A LOWER BOUND FOR THE MULTIPLICATION OF POLYNOMIALS MODULO A POLYNOMIAL

Bshouty, Nader H.

http://hdl.handle.net/1880/45742
unknown

Downloaded from PRISM: https://prism.ucalgary.ca
A LOWER BOUND FOR THE MULTIPLICATION OF
POLYNOMIALS

MODULO A POLYNOMIAL

Nader H. Bshouty

Department of Computer Science
University of Calgary
Calgary, Alberta, Canada

ABSTRACT

In [Theoretical Computer Science, 1983], Lempel, Sazonov and Winograd proved the
lower bound

\[
\begin{bmatrix}
2 & 1 \\
q & 1
\end{bmatrix} n - o(n)
\]

for the multiplicative complexity of the multiplication of two polynomials of degree \( n \) 1
modulo an irreducible polynomial \( p \) of degree \( n \) over a finite field \( F \) with \( q \) elements.

In this paper we prove this lower bound holds for any polynomial \( p \) of degree \( n \)

Key Words: multiplicative complexity, quadratic algorithms, linear codes.

1 INTRODUCTION

Let \( F \) be a field and let \( B = \{ B_1, \ldots, B_k \} \) be a set of \( n \times m \)-matrices with entries from \( F \). Let

\( x = (x_1, \ldots, x_m)^t \) and \( y = (y_1, \ldots, y_m t)^t \) be vectors of indeterminates. A quadratic algorithm

over \( F \) that computes the bilinear forms \( x^t B y = (x^t B_1 y, \ldots, x^t B_k y) \) is a straightline algorithm

over \( F \) for \( x^t B y \) such that its nonscalar multiplications are of the shape \( l(x, y) \cdot l'(x, y) \), where

\( l(x, y) \) and \( l'(x, y) \) are linear forms of \( x \) and \( y \). The complexity \( L_F(B) \) of \( B \) is the minimal number

of nonscalar multiplications needed to compute \( x^t B y \) by quadratic algorithms over \( F \). It is known
from [S] that when $F$ is an infinite field, then $L_F(B)$ is the minimal number of non-scalar multiplications/divisions needed to compute $x^TBy$ by straightline algorithms. When $F$ is finite, then it is known from [W] that $L_F(B)$ is the minimal number of non-scalar multiplications needed to compute $x^TBy$ by a straightline algorithm without divisions.

For the vectors $x = (x_0, \ldots, x_{k-1})$ and $y = (y_0, \ldots, y_{k-1})$ we define $x(\alpha) = x_0 + x_1\alpha + \cdots + x_{k-1}\alpha^{k-1}$ and $y(\alpha) = y_0 + y_1\alpha + \cdots + y_{k-1}\alpha^{k-1}$. Let $p(\alpha)$ be a polynomial of degree $n$ over $F$ and define $B(p) = \{B_0, \ldots, B_{k-1}\}$ where

$$\sum_{i=0}^{k-1} (x^T B_i y) \alpha^i = x(\alpha) y(\alpha) \mod p(\alpha).$$

That is, $x^T B_i y$ is the $i+1$ coefficient of $x(\alpha) y(\alpha) \mod p(\alpha)$.

In [LSW] Lemple, Seroussi and Winograd used coding theory to prove the lower bound

$$L_F(B(p)) \geq 2 + \left\lfloor \frac{1}{|F| - 1} \right\rfloor n - o(n),$$

when $p$ is an irreducible polynomial of degree $n$. In [CC], Chudnovsky and Chudnovsky proved the linear upper bound

$$L_F(B(p)) \leq 2 + O\left(\frac{1}{|F|^{1/2}}\right) n.$$

In this paper we generalize the result in [LSW] as follows:

**Theorem.** Let $p \in F[\alpha]$ be any polynomial of degree $n$ over a finite field $F$. Then

$$L_F(B(p)) \geq 2 + \frac{1}{|F| - 1} n - o(n).$$

The method we use involves a combination of the coding method which is used in [LSW], and the substitution method used in [BD].

This paper is organized as follows. In section 2 we give some preliminary results and the connection between linear codes and the complexity of bilinear forms. In section 3 we prove the theorem.

### 2. Preliminary Results

This section contains a survey of some basic concepts that will be employed throughout the
paper.

**Definition 1.** Let $B = \{B_1, \ldots, B_n\}$ be an $n$--set of $n \times n$--matrices and $M, N$ and $K = (K_{i,j})$ be $n \times n$--matrices. We define

$$NB = \{NB_1, \ldots, NB_n\}$$

and $B[K] = \left\{ \sum_{j=1}^{n} K_{i,j} B_j, \ldots, \sum_{j=1}^{n} K_{i,j} B_j \right\}.$

For an $n$--set $C$ of $n \times n$--matrices we write $B = C$ if there exist nonsingular $n \times n$--matrices $N, M$ and $K$ such that

$$NB[K]M = C.$$

**Definition 2.** Let $B = \{B_1, \ldots, B_n\}$ be an $n$--set of $n \times n$ matrices and let $C = \{C_1, \ldots, C_m\}$ be an $m$--set of $m \times m$--matrices. We define

$$B \otimes C = \{\tilde{B}_1, \ldots, \tilde{B}_n, \tilde{C}_1, \ldots, \tilde{C}_m\},$$

where

$$\tilde{B}_i = \begin{bmatrix} B_i & 0_{m \times \alpha} \\ 0_{n \times \alpha} & 0_{m \times \alpha} \end{bmatrix}, \quad \tilde{C}_j = \begin{bmatrix} 0_{n \times \alpha} & 0_{n \times \alpha} \\ 0_{m \times \alpha} & C_j \end{bmatrix}$$

and $0_{s \times r}$ is the $s \times r$ zero matrix.

We also define

$$B \otimes C = \{B_i \otimes C_j \mid i = 1, \ldots, n, j = 1, \ldots, m\},$$

where $\otimes$ is the Kronecker product of matrices.

Let $n$ be an integer. A linear code over $F$ of length $n$ is a linear subspace $C$ of $F^n$. If $\dim C = k$, then $C$ is called an $[n, k]$ code. For $c \in C$ the weight of $c$, denoted by $wt(c)$, is the number of nonzero components of $c$. The minimal weight of $C$ is $\min \{wt(c) \mid c \in C \setminus \{0\}\}$. We say that $C$ is an $[n, k, d]$ code if $C \subseteq F^n$ is a code of dimension $k$ and minimal weight $d$. Let $N_F(k, d)$ be the smallest integer such that there exists an $[N_F(k, d), k, d]$ code. The connection between the linear codes and the complexity of bilinear forms over $F$ is given in the following lemma.

**Lemma 1.** [BD, LW]. Let $B = \{B_1, \ldots, B_n\}$ be a set of $n \times m$ matrices and let $G = \text{Span}_F(B)$ be the linear space over $F$ spanned by the elements of $B$. Let $d = \min_{B' \in G \setminus \{0\}} \text{rank } B'$ and $k = \dim \text{Span}_F(B)$. Then
Lemma 1 is proved in [BD, LW] for the bilinear algorithm model of computation. The proof for the quadratic algorithm model is very similar and will be omitted.

The next lemma gives a lower bound for \( N_F (k, d) \).

**Lemma 2.** (Griesmer Bound [1, p. 59]). We have

\[
N_F (k, d) \geq \sum_{i=0}^{k-1} \left( \frac{d}{|F|^{i+1}} \right) \geq \begin{cases} 
1 + \frac{1}{|F| - 1} - \frac{1}{k - 1 (|F| + 1)} d & \text{if } k \leq \log_{|F|} d \\
1 + \frac{1}{|F| - 1} d + k - \log_{|F|} d - 3, & \text{if } k > \log_{|F|} d.
\end{cases}
\]

Other lower bound techniques known from the literature for the complexity of quadratic algorithms are the following.

**Lemma 3.** [BD]. Let \( B = \{B_1, \ldots, B_{k_1}\} \) and \( C = \{C_1, \ldots, C_{k_2}\} \) be sets of \( n \times m \) matrices. Then

\[
L_F (C \cup B) \geq \dim \text{Span}_F (C) + \min_{\lambda \in \mathbb{F}} L_F \left( \{B_1 + \sum_{j=1}^{k_2} \lambda_{k_1, j} C_j, \ldots, B_{k_1} + \sum_{j=1}^{k_2} \lambda_{k_2, j} C_j\} \right).
\]

**Lemma 4.** Let \( B \) and \( C \) be as in Lemma 3 with \( k_1 = k_2 = n = m \)

1. If \( \text{Span}_F (B) \subseteq \text{Span}_F (C) \), then

\[
L_F (B) \leq L_F (C).
\]

2. If \( B = C \), then

\[
L_F (B) = L_F (C).
\]

(see definition 1 for \( \subseteq \)).

3. PROOF OF THE LOWER BOUND

In this section we prove the theorem stated in section 1.

Let \( B \) be an independent set of matrices. We say that \( B \) is a \((k, l, d)\)-set if \( |B| = k \) and there exists \( k - l \) matrices \( B_1, \ldots, B_{k-l} \in B \) such that for any \( B \in \text{Span}_F (\{B_1, \ldots, B_{k-l}\}) \) we have

\[
\text{rank } B \geq d.
\]

We remind the reader that \( \text{Span}_F (B) \) is the linear space spanned by the elements of \( B \). If \( B \) is a \((k, l, d)\)-set, then \( B \) is a \((k, l', d)\)-set for any \( l' \leq l \). This follows from the fact that,
B ∈ \text{Span}_F (B_1, \ldots, B_{k-1}) \implies B \notin \text{Span}_F (B_1, \ldots, B_{k-1}).

The following lemma will be used to prove the theorem.

**Lemma 5.** Let \(B^{(i)}\), be a \((k_i, l_i, d_i)\)-set for \(i = 1, \ldots, s\) and let \(l = \min_i k_i l_i l\). Then

\[
L_F (B^{(1)} \ominus \cdots \ominus B^{(s)}) \geq \sum_{i=1}^s k_i - s l + N_F \left[ l, \sum_{i=1}^s d_i \right].
\]

**Proof.** Let \(B^{(i)} = \{B_i^{(i)}, \ldots, B_{l_i}^{(i)}, B_{l_i+1}^{(i)}, \ldots, B_{k_i}^{(i)}\} \) for \(i = 1, \ldots, s\), such that, for any \(B \notin \text{Span}_F (\{B_i^{(i)}, \ldots, B_{k_i}^{(i)}\})\),

\[
\text{rank } B \geq d_i. \tag{*}
\]

Consider the following two sets

\[
V_1 = \{\text{diag} (B_i^{(1)}, B_i^{(2)}, \ldots, B_i^{(s)}) \mid i = 1, \ldots, l\}
\]

and

\[
V_2 = (B^{(1)} - D^{(1)}) \ominus \cdots \ominus (B^{(s)} - D^{(s)}),
\]

where

\[
D^{(i)} = \{B_i^{(i)}, \ldots, B_{l_i}^{(i)}\}.
\]

Here, \(\text{diag} (B_i^{(1)}, \ldots, B_{l_i}^{(s)})\) is the block matrix

\[
\begin{pmatrix}
B_i^{(1)} & & \\
& \ddots & \\
& & B_i^{(s)}
\end{pmatrix}
\]

Obviously, \(\text{Span}_F (V_1 \cup V_2)\) is a subset of \(\text{Span}_F (B^{(1)} \ominus \cdots \ominus B^{(s)})\). Therefore, by lemma 4 and lemma 3,

\[
L_F (B^{(1)} \ominus \cdots \ominus B^{(s)}) \geq L_F (V_1 \cup V_2) \geq \dim \text{Span}_F (V_2) + \min_{\lambda \in F^{k_1, \ldots, k_s}} L_F (S_\lambda) \tag{1}
\]

where, for \(\lambda = (\lambda_{i,1}, \ldots, \lambda_{i,k_i})\), we have

\[
S_\lambda = \left\{ \text{diag} \left[ B_i^{(1)} + \sum_{j \in i} \lambda_{i,j} B_j^{(1)}, \ldots, B_i^{(s)} + \sum_{j \in i} \lambda_{i,j} B_j^{(s)} \right] \mid i = 1, \ldots, l \right\}.
\]
Every nonzero element in $\text{Span}_F(S_k)$ is of the form

$$
P = \text{diag} \left( \sum_{i=1}^l \delta_i B_i \lambda_{1,j} B_j \lambda_{2,j} B_j \ldots, \sum_{i=1}^{l_1} \delta_i B_i^{(1)} \sum_{j=1}^{k_1} \lambda_{1,j}^{(1)} B_j^{(1)} \sum_{i=1}^{l_2} \delta_i B_i^{(2)} \sum_{j=1}^{k_2} \lambda_{1,j}^{(2)} B_j^{(2)} \ldots \right),
$$

where not all $\delta_i$ are zero. Since not all $\delta_i$ are zero, we have

$$G_h = \sum_{i=1}^l \delta_i B_i^{(h)} + \sum_{j=1}^{k_h} \lambda_{1,j}^{(h)} B_j^{(h)} \in \text{Span}_F(\{B_1^{(h)}, \ldots, B_{k_h}^{(h)}\}) \quad \text{for} \quad h = 1, \ldots, s.$$ 

Therefore, by (*), for any $P \in \text{Span}_F(S_k)$ we have

$$\text{rank } P = \sum_{i=1}^s \text{rank } G_i \geq \sum_{i=1}^s d_i.$$ 

Thus, by lemma 1, we have

$$L_P(S_k) \geq N_P \left( \sum_{i=1}^s d_i \right). \quad (2)$$

Now, it is obvious that

$$\text{dim } \text{Span}_F(V_2) = \left( \sum_{i=1}^s k_i \right) - s l.$$ 

Combining this with (1) and (2), the result of the lemma follows. □

Let $i_F(n)$ denote the maximum possible number of distinct irreducible factors of a polynomial of degree $n$ over the field $F$. The following lemma is known from [KB].

**Lemma 6.** For sufficiently large $n$ we have

$$i_F(n) \leq \frac{2n}{\log_{11} n}.$$ 

For integers $1 \leq j \leq m$ we define the $m \times m$ Hankel matrix

$$I_m^{(j)} = \begin{bmatrix}
0 & \cdots & 0 & 1 & \cdots & 0 \\
0 & \cdots & 0 & 1 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 1 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\end{bmatrix}.$$ 

That is,

$$I_m^{(j)}[i,k] = \begin{cases} 
1 & \text{if } m-i-k+2 = j, \\
0 & \text{otherwise}.
\end{cases}$$
It is known that,
\[ B(\alpha^n) = \{f^{(i)}_m | j = 1, \ldots, m\}. \] (3)

From [AGW], for any polynomial \( p(\alpha) \) of degree \( n \)
\[ B(p) = \{C_p^0, C_p^1, \ldots, C_p^{n-1}\}, \] (4)
where \( C_p \) is the companion matrix of \( p \). It is well known that when \( p \) is an irreducible polynomial over \( F \) then for any nonzero \( C \in \text{Span}_F(B(p)) \) we have
\[ \text{rank } C = \deg p. \] (5)

Another important property that will be used for the proof of the theorem is the following.

Let \( H_i \) be nonsingular \( n \times n \) matrices for \( i = 1, \ldots, j_0 \). The matrix \( \sum_{j=1}^{j_0} \Theta H_j \) is of the shape
\[
\begin{pmatrix}
0_{n \times n} & \cdots & 0_{n \times n} & H_{j_0} & \cdots & H_3 & H_2 & H_1 \\
& & \ddots & & \ddots & & \ddots & \ddots \\
& & & 0_{n \times n} & & \ddots & \ddots & \ddots \\
& & & & 0_{n \times n} & & \ddots & \ddots \\
\end{pmatrix}
\]
where \( 0_{n \times n} \) is the \( n \times n \) zero matrix. Therefore
\[ \text{rank } \left\{ \sum_{j=1}^{j_0} \Theta H_j \right\} = n j_0. \] (6)

We now prove the theorem.

**Theorem.** Let \( p(\alpha) \) be any polynomial of degree \( n \). Then
\[ L_F(B(p)) \geq \left( 2 + \frac{1}{|F| - 1} \right)n - o(n). \]

**Proof.** Let \( p = p_1^{d_1} \cdots p_k^{d_k} \), where \( p_1, \ldots, p_k \) are distinct irreducible polynomials. Let \( \deg p = n \), \( \deg p_i = r_i, \deg p_i^{d_i} = s_i = r_i d_i \), and \( s_1 \leq \cdots \leq s_k \). It is well known from [AGW1] and [AGW2] that
\[ B(p) = B(\alpha^{d_1}) \Theta B(p_1) \Theta \cdots \Theta B(\alpha^{d_k}) \Theta B(p_k). \]
By (3) and (4),
\[ B(p) = B_1 \Theta \cdots \Theta B_k, \]
where
\[ B_h = \left\{ (E_{d_i}^j) \Theta C_{p_h}^i \mid 1 \leq i \leq d_h, 0 \leq j \leq r_h - 1 \right\} \] for \( h = 1, \ldots, k \),
and \( C_{p_h}^i \) is the companion matrix of \( p_h \). We now prove the following claims.

**Claim 1.** There exists an integer \( w \leq 2 \frac{n^{1/2}}{\log F_1 n} \) such that
\[ s_i = r_i d_i = \deg p_i^j > \frac{1}{2} \frac{n^{1/2}}{\log F_1 n} \text{ for } i = w, \ldots, k. \]

Let \( w \) be an integer such that
\[ \deg p_1^{d_1} \cdots p_w^{d_w} \geq n^{1/2}, \quad \deg p_1^{d_1} \cdots p_{w-1}^{d_{w-1}} < n^{1/2}. \] (7)

By lemma 6, for sufficiently large \( n \),
\[ w \leq 2 \frac{n^{1/2}}{\log F_1 n}. \]

Now, since \( \deg p_1^{d_1} \cdots p_w^{d_w} \geq n^{1/2} \) and \( s_1 \leq s_2 \leq \cdots \leq s_w \), we have that
\[ s_w = \deg p_w^{d_w} \geq \frac{n^{1/2}}{w} \geq \frac{1}{2} \frac{n^{1/2}}{\log F_1 n}. \]

Since \( s_w \leq s_{w+1} \leq \cdots \leq s_k \) we have
\[ s_i = \deg p_i^{d_i} \geq \frac{1}{2} \frac{n^{1/2}}{\log F_1 n} \text{ for } i = w, w + 1, \ldots, k. \]

This completes the proof of claim 1.

Let
\[ m_i = \left\lceil \frac{\log F_1 \log n}{r_i} \right\rceil \text{ for } i = w, \ldots, k, \] (8)
and
\[ W_i = (E_{d_i}^j) \Theta C_{p_i}^i \mid \max (d_i - m_i + 1, 1) \leq j \leq d_i, 0 \leq l \leq r_i - 1 \] for \( i = w, \ldots, k. \)

**Claim 2.** We have
\[ | W_i | \geq \log F_1 \log n \text{ for } i = w, \ldots, k. \]

If \( \max (d_i - m_i + 1, 1) = d_i - m_i + 1 \), then by (8), \( | W_i | = m_i r_i \geq \log F_1 \log n \). If \( \max (d_i - m_i + 1, 1) = 1 \), then \( d_i - m_i + 1 \leq 1 \) and \( | W_i | = r_i \). This implies that
\[ d_i \leq m_i = \left\lceil \frac{\log F_1 \log n}{r_i} \right\rceil < \frac{\log F_1 \log n}{r_i} + 1. \]
We multiply the latter inequality by \( r_i \) and obtain \( r_i > d_i - r_i - \log_{|F|} \log n \). Now, using claim 1, for \( i = m, \ldots, k \),

\[
r_i > \frac{1}{2} \log_{|F|} n - \log_{|F|} \log n > \log_{|F|} \log n.
\]

Now, since \( |W_i| = r_i \) the result of the claim follows.

**Claim 3.** The set \( B_i \) is a \( \{ s_i, |W_i|, s_i - m_i, r_i, r_i \} \) -set for \( i = m, \ldots, k \). \( \square \)

Obviously, \( |B_i| = s_i = r_i d_i \). If \( P \in \text{Span}_F(B_i - W_i) \), then there exist constants \( \{ \delta_V \mid V \in W_i \} \subseteq |F| \) not all zero and \( \{ \eta_V \mid V \in B_i - W_i \} \subseteq |F| \) such that

\[
P = \sum_{V \in W_i} \delta_V V + \sum_{V \in B_i - W_i} \eta_V V.
\]

Then \( P \) can be written as

\[
\sum_{i=0}^{r_i-1} \sum_{j \in B_i} \lambda_{i,j} (I_{d_i})^j \Theta C_{p_i}^j
\]

where not all \( \lambda_{i,j} \) are zero. Therefore

\[
P = \sum_{j=0}^{d_i} \left[ I_{d_i}^j \Theta \sum_{i=0}^{r_i-1} \lambda_{i,j} C_{p_i}^j \right].
\]

Suppose \( j_0 \), where \( \max(d_i - m_i + 1, 1) \leq j_0 \leq d_i \), is the maximal integer such that \( \sum_{i=0}^{r_i-1} \lambda_{i,j} C_{p_i}^j \neq 0 \). By (5) and (6)

\[
\text{rank}(P) = \text{rank}
\left[
I_{d_i}^j \Theta \sum_{i=0}^{r_i-1} \lambda_{i,j} C_{p_i}^j
\right] = j_0 r_i \geq (d_i - m_i + 1) r_i = s_i - m_i, r_i, r_i.
\]

This proves that \( B_i \) is a \( \{ s_i, |W_i|, s_i - m_i, r_i, r_i \} \) -set and claim 3 is proved.

Now, by claim 2, \( \min_{w \in S} |W_i| \geq \log_{|F|} \log n \), and by lemma 5,

\[
L_F(B(p)) \geq \sum_{i=m}^{w \text{ times}} L_F([0] \Theta \cdots \Theta [0] \Theta B_{w+1} \Theta \cdots \Theta B_k) = L_F(B_{w+1} \Theta \cdots \Theta B_k) \geq \log_{|F|} \log n + \sum_{i=m}^{k} (s_i - m_i, r_i, r_i)
\]

We now estimate each term in (9). By (7), we have

\[
\sum_{i=m}^{k} s_i = \deg p_w^d \cdots p_k^d \geq n - n^{1/2}.
\]
By lemma 6,

$$k \leq \frac{2n}{\log_{|F|} n}.$$  

Therefore,

$$k \log_{|F|} \log n = o(n).$$  \hspace{1cm} (11)

By (8), (10) and (11),

$$\sum_{i \in \omega} (s_i - m_i r_i + r_i) \geq n - n^{1/2} - \sum_{i \in \omega} (m_i - 1) r_i \geq n - n^{1/2} - k \log_{|F|} \log n = n - o(n).$$  \hspace{1cm} (12)

Combining this with (9), (10) and (11) we get

$$L_F(B(p)) \geq n + N_F \left( \left\lceil \log_{|F|} \log n \right\rceil, n - o(n) \right) - o(n).$$  \hspace{1cm} (13)

Now, by lemma 2, we have

$$N_F \left( \left\lceil \log_{|F|} \log n \right\rceil, n - o(n) \right) \geq \left[ 1 + \frac{1}{|F| - 1} - \frac{1}{\left\lceil \log_{|F|} \log n \right\rceil} \right] (n - o(n))$$

$$= \left[ 1 + \frac{1}{|F| - 1} \right] n - o(n).$$

Combining this with (13), we obtain the result

$$L_F(B(p)) \geq \left[ 2 + \frac{1}{|F| - 1} \right] n - o(n). \quad \square$$

REFERENCES

[AGW1] A. Averbuch, Z. Galil, S. Winograd, Classification of all minimal bilinear algorithms for computing the coefficients of the product of two polynomials in the algebra $G[u]/<u^n>$, manuscript.

[AGW2] A. Averbuch, Z. Galil, S. Winograd, Classification of all minimal bilinear algorithms for computing the coefficients of the product of two polynomials in the algebra $G[u]/<Q(u)^l>$, $l > 1$, manuscript.


