} \rangle$, and it contains all coefficients of the polynomial $f_t(x)$ in (3.3). The constant term of $\psi_{j,\xi}(t)$ lies in R_2; we denote it by $A := \xi a_{i_{j-1}}^{\frac{1}{d_j}} a_{i_j}^{-\frac{1}{d_j}}$. It is an "approximate root" in the sense that $f_t(A) \in \mathbf{m}$ and $f_t'(A)$ is the sum of a unit in R_2 and an element of $\mathbf{m}$. By Hensel's Lemma there exists a unique element B in $R_2[[t^{\frac{1}{d_j}}]]$ such that $f_t(B) = 0$ and $A - B \in \mathbf{m}$. By repeating the uniqueness part of this application of Hensel's Lemma for the coefficient domain R_1 instead of R_2, we conclude that B must be equal to our Puiseux series $\psi_{j,\xi}(t)$. ■

A case of special interest is the finest triangulation, where $I = \{1, 2, \ldots, n-1\}$, $r = n$, $d_j = 1$ for all j. Theorem 4.1 implies that in this case our construction is characteristic-free. In other words, the series solutions $X_{j,-1}$ have integer coefficients. These series are

$$X_{j,-1} \quad = \quad -\left[\frac{a_{j-1}}{a_j}\right] + \left[\frac{a_{j-2}}{a_{j-1}}\right] \qquad \text{for} \quad j = 1, 2, \ldots, n \tag{4.1}$$

where $\left[\frac{a_{j-1}}{a_j}\right]$ is the sum over all Laurent monomials

$$\frac{(-1)^{i_j}}{i_{j-1}+1} \cdot \binom{i_j}{i_0 \ldots i_{j-1}\, i_{j+1} \ldots i_n} \cdot \frac{a_0^{i_0} a_1^{i_1} \cdots a_{j-2}^{i_{j-2}}\, a_{j-1}^{i_{j-1}+1}\, a_{j+1}^{i_{j+1}} \cdots a_n^{i_n}}{a_j^{i_j+1}} \tag{4.2}$$

where $i_0, i_1, \ldots, i_n$ are non-negative integers satisfying the relations

$$\begin{aligned} i_0 + i_1 + i_2 + i_3 + \cdots + i_{j-1} \ - i_j \ + i_{j+1} + \cdots + i_n &= 0 \\ i_1 + 2i_2 + 3i_3 \cdots + (j-1)i_{j-1} \ - ji_j \ + (j+1)i_{j+1} + \cdots + ni_n &= 0. \end{aligned} \tag{4.3}$$

For generating the series $\left[\frac{a_{j-1}}{a_j}\right]$ on a computer it convenient to rewrite (4.3) as follows.

$$\begin{aligned} i_{j-1} &= -ji_0 - (j-1)i_1 - (j-2)i_2 - \cdots - 2i_{j-2} \ + \ i_{j+1} + 2i_{j+2} + \cdots + (n-j)i_n, \\ i_j &= -(j-1)i_0 - (j-2)i_1 - \cdots - i_{j-2} \ + \ 2i_{j+1} + 3i_{j+2} + \cdots + (n-j+1)i_n. \end{aligned}$$

These equations ensure that the multinomial coefficient in (4.2) is divisible by $i_{j-1} + 1$.

5. Multivariate Outlook

Consider a system of n polynomial equations in n variables, where the terms in the i-th equation have their exponent vectors in a fixed set $\mathcal{A}_i \subset \mathbf{Z}^n$. (This is the meaning of "sparse" in [9]). By Bernstein's Theorem, the system has mixed volume many roots $(X_1, \ldots, X_n)$. Each coordinate X_i is an algebraic function in all the coefficients. It is natural to wonder whether X_i satisfies the $\mathcal{A}$-hypergeometric differential equations where $\mathcal{A}$ is the configuration arising from the "Cayley trick" (cf. [7, 2.5], [8, page 273]):

$$\mathcal{A} \quad := \quad \mathcal{A}_1 \times \{e_1\} \ \cup \ \mathcal{A}_2 \times \{e_2\} \ \cup \ \cdots \ \cup \ \mathcal{A}_n \times \{e_n\} \quad \subset \quad \mathbf{Z}^{2n}.$$

The answer is **no**. The coordinates of the roots $(X_1, \ldots, X_n)$ are not $\mathcal{A}$-hypergeometric. For example, let (X_1, X_2) be the unique solution of the system of two linear equations

$$a_0 \ + \ a_1x_1 \ + \ a_2x_2 \quad = \quad b_0 \ + \ b_1x_1 \ + \ b_2x_2 \quad = \quad 0.$$

Then we find

$$\frac{\partial^2 X_2}{\partial a_0 \partial b_1} \quad \neq \quad \frac{\partial^2 X_2}{\partial a_1 \partial b_0}.$$

Note also that Mayr's differential equations (40) and (45) in [10] are no longer binomial.

The correct generalization of Proposition 2.1 to higher dimensions is the following. Let $J(x_1, \ldots, x_n)$ denote the *Jacobian* $det(\frac{\partial f_i}{\partial x_j})$ of the given equations $f_1 = \cdots = f_n = 0$.

Proposition 5.1. *For any integers $u_1, u_2, \ldots, u_n$ the algebraic function*

$$\frac{X_1^{u_1} X_2^{u_2} \cdots X_n^{u_n}}{J(X_1, X_2, \ldots, X_n)} \tag{5.1}$$

satisfies the $\mathcal{A}$-hypergeometric differential equations arising from the Cayley trick.

Over the field of complex numbers, we can derive Proposition 5.1 from Theorem 2.7 in [7] using Cauchy's formula in several variables. An alternative algebraic proof follows from the construction in [1, §2]. The quantity in (5.1) should be thought of as a *local residue* on a toric variety [4]. The sum over all (mixed volume many) local residues is the *global residue* associated with the monomial $x_1^{u_1} \cdots x_n^{u_n}$ relative to the given equations $f_1 = \cdots = f_n = 0$. The global residue is a rational function. It is of importance in elimination theory. We plan to extend the techniques in Section 3 to the setting of Proposition 5.1 in a subsequent joint work with E. Cattani and A. Dickenstein. The goal of that project is to develop new formulas and algorithms for computing local and global residues. Another interesting question is whether explicit Puiseux series expansions of (5.1) might be useful to improve the numerical component in the homotopy algorithm proposed in [9].

Acknowledgements

I am grateful to Alicia Dickenstein, Hendrik Lenstra and Michael Trott for helpful comments on earlier drafts of this paper. This project was partially supported by grants from the National Science Foundation (NYI) and the David and Lucile Packard Foundation.

References

[1] A. Adolphson and S. Sperber: Differential modules defined by systems of equations, *Seminario Rendiconti di Universita di Padova*, to appear.

[2] R. Birkeland: Über die Auflösung algebraischer Gleichungen durch hypergeometrische Funktionen, *Mathematische Zeitschrift* **26** (1927) 565–578.

[3] J.-L. Brylinski: Radon transform and functionals on the space of curves, Manuscript, 1993.

[4] E. Cattani, D. Cox and A. Dickenstein: Residues in toric varieties, *Compositio Math.*, to appear.

[5] D. Eisenbud: *Commutative Algebra with a View Towards Algebraic Geometry*, Graduate Texts in Mathematics, Springer Verlag, New York, 1995.

[6] I.M. Gel'fand, A.V. Zelevinsky and M.M. Kapranov: Hypergeometric functions and toral manifolds, *Functional Analysis and its Applications* **23** (1989) 94–106.

[7] I.M. Gel'fand, A.V. Zelevinsky and M.M. Kapranov: Generalized Euler integrals and $\mathcal{A}$-hypergeometric functions, *Advances in Mathematics* (1990) **84** 255–271.

[8] I.M. Gel'fand, A.V. Zelevinsky and M.M. Kapranov: *Discriminants, Resultants, and Multidimensional Determinants*, Birkhäuser, Boston, 1994.

[9] B. Huber and B. Sturmfels: A polyhedral method for solving sparse polynomial systems, *Mathematics of Computation* **64** (1995) 1541–1555.

[10] K. Mayr: Über die Auflösung algebraischer Gleichungssysteme durch hypergeometrische Funktionen, *Monatshefte für Mathematik und Physik* **45** (1937) 280–313.

[11] J. McDonald: Fiber polytopes and fractional power series, *Journal of Pure and Applied Algebra*, **104** (1995) 213–233.

[12] *Solving the Quintic with Mathematica*, poster distributed by Wolfram Research, 1994.