ADDITION REQUIREMENTS FOR RATIONAL FUNCTIONS

David G. Kirkpatrick[†]
Cornell University

Zvi M. Kedem[§]
M.I.T.

TR 75-255

August 1975

Department of Computer Science
Cornell University
Ithaca, New York 14853

[†] This research was done in part at the University of Toronto, with the support of the National Research Council of Canada.

[§] This research was supported in part by the National Science Foundation grant GP 22796 A2.

Addition Requirements for Rational Functions

by David G. Kirkpatrick (Cornell University) †
and Zvi M. Kedem (Massachusetts Inst. of Technology) §

Abstract

A notion of rank or independence for arbitrary sets of rational functions is developed, which bounds from below the number of additions and subtractions required of all straight-line algorithms which compute those functions. This permits a uniform derivation of the best lower bounds known for a number of familiar sets of rational functions.

The result is proved without the use of substitution arguements. This not only provides an interesting contrast to standard approaches for arithmetic lower bounds, but also allows the algebraic setting to be somewhat generalized.

Keywords: additions, algorithms, analysis of algorithms, arithmetic complexity, computational complexity, dimensionality, lower bounds, matrix multiplication, optimality, polynomials, rational functions.

CR Categories: 5.12, 5.25

† This research was done in part at the University of Toronto, with the support of the National Research Council of Canada.

§ This research was supported in part by National Science Foundation grant GP 22796 A2.

I. Introduction

A central problem in arithmetic complexity is to take some set of rational functions and determine a lower bound on the number of arithmetic operations which are sufficient to compute the functions. It is a symptom of our lack of understanding of the interaction between multiplicative operations (i.e. multiplication and division) and additive operations (i.e. addition and subtraction), that this problem is rarely treated in its full generality. Indeed, most research in arithmetic complexity has focussed on one, most often the multiplicative, operation type.

While there is evidence that multiplication is inherently more difficult than addition, this does not justify the relative lack of attention paid to additive complexity. This lack is perhaps most effectively illustrated by the fact that, prior to the work presented in this paper, there did not exist any general framework for directly proving a non-trivial lower bound on the additive complexity of the simplest of expressions, $a_1 + a_2 + \ldots + a_n$. While it is this lack which motivates our study, we emphasize that both our techniques and our results tend to complement as well as supplement previous work in arithmetic complexity.

We shall first present some basic definitions and a survey of related work. Section III describes the algebraic setting for our work, and introduces our notion of independence. This notion is developed in the context of multivariate polynomials in Section IV, and extended to general rational functions in

Section V. Section VI contains applications of our central result to a number of common arithmetic expressions. Finally, in Section VII, we mention a few open questions related to our work.

II. Related Work

If F is any field and $\underset{\sim}{a} \triangleq a_1, \ldots, a_n$ is a sequence of distinct indeterminates over F, then elements of $F(\underset{\sim}{a})^{\dagger}$ are called rational functions in $a_1, \ldots, a_n$ over F.

Following Winograd [8], we say that $\mathcal{A}$ is a (rational) algorithm over $(F(\underset{\sim}{a}), G)$, computing $\Psi \subseteq F(\underset{\sim}{a})$ given G, if:

(1) $G \subseteq F(\underset{\sim}{a})$

(2) $\mathcal{A}$ is a finite sequence of pairs $(\alpha_1, \beta_1), \ldots, (\alpha_t, \beta_t)$ where either

(a) $\alpha_i = (y)$ and $\beta_i = y$, where $y \in G$

or (b) $\alpha_i = (\circ, j, k)$ where $\circ \in \{+, -, \times, /\}$, $j, k < i$, and $\beta_i = \beta_j \circ \beta_k$. Furthermore, if $\circ = /$ then $\beta_k \neq 0$.

and (3) $\Psi \subseteq \{\beta_1, \ldots, \beta_t\}$.

If we restrict (2)(b) so that $\circ \in \{+, -, \times\}$, then we say that $\mathcal{A}$ is a polynomial algorithm over $(F(\underset{\sim}{a}), G)$.

We will denote by $\nu(\mathcal{A})$ the number of additions and subtractions in $\mathcal{A}$.

---

† $F(\underset{\sim}{a})$ denotes the field extension of F by the indeterminates $a_1, \ldots, a_n$.

In pioneering the area of arithmetic complexity, Ostrowski [7], considered the problem of determining both additive and multiplicative operation requirements for computing a general n-th degree polynomial. Using straightforward substitution techniques, he showed that for polynomial algorithms, n additive operations are necessary.

Techniques which apply to more general classes of functions are presented by both Belaga [1] and Winograd [8]. Belaga employs the notion of degrees of freedom for rational functions in a single indeterminate. The degree of freedom of such an expression corresponds to the number of algebraically independent coefficients.

THEOREM A (Belaga [1])

Any algorithm which contains p additions, computes a rational expression with at most p+1 degrees of freedom.

In a sense, Belaga's result is restricted by a dual theorem due to Motzkin [6] which relates the same degrees of freedom to multiplicative requirements. Combining the two, we find that if a function can be computed using k multiplication/divisions then the best lower bound on additions that can be obtained by degrees of freedom arguements is about 2k. Degrees of freedom arguements provide tight bounds for polynomials whose terms are all algebraically independent, but they can easily fail to do so if this condition is relaxed. For example, the polynomial $a^2x^2+ab^3x+b^5$ requires two additions despite having algebraically dependent coefficients.

Winograd [8] deals with the computation of functions which are linear in the indeterminates $x_1,\ldots,x_n$. A set of such functions can be expressed as a matrix-vector multiplication, $\Phi\underset{\sim}{x}$, where Φ is a $t\times n$ matrix whose elements are drawn from some field F, and $\underset{\sim}{x}$ denotes the vector $(x_1,\ldots,x_n)$. Winograd's theorem can be stated as,

<u>THEOREM B</u> (Winograd [8])

Any algorithm over $(F(\underset{\sim}{x}), F\cup \underset{\sim}{x})$ which computes the product $\Phi\underset{\sim}{x}$, requires at least $N(\Phi)-t$ addition/subtractions, (where $N(\Phi)$ is the column rank of Φ with respect to a rational subfield $G\subset F$).

Actually, this can be strengthened in the case that $F=G(y_1,\ldots,y_t)$ and G is a subfield of the complex numbers. In this case, Theorem B holds for all algorithms over $(F(\underset{\sim}{x}), F\cup G(\underset{\sim}{x}))$, which means that preconditioning of the set $\underset{\sim}{x}$ is not charged. Winograd also has the following unpublished result,

<u>THEOREM C</u> (Winograd [9])

Any polynomial algorithm over $(F(\underset{\sim}{x}), F\cup \underset{\sim}{x})$ which computes the product $\Phi\underset{\sim}{x}$, requires at least $N^*(\Phi)-t$ addition/subtractions, (where $N^*(\Phi)$ is the number of non-zero columns in Φ).

Since $N^*(\Phi)\geqq N(\Phi)$, this gives a uniformly stronger bound than Theorem B, at the cost of restricting the class of algorithms.

The principle advantage of Winograd's framework is that it allows a straightforward analysis of problems which concern the computation of a family of expressions. The obvious drawback is that many problems cannot be advantageously expressed in this framework. For example, it appears that Theorem B would give a lower bound of zero for both the expression $x_1+\ldots+x_n$ and the pair of expressions, ax_1-bx_2, bx_1+ax_2 (complex product).[†]

In this paper we present yet another complexity measure which has its roots in algebraic independence. However, unlike the degrees of freedom measure and Winograd's framework, our notion of independence is not dependent on the structure of expressions. Furthermore, a straightforward application of our central result generalizes both Theorems B and C.

---

† The only straightforward adaptations are

$$(1,\ldots,1)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}a & -b\\ b & a\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}.$$

III. Algebraic Preliminaries

Let D be any integral domain [†] and let F be the quotient field of D. Let $\underset{\sim}{a} \triangleq a_1,\ldots,a_n$ and $\underset{\sim}{b} \triangleq b_1,\ldots,b_m$ be sequences of distinct indeterminates over F. We are interested in finding lower bounds on the number of additions and subtractions required to compute finite subsets of $F(\underset{\sim}{a})$ using algorithms over $(F(\underset{\sim}{a}) , D \cup \underset{\sim}{a})$.

Let N and Q denote the natural and rational numbers respectively. Let $N\langle\underset{\sim}{a},\underset{\sim}{b}\rangle \triangleq \{ a_1^{\lambda_1} \ldots a_n^{\lambda_n} b_1^{\delta_1} \ldots b_m^{\delta_m} \mid \lambda_i,\delta_i \in N \}$. We make use of the injection $\Theta: N\langle\underset{\sim}{a},\underset{\sim}{b}\rangle \to Q^{n+m}$ defined by

$$\Theta(a_1^{\lambda_1} \ldots a_n^{\lambda_n} b_1^{\delta_1} \ldots b_m^{\delta_m}) \triangleq (\lambda_1,\ldots,\lambda_n,\delta_1,\ldots,\delta_m).$$

Θ extends to subsets $X \subseteq N\langle\underset{\sim}{a},\underset{\sim}{b}\rangle$ by $\Theta(X) \triangleq \{ \Theta(X_i) \mid X_i \in X \}$. Thus Θ maps a set of monomials onto a set of vectors in Q^{n+m}. So, if we let $\rho_V(S)$ denote the vector-rank [§] of any finite subset $S \subset Q^{n+m}$, then we can define the monomial-rank (denoted ρ_M) of any subset $X \subset N\langle\underset{\sim}{a},\underset{\sim}{b}\rangle$ as $\rho_M(X) \triangleq \rho_V(\Theta(X))$.

Two important properties of monomial-rank are summarized in the following lemma.

---

† An integral domain is a ring D for which $uv = 0 \Rightarrow u = 0$ or $v = 0$, for all $u,v \in D$.

§ The vector-rank of S is just the dimension of the subspace $\{ \Sigma\, \lambda_i s_i \mid \lambda_i \in Q \text{ and } s_i \in S \} \subseteq Q^{n+m}$.

LEMMA 1. Let $Y \subseteq X \subseteq N\langle \underset{\sim}{a}, b_1, \ldots, b_k \rangle$ and $e \in N\langle \underset{\sim}{a} \rangle$.

Then, (a) $\rho_M(Y) \leqq \rho_M(X) \leqq \rho_M(Y) + |X| - |Y|$ †

and (b) $\rho_M(X \cup \{eb_{k+1}\}) = \rho_M(X) + 1.$

Proof: These follow by straightforward applications of the definitions. □

IV. Computing Multivariate Polynomials

We start by restricting our attention to the computation of arbitrary elements of $D[\underset{\sim}{a}]$ §, using polynomial algorithms over ($F(\underset{\sim}{a})$, $D \cup \{\underset{\sim}{a}\}$). If R is any ring, let R^+ denote the set $R - \{0\}$.

Let $E \in D[\underset{\sim}{a},\underset{\sim}{b}]^+$. We define the term set of E, $\overline{E}$, by

$$\overline{E} \triangleq \left\{ e \;\middle|\; \begin{array}{l} e \in N\langle \underset{\sim}{a},\underset{\sim}{b} \rangle \text{ and } e \text{ appears with} \\ \text{a non-zero coefficient in } E \end{array} \right\}$$

We can now define the expression-rank (denoted ρ_E) of a set of expressions $\{E_1, \ldots, E_t\} \subset D[\underset{\sim}{a}]$ as

$$\rho_E(E_1, \ldots, E_t) \triangleq \rho_M(\overline{E_1 b_1 + \ldots + E_t b_t}) - t.$$

That is, in order to find the expression-rank of a set of expressions one must first combine the expressions, using new indeterminates, and then find the monomial-rank of the term set of the resulting expression.

---

† $|X|$ denotes the cardinality of the set X.

§ $D[\underset{\sim}{a}]$ denotes the ring extension of D by the indeterminates $a_1, \ldots, a_n$.

The following simple properties of the expression-rank of a set of expressions should indicate its potential as a measure of arithmetic complexity.

<u>LEMMA 2</u>. Let $E_1,\ldots,E_{k+1} \in D[\underset{\sim}{a}]^+$ and let $H \in D^+ \cup \{\underset{\sim}{a}\}$.
Then, for all $1 \leqq i,j \leqq k$,

(a) $\rho_E(E_1,\ldots,E_k,E_{k+1}) \geqq \rho_E(E_1,\ldots,E_k)$

(b) $\rho_E(E_1,\ldots,E_k,H) = \rho_E(E_1,\ldots,E_k)$

(c) $\rho_E(E_1,\ldots,E_k,E_i \times E_j) = \rho_E(E_1,\ldots,E_k)$

and (d) $\rho_E(E_1,\ldots,E_k,E_i \pm E_j) \leqq \rho_E(E_1,\ldots,E_k) + 1$

Remark: These statements assert that, (a) expression-rank is not decreased by the addition of new expressions, (b),(c) the addition of "free" expressions or products of earlier expressions does not alter the expression-rank, and (d) the addition of an expression formed by addition/subtraction of earlier expressions can increase the expression-rank by at most one.

<u>Proof</u>: Let e_i° and e_i denote a fixed and an arbitrary element of $\overline{E}_i$, respectively. Then,

(a) $\rho_E(E_1,\ldots,E_k,E_{k+1})$

$$\begin{aligned}
&\triangleq \rho_M(\overline{E_1 b_1 + \ldots + E_{k+1} b_{k+1}}) - (k+1) \\
&= \rho_M(\overline{E_1 b_1 + \ldots + E_k b_k} \cup \overline{E_{k+1} b_{k+1}}) - (k+1) \\
&\geqq \rho_M(\overline{E_1 b_1 + \ldots + E_k b_k} \cup \{ e_{k+1}^\circ b_{k+1} \}) - (k+1) && \text{, LEM 1(a)} \\
&= \rho_M(\overline{E_1 b_1 + \ldots + E_k b_k}) - k && \text{, LEM 1(b)} \\
&\triangleq \rho_E(E_1,\ldots,E_k)
\end{aligned}$$

(b) Since $H \in D^+ \cup \{\underset{\sim}{a}\}$, $\overline{H}$ contains exactly one element which we denote by h. Hence,

$$
\begin{aligned}
&\rho_E(\, E_1,\ldots,E_k,H \,) \\
&\qquad \triangleq \rho_M(\, \overline{E_1b_1+\ldots+E_kb_k+Hb_{k+1}} \,) - (k+1) \\
&\qquad = \rho_M(\, \overline{E_1b_1+\ldots+E_kb_k} \cup \{hb_{k+1}\} \,) - (k+1) \\
&\qquad = \rho_M(\, \overline{E_1b_1+\ldots+E_kb_k} \,) - k \qquad\qquad \text{, LEM 1(b)} \\
&\qquad \triangleq \rho_E(\, E_1,\ldots,E_k \,)
\end{aligned}
$$

(c) Since $e_ie_jb_{k+1} = (e_i^\circ e_j^\circ b_{k+1})(e_ib_i)(e_i^\circ b_i)^{-1}(e_jb_j)(e_j^\circ b_j)^{-1}$ it follows that $\Theta(\, \overline{E_ib_i} \cup \overline{E_jb_j} \cup \{e_i^\circ e_j^\circ b_{k+1}\} \,)$ generates all of $\Theta(\, \overline{(E_i\times E_j)b_{k+1}} \,)$. Hence,

$$
\begin{aligned}
&\rho_M(\, \overline{E_1b_1+\ldots+E_kb_k} \cup \overline{(E_i\times E_j)b_{k+1}} \,) \\
&\qquad = \rho_M(\, \overline{E_1b_1+\ldots+E_kb_k} \cup \{e_i^\circ e_j^\circ b_{k+1}\} \,)
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
&\rho_E(E_1,\ldots,E_k,E_i\times E_j \,) \\
&\qquad \triangleq \rho_M(\, \overline{E_1b_1+\ldots+E_kb_k+(E_i\times E_j)b_{k+1}} \,) - (k+1) \\
&\qquad = \rho_M(\, \overline{E_1b_1+\ldots+E_kb_k} \cup \{e_i^\circ e_j^\circ b_{k+1}\} \,) - (k+1) \\
&\qquad = \rho_M(\, \overline{E_1b_1+\ldots+E_kb_k} \,) - k \qquad\qquad \text{, LEM 1(b)} \\
&\qquad \triangleq \rho_E(\, E_1,\ldots,E_k \,)
\end{aligned}
$$

(d) Since $e_ib_{k+1} = (e_i^\circ b_{k+1})(e_ib_i)(e_i^\circ b_i)^{-1}$ and $e_jb_{k+1} = (e_j^\circ b_{k+1})(e_jb_j)(e_j^\circ b_j)^{-1}$ it follows that $\Theta(\, \overline{E_ib_i} \cup \overline{E_jb_j} \cup \{e_i^\circ b_{k+1}, e_j^\circ b_{k+1}\} \,)$ generates all of $\Theta(\, \overline{(E_i\pm E_j)b_{k+1}})$.

Hence, $\rho_M(\, \overline{E_1b_1+\ldots+E_kb_k} \cup \overline{(E_i\pm E_j)b_{k+1}})$

$$\leqq \rho_M(\, \overline{E_1b_1+\ldots+E_kb_k} \cup \{e_i^\circ b_{k+1}, e_j^\circ b_{k+1}\} \,)$$

Consequently,

$$\begin{aligned}
&\rho_E(\ E_1,\ldots,E_k,E_i\pm E_j\ ) \\
&\qquad \triangleq \rho_M(\ \overline{E_1b_1+\ldots+E_kb_k+(E_i\pm E_j)b_{k+1}}\ ) - (k+1) \\
&\qquad \leqq \rho_M(\ \overline{E_1b_1+\ldots+E_kb_k} \cup \{e_i^\circ b_{k+1},e_j^\circ b_{k+1}\}\ ) - (k+1) \\
&\qquad \leqq \rho_M(\ \overline{E_1b_1+\ldots+E_kb_k}\ ) - k + 1 \qquad\qquad , \text{LEM } 1(a) \\
&\qquad \triangleq \rho_M(\ E_1,\ldots,E_k\ ) + 1 \qquad\qquad \square
\end{aligned}$$

It is now possible to give a straightforward proof of the following:

THEOREM 1.

Let $\mathcal{C} \triangleq (\alpha_1,\beta_1),\ldots,(\alpha_t,\beta_t)$ be any polynomial algorithm over $(\ F(\underset{\sim}{a}),D \cup\{\underset{\sim}{a}\}\ )$. Then, $\nu(\mathcal{C}) \geqq \rho_E(\ \beta_1,\ldots,\beta_t\ )$

Remark: This asserts that the number of addition/subtractions in a polynomial algorithm $\mathcal{C}$ is at least the expression-rank of the set of expressions computed by $\mathcal{C}$.

Proof: (by induction on t)

(t=1) In this case $\beta_1 \in D^+ \cup \{\underset{\sim}{a}\}$, and hence $\rho_E(\ \beta_1\ ) = 0$. Thus the theorem holds trivially.

($t \leqq s$) Assume that the theorem holds for all $t \leqq s$.

(t=s+1) Let $\mathcal{C}' = (\alpha_1,\beta_1),\ldots,(\alpha_s,\beta_s)$. It follows from the induction hypothesis that $\nu(\mathcal{C}') \geqq \rho_E(\ \beta_1,\ldots,\beta_s\ )$.

There are three cases to consider:

(i) The t-th step introduces a new input. That is,

$$\alpha_t = (H) \quad \text{and} \quad \beta_t = H \in D^+ \cup \{\underset{\sim}{a}\}.$$

Then, $\rho_E(\beta_1,\ldots,\beta_t)$

$$= \rho_E(\beta_1,\ldots,\beta_s) \qquad \text{, LEM 3(b)}$$

$$\leqq \nu(\mathcal{A}') = \nu(\mathcal{A})$$

(ii) The t-th step is a multiplication. That is,

$$\alpha_t = (\times,i,j) \text{ and } \beta_t = \beta_i \times \beta_j$$

Then, $\rho_E(\beta_1,\ldots,\beta_t)$

$$= \rho_E(\beta_1,\ldots,\beta_s) \qquad \text{, LEM 3(c)}$$

$$\leqq \nu(\mathcal{A}') = \nu(\mathcal{A})$$

(iii) The t-th step is an addition/subtraction. That is,

$$\alpha_t = (\pm,i,j) \text{ and } \beta_t = \beta_i \pm \beta_j$$

Then, $\rho_E(\beta_1,\ldots,\beta_t)$

$$\leqq \rho_E(\beta_1,\ldots,\beta_s) + 1 \qquad \text{, LEM 3(d)}$$

$$\leqq \nu(\mathcal{A}') + 1 = \nu(\mathcal{A})$$

Hence, the hypothesis holds for t=s+1, and by induction the theorem is true for all $t \geqq 1$. □

<u>COROLLARY 1.</u> If $\mathcal{A}$ is any polynomial algorithm over $(F(\underset{\sim}{a}) , D \cup \{\underset{\sim}{a}\})$ which computes the expressions $E_1,\ldots,E_k \in D[\underset{\sim}{a}]^+$, then $\nu(\mathcal{A}) \geqq \rho_E(E_1,\ldots,E_k)$.

<u>Proof</u>: If $\mathcal{A} = (\alpha_1,\beta_1),\ldots,(\alpha_t,\beta_t)$, then by the theorem $\nu(\mathcal{A}) \geqq \rho_E(\beta_1,\ldots,\beta_t)$. But, by definition $\{E_1,\ldots,E_k\} \subseteq \{\beta_1,\ldots,\beta_t\}$, so by lemma 3(a),

$$\rho_E(\beta_1,\ldots,\beta_t) \geqq \rho_E(E_1,\ldots,E_k)$$ □

V. Computing Rational Functions

We now remove our earlier restrictions and consider the computation of arbitrary finite subsets of $F(\underset{\sim}{a})^{+}$, using general algorithms over $(F(\underset{\sim}{a}) , D \cup \{\underset{\sim}{a}\})$.

Given any rational algorithm, we can construct a polynomial algorithm which simulates the first by keeping track of the numerator and denominator of every intermediate expression. The following lemma shows that this can be done without increasing the number of addition/subtractions.

LEMMA 3. Let $\mathcal{a} = (\alpha_1,\beta_1),\ldots,(\alpha_t,\beta_t)$ be any rational algorithm over $(F(\underset{\sim}{a}) , D \cup \{\underset{\sim}{a}\})$. Then, there exists a polynomial algorithm, $\mathcal{a}' = (\alpha_1',\beta_1'),\ldots,(\alpha_{2t}',\beta_{2t}')$, where $\nu(\mathcal{a}') = \nu(\mathcal{a})$ and for all i satisfying $1 \leqq i \leqq t$, $(\beta_{2i-1}'/\beta_{2i}') = \beta_i$.

Proof: (by induction on t)

We shall first make two assumptions concerning the form of $\mathcal{a}$, neither of which affect the generality of our arguements:

(i) We assume that $(\alpha_1,\beta_1) = ((1),1)$. That is, the first step of $\mathcal{a}$ introduces the constant 1.

(ii) We assume that all addition/subtraction steps are of the form (α_i,β_i) where $\alpha_i = (\pm,j,1)$ and $\beta_i = \beta_j \pm 1$, for some $j < i$.

Given the identity, $\beta_j \pm \beta_k = ((\beta_j/\beta_k) \pm 1) \times \beta_k$, it follows that both of these assumptions can be ensured without modifying the number of addition/subtractions in an algorithm.

(t=1) Since $(\alpha_1,\beta_1) = ((1),1)$, it suffices to make

$$\alpha'_1 = \alpha'_2 = (1) \quad \text{and} \quad \beta'_1 = \beta'_2 = 1$$

($t \leqq s$) Assume that the theorem holds for all $t \leqq s$.

(t=s+1) Let $\mathcal{B} = (\alpha_1,\beta_1),\ldots,(\alpha_s,\beta_s)$ and let $\mathcal{B}' = (\alpha'_1,\beta'_1),\ldots,(\alpha'_{2s},\beta'_{2s})$ satisfy the induction hypothesis. There are four cases to consider:

(i) The t-th step introduces a constant. That is,

$$(\alpha_t,\beta_t) = ((H),H) \quad \text{where} \quad H \in D^+ \cup \{\underset{\sim}{a}\}.$$

Then, let $(\alpha'_{2t-1},\beta'_{2t-1}) = ((H),H)$ and $(\alpha'_{2t},\beta'_{2t}) = ((1),1)$.

(ii) The t-th step is a multiplication. That is,

$$(\alpha_t,\beta_t) = ((\times,i,j),\beta_i\times\beta_j).$$

Then, let $(\alpha'_{2t-1},\beta'_{2t-1}) = ((\times,2i-1,2j-1),\beta'_{2i-1}\times\beta'_{2j-1})$

and $(\alpha'_{2t},\beta'_{2t}) = ((\times,2i,2j),\beta'_{2i}\times\beta'_{2j})$

(iii) The t-th step is a division. That is,

$$(\alpha_t,\beta_t) = ((/,i,j),\beta_i/\beta_j).$$

Then, let $(\alpha'_{2t-1},\beta'_{2t-1}) = ((\times,2i-1,2j),\beta'_{2i-1}\times\beta'_{2j})$

and $(\alpha'_{2t},\beta'_{2t}) = ((\times,2i,2j-1),\beta'_{2i}\times\beta'_{2j-1})$

(iv) The t-th step is an addition/subtraction. That is,

$$(\alpha_t,\beta_t) = ((\pm,i,1),\beta_i\pm 1).$$

Then, let $(\alpha'_{2t-1},\beta'_{2t-1}) = ((\pm,2i-1,2i),\beta'_{2i-1}\pm\beta'_{2i})$

and $(\alpha'_{2t},\beta'_{2t}) = ((\times,2i,1),\beta'_{2i}\times 1)$

In all cases let $\mathcal{A}' = \mathcal{B}',(\alpha'_{2t-1},\beta'_{2t-1}),(\alpha'_{2t},\beta'_{2t})$.
It follows from our construction that $\nu(\mathcal{A}') = \nu(\mathcal{A})$, and

$(\beta'_{2t-1}/\beta'_{2t}) = \beta_t$. Hence, the hypothesis holds for t=s+1, and by induction the lemma holds for all $t \geqq 1$. □

<u>COROLLARY 2</u>. Let $\mathcal{Q} = (\alpha_1,\beta_1),\ldots,(\alpha_t,\beta_t)$ be any rational algorithm over ($F(\underset{\sim}{a})$, $D \cup \{\underset{\sim}{a}\}$), and suppose $\beta_i = A_i/B_i$, where $A_i, B_i \in D[\underset{\sim}{a}]^+$ are relatively prime. Then, there exist polynomials $C_1,\ldots,C_t \in D[\underset{\sim}{a}]^+$ and a polynomial algorithm $\mathcal{Q}'$, over ($F(\underset{\sim}{a})$, $D \cup \{\underset{\sim}{a}\}$), such that $\nu(\mathcal{Q}') = \nu(\mathcal{Q})$ and $\mathcal{Q}'$ computes the polynomials $A_1C_1, B_1C_1, \ldots, A_tC_t, B_tC_t$.

<u>Proof</u>: Let $\mathcal{Q}'$ be constructed as in the lemma. Then $\beta'_{2i-1}/\beta'_{2i} = A_i/B_i$. But, A_i and B_i relatively prime implies that, for some $C_i \in D[\underset{\sim}{a}]^+$, $\beta'_{2i-1} = A_iC_i$ and $\beta'_{2i} = B_iC_i$. □

As a result of the preceding corollary, we are led to ask whether it is possible for the expression-rank of a sequence of expressions to be decreased through multiplication by non-zero polynomials. The following two lemmas provide the desired answer.

Suppose $X, Y \in D[\underset{\sim}{a}]^+$. We denote by $\overline{X}\cdot\overline{Y}$ the set $\{ xy \mid x \in \overline{X} , y \in \overline{Y} \}$. Note, $\overline{XY}$ is contained in but not necessarily equal to $\overline{X}\cdot\overline{Y}$, since some terms may cancel in the product XY.

If $W \subset N\langle\underset{\sim}{a},\underset{\sim}{b}\rangle$, then we define the convex interior of W,

$$I(W) \triangleq \{ z \mid z = \Pi\, w_j^{\lambda_j} ,\ w_j \in W \text{ and } 0 \leqq \lambda_j < 1 \}.$$

We call the set $V(W) \triangleq \{ w \in W \mid w \notin I(W) \}$, the set of vertices of W. Clearly, $I(W) = I(V(W))$.

LEMMA 4.

If $X, Y \in D[\underset{\sim}{a}]^+$, then $V(\overline{X}\cdot\overline{Y}) \subseteq \overline{XY}$

Proof: By definition, $V(\overline{X}\cdot\overline{Y}) \subseteq \overline{X}\cdot\overline{Y}$. Assume $\overline{X}\cdot\overline{Y} - \overline{XY} \neq \phi$, (otherwise there is nothing to prove). Let $e \in \overline{X}\cdot\overline{Y} - \overline{XY}$, i.e., e is cancelled in the product XY. Since D is an integral domain, there must exist distinct $x, x' \in \overline{X}$, and distinct $y, y' \in \overline{Y}$, such that $e = xy = x'y'$. Hence,

$e = (xy')^{1/2}(x'y)^{1/2} \in I(\overline{X}\cdot\overline{Y})$. Thus, $e \notin V(\overline{X}\cdot\overline{Y})$,

and in general $V(\overline{X}\cdot\overline{Y}) \subseteq \overline{XY}$. □

This technical lemma allows us to give a very simple proof the following:

LEMMA 5. Let $X_1,\ldots,X_t,Y_1,\ldots,Y_t \in D[\underset{\sim}{a}]^+$.

Then, $\rho_E(X_1Y_1,\ldots,X_tY_t) \geqq \rho_E(X_1,\ldots,X_t)$.

Proof: Let $W_i \triangleq \overline{X}_i\cdot\overline{Y_ib_i}$, and let $V_i \triangleq V(W_i)$.

Since $V_i \subseteq W_i \subseteq I(W_i) \cup V(W_i) = I(V_i) \cup V_i$, it follows that $\Theta(V_i)$ generates all of $\Theta(W_i)$, and hence

$$\rho_M(V_1 \cup \ldots \cup V_t) = \rho_M(W_1 \cup \ldots \cup W_t) .$$

But, by lemma 4, we know that $V_i \subseteq \overline{X_iY_ib_i}$, and so, by lemma 1(a), $\rho_M(V_1 \cup \ldots \cup V_t) \leqq \rho_M(\overline{X_1Y_1b_1+\ldots+X_tY_tb_t})$.

Thus, if y_i denotes an arbitrary element of $\overline{Y}_i$,

we have $\rho_E(X_1Y_1,\ldots,X_tY_t)$

$$\begin{aligned} &\triangleq \rho_M(\overline{X_1Y_1b_1+\ldots+X_tY_tb_t}) - t \\ &\geqq \rho_M(V_1 \cup \ldots \cup V_t) - t \\ &= \rho_M(W_1 \cup \ldots \cup W_t) - t \end{aligned}$$

$$\triangleq \rho_M(\ \overline{X}_1 \cdot \overline{Y_1 b_1} \cup \ldots \cup \overline{X}_t \cdot \overline{Y_t b_t}\ ) - t$$
$$\geqq \rho_M(\ \overline{X_1 y_1 b_1} \cup \ldots \cup \overline{X_t y_t b_t}\ ) - t \qquad \text{, LEM 1(a)}$$
$$= \rho_M(\ \overline{X_1 b_1' + \ldots + X_t b_t'}\ ) - t$$
$$\triangleq \rho_E(\ X_1, \ldots, X_t\ ) \qquad \square$$

Finally, we have,

<u>THEOREM 2.</u>

Let $\mathcal{Q}$ be any rational algorithm , over ($F(\underset{\sim}{a})$, $D \cup \{\underset{\sim}{a}\}$), which computes the rational functions $A_1/B_1, \ldots, A_k/B_k$, where $A_i, B_i \in D[\underset{\sim}{a}]^+$ are relatively prime. Then ,

$$\nu(\mathcal{Q}) \geqq \rho_E(\ A_1, B_1, \ldots, A_k, B_k\ ).$$

<u>Proof</u>: By corollary 2 , we know that there exist polynomials $C_1, \ldots, C_k \in D[\underset{\sim}{a}]^+$, and a polynomial algorithm $\mathcal{Q}'$, with $\nu(\mathcal{Q}') = \nu(\mathcal{Q})$, such that $\mathcal{Q}'$ computes $A_1C_1, B_1C_1, \ldots, A_kC_k, B_kC_k$. But ,

$$\nu(\mathcal{Q}') \geqq \rho_E(\ A_1C_1, B_1C_1, \ldots, A_kC_k, B_kC_k\ ) \qquad \text{, COR 1}$$
$$\geqq \rho_E(\ A_1, B_1, \ldots, A_k, B_k\ ) \qquad \text{, LEM 5}$$

Hence , $\nu(\mathcal{Q}) \geqq \rho_E(\ A_1, B_1, \ldots, A_k, B_k\ ).$ $\square$

<u>COROLLARY 3.</u> Let $\mathcal{Q}$ be any rational algorithm over ($F(\underset{\sim}{a})$, $D \cup \{\underset{\sim}{a}\}$), which computes the rational functions, $A_1/B_1, \ldots, A_k/B_k$, where $A_i, B_i \in D[\underset{\sim}{a}]^+$ are relatively prime. Then , $\nu(\mathcal{Q}) \geqq \rho_V(\ (\ \overline{A_1 b_1 + B_1 b_2 + \ldots + A_k b_{2k-1} + B_k b_{2k}}\ )\ ) - 2k.$

In the case that the functions to be computed are all multivariate polynomials, the following corollary provides the same bound as corollary 3, while being somewhat less cumbersome to apply.

COROLLARY 4. Let $\mathfrak{a}$ be any rational algorithm over $(F(\underset{\sim}{a}) , D \cup \{\underset{\sim}{a}\})$, which computes the polynomials , $A_1,\ldots,A_k \in D[\underset{\sim}{a}]^+$. Then, $\nu(\mathfrak{a}) \geqq \rho_V(\theta(\overline{A_1 b_1+\ldots+A_k b_k})) - k$.

Proof: This follows directly from Theorem 2 and the definitions, given the equality,

$$\rho_E(A_1,1,A_2,1,\ldots,A_k,1) = \rho_E(A_1,\ldots,A_k)$$

which was established by lemma 2(b). $\square$

VI. Applications

Corollaries 3 and 4 provide straightforward procedures for reducing the problem of determining addition/subtraction requirements for arbitrary sets of rational functions, to the problem of determining the rank of a set of vectors in Q^t. In this way, we can generate the best lower bounds known, in a number of cases optimal bounds, for a large number of familiar arithmetic expressions.†

As before, we let D denote an arbitrary integral domain,

---

† Most of these first appeared in [4] or [5]. Others were given in [3].

and F the quotient field of D. For the sake of uniformity, let I denote the set of symbols formed from $\{a,x,y\}$ by the possible addition of primes or subscripts. I can be thought of as an arbitrarily large pool of distinct indeterminates, and will take the place of the set $\{\underset{\sim}{a}\}$ of the preceding development.

Let $\mathfrak{a}$ be any rational algorithm over ($F(I)$, $D \cup I$). As before, $\nu(\mathfrak{a})$ denotes the number of addition/subtraction steps in $\mathfrak{a}$.

<u>A1.</u> If $\mathfrak{a}$ computes the expression $a_1 + a_2 + \ldots + a_n$, then $\nu(\mathfrak{a}) \geqq n-1$.

<u>Proof:</u> We know, by corollary 4, that

$$\nu(\mathfrak{a}) \geqq \rho_V(\,\Theta(\,\overline{(a_1+\ldots+a_n)b_1}\,)\,) - 1 .$$

But, it is easy to verify that the vectors in $\Theta(\,\overline{(a_1+\ldots+a_n)b_1}\,)$ are all independent, over Q, and hence $\rho_V(\,\Theta(\,\overline{(a_1+\ldots+a_n)b_1}\,)\,) = n$.

□

More generally,

<u>A2.</u> If $\mathfrak{a}$ computes the expression $\dfrac{a_1 + a_2 +\ldots+ a_n}{a_{n+1}+a_{n+2}+\ldots+a_{n+m}}$, then $\nu(\mathfrak{a}) \geqq n+m-2$.

<u>A3.</u> If $\mathfrak{a}$ computes the pair of expressions, $a_1a_3 - a_2a_4$ and $a_1a_4 + a_2a_3$, (the real and imaginary parts of the complex product $(a_1+a_2i)(a_3+a_4i)$), then $\nu(\mathfrak{a}) \geqq 2$.

<u>Proof:</u> By corollary 4, it suffices to verify that

$$\rho_V(\,\Theta(\,\overline{(a_1a_3 - a_2a_4)b_1 + (a_1a_4 + a_2a_3)b_2}\,)\,) = 4 . \qquad \square$$

A4. If $\mathcal{A}$ computes the general rational function,

$$\frac{a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0}{a'_m x^m + a'_{m-1}x^{m-1} + \ldots + a'_0}, \quad \text{then} \quad \nu(\mathcal{A}) \geqq n+m .$$

The following result generalizes both Theorems B and C.

A5. Let Φ be any $t \times n$ matrix over D , and let $N^*(\Phi)$ denote the number of columns of Φ which are not identically zero. If $\mathcal{A}$ computes the matrix-vector product ,

$$\Phi \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \text{then} \quad \nu(\mathcal{A}) \geqq N^*(\Phi) - t .$$

Proof: Let $k = N^*(\Phi)$, and assume, without loss of generality, that the first k columns of Φ are not identically zero. Let $\Phi = \begin{pmatrix} \phi_{11} & \cdots & \phi_{1n} \\ \vdots & & \\ \phi_{t1} & \cdots & \phi_{tn} \end{pmatrix}$ and for $1 \leqq i \leqq k$, define $[i] \triangleq \min \{ j \mid \phi_{ji} \neq 0 \}$. By corollary 4, it suffices to show that $\rho_V(\Theta(\bar{E})) \geqq k$, where

$$\begin{aligned} E \triangleq {} & (\phi_{11}a_1 + \ldots + \phi_{1n}a_n)b_1 \\ & + \cdots + (\phi_{t1}a_1 + \ldots + \phi_{tn}a_n)b_t \end{aligned}$$

But, by the definition of $[i]$, we know that

$$\{ a_1 b_{[1]}, a_2 b_{[2]}, \ldots, a_k b_{[k]} \} \subseteq \bar{E} .$$

Hence, $\rho_V(\Theta(\bar{E})) \geqq \rho_V(\Theta(\{ a_1 b_{[1]}, \ldots, a_k b_{[k]} \}) = k$.

□

A6. Let $A \triangleq (a_{ij})$ and $X \triangleq (x_{ij})$ be $m \times n$ and $n \times p$ matrices, respectively. If $\mathfrak{a}$ computes the matrix product $A \cdot X$, then $\nu(\mathfrak{a}) \geqq (m+p-1)(n-1)$.

In particular, this gives addition/subtraction lower bounds of $m(n-1)$, for the product of an $m \times n$ matrix with an n-vector, and $2n^2-3n+1$, for the product of two $n \times n$ matrices.

A7. Let $X \triangleq \begin{pmatrix} x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_m^n & x_m^{n-1} & \cdots & x_m & 1 \end{pmatrix}$

If $\mathfrak{a}$ computes the matrix-vector product $X \cdot \begin{pmatrix} a_n \\ \vdots \\ a_1 \\ a_0 \end{pmatrix}$, then $\nu(\mathfrak{a}) \geqq n+m-1$.

The obvious interpretation of the above is that it requires at least $n+m-1$ addition/subtractions to evaluate an n-th degree polynomial at m arbitrary points.

A8. Let $P_i \triangleq \sum_{j=0}^{n(i)} a_{ij} x^j$, for $i = 1,\ldots,t$.

If $\mathfrak{a}$ computes the set $P_1,\ldots,P_t$, then $\nu(\mathfrak{a}) \geqq \sum_{i=1}^{t} n(i)$.

A9. If $\mathfrak{a}$ computes the expression $\sum_{i=0}^{n} \sum_{j=0}^{n} a_{ij} x^i y^j$, then $\nu(\mathfrak{a}) \geqq (n+1)^2$

VII. Open Questions

The results and techniques of this paper leave unanswered a number of interesting questions. Some of these, including both a related development using substitution techniques and observations on the structure of algorithms (if they exist) which achieve the lower bounds given by our independence measure, will be considered in future papers.

Of major importance is the problem of developing techniques for lower bounds on additions, which are non-linear in the number of indeterminants. We should mention the initial success of Borodin and Cook [2] in this direction. It is hoped that, perhaps through consideration of the geometrical interpretations employed in lemma 4, our techniques might be modified to provide this kind of bound.

Acknowledgment We would like to thank A.B.Borodin for his encouragement and numerous helpful comments.

References

[1] E.C.Belaga, " On computing polynomials in one variable with initial preconditioning of coefficients," Problemi Kibernetiki 5 (1961), 7-15.

[2] A.B.Borodin and S.A.Cook, "On the number of additions to compute specific polynomials," Proc. 6th Annual ACM Symposium on Theory of Computing (1974), 342-347.

[3] Z.M.Kedem, "Studies in algebraic computational complexity," Doctor of Science thesis, Israel Institute of Technology (December 1973).

[4] D.G.Kirkpatrick, "On the additions necessary to compute certain functions," Proc. 4th Annual ACM Symposium on Theory of Computing (1972), 94-101.

[5] D.G.Kirkpatrick, Technical Report 39, University of Toronto (1972).
- an expanded version of [4].

[6] T.S.Motzkin, "Evaluation of polynomials and evaluation of rational functions," Bulletin of American Mathematical Society 61 (1955), 163.

[7] A.M.Ostrowski, "On two problems in abstract algebra connected with Horner's rule," Studies Presented to R. von Mises (1954), Academic Press, N.Y.

[8] S.Winograd, "On the algebraic complexity of functions," Actes, Congres intern. Math. 3 (1970), 283-288.

[9] S.Winograd, private communication (1971).