Multiplication of polynomials over finite fields

Nader H. Bshouty and Michael Kaminski

Department of Computer Science
Technion - Israel Institute of Technology
Haifa 32000
Israel

Abstract. We prove the $2.5n - o(n)$ lower bound on the number of multiplications/divisions required to compute the coefficients of the product of two polynomials of degree n over a finite field by means of straight-line algorithms.

1. Introduction

The number of multiplications/divisions required for computing the product of two degree-n polynomials over an infinite field is known to be $2n+1$. The method is to evaluate both polynomials at each of $2n+1$ distinct points (allowing ∞), multiplying and interpolating the result. This method fails for the fields with the number of elements less than $2n$. The bilinear and quadratic complexity of polynomial multiplication over finite fields has been widely studied in the literature, cf. [BD1], [BD2], [J2], [KA], [KB2], [LSW] and [LW]. The best known lower bounds on the bilinear complexity of multiplying two degree-n polynomials over finite fields are as follows. For the binary field the bound is $3.52n$, cf. [BD2]. The same bound holds for the quadratic complexity, cf. [LSW] and [KA], where the proofs can be applied to quadratic algorithms as well. For the fields with more than 2 elements the bound is $3n - o(n)$, cf [KB2]. Lemma 1 in this paper together with the results form [KB2] shows that the same bounds also hold for the quadratic complexity.

The proofs of the above bounds are based on a special structure of quadratic algorithms, namely, on the fact that the multiplications in quadratic algorithms are independent each of other. It is known from [ST] that if a set of quadratic forms over an infinite field can be computed in t multiplications/divisions, then it can be computed in t multiplications by an quadratic algorithm whose total number of operations differs from that of the original one by a factor of a small constant. But it is unknown whether a similar result holds for finite fields. Also no example of a set of bilinear forms with a nontrivial lower bound on the number of multiplications/divisions required for its computation is known from the literature. In this paper we prove the $2.5n - o(n)$ lower bound on the number of multiplications/divisions required for computing the product of two degree-n polynomials over a finite field by means of straight-line algorithms.

Let $\mathbf{F}_q$ denote the q-element field and let $\mu_q(n)$ denote the number of multiplications/divisions required to compute the coefficients of the product of a polynomial of degree $n - 1$ and a polynomial of degree n over $\mathbf{F}_q$ by means of straight-line algorithms. A straightforward substitution argument shows that the number of multiplications/divisions required for computing the product of two polynomials of degree n exceeds $\mu_q(n)$ at least by 1. The product of polynomials of degrees $n - 1$ and n is considered for a technical reason explained in the next section.

Theorem. *For any q we have* $\mu_q(n) > 5n/2 - n/4\lg_q n - O(n/\lg_q^2 n)$.

The rest of the paper is organized as follows. In the next section we introduce some notation and definitions, and prove the major auxiliary technical lemmas. The proof of the lower bound is presented in Section 3.

2. Notation and auxiliary lemmas

In this section we introduce some notation and prove the major auxiliary lemmas we shall need for the proof of the theorem.

Let k be a positive integer and let $a_0, \ldots, a_{k-1}$ be given elements of a field F. A sequence $\sigma = s_0, s_1, \ldots, s_l$ of elements of F satisfying the relation

$$s_{m+k} = a_{k-1}s_{m+k-1} + a_{k-2}s_{m+k-2} + \cdots + a_0 s_m\ , \quad m = 0, 1, \ldots, l - k$$

is called a (finite k-th-order homogeneous) *linear recurring sequence* in F. The terms $s_0, s_1, \ldots, s_{k-1}$ are referred as *initial values*. The polynomial

$$f(\alpha) = \alpha^k - a_{k-1}\alpha^{k-1} - a_{k-2}\alpha^{k-2} - \cdots - a_0 \in F[\alpha]$$

is called a *characteristic polynomial* of σ. Let (α) be a characteristic polynomial of σ of the minimal degree. If $\mathbf{deg}(\alpha) + \mathbf{deg}\, f(\alpha) \le l + 1$, then (α) divides $f(\alpha)$, cf. [KB2, Proposition 1]. Thus if $\mathbf{deg}(\alpha) \le (l + 1)/2$, then (α) is a unique characteristic polynomial of the minimal degree. It is called the *minimal polynomial* of σ and denoted $f_{\sigma}(\alpha)$.

For a sequence $\sigma = \{s_0, \ldots, s_{2n-1}\}$ we define the $(n+1) \times n$ *Hankel matrix* $H(\sigma)$ by

$$\begin{pmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ s_n & s_{n+1} & \cdots & s_{2n-1} \end{pmatrix}.$$

Let H^i denote the $(i+1)$-st row of H, $i = 0, 1, \ldots, n$. Let k be the minimal positive integer such that there exist $a_0, \ldots, a_{k-1} \in F$ satisfying

$$\sum_{i=0}^{k-1} a_i H^i = H^k .$$

The existence of k is provided by the fact that H has n columns and $n+1$ rows. We define $\tilde{\sigma} = \{\tilde{s}_0, \tilde{s}_1, \ldots, \tilde{s}_{2n-1}\}$ by the recurrence

$$\tilde{s}_{i+k} = a_{k-1}\tilde{s}_{i+k-1} + a_{k-2}\tilde{s}_{i+k-2} + \cdots + a_0\tilde{s}_i ,$$

with initial values $\tilde{s}_i = s_i$, $i = 0, \ldots, k-1$.

Let $\bar{\sigma} = \sigma - \tilde{\sigma}$. We shall denote $H(\tilde{\sigma})$ and $H(\bar{\sigma}) = H - H(\tilde{\sigma})$ by $\tilde{H}$ and $\bar{H}$, respectively. Let $f_H(\alpha) = \alpha^k - \sum_{i=0}^{k-1} a_i \alpha^i$, i.e., $f_H(\alpha)$ is a characteristic polynomial of $\tilde{\sigma}$. (In fact, $f_H(\alpha) = f_{\tilde{\sigma}}(\alpha)$, since, by definition, $f_H(\alpha)$ is a characteristic polynomial of the minimal degree.)

It can be easily verified that Lemmas 1-3 in [KB2] stated for $(n+1) \times (n+1)$ Hankel matrices hold for $(n+1) \times n$ Hankel matrices as well. The reason for dealing with $(n+1) \times n$ Hankel matrices is that for an $(n+1) \times (n+1)$ Hankel matrix H of rank $n+1$ the polynomial $f_H(\alpha)$ is not defined. Thus that case has to be treated separately, whereas dealing with $(n+1) \times n$ Hankel matrices enables a uniform treatment.

Let **S** be a finite set of $(n+1) \times n$ Hankel matrices. Define $f_{\mathbf{S}}(\alpha) = lcm\{f_H(\alpha)\}_{H \in \mathbf{S}}$, where lcm is an abbreviation for ‘‘the least common multiple’’, $d_{\mathbf{S}} = \mathbf{deg}\, f_{\mathbf{S}}(\alpha)$ and $r_{\mathbf{S}} = \mathbf{max}\{\mathbf{rank}\, \bar{H}\}_{H \in \mathbf{S}}$.

Below $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1})^T$ and $\mathbf{y} = (y_0, y_1, \ldots, y_n)^T$ denote column vectors of indeterminates. We remind the reader that a *quadratic* (respectively *bilinear*) algorithm for computing a set of bilinear forms of **x** and **y** is a straight-line algorithm whose non-scalar multiplications are of the shape $L * L'$, where

L and L' are linear forms in $\mathbf{x}$ and $\mathbf{y}$ (respectively L is a linear form in $\mathbf{x}$ and L' is a linear form in $\mathbf{y}$) and each bilinear form is obtained by computing a linear combination of these products.

Lemma 1. *Let* $\mathbf{S}=\{H_0, H_1, \dots, H_s\}$ *be a set of* $(n+1)\times n$ *Hankel matrices. Then computing the set of bilinear forms* $\{\mathbf{x}^T H_i \mathbf{y}\}_{i=1,\dots,s}$ *by means of a quadratic algorithm requires at least* $\mathbf{min}\{d_{\mathrm{S}}+r_{\mathrm{S}}, n+1\}$ *multiplications.*

Proof. Assume that quadratic and bilinear complexities of the set of bilinear forms $\{\mathbf{x}^T H_i \mathbf{y}\}_{i=1,\dots,s}$ are equal to t_Q and t_B, respectively, and let $H(\mathbf{z})=\sum_{j=0}^{s} z_j H_j$ be the characteristic matrix of the above set. Let r be the row rank of $H(\mathbf{z})$. Then $t_Q \ge (t_B+r)/2$, cf. [J3, Theorem 3.5]. Since $t_B \ge r = \mathbf{min}\{d_{\mathrm{S}}+r_{\mathrm{S}}, n+1\}$, cf. [KB2, Lemmas 2 and 3], we have $t_Q \ge \mathbf{min}\{d_{\mathrm{S}}+r_{\mathrm{S}}, n+1\}$. □

Let V be a vector space over F and let $W \subset V$. We shall denote the linear subspace of V spanned by all the vectors from W by $[W]$.

Lemma 2. *Let* $\mathbf{S}$ *and* $\mathbf{S}'$ *be finite sets of* $(n+1)\times n$ *Hankel matrices such that* $[\mathbf{S}]=[\mathbf{S}']$. *If* $d_{\mathrm{S}}+r_{\mathrm{S}} \le n$, *then* $f_{\mathrm{S}}(\alpha)=f_{\mathrm{S}'}(\alpha)$ *and* $r_{\mathrm{S}}=r_{\mathrm{S}'}$.

Proof. Let $\mathbf{S}=\{H(\sigma_i)\}_{i=1,\dots,k}$ and let $H(\sigma)\in\mathbf{S}'$, where $\sigma=\sum_{i=1}^{k}\lambda_i\sigma_i$. It suffices to prove that $f_H(\alpha)$ divides $f_{\mathrm{S}}(\alpha)$ and $\mathbf{rank}\,\bar{H} \le r_{\mathrm{S}}$. By [LN, Theorem 8.55, p. 425], $f_{\mathrm{S}}(\alpha)=\alpha^{d_{\mathrm{S}}}-\sum_{i=1}^{d_{\mathrm{S}}-1} a_i\alpha^i$ is a characteristic polynomial of $\tilde{\sigma}=\sum_{i=1}^{k}\lambda_i\tilde{\sigma}_i$. Since $d_{\mathrm{S}}+r_{\mathrm{S}} \le n$, the sequences σ and $\sum_{i=1}^{k}\lambda_i\tilde{\sigma}_i$ have the same first $2n-r_{\mathrm{S}}$ elements. Therefore $H^{d_{\mathrm{S}}}=\sum_{i=1}^{d_{\mathrm{S}}-1} a_i H^i$ which implies that $f_H(\alpha)$ divides $f_{\mathrm{S}}(\alpha)$. This, in turn, implies the equality $\bar{\sigma}=\sum_{i=1}^{k}\lambda_i\bar{\sigma}_i$. Thus the first $2n-r_{\mathrm{S}}$ elements of $\bar{\sigma}$ are zero and the inequality $\mathbf{rank}\,\bar{H} \le r_{\mathrm{S}}$ follows. □

Lemma 3. *Let* $\mathbf{S}=\{H_1, H_2, \dots, H_m\}$ *be a set of linearly independent* $(n+1)\times n$ *Hankel matrices such that* $d_{\mathrm{S}}+r_{\mathrm{S}} \le n$. *Let* l *be the number of distinct irreducible factors of* $f_{\mathrm{S}}(\alpha)$. *Then computing the set of bilinear forms* $\{\mathbf{x}^T H_i \mathbf{y}\}_{i=1,\dots,m}$ *by means of straight-line algorithms requires at least* $m+d_{\mathrm{S}}+r_{\mathrm{S}}-l-1$ *multiplications/divisions.*

Proof. Let $F[u]$ be the ring of univariate polynomials over the field F and let $F(u)$ be the field of fractions of $F[u]$. I.e., $F(u)$ is the extension of F with a transcendental element u. Then any straight-line algorithm over F is also a straight-line algorithm over $F(u)$ and polynomials irreducible over F remain irreducible over $F(u)$. In particular, the number of irreducible factors of $f_{\mathrm{S}}(\alpha)$ over $F(u)$ is equal to l.

Thus, extending F with a transcendental element, if necessary, we may assume that the field of constants F is infinite. Thus we may restrict ourselves to quadratic algorithms, cf. [ST]. Assume that all the bilinear forms defined by the matrices from $\mathbf{S}$ can be computed in t multiplications. Then there exist $2t$ linear forms $L_1(\mathbf{x},\mathbf{y}), , \ldots , L_t(\mathbf{x},\mathbf{y})$ and $L'_1(\mathbf{x},\mathbf{y}), \ldots , L'_t(\mathbf{x},\mathbf{y})$ in $\mathbf{x}$ and $\mathbf{y}$ such that each $\mathbf{x}^T H_i \mathbf{y}$ is a linear combination of the products $\{L_i(\mathbf{x},\mathbf{y}) L'_i(\mathbf{x},\mathbf{y})\}_{i=1,\ldots,t}$. Let $\mathbf{p} = (L_1(\mathbf{x},\mathbf{y}) L'_1(\mathbf{x},\mathbf{y}), \ldots , L_t(\mathbf{x},\mathbf{y}) L'_t(\mathbf{x},\mathbf{y}))^T$ and let $\mathbf{q} = (\mathbf{x}^T H_1 \mathbf{y}, \ldots , \mathbf{x}^T H_s \mathbf{y})^T$. By the definition of quadratic algorithms there exists an $m \times t$ matrix U whose entries are constants from F such that $\mathbf{q} = U \mathbf{p}$. Since the matrices $\{H_i\}_{i=1,\ldots,m}$ are linearly independent, $\mathbf{rank}\, U = m$.

Permuting the components of $\mathbf{p}$, if necessary, we may assume that the first m columns of U are linearly independent. Hence there exist a non-singular $m \times m$ matrix W and an $m \times (t-m)$ matrix V such that $W\mathbf{q} = (\, I_m , V\,)\mathbf{p}$, where I_m denotes the $m \times m$ identity matrix.

Let $W\mathbf{q} = (\mathbf{x}^T H'_1 \mathbf{y}, \ldots , \mathbf{x}^T H'_m \mathbf{y})$ and let $\mathbf{S}' = \{H'_1, \ldots , H'_m\}$. Since W is a nonsingular matrix, we have $[\mathbf{S}] = [\mathbf{S}']$. Therefore, by Lemma 2, $f_{\mathrm{S}}(\alpha) = f_{\mathrm{S}'}(\alpha)$ and $r_{\mathrm{S}} = r_{\mathrm{S}'}$. Let $f_{\mathrm{S}} = \prod_{i=1}^{l} f_i^{d_i}(\alpha)$ be the decomposition of $f_{\mathrm{S}}(\alpha)$ into its irreducible factors. Then there exists a sequence $H'_{j_0}, H'_{j_1}, \ldots , H'_{j_l}$ of matrices from $\mathbf{S}'$ such that $\mathbf{rank}\, H'_{j_0} = r_{\mathrm{S}}$ and $f_i^{d_i}(\alpha)$ divides $f_{H'_{j_i}}(\alpha)$, $i = 1, 2, \ldots , l$. Notice that the elements in the above sequence are not necessarily distinct. Permuting the components of $\mathbf{p}$, if necessary, we may assume that $j_i \geq m - l$, $i = 1, 2, \ldots , l$. Then $lcm\{f_{H_{m-i}}(\alpha)\}_{i=0,1,,\ldots,l} = f_{\mathrm{S}}(\alpha)$ and $\mathbf{max}\{\overline{H}_{m-i}\}_{i=0,1,,\ldots,l} = r_{\mathrm{S}}$. By Lemma 1, computing the bilinear forms defined by the last $l+1$ components of $W\mathbf{q}$ requires at least $d_{\mathrm{S}} + r_{\mathrm{S}}$ multiplications. Since the first $m-l-1$ components in the each of the last $l-1$ rows of $(\, I_m , V\,)$ are zero and the products $\{L_i(\mathbf{x},\mathbf{y}) L'_i(\mathbf{x},\mathbf{y})\}_{i=1,\ldots,t}$ are computed independently each of other, we have $t - (m - l - 1) \geq d_{\mathrm{S}} + r_{\mathrm{S}}$. Hence $t \geq m + d_{\mathrm{S}} + r_{\mathrm{S}} - l - 1$ which completes the proof. □

3. Proof of the lower bound

We shall need the following definition.

Let $M = (m_{i,j})$ be a $u \times v$ matrix. We shall say that M is in *echelon form* if there exists a $k \leq u$ and a sequence $j_1 < j_2 < \cdots < j_k$ such that the following conditions are satisfied.

(i) All the entries in the last $u-k$ rows of M are zero.

(ii) For each $i=1, \ldots, k$ the entry m_{i,j_i} is not equal to zero.

(iii) For each $i=1, \ldots, k$ and $j<j_i$ the entry $m_{i,j}$ is zero.

It is wellknown that each matrix can be transformed into echelon form by a sequence of elementary operations on its rows.

Proof of the theorem. We have to compute $z_k = z_k(\mathbf{x}, \mathbf{y}) = \sum\limits_{i+j=k} x_i y_j$, $k=0, \ldots, 2n-1$. Let $\mathbf{z} = (z_0, z_1, \ldots, z_{2n-1})^T$. Assume that $\mu_q(n) = t$, i.e. all the bilinear forms defined by the components of $\mathbf{z}$ can be computed in t multiplications/divisions. It is known from [FZ] that $t \geq 2n$. Let $m_1, m_2, \ldots, m_t$ be all the multiplications/divisions of an algorithm that computes $\{z_k\}_{k=0,\ldots,2n-1}$. Let $\mathbf{p} = (m_t, m_{t-1}, \ldots, m_1)^T$. Then there exist a $2n \times t$ matrix U whose entries are constants from $\mathbf{F}_q$ and a t-dimensional column vector $\mathbf{q}$ whose components are affine forms in $\mathbf{x}$ and $\mathbf{y}$ such that $\mathbf{z} = U\mathbf{p} + \mathbf{q}$. There exist a non-singular $2n \times 2n$ matrix W such that the matrix WU in echelon form. Multiplying $\mathbf{z}$ by W we obtain $W\mathbf{z} = WU\mathbf{p} + W\mathbf{q}$.

Let $W\mathbf{z} = (\mathbf{x}^T H_{2n}\mathbf{y}, \mathbf{x}^T H_{2n-2}\mathbf{y}, \ldots, \mathbf{x}^T H_1 \mathbf{y})^T$ and let $\mathbf{S}_m = \{H_1, H_2, \ldots, H_m\}$, $m=1,2,\ldots,2n$. Since the matrices $H_1, H_2, \ldots, H_{2n}$ are linearly independent, $d_{\mathbf{S}_{2n}} + r_{\mathbf{S}_{2n}} \geq 2n$, cf. [KB2, Lemma 1]. Let $\mathbf{i}_q(n)$ denote the maximal possible number of distinct factors of a polynomial of degree n over $\mathbf{F}_q$. Obviously, $\mathbf{i}_q(n) < n$. Therefore there is an integer m such that $d_{\mathbf{S}_{m-1}} + r_{\mathbf{S}_{m-1}} \leq (n + \mathbf{i}_q(n/2))/2$ and $d_{\mathbf{S}_m} + r_{\mathbf{S}_m} > (n + \mathbf{i}_q(n/2))/2$. We shall consider the cases of $d_{\mathbf{S}_m} + r_{\mathbf{S}_m} > n$ and $d_{\mathbf{S}_m} + r_{\mathbf{S}_m} \leq n$ separately.

If $d_{\mathbf{S}_m} + r_{\mathbf{S}_m} > n$, then, by Lemma 1, computing the set of bilinear form defined by the last m components of $W\mathbf{z}$ requires at least $n+1$ multiplications/divisions. By the definition of echelon form, at least $2n-m$ first components in the last m rows of WU are zero. Thus we have $t-(2n-m) \geq n+1$. Since the matrices $H_1, H_2, \ldots, H_{m-1}$ are linearly independent, $m-1 \leq (n + \mathbf{i}_q(n/2))/2$, cf. [KB2, Lemma 1]. Therefore $t \geq (5n - \mathbf{i}_q(n/2))/2$.

If $d_{\mathbf{S}_m} + r_{\mathbf{S}_m} \leq n$, then, by Lemma 3, computing the set of bilinear form defined by the last m components of $W\mathbf{z}$ requires at least $m + d_{\mathbf{S}_m} + r_{\mathbf{S}_m} - \mathbf{i}_q(\mathbf{deg}\, f_{\mathbf{S}_m}(\alpha)) - 1$ multiplications/divisions. Since at least $2n-m$ first components in the last m rows of WU are zero, we have

$$t-(2n-m)\geq m+d_{\mathbf{s}_m}+r_{\mathbf{s}_m}-\mathbf{i}_q(\mathbf{deg}\, f_{\mathbf{s}_m}(\alpha))-1\,.$$

Since $d_{\mathbf{s}_m}+r_{\mathbf{s}_m}>(n+\mathbf{i}_q(n/2))2$ and the function $n-\mathbf{i}_q(n)$ is non-decreasing, it follows that $t\geq\dfrac{5n-\mathbf{i}_q(n/2+\mathbf{i}_q(n/2))}{2}$. Obviously, $\mathbf{i}_q(n_1+n_2)\leq\mathbf{i}_q(n_1)+\mathbf{i}_q(n_2)$. Thus in either of the cases treated above we have $t\geq\dfrac{5n-\mathbf{i}_q(n/2)-\mathbf{i}_q(\mathbf{i}_q(n/2))}{2}$.

It can be easily shown that $\mathbf{i}_q(n)<n/\lg_q n+O(n/\lg_2^2 n)$. In particular, the proof for $q\geq3$ can be found in [KB2, Appendix 1]. Thus $\mu_q(n)=t\geq5n/2-n/4\lg_q n-O(n/\lg_2^2 n)$. □

References

[BD1] R.W. Brockett, D. Dobkin, On the Optimal Evaluation of a set of Bilinear Forms, *Linear Algebra and Its Applications* **19** (1978), 207-235.

[BD2] M.R. Brown, D.P. Dobkin, An Improved Lower Bound on Polynomial Multiplication, *IEEE Transactions on Computers* **29** (1980), 337-340.

[FZ] C.M. Feduccia, Y. Zalcstein, Algebras having linear multiplicative complexity, *J. ACM* **24** (1977), 311-331.

[HM] J. Hopcroft, J. Munsinski, Duality applied to the complexity of matrix multiplication, *SIAM J. Comput.* **2** (1973), 159-173.

[J1] J. Ja’ Ja’, Optimal evaluation of pairs of bilinear forms, *SIAM J. Comput.* **8** (1979), 443-462.

[J2] J. Ja’ Ja’, Computation of Bilinear Forms over Finite Fields, *J. ACM* **27** (1980), 822-830.

[J3] J. Ja’ Ja’, On the complexity of bilinear forms with commutativity. *SIAM J. Comput.* **9** (1979), 713-728.

[KA] M. Kaminski, A lower bound for polynomial multiplication, *Theoret. Comput. Sci.* **40** (1985), 319-322.

[KB1] M. Kaminski, N.H. Bshouty, Multiplicative complexity of polynomial multiplication over finite fields, in *Proceedings of 28th Annual IEEE Symposium on Foundations of Computer Science* pp. 138-140, The Institute of Electrical and Electronic Engineers, New York, 1987.

[KB2] M. Kaminski, N.H. Bshouty, Multiplicative complexity of polynomial multiplication over finite fields, submitted.

[LSW] A. Lempel, G. Seroussi, S. Winograd, On the Complexity of Multiplication in Finite Fields, *Theoret. Comput. Sci.* **22** (1983), 285-296.

[LW] A. Lempel, S. Winograd, A New Approach to Error-Correcting Codes, *IEEE Transactions on Information Theory* **23** (1977), 503-508.

[LN] R. Lidl, H. Niederreiter, *Finite Fields*, Encyclopedia of Mathematics and its Applications, Vol. 20, G.-C. Rota, ed., Addison-Wesley, Reading, Massachusetts, 1983.

[ST] V. Strassen, Vermeidung von Divisionen, *J. Reine Angew. Math.* **264** (1973), 184-202.