Expressing a Fraction of Two Determinants as a Determinant.

Pascal Koiran, Erich Kaltofen

Expressing a Fraction of Two Determinants as a Determinant*

Erich Kaltofen[1] and Pascal Koiran[2]

[1]Dept. of Mathematics, North Carolina State University,
Raleigh, North Carolina 27695-8205 USA
kaltofen@math.ncsu.edu; http://www.kaltofen.us

[2]Laboratoire de l'Informatique du Parallélisme (LIP)†,
École Normale Supérieure de Lyon
Université de Lyon, France
Pascal.Koiran@ens-lyon.fr

Abstract

Suppose the polynomials f and g in $\mathsf{K}[x_1,\ldots,x_r]$ over the field K are determinants of non-singular $m\times m$ and $n\times n$ matrices, respectively, whose entries are in $\mathsf{K}\cup\{x_1,\ldots,x_r\}$. Furthermore, suppose $h=f/g$ is a polynomial in $\mathsf{K}[x_1,\ldots,x_r]$. We construct an $s\times s$ matrix C whose entries are in $\mathsf{K}\cup\{x_1,\ldots,x_r\}$, such that $h=\det(C)$ and $s=\gamma(m+n)^6$, where $\gamma=O(1)$ if K is an infinite field or if for the finite field $\mathsf{K}=\mathbb{F}_q$ with q elements we have $m=O(q)$, and where $\gamma=(\log_q m)^{1+o(1)}$ if $q=o(m)$. Our construction utilizes the notion of skew circuits by Toda and weakly-skew circuits by Malod and Portier. Our problem was motivated by resultant formulas derived from Chow forms.

Additionally, we show that divisions can be removed from formulas that compute polynomials in the input variables over a sufficiently large field within polynomial formula size growth.

1. Introduction

1.1. Motivation

Our investigated problem was motivated by the question of resultant formulas without division. Originally, the resultant of a set of t homogeneous polynomial equations $f_1=\cdots=f_t=0$ in t variables has been expressed as a GCD of determinants, whose matrices have the coefficients of the polynomials as entries. Macaulay [1916] gave a formula of a quotient of two such

*This material is based on work supported in part by the National Science Foundation under Grant CCF-0514585 (Kaltofen).

†Unité mixte de recherche (UMR) 5668 ENS Lyon, CNRS, Université Claude Bernard Lyon 1 (UCBL), INRIA.

determinants. In special cases, one can remove the division and construct a single determinant that is the resultant [Khetan 2002; Eisenbud et al. 2003; Khetan et al. 2004]. Those constructions use properties of exact sequences of exterior algebras. However, an algebraic complexity theoretic approach can remove the division in the general case in an entirely different manner. Take Macaulay's formula resultant$(f_1, \ldots, f_t) = \det(A)/\det(B)$, where the larger matrix A has dimensions $m \times m$. One converts the determinants to straight-line programs, removes the division by Strassen's [1973] method, or computes their GCD [Kaltofen 1988], parallelizes the straight-line program to $O((\log m)^2)$ depth [Valiant et al. 1983; Miller et al. 1988], converts the resulting division-free straight-line program to a formula of quasi-polynomial size $m^{O(\log m)}$, and finally writes the resultant formula as the projection of a determinant of a matrix C of dimensions $k \times k$ where $k = m^{O(\log m)}$ [Valiant 1979]. Note that C with $\det(C) = \det(A)/\det(B)$ has as entries the coefficients of the original polynomials or constants. Here we shall show that there is a matrix $\bar{C}$ of dimensions $O(m^6)$ whose entries are the coefficients of the original polynomials or constants with $\det(\bar{C}) = \det(A)/\det(B)$.

1.2. Results and Used Approach

Our main result is the following theorem.

Theorem 1. *Let $f, g, h \in \mathsf{K}[x_1, \ldots, x_r] \setminus \{0\}$, where K is a field, such that $f/g = h$ and f is a projection à la Valiant [1979] of an $m \times m$ determinant and g is a projection of an $n \times n$ determinant ($n \leq m$ or $n > m$), meaning that there are matrices $A \in \mathsf{K}[x_1, \ldots, x_r]^{m \times m}$ and $B \in \mathsf{K}[x_1, \ldots, x_r]^{n \times n}$, whose entries are in $\mathsf{K} \cup \{x_1, \ldots, x_r\}$ with $f = \det(A)$ and $g = \det(B)$. Then there exists an $s \times s$ matrix C whose entries are in $\mathsf{K} \cup \{x_1, \ldots, x_r\}$ such that $h = f/g = \det(C)$ and s is polynomial in $m + n$, that is, the exact quotient of f and g is a projection of a determinant of polynomial dimension. More precisely,*

- *(i) if K is infinite or if $m = O(|K|)$ we can take $s = O((m+n)^6)$;*
- *(ii) if K is a small finite field, we can take $s = O((m+n)^6 \cdot \mathrm{M}(\log_{|\mathsf{K}|} m))$, where $\mathrm{M}(l) = l \cdot (\log l) \cdot (\operatorname{loglog} l)$.*

We prove our result via the notion of *weakly-skew* division-free arithmetic circuits by Malod and Portier [2007]. We consider division-free arithmetic circuits (straight-line programs), which are directed acyclic graphs (DAGs) whose nodes have fan-in at most two and which perform addition, subtraction and multiplication. The operands are the values in previous nodes, constant scalars or input variables. The values of designated output nodes are multivariate polynomials in the input variables. The size of the graph is the number of arithmetic operations (sequential complexity) performed. This definition of size is consistent with [Valiant 1979], for instance. Malod and Portier [2007] work with a slightly different definition of size, in which input nodes (variables or constants) are counted along with arithmetic nodes. To avoid any confusion, we will call this second notion *fat size*. The fat size of a circuit is therefore equal to the sum of its size and of the number of input nodes. It is bounded by 3 times the size since the number of input nodes is equal to at most twice the number of arithmetic nodes.

Toda [1992] introduces *skew* division-free arithmetic circuits, which have the property that at least one of the two operands in each multiplication node is either a scalar constant or an input variable. Toda proves that the determinant polynomial of an $m \times m$ matrix can be computed by a skew circuit of size $O(m^{20})$. In weakly-skew circuits at least one of the two operands to a multiplication node must be computed by a separate circuit. Figure 1 shows an example of a weakly-skew circuits. The separate circuits for the operands of multiplication nodes are marked by dashed boxes.

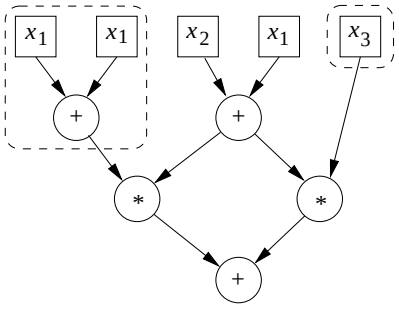

Figure 1: A weakly-skew circuit [Malod and Portier 2007, Fig. 4]

Valiant [1979] proves that every formula of size s is the projection of an $(s+2) \times (s+2)$ determinant. In formulas, both operands to all nodes are computed by separate formulas. Valiant's proof can be generalized to show that every polynomial that is computed by a weakly-skew circuit of fat size s is the projection of an $(s+1) \times (s+1)$ determinant [Malod and Portier 2007, Lemma 6], also using negated input variables. Furthermore, the division-free parallel circuits by Berkowitz [1984] for the characteristic polynomial of an $m \times m$ matrix can be converted to weakly-skew circuits of size $O(m^5)$ [Malod and Portier 2007, Proposition 5]. In Section 3 we apply Strassen's [1973] technique for elimination of divisions directly to the characteristic polynomials derived from A and B in Theorem 1 above, thus obtaining a division-free weakly-skew circuit for $h = \det(A)/\det(B)$, which then is the projection of a determinant. Note that the technique in [Canny 1990], which is not directly applicable, assumes that the diagonals of A and B hold a single separate variable, implying also $n < m$. Along the way we show in Section 2 that weakly-skew circuits can be simulated by skew circuits with an increase in size by a constant factor only; and that Toda's $O(m^{20})$ bound can be reduced to $O(m^5)$.

Since weakly-skew circuits are projections of determinants, we have the following corollary to Theorem 1 and our transformation results to skew circuits.

Corollary 1. *Let $f, g, h \in \mathsf{K}[x_1, \ldots, x_r] \setminus \{0\}$, where K is a field with $\deg(g) = O(|\mathsf{K}|)$ such that $f/g = h$ and f and g are computed by a weakly-skew circuit of size s with inputs $x_1, \ldots, x_r$. Then h can be computed by a skew circuit of size $O(s^6)$.*

Suppose in the above corollary that f and g are computed by division-free formulas of size $\leq s$. Then $h = f/g$ is a projection of a determinant of a matrix of dimension $O(s^6)$. In Section 4 we show that there exists a division-free formula of size (s^5) that computes $h = f/g$. As a consequence of the latter result, we can show that a polynomial in $\mathsf{K}[x_1, \ldots, x_r]$ that is computed by a formula of size s with additions, subtractions, multiplications *and divisions*,

where K is a sufficiently large field and $x_1, \ldots, x_r$ are input variables, can be computed by a division-free formula of size $s^{O(1)}$ (see Theorem 3 on page 10).

2. From Weakly-skew to Skew Circuits

In this section we show that weakly-skew circuits can be efficiently simulated by skew circuits. This fact will be used in the proof of Theorem 1, part (ii) given in Section 3.2.

Two algorithms for converting weakly-skew circuits into skew circuits are already described by Malod and Portier [2007]. The polynomial computed by a weakly-skew circuit of fat size m can be converted into a determinant of size $m+1$ by [Malod and Portier 2007, Lemma 6]. One can then apply Toda's algorithm, which evaluates a determinant of an $m \times m$ matrix by a skew circuit of size $O(m^{20})$. It is observed in [Malod and Portier 2007, Section 5.2] that the role of the determinant in this first algorithm can be played by the polynomial $F_m = \mathrm{Trace}(X^m)$. The method described below is even simpler, and more efficient. As a byproduct, we obtain an improvement to $O(m^5)$ of Toda's original $O(m^{20})$ bound.

Remark 1. In the determinant constructed in [Malod and Portier 2007, Lemma 6], all input variables are negated. This is not a problem for converting a weakly-skew circuit into a skew circuit with the algorithm described in the paragraph above. However, the occurence of negated input variables makes their construction unsuitable for the proof of Theorem 1. To circumvent this difficulty, one can modify slightly Malod and Portier's condition to obtain a matrix of size $m+1$ *without negated variables*. This can be done in two steps:

1. In their proof of Lemma 6 Malod and Portier first construct a matrix B of size m such that $\det B = -f$, where f is the polynomial computed by a weakly-skew circuit of size m. In this matrix all variables are negated (and all diagonal entries are equal to 1 except the first, which is 0). Now let $C = -B$: we have $\det C = \pm f$, and there are no negated variables in C.

2. If m is odd we have $\det C = -f$. As Malod and Portier, we can add one last row and one last column full of 0's (except for an entry equal to -1 in the bottom right corner) to obtain a matrix of size $m+1$ whose determinant is equal to f.

We shall work with acyclic edge-weighted directed graphs. We recall that the weight of a path in such a graph is defined as the product of the weights of the edges appearing in the path. If s and t are two vertices of G, the weight of (s,t) in G is defined as the sum of the weights of all paths from s to t.

Lemma 1. *Let W be a weakly-skew circuit of fat size m. There exists an acyclic directed graph G, with two distinguished vertices s and t such that:*

(*i*) *The weight of (s,t) in G is the polynomial computed by W, and G is of size at most $m+1$.*

(*ii*) *Every vertex in G other than s has either a single incoming edge, of weight equal to an input of W or to the constant 2, or two incoming edges, each of weight 1.*

This is essentially Lemma 5 of Malod and Portier [2007]. In that lemma the authors prove the existence of a graph G satisfying (i). An inspection of their proof shows that the graph that they construct also satisfies the second property.

Proposition 1. *Let W be a weakly-skew circuit of fat size m. There exists a skew circuit W' which is equivalent to W (i.e., computes the same polynomial), has the same number of input nodes as W, and has at most m arithmetic nodes (W' is therefore of fat size at most $2m$).*

The algorithmic idea behind this result is quite simple: for each vertex v in the graph G of Lemma 1 we compute the weight $\omega(v)$ of the pair (s, v). If v has a single incoming edge of weight x connecting s to v then of course $\omega(v) = x$. If v has a single incoming edge of weight x connecting a vertex $v' \neq s$ to v, we can apply the formula $\omega(v) = x \times \omega(v')$ if x is an input of W. If x is the constant 2, we apply the formula $\omega(v) = \omega(v') + \omega(v')$ (one could of course use the the formula $\omega(v) = 2 \times \omega(v')$ instead, at the cost of introducing one additional constant input in W'). Finally, if v has two incoming edges of weight 1, connecting the vertices v_1 and v_2 to v, we have $\omega(v) = \omega(v_1) + \omega(v_2)$. The resulting circuit satisfies the requirements of Proposition 1.

Corollary 2. *The determinant of a $m \times m$ matrix can be computed by a skew circuit of size $O(m^5)$.*

This is a significant improvement over the $O(m^{20})$ bound given by Toda [1992]. Corollary 2 is an immediate consequence of Proposition 1 since, as pointed out in section 1, the determinant of a $m \times m$ matrix can be computed by a weakly-skew circuit of size $O(m^5)$.

Skew circuits will be useful in Section 3.2 due to the following proposition and the subsequent remarks.

Proposition 2. *Let K be a field and let $\mathsf{L} = \mathsf{K}(\theta)$ be an algebraic extension of K of degree $d = [\mathsf{L} : \mathsf{K}]$. Let $f \in \mathsf{L}[x_1, \ldots, x_r]$ be a polynomial computed by a skew circuit W_L of size m, with input nodes labeled by variables from $\{x_1, \ldots, x_r\}$ or constants from L. Let us expand f according to the powers of θ: one can write $f = \sum_{j=0}^{d-1} \theta^j f_j$, where $f_j \in \mathsf{K}[x_1, \ldots, x_r]$.*

There exists a skew circuit W_K of size $O(d^2 m)$ with at most d output nodes which computes simultaneously all the polynomials f_j. Moreover, W_K uses only constants from K.

Proof. Let α be a node of W_L computing a polynomial $f_\alpha \in \mathsf{L}[x_1, \ldots, x_r]$. We use a standard technique: in W_K, f_α will be represented by d nodes computing polynomials $f_{0,\alpha}, \ldots, f_{d-1,\alpha} \in \mathsf{K}[x_1, \ldots, x_r]$ such that $f_\alpha = \sum_{j=0}^{d-1} \theta^j f_{j,\alpha}$. If α is an input node labeled by some variable x_i we can take $(f_{0,\alpha}, \ldots, f_{d-1,\alpha}) = (x_i, 0, \ldots, 0)$; if α is labeled by the constant $\sum_{j=0}^{d-1} a_j \theta^j$ we can take $(f_{0,\alpha}, \ldots, f_{d-1,\alpha}) = (a_0, a_1, \ldots, a_{d-1})$. Assume now that α is an addition node with inputs coming from nodes β and γ. In this case we simply perform componentwise additions since $f_{j,\alpha} = f_{j,\beta} + f_{j,\gamma}$.

Finally, the case where α is a multiplication node can be split in two subcases since W_L is skew: multiplication by a variable x_i, or multiplication by a constant from L. In the first subcase, assume that α multiplies the output of node β by x_i. We have $f_{j,\alpha} = x_i f_{j,\beta}$. Observe that the d resulting multiplications are skew. In the second subcase, assume that α

multiplies the output of node β by a constant $\eta \in \mathsf{L}$. Multiplication by a constant is a K-linear operation. The tuple $(f_{0,\alpha}, \ldots, f_{d-1,\alpha})$ can therefore be obtained from $(f_{0,\beta}, \ldots, f_{d-1,\beta})$ by multiplication by an appropriate $d \times d$ matrix A_η with entries in K. The corresponding matrix-vector product can be computed in $O(d^2)$ arithmetic operations, and once again the resulting multiplications are all skew. □

Remark 2. We will apply this result in Section 3.2 in a situation where we know that the output of W_L lies in fact in $\mathsf{K}[x_1, \ldots, x_r]$. In this case, the subcircuit associated to the first output node of W_K computes the same polynomial as W_L.

Remark 3. Proposition 2 also applies to weakly-skew rather than skew circuits. First one converts the weakly-skew circuit into a skew circuit using Proposition 1. One can then apply Proposition 2 to the skew circuit.

Remark 4. The size of W_K in Proposition 2 can be reduced to $O(m\, d\, (\log d) \operatorname{loglog} d)$ by fast polynomial multiplication algorithms [Cantor and Kaltofen 1991] and fast division with remainder algorithms [von zur Gathen and Gerhard 1999, Section 9.1].

3. Elimination of Divisions

3.1. Large Coefficient Fields

Our symbolic determinants live in a multivariate polynomial domain $\mathsf{K}[x_1, \ldots, x_r]$, where K is a sufficiently large field. Consider we have non-singular matrices

$$A(x_1, \ldots, x_r) \in \mathsf{K}[x_1, \ldots, x_r]^{m \times m} \text{ and } B(x_1, \ldots, x_r) \in \mathsf{K}[x_1, \ldots, x_r]^{n \times n},$$

whose entries are either variables or constants, i.e.,

$$\forall i, j, k, l \text{ with } 1 \le i, j \le m, 1 \le k, l \le n\colon (A)_{i,j}, (B)_{k,l} \in \mathsf{K} \cup \{x_1, \ldots, x_r\}.$$

Here $(M)_{i,j}$ denotes the element in row i and column j in the matrix M.

We suppose now that $\det(A)/\det(B) \in \mathsf{K}[x_1, \ldots, x_r]$, that is the polynomial division by $\det(B)$ is exact. We construct a division-free weakly-skew arithmetic circuit W of size $O((m+n)^6)$, i.e, *polynomial* in the dimensions of A and B, that computes the polynomial $\det(A)/\det(B)$.

The construction follows Strassen's [1973], using Berkowitz's [1984][‡] and Chistov's [1985] weakly-skew arithmetic circuits for the characteristic polynomial. Let $u_1, \ldots, u_r \in \mathsf{K}$ be such that both $U_A = A(u_1, \ldots, u_r)$ and $U_B = B(u_1, \ldots, u_r)$ are non-singular. Such values always exist if $|\mathsf{K}| > m \ge \deg(\det(A))$ [Schwartz 1980; Zippel 1979].

[‡]Sasaki and Murao [1982] compute the characteristic polynomial of an $n \times n$ matrix with entries in a commutative ring in $n^{\omega+1+o(1)}$ ring operations, and Berkowitz [1984, Section 4] conjectures that there exists a division-free arithmetic circuits of size $O(n^\omega)$ for the characteristic polynomial, given that $n \times n$ matrices can be multiplied with $O(n^\omega)$ operations. We can set $\omega = 2.375477$ [Coppersmith and Winograd 1990]. The Sasaki&Murao/Berkowitz problem remains open. The best division-free complexity for the characteristic polynomial is $O(n^{2.697263})$ [Kaltofen and Villard 2004].

Now consider

$$\det(U_B - \lambda(U_B - B)) = \det(U_B) \cdot \det(I_n - \underbrace{U_B^{-1}\,\lambda(U_B - B)}_{\lambda\bar{B}}) \in \mathsf{K}[x_1, \ldots, x_r][\lambda], \qquad (1)$$

where I_n denotes an n-dimensional identity matrix. Now the coefficient of λ^i in (1) is the homogeneous part, $\hom_T(\det(B), i)$, of total degree i of $\det(B)$ represented in the term basis

$$T = \{(u_1 - x_1)^{d_1} \cdots (u_r - x_r)^{d_r} \mid d_j \geq 0\}, \qquad (2)$$

namely

$$\det(B) = \sum_{i=0}^{n} \underbrace{\sum_{d_1+\cdots+d_r=i} c_{d_1,\ldots,d_r}(u_1 - x_1)^{d_1} \cdots (u_r - x_r)^{d_r}}_{\hom_T(\det(B), i)} \text{ where } c_{d_1,\ldots,d_r} \in \mathsf{K}.$$

Note that evaluation (1) at $\lambda = 1$ gives $\det(B)$. We compute $1/\det(U_B - \lambda(U_B - B))$ as a truncated power series in $\mathsf{K}(u_1 - x_1, \ldots, u_r - x_r)[[\lambda]]$. Because the constant coefficient of (1) as a polynomial in λ is in K, the coefficients of λ^i in the power series for the reciprocal are homogeneous polynomials of degree i in the basis (2). We present a weakly-skew circuit for the coefficients of λ^i.

Let $M_{1\ldots l,1\ldots l}$ denote the top left $l \times l$ principal submatrix of a matrix M. Chistov's algorithm is based on the identities

$$\begin{aligned}
\frac{1}{\det(I_n - \lambda\bar{B})} &= \prod_{j=1}^{n} \frac{\det(I_{j-1} - \lambda M_{1\ldots j-1,1\ldots j-1})}{\det(I_j - \lambda M_{1\ldots j,1\ldots j})} \\
&= \prod_{l=1}^{n} \left((I_l - \lambda\bar{B}_{1\ldots l,1\ldots l})^{-1}\right)_{l,l} \\
&= \prod_{l=1}^{n} \left(\sum_{k=0}^{\infty} \lambda^k \bar{B}^k_{1\ldots l,1\ldots l}\right)_{l,l} \\
&\equiv \prod_{l=1}^{n} \left(\sum_{k=0}^{m} (\bar{B}^k_{1\ldots l,1\ldots l})_{l,l}\lambda^k\right) \pmod{\lambda^{m+1}} \qquad (3) \\
&\equiv 1 + q_1\lambda + \cdots + q_m\lambda^m \pmod{\lambda^{m+1}}. \qquad (4)
\end{aligned}$$

We use weakly-skew circuits to compute the coefficients q_k. Each $(\bar{B}^k_{1\ldots l,1\ldots l})_{l,l}$ in (3) is computed as

$$\left(\bar{B}_{1\ldots l,1\ldots l} \cdot (\bar{B}_{1\ldots l,1\ldots l} \cdot (\cdots (\bar{B}_{1\ldots l,1\ldots l} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}) \ldots))\right)_l$$

by weakly-skew circuits of size $O(mn^2)$. The weakly-skew circuits for carrying out the modular product (4) require no more than $m + 1$ copies of the circuits for each coefficient $(\bar{B}^k_{1\ldots l,1\ldots l})_{l,l}$. Thus a weakly-skew circuit W_1 of size $O(m^2n^3)$ computes all q_k in (4).

Malod and Portier [2007, Proposition 5] compute the coefficients of

$$1 + p_1\lambda + \cdots + p_m\lambda^m = \det(I_m - U_A^{-1}\lambda(U_A - A))$$

via a weakly-skew circuit W_2 of size $O(m^5)$ by Berkowitz's algorithm. Alternatively and less efficiently, one could as above compute

$$\frac{1}{\det(I_m - U_A^{-1}\lambda(U_A - A))} \equiv 1 + \bar{p}_1\lambda + \cdots + \bar{p}_m\lambda^m \pmod{\lambda^{m+1}},$$

and compute the truncated power series of the reciprocal as

$$\begin{aligned}\frac{1}{1 + \bar{p}_1\lambda + \cdots + \bar{p}_m\lambda^m} &\equiv \sum_{l=0}^{m}(-\bar{p}_1\lambda - \cdots - \bar{p}_m\lambda^m)^l \pmod{\lambda^{m+1}} \\ &\equiv 1 + p_1\lambda + \cdots + p_m\lambda^m \pmod{\lambda^{m+1}},\end{aligned} \tag{5}$$

again repeating the circuits which compute the coefficients $\bar{p}_i$ in the truncated polynomial powers.

A weakly-skew circuit then carries out the multiplication

$$\begin{aligned}\frac{\det(U_A - \lambda(U_A - A))}{\det(U_B - \lambda(U_B - B))} &\equiv \frac{\det(U_A)}{\det(U_B)}(1 + p_1\lambda + \cdots + p_m\lambda^m) \\ &\quad \cdot (1 + q_1\lambda + \cdots + q_m\lambda^m) \pmod{\lambda^{m+1}} \\ &\equiv \varrho_0 + \varrho_1\lambda + \cdots + \varrho_m\lambda^m \pmod{\lambda^{m+1}}.\end{aligned} \tag{6}$$

Here and before, the truncation could be performed at $\deg(\det(A)/\det(B)) + 1 \leq m + 1$. Again, we need no more than $m + 1$ copies of the circuit W_1 for the coefficients q_k or of W_2 for p_k. Because $h = \det(A)/\det(B)$ has total degree $\leq \det(A) \leq m$ and $\rho_i = \hom_T(h, i)$, we can compute

$$\frac{\det(A)}{\det(B)} = \varrho_0 + \varrho_1 + \cdots + \varrho_m$$

by a division-free weakly-skew arithmetic circuit of size $O(\min\{m^3n^3 + m^5, m^2n^3 + m^6\})$.

Remark 5. The above techniques also yield a single determinant of a matrix of polynomially-sized dimensions for a fraction of products of determinants $(\prod_i A^{[i]})/(\prod_j B^{[j]}) \in \mathsf{K}[x_1, \ldots, x_r]$, with and without using block diagonal matrices. Such fractions occur when computing the resultant via Koszul complexes[§].

3.2. Small Coefficient Fields

When the coefficient field K has few elements, divisions by zero may occur at all values $u_1, \ldots, u_r \in \mathsf{K}$ and there is no no-singular U_B in Section 3. We can nonetheless obtain the bound of Theorem 1.(ii) using a field extension.

Indeed, from Section 3 we know that the quotient $h = f/g$ can be computed by a polynomial size weakly-skew circuit W_{L} with constants from L if L is an extension of K with

[§] Ágnes Szántó has pointed this application out to us.

at least $m+1$ elements. We can therefore work with an extension of degree $O(\log_{|\mathsf{K}|}(m))$. Since h is actually a polynomial with coefficients in K, by applying Proposition 2 and the two subsequent remarks to W_{L} we obtain a skew circuit W_{K} which computes h using constants from K only. Finally, W_{K} can be transformed into a determinant by [Malod and Portier 2007, Lemma 6] as in the proof of Theorem 1(i).

4. Formulas With Divisions

We shall prove the following theorem.

Theorem 2. *Let $f, g, h \in \mathsf{K}[x_1, \ldots, x_r] \setminus \{0\}$, where K is a field with $|\mathsf{K}| > \deg(f) + \deg(g) \times \deg(h)$, such that $f/g = h$ and f is computed by a division-free formula of size s_f and g is computed by a division-free formula of size s_g with inputs $x_1, \ldots, x_r$. Then h can be computed by division-free formula of size $O((s_f + s_g)^5)$.*

The proof uses interpolation (cf. [Kaltofen 1988, Section 5]). Denote by the total degrees $\delta_f = \deg(f)$, $\delta_g = \deg(g)$ and $\delta_h = \deg(h)$. Note that $\delta_f \leq s_f$ and $\delta_g \leq s_g$. As in Section 3, let $u_1, \ldots, u_r \in \mathsf{K}$ such that $g_0 = g(u_1, \ldots, u_r) \neq 0$. Again as in (5) and (6) using the term basis (2), we compute

$$\sum_{i=0}^{\delta_h} \hom_T(h,i)\lambda^i \equiv \Big(\sum_{i=0}^{\delta_f} \hom_T(f,i)\lambda^i\Big) \Big/ \Big(\sum_{i=0}^{\delta_g} \hom_T(g,i)\lambda^i\Big) \pmod{\lambda^{\delta_h+1}}$$

$$\equiv \underbrace{\frac{1}{g_0}\Big(\sum_{i=0}^{\delta_f} \hom_T(f,i)\lambda^i\Big) \cdot \sum_{l=0}^{\delta_h} \Big(-\frac{1}{g_0}\Big)^l \Big(\sum_{i=1}^{\delta_g} \hom_T(g,i)\lambda^i\Big)^l}_{H(x_1,\ldots,x_r,\lambda)} \pmod{\lambda^{\delta_h+1}}. \quad (7)$$

We compute the polynomial H in (7) of degree in λ of no more than $D = \delta_f + \delta_g\delta_h$ by interpolation at $\lambda = v_0, \ldots, v_D \in \mathsf{K}$ as an exact polynomial before truncating modulo λ^{δ_h+1}. Note that for all j we have

$$\sum_{i=0}^{\delta_f} \hom_T(f,i)v_j^i = f(u_1 - v_j(u_1 - x_1), \ldots, u_r - v_j(u_1 - x_r))$$

and

$$\sum_{i=1}^{\delta_g} \hom_T(g,i)v_j^i = g(u_1 - v_j(u_1 - x_1), \ldots, u_r - v_j(u_1 - x_r)) - g_0, \quad (8)$$

and therefore can obtain the values in the inputs $x_1, \ldots, x_r$ by formulas. We have formulas for each $H(x_1, \ldots, x_r, v_j)$ by repeating the formulas (8) no more than $\delta_h(\delta_h+1)/2$ many times. Interpolation is a matrix times vector product and again is done by repeating the formulas for $H(x_1, \ldots, x_r, v_j)$ no more than $D+1$ times. Finally, we add the thus obtained first δ_h+1 coefficients in λ of H. Note that all divisions are by scalars independent on $x_1, \ldots, x_r$.

Theorem 2 together with the well-known parallel circuits for formula evaluation allows the removal of divisions in formulas altogether. When there are division nodes in formulas

with inputs $x_1, \ldots, x_r$, it is assumed that all rational functions in $\mathsf{K}(x_1, \ldots, x_r)$ by which is divided are non-zero. Formulas with a divisions by a generic 0 are naturally excluded. For certain values in K for the inputs $x_1, \ldots, x_r$ a zero division can occur.

Theorem 3. *There exists a real constant $\gamma > 0$ with the following properties. Let $h \in \mathsf{K}[x_1, \ldots, x_r]$ be computed by a formula (with divisions) of size s with inputs $x_1, \ldots, x_r$. Assume K is a field with $|\mathsf{K}| > s^\gamma$. Then h can be computed by division-free formula of size $O(s^\gamma)$.*

In the following proof γ_1, γ_2 and γ_3 are fixed positive real constants. The proof observes that h is computed by a circuit V_1 of fan-in at most 2 with divisions of depth $\leq \gamma_1 \log(s)$ ([Kosaraju and Delcher 1988] and the references there). By computing unreduced numerator and denominator polynomials for each node separately, we have two division-free circuits V_2 and V_3 of depth $\leq \gamma_2 \log(s)$ that compute polynomials f and g such that $f/g = h$. We can convert V_2 and V_3 into division-free formulas of depth $\leq \gamma_2 \log(s)$, hence of size $< 2^{\gamma_2 \log(s)+1} = O(s^{\gamma_3})$, which also bounds the degrees of f and g. Applying Theorem 2 to both formulas yields Theorem 3.

References

Berkowitz, Stuart J. On computing the determinant in small parallel time using a small number of processors. *Inform. Process. Letters*, 18:147–150, 1984.

Canny, J. Generalized characteristic polynomials. *J. Symbolic Comput.*, 9(3):241–250, 1990.

Cantor, D. G. and Kaltofen, E. On fast multiplication of polynomials over arbitrary algebras. *Acta Inform.*, 28(7):693–701, 1991.

Chistov, A. L. Fast parallel calculation of the rank of matrices over a field of arbitrary characteristic. In *Proc. FCT '85*, volume 199 of *Lect. Notes Comput. Sci.*, pages 63–69, Heidelberg, Germany, 1985. Springer Verlag.

Coppersmith, D. and Winograd, S. Matrix multiplication via arithmetic progressions. *J. Symbolic Comput.*, 9(3):251–280, 1990. Special issue on complexity theory.

Eisenbud, D., Schreyer, F.-O., and Weyman, J. Resultants and Chow forms via exterior syzygies. *J. Amer. Math. Soc.*, 16:537–579, 2003.

von zur Gathen, Joachim and Gerhard, J. *Modern Computer Algebra*. Cambridge University Press, Cambridge, New York, Melbourne, 1999. ISBN 0-521-64176-4. Second edition 2003.

Kaltofen, E. Greatest common divisors of polynomials given by straight-line programs. *J. ACM*, 35(1):231–264, 1988.

Kaltofen, Erich and Villard, Gilles. On the complexity of computing determinants. *Computational Complexity*, 13(3-4):91–130, 2004.

Khetan, Amit. Determinantal formula for the Chow form of a toric surface. In Mora, T., editor, *Proc. 2002 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'02)*, pages 145–150, New York, N. Y., 2002. ACM Press. ISBN 1-58113-484-3.

Khetan, Amit, Song, Ning, and Goldman, Ron. Sylvester A-resultants for bivariate polynomials with planar newton polygons. In Gutierrez, Jaime, editor, *ISSAC 2004 Proc. 2004 Internat. Symp. Symbolic Algebraic Comput.*, pages 205–212, New York, N. Y., 2004. ACM Press. ISBN 1-58113-827-X.

Kosaraju, S. R. and Delcher, A. L. Optimal parallel evaluation of tree-structured computations by raking. In *Proc. AWOC 88*, volume 319 of *Lect. Notes Comput. Sci.*, pages 101–110, Heidelberg, Germany, 1988. Springer Verlag.

Macaulay, F. S. *The Algebraic Theory of Modular Systems*. Number 19 in Cambridge Tracts. The University Press, Cambridge, Great Britain, 1916. Reissued in the Cambridge Mathematical Library with an Introduction by Paul Roberts, 1994.

Malod, Guillaume and Portier, Natacha. Characterizing Valiant's algebraic complexity classes. *Journal of Complexity*, 2007. to appear
URL http://www.ens-lyon.fr/LIP/Pub/Rapports/RR/RR2005/RR2005-44.pdf.

Miller, G. L., Ramachandran, V., and Kaltofen, E. Efficient parallel evaluation of straight-line code and arithmetic circuits. *SIAM J. Comput.*, 17(4):687–695, 1988.

Sasaki, T. and Murao, H. Efficient Gaussian elimination method for symbolic determinants and linear systems. *ACM Trans. Math. Software*, 8(3):277–289, 1982.

Schwartz, J. T. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27:701–717, 1980.

Strassen, V. Vermeidung von Divisionen. *J. reine u. angew. Math.*, 264:182–202, 1973. In German.

Toda, Seinosuke. Classes of arithmetic circuits capturing the complexity of computing the determinant. *IEICE Transactions on Information and Systems*, E75-D(1):116–124, January 1992.

Valiant, L., Skyum, S., Berkowitz, S., and Rackoff, C. Fast parallel computation of polynomials using few processors. *SIAM J. Comp.*, 12(4):641–644, 1983.

Valiant, L. G. Completeness classes in algebra. In *Proc. 11th Annual ACM Symp. Theory Comput.*, pages 249–261, New York, N.Y., 1979. ACM.

Zippel, Richard. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation*, volume 72 of *Lect. Notes Comput. Sci.*, pages 216–226, Heidelberg, Germany, 1979. Springer Verlag. Proc. EUROSAM '79.