A new faster algorithm for factoring skew polynomials over finite fields

Xavier Caruso* and Jérémy Le Borgne†

Abstract

In this paper, we provide an algorithm for the factorization of skew polynomials over finite fields. It is faster than the previously known algorithm, which was due to Giesbrecht ([Gie98]). There are two main improvements. The first one is obtained through a careful study of the structure of the quotients of a skew polynomial ring, using theoretical results relating skew polynomial rings and Azumaya algebras. The second improvement is provided by giving faster sub-algorithms for the arithmetic in skew polynomial rings, such as multiplication, division, and extended Euclidean division.

Contents

*Université Rennes 1, IRMAR, Campus de Beaulieu, 35042 Rennes Cedex, France, xavier.caruso@normalesup.org

†IRMAR & ENS Rennes, Campus de Ker Lann, Avenue Robert Schuman, 35170 Bruz, jeremy.leborgne@ens-rennes.fr

Introduction

The aim of this paper is to present a new algorithm for factorization in rings of skew polynomials over finite fields. These noncommutative rings have been widely studied, including from an algorithmic point of view, since they were first introduced by Ore in 1933. Today, one important application for the study of skew polynomials over finite fields is related to some error-correcting codes introduced in [Gab85].

The first significant results in terms of effective arithmetics in these rings, including an algorithm for factoring a skew polynomial as a product of irreducible elements, appear in Giesbrecht's paper [Gie98]. In the present paper, we give a factorization algorithm whose complexity improves on Giesbrecht's. We also describe various fast-multiplication algorithms for skew polynomials, and some additional algorithms such as Euclidean divison and gcd.

Let k be a finite field of characteristic p, and let σ be an automorphism of k. We denote by k^σ the subfield of k fixed by σ, and by q its cardinality. Let also r denote the order of σ on k; the extension k/k^σ is then cyclic of degree r. The ring $k[X,\sigma]$ of skew polynomials with coefficients in k is a noncommutative ring, on which multiplication is determined by $X \cdot a = \sigma(a) \cdot X$ for all $a \in k$. As we will see in the first section, a skew polynomial can always be factored as a product of irreducible skew polynomials. However, such a factorization is not unique in general.

In the second section, we will study more carefully the structure of skew polynomial rings, by putting them in the framework of Azumaya algebras. The structure theorem we will rely on is the following:

Theorem ([Ike84] Theorem 2, *cf* Theorem 2.1.2)**.** *The ring $k[X,\sigma][1/X]$ is an Azumaya algebra over its centre $k^\sigma[X^r][1/X^r]$.*

This Theorem appears in [Ike84]. We will give a relatively short proof of this result, which makes this paper self-contained.

This Theorem has many important consequences for our purpose. The first one is the existence of a *reduced norm* map $k[X,\sigma] \to k^\sigma[X^r]$, which turns out to have very nice properties related to factorizations. More precisely, we shall explain how it can be used to establish a close link between factorizations of a skew polynomial and basic linear algebra over finite extensions of k^σ.

The third section of the paper deals with algorithmic aspects of arithmetic in skew polynomial rings. We start by giving various fast-multiplication algorithms and, as usual, we derive from them efficient algorithms to compute Euclidean division and gcd.

Then, we reach the core algorithm of this paper: the factorization algorithm, which is presented in the fourth section. Making an intensive use of the theory developped previously, we obtain a very efficient probabilistic algorithm to factor a skew polynomial as a product of irreducible skew polynomials, `SkewFactorization`. Before stating our complexity theorem, we recall the soft-O notation : if u_n and v_n are two sequences, the notation $u_n = \tilde{O}(v_n)$ means that there exists a positive integer k such that $u_n = O(v_n \log^k v_n)$.

Theorem (*cf* Theorem 4.3.4)**.** *The algorithm `SkewFactorization` factors a skew polynomial of degree d in $k[X,\sigma]$ with average complexity*

$$\tilde{O}(dr^3 \log q + d \log^2 q + d^{1+\varepsilon} (\log q)^{1+o(1)}) + F(d, k^\sigma)$$

bit operations, for all $\varepsilon > 0$. Here $F(d,K)$ denotes the complexity of the factorization of a (commutative) polynomial of degree d over the finite field K.

Remark 0.0.1. In the above Theorem, the computation model we use is the computation tree model (see [BCS97], §4.4).

Remark 0.0.2. Let ω be an exponent strictly greater that 2 such that the complexity of the matrix multiplication is $\tilde{O}(n^\omega)$ for input matrices of size $n \times n$. If we assume further that $\log q$ remains bounded, there is a variant of Theorem 4.3.4 stating that `SkewFactorization` runs with complexity $\tilde{O}(dr^\omega + d^{1+\varepsilon}) + F(d, k^\sigma)$ bit operations. We note that this version, when it applies, is generally stronger (the factor r^3 is replaced by r^ω).

Today, the best (average) complexity known for polynomial factorization, due to Kedlaya and Umans [KU08] (improving a former algorithm by Kaltofen and Shoup [KS98]), is:

$$F(d, K) = (d^{3/2+o(1)} + d^{1+o(1)} \log q) \cdot (\log q)^{1+o(1)}$$

bit operations, where q is the cardinality of K. Assuming this value for $F(d, K)$, we see that the terms $d \log^2 q$ and $d^{1+\varepsilon}(\log q)^{1+o(1)}$ are negligible compared to $F(d, K)$. If furthermore $r^3 \ll d$, so is the term $dr^3 \log q$. With this extra assumption, the complexity of our algorithm is then comparable to the complexity of the factorization of a *commutative* polynomial of the same degree.

The complexity of our algorithm should be compared to the complexity of Giesbrecht's algorithm, which is:

$$\tilde{O}(d^4 r^2 \log q + d^3 r^3 \log q + d \cdot \mathrm{MM}(dr) \log q + d^2 r \cdot \log^2 q)$$

bit operations[1] where $\mathrm{MM}(n)$ is the complexity of the multiplication of two $n \times n$ matrices.

The strategy of our algorithm is roughly comparable to the one of Giesbrecht's: in order to factor P, we find a multiple N of P lying in the centre of $k[X, \sigma]$, we factor N in the centre (which is a commutative polynomial ring) and we recover a factorization of P from the factorization of N we have just computed. The two main improvements are the following. First, we obtain better algorithms to achieve basic operations (like multiplication, Euclidean division and gcd's). Using them as subroutines significantly improves the complexity of the factorization. The second improvement (which is the most important) is of theoretical nature: it strongly relies on the nice properties of Azumaya algebras (*e.g.* Morita equivalence and the existence of the reduced norm). For instance, in order to obtain the central multiple N, it will be enough to compute the reduced norm, for which efficient algorithms exist. In the same way, our theoretical results imply that, for some particular P, the quotient $k[X, \sigma]/k[X, \sigma]P$ is endowed with a rich structure and we use it to replace computations with large matrices over k^σ by computations with matrices of size at most r defined over a bigger field. Since usual arithmetics in field extensions is more efficient than computations with matrices (quasilinear vs subcubic), we gain a lot.

Our algorithm has been implemented in SAGE and MAGMA. We discuss briefly about the implementation and provide benchmarks in the fifth section.

This work was supported by the *Agence Nationale de la Recherche*, CETHop project referenced ANR-09-JCJC-0048-01.

[1] In Giesbrecht's paper, the complexity is given in number of operations in k^σ. Since any operation in k^σ requires $\tilde{O}(\log q)$ bit operations (using fast algorithms), the complexity we have given is just obtained from Giesbrecht's one by multiplying by $\tilde{O}(\log q)$.

1 Problem: Factoring skew polynomials over finite fields

1.1 Some facts about $k[X, \sigma]$

Let k be a finite field of characteristic p and let σ be an automorphism of k. We denote by k^σ the subfield of k fixed by σ. Let r be the order of σ: r is also the degree of the extension k/k^σ. We denote by $k[X, \sigma]$ the ring of skew polynomials with coefficients in k. By definition, its underlying additive group is $k[X]$ and the multiplication on it is non-commutative and ruled by the formula:

$$\forall a \in k,\ Xa = \sigma(a)X.$$

Applying this rule n times, we find the relation $X^n a = \sigma^n(a)X^n$ for all $a \in k$. Using distributivity, it completely determines the multiplication on $k[X, \sigma]$.

Example 1.1.1. Let p be a prime number, $k = \mathbf{F}_{p^3}$, and let σ be the canonical Frobenius endomorphism on k. If $P = X^2 + a_1X + a_0 \in k[X, \sigma]$, and $Q = X + b_0 \in k[X, \sigma]$, we have:

$$\begin{aligned} PQ &= X^3 + (\sigma^2(b_0) + a_1)X^2 + (a_1\sigma(b_0) + a_0)X + a_0b_0, \\ QP &= X^3 + (\sigma(a_1) + b_0)X^2 + (\sigma(a_0) + a_1b_0)X + a_0b_0. \end{aligned}$$

We recall some notions from [Jac96], Chapter 1 (mainly §1.1 and 1.2). The centre of $k[X, \sigma]$ is $k^\sigma[X^r]$. The ring $k[X, \sigma]$ is endowed with left- and right-euclidean division algorithms. Hence, there are also notions of right- and left-greatest common divisor, and left- and right-lowest common multiple (denoted respectively by rgcd, lgcd, llcm, rlcm). Of course, every element of $k[X, \sigma]$ can be written as a product of irreducible elements of $k[X, \sigma]$. However, such a factorization is not unique in general (see Example 1.1.3). The first result that describes how two factorizations of a skew polynomial as a product of irreducible factors are related is due to Ore. Before stating it, we need a definition:

Definition 1.1.2. Let $P, Q \in k[X, \sigma]$ be two skew polynomials. Then P and Q are *similar* if there exist $U, V \in k[X, \sigma]$ such that $\mathrm{rgcd}(P, V) = 1$, $\mathrm{lgcd}(Q, U) = 1$ and $UP = QV$.

Even though it may not be clear at first glance, this is an equivalence relation. Remark that in the case $\sigma = \mathrm{id}$, this just means that P and Q are equal up to multiplication by an element of $k^\times$. We then have the following theorem:

Theorem (Ore, [Ore33])**.** *Let $P_1, \dots, P_n$ and $Q_1, \dots, Q_m$ be irreducible skew polynomials. If $P_1 \cdots P_n = Q_1 \cdots Q_m$, then $m = n$ and there exists a permutation τ of $\{1, \dots, n\}$ such that for all $1 \leq i \leq n$, P_i is similar to $Q_{\tau(i)}$.*

We insist on the fact that the converse of Ore's theorem is false: in general, if the P_i and Q_i are pairwise similar, $\prod P_i$ and $\prod Q_i$ are not even similar, let alone equal.

Example 1.1.3. Consider $k = \mathbf{F}_8$ presented as $\mathbf{F}_2[\alpha]$ where α is solution of the polynomial equation $\alpha^3 + \alpha + 1 = 0$. Endow k with the usual Frobenius $\sigma : t \mapsto t^2$. The polynomial

$$\begin{aligned} F(X) &= X^5 + X^4 + \alpha X^3 + \alpha^6 X^2 + \alpha^3 X + \alpha^2 \\ &= X^5 + X^4 + \alpha X^3 + (\alpha^2 + 1)X^2 + (\alpha + 1)X + \alpha^2 \end{aligned}$$

has twenty different factorizations recorded on Figure 1. It turns out that all polynomials of degree 1 (resp. of degree 2) appearing in this list are similar to each other[2]. We can then check that Ore Theorem is indeed true on this particular example.

[2] In fact, according to the third assertion of Proposition 2.1.17, all skew polynomials in $\mathbf{F}_8[X, \sigma]$ of degree 1 with a nonzero constant term are similar to each other.

N°	1st factor	2nd factor	3rd factor	4th factor
1	$X^2+\alpha^5X+\alpha$	$X+\alpha$	$X+1$	$X+1$
2	$X+\alpha^5$	$X^2+\alpha^2X+\alpha^4$	$X+1$	$X+1$
3	$X^2+\alpha^5X+\alpha$	$X+\alpha^6$	$X+\alpha^2$	$X+1$
4	$X+\alpha^6$	$X^2+X+\alpha$	$X+\alpha^2$	$X+1$
5	$X^2+\alpha^5X+\alpha$	$X+\alpha^2$	$X+\alpha^6$	$X+1$
6	$X+\alpha^3$	$X^2+\alpha^2X+1$	$X+\alpha^6$	$X+1$
7	$X+\alpha^6$	$X+\alpha$	$X^2+X+\alpha^2$	$X+1$
8	$X+\alpha^5$	$X+\alpha^2$	$X^2+X+\alpha^2$	$X+1$
9	$X+\alpha^3$	$X+\alpha^4$	$X^2+X+\alpha^2$	$X+1$
10	$X^2+\alpha^5X+\alpha$	$X+\alpha^6$	$X+\alpha^5$	$X+\alpha^4$
11	$X+\alpha^6$	$X^2+X+\alpha$	$X+\alpha^5$	$X+\alpha^4$
12	$X+\alpha^6$	$X+\alpha^3$	$X^2+\alpha^6X+\alpha^3$	$X+\alpha^4$
13	$X^2+\alpha^5X+\alpha$	$X+\alpha^6$	$X+\alpha^4$	$X+\alpha^5$
14	$X+\alpha^6$	$X^2+X+\alpha$	$X+\alpha^4$	$X+\alpha^5$
15	$X+\alpha^6$	$X+\alpha^4$	$X^2+\alpha^5X+\alpha$	$X+\alpha^5$
16	$X+\alpha^6$	$X+\alpha^4$	$X+1$	$X^2+\alpha^5X+\alpha^6$
17	$X+\alpha^6$	$X+\alpha^3$	$X+\alpha$	$X^2+\alpha^5X+\alpha^6$
18	$X+\alpha^6$	$X+\alpha$	$X+\alpha^3$	$X^2+\alpha^5X+\alpha^6$
19	$X+\alpha^5$	$X+\alpha^2$	$X+\alpha^3$	$X^2+\alpha^5X+\alpha^6$
20	$X+\alpha^3$	$X+\alpha^4$	$X+\alpha^3$	$X^2+\alpha^5X+\alpha^6$

Figure 1: The 20 factorizations of $X^5+X^4+\alpha X^3+\alpha^6X^2+\alpha^3X+\alpha^2$

An interesting point of view on skew polynomials is that of *φ-modules* that we shall elaborate on later. For now, it is enough to say that a φ-module over k is a $k[X,\sigma]$-module of finite type. If $P \in k[X,\sigma]$ is nonzero, a typical example of a φ-module over k is $k[X,\sigma]/k[X,\sigma]P$, which is "the φ-module associated to P". Then, two skew polynomials are similar if and only if the associated φ-modules are isomorphic, and Ore's theorem is just a restatement of the Jordan-Hölder Theorem in the category of φ-modules.

1.2 Statement of the problem

The main problem we are interested in in this paper is to design an algorithm for finding a factorization of a given skew polynomial into irreducible factors. Its input would then be a skew polynomial $P \in k[X,\sigma]$, while its output has to be a list of irreducible polynomials $(P_1, \dots, P_m)$ such that $P = P_m \cdots P_1$. Since, according to Ore's theorem, there is generally not a unique solution to this problem, we only require that the algorithm returns one solution.

Example 1.2.1. On the input

$$F(X) = X^5 + X^4 + \alpha X^3 + \alpha^6 X^2 + \alpha^3 X + \alpha^2$$

(*cf* Example 1.1.3) the algorithm we want to design should output one of the twenty factorizations recorded in Fig. 1. We underline that the algorithm we are going to design is probabilistic: as a consequence its output may vary among all possible factorizations (with a given distribution on which we do not require anything).

2 Structure of skew polynomial rings

The aim of this section is to make first the link between skew polynomials on the one hand and Azumaya algebras on the other hand and secondly to derive some consequences about factorizations.

2.1 Azumaya algebra and reduced norm

We first recall informally the definition of the latter: if C is a commutative ring, a C-algebra A is an Azumaya algebra if it becomes isomorphic to a matrix algebra after an étale extension of C. In particular, if C is a field, an algebra A is Azumaya over C if and only if A becomes isomorphic to a matrix algebra over a separable closure of C. Thus Azumaya algebras over a field are exactly *central simple algebras*. In general, Azumaya algebras appear as the natural[3] generalisation of central simple algebras over general rings (or even schemes). In first approximation, Azumaya algebra can be thought of as a matrix algebra because they share some properties, as the Morita equivalence or the existence of a determinant map. There are nonetheless important differences in their behaviours. An easy example is given by the quaternion algebra $\mathbf{H}$ over $\mathbf{R}$. It is indeed an Azumaya algebra (*i.e.* a simple central algebra) thanks to the isomorphism $\mathbf{H} \otimes_{\mathbf{R}} \mathbf{C} \simeq \mathcal{M}_2(\mathbf{C})$:

$$a + bi + cj + dk \mapsto \begin{pmatrix} a + b\sqrt{-1} & -c - d\sqrt{-1} \\ c - d\sqrt{-1} & a - b\sqrt{-1} \end{pmatrix} \quad \text{with } a, b, c, d \in \mathbf{C}.$$

[3] At least to the mathematician working in algebraic geometry...

However, we know that every nonzero element in $\mathbf{H}$ is invertible, which is certainly not the case in matrix algebras! Still there are positive points. For instance the usual norm of a quaternion — which is $a^2+b^2+c^2+d^2$ — is equal to the determinant to the associated complex matrix. As quickly mentionned above, this construction extends to all Azumaya algebras and defines what we call the *reduced norm*. We will make an extensive use of it and its properties in this paper.

2.1.1 Definitions and first properties

Let C be a commutative ring. Let us agree that, for any prime ideal $\mathfrak{P}$ of C, we denote by $C_{\mathfrak{P}}$ the fraction field of $C/\mathfrak{P}$. An alternative definition of Azumaya algebra is the following.

Definition 2.1.1 (Azumaya algebra)**.** Let C be a commutative ring, and let A be an algebra over C. Then A is an *Azumaya algebra* if, for every prime ideal $\mathfrak{P}$ of C, the $C_{\mathfrak{P}}$-algebra $A \otimes_C C_{\mathfrak{P}}$ is simple (*i.e.* it has no nontrivial two-sided ideals) and central (*i.e.* its centre is $C_{\mathfrak{P}}$).

We refer to [Azu51] and [Gro95] for a complete study of Azumaya algebras. Their relations with skew polynomial rings have also been studied, initially by Ikehata [Ike81, Ike84] and then by various authors. We recall the following Theorem of Ikehata, for which we also provide a proof:

Theorem 2.1.2 ([Ike84], Theorem 2)**.** *The ring $k[X,\sigma][1/X]$ is an Azumaya algebra over $k^\sigma[X^r][1/X^r]$.*

Proof. Let us denote by $\mathcal{R}$ the ring $k[X,\sigma][1/X]$ and by $\mathcal{C}$ its centre $k^\sigma[X^r][1/X^r]$. By definition, it is enough to show that for every prime ideal $\mathfrak{P}$ of $\mathcal{C}$, $\mathcal{R} \otimes_{\mathcal{C}} C_{\mathfrak{P}}$ is a central simple algebra over $\mathcal{C}_{\mathfrak{P}}$. The case $\mathfrak{P} = (0)$ is exactly [Jac96], Theorem 1.4.6. The other prime ideals of $\mathcal{C}$ are of the form (N) with $N \in k^\sigma[X^r]$ monic irreducible and different from X^r. Fix such an irreducible polynomial N and set $E = \mathcal{C}/N\mathcal{C}$. Let us first show that $\mathcal{R}_N = \mathcal{R} \otimes_{\mathcal{C}} E$ is simple. Let $I \subset \mathcal{R}_N$ be a two-sided ideal. Assume that $I \neq (0)$, and let $x \in \mathcal{R}_N$ be a nonzero element of I. First remark that every element $x \in \mathcal{R}_N$ can be written as $P \otimes 1$ (indeed, if t is the class of X^r is $E = \mathcal{C}/(N)$, then $1 \otimes t = X^r \otimes 1$). Now assume that x and P are chosen such that the number of nonzero coefficients of P is minimal (with $x \in I \setminus \{0\}$). We can assume that P is monic of degree d. We have $P - XPX^{-1} \in I$, and this polynomial has less nonzero coefficients than P, so that it is zero. Similarly, if $a \in k^\times$, $P - \sigma^d(a)^{-1}Pa = 0$. This shows that x is central. Since the centre of $\mathcal{R}_N$ is a commutative finite integral E-algebra, it is a field. Thus x is invertible and $I = \mathcal{R}_N$.

It remains to prove that this centre is exactly E. We just need to solve the equations $X\sum_{i=0}^{\deg(N)-1} a_i X^i = \sum_{i=0}^{\deg(N)-1} a_i X^{i+1}$ and $\alpha\sum_{i=0}^{\deg(N)-1} a_i X^i = \sum_{i=0}^{\deg(N)-1} a_i X^i \alpha$ for α a generator of k/k^σ. It is easy to see that the solutions are exactly (the reduction modulo N of) elements of $k^\sigma[X^r]$, so that the centre of $\mathcal{R}_N$ is E. □

This result has various corollaries that are interesting for questions about factoring skew polynomials.

Corollary 2.1.3. *Let $N \in k^\sigma[X^r]$ be an irreducible polynomial different from X^r. Let E_N be the quotient field $\mathcal{C}/N\mathcal{C}$. Then*

$$\mathcal{R}/N\mathcal{R} \simeq \mathcal{M}_r(E_N),$$

the ring of $r \times r$ matrices with coefficients in E_N.

Proof. It follows from the fact that any simple central algebra over a finite field E is isomorphic to a matrix algebra $\mathcal{M}_n(E)$ for some n. In our case, the equality $n = r$ follows from the fact that $\mathcal{R}/N\mathcal{R}$ has dimension $r^2 \deg N$ over k^σ. □

Corollary 2.1.4. *Let $N \in k^\sigma[X^r]$ be an irreducible polynomial different from X^r. Then the category of left-modules over $\mathcal{R}/N\mathcal{R}$ is equivalent to the category of vector spaces over $\mathcal{C}/N\mathcal{C}$.*

Proof. It follows from Corollary 2.1.3 and the equivalence theorem of Morita (see for example [AF92], §§21, 22). □

One of the usual objects associated to Azumaya algebras is the notion of *reduced norm*.

Definition 2.1.5 (Reduced norm)**.** Let A be an Azumaya algebra over a commutative ring C. Fix an étale extension C' of C, and a map τ such that $\tau : A \otimes_C C' \simeq \mathcal{M}_n(C')$ for some n. The the reduced norm

$$\mathcal{N}_{A/C} \ : \ A \to C,$$

is defined as $\mathcal{N}_{A/C}(x) = \det_C(x \otimes 1)$.

In our situation, it is a multiplicative morphism $\mathcal{N} \ : \ k[X,\sigma][1/X] \to k^\sigma[X^r][1/X^r]$ which can be alternatively defined as follows. Consider $k[X^r][1/X^r] \subset k[X,\sigma][1/X]$; it is a maximal étale subalgebra over the centre $k^\sigma[X^r][1/X^r]$ and it corresponds to the subalgebra of diagonal matrices through the isomorphism $k[X,\sigma][1/X] \simeq \mathcal{M}_r(k^\sigma[X^r][1/X^r])$. We deduce from this that $\mathcal{N}(P)$ is equal to the determinant of the right-multiplication by P on $k[X,\sigma][1/X]$ considered as a free $k[X^r][1/X^r]$-module (a basis of it being $(1, X, \dots, X^{r-1})$ for example). Using this, we deduce that the image of $k[X,\sigma]$ under $\mathcal{N}$ lies inside in $k[X^r]$ and hence in $k^\sigma[X^r]$. We furthermore note that, if P is a central skew polynomial (*i.e.* $P \in k^\sigma[X^r]$), the multiplication by P acts on a $k[X,\sigma]$ by the scalar matrix $P \cdot \mathrm{I}_r$ which has determinant P^r. Therefore $\mathcal{N}(P) = P^r$ provided that $P \in k^\sigma[X^r]$.

Example 2.1.6. Consider again the skew polynomial $F(X)$ of Example 1.1.3. Here $r = 3$ and the matrix of the right-multiplication by $F(X)$ in the $k[X^3][1/X^3]$-basis $(1, X, X^2)$ of $k[X,\sigma][1/X]$ is:

$$\begin{aligned} M_F &= \begin{pmatrix} \alpha X^3 + \alpha^2 & X^6 + \sigma(\alpha^6)X^3 & X^6 + \sigma^2(\alpha^3)X^3 \\ X^3 + \alpha^3 & \sigma(\alpha)X^3 + \sigma(\alpha^2) & X^6 + \sigma^2(\alpha^6)X^3 \\ X^3 + \alpha^6 & X^3 + \sigma(\alpha^3) & \sigma^2(\alpha)X^3 + \sigma^2(\alpha^2) \end{pmatrix} \\ &= \begin{pmatrix} \alpha X^3 + \alpha^2 & X^6 + \alpha^5 X^3 & X^6 + \alpha^5 X^3 \\ X^3 + \alpha^3 & \alpha^2 X^3 + \alpha^4 & X^6 + \alpha^3 X^3 \\ X^3 + \alpha^6 & X^3 + \alpha^6 & \alpha^4 X^3 + \alpha^8 \end{pmatrix} \end{aligned}$$

Its determinant is $(X^3)^5 + (X^3)^3 + (X^3)^2 + (X^3)$. It is then the reduced norm of $F(X)$. We remark that the reduced norm lies in the centre $\mathbf{F}_2[X^3]$ whereas the matrix M_F itself has only coefficients in $\mathbf{F}_8[X^3]$. This is a general phenomenon.

Remark 2.1.7. The property of being an Azymaya algebra remains true with some other noncommutative polynomial rings, as $k[X,\partial]$ where $\partial f = f\partial + f'$ [Rev73, BCS14]. However in general Corollary 2.1.3 fails.

2.1.2 Reinterpretation in terms of Galois representations

The Morita equivalence we have just stated can be reinterpreted in terms of Galois representations, recovering this way a variation of a theorem of Katz. In order to do so, let us first give one definition.

Definition 2.1.8. A *φ-module* over k is a finite dimensional k-vector space D endowed with an endomorphism $\varphi : D \to D$ that is semilinear with respect to σ, *i.e.* for all $x \in D$ and $a \in k$, $\varphi(\lambda x) = \sigma(\lambda)\varphi(x)$. A φ-module is said to be *étale* if the map φ is injective.

By definition, a φ-module (resp. an étale φ-module) over k is exactly a left-$k[X,\sigma]$-module having finite dimension over k.

Definition 2.1.9. If $P \in k[X,\sigma]$, the φ-module D_P associated to P is $k[X,\sigma]/k[X,\sigma]P$, endowed with the semilinear map φ given by left-multiplication by X. We say that P is étale if D_P is étale. (It exactly means that the constant coefficient of P is nonzero.)

Remark 2.1.10. Two skew polynomials P and Q are similar if and only if $D_P \simeq D_Q$.

The Morita equivalence shows the following:

Corollary 2.1.11. *The category of étale φ-modules over k is equivalent to the category of finite dimensional k^σ-vector spaces endowed with an invertible endomorphism.*

Proof. Let D be an étale φ-module over k. Since D has finite dimension over k, it is annihilated by some ideal (N) of $\mathcal{C}$. By Corollary 2.1.4, the categories of left-$\mathcal{R}/N\mathcal{R}$-modules and $\mathcal{C}/N\mathcal{C}$-modules are equivalent and we are done. □

This Corollary can also be seen as a variation of the following Theorem due to Katz in a quite larger generality.

Theorem 2.1.12 ([Kat73], Proposition 4.1.1)**.** *Let K be a field of characteristic $p > 0$ endowed with a power of the Frobenius endomorphism σ. Then the category of étale φ-modules over K is equivalent to the category of K^σ-representations of the absolute Galois group of K.*

Indeed, if $\sigma(a) = a^{p^s}$, let $K = k\mathbf{F}_{p^s}$. Then $K^\sigma = k^\sigma$, and the absolute Galois group of K is a procyclic group, so that a representation of this group is just the data of an invertible endomorphism of a k^σ-vector space of finite dimension (giving the action of a generator of the group). The functor giving this equivalence is explicit: the representation corresponding to an étale φ-module D over k is $\mathrm{Hom}_\varphi(D, K^{\mathrm{sep}})$.

Proposition 2.1.13. *Let (D,φ) be a φ-module over k, and let σ^r be the generator of the absolute Galois group of $k\mathbf{F}_{p^s}$. Then the action of σ^r on the k^σ-representation V corresponding to D is isomorphic to φ^r:*

$$(V \otimes_{k^\sigma} k, \sigma \otimes 1) \simeq (D, \varphi^r).$$

Proof. It is enough to prove the result when φ^r is cyclic. Let $f \in V = \mathrm{Hom}_\varphi(D, K^{\mathrm{sep}})$. Then for $x \in D$, $\sigma^r f(x) = f(\varphi^r(x))$. This shows that the polynomials annihilating σ^r and φ^r are the same. The characteristic and minimal polynomials of σ^r are the same, and equal to the characteristic polynomial of φ^r, so these two endomorphisms are conjugate. □

Using the fact that two skew polynomials are similar if and only if the corresponding φ-modules are isomorphic, we immediately get:

Corollary 2.1.14. *Let $P, Q \in k[X,\sigma]$. The skew polynomials P and Q are similar if and only if the $k^\sigma[X^r]$-modules (D_P, φ^r) and (D_Q, φ^r) are isomorphic.*

Since φ^r is a k-linear map, testing whether the aforementioned $k^\sigma[X^r]$-modules are isomorphic is straightforward. The notion of reduced norm (*cf* Definition 2.1.5) has also a nice interpretation in this context, as precised by the following lemma.

Lemma 2.1.15. *Let $P \in k[X,\sigma]$ be monic and let (D_P, φ) be the corresponding φ-module. Then the norm $\mathcal{N}(P)$ (viewed as a commutative polynomial in the variable X^r) is the characteristic polynomial of φ^r. If $P = a\tilde{P}$ with $\tilde{P}$ monic, then $\mathcal{N}(P) = N_{k/k^\sigma}(a) \cdot \mathcal{N}(\tilde{P})$.*

Proof. Let m_P be the right-multiplication by P acting on $k[X,\sigma]$. Since both $P \mapsto \mathcal{N}(P) = \det m_P$ and $P \mapsto \chi_{\varphi^r}$ are multiplicative, it is enough to prove the Lemma when P is monic irreducible. Let $\pi : k[X,\sigma] \to D_P$ be the canonical projection. We have $\pi \circ m_P = 0$. Since π is surjective, the multiplication by $\det m_P$ is also zero in D_P. This means that the minimal polynomial of the multiplication by X^r on D_P is a divisor of $\det m_P$. Since P is irreducible, this minimal polynomial is the characteristic polynomial χ of φ^r. It is then enough to show (1) that the degree of $\mathcal{N}(P)$ is the same as the degree of χ and (2) that $\mathcal{N}(P)$ is monic. Write $P = P_0 + XP_1 + \cdots + X^{r-1}P_{r-1}$ with the P_i's in $k[X^r]$. In the basis $(1, X, \ldots, X^{r-1})$, the matrix of m_P is:

$$\begin{pmatrix} P_0 & X^r\sigma(P_{r-1}) & \cdots & \cdots & X^r\sigma^{r-1}(P_1) \\ P_1 & \sigma(P_0) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & X^r\sigma^{r-1}(P_{r-1}) \\ P_{r-1} & \cdots & \cdots & \cdots & \sigma^{r-1}(P_0) \end{pmatrix}.$$

Let $0 \le i \le r-1$ be the greatest integer such that the degree of P_i is maximal, and denote by δ this degree. In the sum giving the determinant of this matrix, we have the term

$$P_i\sigma(P_i)\cdots\sigma^{r-i-1}(P_i)X^r\sigma^{r-i}(P_i)\cdots\sigma^{r-1}(P_i),$$

whose degree is $\delta(r-i) + (\delta+1)i = \delta r + i$ (as a polynomial in X^r). All the other terms of the determinant have degree less than this, so $\mathcal{N}(P) = \det m_P$ has degree $\delta r + i = \deg P = \deg \chi$ and is monic. □

Example 2.1.16. We illustrate the above Lemma with the skew polynomial $F(X)$ of Example 1.1.3 (see also Example 2.1.6). Since $F(X)$ has degree 5, a k-basis of D_F is $(1, X, X^2, X^3, X^4)$. The matrix of φ^3 (recall that $r = 3$ here) on D_F — which is nothing but the multiplication by X^3 on D_F — with respect to the above basis is:

$$N_F = \begin{pmatrix} 0 & 0 & \alpha^2 & \alpha^2 & 1 \\ 0 & 0 & \alpha^3 & \alpha^6 & \alpha^2 \\ 0 & 0 & \alpha^6 & 0 & 1 \\ 1 & 0 & \alpha & \alpha^6 & \alpha^6 \\ 0 & 1 & 1 & \alpha^6 & 0 \end{pmatrix}.$$

We note that the i-th column ($0 \leq i < r$) of this matrix gathers the coefficients of the remainder in the Euclidean division of X^{r+i} by $F(X)$. The characteristic polynomial of N_F is $\chi_F(T) = T^5 + T^3 + T^2 + 1$. The reduced norm of $F(X)$ is then $\chi_F(X^3)$. We can compare this result with that obtained (in a quite different way) in Example 2.1.6. Again, we may remark that $\chi_F(T)$ has coefficients in $\mathbf{F}_2$, whereas the entries of the matrix N_F do not lie in $\mathbf{F}_2$ but in $\mathbf{F}_8$.

2.1.3 Reduced norm and factorizations

The reduced norm appears as a very powerful tool to study factorizations. The next Proposition makes this remark more precise.

Proposition 2.1.17. *Let $\mathcal{N}$ be the reduced norm map on $k[X,\sigma]$. Then the following properties hold:*

- *$\forall P \in k[X,\sigma]$, P is a right- and left-divisor of $\mathcal{N}(P)$ in $k[X,\sigma]$,*
- *$\forall P \in k[X,\sigma]$, P is irreducible if and only if $\mathcal{N}(P)$ is irreducible in $k^\sigma[X^r]$,*
- *If $P, Q \in k[X,\sigma]$ and P is irreducible, then P and Q are similar if and only if $\mathcal{N}(P) = \mathcal{N}(Q)$ (up to multiplicative constant).*

Proof. The first fact is well-known (see for instance [Jac96], Proposition 1.7.1). It can be seen easily from the fact that if (D_P, φ) is the φ-module associated to P, then $\mathcal{N}(P)(\varphi) = 0$. Indeed, the left-ideal $\{R \in k[X,\sigma] \mid R(\varphi) = 0\}$ is exactly $k[X,\sigma]P$.
For the second assertion, remark that P is irreducible if and only if D_P is simple, which holds if and only if the corresponding representation is irreducible. This is true if and only if the characteristic polynomial of φ^r is irreducible in $k^\sigma[X^r]$.
Finally, we have already seen that the similarity class of a skew polynomial is determined by the conjugacy class of the action of φ^r on the corresponding φ-module (Corollary 2.1.14). For irreducible elements, this is completely determined by the characteristic polynomial of φ^r, *i.e.* the reduced norm. □

Since P is a divisor of $\mathcal{N}(P)$, we can expect that if $\tilde{N}$ is some irreducible factor of $\mathcal{N}(P)$ in $k^\sigma[X^r]$, then $\mathrm{rgcd}(\tilde{N}, P)$ would be a nonconstant right-divisor of P. This is actually always true and formalized by the following lemma:

Lemma 2.1.18. *Let $P \in k[X,\sigma]$ be étale and monic. Let $N = \mathcal{N}(P)$. If $N = N_1 \cdots N_m$ with all N_i's irreducible. Then there exist $P_1, \ldots, P_m \in k[X,\sigma]$ such that $P = P_1 \cdots P_m$ and for all $1 \leq i \leq m$, $\mathcal{N}(P_i) = N_i$.*
Moreover, P_m can be chosen as an irreducible right-divisor of $\mathrm{rgcd}(P, N_m)$.

Proof. By induction on m, it is enough to prove the last assertion. Let V_P be the Galois representation corresponding to the φ-module D_P *via* Katz's equivalence of categories (*cf* Theorem 2.1.12). Using Proposition 2.1.13, we find that V_P has a subrepresentation which is isomorphic to the quotient $k^\sigma[X^r]/N_m$ (where σ^r acts by multiplication by X^r). Hence, there exists a surjective map $D_P \to D_{P_m}$ where P_m is some skew polynomial of reduced norm N_m. It implies that P_m is a right divisor P and then also a right divisor of $\mathrm{rgcd}(P, N_m)$. This concludes the proof. □

Remark 2.1.19. This result shows how to determine the similarity classes of irreducible skew polynomials appearing in a factorization of P. It also shows that any order is possible for the appearance of these similarity classes in a factorization of P.

Assume that $\mathcal{N}(P)$ is the product of m *distinct* irreducible factors. If $\tilde{N}$ is such a factor, the right greatest common divisor $\mathrm{rgcd}(\tilde{N}, P)$ has to be irreducible as well. Indeed, it is a product of irreducible factors whose reduced norm is $\tilde{N}$ and P has at most one such factor since $\tilde{N}^2$ does not divide $\mathcal{N}(P)$. Combining this with the above Lemma, we find that P admits exactly $m!$ factorizations corresponding to each possible ordering of the factors of $\mathcal{N}(P)$. When $\tilde{N}$ is no longer assumed to be separable, it remains true that $\mathrm{rgcd}(\tilde{N}, P)$ is never constant. Therefore, in order to study the possible factorizations of P, we shall often assume that P divides $\tilde{N}$.

2.2 On the structure of D_P

In this subsection, we begin a close study of the structure of D_P equipped with the linear endomorphism φ^r. All the results we are going to prove will play a very important role in the correctness of our algorithm for factorization of skew polynomials, as well as for the probabilistic aspects of the algorithm.

2.2.1 Some remarks about rgcd's and llcm's

To begin with, we would like to clearly state the relations between rgcd's and llcm's in skew polynomial rings on the one hand and sums and intersections of $k[X,\sigma]$-modules on the other hand. We fix a skew polynomial P and set $D_P = k[X,\sigma]/k[X,\sigma]P$. To any divisor D of P, we can attach the submodule $D \cdot D_P$ of D_P consisting of the left multiples of D.

Lemma 2.2.1. *Let P be a skew polynomial. Let D_1 and D_2 be two right divisors of P. Then $R = rgcd(D_1, D_2)$ and $M = llcm(D_1, D_2)$ are right divisors of P and:*

- $D \cdot D_1 + D \cdot D_2 = k[X,\sigma]R/k[X,\sigma]P$,
- $D_1 \cap D_2 = k[X,\sigma]M/k[X,\sigma]P$.

Proof. Left to the reader. □

An important case of the above Lemma occurs when P is a divisor of an irreducible central element N. Set $E = k^\sigma[X^r]/N$ and let F be the E-vector space corresponding to P under the Morita equivalence of Corollary 2.1.4. The reduced norm of P is the equal to $N^{\dim_E F}$. Furthermore there is a one-to-one correspondance $D \mapsto F_D$ between the monic divisors of P and the sub-E-vector spaces of F. This correspondance is decreasing in the sense that if D_1 divides D_2 (both of them dividing P), then $F_{D_1} \subset F_{D_2}$. For D_1 and D_2 general, this implies that:

- the subspace corresponding to $\mathrm{rgcd}(D_1, D_2)$ is $F_{D_1} + F_{D_2}$, and
- the subspace corresponding to $\mathrm{llcm}(D_1, D_2)$ is $F_{D_1} \cap F_{D_2}$.

Similarly, we find that irreducible monic divisors of P (resp. factorizations of P into monic irreducible factors) correspond to hyperplanes in F (resp. complete flags of F). As a consequence, we get the following.

Proposition 2.2.2. *Let P be a skew polynomial dividing a central irreducible polynomial N. Define the integer e by $\mathcal{N}(P) = N^e$ and set $\delta = \deg_{k^\sigma[X^r]} N$. Then*

- *the number of irreducible monic right divisors of P is* $[e]_{q^\delta} = \frac{q^{e\delta}-1}{q^\delta-1}$
- *the number of factorizations of P into monic irreducible factors is*

$$[e]_{q^\delta}! = \frac{q^{e\delta}-1}{q^\delta-1} \cdot \frac{q^{(e-1)\delta}-1}{q^\delta-1} \cdots \frac{q^{2\delta}-1}{q^\delta-1} \cdot \frac{q^\delta-1}{q^\delta-1}.$$

We note that similar results already appear in [vzGGZ10, Section 4].

2.2.2 The notion of type

In order to study (D_P, φ^r), one thing we can do is to consider the Jordan form of φ^r. This actually can be easily determined using rgcd's.

Definition 2.2.3 (Type)**.** Let $P \in k[X,\sigma]$ and let N be an irreducible polynomial in the centre $k^\sigma[X^r]$ of degree d (in the variable X^r). The *N-type* — or simply the *type* if N is clear by the context — of P is the sequence $(e_1, e_2, \ldots)$ defined by:

$$d \cdot (e_1 + \cdots + e_i) = \deg \operatorname{rgcd}(P, N^i).$$

The *complete type* of P is the collection of couples $(N, N\text{-type of } P)$ for N running over all irreducible polynomials of $k^\sigma[X^r]$.

We derive from the fact that the degree of $\operatorname{rgcd}(P, N^i)$ is necessaly not greater than the degree of P that the e_i's must vanish when i is large enough. In the sequel, we shall often omit the final 0's in the type. It follows easily from the definition that $(e_1, e_2, \ldots)$ is the Jordan type of φ^r acting on the subspace $D_P[N^\infty]$ defined as characteristic subspace of D_P with respect to the endomorphism φ^r and any eigenvalue λ (lying in an algebraic closure of k) with $N(\lambda) = 0$. In particular, the sequence (e_i) is nonincreasing. Moreover, if a_i denotes the size of the i-th largest Jordan block of φ^r acting on $D_P[N^\infty]$, the Young diagram associated to $(a_1, a_2, \ldots)$ is dual to the one associated to $(e_1, e_2, \ldots)$. We shall say that $(a_1, a_2, \ldots)$ is the *N-dual type* of P.

Remark 2.2.4. Still denoting by $(e_1, e_2, \ldots)$ the N-type of a skew polynomial P, it follows directly from the definition that e_1 must be at most r (since N has degree dr when it is considered as a skew polynomial in $k[X,\sigma]$). In other words, the subspace $D_P[N^\infty]$ always admits at most r Jordan blocks. This implies in particular that the φ-module $k \oplus k \oplus \cdots \oplus k$ with $r+1$ summands is not isomorphic to some D_P.

There are close relationships between the complete type of a skew polynomial P and its reduced norm. Indeed, we first remark that the N-type of P is $(0, 0, \ldots)$ as soon as N does not divide the reduced norm of P because $\mathcal{N}(P)$ is the characteristic polynomial of φ^r action on D_P by Lemma 2.1.15. Conversely, the irreducible polynomials N for which P has a nonzero N-type are exactly the irreducible divisors of $\mathcal{N}(P)$. We can actually be more precise: if $(e_1, e_2, \ldots)$ is the N-type of P then the N-adic valuation of $\mathcal{N}(P)$ is exactly $\sum_i e_i$.

Example 2.2.5. In Examples 2.1.6 and 2.1.16, we have computed the reduced norm of the skew polynomial

$$F(X) = X^5 + X^4 + \alpha X^3 + \alpha^6 X^2 + \alpha^3 X + \alpha^2 \in \mathbf{F}_8[X,\sigma] \quad \text{with } \sigma : t \mapsto t^2$$

and found $\mathcal{N}(F) = (X^3)^5 + (X^3)^3 + (X^3)^2 + (X^3) \in \mathbf{F}_2[X^3]$ as a result. The factorization of $\mathcal{N}(F)$ into irreducible factors is

$$\mathcal{N}(F) = \big[(X^3) + 1\big]^3 \cdot \big[(X^3)^2 + (X^3) + 1)\big].$$

The factor $N_2 = (X^3)^2 + (X^3) + 1$ is simple. Therefore the N_2-type of P must be (1). Set now $N_1 = X^3 + 1$. In order to determine the N_1-type of P, we have to compute the rgcd's of F with the successive powers of N_1 until the result has degree 3 (which is the N_2-valuation of $\mathcal{N}(P)$). We get:

$$\begin{aligned}\operatorname{rgcd}(F, N_1) &= X^2 + \alpha^3 X + \alpha^2 \\ \operatorname{rgcd}(F, N_1^2) &= X^3 + \alpha X^2 + X + \alpha.\end{aligned}$$

Applying the definition, we find that the N_1-type of F is $(2, 1)$.

In the sequel, we shall often work with N-*isotopic* skew polynomials, that are skew polynomials P such that $\mathcal{N}(P)$ is a power of N. For such skew polynomials, the only non-trivial type is the N-type and we shall often abuse notations by calling it the *type* of P. Assuming that P is N-isotropic, it is clear from the definition that P has type (e) if and only if it divises N.

2.2.3 The case of a divisor of a central irreducible element

We have already seen several times that skew polynomials appearing as divisors of central irreducible elements play a singular role. The aim of this paragraph is to study them more carefully. Let $N \in k^\sigma[X^r]$ be a monic irreducible polynomial. We set $E = k^\sigma[X^r]/(N)$ as usual. Let $P \in k[X, \sigma]$ be a right-divisor of N. We first remark that, since $\mathcal{N}(N) = N^r$, the norm of P is N^e for some integer $e \in \{1, \dots, r\}$.

Lemma 2.2.6. *The φ-module D_N is isomorphic to a direct sum of r copies of a simple φ-module.*

Proof. It follows directly from Corollary 2.1.4. □

The Lemma implies that if P is a right-divisor of N with $\mathcal{N}(P) = N^e$, then the φ-module $D_P = k[X, \sigma]/k[X, \sigma]P$ is isomorphic to a direct sum of e copies of a simple φ-module. From this, we deduce that $\operatorname{End}_\varphi(D_P) \simeq \mathcal{M}_e(E)$.

From now on, we write $N = PQ$ for some $Q \in k[X, \sigma]$. Note that it implies that $QN = QPQ$; therefore $NQ = QPQ$ (since N lies in the centre) and, simplifying by Q, we get $N = QP$. In other words P and Q commute. The following proposition compares the φ-module $D_P = k[X, \sigma]/k[X, \sigma]P$ and its ring of endomorphisms.

Proposition 2.2.7. *The map*

$$\begin{array}{ccccc} D_P & \to & \operatorname{End}_\varphi(D_P) & & \\ R & \mapsto & m_{QR} & : & \left|\begin{array}{ccc} D_P & \to & D_P \\ x & \mapsto & xQR \end{array}\right. \end{array}$$

is a surjective additive group homomorphism.

Note that since $PQ = QP = N$ is central in $k[X,\sigma]$, the map above is well-defined. Indeed, we have to check that if $x \equiv x' \pmod P$ and $R \equiv R' \pmod P$ then $xQR \equiv x'QR' \pmod P$. Writing $x' = x + SP$ and $R' = R + TP$, we have:

$$\begin{aligned} x'QR' &= xQR + SPQR + (xQT + SPQ)P \\ &\equiv xQR + SNR \equiv xQR + SRN \equiv xQR + SRQP \equiv xQR \pmod P \end{aligned}$$

which is exactly what we want. In order to prove the proposition, we will need the following lemma, that states that in the case $P = N$, that map is in fact an isomorphism.

Lemma 2.2.8. *Let $N \in k[X,\sigma]$. Then the map:*

$$\begin{array}{ccccc} D_N & \to & \mathrm{End}_\varphi(D_N) & & \\ R & \mapsto & m_R : & \left| \begin{array}{ccc} D_N & \to & D_N \\ x & \mapsto & xR \end{array} \right. \end{array}$$

is an isomorphism of rings.

Proof. The fact that our map is a morphism of rings is straightforward. It is injective because $R = m_R(1)$. For the surjectivity, we remark that if N is a commutative polynomial of degree δ, D_N has dimension δr^2 over k^σ and, on the other hand, that if E is the field $k^\sigma[X^r]/(N)$, $\mathrm{End}_\varphi(D_N)$ is isomorphic to $\mathcal{M}_r(E)$, so it also has dimension δr^2. □

Proof of Proposition 2.2.7. We have the exact sequence of φ-modules:

$$0 \to k[X,\sigma]P/k[X,\sigma]N \to D_N \to D_P \to 0,$$

and D_Q is isomorphic to $k[X,\sigma]P/k[X,\sigma]N$ *via* the multiplication by P. Since $D_N \simeq D_P^{\oplus r}$, this sequence is split. Let $s : D_P \to D_N$ be a section. We have $Ps(1) = s(P) \equiv 0 \pmod N$, so there exists $S \in D_N$ such that $Ps(1) = NS$. Thus $s(1) = QS$. On the other hand, $QS = s(1) \equiv 1 \pmod P$. Hence there exists some $V \in k[X,\sigma]$ such that

$$QS + VP = 1.$$

It implies that D_P is isomorphic to $k[X,\sigma]QS/k[X,\sigma]N$ *via* the multiplication by QS.

Let $u \in \mathrm{End}_\varphi(D_P)$, and let $A = u(1) \in D_P$. For all $x \in k[X,\sigma]$, $u(x) = xu(1) = xA$. In other words, u is the mulitplication by A, *i.e.* $u = m_A$. We then want to show that m_A is of the form m_{QR} for some $R \in D_P$. Let $\tilde{u}$ the endomorphism of $k[X,\sigma]QS/k[X,\sigma]N$ deduced from u: we have $\tilde{u}(QS) = AQS$.

Since $D_N = k[X,\sigma]QS/k[X,\sigma]N \oplus k[X,\sigma]P/k[X,\sigma]N$ (decomposition of φ-modules), we can extend $\tilde{u}$ to D_N by setting $\tilde{u}(P) = 0$. By Lemma 2.2.8, there exists $T \in D_N$ such that for all $x \in D_N$, $\tilde{u}(x) = xT$. In particular:

$$\left\{ \begin{array}{rcll} PT & \equiv & 0 & \pmod N \\ QST & \equiv & AQS & \pmod N. \end{array} \right.$$

Since $VPT + QST = T$, we have $QST \equiv T \pmod N$. So, for $x \in D_N$, we get $\tilde{u}(xQS) = xQST = xQSAQS = (xQSA)QS$. Hence, for $x \in D_P$, $u(x) = xQSA$. Setting $R = SA$, we have $u = m_{QR}$. □

Corollary 2.2.9. *Let R be a random variable uniformly distributed on D_P. Then the right multiplication by QR, m_{QR}, is uniformly distributed on $\mathrm{End}_\varphi(D_P) \simeq \mathcal{M}_e(E)$.*

Proof. Since $R \mapsto m_{QR}$ is surjective, the probability that m_{QR} is equal to $u \in \mathrm{End}_\varphi(D_P)$ is proportional to the cardinality of the fibre above u. We conclude the proof by remarking that k-linearity together with surjectivity implies that all fibres have the same cardinality. □

3 Algorithms for arithmetics in skew polynomial rings

In this section, we describe algorithms for arithmetics in $k[X,\sigma]$: multiplication, Euclidean division, gcd's and lcm's, and we give their complexities. Throughout the rest of the paper, we will use the following notations:

- $\mathrm{MM}(n)$ is the number of operations (in k^σ) needed to compute the product of two $n \times n$ matrices with coefficients in k^σ.
- $\mathrm{SM}(d,r)$ is the number of operations (in k^σ) needed to multiply two skew polynomials with coefficients in k of degree at most d.

Below, we shall prove that one can take $\mathrm{SM}(d,r) = \tilde{O}(dr^2)$. Regarding matrix multiplication, the naive algorithm gives $\mathrm{MM}(n) = O(n^3)$ but it is well known that this complexity can be improved. For instance, using Strassen's algorithm, one have $\mathrm{MM}(n) = O(n^{\log_2 7})$. Today, the best known asymptotic complexity for matrix multiplication is due to Le Gall [LG14] and is about $O(n^{2.373})$.

We also assume that all usual arithmetics with polynomials can be done in quasilinear time. In particular, we assume that all usual operations (basically addition, multiplication and inverse) in an extension of k^σ of degree d requires $\tilde{O}(d)$ operations in k^σ. We refer to [GG03] for a presentation of algorithms having these complexities. Regarding the Frobenius morphism on k, we assume that all the conjugates of an element $a \in k$ can be computed in $O(r^2)$ operations in k^σ.

3.1 Fast arithmetics in skew polynomial rings

This section is dedicated to basic algorithms for arithmetics in skew polynomial rings.

3.1.1 Multiplication

Let $A, B \in k[X,\sigma]$, both of degree $\leq d$. We give several algorithms to compute the product AB and we compare their complexities.

The classical algorithm Let us recall that the classical algorithm of [Gie98], Lemma 1.1 (which throughout this section will be referred to as "Giesbrecht's algorithm") has complexity $\tilde{O}(d^2r + dr^2)$. This algorithm uses the explicit formula for the coefficients of the product of two skew polynomials: if $A = \sum_{i=0}^{d_1} a_i X^i$ and $B = \sum_{i=0}^{d_2} b_i X^j$, then their product is $\sum_{i=0}^{d_1+d_2} \left(\sum_{j=0}^{i} a_j \sigma^j(b_{i-j}) \right) X^i$. For each coefficient b_i of B, the list of the images of b_i under all the powers of σ can be computed in $O(r^2)$ operations in k^σ. Hence, all the $\sigma^j(b_{i-j})$ that may appear in the above formula can be computed in $\tilde{O}(d_2 r^2)$. Once we have these coefficients, it remains to compute the product, which is done with $O(d_1 d_2)$ operations in k, so the total complexity is $\tilde{O}(d_2 r^2 + d_1 d_2 r)$. To write it more simply, if both polynomials have degree less than d, then their product can be computed in $\tilde{O}(d^2 r + dr^2)$ operations in k^σ.

If r is large compared to d, we certainly do not need all images of the coefficients b_i's under the iterates of the Frobenius but only those of order less than d. In other words, the algorithm described just above does a lot of useless computations. In order to avoid them and desing nevertheless a competitive algorithm, we shall use fast modular composition as

it is described in [KU08]. Recall that, given three polynomials f, g and h over a finite field E, this algorithm compute $f \circ g \bmod h$ with complexity

$$\tilde{O}(n^{1+\varepsilon} \cdot (\log \#E)^{1+O(1)}) \quad (\text{for all } \varepsilon > 0)$$

bit operations where n denotes the maximal degree of f, g and h. Let us now explain how to use this to compute efficiently the iterates of σ. We denote by q the cardinality of k^σ and by e an integer such that σ acts on k by raising to the q^e-power. Let also $h(t)$ be the polynomial over k^σ be the polynomial defining the extension k/k^σ. We then have the identification $k \simeq k^\sigma[t]/h(t)$. Given $x(t) \in k^\sigma[t]$ (corresponding to an element $x \in k$) and an integer m, we can compute $\sigma^m(x(t))$ as follows: we first compute $f(t) = x(t)^q \bmod h(t)$ and then, using Kedlaya and Umans' algorithm, we compute $f \circ f \circ \cdots \circ f(t) \bmod h(t)$ where f is repeated $(em \bmod r)$ times. This gives a total complexity of:

$$\tilde{O}(r \log^2 q + r^{1+\varepsilon} \cdot (\log q)^{1+O(1)}) \quad (\text{for all } \varepsilon > 0)$$

bit operations.

Proposition 3.1.1. *Using these technics, the product of a skew polynomial of degree d_1 by a skew polynomial of degree d_2 can be computed in*

$$\tilde{O}(d_1 d_2 r \log q \cdot (\log q + r^\varepsilon \cdot (\log q)^{O(1)})) \quad (\textit{for all } \varepsilon > 0)$$

bit operations.

This complexity corresponds roughly speaking to

$$\tilde{O}(d_1 d_2 r \cdot (\log q + r^\varepsilon \cdot (\log q)^{O(1)})) \quad (\text{for all } \varepsilon > 0)$$

operations in k^σ. If r is large compared to d_2 and q remains small, this is betten than the complexity $\tilde{O}(d_2 r^2 + d_1 d_2 r)$ obtained above.

Algorithm 1: `SkewMultiplicationClassical`

Input: $(A, B) \in k[X, \sigma]^2$
Output: The product $AB \in k[X, \sigma]$

$d_1 = \deg A$, $d_2 = \deg B$;
$q = \#k^\sigma$;
Compute e such that $\sigma = (x \mapsto x^{q^e})$;
$h(t) =$ `DefiningPolynomial`(k/k^σ);
for $0 \leq j \leq d_2$ **do**
    $x(t) = B_j$ seen as an element of $k^\sigma[t]/(h(t))$;
    $f(t) = x(t)^q \pmod{h(t)}$;
    for $0 \leq m \leq d_1$ **do**
        $\beta_j^{(m)} = f \circ \cdots \circ f(t) \pmod{h(t)}$, where f is repeated $em \bmod r$ times
return $\sum_{i=0}^{d_1+d_2} \left(\sum_{j=0}^{i} A_j \beta_{i-j}^{(j)} \right) X^i$

The Karatsuba method Let $A, B \in k[X, \sigma]$. Write $A = A_0 + X^{mr} A_1$ and $B = B_0 + X^{mr} B_1$, with $m = \lfloor \frac{\max\{\deg A, \deg B\}}{2r} \rfloor$. We can then write:

$$AB = C_0 + X^{mr} C_1 + X^{2mr} C_2,$$

with $C_0 = A_0B_0$, $C_1 = A_0B_1 + A_1B_0$ and $C_2 = A_1B_1$, because X^{mr} lies in the center of $k[X, \sigma]$. If we set $P = (A_0 + A_1)(B_0 + B_1)$, we get the fact that $C_1 = P - C_0 - C_2$. Hence, we can recover the product AB doing the 3 multiplications $C_0 = A_0B_0$, $C_2 = A_1B_1$ and $P = (A_0 + A_1)(B_0 + B_1)$. Let $\mathrm{SM}_{\mathrm{Kar}}(d, r)$ be the number of multiplications needed in k^σ to multiply two elements of $k[X, \sigma]$ of degree $\leq d$ using this method. We get:

$$\mathrm{SM}_{\mathrm{Kar}}(d, r) \leq 3 \cdot \mathrm{SM}_{\mathrm{Kar}}\left(\frac{d}{2}, r\right) \leq 3^{\frac{\log(d/r)}{\log 2}} \cdot \mathrm{SM}(r, r).$$

Hence, this method allows to multiply polynomials of degree $\leq d$ in time $O\Big((\frac{d}{r})^{\frac{\log 3}{\log 2}}\mathrm{SM}(r, r)\Big)$ provided that $d > r$. Using Giesbrecht's algorithm for multiplication of skew polynomials of degree $\leq r$, we get a complexity of $O\big(d^{\frac{\log 3}{\log 2}} r^{3-\frac{\log 3}{\log 2}}\big)$, which is around $O\left(d^{1.58} r^{1.41}\right)$.

Algorithm 2: `SkewMultiplicationKaratsuba`

Input: $(A, B) \in k[X, \sigma]^2$
Output: The product $AB \in k[X, \sigma]$

$d_1 = \deg A$, $d_2 = \deg B$;
$d = \max\{d_1, d_2\}$;
$r = [k : k^\sigma]$;
if $d \leq r$ **then**
  return `SkewMultiplicationClassical`(A, B)
$m = \lfloor d/2r \rfloor$;
Write $A = X^{mr}A_1 + A_0$ and $B = X^{mr}B_1 + B_0$;
$C_0 =$ `SkewMultiplicationKaratsuba`(A_0, B_0);
$C_2 =$ `SkewMultiplicationKaratsuba`(A_1, B_1);
$C_1 =$ `SkewMultiplicationKaratsuba`$(A_0 + A_1, B_0 + B_1) - C_0 - C_2$;
return $X^{2mr}C_2 + X^{mr}C_1 + C_0$

Reduction to the commutative case Here, we use fast multiplication for commutative polynomials to multiply skew polynomials. Write $A = \sum_{i=0}^{r-1} A_i X^i$, with each A_i in $k[X^r]$. For $0 \leq i \leq r-1$, we denote by $B^{(i)}$ the skew polynomial deduced from B by applying σ^i to all coefficients. Then we have:

$$AB = \sum_{i=0}^{r-1} A_i B^{(i)} X^i.$$

Since $A_i \in k[X^r]$, it is easy to see that the product $A_iB^{(i)}$ is the same as the product of these polynomials computed in $k[X]$. The algorithm is the following:

1. Compute the $B^{(i)}$.
2. Compute all the products $A_iB^{(i)}$.
3. Compute the sum $AB = \sum_{i=0}^{r-1} A_i B^{(i)} X^i$.

Lemma 3.1.2. *The number of operations needed in k^σ for the multiplication of two skew polynomials of degree at most d by the above algorithm is $\tilde{O}(dr^2)$.*

Proof. We may assume that both A and B have degree d. For step 1, we need to compute all the conjugates of the d coefficients of B, which can be done in $O(dr^2)$ operations in k^σ. The multiplications of step 2 (as multiplications of elements of $k[X]$) can be done in $\tilde{O}(d)$ multiplications of elements of k, which corresponds to $\tilde{O}(dr)$ operations in k^σ. The total complexity of this step is then $\tilde{O}(dr^2)$ operations in k^σ. Finally, there are less than $2dr$ additions of elements of k to do in step 3., which is done in $O(dr^2)$. The global complexity is therefore $\tilde{O}(dr^2)$. □

Remark 3.1.3. This complexity is apparently better than that of Giesbrecht's algorithm (the term d^2r has gone) but we want to note that Giesbrecht's algorithm can beat this "commutative method" if the degree of B is much less than the degree of A. Indeed, in that case the dominant term in Giesbrecht's complexity is d_1d_2r which can be competitive with d_1r^2 if r is large compared to d_2.

Algorithm 3: `SkewMultiplicationCommutative`

Input: $(A, B) \in k[X, \sigma]^2$
Output: The product $AB \in k[X, \sigma]$

1 **for** $0 \leq i \leq r-1$ **do**
2 $\quad A_i = \sum_{j \equiv i \pmod r} a_j X^{j-i}$;
3 $\quad C_i = A_i \cdot \left(\sum_{j=0}^{\deg B} b_j X^j\right)$ computed in $k[X]$;
4 $\quad$ **for** $0 \leq j \leq \deg B$ **do**
5 $\quad\quad b_j = \sigma(b_j)$
6 **return** $\sum_{i=0}^{r-1} C_i X^i$

The matrix method Pick δ some positive integer and consider E/k the unique extension of degree δ. We fix an element $t \in E$ which generates this extension and denote by N its minimal polynomial over k^σ. We shall consider N as a polynomial in the variable X^r; it then lies in $k^\sigma[X^r]$, that is the centre of $k[X, \sigma]$. Considered as a skew polynomial, N has degree $r^2\delta$.

By Lemma 2.1.3, the algebra $k[X, \sigma]/N$ is isomorphic to $\mathcal{M}_r(E)$. Under our assumptions, the isomorphism can be made explicit. It maps an element $a \in k$ to the matrix:

$$M_a = \begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & \sigma(a) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sigma^{r-1}(a) \end{pmatrix}.$$

and the variable X to the matrix:

$$M_X = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ t & \cdots & 0 & 0 \end{pmatrix}.$$

More generally, if A is any skew polynomial, the image of A mod N in $\mathcal{M}_r(E)$ is the

matrix:

$$M_A = \begin{pmatrix} A_0(t) & \sigma(A_{r-1})(t) & \cdots & \sigma^{r-1}(A_1)(t) \\ tA_1(t) & \sigma(A_0)(t) & \cdots & \vdots \\ \vdots & \ddots & \ddots & \sigma^{r-1}(A_{r-1})(t) \\ tA_{r-1}(t) & \cdots & t\sigma^{r-2}(A_1)(t) & \sigma^{r-1}(A_0)(t) \end{pmatrix}.$$

where we have written $A \equiv \sum_{i=0}^{r-1} A_i X^i \pmod N$ with $A_i \in k[X^r]$ and $\deg A_i < \delta r$ for all i. The matrix M_A can be computed as follows. We first evaluate all the polynomials A_i's at all $\sigma^j(t)$ for $j \in \{0, \dots, r\}$. Using efficient algorithms (see [GG03], §10), it requires $\tilde{O}(\delta r^3)$ operations in k^σ. Let us extend σ to an automorphism of E. We can then compute $\sigma^j(A_i)(t)$ by applying σ^j to $A_i(\sigma^{-j}(t))$. Computing all these quantites requires

$$\tilde{O}(\delta r^3 \log^2 q + \delta^{1+\varepsilon} r^{3+\varepsilon} \cdot (\log q)^{1+o(1)}) \quad \text{(for all } \varepsilon > 0)$$

further bit operations as explain at the end of Paragraph 3.1.1. Then to obtain M_A, it remains to multiply some of the previous coefficients by t, which requires at most $\tilde{O}(\delta r^3)$ further operations in k^σ.

We can go in the other direction following the same ideas. We first divide by t all coefficients below the diagonal of M_A. We then apply σ^0 to the first column, σ^{r-1} to the second column, ..., σ to the last column and, finally, recover the A_i's by interpolation. The complexity is the same as before.

The sub-algorithms are the following:

Algorithm 4: `SkewPolynomialToMatrix`

Input: $(A, N) \in k[X, \sigma] \times k^\sigma[t]$
Output: The matrix M_A of multiplication by A modulo N

```
E = k[t]/(N);
for 0 ≤ i ≤ r − 1 do
    A_i = Σ_{j≡i (mod r)} a_i X^{j−i};
    for 0 ≤ j ≤ r − 1 do  a_{i,j} = A_i(σ^j(t));
for 0 ≤ i, j ≤ r − 1 do
    if i < j then
        m_{i,j} = σ^j(a_{i−j,−j})
    else
        m_{i,j} = tσ^j(a_{i−j,−j})
return (m_{i,j})
```

Once noticed these facts, the idea is quite simple. Let $A, B \in k[X, \sigma]$ of degree $< \frac{1}{2}\delta r^2$. We compute the corresponding matrices M_A, M_B, then the product $M_A M_B$ and finally recover the coefficients of (the reduction modulo N) of AB. This whole algorithm runs with complexity:

$$\tilde{O}(\delta r \cdot \mathrm{MM}(r) \log q + \delta r^3 \log^2 q + \delta^{1+\varepsilon} r^{3+\varepsilon} \cdot (\log q)^{1+o(1)}) \quad \text{(for all } \varepsilon > 0)$$

Since we are multiplying this way polynomials of degree at most $\frac{1}{2}\delta r^2$, we get:

$$\mathrm{SM}(d, r) = \tilde{O}\Big(d \tfrac{\mathrm{MM}(r)}{r} + dr \log q + d^{1+\varepsilon} r \cdot (\log q)^{o(1)}\Big) \quad \text{(for all } \varepsilon > 0)$$

by this method. If $\log q$ remains bounded, this order of magnitude of this complexity is comparable to $\tilde{O}(d \frac{\mathrm{MM}(r)}{r})$ operations in k^σ. If $\mathrm{MM}(r) \ll r^3$ (which is a usual assumption), this method beats the "reduction to the commutative case".

Complexity comparison The following chart recalls the complexities obtained with the different multiplication algorithms., for skew polynomials of degree $\leq d$, given in operations in k^σ.

Algorithm	Classical	Karatsuba
Complexity	$\tilde{O}\big(d^2 r \cdot (\log q + r^\varepsilon \cdot (\log q)^{O(1)})\big)$	$\tilde{O}\big(d^{1.58} r^{1.41}\big)$

Algorithm	Commutative	Matrix
Complexity	$\tilde{O}\big(dr^2\big)$	$\tilde{O}(d\,\frac{\mathrm{MM}(r)}{r} + dr\log q + d^{1+\varepsilon} r \cdot (\log q)^{o(1)})$

The better asymptotical theoretical complexity is achieved by the "matrix" method: it would even be quasi-optimal if matrix multiplication was. There are nevertheless counterparts. First of all, we notice that it is inefficient if $d < r^2$ (since the parameter δ should be an integer). Furthermore implementing it is far from being trivial for (at least) two reasons: first, it relies on fast matrix multiplication and second, the choice of a good extension E might be quite subtle. That is why the other algorithms we have presented might have some interest. The "commutative" method is very promising but again the announced complexity supposes that we are able to perform multiplication of commutative polynomials in quasi-linear time. Thus, in the range where Karatsuba algorithm is competitive for commutative polynomials, using the "Karatsuba" method might be the better choice.

3.2 Euclidean divisions

3.2.1 Euclidean division

Let $A, B \in k[X,\sigma]$ with $\deg A \geq \deg B$. We want to compute the right-Euclidean division of A by B:

$$A = QB + R,$$

with $\deg R < \deg B$. The following algorithm is based on the Newton iteration process presented for example in [GG03], §9.1, which uses reciprocal polynomials. Our algorithm is an almost direct adaptation of it, the only subtlety here is that the map sending a skew polynomial to its reciprocal polynomial is not a morphism.

Lemma 3.2.1. *For $\delta \geq 0$, we denote by $k[X,\sigma]_{\leq\delta}$ the subspace of skew polynomials of degree at most δ. Let*

$$\begin{array}{rccc} \tau_\delta & : & k[X,\sigma]_{\leq\delta} & \to & k[X,\sigma^{-1}]_{\leq\delta} \\ & & \sum_{i=0}^{\delta} a_i X^i & \mapsto & \sum_{i=0}^{\delta} a_{\delta-i} X^i \end{array}.$$

Then τ_δ is k-linear, bijective, and for all $P, Q \in k[X,\sigma]$, with $\deg P \leq \delta_1$ and $\deg Q \leq \delta_2$, we have:

$$\tau_{\delta_1}(P)\tau_{\delta_2}(Q^{(\delta_1)}) = \tau_{\delta_1+\delta_2}(PQ).$$

Proof. The k-linearity is trivial, as well as bijectivity. Let $P = \sum_{i=0}^{\delta_1} a_i X^i$ and $Q = \sum_{j=0}^{\delta_2} b_j X^j$. Then the coefficient of X^l in the product PQ is

$$c_l = \sum_{i+j=l} a_i \sigma^i(b_j).$$

Hence, the coefficient of X^l in $\tau_{\delta_1+\delta_2}(PQ)$ is $c_{\delta_1+\delta_2-l} = \sum_{i+j=l} a_{\delta_1-i}\sigma^{\delta_1-i}(b_{\delta_2-j})$. This is clearly the coefficient of X^l in the product $\tau_{\delta_1}(P)\tau_{\delta_2}(Q^{(\delta_1)})$, computed in $k[X,\sigma^{-1}]$. □

Let us now describe the Euclidean division algorithm. Let $d_1 = \deg A$ and $d_2 = \deg B$. According to the previous formula, if $A = QB + R$ is the right-Euclidean division of A by B, we have:

$$\tau_{d_1}(A) = \tau_{d_1-d_2}(Q)\tau_{d_2}(B^{(d_1-d_2)}) + \tau_{d_1}(R).$$

Since $\deg R < d_2$, $\tau_{d_1}(R)$ is divisible by $X^{d_1-d_2+1}$. The idea is to compute an approximation of the left-inverse of $\tilde{B} = \tau_m(B^{(d_1-d_2)})$ in $k[\![X, \sigma^{-1}]\!]$ (the ring of skew power series, which is defined in the obvious way, and is only used here to sketch the idea of the algorithm). Once we get such an approximation $\tilde{Q}$, truncated at precision $X^{d_1-d_2}$, we know that $\tau_{d_1}(A)\tilde{Q}\tilde{B} - \tau_{d_1}(A) \in X^{d_1-d_2}k[X, \sigma^{-1}]$, and by applying $\tau_{d_1}^{-1}$, we get the quotient Q.

Computing successive approximations of $\tilde{Q}$ is done by Newton iteration: let B_0 be the constant coefficient of $\tilde{B}$, we define $\tilde{Q}_0 = B_0^{-1}$, and $\tilde{Q}_{i+1} = 2\tilde{Q}_i - \tilde{Q}_i\tilde{B}\tilde{Q}_i$, truncated at X^{2^i}.

Lemma 3.2.2. *For all $i \geq 0$, $\tilde{Q}_i\tilde{B} - 1 \in X^{2^i}k[X, \sigma^{-1}]$.*

Proof. The proof goes by induction on i. By construction, $\tilde{Q}_0\tilde{B} - 1 \in Xk[X, \sigma^{-1}]$. Now assuming that the result is true for some $i \geq 0$, we have:

$$\tilde{Q}_{i+1}\tilde{B} - 1 = 2\tilde{Q}_i\tilde{B} - \tilde{Q}_i\tilde{B}\tilde{Q}_i\tilde{B} - 1 = -(1 - \tilde{Q}_i\tilde{B})^2 \in X^{2^{i+1}}k[X, \sigma^{-1}]$$

and we are done. □

Algorithm 5: `REuclideanDivision`

Input: $A, B \in k[X, \sigma]$ with $\deg A \geq \deg B$
Output: $Q, R \in k[X, \sigma]$ with $\deg R < \deg B$ such that $A = QB + R$

$d_1 = \deg A$; $d_2 = \deg B$;
$\tilde{B} = \tau_{d_2}(B^{(n)})$;
$\tilde{Q} = \text{Coefficient}(\tilde{B}, 0)^{-1}$;
$i = 1$;
while $i < d_1 - d_2 + 1$ **do**
  $\tilde{Q} = 2\tilde{Q} - \tilde{Q}(\tilde{B} \pmod{X^i})\tilde{Q} \pmod{X^{2i}}$;
  $i = 2i$;
$\tilde{Q} = (\tau_{d_1}(A) \pmod{X^{d_1-d_2}})\tilde{Q} \pmod{X^{d_1-d_2}}$;
$Q = \tau_{d_1-d_2}^{-1}(\tilde{Q})$;
$R = A - QB$;
return Q,R;

Proposition 3.2.3. *The algorithm* `REuclideanDivision` *returns the quotient and remainder of the right-division of A of degree d_1 by B of degree d_2 in $\tilde{O}(\mathrm{SM}(d_1, r))$ operations in k^σ.*

Proof. We have already seen that the result of this algorithm is correct. In order to compute $\tilde{B}$, $O(d_2r^2)$ operations are needed. The `while` loop in the algorithm has $\log_2(d_1 - d_2 + 1)$ steps. Moreover at the i-th step, we compute the product of skew polynomials of degree 2^i. Therefore the total complexity of this is $\sum_{i=0}^{\log_2(d_1-d_2+1)} \mathrm{SM}(2^i, r) = \tilde{O}(\mathrm{SM}(d_1 - d_2, r))$. Computing $(\tau_{d_1}(A) \pmod{X^{d_1-d_2}})\tilde{Q}$ has the same complexity. Finally, we compute the product QB in $\mathrm{SM}(\max d_2, d_1 - d_2, r)$ operations, and $R = A - QB$ in $\tilde{O}(\mathrm{SM}(d_1, r))$ operations. □

3.2.2 Greatest common divisors and lowest common multiples

This section describes an algorithm adapted directly from Algorithm 11.4 of [GG03], to compute the right-gcd R of two skew polynomials A and B, together with skew polynomials U, V such that $UA + VB = R$. As we have seen before, this also gives almost directly the left-lcm of A and B.

This algorithm relies on the fact that in the Euclidean division, the highest-degree terms of the quotient only depend on the highest-degree terms of the dividend and divisor. If $A \in k[X, \sigma]$ and $n \in \mathbf{N}$, with $A = \sum_{i=0}^{d} a_i X^i$ of degree d, we set $A_{(n)} = \sum_{i=0}^{n} a_{d-i} X^{n-i}$, with the convention that $a_j = 0$ for $j \notin \{0, \dots, d\}$. Then, for $n \geq 0$, $A_{(n)}$ is a skew polynomial of degree n, and for $n < 0$. Note that for all $i \geq 0$, $(AX^i)_{(n)} = A_{(n)}$.

Definition 3.2.4. If $A, B, A^*, B^* \in k[X, \sigma]$ with $\deg A \geq \deg B$ and $\deg A^* \geq \deg B^*$, and $\delta \in \mathbf{Z}$, we say that (A, B) and (A^*, B^*) *coincide up to* δ if

1. $A_{(\delta)} = A^*_{(\delta)}$,
2. $B_{(\delta - (\deg P - \deg Q))} = B^*_{(\delta - (\deg P^* - \deg Q^*))}$

Then we have the following:

Lemma 3.2.5 ([GG03], Lemma 11.1.)**.** *Let $n \in \mathbf{Z}$, (A, B) and $(A^*, B^*) \in (k[X, \sigma] \setminus \{0\})^2$ that coincide up to 2δ, with $\delta \geq \deg A - \deg B \geq 0$. Define Q, R, Q^*, R^* as the quotient and remainder in the right-divisions:*

$$\begin{array}{rcll} A & = & QB + R, & \text{with } \deg R < \deg B, \\ A^* & = & Q^*B^* + R^*, & \text{with } \deg R^* < \deg B^*. \end{array}$$

Then $Q = Q^*$, and either (B, R) and (B, R^*) coincide up to $2(\delta - \deg Q)$ or $R = 0$ or $\delta - \deg Q < \deg B - \deg R$.

Now, we want to carry this approximation further down the sequence of quotients when doing the Euclidean algorithm. For $A_0, A_1, A_0^*, A_1^* \in k[X, \sigma]$ monic, with $\deg A_0 > \deg A_1$ and $\deg A_0^* > \deg A_1^*$, we write:

$$\begin{array}{rclcrcl} A_0 & = & Q_1 A_1 + \rho_2 A_2, & & A_0^* & = & Q_1^* A_1^* + \rho_2^* A_2^*, \\ & \vdots & & & & \vdots & \\ A_{i-1} & = & Q_i A_i + \rho_{i+1} A_{i+1}, & & A_{i-1}^* & = & Q_i^* A_i^* + \rho_{i+1}^* A_{i+1}^*, \\ & \vdots & & & & \vdots & \\ A_{\ell-1} & = & Q_\ell A_\ell, & & A_{\ell^*-1}^* & = & Q_{\ell^*}^* A_{\ell^*}^*, \end{array}$$

with for all i, $\deg A_{i+1} < \deg A_i$, with $\rho_i \in k^\times$ and A_i monic. From this sequence, we define for $1 \leq i \leq \ell$, $m_i = \deg Q_i$, $d_i = \deg A_i$, and for $\delta \in \mathbf{N}$,

$$\eta(\delta) = \max \left\{ 0 \leq j \leq \ell \mid \sum_{1 \leq i \leq j} m_i \leq \delta \right\}.$$

We define analogously m_i^*, d_i^* and η^*. Then the following lemma quantifies how much the first results in the Euclidean algorithm only depend on the highest-power terms of the entires;

Lemma 3.2.6 ([GG03], Lemma 11.3.). *Let $\delta \in \mathbf{N}$, $h = \eta(\delta)$ and $h^* = \eta^*(\delta)$. If (A_0, A_1) and (A_0^*, A_1^*) coincide up to 2δ, then $h = h^*$, $Q_i = Q_i^*$ and $\rho_{i+1} = \rho_{i+1}^*$ for $1 \leq i \leq h$.*

Let us now describe the extended Euclidean Algorithm.

Algorithm 6: `FastExtendedRGCD`

Input: $A_0, A_1 \in k[X, \sigma]$ monic, $d_0 = \deg A_0 \geq \deg A_1 = d_1$ and $d \in \mathbf{N}$ with $0 \leq d \leq d_0$.

Output: $M \in \mathcal{M}_2(k[X, \sigma])$ such that $M \begin{pmatrix} A_0 \\ A_1 \end{pmatrix} = \begin{pmatrix} A_h \\ A_{h+1} \end{pmatrix}$ with $h = \eta(d)$.

1 **if** $A_1 = 0$ *or* $d < d_0 - d_1$ **then return** $0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;
2 $\delta = \lfloor d/2 \rfloor$;
3 $R = \texttt{FastExtendedRGCD}(R_{0(2\delta)}, R_{1(2\delta-(d_0-d_1))}, 2\delta, 2\delta - (d_0 - d_1), \delta)$;
4 $\begin{pmatrix} A_0' \\ A_1' \end{pmatrix} = R \begin{pmatrix} A_0 \\ A_1 \end{pmatrix}$; $\begin{pmatrix} d_0' \\ d_1' \end{pmatrix} = \begin{pmatrix} \deg A_0' \\ \deg A_1' \end{pmatrix}$;
5 **if** $A_1' = 0$ *or* $d < d_0 - d_j$ **then return** R;
6 $Q_j = A_0'/A_1'$; $\rho_2' = \text{LeadingCoefficient}(A_0' \bmod A_1')$;
7 $A_2' = (\rho_2')^{-1}(A_0' \bmod A_1')$; $d_2' = \deg A_2'$;
8 $\delta^* = d - (d_0 - d_1')$;
9 $S = \texttt{FastExtendedRGCD}(A_1', A_2', 2d^*, 2\delta^* - (d_1' - d_2'), \delta^*)$;
10 $M_j = \begin{pmatrix} 0 & 1 \\ (\rho_2')^{-1} & (\rho_2')^{-1} Q_j \end{pmatrix}$;
11 **return** $S \cdot M_j \cdot R$;

When executed for $d = d_0$, the above algorithm gives an immediate way to compute the right-gcd and left-lcm of A_0 and A_1. Indeed, in this case, we get a matrix $M = \begin{pmatrix} U_0 & U_1 \\ V_0 & V_1 \end{pmatrix}$ such that $U_0 A_0 + U_1 A_1 = \text{rgcd}(A_0, A_1)$, and $V_0 A_0 = -V_1 A_1 = \text{llcm}(A_0, A_1)$.

Theorem 3.2.7 ([GG03], Theorem 11.5.). *The algorithm* `FastExtendedRGCD` *works correctly and uses at most $O(\text{SM}(\delta, r) \log n)$ operations in k^σ if A_0 has degree $d \leq 2\delta$. In particular, it allows to compute the rgcd and llcm with $\tilde{O}(\text{SM}(d, r))$ operations in k^σ.*

Proof. The proof for correctness is exactly the same as the one in [GG03] and relies on the previous two lemmas. Let us give more details about the complexity of the algorithm. Denote by $T(d_0, d_1, \delta)$ the time needed to call `FastExtendedRGCD` on two skew polynomials A_0, A_1 of degrees d_0, d_1, with parameter d. Set $\delta = \lfloor d_0/2 \rfloor$. Then we have:

$$T(d_0, d_1, \delta) \leq T(2\delta, 2\delta - (d_0 - d_1), \delta) + T(2\delta^*, 2\delta^* - (d_j - d_{j-1}), \delta^*) + O(\text{SM}(d_0, r)).$$

The term $\text{SM}(d_0, r)$ here comes from the multiplications needed from matrix multiplications (all the polynomials in these matrices have degree at most d_0) and one due to the Euclidean division algorithm. The result follows by induction from the fact that $\delta^* = \lceil d_0/2 \rceil$. □

4 Algorithm for factorization in skew polynomial rings

We now enter into the main contribution of this paper: the factorization algorithm. As preliminaries, we first describe efficient algorithms for computing the reduced norm of a skew polynomial as defined in the theoretical part. Using them, together with still other theoretical results, we then design our factorization algorithm. We include a detailed analysis of its complexity.

4.1 Computing the reduced norm

In this section, we address the question to compute explicitely the reduced norm of a given skew polynomial of degree d. We shall prove the following Theorem.

Theorem 4.1.1. *There exists a probabilistic algorithm that computes the reduced norm of a skew polynomial P of degree d with complexity $\tilde{O}(dr^2 \min(d,r))$ operations in k^σ on average.*

To prove this Theorem, we shall describe two different algorithms to compute the reduced norm of P, the first one (resp. the second one) having complexity $\tilde{O}(d\mathrm{MM}(r))$ (resp. $\tilde{O}(\mathrm{MM}(d)r + d\mathrm{SM}(d,r))$ on average. Using the first algorithm if $d > r$ and the second one otherwise, and taking that $\mathrm{SM}(d,r) = dr^2$ (achieved by the "reduction to the commutative case" method) and $\mathrm{MM}(n) = O(n^3)$, we get the announced complexity in Theorem 4.1.1.

Remark 4.1.2. If we use the "matrix" method instead of the "reduction to the commutative case" method for multiplying skew polynomial, we may get a better complexity in Theorem 4.1.1 under some hypothesis ($\log q$ small essentialy) whose order of magnitude is about:

$$\tilde{O}\Big(dr\,\frac{\mathrm{MM}(\min(d,r))}{\min(d,r)}\Big)$$

operations in k^σ.

First algorithm

We use the fact that $\mathcal{N}(P)$ is the determinant of multiplication by X^r on D_P, seen as a $k[X^r]$-module. Let $t \in k$ be a primitive element over k^σ, and let $\pi_t \in k^\sigma[X^r]$ be its minimal polynomial over k^σ. Let $R_0 \in k^\sigma[X^r]$ be a polynomial of degree $n > d/r$. Let R be the polynomial obtained by composition: $R = \pi_t \circ R_0$. We work in the ring $\mathcal{A} = k^\sigma[X^r]/R$.

The idea is the following: if R is irreducible, then $\mathcal{A}$ is a field extension of k^σ, and there is a natural embedding of k into $\mathcal{A}$, mapping t to R_0. Then we can write the matrix of multiplication by P in $k[X,\sigma]$ seen as a module over $k[X^r]$, and map it to a matrix with coefficients in $\mathcal{A}$. Then we can compute the determinant of this matrix, which is the image ν of the norm of P by the map $k[X^r] \to \mathcal{A}$. Since it is known to be a polynomial with coefficients in k^σ of degree d, and since $[\mathcal{A} : k^\sigma] > d$, the coefficients of the $\mathcal{N}(P)$ are exactly the coefficients of ν written in the canonical basis of $\mathcal{A}$.

Actually, all of the above still holds if $\mathcal{A}$ is not a field, except that we may not use algorithms for determinants over fields to compute ν. However, we can still obtain this determinant efficiently by computing the Hermite normal form of the matrix of multiplication by P in the Euclidean domain $\mathcal{A}$. So in practice, all we have to do is write the matrix of multiplication by P as a matrix with coefficients in $k[X^r]$. Write $P = P_0 + P_1 X + \cdots + P_{r-1}X^{r-1}$. As stated in the proof of Lemma 2.1.15, in the canonical basis $1, X, \ldots, X^{r-1}$, the matrix of multiplication by P is:

$$\begin{pmatrix} P_0 & X^r\sigma(P_{r-1}) & \cdots & \cdots & X^r\sigma^{r-1}(P_1) \\ P_1 & \sigma(P_0) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & X^r\sigma^{r-1}(P_{r-1}) \\ P_{r-1} & \cdots & \cdots & \cdots & \sigma^{r-1}(P_0) \end{pmatrix}.$$

We map this matrix to $\mathcal{A}$ by taking X^r to its residue class modulo R, and t to $R_0 \pmod R$. Then, we compute its determinant ν (using Smith normal form), and we can read the coefficients of $\mathcal{N}(P)$ on ν.

If P has degree d, the complexity of these operations is determined as follows. We have $\tilde{O}(dr^2)$ operations to compute all the conjugates of the P_i's under the action of the Frobenius. Multiplication by X^r is free in $k[X^r]$. This yields a total of $\tilde{O}(dr^2)$ operations to compute the matrix, and then $\tilde{O}(d\mathrm{MM}(r))$ operations in k^σ to get its determinant. Hence, P can be computed in $\tilde{O}(d\mathrm{MM}(r))$ operations in k^σ.

Second algorithm

The second algorithm we would like to present relies on the characterization of the reduced norm given by Lemma 2.1.15. Given a skew polynomial P of degree d, the aforementioned Lemma reduces the computation of $\mathcal{N}(P)$ to the computation of the caracteristic polynomial of the application $m : R \mapsto X^r R$ acting on the quotient $k[X,\sigma]/k[X,\sigma]P$. To achieve this computation, we will write down the matrix M of m in the standard basis $(1, X, \ldots, X^{d-1})$ of $k[X,\sigma]/k[X,\sigma]P$. We note that the entries on the j-th column of M are exactly the coefficients of the remainder in the right Euclidean division of X^{r-1+j} by P. We are then reduced to compute these remainders when j varies between 1 and d. We now explain how to do this efficiently (using a noncommutative version of the fast modular exponentiation).

We recall that, given a skew polynomial R and an integer n, we have defined the skew polynomial $R^{(n)}$ obtained from R by applying σ^n to its coefficients. We remark that the relation $X^n R = R^{(n)} X^n$ holds and that $R^{(n)}$ can be computed in $\tilde{O}(\deg R \cdot r^2)$ operations in k^σ. In the following, we shall denote by $R \bmod P$ the remainder in the right Euclidean division of R by P. We furthermore set $R_n = X^n \bmod P$ for all n. The following Lemma is the key to perform fast exponentiations modulo P.

Lemma 4.1.3. *If n and m are two nonnegative integers, the relation:*

$$R_{n+m} = (R_n^{(-m)} R_m) \bmod P$$

holds.

Proof. By definition, there exist two skew polynomials Q_n and Q_m such that $R_n = X^n + Q_n P$ and $R_m = X^m + Q_m P$. Furthermore, in the localized ring $k[X,\sigma][\frac{1}{X}]$, the following computation makes sense:

$$\begin{aligned} R_n^{(-m)} R_m &= X^m R_n X^{-m} R_m \\ &= X^m (X^n + Q_n P) X^{-m} (X^m + Q_m P) \\ &= X^{n+m} + \big(X^n Q_m + X^m Q_n + (Q_n P)^{(-m)} Q_m\big) P. \end{aligned}$$

The Lemma follows from this. □

Lemma 4.1.3 yields Algorithm 7 below whose complexity is $O(\mathrm{SM}(d,r)\ \log n)$ if P has degree d. Therefore we can compute the whole matrix M with complexity $\tilde{O}(d\,\mathrm{SM}(d,r))$. Finally, keeping in mind that M is a $d \times d$ matrix with coefficients in the field k, we find that one can compute its characteristic polynomial in average time $\tilde{O}(\mathrm{MM}(d)r)$. The total complexity of our algorithm is then $\tilde{O}(d\,\mathrm{SM}(d,r) + \mathrm{MM}(d)r)$.

Algorithm 7: `FastModularExponentiation`

Input: an integer n, a skew polynomial P
Output: the remainder of the right Euclidean division of X^n by P

$m =$ integer part of $n/2$;
$R_m =$ `FastModularExponentiation`(m, P);
$R_{2m} = (R_m^{(-m)} R_m) \bmod P$;
if *n is even* **then**
  return R_{2m};
else
  return $(X R_{2m}) \bmod P$;

4.2 A fast factorization algorithm

Let $P \in k[X, \sigma]$ be a monic polynomial. Our aim is to give an algorithm for computing a factorization of P as a product of irreducible skew polynomials. The idea of the algorithm is to reduce this problem to that of factoring polynomials of type (e) (using rgcd's with factors of the norm of P) and then to factor polynomials of type (e). (See Definition 2.2.3 as a reminder of the definition of the type.) For the sake of brevity, in the algorithms we will use the notation A/B for the quotient of the right-division of A by B.

Reduction to the type-(e) case Given a skew polynomial P, one can factor it as a product of polynomials of type (e) by taking successive right gcds with the irreducible factors of the norm. More precisely, suppose that the reduced norm of P factors as follows:

$$\mathcal{N}(P) = N_1 \cdot N_2 \cdots N_n$$

where the N_i's are irreducible in the centre. Then $D = \mathrm{rgcd}(N_m, P)$ is a right factor of P having type (e) for some e. We now perform the division of P by D, *i.e.* we consider the unique skew polynomial P' such that $P = P'D$. The degree of P' is strictly less than that of P and we now continue our computation by replacing P by P'. Such a factorization is called a *type* (e) *factorization* of P.

Example 4.2.1. A type (e) factorization of our canonical example

$$F(X) = X^5 + X^4 + \alpha X^3 + \alpha^6 X^2 + \alpha^3 X + \alpha^2 \in \mathbf{F}_8[X, \sigma] \quad \text{with } \sigma : t \mapsto t^2$$

(see Example 1.1.3 for instance) is obtained as follows. Recall from Example 2.2.5 that the reduced norm of P can be written as $\mathcal{N}(F) = N_1^3 N_2$ with $N_1 = (X^3)+1$ and $N_2 = (X^3)^2 + (X^3) + 1$. The right factor of a type (e) factorization of $F(X)$ is given by $\mathrm{rgcd}(F, N_2) = X^2 + \alpha^5 X + \alpha^3$, which has type (1). The resulting factorization is:

$$F(X) = \left(X^3 + \alpha^4 X^2 + \alpha^2 X + \alpha^3\right) \cdot \left(X^2 + \alpha^5 X + \alpha^6\right).$$

Note that this factorization is not the final one because the first factor has type $(2, 1)$. We therefore continue with it. Its reduced norm is N_1^3. So, we have to compute its rgcd with N_1. We get this way the following refined factorization of $F(X)$:

$$F(X) = \left(X + \alpha^6\right) \cdot \left(X^2 + \alpha^5 X + \alpha^4\right) \cdot \left(X^2 + \alpha^5 X + \alpha^6\right)$$

which is a complete type (e) factorization. Indeed, on the one hand we have already seen that the last factor has type (2) and on the other hand, one can directly check that the first (resp. the second) factor has norm N_1 (resp. N_2) and thus has type (1).

We note that a type (e) factorization is not unique since it strongly depends on the order we have chosen on the N_i's. We present below a recursive algorithm (Algorithm 8) for computing a type (e) factorization of an input skew polynomial P. The recursivity allows us to work with skew polynomials having balanced degrees and leads to an algorithm which has a better cost than the naive one.

Algorithm 8: `Type_e_Factorization`

Input: $P \in k[X,\sigma]$, $(N_1,\dots,N_n)$ irreducible such that $\mathcal{N}(P) = \prod N_i$, ordered by nondecreasing degree

Output: $P_{1,1}, P_{1,2}, \dots, P_{1,m_1}, \dots, P_{n,1}, \dots, P_{n,m_n} \in k[X,\sigma]$ and $N_1,\dots,N_n \in k^\sigma[X^r]$ irreducible such that $P = \prod_i \prod_j P_{i,j}$ and each $P_{i,j}$ has type $(e_{i,j})$ and norm $N_i^{e_{i,j}}$

```
d_1 = deg N_1;
for 1 ≤ i ≤ m − 1 do d_{i+1} = d_i + deg N_{i+1};
d = d_m;  δ = d/log d;
i = min{1 ≤ j ≤ m − 1 | d_j > d + δ/2};
if [d − δ/2, d + δ/2] ∩ {d_1, ..., d_{m−1}} = ∅ then
    j = m;
    while j ≥ i do
        P_j = rgcd(P, N_j);
        j = j − deg P_j / deg N_j;
        P = P/P_j;
    return Type_e_Factorization(P, (N_1, ..., N_{i−1})), {P_j | i ≤ j ≤ m};
else
    M = N_i ··· N_m;
    Q_1 = rgcd(P, M);  Q_2 = P/M;
    return Type_e_Factorization(Q_2, (N_1, ..., N_i)),
    Type_e_Factorization(Q_1, (N_i, ..., N_m));
```

Factoring a polynomial of type (e) Let us now explain how to factor a polynomial P of type (e). Clearly, $\mathcal{N}(P) = N^e$ with $N \in k^\sigma[X^r]$ irreducible. In addition, we know that P is a divisor of N. From this, we deduce that $e \leq r$.

Let Q be a skew polynomial such that $PQ = N$. Let $R \in D_P = k[X,\sigma]/k[X,\sigma]P$ be a random element. By Proposition 2.2.7 and Corollary 2.2.9, the endormorphism m_{QR} (defined as the right-multiplication by QR on D_P) is uniformly distributed in $\mathrm{End}_\varphi(D_P)$. By Lemma 2.2.6, we know that $\mathrm{End}_\varphi(D_P)$ is isomorphic to $\mathcal{M}_e(E)$ with $E = k^\sigma[X^r]/N$. Hence, there exist $\lambda_0,\dots,\lambda_{e-1} \in E$ such that $(QR)^e = \sum_{i=0}^{e-1} \lambda_i (QR)^i \pmod P$ which, by multiplying by Q on the right, is equivalent to:

$$(QR)^e Q = \sum_{i=0}^{e-1} \lambda_i \cdot (QR)^i Q \pmod N.$$

The latter formulation is better suited for computation because reducing modulo N is simpler thanks to the fact that N is central. Now assume that $C(T) = T^e - \sum_{i=0}^{e-1} \lambda_i T^i \in E[T]$ has a root $\xi \in E$. Then $QR - \xi$ induces a noninjective linear mapping on D_P. This implies that $\mathrm{rgcd}(QR - \xi, P)$ is nontrivial. Moreover it is irreducible as soon as ξ is a simple eigenvalue of m_{QR}. We shall see in the next section (*cf* Proposition 4.4.1) that m_{QR} admits a simple eigenvalue with good probability.

Remark 4.2.2. Unless $e = 2$, there is a positive probability that m_{QR} has a *unique* eigenvalue in E which is simple in addition. If this property holds, we can simply find ξ by computing the gcd of $C(T)$ and $T^{\#E} - T$; indeed, under our assumption, this gcd is nothing but $T - \xi$.

Once we get an irreducible factor, we can proceed recursively to factor P. However, we can also use a slightly more efficient trick relying on the knowledge of an irreducible factor. Assume that we know an irreducible right factor P_1 of P, and write $P = P_2 P_1$. Let $R \in k[X, \sigma]$ and let $A = \mathrm{rgcd}(P, P_1 Q R)$. Now let $B = \mathrm{llcm}(A, P_1) = \tilde{B} P_1$. Since P is a right multiple of both P_1 and A, B is a divisor of P. Hence, $\tilde{B}$ is a divisor of P_2. In general, A and $\tilde{B}$ should have the same degree as P_1, yielding an irreducible factor of P_2. The precise probability study will appear in §4.3.4. The following two algorithms describe how to factor a polynomial P of type (e): the first one finds one irreducible factor of P, and the second one performs the "lcm trick" to factor P as a product of irreducibles given

one irreducible right factor.

Algorithm 9: `FirstFactor`

Input: $(P, N) \in k[X, \sigma] \times k^\sigma[X^r]$
such that P has type (e), $\mathcal{N}(P) = N^e$ and N is irreducible
Output: An irreducible right-divisor of P

$E = k^\sigma[X^r]/(N)$;
$R_0 = Q = N/P$;
while *true* **do**
  $R =$ `RandomElement`$(k[X, \sigma]/k[X, \sigma]P)$;
  for $0 \leq i \leq e-1$ **do** $R_{i+1} = (QRR_i) \,\%\, N$;
  Find $\lambda_0, \ldots, \lambda_{e-1} \in E$ such that $R_e \equiv \sum_{i=0}^{e-1} \lambda_i R_i \pmod N$;
  $C(T) = T^e - \sum_{i=0}^{e-1} \lambda_i T^i$;
  if *F has a simple root ξ in E* **then**
    $P_1 = \mathrm{rgcd}(P,\, QR - \xi)$;
    if $\deg_{k[X,\sigma]} P_1 = \deg_{k^\sigma[X^r]} N$ **then return** P_1;

Algorithm 10: `FactorStep`

Input: $(P, N, P_1) \in k[X, \sigma] \times k^\sigma[X^r]$
such that P has type (e), $\mathcal{N}(P) = N^e$, N is irreducible and P_1 is an irreducible right factor of P
Output: Irreducible polynomials $P_1, \ldots, P_e$ such that $P = P_e \cdots P_1$

$Q = N/P$;
for $1 \leq i \leq e-1$ **do**
  while *true* **do**
    $R =$ `RandomElement`$(k[X, \sigma]/k[X, \sigma]P)$;
    $A = \mathrm{rgcd}(P, P_1 QR)$;
    $\tilde{B} = \mathrm{llcm}(P_1, A)/P_1$;
    if $\deg B = \deg P_1$ **then**
      $P_i = \tilde{B}$;
      break;
return $P_1, \ldots, P_e$;

Glueing together the previous algorithms, we get a complete factorization algorithm. We assume that the function `Factorization` returns the factorization of a (commutative)

polynomial as a product of irreducible polynomials ordered by their degrees.

Algorithm 11: `SkewFactorization`

Input: $P \in k[X, \sigma]$
Output: A list of irreducible polynomials $(P_1, \ldots, P_m)$ such that $P = P_m \cdots P_1$

$N = \mathcal{N}(P)$;
$N_1 \cdots N_m =$ `Factorization`(N);
$(G_{1,1}, \ldots, G_{n,m_n}) =$ `Type_e_Factorization`$(P, (N_1, \ldots, N_m))$;
for $1 \leq i \leq n$ **do**
  for $1 \leq j \leq m_i$ **do**
    $F =$ `FirstFactor`$(G_{i,j}, N_i)$;
    $P_{i,j,1}, \ldots, P_{i,j,e_{ij}} =$ `FactorStep`$(G_{i,j}, N_i, F)$;
return $(P_{i,j,l})$;

Example 4.2.3. We continue Example 4.2.1. Recall that we have already computed a type (e) decomposition of $F(X)$ and found:

$$F(X) = \underbrace{\left(X + \alpha^6\right)}_{\text{type (1)}} \cdot \underbrace{\left(X^2 + \alpha^5 X + \alpha^4\right)}_{\text{type (2)}} \cdot \underbrace{\left(X^2 + \alpha^5 X + \alpha^6\right)}_{\text{type (1)}}.$$

Since the first and the last factors have type (1), they are irreducible. It then just remains to complete the factorization of the second factor $P = X^2 + \alpha^5 X + \alpha^4$.

To do this, we apply Algorithm `FirstFactor`. The norm of P is $\mathcal{N}(P) = N^2$ with $N = (X^3) + 1$. We thus have $E = \mathbf{F}_2$ (recall that E is the extension of $k^\sigma = \mathbf{F}_2$ defined by N). The skew polynomial Q defined by $PQ = N$ is then $Q = X + \alpha^3$. We have a pick a random polynomial $R \in D_P$. Suppose that we take $R \equiv X + 1 \pmod{P}$. Continuing the execution of Algorithm `FirstFactor`, we find:

$$\begin{cases} R_0 = X + \alpha^3, \\ R_1 = \alpha^6 X^2 + \alpha X + \alpha^2, \\ R_2 = \alpha^6 X^2 + \alpha X + \alpha^2. \end{cases}$$

The relation between these polynomials is then $R_2 = R_1$, *i.e.* $C(T) = T^2 - T$. This (commutative) polynomial has apparently two simple roots, namely 0 and 1. Therefore, we can choose either $\xi = 0$ or $\xi = 1$. If we take $\xi = 0$ (for example), we obtain the right divisor $\operatorname{rgcd}(A, P) = X + 1$. This leads to the factorization $P = (X + \alpha^4) \cdot (X + 1)$ and finally to the following complete factorization of $F(X)$:

$$F(X) = \left(X + \alpha^6\right) \cdot \left(X + \alpha^4\right) \cdot \left(X + 1\right) \cdot \left(X^2 + \alpha^5 X + \alpha^6\right).$$

This is the 16^{th} factorization of Figure 1. If we had preferred to choose $\xi = 1$, we would have ended up with the factorization:

$$F(X) = \left(X + \alpha^6\right) \cdot \left(X + \alpha\right) \cdot \left(X + \alpha^3\right) \cdot \left(X^2 + \alpha^5 X + \alpha^6\right)$$

which is the 18^{th} factorization of Figure 1.

4.3 Complexity

In this section, we analyze the complexity of the factorization algorithm. The complexity will be expressed in terms of the degree d of the skew polynomial that is to be factored, the degree r of k/k^σ, and the cardinal q of k^σ.

4.3.1 Complexity of the steps

Let us detail the complexity of the steps of our factorization algorithm.

Type-(e)-factorization We have the following lemma, giving the complexity of the algorithm `Type_e_Factorization`.

Lemma 4.3.1. *Let $P \in k[X,\sigma]$ and let $N_1,\dots,N_m \in k^\sigma[X^r]$ be irreducible polynomials such that P divides $N_1\cdots N_m$ in $k[X,\sigma]$. Then the algorithm* `Type_e_Factorization` *applied to P and $N_1,\dots,N_m$ returns a correct result with $\tilde{O}(dr^3)$ operations in k^σ.*

Proof. Let us prove the result by induction on d. Let $(N_1,a_1),\dots,(N_m,a_m)$ be the irreducible polynomials that are given as arguments, and $\delta_i = \deg N_i$ for $1 \le i \le m$. We assume that the N_i's are ordered so that the sequence of δ_i is nondecreasing. There are two cases to look at.

If there exists $1 \le i \le m$ and $1 \le a \le a_i$ such that

$$\sum_{j=1}^{i-1} a_j\delta_j + a\delta_i \in \left[\frac{d}{2}\left(1-\frac{1}{\log d}\right),\ \frac{d}{2}\left(1+\frac{1}{\log d}\right)\right],$$

then we choose the minimal (i,a) (for the lexicographical order) having this property. We write $N_l = N_i^a\prod_{j=1}^{i-1}N_j^{a_j}$, and $N_r = N/N_l$. Then we write $P_r = \mathrm{rgcd}(P,N_r)$, and define P_l as the quotient in the right-division of P by P_r. The algorithm is then applied to $(P_l,N_l,(N_1,a_1),\dots,(N_i,a))$ and $(P_r,N_r,(N_i,a_i-a),\dots,(N_m,a_m))$.

The number of operations needed for this is denoted by $C(d,r)$. In this case, we have:

$$C(d,r) \le 2C\left(d\left(1+\frac{1}{\log d}\right),r\right) + \tilde{O}(\mathrm{SM}(dr,r)).$$

Indeed, the operations we have to do before starting the recursive steps are: computing a product of (commutative) polynomials in $k^\sigma[X^r]$ such that the sum of their degrees is less than $d\left(1+\frac{1}{\log d}\right)$, computing the right gcd of P with a polynomial of degree less than dr, and dividing P by this gcd. The most expensive part is the computation of the gcd, and it costs $\tilde{O}(\mathrm{SM}(dr,r))$.

In the other case, there is no (i,a) such that

$$\sum_{j=1}^{i-1} a_j\delta_j + a\delta_i \in \left[\frac{d}{2}\left(1-\frac{1}{\log d}\right),\frac{d}{2}\left(1+\frac{1}{\log d}\right)\right].$$

Hence, for (i,a) such that $\sum_{j=1}^{i-1} a_j\delta_j + a\delta_i > \frac{d}{2}\left(1+\frac{1}{\log d}\right)$, we know that $\delta_i > \frac{d}{\log d}$, and there are at most $\log d$ such couples (i,a). In this case, the algorithm is to compute N_l, N_r as before, and then the successive gcd's of P the N_i's having the previous property, and apply the algorithm with the last quotient P_l and N_l.

There are at most $\log d$ rgcd's of a skew polynomial of degree at most d with skew polynomials of degree at most dr, which takes $\tilde{O}(\mathrm{SM}(dr,r))$ operations, and all the other computations are cheaper than this. Again, we have:

$$\begin{aligned} C(d,r) &\le C\left(d\left(1+\frac{1}{\log d}\right),r\right) + \tilde{O}(\mathrm{SM}(dr,r)) \\ &\le 2\cdot C\left(d\left(1+\frac{1}{\log d}\right),r\right) + \tilde{O}(\mathrm{SM}(dr,r)). \end{aligned}$$

Let us assume that the $\tilde{O}(\mathrm{SM}(dr,r))$ appearing in the above inequality is $\leq cdr^3 \log^\alpha d$ for some constants c, α (we use the fact that $\mathrm{SM}(d,r) = \tilde{O}(dr^2)$). We are going to show that there exists a constant c' such that

$$C(d,r) \leq c' dr^3 \log^{\alpha+1} d.$$

We want to have:

$$C(d,r) \leq 2c' \frac{d}{2}\left(1+\frac{1}{\log d}\right) r^3 \log^{\alpha+1}\left(\frac{d}{2}\left(1+\frac{1}{\log d}\right)\right) + cdr^3 \log^\alpha d.$$

This implies that:

$$C(d,r) \leq c' dr^3 \log^{\alpha+1} d \left(1 - \frac{\log 2}{\log d} + O\left(\frac{1}{\log^2 d}\right)\right)^{\alpha+1} + c' dr^3 \log^\alpha d + cdr^3 \log^{\alpha+1} d.$$

If we choose c' such that $c' + c - c'(\alpha+1)\log 2 + O\left(\frac{1}{\log^2 d}\right) \leq 0$ for d large enough, then induction shows that for d large enough,

$$C(d,r) \leq c' dr^3 \log^{\alpha+1} d.$$

Since it is possible to choose such a c', the proof is complete. □

Remark 4.3.2. If we use the "matrix" method instead of the "reduction to the commutative case" for multiplying skew polynomials, we may get a better complexity in Lemma 4.3.1.

FirstFactor We study here the complexity of the algorithm `FirstFactor` described in §4.2. We shall only deal with the case $r > 2$, letting let the case $r = 2$ (which is similar and simpler) to the reader.

In the following, we assume that P has type (e) and norm N^e, with $e \leq r$. The degree of N as an element of $k^\sigma[X^r]$ is denoted by δ. The degree of P then equals δe. We now detail the complexity of each invidual step of the algorithm:

1. *Compute $Q \in k[X,\sigma]$ such that $PQ = N$.* This Euclidean division can be done with complexity $\mathrm{SM}(dr,r)$. Note that this step is done only once even if the loop fails to find a divisor.

2. *Choose a random element $R \in k[X,\sigma]/k[X,\sigma]P$ and compute $RQ, \ldots, (RQ)^e$ modulo N.* This requires e multiplications of skew polynomials of degree δr plus one reduction modulo N at each step. After having remarked the reduction modulo N of a skew polynomial is equal to its reduction modulo N in the ring of usual polynomials, we see that it costs only $\tilde{O}(\delta^2 r)$ operations in k^σ. The whole cost of this step is then $O(e \cdot \mathrm{SM}(\delta r, r))$.

3. *Find a linear dependence between the $R_i = (QR)^i Q$ of the following form:*

$$\sum_{i=0}^{e} a_i (QR)^i Q \equiv 0 \pmod{N}. \tag{1}$$

   where all a_i's are in E. Even though the R_i's naturally live in a space of dimension r^2 over E, we know the first e of them are linearly dependent, and we can work in a vector space of dimension e over E by projection. Hence, the complexity of this step is $\delta \cdot \mathrm{MM}(e)$.

4. *Check whether the polynomial* $C(T) = \sum_{i=0}^{e} a_i T^i$ *(where the* a_i*'s are defined by formula* (1)*) has a root in* E. For this, it is enough to compute the gcd of F with $T^{\#E} - T$. Noting that $\#E = q^\delta$ with $q = \#k^\sigma$, we can first compute $T^{\#E}$ modulo $C(T)$ by first raising T to the q-th power modulo $C(T)$ (using classical fast exponentiation) and then performing $O(\log \delta)$ modular compositions. Using Corollary 5.2 of [KU08], this can be done in $\tilde{O}(\delta \log^2 q + e^{1+\varepsilon}(\delta \log q)^{1+o(1)})$ bit operations, for all $\varepsilon > 0$ (the first term corresponds to the fast exponentiation and the second to the modular compositions). It then remains to compute the gcd of two polynomials over E of degree $\leq e$, which can be achieved with $\tilde{O}(e\delta)$ more operations in k^σ.

5. *Compute the right gcd of* P *with a skew polynomial of degree* δr. This costs $\tilde{O}(\mathrm{SM}(\delta r, r))$ operations in k^σ. Note that this step is done only once even if the loop fails to find a divisor.

Since any operation in k^σ requires $\tilde{O}(\log q)$ bit operations, the total complexity of `FirstFactor` is on average

$$\tilde{O}\big(\mathrm{SM}(\delta r, r) \cdot e \log q + \mathrm{MM}(e) \cdot \delta \log q + \delta \log^2 q + e^{1+\varepsilon}(\delta \log q)^{1+o(1)}\big) \tag{2}$$

bit operations provided that we can prove that the main loop succeeds with a probability which is bounded from below by a positive constant which does not depend on q, r and e. We postpone the proof of this latest statement to §4.4.1 (see Proposition 4.4.1).

Using $\mathrm{SM}(n, r) = \tilde{O}(nr^2)$ (achieved by the "reduction to the commutative case" method) and $\mathrm{MM}(n) = O(n^3)$ and noting that $e \leq r$, the above complexity becomes:

$$\tilde{O}(\delta e r^3 \log q + \delta \log^2 q + e^{1+\varepsilon}(\delta \log q)^{1+o(1)})$$

bit operations for all $\varepsilon > 0$.

Remark 4.3.3. Using instead the "method" matrix of skew multiplication and taking $\mathrm{MM}(n) = O(n^\omega)$ for some exponant $\omega > 2$, one can replace r^3 by r^ω in the complexity above if we assume further that $\log q$ remains bounded (otherwise other terms involving a factor $\log^2 q$ appear).

FactorStep We recall that this algorithm computes a factorization of P (still of type (e)) knowing a factor of P. We shall prove in §4.4.2 (see Corollary 4.4.5) that the condition in the "if" statement on line 7 is true with very high probability. It follows from this that the inner "while" loop will be executed in average $O(1)$ times for each execution of the "for" loop. Keeping in mind the results of §3.2.7, we deduce that Algorithm `FactorStep` runs in average with complexity $\tilde{O}(\delta r^3)$.

4.3.2 Total complexity

Let us sum up all the previous step complexities to give the complexity of the whole factorization algorithm.

Theorem 4.3.4. *The algorithm* `SkewFactorization` *factors a skew polynomial of degree* d *with average complexity:*

$$\tilde{O}(dr^3 \log q + d \log^2 q + d^{1+\varepsilon}(\log q)^{1+o(1)}) + F(d, k^\sigma)$$

bit operations where $F(d, K)$ *denotes the complexity of the factorization of a (commutative) polynomial of degree* d *over the finite field* K.

Remark 4.3.5. Once again, if $\log q$ remains bounded, it is possible to replace r^3 by r^ω in the above complexity.

Proof. Computing the norm of $P \in k[X,\sigma]$ of degree d takes $\tilde{O}(dr^3)$ operations in k^σ. Factoring the norm $\mathcal{N}(P)$ (that has degree d as an element of $k^\sigma[X^r]$) takes by definition $F(d,k^\sigma)$ operations in k^σ. Then, the algorithm `Type_e_Factorization` runs in $\tilde{O}(dr^3)$ operations in k^σ. Let $P_1,\dots,P_m$ be the factors of P obtained after `Type_e_Factorization`. Assume that P_i has type e_i and degree $\delta_i e_i$. Then for each i, the factorization of P_i takes $\tilde{O}(\delta_i e_i r^3 \log q + \delta_i e_i \log^2 q + e^{1+\varepsilon}(\delta \log q)^{1+o(1)})$ bit operations (it uses `FirstFactor` and `FactorStep`). Hence factor P given its "type-(e)-factorization" requires $\tilde{O}(dr^3 \log q + d\log^2 q + d^{1+\varepsilon}(\log q)^{1+o(1)})$ bit operations. Putting all the steps together, we get the desired complexity. □

4.4 Probabilistic aspects

4.4.1 Probability of finding a factor

In this subsection, we study the probability that the main loop of Algorithm `FactorStep` succeeds. We have already noticed that this probability is the probability that a random square matrix in $\mathcal{M}_e(E)$ has a simple eigenvalue in E. We recall that E is the finite extension of k^σ defined by the irreducible polynomial N. It is then a finite field of cardinality q^d where d is the degree of N.

Proposition 4.4.1. *A random matrix in $\mathcal{M}_e(E)$ has a single eigenvalue in E with probability at least* 0.15.

Moreover if $e \geq 3$, this is the only eigenvalue in E with probability at least 0.15 *also.*

The cases $e \in \{1,2\}$ can be easily checked by hand and are left to the reader. We now assume that $e \geq 3$ and we are going to prove the second statement of the proposition. Let B_e be the probability that a $e \times e$ matrix with coefficients in E has 0 as a simple, unique eigenvalue in E. Since E has cardinality q^d, this is $\frac{1}{q^d}$ times the probability that a $r \times r$ matrix has a simple, unique eigenvalue in E. Denoting by A_i the probability that a $i \times i$ matrix with coefficients in E has *no* eigenvalue in E, we have:

$$q^{de^2} B_e = \#\mathbf{P}(E^e) \cdot q^{d(e-1)} q^{d(e-1)^2} \cdot A_{e-1}$$

from what we get:

$$q^d B_e = \frac{1 - \frac{1}{q^{de}}}{1 - \frac{1}{q^d}} \cdot A_{e-1}.$$

Since $q^d B_e$ is the probability we are interested in, it is enough to bound from below the quantity A_{e-1}. By [NP98], Theorems 4.1 and 4.2, we get the formula for the generating series:

$$\sum_{i=0}^{+\infty} A_i z^i = \frac{1}{1-z} G(z)$$

where $G(z) = \prod_{i=1}^{+\infty} \left(1 - \frac{z}{q^{di}}\right)^{q^d - 1}$. If we write $G(z) = \sum_i C_i z^i$, then for all $i \geq 0$, $A_i = \sum_{j=0}^{i} (-1)^j C_j$.

Lemma 4.4.2. *We have the following formulas:*

- $A_0 = C_0 = 1$,
- $C_1 = 1$,
- $C_2 = \frac{q^d}{2(q^d+1)}$.

Proof. The first two assertions follow easily from identifying the coefficients of 1 and z in the power series G. For the third formula, identifiying the coefficient of z^2 gives:

$$C_2 = \sum_{i=1}^{+\infty} \frac{(q^d-1)(q^d-2)}{2q^{2di}} + \sum_{i<j} \frac{(q^d-1)^2}{q^{d(i+j)}}.$$

The result then follows from the usual formulas for sums of geometric progressions. □

Next, remark that:

$$\sum_{i=0}^{+\infty} C_i = \prod_{i\geq 1} \left(1 + \frac{1}{q^{di}}\right)^{q^d-1} \quad \text{and} \quad \sum_{i=0}^{+\infty} (-1)^i C_i = \prod_{i\geq 1} \left(1 - \frac{1}{q^{di}}\right)^{q^d-1}.$$

Combining both expressions, we get:

$$2 \cdot \sum_{i=0}^{+\infty} C_{2i+1} = \prod_{i\geq 1} \left(1 + \frac{1}{q^{di}}\right)^{q^d-1} - \prod_{i\geq 1} \left(1 - \frac{1}{q^{di}}\right)^{q^d-1}.$$

Studying the function $x \mapsto \prod_{i\geq 1}(1+x^{-i})^{x-1} - \prod_{i\geq 1}(1-x^{-i})^{x-1}$, we find that the sum $\sum_{i=0}^{+\infty} C_{2i+1}$ is smaller than its limit when q^d goes to infinity, which is $\sinh(1)$. Now, it is clear that for all $i \geq 2$:

$$A_i \geq C_0 + C_2 - \sum_{i=0}^{+\infty} C_{2i+1} = 1 + \frac{q^d}{2(q^d+1)} - \sinh(1) > 0.324 - \frac{1}{2(q^d+1)}$$

Proposition 4.4.1 follows from this.

Remark 4.4.3. If q^d is at least 23, the probability of success is actually at least 0.3.

4.4.2 Probability of finding another factor

As usual, we assume that P is a right-divisor of $N \in k^\sigma[X^r]$ irreducible, with $\mathcal{N}(P) = N^e$ and $\deg N = \delta$. We have seen that once we know an irreducible factor of P, there is an easy way to factor it without using FirstFactor again. The following lemma makes this more precise:

Lemma 4.4.4. *Let $P = P_2P_1$ with P_1 irreducible and P_2 reducible, and let R be a random variable following the uniform distribution on $k[X,\sigma]$. Let $A = rgcd(P, P_1QR)$ and $B = llcm(A, P_1) = \tilde{B}P_1$. Then the probability that $\tilde{B}$ is an irreducible right factor of P_2 is at least $1 - \frac{1}{q^{\delta(e-1)}}$.*

Proof. We work in $k[X,\sigma]/N$. Remark then that $AQ = \mathrm{rgcd}(N, P_1QRQ)$ and that $B = \mathrm{llcm}(AQ, P_1Q)$. We see the multiplication by RQ as an endomorphism m_{RQ} of $k[X,\sigma]Q/N$. Since R follows the uniform distribution, so does m_{RQ}. Remark that $m_{RQ}(k[X,\sigma]P_1Q/N)$ is a sub-φ-module of $k[X,\sigma]/Q$. It is actually equal to $k[X,\sigma]AQ/N$.

Indeed, $k[X,\sigma]P_1QRQ \subset k[X,\sigma]AQ$, and $AQ \in k[X,\sigma]P_1QRQ/N$ by definition. Then, we remark that the projection along $k[X,\sigma]P_2$ onto $k[X,\sigma]P_1Q/N$ maps the sub-φ-module $UQk[X,\sigma]/N$ to $\mathrm{llcm}(U,P_1)Qk[X,\sigma]/N$. In particular, $BQk[X,\sigma]/N$ is the projection of $m_{RQ}(P_1Qk[X,\sigma]/N)$ onto $k[X,\sigma]P_1Q/N$. Therefore, $\tilde{B}$ is an irreducible right-factor of P_2 unless $m_{RQ}(k[X,\sigma]P_1Q/N) = k[X,\sigma]P_1Q/N$. Since m_{RQ} is uniformly distributed in the endomorphisms of D_P and $k[X,\sigma]P_1Q/N$ has cardinal $q^{de(e-1)}$ while D_P has cardinal q^{de^2}, this happens with probability $\frac{1}{q^{d(e-1)}}$. □

Corollary 4.4.5. *The "if" statement on line 7 of Algorithm 10 is true with probability at least* $1 - \frac{1}{q^{\delta(e-1)}}$.

5 Implementation and performance

The factorization algorithm presented in this article has been implemented in SAGE and MAGMA. The source code is available from this page:

http://cethop.math.cnrs.fr/prodscient/algos.html

Thanks to the `sage` notebook, online demonstrations are also available:

https://cethop.math.cnrs.fr:8443/pub/

The first table (resp. the second table) presented in Figure 2 (page 38) gives running times for random skew polynomials (resp. random skew polynomials *lying in the center*) of various degrees over various base rings. We first observed that the behaviours depends a lot on the input distribution we consider. Indeed when working with random central skew polynomials, on the one hand the computation of the reduced norm is simplified (because $\mathcal{N}(P)$ is just P^r when P is central) but, on the other hand the type of any irreducible divisor of the norm is always $(r,\dots,r)$, so that we always need to get into the part *Factoring a polynomial of type (e)* which is undoubtedly the most intricated.

Actually, the timings we obtained do not quite reflect the predicted complexity. This is due to a subtle combination of many different reasons, among them we can cite:

- Due to the probabilistic nature of our algorithm, the running time may vary a lot (by a factor 3, say) under the *same* input. In the same spirit, it may happen that certain skew polynomials are difficult to factorize whereas others are easier (because for instance their reduced norm is separable).

- We have implemented Karatsuba multiplication because we observed that it was often faster than others in the considered range[4]. Nevertheless, its theoretical complexity is already complicated (namely $\tilde{O}(d^{1.58}r^{1.41})$) and our implementation involves arbitrary thresholds (which probably explains the bad timings for $d = 54$ in the first table).

- We are using FLINT for factorizing commutative polynomials; this part is then highly optimized and, although it is theoretically the most expansive, it runs very fast in practice and is definitely *not* the bottleneck.

[4]This is probably due to some artefacts (*e.g.* costly conversions between SAGE and PARI objects) coming from the framework of SAGE.

Base ring	d				
	6	**18**	**54**	**162**	**486**
$\mathbf{F}_{2^3}$	7 ms	11 ms	30 ms	87 ms	373 ms
$\mathbf{F}_{2^7}$	7 ms	123 ms	679 ms	804 ms	2810 ms
$\mathbf{F}_{2^{11}}$	7 ms	942 ms	6229 ms	5833 ms	18525 ms
$\mathbf{F}_{2^{15}}$	10 ms	4934 ms	29277 ms	26724 ms	86494 ms
$\mathbf{F}_{3^3}$	5 ms	10 ms	30 ms	94 ms	522 ms
$\mathbf{F}_{3^7}$	4 ms	133 ms	708 ms	877 ms	3322 ms
$\mathbf{F}_{3^{10}}$	5 ms	658 ms	4074 ms	4014 ms	12853 ms
$\mathbf{F}_{5^3}$	5 ms	9 ms	30 ms	97 ms	606 ms
$\mathbf{F}_{5^6}$	3 ms	67 ms	379 ms	525 ms	2189 ms
$\mathbf{F}_{7^3}$	4 ms	9 ms	30 ms	99 ms	656 ms
$\mathbf{F}_{7^5}$	12 ms	37 ms	173 ms	317 ms	1525 ms

Input distribution: uniform among all skew polynomials of degree d

Base ring	d				
	6	**18**	**54**	**162**	**486**
$\mathbf{F}_{2^3}$	35 ms	111 ms	267 ms	693 ms	2069 ms
$\mathbf{F}_{2^7}$	141 ms	396 ms	1057 ms	2364 ms	5899 ms
$\mathbf{F}_{2^{11}}$	413 ms	991 ms	2351 ms	6624 ms	15309 ms
$\mathbf{F}_{2^{15}}$	1200 ms	2662 ms	5686 ms	13142 ms	32552 ms
$\mathbf{F}_{3^3}$	45 ms	92 ms	232 ms	626 ms	2233 ms
$\mathbf{F}_{3^7}$	144 ms	425 ms	816 ms	2274 ms	6605 ms
$\mathbf{F}_{3^{10}}$	387 ms	694 ms	1611 ms	3852 ms	11400 ms
$\mathbf{F}_{5^3}$	52 ms	117 ms	307 ms	755 ms	2681 ms
$\mathbf{F}_{5^6}$	129 ms	316 ms	663 ms	2004 ms	6222 ms
$\mathbf{F}_{7^3}$	67 ms	162 ms	329 ms	907 ms	3094 ms
$\mathbf{F}_{7^5}$	169 ms	310 ms	650 ms	1895 ms	6764 ms

Input distribution: uniform among central polynomials of degree $r\lceil\frac{d}{r}\rceil$

These benchmarks were obtained with our SAGE implementation on an
AMD Opteron 6272 machine with 4 cores at 2GHz and 8GB RAM, running Linux.

Figure 2: Average running times for the factorization algorithm

References

[AF92] Frank W. Anderson and Kent R. Fuller, *Rings and categories of modules*, second ed., Graduate Texts in Mathematics, vol. 13, Springer-Verlag, New York, 1992. MR 1245487 (94i:16001)

[Azu51] Gorô Azumaya, *On maximally central algebras*, Nagoya Math. J. **2** (1951), 119–150. MR 0040287 (12,669g)

[BCS97] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi, *Algebraic complexity theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 315, Springer-Verlag, Berlin, 1997, With the collaboration of Thomas Lickteig. MR 1440179 (99c:68002)

[BCS14] Alin Bostan, Xavier Caruso, and Éric Schost, *A fast algorithm for computing the characteristic polynomial of the p-curvature*, ISSAC'14, ACM, New York, 2014, pp. 59–66.

[Gab85] È. M. Gabidulin, *Theory of codes with maximum rank distance*, Problemy Peredachi Informatsii **21** (1985), no. 1, 3–16. MR 791529 (87f:94036)

[GG03] Joachim Von Zur Gathen and Jurgen Gerhard, *Modern computer algebra*, 2 ed., Cambridge University Press, New York, NY, USA, 2003.

[Gie98] Mark Giesbrecht, *Factoring in skew-polynomial rings over finite fields*, J. Symbolic Comput. **26** (1998), no. 4, 463–486. MR 1646671 (99i:16053)

[Gro95] Alexander Grothendieck, *Le groupe de Brauer. I. Algèbres d'Azumaya et interprétations diverses*, Séminaire Bourbaki, Vol. 9, Soc. Math. France, Paris, 1995, pp. Exp. No. 290, 199–219. MR 1608798

[Ike81] Shûichi Ikehata, *Azumaya algebras and skew polynomial rings*, Math. J. Okayama Univ. **23** (1981), no. 1, 19–32. MR 620719 (82j:16013)

[Ike84] ———, *Azumaya algebras and skew polynomial rings. II*, Math. J. Okayama Univ. **26** (1984), 49–57. MR 779774 (86e:16001)

[Jac96] Nathan Jacobson, *Finite-dimensional division algebras over fields*, Springer-Verlag, Berlin, 1996. MR 1439248 (98a:16024)

[Kat73] Nicholas M. Katz, *p-adic properties of modular schemes and modular forms*, Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), Springer, Berlin, 1973, pp. 69–190. Lecture Notes in Mathematics, Vol. 350. MR 0447119 (56 #5434)

[KS98] Erich Kaltofen and Victor Shoup, *Subquadratic-time factoring of polynomials over finite fields*, Math. Comp. **67** (1998), no. 223, 1179–1197. MR 1459389 (99m:68097)

[KU08] Kiran S. Kedlaya and Christopher Umans, *Fast modular composition in any characteristic*, Foundations of Computer Science, IEEE Annual Symposium on **0** (2008), 146–155.

[LG14] François Le Gall, *Powers of tensors and fast matrix multiplication*, ISSAC 2014—Proceedings of the 2014 International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2014, pp. 296–303.

[NP98] Peter M. Neumann and Cheryl E. Praeger, *Derangements and eigenvalue-free elements in finite classical groups*, J. London Math. Soc. (2) **58** (1998), no. 3, 564–586. MR 1678151 (2000a:20153)

[Ore33] Oystein Ore, *Theory of non-commutative polynomials*, Ann. of Math. (2) **34** (1933), no. 3, 480–508. MR 1503119

[Rev73] Philippe Revoy, *Algèbres de Weyl en caractéristique p*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A225–A228.

[vzGGZ10] Joachim von zur Gathen, Mark Giesbrecht, and Konstantin Ziegler, *Composition collisions and projective polynomials*, ISSAC 2010—Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2010, pp. 123–130. MR 2920545