Zeitschrift: L'Enseignement Mathématique
Herausgeber: Commission Internationale de l'Enseignement Mathématique
Band: 26 (1980)
Heft: 1-2: L'ENSEIGNEMENT MATHÉMATIQUE

Artikel: LINEAR DISJOINTNESS AND ALGEBRAIC COMPLEXITY
Autor: Baur, Walter / Rabin, Michael O.
Kapitel: 1. Introduction
DOI: https://doi.org/10.5169/seals-51078

LINEAR DISJOINTNESS AND ALGEBRAIC COMPLEXITY

by Walter BAUR and Michael O. RABIN

Dedicated to Ernst Specker on the occasion of his 60th birthday.

1. INTRODUCTION

It is well known that any algorithm for the evaluation of a polynomial

$$f(y) = x_0 + x_1 y + \ldots + x_n y^n , \tag{1}$$

or of an inner product of two vectors

$$(x, y) = x_1 y_1 + \ldots + x_n y_n , \tag{2}$$

requires, under certain natural assumptions such as that $y, x_1, \ldots, x_n$ are algebraically independent over some ground-field F, at least n multiplications. This number of multiplications can of course be achieved by an appropriate algorithm.

Motzkin [3] has introduced the idea of preprocessing the coefficients of a polynomial. In certain situations, for example when we have to evaluate f for many values $y = c_1, y = c_2, \ldots$ of the argument, though these values are not given in advance, it makes sense to compute once and for all certain functions $\alpha_1 (x_1, \ldots, x_n), \ldots, \alpha_n (x_1, \ldots, x_n)$ of the coefficients and use these $\alpha_1, \ldots, \alpha_n$ later on in an algorithm for the calculation of the $f(c_i)$, i.e. $f(y)$. The $\alpha_1, \ldots, \alpha_n$ and the algorithm should be so chosen that the evaluation of $f(y)$ now requires fewer than n multiplications. The cost of this "preprocessing" of the coefficients $x_1, \ldots, x_n$ is then absorbed in the saving in the computations $f(c_1), f(c_2), \ldots$.

Motzkin has shown that preprocessing of the coefficients can lead to evaluation of $f(y)$ in $\left\lceil \frac{n}{2} \right\rceil + 2$ multiplications and $n + 2$ additions. From now on we shall concentrate our attention on the number of multiplications or divisions used in an algorithm. The notation $n\ M/D$ means n multiplications or divisions. We must take into account divisions as well as multiplications because a product xy can be computed by doing two divisions.

Winograd [6] has noted that if in (2) we allow preprocessing on both the x and y then $\lceil \frac{n}{2} \rceil M$ are sufficient. Namely, assume n even and define

$$w(x) = x_1x_2 + x_3x_4 + .. + x_{n-1}x_n .$$

If $w(x)$ and $w(y)$ have been precomputed then

$$(x, y) = (x_1 + y_2)(x_2 + y_1) + \ldots + (x_{n-1} + y_n)(x_n + y_{n-1}) - w(x) - w(y)$$

computes (2) with $\frac{n}{2} M$. There are situations, involving many vectors $x, y, z, \ldots$, and many scalar products, say, $(x, y), (y, z), (x, z), \ldots$, where this idea makes computational sense.

Can the upper bound $\frac{n}{2}$ in the algorithms for $f(y)$ and (x, y) with preprocessing be improved. Can we get lower bound results for these and more general computational problems. We have, of course, to be careful about the preprocessing that we permit. For example, if we permit to form products $x_i y_i$ then no multiplications will later be needed in computing (x, y). Thus preprocessing for (2) should not involve multiplications "mixing" the $x_1, \ldots, x_n$ with the $y_1, \ldots, y_n$, or with y in the case of $f(y)$. It will be seen later that the crux of this paper is a precise determination of the sort of "mixing" that should be avoided so as to yield a good lower-bound result.

In [3] (see also [4]) it is shown that if $F \subseteq K \subseteq K(y)$ and if $x_1, \ldots, x_n \in K$ are algebraically independent over F, then any computation of $f(y)$ which allows the use of any $\alpha_1, \alpha_2, \ldots \in K$ must involve $\frac{n}{2} M/D$, even if a multiplication step $a \cdot b$ is not counted if $a \in F$ or $b \in F$, and a step a/b is not counted when $b \in F$. Similar results hold for polynomials in several variables $y_1, y_2, \ldots$.

Winograd [6] has introduced another lower bound theorem for the case of computations with preprocessing. His theorem involves restrictions on the fields in question, and the conditions (involving topology) for the theorem to hold are difficult to interpret or check in specific cases. The proof in [6] employs topological methods.

In the present paper we observe that the concept of linear disjointness of two fields over a common subfield provides a proper framework for a very general result, Theorem 1, on lower bounds for the number of M/D operations in computations with preprocessing. The result and its simple

proof are expressed in purely algebraic terms. In section 4 we apply Theorem 1 to obtain the known results on lower bounds, as well as new results which do not fall within the scope of previous methods.

2. Basic concepts and the Main Theorem

Let Ω be a field and S a subset of its elements. Following [5, 6], a (straight-line) algorithm or computation in (Ω, S) is a sequence π: $\pi(1), \ldots, \pi(l)$ where for each $1 \leqslant k \leqslant l$ we have $\pi(k) \in S$, or for some $i, j < k$, $\pi(k) = (+, i, j)$ or $(-, i, j)$ or $(\cdot, i, j)$ or $(/, i, j)$.

With π we associate the sequence $r(1), \ldots, r(l)$ of the results of the computation π. The $r(k)$ are all elements of $\Omega \cup \{u\}$. Define $r(1) = \pi(1) \in S$. Inductively, if $r(1), \ldots, r(k-1)$ are already defined we set $r(k) = \pi(k)$ if $\pi(k) \in S$, $r(k) = r(i) + r(j)$ if $\pi(k) = (+, i, j)$, etc. By convention, $r/0 = u + r = u \cdot r = \ldots = u$ for $r \in \Omega \cup \{u\}$.

We say that π computes the elements $a_1, \ldots, a_m \in \Omega$ if there exist $1 \leqslant i_j \leqslant l$, $1 \leqslant j \leqslant m$, so that for the results of π we have $r(i_j) = a_j$, $1 \leqslant j \leqslant m$.

In the sequel we shall be interested in fields $F \subseteq \Omega$ and two intermediate fields E, K. Thus

$$\begin{array}{ccc} & \Omega & \\ \diagup & & \diagdown \\ E & & K \\ \diagdown & & \diagup \\ & F & \end{array} \tag{3}$$

The following concept comes from the theory of fields and from algebraic geometry, see [1, 2].

Definition. The fields E and K are linearly disjoint over F if any $e_1, \ldots, e_m \in E$ which are linearly independent over F are also linearly independent over K, i.e. $\Sigma a_i e_i = 0$, $a_i \in K$, only if $a_i = 0$, $1 \leqslant i \leqslant m$.

As the definition stands, the fields E and K play different roles. It is however easy to see that the above definition implies the analogous statement with the roles of E and K interchanged. (See e.g. [1].)

Our theorem will be about computations π in $(\Omega, E \cup K)$. The fact that we permit using any $\alpha \in E \cup K$ at no computational cost captures, in an algebraic form, the idea of preprocessing.