Withdrawn paper: fast multiplication of integer matrices

Joris van der Hoeven

Fast multiplication of integer matrices
Withdrawn paper

Joris van der Hoeven

Laboratoire d'informatique
UMR 7161 CNRS
École polytechnique
91128 Palaiseau Cedex
France

Email: vdhoeven@lix.polytechnique.fr
Web: http://www.lix.polytechnique.fr/~vdhoeven

November 2, 2012

In this paper we will show that dense $n \times n$ matrices with integer coefficients of bit sizes $\leqslant b$ can be multiplied in quasi-optimal time. This shows that the exponent $\omega_{\mathbb{Z}}$ for matrix multiplication over $\mathbb{Z}$ is equal to two. Moreover, there is hope that the exponent can be observed in practice for a sufficiently good implementation.

Keywords: matrix multiplication, FFT, skew polynomials
A.M.S. subject classification: 15-04, 68Q25, 68W30

Erratum. Eric Schost discovered a bug in this paper: by construction, formula (7) does not hold, since the matrix $V_{a,n}$ is not invertible whenever $l \geqslant 2$. Unfortunately, we do not see how to repair this bug: not only $V_{a,n}$ is not invertible, but the rank of $V_{a,n}$ is actually very small (and equal to $\max\{q_1, ..., q_l\}$). This makes the proof collapse.

We have also tried a few other things, such as fixing $a = 2$ and trying to take many q_i for which a is a primitive root of unity of a small order. This also does not work, since the product of two operators in $\mathbb{Z}[X, Q]_2$ makes the coefficient sizes increase by n^2 (and not merely n) bits. Going until "small order n", we must have $q_1 \cdots q_l \mid \operatorname{lcm}(2^1 - 1, 2^2 - 1, ..., 2^n - 1)$. This lcm is of the same order as 2^{n^2}, but strictly smaller than 2^{n^2}.

1. Introduction

1.1. Main results

Let $\mathbb{A}$ be an effective ring, which means that there exist algorithms for performing all ring operations. It is classical that there exist quasi-optimal algorithms for many kinds of computations with polynomials over $\mathbb{A}$. For instance, two polynomials of degrees $<n$ can be multiplied using $\mathsf{M}_{\mathbb{A}}(n) = \mathcal{O}_{\mathbb{A}}(n \log n \log\log n)$ operations in $\mathbb{A}$ [5, 17, 6, 13]. If $\mathbb{A}$ admits primitive 2^p-th roots of unity for any p, then we even have $\mathsf{M}_{\mathbb{A}}(n) = \mathcal{O}_{\mathbb{A}}(n \log n)$. From the complexity point of view, multiplication is the central operation: good bounds for the complexities of division, g.c.d., multipoint evaluation, etc. are known in terms of the complexity of multiplication.

In a similar way, one may study the asymptotic complexity $\mathsf{MM}_{\mathbb{A}}(n)$ of $n \times n$ matrix multiplication. In particular, an open problem in algebraic complexity theory is to find the infemum $\omega \geqslant 2$ of all exponents α such that $\mathsf{MM}_{\mathbb{A}}(n) = \mathcal{O}_{\mathbb{A}}(n^{\alpha})$. The naive schoolbook multiplication admits a complexity $\mathcal{O}_{\mathbb{A}}(n^3)$, which shows that $\omega \leqslant 3$. The first non trivial algorithm for matrix multiplication was given by Strassen [18], who showed that $\omega \leqslant \log_2 7$. Subsequently, a series of better and better algorithms were given [16, 7, 5, 19], and the best current known bound is $\omega \leqslant 2.3727$. However, apart from Strassen's exponent $\log_2 7$, none of these theoretical exponents have been observed in practice.

One important feature of the currently known bounds is that they work uniformly for all rings $\mathbb{A}$; that is, the corresponding algorithms are all algebraic, and only make use of the ring operations in $\mathbb{A}$. Despite its elegance, this approach has two drawbacks:

- For more specific rings $\mathbb{A}$, such as $\mathbb{A} = \mathbb{Z}$ or $\mathbb{A} = \mathbb{Z}/p\,\mathbb{Z}$, there might exist non algebraic algorithms with a much better complexity.
- Algebraic complexity bounds do not take into account the bit sizes of coefficients. In practice, the actual bit complexity of an operation in $\mathbb{A}$ usually depends on the particular arguments.

In this paper, we will focus on the bit complexity of integer matrix multiplication. Let $\mathsf{MI}(n, b)$ denote the bit complexity of multiplying two $n \times n$ matrices whose coefficients are integers of bit sizes at most b, and let $\mathsf{I}(b) = \mathsf{MI}(1, b)$. Our main result is the following:

THEOREM 1. *There exists a universal constant C such that for each $n \in \mathbb{N}$, there exists a $B_n \in \mathbb{N}$ such that for all $b \geqslant B_n$, we have*

$$\mathsf{MI}(n, b) \;\leqslant\; C n^2 \log^3 n \log\log n\, \mathsf{I}(b).$$

The dependence of the offset B_n on n involves number theoretic issues about the existence of primes which satisfy certain properties. Under the additional assumption that the number Γ_N of Sophie Germain primes (i.e. primes p such that $2\,p + 1$ is also prime) less than N behaves in the expected way (as observed in practice up to very large N), our main theorem can be sharpened:

THEOREM 2. *Assume that $\Gamma_N \geqslant C\,N/\log^2 N$ for some constant C. Then*

$$\mathsf{MI}(n, b) \;=\; \mathcal{O}(n^2 \log^3 n \log\log n\, \mathsf{I}(b)),$$

where the bound is uniform in n and b, under the condition that $\log n = \mathcal{O}(b)$.

Modulo giving in on the sharpness of the complexity bound, we will also show that the number theoretic assumption can be removed altogether. Our main result is the following theorem, which can be summarized as "$\omega_{\mathbb{Z}} = 2$":

THEOREM 3. *For $\log n = \mathcal{O}(b)$, we have*

$$\mathsf{MI}(n, b) \;=\; \mathcal{O}(n^2 \log^4 n \log\log n\, \mathsf{I}(b)).$$

1.2. Outline of the paper

There are three main ideas behind our algorithms. First of all, we exploit the strong connection between matrix multiplication and multiplication of skew polynomials [9, 3, 1]. More specifically, let $\mathbb{A}[X, Q]_a$ with $a \in \mathbb{A}$ denote the ring of skew polynomials such that X and Q commute with all elements in $\mathbb{A}$, and such that $Q\,X = a\,X\,Q$. Given $n \in \mathbb{N}$, we also denote $\mathbb{A}[X, Q]_{a,<n,<n} = \{L \in \mathbb{A}[X, Q]_a : \deg_X L < n, \deg_Q L < n\}$. In section 2, we will recall how the product of two operators in $\mathbb{A}[X, Q]_{a,<n,<n}$ reduces to a finite number of $n \times n$ matrix multiplications. In section 3, we will also study a variant of this reduction. Assuming that a is a primitive n-th root of unity and denoting by $\mathbb{A}[X, Q]_{a,n}$ the quotient of $\mathbb{A}[X, Q]_a$ by the relations $X^n = Q^n = 1$, we will show that multiplication in $\mathbb{A}[X, Q]_{a,n}$ is essentially equivalent to matrix multiplication.

The second main idea is to use the observation that multiplication in $\mathbb{A}[X, Q]_{a,<n,<n}$ can actually be done much faster if the multiplicative order q of a is small compared to n. Indeed, in that case X^q and Q^q lie in the center of $\mathbb{A}[X, Q]_a$, so skew polynomials in $\mathbb{A}[X, Q]_{a,<n,<n}$ can be considered as bivariate polynomials in X^q and Q^q with coefficients in $\mathbb{A}[X, Q]_{a,<q,<q}$. An analogue of this observation in the setting of differential operators in $\mathbb{F}_p[X, \partial/\partial X]$ was first made in [3].

In itself, the above observation does not imply the existence of a fast algorithm for matrix multiplication, because the problem of multiplying two $n \times n$ matrices can not directly be reduced to the multiplication problem in $\mathbb{A}[X, Q]_{a,<n,<n}$, when a has small order. The third idea (see section 4) is to force such a reduction by using multi-modular techniques.

More precisely, we slightly increase n (if necessary) so that $n = q_1 \cdots q_l$ for small numbers $q_1, ..., q_l$ which are pairwise coprime. We also pick numbers $p_1, ..., p_l$ (again as small as possible), such that $p_1, ..., p_l, q_1, ..., q_l$ are pairwise coprime and $q_i \mid \phi(p_i)$ for each i. This latter condition ensures the existence of a primitive q_i-th root of unity a_i in $\mathbb{Z}/p_i\,\mathbb{Z}$, whence the existence of a primitive n-th root of unity a in $\mathbb{Z}/m\,\mathbb{Z}$ with $m = p_1 \cdots p_l$. Now the problem of $n \times n$ matrix multiplication over $\mathbb{Z}/m\,\mathbb{Z}$ first reduces to the multiplication problem in $\mathbb{Z}/m\,\mathbb{Z}[X, Q]_{a,n}$. We next reduce this problem to l multiplication problems in the rings $\mathbb{Z}/p_i\,\mathbb{Z}[X, Q]_{a_i,n}$. Since the a_i have small orders q_i in the rings $\mathbb{Z}/p_i\,\mathbb{Z}$, these latter problems can be solved fast. We finally reconstruct the actual product using Chinese remaindering.

Altogether this proves the existence of a fast matrix product in $(\mathbb{Z}/m\,\mathbb{Z})^{n\times n}$, which can be further lifted into a fast matrix product in $(\mathbb{Z}/m^k\,\mathbb{Z})^{n\times n}$ for any k, whence into a fast matrix product in $\mathbb{Z}^{n\times n}$. In order to gain a control over the offset $B_n \approx \log_2 m$ in theorem 1, it is important to construct "nice" sequences $p_1, ..., p_l, q_1, ..., q_l$, for which $p_1 \cdots p_l$ remains small as a function of n. This topic will be discussed in section 5. In practice, there is no real problem, because prime numbers abound. In theory however, proving the existence of small $p_1, ..., p_l, q_1, ..., q_l$ with all required properties relies on non trivial theorems from number theory. We will first prove the existence of a quite sharp complexity bound (theorem 2) under the plausible hypothesis that there are many Sophie Germain prime numbers. We will also prove the existence of a somewhat less sharp bound which holds in general (theorem 3).

2. Skew polynomial and matrix products

2.1. Skew polynomials and their matrices

Let $\mathbb{A}$ be a ring. Given a constant $a \in \mathbb{A}$, we will denote by $\mathbb{A}[X, Q]_a$ the ring of operators

$$L = \sum_{i,j} L_{i,j}\, X^i\, Q^j, \tag{1}$$

where X and Q commute with all elements in $\mathbb{A}$, and

$$Q\,X \;=\; a\,X\,Q.$$

An operator L of the form (1) naturally acts on $\mathbb{A}[X]$ using

$$L(X^k) \;=\; \sum_{i,j} L_{i,j}\,a^{jk}\,X^{i+k}. \tag{2}$$

Given $k, n, r \in \mathbb{N}$, we denote

$$\begin{aligned}\mathbb{A}[X]_{<k} &= \{P \in \mathbb{A}[X] \colon \deg P < k\}\\ \mathbb{A}[X\,,Q]_{a,<n,<r} &= \{L \in \mathbb{A}[X\,,Q]_a \colon \deg_X L < n, \deg_Q L < r\}.\end{aligned}$$

Any operator $L \in \mathbb{A}[X\,,\,Q]_{a,<n,<r}$ induces a linear mapping of $\mathbb{A}[X]_{<k}$ into $\mathbb{A}[X]_{<k+n}$, and the matrix of this mapping with respect to the canonical bases $(1, X, ..., X^{k+n-1})$ and $(1, X, ..., X^{k-1})$ will be denoted by

$$\mathrm{M}_{L,k+n,k} \;=\; \begin{pmatrix} L(1)_0 & \cdots & L(X^{k-1})_0 \\ \vdots & & \vdots \\ L(1)_{k+n-1} & \cdots & L(X^{k-1})_{k+n-1} \end{pmatrix}.$$

In view of (2), we have

$$\mathrm{M}_L \;=\; \mathrm{Tw}_{n,k}(\Lambda_{L,n,r}\,V_{a,r,k}), \tag{3}$$

where

$$\Lambda_{L,n,r} \;=\; \begin{pmatrix} L_{0,0} & \cdots & L_{0,r-1} \\ \vdots & & \vdots \\ L_{n-1,0} & \cdots & L_{n-1,r-1} \end{pmatrix} \qquad V_{a,r,k} \;=\; \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & a & \cdots & a^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & a^{r-1} & \cdots & a^{(k-1)(r-1)} \end{pmatrix},$$

and where $\mathrm{Tw}_{n,k}$ twists matrices as follows:

$$\mathrm{Tw}_{n,k}\begin{pmatrix} c_{0,0} & \cdots & c_{0,k-1} \\ \vdots & & \vdots \\ c_{n-1,0} & \cdots & c_{n-1,k-1} \end{pmatrix} \;=\; \begin{pmatrix} c_{0,0} & & \\ \vdots & \ddots & \\ c_{n-1,0} & & c_{0,k-1} \\ & \ddots & \vdots \\ & & c_{n-1,k-1} \end{pmatrix}$$

2.2. Multiplying skew polynomials

Given two operators $K, L \in \mathbb{A}[X, Q]_{a,<n,<n}$, we have

$$\mathrm{M}_{KL,4n,2n} \;=\; \mathrm{M}_{K,4n,3n}\,\mathrm{M}_{L,3n,2n}.$$

Under the assumption that the multiplicative order of a in $\mathbb{A}$ is at least $2\,n$, the matrix $V_{a,2n,2n}$ is invertible. In that case, we may therefore multiply K and L using

$$\Lambda_{KL,2n,2n} \;=\; \mathrm{Tw}^{-1}_{2n,2n}(\mathrm{Tw}_{n,3n}(\Lambda_{K,n,n}\,V_{n,3n})\,\mathrm{Tw}_{n,2n}(\Lambda_{L,n,n}\,V_{a,n,2n}))\,V^{-1}_{a,2n,2n}. \tag{4}$$

Assuming that $\mathbb{A}$ is an effective ring (i.e. we have algorithms for performing the ring operations), let us analyze the complexity of this algorithm in terms of the number of required operations in $\mathbb{A}$.

We first observe that the operation of multiplying a vector by the matrix $V_{a,r,k}$ with $k \geqslant r$ corresponds to multipoint evaluation of a polynomial of degree $<k$ on the geometric sequence $1, a, ..., a^{r-1}$. It is known [2] that this operation can be performed in time $\mathcal{O}_{\mathbb{A}}(\mathsf{M}_{\mathbb{A}}(r)\, k/r)$, where $\mathsf{M}_{\mathbb{A}}(r)$ denotes the cost of multiplying two polynomials in $\mathbb{A}[X]_{<r}$. Similarly, the operation of multiplying a vector by the matrix $V_{a,r,r}^{-1}$ corresponds to multipoint interpolation of a polynomial of degree $<r$ on the geometric sequence $1, a, ..., a^{r-1}$. Again, this operation can be performed in time $\mathcal{O}_{\mathbb{A}}(\mathsf{M}_{\mathbb{A}}(r))$ [2]. Clearly, the twisting operations can be performed in linear time.

Denoting by $\mathsf{MM}_{\mathbb{A}}(n)$ the cost of multiplying two $n \times n$ matrices in $\mathbb{A}^{n\times n}$, and by $\mathsf{SM}_{\mathbb{A},a}(n)$ the cost of multiplying two skew polynomials in $\mathbb{A}[X, Q]_{a,<n,<n}$, we thus obtain:

Lemma 4. *With the above notations, and under the assumption that a has order at least $2\,n$, we have*

$$\mathsf{SM}_{\mathbb{A},a}(n) \;=\; \mathcal{O}_{\mathbb{A}}(\mathsf{MM}_{\mathbb{A}}(n) + \mathsf{M}_{\mathbb{A}}(n)\, n).$$

We may view (4) as an evaluation-interpolation strategy, written symbolically as

$$K\, L \;=\; \mathrm{Eval}_{2n,2n}^{-1}(\mathrm{Eval}_{n,3n}(K)\, \mathrm{Eval}_{n,2n}(L)).$$

The evaluation and interpolation steps require $\mathcal{O}_{\mathbb{A}}(\mathsf{M}_{\mathbb{A}}(n)\, n)$ operations, whereas the inner product requires $\mathcal{O}_{\mathbb{A}}(\mathsf{MM}_{\mathbb{A}}(n))$ operations.

2.3. The case when the order of a is comparable to n

Let us now show how to generalize the lemma from the previous section to the case when the order of a is a constant times smaller than $2\,n$. More precisely, let k be a fixed constant, and assume that a has order at least $r = 2\,\lceil n/k \rceil$.

As a preliminary, let us first consider an operator $L \in \mathbb{A}[X, Q]_{a,<s,<s}$ with $s \in \mathbb{N}$ and an integer $p \in \mathbb{N}$. Then there exists a unique operator $L' \in \mathbb{A}[X, Q]_{a,<s,<s}$ with

$$Q^p\, L \;=\; L'\, Q^p,$$

and the coefficients of L' may be computed using $\mathcal{O}(s^2)$ operations in $\mathbb{A}$ using the formula

$$L'_{i,j} \;=\; a^{ip}\, L_{i,j}.$$

Similarly, there exist unique operators $L'', L''', L'''' \in \mathbb{A}[X, Q]_{a,<s,<s}$ with $L\, Q^p = Q^p\, L''$, $X^p\, L = L'''\, X^p$ and $L\, X^p = X^p\, L''''$, whose coefficients can be computed efficiently.

Now consider two operators $K, L \in \mathbb{A}[X, Q]_{a,<n,<n}$. We may decompose them

$$\begin{aligned} K &= \sum_{0 \leqslant i,j < k} X^{ir}\, K_{[i,j]}\, Q^{jr} \\ L &= \sum_{0 \leqslant i,j < k} X^{ir}\, L_{[i,j]}\, Q^{jr}, \end{aligned}$$

with $K_{[i,j]},\ L_{[i,j]} \in \mathbb{A}[X,\ Q]_{a,<r,<r}$. By what precedes, it takes $\mathcal{O}(r^2\ k^2)$ operations to compute operators $K'_{[i,j]}$ and $L'_{[i,j]}$ with

$$\begin{aligned} K &= \sum_{0 \leqslant i,j<k} X^{ir}\, Q^{jr}\, K'_{[i,j]} \\ L &= \sum_{0 \leqslant i,j<k} L'_{[i,j]}\, X^{ir}\, Q^{jr}. \end{aligned}$$

Using lemma 4, we may compute the k^4 products $P'_{[i,j,i',j']} = K'_{[i,j]}\ L'_{[i',j']}$ in time $\mathcal{O}_{\mathbb{A}}((\mathsf{MM}_{\mathbb{A}}(r) + \mathsf{M}_{\mathbb{A}}(r)\ r)\ k^4)$. Again using our preliminary, it takes $\mathcal{O}(k^4\ r^2)$ operations to compute all operators $P_{[i,j,i',j']}$ with

$$X^{ir}\, Q^{jr}\, P'_{[i,j,i',j']}\, X^{i'r}\, Q^{j'r} \;=\; X^{(i+i')r}\, P_{[i,j,i',j']}\, Q^{(j+j')r}.$$

Adding up the various complexity bounds, we have proved:

Lemma 5. *With the above notations, and under the assumption that a has order at least $2\lceil n/k\rceil$, we have*

$$\mathsf{SM}_{\mathbb{A},a}(n) \;=\; \mathcal{O}_{\mathbb{A}}(\mathsf{MM}_{\mathbb{A}}(n)\, k^2 + \mathsf{M}_{\mathbb{A}}(n)\, n).$$

We may still view the present algorithm as a generalized evaluation-interpolation strategy, written symbolically as

$$K\, L \;=\; \mathrm{Eval}^{-1}_{2n,2n;k}(\mathrm{Eval}_{n,3n;k}(K)\, \mathrm{Eval}_{n,2n;k}(L)). \tag{5}$$

This time, the evaluation and interpolation steps require $\mathcal{O}_{\mathbb{A}}\big(\mathsf{M}_{\mathbb{A}}\big(\frac{n}{k}\big)\ \frac{n}{k}\, k^2\big) = \mathcal{O}_{\mathbb{A}}(\mathsf{M}_{\mathbb{A}}(n)\ n)$ operations (the evaluated structures being the evaluations in the previous sense of the operators $K'_{[i,j]}$ and $L'_{[i',j']}$), whereas the inner product requires $\mathcal{O}_{\mathbb{A}}\big(\mathsf{MM}_{\mathbb{A}}\big(\frac{n}{k}\big)\ k^4\big) = \mathcal{O}_{\mathbb{A}}(\mathsf{MM}_{\mathbb{A}}(n)\, k^2)$ operations.

2.4. The case when a has a small order

The case when the order q of a is small with respect to n is actually even more favourable from the complexity point of view. Indeed, in this case, both X^q and Q^q lie in the center of $\mathbb{A}[X, Q]_a$. Setting $r = \lceil n/q\rceil$ and

$$\mathbb{S} \;=\; \mathbb{A}[X, Q]_{a,<q,<q},$$

we may thus rewrite any operator $L \in \mathbb{A}[X,\ Q]_{a,<n,<n}$ as a polynomial in $\mathbb{S}[X^q,\ Q^q]$ of degrees $<r$ in both X^q and Q^q. Moreover, the evaluation-interpolation mappings for operators $L \in \mathbb{S}$ extend to $\mathbb{S}[X^q, Q^q]$ in a coefficientwise manner:

$$\mathrm{Eval}_{n_1,n_2;k}\left(\sum_{i,j} L_{i,j}\, X^{iq}\, Q^{jq}\right) \;=\; \sum_{i,j} \mathrm{Eval}_{n_1,n_2;k}(L_{i,j})\, X^{iq}\, Q^{jq}.$$

In our specific case, we may take $k = 3$ and $n_1,\ n_2 \in \{1,\ 2,\ 3\}\ r$. Given two operator polynomials $K, L \in \mathbb{S}[X^q, Q^q]$, we may thus multiply them using

$$K\, L \;=\; \mathrm{Eval}^{-1}_{2q,2q;3}(\mathrm{Eval}_{q,3q;3}(K)\, \mathrm{Eval}_{q,2q;3}(L)).$$

It takes a time

$$\begin{aligned}\mathcal{O}_{\mathbb{A}}(\mathsf{M}_{\mathbb{A}}(q)\, q\, (n/q)^2) &= \mathcal{O}_{\mathbb{A}}\left(\frac{\mathsf{M}_{\mathbb{A}}(q)}{q}\, n^2\right)\\ &= \mathcal{O}_{\mathbb{A}}(\mathsf{M}_{\mathbb{A}}(n)\, n)\end{aligned}$$

to compute the evaluations and interpolation in this formula. The inner multiplication $\mathrm{Eval}_{q,3q;3}(K)\,\mathrm{Eval}_{q,2q;3}(L)$ really consists of 81 multiplications of polynomials with $\mathcal{O}(q)\times\mathcal{O}(q)$ matrix coefficients. Using a fast algorithm for polynomial multiplication, this can be done in time

$$\mathcal{O}_{\mathbb{A}}(\mathsf{MM}_{\mathbb{A}}(q)\,\mathsf{M}_{\mathbb{A}}(n/q)\, n/q) = \mathcal{O}_{\mathbb{A}}\left(\frac{\mathsf{MM}_{\mathbb{A}}(q)}{q^2}\,\mathsf{M}_{\mathbb{A}}(n)\, n\right).$$

Putting everything together, this proves:

Lemma 6. *With the above notations, and assuming that a has order $q<2\,n$, we have*

$$\mathsf{SM}_{\mathbb{A},a}(n) = \mathcal{O}_{\mathbb{A}}\left(\frac{\mathsf{MM}_{\mathbb{A}}(q)}{q^2}\,\mathsf{M}_{\mathbb{A}}(n)\, n\right).$$

Remark 7. When exploiting the fact that $L\in\mathbb{A}[X,Q]_{a,<n,<r}$ actually maps $\mathbb{A}[X]_{<l}$ into $\mathbb{A}[X]_{<l+n-1}$ for any l, it is actually possible to take $k=2$ instead of $k=3$, so that the inner multiplication step amounts to only 16 polynomial matrix products instead of 81.

3. Cyclic skew operators and matrix multiplication

3.1. Cyclic skew polynomials and their matrices

Consider the ring $\mathbb{A}[X,Q]_a$ from section 2.1 and assume that $n\in\mathbb{N}$ is such that $a^n=1$. In that case, we may define the quotient operator algebra

$$\mathbb{A}[X,Q]_{a,n} = \mathbb{A}[X,Q]_a/(X^n-1,Q^n-1).$$

Any operator $L\in\mathbb{A}[X,Q]_{a,n}$ naturally acts on the space $\mathbb{A}[X]_n=\mathbb{A}[X]/(X^n-1)$, so we may again consider its matrix

$$\mathrm{M}_L = \begin{pmatrix} L(1)_0 & \cdots & L(X^{n-1})_0\\ \vdots & & \vdots\\ L(1)_{n-1} & \cdots & L(X^{n-1})_{n-1}\end{pmatrix}$$

with respect to the canonical basis $1,\ldots,X^{n-1}$. This time, we have

$$\mathrm{M}_L = \mathrm{Rot}_n(\Lambda_L\, V_{a,n}), \tag{6}$$

where

$$\Lambda_L = \begin{pmatrix} L_{0,0} & \cdots & L_{0,n-1}\\ \vdots & & \vdots\\ L_{n-1,0} & \cdots & L_{n-1,n-1}\end{pmatrix} \qquad V_{a,n} = \begin{pmatrix} 1 & 1 & \cdots & 1\\ 1 & a & \cdots & a^{n-1}\\ \vdots & \vdots & & \vdots\\ 1 & a^{n-1} & \cdots & a^{(n-1)^2}\end{pmatrix},$$

and where Rot_n rotates the coefficients of a matrix as follows:

$$\mathrm{Rot}_n\begin{pmatrix} c_{0,0} & \cdots & c_{0,n-1} \\ \vdots & & \vdots \\ c_{n-1,0} & \cdots & c_{n-1,n-1} \end{pmatrix} = \begin{pmatrix} c_{0,0} & c_{n-1,1} & \ddots & & c_{1,n-1} \\ c_{1,0} & c_{0,1} & \ddots & c_{n-2,n-2} & \\ & c_{1,1} & \ddots & c_{n-1,n-2} & c_{n-2,n-1} \\ c_{n-2,0} & & \ddots & c_{0,n-2} & c_{n-1,n-1} \\ c_{n-1,0} & c_{n-2,1} & \ddots & c_{1,n-2} & c_{0,n-1} \end{pmatrix}$$

If a has order n and n is invertible in $\mathbb{A}$, then $V_{a,n}$ is an FFT matrix of order n, whose inverse is given by ${}^1\!/_n V_{a^{-1},n}$. We may thus compute Λ_L as a function of M_L using

$$\Lambda_L \;=\; \tfrac{1}{n}\,\mathrm{Rot}_n^{-1}(\mathrm{M}_L)\,V_{a^{-1},n}. \tag{7}$$

3.2. The fundamental equivalence

Assume now that $\mathbb{A}$ is an effective ring. Assume still that a has order n and that n is invertible in $\mathbb{A}$. Given $K, L \in \mathbb{A}[X, Q]_{a,n}$, we may then compute KL using

$$KL \;=\; \tfrac{1}{n}\,\mathrm{Rot}^{-1}(\mathrm{Rot}(\Lambda_K V_{a,n})\,\mathrm{Rot}(\Lambda_L V_{a,n}))\,V_{a^{-1},n}. \tag{8}$$

Indeed, for monomial operators K and L of the form $X^i Q^j$, it is easily checked that

$$\mathrm{M}_{KL} \;=\; \mathrm{M}_K\,\mathrm{M}_L,$$

and this identity extends to the general case by bilinearity. In the formula (8), the matrix multiplications by $V_{a,n}$ and $V_{a^{-1},n}$ can be done using n FFTs of length n. Inversely, given two n by n matrices $M, N \in \mathbb{A}^{n\times n}$, we may reduce the computation of their product to the multiplication of two operators in $\mathbb{A}[X, Q]_{a,n}$:

$$MN \;=\; \tfrac{1}{n^2}\,\mathrm{Rot}((\mathrm{Rot}^{-1}(M)\,V_{a^{-1},n})\,(\mathrm{Rot}^{-1}(N)\,V_{a^{-1},n})\,V_{a,n}).$$

Denoting by $\mathsf{MM}_{\mathbb{A}}(n)$ the cost of a matrix product in $\mathbb{A}^{n\times n}$ and by $\mathsf{CM}_{\mathbb{A}}(n)$ the cost of multiplication in $\mathbb{A}[X, Q]_{a,n}$, we thus obtain the following equivalence:

Lemma 8. *With the above notations, we have*

$$\begin{aligned} \mathsf{CM}_{\mathbb{A}}(n) &= \mathsf{MM}_{\mathbb{A}}(n) + \mathcal{O}_{\mathbb{A}}(\mathsf{M}_{\mathbb{A}}(n)\,n) \\ \mathsf{MM}_{\mathbb{A}}(n) &= \mathsf{CM}_{\mathbb{A}}(n) + \mathcal{O}_{\mathbb{A}}(\mathsf{M}_{\mathbb{A}}(n)\,n). \end{aligned}$$

4. Fast multiplication through modular reduction

4.1. The main algorithm

Let $p_1, ..., p_l, q_1, ..., q_l$ be numbers which are pairwise coprime, and such that

$$q_i \;\mid\; \phi(p_i) \tag{9}$$

for all i, where ϕ stands for Euler's totient function. We will call the sequence $p_1, ..., p_l$, $q_1, ..., q_l$ a *reductor sequence*. Let a_i be a primitive q_i-th root of unity in $\mathbb{Z}/p_i\mathbb{Z}$ for each i (such roots exist because of the assumption (9)). In practice, we will usually take $p_1, ..., p_l$, $q_1, ..., q_l$ to be prime powers. If the p_i are primes, then (9) reduces to $q_i \mid p_i - 1$.

Setting $m = p_1 \cdots p_l$ and $n = q_1 \cdots q_l$, there exists an $a \in \mathbb{Z}/m\,\mathbb{Z}$ with $a \bmod p_i = a_i$ for each i. Since the q_i are pairwise coprime, this element a is a primitive n-th root of unity in $\mathbb{Z}/m\,\mathbb{Z}$. Since the p_i and q_j are also pairwise coprime, the number n is invertible in $\mathbb{Z}/m\,\mathbb{Z}$.

Now consider the operator algebra $(\mathbb{Z}/m\,\mathbb{Z})[X, Q]_{a,n}$. The modular reductions

$$\begin{array}{rcl} (\mathbb{Z}/m\,\mathbb{Z})[X, Q]_{a,n} & \xrightarrow{\pi_i} & \mathbb{Z}/p_i\,\mathbb{Z}[X, Q]_{a_i,n} \\ i \bmod m & \longmapsto & i \bmod p_i \\ X & \longmapsto & X \\ Q & \longmapsto & Q \end{array}$$

are well-defined ring homomorphisms and we have the Chinese remaindering isomorphism

$$\begin{array}{rcl} (\mathbb{Z}/m\,\mathbb{Z})[X, Q]_{a,n} & \xrightarrow{\pi = \pi_1 \times \cdots \times \pi_l} & \mathbb{F}_{p_1}[X, Q]_{a_1,n} \times \cdots \times \mathbb{F}_{p_l}[X, Q]_{a_l,n} \\ L & \longmapsto & (\pi_1(L), ..., \pi_l(L)). \end{array}$$

Given two operators $K, L \in (\mathbb{Z}/m\,\mathbb{Z})[X, Q]_{a,n}$, we now compute their product using

$$K\,L \;=\; \pi^{-1}(\pi(K)\,\pi(L)).$$

Let us analyze the complexity of this method in terms of the number of required operations in $\mathbb{A} = \mathbb{Z}/m\,\mathbb{Z}$. In order to simplify our analysis, we will not exploit the fact that operations in $\mathbb{Z}/p_i\,\mathbb{Z}$ are usually easier than operations in $\mathbb{A}$. We will rather use the facts that elements in $\mathbb{Z}/p_i\,\mathbb{Z}$ can be represented by elements in $\mathbb{A}$ and that the ring operations in $\mathbb{Z}/p_i\,\mathbb{Z}$ can be "mimicked" by the corresponding operations in $\mathbb{A}$.

Theorem 9. *Under the above assumptions, we have*

$$\begin{array}{rcl} \mathsf{CM}_{\mathbb{A}}(n) & = & \mathcal{O}_{\mathbb{A}}\left(\left[\dfrac{\mathsf{MM}_{\mathbb{A}}(q_1)}{q_1^2} + \cdots + \dfrac{\mathsf{MM}_{\mathbb{A}}(q_l)}{q_l^2}\right] \mathsf{M}_{\mathbb{A}}(n)\, n\right) \\ & = & \mathcal{O}_{\mathbb{A}}((q_1 + \cdots + q_l)\, n^2 \log n \log\log n). \end{array}$$

Remark. Of course, we understand the complexity bounds $\mathcal{O}_{\mathbb{A}}(T)$ as follows: there exists a universal constant K such that for any p_i, q_i and a_i which satisfy our assumptions, we have a multiplication algorithm of complexity $\leqslant K\,T$.

Proof. The computation of $\pi(K)$ and $\pi(L)$ requires n^2 reductions modulo p_i for each i, which can be performed in time $\mathcal{O}(n^2\, l)$. In fact, using our representation of elements in $\mathbb{Z}/p_i\,\mathbb{Z}$, these reductions are really no-operations.

The inner product $\pi(K)\ \pi(L)$ comes down to the computation of the product $\pi_i(K)\ \pi_i(L)$ for each i. By lemma 6, this product $\pi_i(K)\ \pi_i(L)$ can be computed in time $\mathcal{O}_{\mathbb{A}}((\mathsf{MM}_{\mathbb{A}}(q_i)/q_i^2)\ \mathsf{M}_{\mathbb{A}}(n)\ n)$. Indeed, we may multiply $\pi_i(K)$ and $\pi_i(L)$ as operators in $(\mathbb{Z}/p_i\,\mathbb{Z})[X, Q]_{a,<n,<n}$ (i.e. forgetting about the relations $X^n = Q^n = 1$), which yields a product in $(\mathbb{Z}/p_i\,\mathbb{Z})[X, Q]_{a,<2n,<2n}$, and then substitute $X^n = Q^n = 1$.

The final reconstruction of $K\,L$ from $\pi(K)\ \pi(L)$ corresponds to the reconstruction of n^2 coefficients in $\mathbb{Z}/m\,\mathbb{Z}$ from their reductions modulo the p_i. Each individual reconstruction can be done in time $\mathcal{O}_{\mathbb{A}}(l \log l) = \mathcal{O}_{\mathbb{A}}(l \log n) = \mathcal{O}_{\mathbb{A}}(l\, \mathsf{M}_{\mathbb{A}}(n)/n)$ using binary splitting.

Adding up the various complexities, we obtain the result. The extra simplifications of the main bound occur by using the crude bound $\mathsf{MM}_{\mathbb{A}}(q) = \mathcal{O}_{\mathbb{A}}(q^3)$ and the classical bound $\mathsf{M}_{\mathbb{A}}(n) = \mathcal{O}_{\mathbb{A}}(n \log n \log\log n)$. □

4.2. Application to matrix multiplication

The combination of theorem 9 and lemma 8 immediately implies the existence of an efficient algorithm for $n \times n$ matrix multiplication. Of course, this also implies an efficient algorithm for $n' \times n'$ matrix multiplication for any $n' \leqslant n$, using zero padding.

Corollary 10. *Under the above assumptions, we have*

$$\begin{aligned} \mathsf{MM}_{\mathbb{A}}(n') &= \mathcal{O}_{\mathbb{A}}\left(\left[\frac{\mathsf{MM}_{\mathbb{A}}(q_1)}{q_1^2} + \cdots + \frac{\mathsf{MM}_{\mathbb{A}}(q_l)}{q_l^2}\right] \mathsf{M}_{\mathbb{A}}(n')\, n'\right) \\ &\quad \mathcal{O}_{\mathbb{A}}((q_1 + \cdots + q_l)\,(n')^2 \log n' \log\log n'), \end{aligned}$$

for any $n' \leqslant n$.

As a next point, we notice that the numbers $p_1^t, \ldots, p_l^t, q_1, \ldots, q_l$ are still pairwise coprime, and (9) implies

$$q_i \;\mid\; \phi(p_i^t) = p_i^{t-1}\,\phi(p_i).$$

Moreover, the primitive q_i-th roots of unity a_i in $\mathbb{Z}/p_i\,\mathbb{Z}$ can be Newton-Hensel lifted into primitive q_i-th roots of unity $\hat{a}_i$ in $\mathbb{Z}/p_i^t\,\mathbb{Z}$ using $\mathcal{O}(l \log t)$ operations in $\mathbb{Z}/p_i^t\,\mathbb{Z}$. With the numbers $p_1, \ldots, p_l, q_1, \ldots, q_l$ and the roots $a_1, \ldots, a_l$ as our only prior knowledge, this means that corollary 10 extends to the case when $\mathbb{A}$ is replaced by $\mathbb{Z}/m^t\,\mathbb{Z}$.

Now let $\mathsf{I}(b)$ denote the bit complexity of multiplying two integers in $\mathbb{Z}_{<b} = \{i \in \mathbb{Z}: 2\,|i| < 2^b\}$, and let $\mathsf{MI}(n, b)$ denote the bit complexity of multiplying two matrices in $\mathbb{Z}_{<b}^{n \times n}$. Taking t such that $m^t > n\,4^b$, we thus obtain:

Corollary 11. *With the above notations, we have*

$$\mathsf{MI}(n', b) = \mathcal{O}((q_1 + \cdots + q_l)\,(n')^2 \log n' \log\log n'\, \mathsf{I}(b))$$

for any $n' \leqslant n$.

Assume now that we are given an arbitrary number $n \in \mathbb{N}$. We may take $q_1, \ldots, q_l$ to be the sequence of the first l prime numbers, where l is minimal with $q_1 \cdots q_l \geqslant n$. Since $n > q_1 \cdots q_{l-1} \geqslant (l-1)!$, we must have $l = \mathcal{O}(\log n / \log\log n)$. Now the prime number theorem also implies that $q_l = \mathcal{O}(l \log l) = \mathcal{O}(\log n)$, whence

$$q_1 + \cdots + q_l = \mathcal{O}(l \log n) = \mathcal{O}(\log^2 n)$$

For fixed $q_1, \ldots, q_l$, Dirichlet's theorem also implies the existence of suitable prime numbers $p_1, \ldots, p_l$ which satisfy our assumptions. We thus conclude:

Theorem 12. *There exists a universal constant C such that for each $n \in \mathbb{N}$, there exists a $B_n \in \mathbb{N}$ such that for all $b \geqslant B_n$, we have*

$$\mathsf{MI}(n, b) \leqslant C n^2 \log^3 n \log\log n\, \mathsf{I}(b).$$

5. Computation of reductor sequences

5.1. Optimal reductor sequences for practical purposes

Assume now that we want to optimize corollary 11 for practical purposes. For a given value

of n, we then have to minimize the sum $q_1 + \cdots + q_l$ under the side conditions that the q_i are pairwise coprime and $q_1 \cdots q_l \geqslant n$. We have performed an exhaustive computation up to $n = 10000$, and for a few other large values of n. The results are displayed in tables 1 and 2.

In the statement of theorem 12, it is still somewhat unsatisfactory that we have no control over the dependence of the offset B_n on n. In order to keep the offset B_n as small as possible, we have to find sequences $p_1, ..., p_l$, $q_1, ..., q_l$ such that $q_1 + \cdots + q_l$ is small, but $p_1 \cdots p_l$ also does not grow too fast. For fixed values of $q_1, ..., q_l$, we have searched for corresponding values of $p_1, ..., p_l$ which minimize the product $p_1 \cdots p_l$. The results are again displayed in tables 1 and 2. Of course, non optimal sequences $q_1, ..., q_l$ might give rise to sequences $p_1, ..., p_l$ for which the product $p_1 \cdots p_l$ is even lower.

n	Best $q_1, ..., q_l$	$q_1 + \cdots + q_l$	Best $p_1, ..., p_l$	$p_1 \cdots p_l$
2	2	2	3	3
3	3	3	7	7
4	4	4	5	5
5	2, 3	5	5, 7	35
7	2, 5	7	3, 11	33
11	3, 4	7	7, 5	35
13	3, 5	8	7, 11	77
16	4, 5	9	13, 11	143
21	2, 3, 5	10	7, 13, 11	1001
31	2, 3, 7	12	5, 13, 29	1885
43	3, 4, 5	12	7, 13, 11	1001
61	2, 5, 7	14	3, 11, 29	957
71	3, 4, 7	14	13, 5, 29	1885
85	3, 5, 7	15	13, 11, 29	4147
106	3, 5, 8	16	47, 11, 17	1309
121	4, 5, 7	16	13, 11, 29	4147
141	2, 3, 5, 7	17	17, 13, 11, 29	70499
211	3, 4, 5, 7	19	13, 17, 11, 29	70499
421	2, 3, 7, 11	23	5, 13, 29, 23	43355
463	2, 5, 7, 9	23	13, 11, 29, 19	78793
631	3, 4, 5, 11	23	7, 13, 31, 23	64883
661	3, 5, 7, 8	23	13, 11, 29, 17	70499
841	3, 4, 7, 11	25	13, 5, 29, 23	43355
925	4, 5, 7, 9	25	13, 11, 29, 19	78793
1261	3, 5, 8, 11	27	7, 31, 17, 23	84847
1321	4, 5, 7, 11	27	13, 31, 29, 23	268801
1541	2, 3, 5, 7, 11	28	17, 13, 31, 29, 23	4569617
2311	5, 7, 8, 9	29	11, 29, 17, 19	103037
2521	2, 3, 5, 7, 13	30	17, 19, 11, 29, 53	5460961
2731	3, 4, 5, 7, 11	30	13, 17, 31, 29, 23	4569617
4621	3, 4, 5, 7, 13	32	19, 17, 11, 29, 53	5460961
5461	2, 5, 7, 9, 11	34	13, 31, 29, 19, 23	5107219
6931	3, 5, 7, 8, 11	34	13, 31, 29, 17, 23	4569617
9241	3, 5, 7, 8, 13	36	19, 11, 29, 17, 53	5460961

Table 1. For each $n \geqslant 2$, one example of a sequence of pairwise coprime numbers $q_1, ..., q_l$, which minimizes $q_1 + \cdots + q_l$ under the side condition $q_1 \cdots q_l \geqslant n$. On ranges between two successive values of n in the table, we may keep using the same best sequence as the latest smaller value of n in the table. For each fixed sequence $q_1, ..., q_l$ we also displayed a corresponding sequence $p_1, ..., p_l$ which minimizes the product $p_1 \cdots p_l$.

n	Best $q_1, ..., q_l$	$q_1 + \cdots + q_l$	Best $p_1, ..., p_l$	$p_1 \cdots p_l$
10	$2, 5$	7	$3, 11$	33
20	$4, 5$	9	$13, 11$	143
50	$3, 4, 5$	12	$7, 13, 11$	1001
100	$3, 5, 7$	15	$13, 11, 29$	4147
200	$2, 3, 5, 7$	17	$17, 13, 11, 29$	70499
500	$2, 5, 7, 9$	23	$13, 11, 29, 19$	78793
1000	$4, 5, 7, 9$	25	$13, 11, 29, 19$	78793
2000	$2, 3, 5, 7, 11$	28	$17, 13, 31, 29, 23$	4569617
5000	$3, 4, 5, 7, 13$	32	$19, 17, 11, 29, 53$	5460961
10000	$3, 5, 7, 8, 13$	36	$19, 11, 29, 17, 53$	5460961
20000	$4, 5, 7, 11, 13$	40	$17, 31, 29, 23, 53$	18629977
50000	$3, 4, 5, 7, 11, 13$	43	$19, 17, 31, 29, 23, 53$	353969563
100000	$3, 5, 7, 8, 11, 13$	47	$19, 31, 29, 17, 23, 53$	353969563
200000	$4, 5, 7, 9, 11, 17$	53	$13, 31, 29, 19, 23, 103$	526043557
500000	$2, 3, 5, 7, 11, 13, 17$	58	$19, 37, 31, 29, 23, 53, 103$	79351647329
1000000	$3, 4, 5, 7, 11, 13, 17$	60	$19, 37, 31, 29, 23, 53, 103$	79351647329
2000000	$3, 5, 7, 8, 11, 13, 17$	64	$19, 31, 29, 41, 23, 53, 103$	87930203797
5000000	$5, 7, 8, 9, 11, 13, 17$	70	$31, 29, 41, 19, 23, 53, 103$	87930203797
10000000	$5, 7, 8, 9, 13, 17, 19$	78	$11, 29, 41, 37, 53, 103, 191$	504571510487
20000000	$2, 5, 7, 9, 11, 13, 17, 19$	83	$31, 41, 29, 37, 23, 53, 103, 191$	32705407907021
50000000	$4, 5, 7, 9, 11, 13, 17, 19$	85	$41, 31, 29, 37, 23, 53, 103, 191$	32705407907021
100000000	$5, 7, 8, 9, 11, 13, 17, 19$	89	$31, 29, 41, 37, 23, 53, 103, 191$	32705407907021

Table 2. A table similar to table 1, but only for some selected values of n. Experimentally speaking, we notice that the bound $p_1 \cdots p_l \leqslant 4\, n^2$ is always satisfied.

5.2. Theoretical complexity bounds modulo a conjecture

One attractive approach to show the theoretical existence of good reductor sequences $p_1, ..., p_l$, $q_1, ..., q_l$ is to only consider sequences of odd prime numbers with $p_i = 2\, q_i + 1$ for each i. Prime numbers q_i of this type are called Sophie Germain primes and a famous conjecture in number theory states that there are infinitely many such primes. A quantative version of this conjecture due to Hardy and Littlewood states that the number Γ_N of Sophie Germain primes smaller than a given number N is asymptotic to

$$\Gamma_N \;\sim\; C_2 \frac{N}{\log^2 N}$$

where C_2 is the "twin prime constant", approximately 0.660161. Although this conjecture has not been proved, it is supported by numeric evidence up till $N = 10^{12}$ (see [4] for a nice survey on this kind of topic).

Proposition 13. *Assume that $\Gamma_N \geqslant C\, N/\log^2 N$ for some constant C. Then there exists a constant K such that for every $n \geqslant 2$, there exist Sophie Germain primes $q_1, ..., q_l$ such that $q_1, ..., q_l$ and $p_1 = 2\, q_1 + 1, ..., p_l = 2\, q_l + 1$ are pairwise distinct, and such that $q_1 \cdots q_l \geqslant n$, $q_1 + \cdots + q_l \leqslant K \log^2 n$ and $p_1 \cdots p_l \leqslant K\, 3^l\, n \log n \log\log n = n^{1+o(1)}$.*

Proof. Let $\gamma_1, \gamma_2, ...$ be the sequence of Sophie Germain primes in increasing order and $\psi(N) = C\, N/\log^2 N$. From the bound $\Gamma_N \geqslant \psi(N)$, it follows that

$$\gamma_N \;\leqslant\; \lceil \psi^{\text{inv}}(N) \rceil \;\sim\; \tfrac{1}{C}\, N \log^2 N.$$

Indeed, assume for contradiction that $\gamma_N > \lceil \psi^{\text{inv}}(N) \rceil$ and let $N' = \lceil \psi^{\text{inv}}(N) \rceil$. Then $\Gamma_{N'} \leqslant N - 1 < N$. But $\Gamma_{N'} \geqslant \psi(N') = \psi(\lceil \psi^{\text{inv}}(N) \rceil) \geqslant \psi(\psi^{\text{inv}}(N)) = N$.

Let us first prove the proposition except for the condition on $p_1 \cdots p_l$. Let $l \asymp \log n / \log\log n$ be minimal such that $l! \geqslant n$. Whatever we take for the values of $q_1, \ldots, q_l$, this will ensure that $q_1 \cdots q_l \geqslant n$. In fact, we can always take $q_1, \ldots, q_l$ in the set $\{\gamma_1, \ldots, \gamma_{2l}\}$: just keep taking elements q_i in this set, while removing both q_i and $2\,q_i + 1$ from it. It follows that

$$\begin{aligned} q_1 + \cdots + q_l &\leqslant \gamma_1 + \cdots + \gamma_{2l} \\ &\leqslant 2\,l\,\gamma_{2l} \\ &= \mathcal{O}(\tfrac{1}{C}(2\,l)^2 \log^2(2\,l)) \\ &= \mathcal{O}(\log^2 n). \end{aligned}$$

This proves the proposition except for the last condition.

Now let $\ell \leqslant l$ be minimal such that $q_1 \cdots q_\ell \geqslant n$. Then the sequence $q_1, \ldots, q_\ell$ still satisfies the same conditions as before. Moreover,

$$\begin{aligned} p_1 \cdots p_\ell &\leqslant 3^\ell\, q_1 \cdots q_\ell \\ &\leqslant 3^\ell\, n\, q_\ell \\ &\leqslant 3^\ell\, n\, \gamma_{2l} \\ &= \mathcal{O}(3^\ell\, n \log n \log\log n). \end{aligned}$$

This completes the proof of our proposition. □

Theorem 14. *Assume that $\Gamma_N \geqslant C\,N/\log^2 N$ for some constant C. Then*

$$\mathsf{MI}(n, b) = \mathcal{O}(n^2 \log^3 n \log\log n\, \mathsf{I}(b)),$$

where the bound is uniform in n and b, under the condition that $\log n = \mathcal{O}(b)$.

Proof. For a given value $n \in \mathbb{N}$, let $p_1, \ldots, p_l, q_1, \ldots, q_l$ be a reductor sequence as constructed in the above proposition. On the one hand, we have $q_1 + \cdots + q_l = \mathcal{O}(\log^2 n)$. On the other hand, we may use corollary 10 in order to compute an $n \times n$ matrix product with coefficients of bit sizes $\leqslant b$ as soon as the coefficients of the product fit in $\log m$ bits (or more bits, by using a power of m instead). If $\log n = \mathcal{O}(b)$, then our construction also ensures that $\log m = \log(p_1 \cdots p_l) = \mathcal{O}(\log n) = \mathcal{O}(b)$. Modulo using a power of m if necessary, this means that we may take $\log m \asymp b$. The theorem now follows from corollary 10. □

5.3. An unconditional uniform complexity bound

Let us now show that we can actually remove the conjectural hypothesis on Γ_N altogether from theorem 14, although we have to give in slightly on the sharpness of the estimate for $\mathsf{MI}(n, b)$. We will need the following effective version of Dirichlet's theorem due to Linnik [14, 15]:

Theorem 15. *Given two numbers $a, d \in \mathbb{N} \setminus \{0\}$ with $a < d$ and $\gcd(a, d) = 1$, let*

$$p(a, d) = \min\{a + n\,d : n \in \mathbb{N}, a + n\,d \text{ prime}\}.$$

Then there exist positive constants c and L such that

$$p(a, d) \leqslant c\,d^L,$$

for all a and d of the above kind.

It has been shown by Heath-Brown [8] that we may actually take $L = 11/2$.

LEMMA 16. *There exists a constant c with the following property: for every $n \geqslant 2$, there exists a reductor sequence $p_1, ..., p_l, q_1, ..., q_l$ such that*

$$\begin{aligned} n &\leqslant q_1 \cdots q_l \leqslant c\, n \log^2 n \\ p_1 \cdots p_l &\leqslant c\, n^6 \\ q_1 + \cdots + q_l &\leqslant c \log^3 n. \end{aligned}$$

Proof. We compute the sequences q_1, ..., q_l and p_1, ..., p_l as a function of n using the following algorithm:

1. Let $l := 0$ and $\mathcal{F}_0 := \varnothing$.
2. If $q_1 \cdots q_l \geqslant n$, then return $p_1, ..., p_l, q_1, ..., q_l$.
3. Let $q_{l+1} = \min \{q : q \text{ prime}, q > q_l, q \notin \mathcal{F}_l\}$.
4. Let $p_{l+1} = p(1, q_{l+1}) = \min \{p : p \text{ prime}, q_{l+1} \mid p - 1\}$.
5. Let $\mathcal{F}_{l+1} := \mathcal{F}_l \cup \{p_l\} \cup \{q : q \text{ prime}, q > q_{l+1}, q \mid p_{l+1} - 1\}$.
6. Set $l := l + 1$ and go to step 2.

In this algorithm, $\mathcal{F}_l$ stands for the set of "forbidden prime numbers" after stage l. We claim that the way we construct these sets ensures that $p_{l+1}, q_{l+1} \notin \{p_1, ..., p_l, q_1, ..., q_l\}$. Since $q_1, q_2, ...$ is increasing by construction, and $\mathcal{F}_l \supseteq \{p_1, ..., p_l\}$, we clearly have $q_{l+1} \notin \{p_1, ..., p_l, q_1, ..., q_l\}$. Assume for contradiction that $p_{l+1} = p_i$ for some i, so that $q_{l+1} \notin \mathcal{F}_i$. By construction,

$$\mathcal{F}_i \supseteq \{q : q \text{ prime}, q > q_i, q \mid p_i - 1\}.$$

Hence, for all primes $q > q_i$ with $q \notin \mathcal{F}_i$, we have $q \nmid p_i - 1$. In particular, $q_{l+1} \nmid p_i - 1$. By construction, we also have $q_{l+1} \mid p_{l+1} - 1$, whence $q_{l+1} \mid p_i - 1$. This contradiction proves our claim.

By Heath-Brown's theorem, there exists a c_1 for which $p_i \leqslant c_1\, q_i^{11/2}$ for all i. In particular, the set $\{p_i\} \cup \{q : q \text{ prime}, q > q_i, q \mid p_i - 1\}$ contains at most $c_2 := \log_2 c_1 + 13/2$ elements for all i. Consequently, $\mathcal{F}_i$ contains at most $i\, c_2$ elements, for all i. In other words, for every $i \in \mathbb{N}$, there exists a $j \leqslant (i\, c_2)^2$ for which q_i is the j-th prime number. By the prime number theorem, it follows that there exists a constant c_3 such that

$$q_i \leqslant c_3\, i^2 \log i,$$

for all $i \in \mathbb{N}$. On the other hand, we must have $q_i \geqslant i$, so the minimal l for which $q_1 \cdots q_l \geqslant n$ satisfies $l = \mathcal{O}(\log n / \log \log n)$. In particular, it follows that $q_l = \mathcal{O}(\log^2 n)$ and $q_1 \cdots q_l = \mathcal{O}(n \log^2 n)$. Using Heath-Brown's theorem a second time, it follows that $p_1 \cdots p_l = \mathcal{O}(n^6)$.

Let us finally estimate the sum $q_1 + \cdots + q_l$. On the one hand, we have already observed that $l = \mathcal{O}(\log n / \log \log n)$. On the other hand, we recall that $q_1, ..., q_l$ is increasing, with $q_l \leqslant c_3\, l^2 \log l$. Consequently, $q_1 + \cdots + q_l = \mathcal{O}(l^3 \log l) = \mathcal{O}(\log^3 n)$. This completes the proof of our lemma. □

We are now in a position to make the complexity bound in theorem 12 uniform in b.

THEOREM 17. *For $\log n = \mathcal{O}(b)$, we have*

$$\mathsf{MI}(n, b) \leqslant \mathcal{O}(n^2 \log^4 n \log \log n\, \mathsf{I}(b)).$$

Proof. Similar to the proof of theorem 14. □

Remark 18. Please recall that, given two integer matrices whose coefficients have bit sizes $\leqslant b$, the coefficients of their product have bit sizes $\leqslant 2\,b + \log_2 b$ (and this bound is sharp). This makes the condition $\log n = \mathcal{O}(b)$ in the theorem quite natural.

6. Conclusion

We have given a fast algorithm for matrix multiplication over the integers. Although the algorithm is not algebraic, it does imply the existence of fast multiplication algorithms over several other rings, such as finite fields, rings p-adic numbers, the rationals, floating point numbers (modulo preconditioning), etc. A more detailed discussion of such generalizations is planned for an upcoming paper, but the ideas are similar to those at the end of section 5 in [12].

Another interesting question is whether the quasi-optimal complexity can be observed in practice. This might actually be feasible for large but still realistic sizes $n \geqslant 1000$. However, this will require an implementation in which all subalgorithms are highly optimized and further improved so as to gain constant or logarithmic factors wherever possible. Let us briefly mention a few ideas in this direction:

- In section 2.4, instead of viewing elements of $\mathbb{A}[X,\, Q]_a$ as polynomials in X^q and Q^q, we might also view them as skew polynomials in $\mathbb{B}[X,\, Q]_{a,<q,<q}$, where $\mathbb{B} = \mathbb{A}[X^q, Q^q]$. This is especially attractive if there exists a fast evaluation-interpolation scheme [11, sections 2.1–2.3] for univariate polynomials over $\mathbb{A}$ of degrees $<n/q$.
- The existence of a fast evaluation-interpolation scheme as above can be forced by taking the p_i in section 4 not too small, so as to guarantee the existence of geometric progressions of size n/q_i modulo p_i, and then use [2]. Although this may be problematic if the coefficients of our original matrices are small, this should not be a problem as soon as these coefficients get larger.
- Yet another strategy for fast evaluation-interpolation is to pick the p_i in $2^e\,\mathbb{N} + 1$ for some exponent e with $2^e \geqslant n$. In that case, we guarantee the existence of primitive 2^e-th roots of unity modulo each p_i, and we may use evaluation-interpolation on TFT-points [10].
- In the proof of theorem 9, it is a pity that we have to consider operators in the algebra $\mathbb{A}[X, Q]_{a,<n,<n}$ and not remain in a "cyclic algebra" all the way along. When using an evaluation-interpolation scheme for polynomials in X^q and Q^q, it should be possible to consider operators in the algebra $\mathbb{A}[X, Q]_a$ quotiented by relations of the form $X^q - \xi$ and $Q^q - \chi$. For the X-variable, this extension should work nicely and give rise to "homothetic rotations"

$$\mathrm{Rot}_{n,\xi}\begin{pmatrix} c_{0,0} & \cdots & c_{0,n-1} \\ \vdots & & \vdots \\ c_{n-1,0} & \cdots & c_{n-1,n-1} \end{pmatrix} = \begin{pmatrix} c_{0,0} & \xi\, c_{n-1,1} & \ddots & & \xi\, c_{1,n-1} \\ c_{1,0} & c_{0,1} & \ddots & \xi\, c_{n-2,n-2} & \\ & c_{1,1} & \ddots & \xi\, c_{n-1,n-2} & \xi\, c_{n-2,n-1} \\ c_{n-2,0} & & \ddots & \xi\, c_{0,n-2} & \xi\, c_{n-1,n-1} \\ c_{n-1,0} & c_{n-2,1} & \ddots & c_{1,n-2} & c_{0,n-1} \end{pmatrix}.$$

  The Q-variable requires more care. Maybe one can do something by considering both quotients by $Q^q - \chi$ and by $Q^q + \chi$.
- In our complexity bounds, we have not fully taken advantage of the fact that operations in $\mathbb{Z}/p_i\,\mathbb{Z}$ are faster than operations in $\mathbb{Z}/m\,\mathbb{Z}$. It should be possible to gain a logarithmic factor by performing a more detailed analysis. In a similar way, using a sharper bound for $\mathsf{MI}(n, b)$ could lead to further improvements.

From a more theoretical point of view, we also used the simple bound

$$\frac{\mathsf{MM}_{\mathbb{A}}(q_1)}{q_1^2} + \cdots + \frac{\mathsf{MM}_{\mathbb{A}}(q_l)}{q_l^2} \leqslant q_1 + \cdots + q_l$$

in many of our complexity estimates. Of course, using the new algorithm, these bounds can actually be improved so as to remove one factor $\mathcal{O}_{\mathbb{A}}(\log n)$ from the final bounds (up to multiplication by iterated logarithms of n). However, we do not think that such improved bounds are realistic from a practical point of view.

Bibliography

[1] A. Benoit, A. Bostan and J. van der Hoeven. Quasi-optimal multiplication of linear differential operators. Technical Report, HAL, 2012. `Http://hal.archives-ouvertes.fr/hal-00685401`, accepted for FOCS '12.

[2] A. Bostan and É. Schost. Polynomial evaluation and interpolation on special sets of points. *Journal of Complexity*, 21(4):420–446, August 2005. Festschrift for the 70th Birthday of Arnold Schönhage.

[3] Alin Bostan, Frédéric Chyzak and Nicolas Le Roux. Products of ordinary differential operators by evaluation and interpolation. In J. Rafael Sendra and Laureano González-Vega, editors, *ISSAC*, pages 23–30. Linz/Hagenberg, Austria, July 2008. ACM Press.

[4] C. Caldwell. An amazing prime heuristic. Technical Report, Univ. of Tennessee at Martin, 2000. Available from `http://www.utm.edu/~caldwell/preprints/Heuristics.pdf`.

[5] D.G. Cantor and E. Kaltofen. On fast multiplication of polynomials over arbitrary algebras. *Acta Informatica*, 28:693–701, 1991.

[6] J.W. Cooley and J.W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Computat.*, 19:297–301, 1965.

[7] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. In *Proc. of the* 19^{th} *Annual Symposium on Theory of Computing*, pages 1–6. New York City, may 25–27 1987.

[8] D.R. Heath-Brown. Zero-free regions for dirichlet l-functions, and the least prime in an arithmetic progression. *Proc. London Math. Soc.*, 64(3):265–338, 1992.

[9] J. van der Hoeven. FFT-like multiplication of linear differential operators. *JSC*, 33(1):123–127, 2002.

[10] J. van der Hoeven. Notes on the Truncated Fourier Transform. Technical Report 2005-5, Université Paris-Sud, Orsay, France, 2005.

[11] J. van der Hoeven. Newton's method and FFT trading. *JSC*, 45(8):857–878, 2010.

[12] J. van der Hoeven and G. Lecerf. On the bit-complexity of sparse polynomial multiplication. Technical Report, HAL, 2010. `Http://hal.archives-ouvertes.fr/hal-00476223`, accepted for publication in JSC.

[13] A. Karatsuba and J. Ofman. Multiplication of multidigit numbers on automata. *Soviet Physics Doklady*, 7:595–596, 1963.

[14] Yu. V. Linnik. On the least prime in an arithmetic progression I. the basic theorem. *Rec. Math. (Mat. Sbornik) N.S.*, 15(57):139–178, 1944.

[15] Yu. V. Linnik. On the least prime in an arithmetic progression II. the Deuring-Heilbronn phenomenon. *Rec. Math. (Mat. Sbornik) N.S.*, 15(57):347–368, 1944.

[16] V. Pan. *How to multiply matrices faster*, volume 179 of *Lect. Notes in Math.* Springer, 1984.

[17] A. Schönhage and V. Strassen. Schnelle Multiplikation grosser Zahlen. *Computing*, 7:281–292, 1971.

[18] V. Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 13:352–356, 1969.

[19] V. Vassilevska Williams. Multiplying matrices faster than Coppersmith-Winograd. In 44^{th} *ACM Symp. on Theory of Computing (STOC'12)*. 2012. To appear. Preliminary version available at `http://cs.berkeley.edu/`.