Bivariate Polynomial Multiplication*

Markus Bläser
Institut für Informatik II, Universität Bonn
Römerstr. 164, 53117 Bonn, Germany
blaeser@cs.uni-bonn.de

Abstract

Inspired by the discussion in [5], we study the multiplicative complexity and the rank of the multiplication in the local algebras $\mathcal{R}_{m,n} = k[x,y]/(x^{m+1}, y^{n+1})$ *and* $\mathcal{T}_n = k[x,y]/(x^{n+1}, x^n y, \dots, y^{n+1})$ *of bivariate polynomials. We obtain the lower bounds* $(2\frac{1}{3} - o(1)) \cdot \dim \mathcal{R}_{m,n}$ *and* $(2\frac{1}{2} - o(1)) \cdot \dim \mathcal{T}_n$ *for the multiplicative complexity of the multiplication in* $\mathcal{R}_{m,n}$ *and* $\mathcal{T}_n$*, respectively. On the other hand, we derive the upper bounds* $3 \cdot \dim \mathcal{T}_n - 2n - 2$ *and* $3 \cdot \dim \mathcal{R}_{m,n} - m - n - 3$ *for the rank of the multiplication in* $\mathcal{T}_n$ *and* $\mathcal{R}_{m,n}$*, respectively, provided that the ground field* k *admits "fast" univariate polynomial multiplication* $\bmod\ x^N - 1$*. Our results are also applicable to arbitrary finite dimensional algebras of truncated bivariate polynomials* $k[x,y]/I$*, where the ideal* $I = (x^{d_0+1}, x^{d_1+1}y, \dots, x^{d_n+1}y^n, y^{n+1})$ *is described by a degree pattern* $d_0 \geq d_1 \geq \cdots \geq d_n \geq 0$*.*

1. Introduction

Motivated by the multiplication of numerical univariate polynomials with arbitrary precision, Schönhage [5] started investigating the computational complexity of the multiplication in the algebras $\mathcal{R}_{m,n} = k[x,y]/(x^{m+1}, y^{n+1})$ and $\mathcal{T}_n = k[x,y]/(x^{n+1}, x^n y, x^{n-1}y^2, \dots, y^{n+1})$. Due to the shape of the truncation pattern, we will also speak of the rectangular resp. triangular case in the following. These algebras are the result of modelling the multiplication of univariate polynomials with arbitrary precision and truncation both in degree and precision in algebraic terms. The first indeterminate x refers to the indeterminate of the univariate polynomials, and y models the precision management. For more details as well as applications the reader might consult [5].

However, we view these problems purely as a task in algebraic complexity theory. We discuss the multiplicative complexity of computing the product pq in these algebras, i.e., counting essential multiplications of linear forms in the coefficients of p and q only, while additions, subtractions and multiplications with scalars from the ground field k are taken for free. If we only allow essential multiplications of linear forms in the coefficients of p with linear forms in the coefficients of q, we get the bilinear complexity or rank of the above computational problems. The reader is referred to [3, 2] for the standard facts of this model and further background. For convenience, we give the basic definitions at the beginning of the next section.

Let us first turn to the algebra $k[x]/(x^{n+1})$ of univariate polynomials. By the Alder–Strassen theorem [1], the multiplicative complexity of computing the product $pq \mod x^{n+1}$ of two polynomials p, q of degree n has the lower bound $2n + 1$. If $\#k \geq 2n$, this bound is sharp. Even the untruncated product pq can be computed with $2n + 1$ bilinear multiplications. This means there is nothing better than computing the untruncated product pq and then throwing away the upper n coefficients. Computing the product pq is done by evaluating p and q at $2n + 1$ different points $\xi_0, \dots, \xi_{2n} \in k$, which is free of costs, then performing $2n+1$ essential bilinear multiplications $p(\xi_\nu)q(\xi_\nu)$, and recovering the coefficients of pq from these products using interpolation, which is again free of costs.

Things are different in the bivariate case. The Alder–Strassen theorem yields the lower bounds $2 \cdot \dim \mathcal{R}_{m,n} - 1$ and $2 \cdot \dim \mathcal{T}_n - 1$ for the multiplicative complexity of the multiplication in $\mathcal{R}_{m,n}$ and $\mathcal{T}_n$, respectively. On the other hand, the foregoing evaluation-multiplication-interpolation process for computing the untruncated product requires interpolation at $4mn + 2m + 2n + 1$ points for polynomials in $\mathcal{R}_{m,n}$ and at $2n^2 + 3n + 1$ points for polynomials in $\mathcal{T}_n$ and hence as many multiplications. The lower bound and the upper bound differ by a factor of about 2. So what is the true factor of n^2?

By a skillful reduction to univariate polynomial multiplication $\bmod\ x^N - 1$, Schönhage [5] reduces the upper bounds to $3mn + 2m + 2n + 1$ for $\mathcal{R}_{m,n}$ and $1.87(n + 1)^2$ for $\mathcal{T}_n$ over suitable ground fields. Here, as the first of our

*This work was supported through funds provided by the Deutsche Forschungsgemeinschaft under Grant No. Scho 415/3-1.

main results, we obtain lower bounds significantly above the Alder–Strassen bound, namely, $(2\frac{1}{3} - o(1)) \cdot \dim \mathcal{R}_{m,n}$ and $(2\frac{1}{2} - o(1)) \cdot \dim \mathcal{T}_n$ for the multiplicative complexity of the multiplication in $\mathcal{R}_{m,n}$ and $\mathcal{T}_n$, respectively. Secondly, we derive $3 \cdot \dim \mathcal{T}_n - 2n - 2 = \frac{3}{2}n(n+1) + n + 1$ as an upper bound for the rank of the multiplication in $\mathcal{T}_n$. For $\mathcal{R}_{m,n}$, we are able to decrease the upper bound by one, namely $3 \cdot \dim \mathcal{R}_{m,n} - m - n - 3 = 3mn + 2m + 2n$.

2. Lower Bounds for Local Algebras

We begin this section with some algebraic preliminaries: If ϕ is the multiplication map of an algebra A, then ϕ is concise, that is, its left and right kernel are both equal to $\{0\}$ and the span of its image is A. An algebra is called local, if it is commutative and has a unique maximal ideal. The maximal ideal of a local algebra is exactly the set of non-units. $\mathcal{R}_{m,n}$ and $\mathcal{T}_n$, for instance, are both local.

To specify the model of computation, let U, V, and W be finite-dimensional vector spaces over a field k, and $\phi : U \times V \to W$ a bilinear mapping. A sequence $\beta = (f_1, g_1, w_1, \ldots, f_l, g_l, w_l)$ with $f_\lambda, g_\lambda \in (U \times V)^*$ and $w_\lambda \in W$ is called a quadratic computation of length l for ϕ, if

$$\phi(u,v) = \sum_{\lambda=1}^{l} f_\lambda(u,v) g_\lambda(u,v) w_\lambda \quad \text{for all } u \in U, v \in V.$$

Here $(U \times V)^*$ denotes the dual space of $U \times V$. The length of a shortest quadratic computation for ϕ is called the multiplicative complexity of ϕ and is denoted by $C(\phi)$. If we require more restrictively that $f_\lambda \in U^*$ and $g_\lambda \in V^*$, we get bilinear computations for ϕ. The length of a shortest bilinear computation for ϕ is called the rank $R(\phi)$ of ϕ. The multiplicative complexity $C(A)$ resp. the rank $R(A)$ of a finite-dimensional associative algebra is defined by the multiplicative complexity resp. rank of its multiplication map.

Let us first consider the rank, which is easier to discuss. Let A be a local algebra with maximal ideal M, $\dim A = N$, and $R(A) = r$. Let $\beta = (f_1, g_1, w_1, \ldots, f_r, g_r, w_r)$ be an optimal bilinear computation for A. Since the multiplication map of A is concise, $f_1, \ldots, f_r$ generate A^*. After a suitable renumbering, we may assume that $f_1, \ldots, f_N$ form a basis of A^*. If $U \subseteq M$ is a vector space, $\dim U = m$, then, after another renumbering,

$$\bigcap_{\mu=1}^{m} \ker f_\mu \oplus U = A.$$

Let $X = \bigcap_{\mu=1}^{m} \ker f_\mu$. Since $U \subseteq M$, there exists a unit $x_0 \in X$. Now $g_{m+1}, \ldots, g_r$ generate A^*, because otherwise there were some $0 \neq y \in \bigcap_{\nu=m+1}^{r} \ker g_\nu$, hence

$$x_0 y = \sum_{\rho=1}^{r} f_\rho(x_0) g_\rho(y) w_\rho = 0,$$

a contradiction. Once again, after a suitable renumbering, $g_{m+1}, \ldots, g_{N+m}$ form a basis of A^*. If $V \subseteq M$ is a vector space, $\dim V = n$, then, after another renumbering, we may also assume that

$$\bigcap_{\nu=m+1}^{m+n} \ker g_\nu \oplus V = A.$$

Let $Y = \bigcap_{\nu=m+1}^{m+n} \ker g_\nu$. For $x \in X$ and $y \in Y$,

$$xy = \sum_{\rho=1}^{r} f_\rho(x) g_\rho(y) w_\rho \in \langle w_{m+n+1}, w_{m+n+2}, \ldots, w_r \rangle,$$

where the brackets $\langle\ \rangle$ denote the linear span. Therefore we have for the linear span of $XY = \{xy \mid x \in X, y \in Y\}$,

$$\dim \langle XY \rangle \leq r - n - m$$

or, equivalently,

$$r \geq \dim U + \dim V + \dim \langle XY \rangle.$$

Since β is optimal, the above considerations prove

Lemma 1 *Let A be a local algebra with maximal ideal M, $\dim A = N$. If $U, V \subseteq M$ are vector spaces, then*

$$R(A) \geq \dim U + \dim V + \min_{\substack{X \oplus U = A,\\ Y \oplus V = A}} \{\dim \langle XY \rangle\}.$$

Since we do not have any further information about $f_1, \ldots, f_m$ and $g_{m+1}, \ldots, g_{m+n}$, there does not seem to be anything better than taking the minimum in Lemma 1.

Lemma 1 carries over to the multiplicative complexity and to arbitrary algebras, if we assume that $U, V \subseteq \operatorname{rad} A$, where $\operatorname{rad} A$ denotes the radical of A. If $\beta = (f_1, g_1, w_1, \ldots, f_l, g_l, w_l)$ is an optimal quadratic computation for A, then, by a result within the proof of the Alder–Strassen theorem [1] (see also [3, Lemma IV.13] or [2, Proposition 17.20]), after a suitable renumbering and exchanging some g_λ with f_λ we may assume that $f_1, \ldots, f_{2N-2d}$ separate the points of $\operatorname{rad} A \times \operatorname{rad} A$, i.e.,

$$\bigcap_{\lambda=1}^{2N-2d} \ker f_\lambda \cap (\operatorname{rad} A \times \operatorname{rad} A) = \{0\},$$

where $d = \dim(A/\operatorname{rad} A)$. Therefore, after a suitable permutation,

$$\bigcap_{\lambda=1}^{2N-n-m} \ker f_\lambda \cap (U \times V) = \{0\}.$$

Let $X = \bigcap_{\lambda=1}^{2N-n-m} \ker f_\lambda$. Again, for all $(x,x') \in X$,

$$\begin{aligned} xx' &= \sum_{\lambda=1}^{l} f_\lambda(x,x') g_\lambda(x,x') w_\lambda \\ &\in \langle w_{n+m+1}, w_{n+m+2}, \dots, w_l \rangle . \end{aligned}$$

Since β is optimal, we obtain

Lemma 2 *Let A be a an associative algebra,* $\dim A = N$. *If $U, V \subseteq \operatorname{rad} A$ are vector spaces, then*

$$\begin{aligned} C(A) \geq{} & \dim U + \dim V \\ & + \min_{\substack{X \oplus (U \times V) \\ = A \times A}} \{\dim \langle xx' \mid (x,x') \in X \rangle\}. \end{aligned}$$

Basically, these new bounds are achieved by the substitution method together with a linear independence argument. Still, they do not exceed $3 \cdot \dim A$. For better bounds, new methods have to be encountered.

2.1. The Triangular Case

Starting with $\mathcal{T}_n$, we apply the foregoing result to our examples. Again, we first do so for the rank.

Let M denote the maximal ideal of $\mathcal{T}_n$. We choose $U = M^{h+1}$ and $V = M^{H+1}$, where $h = \lfloor \frac{n}{2} \rfloor$ and $H = \lceil \frac{n}{2} \rceil$, that is, U consists of all polynomials whose monomials all have total degree greater than h. The same holds for V and H, respectively. If X is a subspace of $\mathcal{T}_n$ such that $X \oplus U = \mathcal{T}_n$, then there are polynomials $p_{i,j} \in U$ such that

$$X = \left\langle x^i y^j + p_{i,j} \mid 0 \leq i + j \leq h \right\rangle .$$

In the same way, if $Y \oplus V = \mathcal{T}_n$, there are polynomials $q_{i,j} \in V$ such that

$$Y = \left\langle x^i y^j + q_{i,j} \mid 0 \leq i + j \leq H \right\rangle .$$

We have

$$\begin{aligned} \dim U &= \tfrac{1}{2}((n+1)(n+2) - (h+1)(h+2)), \\ \dim V &= \tfrac{1}{2}((n+1)(n+2) - (H+1)(H+2)). \end{aligned}$$

So it remains to determine $\dim \langle XY \rangle$: Let $0 \leq \nu \leq n$. For all $i, j \geq 0$ with $i + j = \nu$ there are $i_1, i_2, j_1, j_2 \geq 0$ such that $i_1 + i_2 = i$, $j_1 + j_2 = j$ and $i_1 + j_1 \leq h$, $i_2 + j_2 \leq H$. For instance, if ν is even, $i_1 = \lfloor \frac{i}{2} \rfloor$, $i_2 = \lceil \frac{i}{2} \rceil$, $j_1 = \lceil \frac{j}{2} \rceil$, $j_2 = \lfloor \frac{j}{2} \rfloor$ is a possible choice. For odd ν, $i_1 = \lfloor \frac{i}{2} \rfloor$, $i_2 = \lceil \frac{i}{2} \rceil$, $j_1 = \lfloor \frac{j}{2} \rfloor$, $j_2 = \lceil \frac{j}{2} \rceil$ will do, since then either i or j is even and $i_2 + j_2 = \lceil \frac{\nu}{2} \rceil \leq H$. Hence, with $a_{i,j} := (x^{i_1} y^{j_1} + p_{i_1,j_1})(x^{i_2} y^{j_2} + q_{i_2,j_2})$,

$$\begin{aligned} a_{i,j} = x^i y^j + & \underbrace{x^{i_1} y^{j_1} q_{i_2,j_2}}_{\in M^{i_1+j_1+H+1}} \\ & + \underbrace{p_{i_1,j_1} x^{i_2} y^{j_2}}_{\in M^{i_2+j_2+h+1}} + \underbrace{p_{i_1,j_1} q_{i_2,j_2}}_{\in M^{h+H+2} = \{0\}} \in \langle XY \rangle . \end{aligned}$$

Since $i_1 + j_1 \leq h$, we have $i + j + 1 \leq i_2 + j_2 + h + 1$ and $M^{\nu+1} = M^{i+j+1} \supseteq M^{i_2+j_2+h+1}$. The same argument yields $M^{\nu+1} \supseteq M^{i_1+j_1+H+1}$. Hence, $a_{i,j} \in M^\nu$ and $\pi_\nu(a_{i,j}) = x^i y^j$, where $\pi_\nu : \mathcal{T}_n \to \mathcal{T}_n / M^{\nu+1}$ is the canonical projection. Therefore $\mathcal{T}_n = \langle a_{i,j} \mid 0 \leq i + j \leq n \rangle \subseteq \langle XY \rangle$.

Now, application of Lemma 1 yields

$$\begin{aligned} R(\mathcal{T}_n) &\geq \dim U + \dim V + \dim \langle XY \rangle \\ &= \tfrac{3}{2}(n+1)(n+2) - \tfrac{1}{2}(\lfloor \tfrac{n}{2} \rfloor + 1)(\lfloor \tfrac{n}{2} \rfloor + 2) \\ &\quad - \tfrac{1}{2}(\lceil \tfrac{n}{2} \rceil + 1)(\lceil \tfrac{n}{2} \rceil + 2). \end{aligned}$$

If $4 | n + 2$, this simplifies to

$$\begin{aligned} R(\mathcal{T}_n) &\geq \tfrac{3}{2}(n+1)(n+2) - \tfrac{1}{4}(n+2)(n+4) \\ &= 2\tfrac{1}{2} \dim \mathcal{T}_n - \tfrac{3}{4}(n+2). \end{aligned}$$

In the case of multiplicative complexity, we deal with subspaces $X \subset \mathcal{T}_n \times \mathcal{T}_n$ fulfilling $X \oplus (U \times V) = \mathcal{T}_n \times \mathcal{T}_n$. Compared with the above problem of determining $\langle XY \rangle$, the problem of determining $\hat{X} := \langle xx' \mid (x,x') \in X \rangle$ only involves some more indices. Again, $\hat{X} = \mathcal{T}_n$. We skip further details here, since they are treated in the proof of the rectangular case. Together with Lemma 2 this proves

Theorem 3 *Let $\mathcal{T}_n = k[x,y]/(x^i y^j \mid i + j = n + 1)$. The multiplicative complexity of the multiplication in $\mathcal{T}_n$ is bounded from below by*

$$C(\mathcal{T}_n) \geq (2\tfrac{1}{2} - o(1)) \cdot \dim \mathcal{T}_n .$$

More precisely,

$$\begin{aligned} C(\mathcal{T}_n) \geq{} & \tfrac{3}{2}(n+1)(n+2) - \tfrac{1}{2}(\lfloor \tfrac{n}{2} \rfloor + 1)(\lfloor \tfrac{n}{2} \rfloor + 2) \\ & - \tfrac{1}{2}(\lceil \tfrac{n}{2} \rceil + 1)(\lceil \tfrac{n}{2} \rceil + 2). \end{aligned}$$

2.2. The Rectangular Case

The same method yields the following lower bound for the rectangular case. To make things easier, we first consider the quadratic algebra $\mathcal{R}_{n,n}$.

Theorem 4 *Let $\mathcal{R}_{n,n} = k[x,y]/(x^{n+1}, y^{n+1})$. The multiplicative complexity of the multiplication in $\mathcal{R}_{n,n}$ has the lower bound*

$$C(\mathcal{R}_{n,n}) \geq (2\tfrac{1}{3} - o(1)) \cdot \dim \mathcal{R}_{n,n} .$$

More precisely,

$$\begin{aligned} C(\mathcal{R}_{n,n}) \geq{} & 3(n+1)^2 - \tfrac{1}{2}(\lfloor \tfrac{s}{2} \rfloor + 1)(\lfloor \tfrac{s}{2} \rfloor + 2) \\ & - \tfrac{1}{2}(\lceil \tfrac{s}{2} \rceil + 1)(\lceil \tfrac{s}{2} \rceil + 2) \\ & - \tfrac{1}{2}(2n - s)(2n - s + 1), \end{aligned}$$

where $s = \lfloor \frac{4n-2}{3} \rfloor$.

Proof. Let M denote the maximal ideal of $\mathcal{R}_{n,n}$. Let $n \le s \le 2n$ be arbitrary but fixed. We choose $U = M^{h+1}$ and $V = M^{H+1}$ where $h = \lfloor \frac{s}{2} \rfloor$ and $H = \lceil \frac{s}{2} \rceil$. If $X \subseteq \mathcal{R}_{n,n} \times \mathcal{R}_{n,n}$ is a vector space such that $X \oplus (U \times V) = \mathcal{R}_{n,n} \times \mathcal{R}_{n,n}$, then there are polynomials $p_{i_1,j_1,i_2,j_2} \in U$ and $q_{i_1,j_1,i_2,j_2} \in V$ such that X is the linear span of the $(x^{i_1}y^{j_1} + p_{i_1,j_1,i_2,j_2}, x^{i_2}y^{j_2} + q_{i_1,j_1,i_2,j_2})$ for $0 \le i_1 + j_1 \le h$, $0 \le i_2 + j_2 \le H$. For convenience, we will skip the indices of p_{i_1,j_1,i_2,j_2} and q_{i_1,j_1,i_2,j_2} in the following. Since

$$\dim M^\nu = \begin{cases} (n+1)^2 - \frac{1}{2}\nu(\nu+1), & 0 \le \nu \le n, \\ \frac{1}{2}(2n-\nu+1)(2n-\nu+2), & n < \nu \le 2n, \end{cases}$$

it remains to determine $\hat{X} := \dim \langle xx' \mid (x,x') \in X \rangle$. Let $0 \le \nu \le s$. As before, for all $i, j \ge 0$ with $i + j = \nu$ there are $i_1, i_2, j_1, j_2 \ge 0$ such that $i_1 + i_2 = i$, $j_1 + j_2 = j$ and $i_1 + j_1 \le h$, $i_2 + j_2 \le H$. We obtain, with $a_{i,j} := (x^{i_1}y^{j_1} + p)(x^{i_2}y^{j_2} + q)$,

$$\begin{aligned} a_{i,j} = x^i y^j + &\underbrace{x^{i_1}y^{j_1}q}_{\in M^{i_1+j_1+H+1}} \\ + &\underbrace{px^{i_2}y^{j_2}}_{\in M^{i_2+j_2+H+1}} + \underbrace{pq}_{\in M^{h+H+2}} \in \hat{X}. \end{aligned}$$

Like in the triangular case, $M^{\nu+1} \supseteq M^{i_2+j_2+h+1}$ and $M^{\nu+1} \supseteq M^{i_1+j_1+H+1}$. In addition, $M^{s+1} \supseteq M^{h+H+2}$. Thus $a_{i,j} \in M^\nu$ and $\pi_\nu(a_{i,j}) = x^i y^j$, where again $\pi_\nu : \mathcal{R}_{n,n} \to \mathcal{R}_{n,n}/M^{\nu+1}$ denotes the canonical projection. Therefore we obtain $\dim \hat{X} \ge (n+1)^2 - \dim M^{s+1}$. By Lemma 2,

$$\begin{aligned} C(\mathcal{R}_{n,n}) &\ge \dim U + \dim V + \dim \hat{X} \\ &\ge 3(n+1)^2 - \tfrac{1}{2}(\lfloor \tfrac{s}{2} \rfloor + 1)(\lfloor \tfrac{s}{2} \rfloor + 2) \\ &\quad - \tfrac{1}{2}(\lceil \tfrac{s}{2} \rceil + 1)(\lceil \tfrac{s}{2} \rceil + 2) \\ &\quad - \tfrac{1}{2}(2n-s)(2n-s+1). \end{aligned}$$

Choosing $s = \lfloor \frac{4n-2}{3} \rfloor$ maximizes the right hand side. □

Similar bounds can also be obtained for $\mathcal{R}_{m,n}$. In this case, if we assume that $m \ge n$, then the ideals $I_\mu := \langle x^i y^j \mid ni + mj \ge n\mu \rangle$, $0 \le \mu \le 2m$, play the roles of the powers of the maximal ideal M in the proof of Theorem 4. The resulting exact bounds are full of sums of rounded quotients of integers, but also yield $C(\mathcal{R}_{m,n}) \ge (2\frac{1}{3} - o(1)) \cdot \dim \mathcal{R}_{m,n}$.

The technique of this section can also be applied to arbitrary finite dimensional algebras of truncated bivariate polynomials $A = k[x,y]/I$, where the ideal $I = (x^{d_0+1}, x^{d_1+1}y, \dots, x^{d_n+1}y^n, y^{n+1})$ is given by a degree pattern $d_0 \ge d_1 \ge \dots \ge d_n \ge 0$. However, depending on the pattern, taking powers of the maximal ideal for U and V is not always the best choice.

3. Upper Bounds by Wraparounds

By a skillful reduction to univariate polynomial multiplication $\bmod\ x^N - 1$, Schönhage [5] obtains the upper bound $R(\mathcal{R}_{m,n}) \le 3mn + 2m + 2n + 1$, provided that the field k supports fast polynomial multiplication $\bmod\ x^N - 1$, more precisely, $R(k[x]/(x^N - 1)) = N$. Since that construction for $\mathcal{R}_{m,n}$ is crucial for our construction in the triangular and the general case, we will first recall it briefly. Following [5] in order to avoid lengthy case discussions, we call a field k “suitable”, if k has characteristic zero and contains all roots of unity. We assume in the following that k is “suitable”. Then $R(k[x]/(x^N - 1)) = N$ for all N.

3.1. The Rectangular Case

Assume we have two polynomials $p, q \in \mathcal{R}_{m,n}$. We want to reduce the multiplication of p, q to the multiplication of some $p', q' \in k[x]/(x^N - 1)$ for suitable N. Both factors are encoded as long univariate polynomials by substituting $y \mapsto x^{3m+2}$, i.e., $p'(x) = p(x, x^{3m+2})$, $q'(x) = q(x, x^{3m+2})$. This substitution groups together the coefficients of all monomials with the same degree in y and leaves $2m + 1$ zero coefficients between these groups. Setting $N = n(3m+2) + 2m + 1$, we can read off the coefficients of the product $pq \in \mathcal{R}_{m,n}$ from the coefficients of the univariate product

$$r'(x) = p'(x)q'(x) \mod x^N - 1.$$

Borrowing the language of [5], a pattern

d d d o o o o o ’ d d d o o o o o ’ d d d o o o o o ’ d d d o o
for the factors p', q'

d d d $*$ $*$ o o o ’ d d d $*$ $*$ o o o ’ d d d $*$ $*$ o o o ’ d d d $*$ $*$ /
o o o $*$ $*$ $*$ $*$ $*$ ’ o o o $*$ $*$ $*$ $*$ $*$ ’ o o o $*$ $*$ $*$ $*$ $*$
for the result r'

(lower coefficients on the left, higher on the right) illustrates this method for $m = 2$, $n = 3$, $N = 29$. Here ‘d’ stands for coefficients of the original polynomials (“data”), ‘o’ for zero coefficients, and the quotes separate the groups of monomials with the same degree in y. The markers ‘$*$’ in the result indicate “garbage”, i.e., coefficients which are not needed to recover the product pq. The second line of the result denotes the wraparound produced by the -1. By inserting the right number of extra zeros, half of the garbage positions are used a second time for the garbage of the wraparound, while the desired coefficients are preserved. Speaking more algebraically, substituting $y \mapsto x^{3m+2}$ defines a vector space monomorphism $h : \mathcal{R}_{m,n} \to k[x]/(x^N - 1)$ such that

$$pq = h'(h(p)h(q)) \qquad \text{for all } p, q \in \mathcal{R}_{m,n},$$

where $h' : k[x]/(x^N - 1) \to \mathcal{R}_{m,n}$ is the vector space epimorphism defined by

$$\sum_{\kappa=0}^{N-1} \xi_\kappa x^\kappa \mapsto \sum_{\mu=0}^{m} \sum_{\nu=0}^{n} \xi_{\mu+(3m+2)\nu} x^\mu y^\nu.$$

Hence $R(\mathcal{R}_{m,n}) \leq R(k[x]/(x^N - 1))$. Over "suitable" fields, we have $R(k[x]/(x^N - 1)) = N$, thus proving

Theorem 5 (Schönhage) *Over "suitable" fields, the rank of the multiplication in $\mathcal{R}_{m,n} = k[x,y]/(x^{m+1}, y^{n+1})$ has the upper bound*

$$R(\mathcal{R}_{m,n}) \leq 3mn + 2m + 2n + 1.$$

3.2. The Triangular Case

Let us turn to the algebra $\mathcal{T}_n$. Adapting the method of the last section, we reduce the multiplication of $p, q \in \mathcal{T}_n$ to the multiplication of some $p', q' \in k[x]/(x^N - 1)$ for suitable N. Here we substitute $y \mapsto x^{2n+1}$ and set $N = n(2n+1)+n+1$. From the product $r' = p'q' \mod x^N - 1$, the coefficients of $pq \in \mathcal{T}_n$ can be read off as depicted in the pattern

d d d d o o o ' d d d o o o o ' d d o o o o o ' d o o o
for the factors p', q'

d d d d * * * ' d d d * * * o ' d d * * * o o ' d * * * /
o o o o * * * ' o o o o * * o ' o o o o *
for the result r'

for $n = 3$, $N = 25$. In contrast to $\mathcal{R}_{m,n}$, the garbage packages of the wraparound consist of at most n coefficients instead of $2m + 1$, hence smaller gaps of extra zeros are sufficient to preserve the desired coefficients of the result. So at a first glance, we get $R(\mathcal{T}_n) \leq R(k[x]/(x^N - 1)) \leq 2n^2 + 2n + 1$ over "suitable" fields. Compared with the upper bound $2n^2 + 3n + 1$ achieved by the evaluation-multiplication-interpolation technique (see [5, Theorem 5]), we have gained only little.

We make progress by observing that not the "full" multiplication in $k[x]/(x^N - 1)$ is needed. In the result pattern, the coefficients of x^κ with

$$\kappa \in \{\mu + (2n+1)\nu \mid 1 \leq \nu < n, 2n - \nu < \mu \leq 2n\} =: I$$

are always zero, i.e., $r' \in \langle x^\kappa \mid \kappa \in \{0, \ldots, N-1\} \setminus I \rangle$.

Now, let $\beta = (f_1, g_1, w_1, \ldots, f_r, g_r, w_r)$ be a bilinear computation of length r for the multiplication map ϕ of $k[x]/(x^N - 1)$. Since ϕ is concise, $w_1, \ldots, w_N$ form a basis of $k[x]/(x^N - 1)$ after a suitable renumbering. If $U \subseteq k[x]/(x^N - 1)$ is a vector space, $\dim U = N - m$, we may assume w.l.o.g. that $\langle w_1, \ldots, w_m \rangle \oplus U = k[x]/(x^N - 1)$. Let π be the projection along $\langle w_1, \ldots, w_m \rangle$ onto U. Then $\beta' := (f_{m+1}, g_{m+1}, \pi(w_{m+1}), \ldots, f_r, g_r, \pi(w_r))$ is a bilinear computation of length $r - m$ for $\pi \circ \phi$. Thus, for any subspace U with $\dim U = N - m$, there is a projection π onto U such that $R(\pi \circ \phi) \leq R(\phi) - m$.

Choosing $U = \langle x^\kappa \mid \kappa \in \{0, \ldots, N-1\} \setminus I \rangle$, we have $\pi(r') = r'$, therefore with $h : \mathcal{T}_n \to k[x]/(x^N - 1)$ being the monomorphism induced by substituting x^{2n+1} for y and $h' : k[x]/(x^N - 1) \to \mathcal{T}_n$ the epimorphism defined by

$$\sum_{\kappa=0}^{N-1} \xi_\kappa x^\kappa \mapsto \sum_{\nu=0}^{n} \sum_{\mu=0}^{n-\nu} \xi_{\mu+(2n+1)\nu} x^\mu y^\nu,$$

the following diagram commutes:

$$\begin{array}{ccc} \mathcal{T}_n \times \mathcal{T}_n & \xrightarrow{\cdot} & \mathcal{T}_n \\ {\scriptstyle h \times h}\big\downarrow & & \big\uparrow{\scriptstyle h'} \\ k[x]/(x^N - 1) \times k[x]/(x^N - 1) & \xrightarrow{\pi \circ \phi} & U \end{array}$$

Thus $R(\mathcal{T}_n) \leq R(\pi \circ \phi) = N - \dim U$. Observing that $\dim U = \frac{1}{2}(n-1)n$, we obtain the following upper bound.

Theorem 6 *Over "suitable" fields, the rank of the multiplication in $\mathcal{T}_n = k[x,y]/(x^\mu y^\nu \mid \mu + \nu = n + 1)$ is bounded from above by*

$$R(\mathcal{T}_n) \leq \tfrac{3}{2}n(n+1) + n + 1 = 3 \cdot \dim \mathcal{T}_n - 2n - 2.$$

3.3. The General Case

The technique of the last subsection can again be applied to any finite dimensional algebra of bivariate truncated polynomials $A = k[x,y]/I$, where the ideal $I = (x^{d_0+1}, x^{d_1+1}y, \ldots, x^{d_n+1}y^n, y^{n+1})$ is described by a degree pattern $d_0 \geq d_1 \geq \cdots \geq d_n \geq 0$.

Theorem 7 *The rank of the multiplication in the algebra $A = k[x,y]/(x^{d_0+1}, x^{d_1+1}y, \ldots, x^{d_n+1}y^n, y^{n+1})$ with degree pattern $d_0 \geq d_1 \geq \cdots \geq d_n \geq 0$ has, over "suitable" fields, the upper bound*

$$R(A) \leq nd_0 + \sum_{\nu=0}^{n-1} \max\{l_\nu - d_0 - 1, l_{n+1+\nu}\} + l_n + n,$$

where $l_\nu = \max\{d_i + d_j + 1 \mid i + j = \nu\}$ is the length of the νth coefficient package, $0 \leq \nu \leq 2n$.

Proof. By substituting $y \mapsto x^{3d_0+2}$, we reduce the multiplication of two bivariate polynomials $p, q \in A$ to the multiplication of the univariate polynomials $p'(x) = p(x, x^{3d_0+2})$ and $q'(x) = q(x, x^{3d_0+2})$ in $k[x]/(x^N - 1)$ with $N = n(3d_0 + 2) + 2d_0 + 1$. Since $d_\nu \leq d_0$ for all $0 \leq \nu \leq n$, it is clear from the rectangular case that the product pq can be obtained from the product $r' = p'q' \in k[x]/(x^N - 1)$. The question is how many zeros the result pattern contains. The following pattern shows the νth coefficient package of the result pattern:

$$\ldots\text{'}\overbrace{\mathrm{d\,d}\ldots\mathrm{d}\,*\,*\ldots*}^{l_\nu}\,\mathrm{o\,o}\ldots\mathrm{o\,o}\ldots\mathrm{o}\text{'}\ldots$$
$$\ldots\text{'}\underbrace{\mathrm{o\,o}\ldots\mathrm{o}}_{d_0+1}\,\underbrace{*\,*\ldots*\,*\,*\ldots*}_{l_{n+1+\nu}}\,\mathrm{o}\ldots\mathrm{o}\text{'}\ldots$$

We have $\min\{3d_0+2-l_\nu, 2d_0+1-l_{n+1+\nu}\}$ zeros in the νth coefficient package of the result pattern for $0 \le \nu \le n-1$, and $2d_0+1-l_n$ zeros in the nth package which is shorter but not affected by the wraparound. As shown in the triangular case, we have $R(A) \le N - \#\text{zeros}$, whence

$$\begin{aligned} R(A) &\le n(3d_0+2)+2d_0+1 \\ &\quad - \sum_{\nu=0}^{n-1} \min\{3d_0+2-l_\nu, 2d_0+1-l_{n+1+\nu}\} \\ &\quad - (2d_0+1-l_n) \\ &\le nd_0 + \sum_{\nu=0}^{n-1} \max\{l_\nu - d_0 - 1, l_{n+1+\nu}\} + l_n + n, \end{aligned}$$

proving the theorem. □

3.4. The Rectangular Case, Revisited

While our bounds yield sharp results for $\mathcal{T}_1$ and $\mathcal{T}_2$, namely $R(\mathcal{T}_1) = C(\mathcal{T}_1) = 5$ and $R(\mathcal{T}_2) = C(\mathcal{T}_2) = 12$, the lower bound and the upper bound for $\mathcal{R}_{1,1}$ differ by one: $7 \le C(\mathcal{R}_{1,1}) \le R(\mathcal{R}_{1,1}) \le 8$. So what is the truth? If char $k \ne 2$, then $\mathcal{R}_{1,1} = \langle 1, x+y, x-y, xy\rangle$. Hence $\mathcal{R}_{1,1}$ is an algebra of minimal rank and $R(\mathcal{R}_{1,1}) = 7$ [4], see also [3, Theorem IV.31] or [2, Theorem 17.38]. How can we achieve similar savings in general? Consider the multiplication in $A := k[x,y]/(x^{m+1}, x^m y^n, y^{n+1})$. By the general case we know that $R(A) \le 3mn + 2m + 2n - 2$. If we take a look at the corresponding pattern,

d d d o o o o o ' d d d o o o o o ' d d d o o o o o ' d d o o o
for the factors

d d d ∗ ∗ o o o ' d d d ∗ ∗ o o o ' d d d ∗ ∗ o o o ' d d $\underline{*}$ ∗ ∗ /
o o o ∗ ∗ ∗ ∗ ∗ ' o o o ∗ ∗ ∗ ∗ o ' o o o ∗ ∗ ∗
for the result

here for $m = 2$, $n = 3$, $N = 29$, one moment's reflection shows that the underlined garbage marker $\underline{*}$ does not mark garbage at all: It corresponds to

$$\sum_{\substack{0\le\mu\le m, 0\le\nu\le n,\\ (\mu,\nu)\ne(0,0),(m,n)}} a_{\mu,\nu} b_{m-\mu,n-\nu},$$

where $a_{\mu,\nu}$ and $b_{\mu,\nu}$ denote the coefficients of the factors. So with two additional multiplications $a_{0,0}{\cdot}b_{m,n}$ and $a_{m,n}{\cdot}b_{0,0}$, we are able to obtain the product pq of $p, q \in \mathcal{R}_{m,n}$, which proves the following bound.

Theorem 8 *Over "suitable" fields, the rank of the multiplication in $\mathcal{R}_{m,n} = k[x,y]/(x^{m+1}, y^{n+1})$ has the upper bound*

$$R(\mathcal{R}_{m,n}) \le 3mn + 2m + 2n = 3{\cdot}\dim R_{m,n} - m - n - 3.$$

References

[1] A. Alder and V. Strassen. On the algorithmic complexity of associative algebras. *Theoret. Comput. Sci.*, 15:201–211, 1981.

[2] P. Bürgisser, M. Clausen, and M. A. Shokrollahi. *Algebraic Complexity Theory*. Springer, 1997.

[3] H. F. de Groote. *Lectures on the Complexity of Bilinear Problems*. LNCS 245. Springer, 1986.

[4] H. F. de Groote and J. Heintz. Commutative algebras of minimal rank. *Lin. Alg. Appl.*, 55:37–68, 1983.

[5] A. Schönhage. Bivariate polynomial multiplication patterns. In *Proceedings of the AAECC 11*, LNCS 948, pages 70–81. Springer, 1995.