Analyse de la sécurité de primitives symétriques dédiées à diverses techniques de preuves

Clémence Bouvier

UNIVERSITÉ DE RENNES 1

RAPPORT DE STAGE

Master Mathématiques de l'Information, Cryptographie

Analyse de la sécurité de primitives symétriques dédiées à diverses techniques de preuves

Clémence BOUVIER
INRIA, Paris 12e, 2 rue Simone Iff
Supervisé par : Anne Canteaut - Léo Perrin

Mars - Juillet 2020

Inria

Remerciements

Je tiens à remercier toutes les personnes ayant contribué à la réussite de ce stage.

Tout d'abord, mes premiers remerciements sont pour Anne Canteaut et Léo Perrin, mes maîtres de stage, pour m'avoir fait confiance et avoir accepté de m'encadrer pendant ces 5 mois de stage.

Je leur suis reconnaissante pour le temps qu'ils m'ont consacré tout au long de cette période. Ils ont su rester disponibles malgré les circonstances particulières dues au confinement.

J'ai également apprécié la liberté et l'autonomie qu'ils m'ont offertes pendant ce stage, cela m'a permis d'orienter le sujet selon mes souhaits.

Enfin, je les remercie pour leur relecture attentive de ce rapport.

Par ailleurs, je souhaite remercier tout particulièrement Delphine Boucher pour son écoute et ses précieux conseils sur mes candidatures, sans lesquels je n'aurais pu trouver ce stage. Son soutien et ses encouragements tout au long de ce master m'ont également permis d'aborder plus sereinement ce stage.

Je remercie aussi Sylvain Duquesne pour m'avoir conseillée lors de ma recherche de stage et pour avoir appuyé ma candidature à l'INRIA.

Table des matières

Table des figures

Liste des tableaux

Chapitre 1

Introduction

1.1 Cryptographie symétrique

La cryptographie permet à deux protagonistes (traditionnellement appelés Alice et Bob) de pouvoir communiquer de façon sûre en assurant la confidentialité, l'authenticité, ainsi que l'intégrité des messages échangés, en présence d'un adversaire. On distingue alors la cryptographie asymétrique de la cryptographie symétrique. Les chiffrements symétriques, ou chiffrements à clé secrète, utilisent une clé partagée par les deux protagonistes. Cette dernière leur permet à la fois de chiffrer et déchiffrer leurs messages.

Les chiffrements symétriques sont répartis en deux catégories : les chiffrements à flot et les chiffrements par blocs. Dans ce rapport, nous nous intéresserons plus particulièrement aux chiffrements par blocs. Ces derniers prennent en entrée un bloc de n bits ainsi qu'une clé de k bits et retournent un message chiffré de même taille que le message clair en entrée. Le plus utilisé des chiffrements par bloc est l'AES (Advanced Encryption Standard) [DR02], qui permet de chiffrer des blocs de 128 bits avec des clés de taille 128, 192 ou 256 bits.

La sécurité des systèmes asymétriques repose sur la difficulté de problèmes mathématiques tels que la factorisation, ou le logarithme discret par exemple, pour lesquels nous ne connaissons pas d'algorithmes de résolution en temps polynomial. En cryptographie symétrique, la cryptanalyse est l'élément essentiel permettant d'évaluer la sécurité d'une primitive. Son rôle est effectivement d'analyser les possibles faiblesses introduites par les nouveaux schémas de chiffrement, afin de s'assurer qu'il n'y a pas d'attaques possibles.

Dans le cas des chiffrements par blocs, on souhaite ainsi vérifier que le nombre de tours et la taille des blocs et de la clef ont été bien choisis par les concepteurs. Plus précisément, un chiffrement par bloc dont la clé a été tirée aléatoirement est dit "sûr" s'il est indistinguable d'une permutation aléatoire. En effet, on peut également voir un chiffrement par bloc comme une famille de 2^k permutations de n bits, et le choix d'une clé correspond donc au choix d'une permutation dans cette famille. Or une permutation aléatoire étant tirée au hasard, il n'y a pas de lien entre l'entrée et la sortie, et le message chiffré ne révèle donc aucune information sur le message clair.

1.2 Usages émergents en cryptographie symétrique

Récemment, de nouvelles primitives de cryptographie symétrique ont été proposées pour être utilisées dans certains protocoles cryptographiques comme le calcul multi-partite, ou en combinaison avec un chiffrement homomorphe ou encore dans divers systèmes de preuve à apport nul de connaissance. Ces protocoles s'inscrivent dans un contexte marqué par le développement du Cloud et des technologies de type Blockchain et doivent ainsi répondre à une préoccupation croissante des utilisateurs en matière de sécurité.

La Blockchain est une technologie numérique qui permet théoriquement de transmettre des données de manière décentralisée, sécurisée, transparente et sans intermédiaire. Cette technologie est notamment utilisée dans les crypto-monnaies comme Bitcoin ou Ethereum. Mais au-delà du secteur de la finance, la Blockchain permet aussi la décentralisation de fonctions sociales, comme c'est par exemple le cas pour les contrats légaux remplacés par des "contrats intelligents".

Afin d'assurer la sécurité de ces chaînes de blocs, de nouveaux systèmes de preuve d'intégrité ont alors fait leur apparition, comme le protocole zk-SNARK développé dans la crypto-monnaie Zcash ou zk-STARK en cours de déploiement dans la blockchain Ethereum. Les SNARKs (Succinct Non-interactive ARgument of Knowledge) [BCG+13] et STARKs (Succinct Transparent ARgument of Knowledge) [BBHR18] sont des

primitives cryptographiques permettant la vérification de l'intégrité des calculs. Dans un modèle client-serveur, lorsqu'un client ayant une faible puissance de calcul délègue une tâche à un serveur avec une forte puissance de calcul, elles permettent ainsi au client de vérifier efficacement si le serveur a bien exécuté la tâche demandée. Ces preuves sont de taille réduite et peuvent être vérifiées rapidement. Dans l'acronyme SNARK, "Non-interactif" signifie qu'il n'y a peu ou pas d'interaction entre les deux parties, et dans l'acronyme STARK, "Transparent" signifie que le système ne requiert pas de configuration initiale de confiance.

Les systèmes SNARKs et STARKs peuvent être équipés d'une propriété de connaissance zéro, on parle alors de zk-SNARKs et zk-STARKs. Les preuves à divulgation nulle de connaissance permettent à "un prouveur" de prouver à un autre individu qu'une déclaration est vraie, sans divulguer aucune autre information que la validité de la déclaration, l'objectif étant de révéler le moins de données possible entre les deux parties.

La complexité de ces systèmes de preuve est régie par différents paramètres tels que la profondeur et la largeur du circuit (c'est-à-dire le nombre de portes dans chaque "niveau" du circuit), mais également la complexité multiplicative, correspondant au nombre de portes de multiplication (la majorité des systèmes de preuve ne sont pas impactés par le nombre de portes d'additions, car cette opération est jugée moins coûteuse). Les preuves à divulgation nulle de connaissance utilisent notamment les circuits arithmétiques, l'idée étant de traduire une affectation de circuit valide en une propriété algébrique de polynômes, en utilisant des programmes arithmétiques quadratiques. Autrement dit, l'arithmétisation consiste à réduire les problèmes de calcul aux problèmes algébriques, impliquant des polynômes "de bas degré" sur un corps fini.

Ces protocoles ont ainsi mis en avant le besoin d'optimiser la complexité multiplicative du circuit décrivant le chiffrement ou la fonction de hachage, c'est-à-dire de minimiser le nombre de multiplications effectuées par la primitive dans des corps finis $\mathbb{F}_q$ de plus grande taille, tel $\mathbb{F}_{2^n}$ où n est de l'ordre de 128, ou des corps premiers. Les algorithmes symétriques classiques sont inappropriés dans ce contexte et les nouveaux protocoles cryptographiques doivent alors être combinés avec des primitives symétriques (chiffrement ou fonction de hachage) ayant des propriétés particulières. Afin de réduire le nombre de multiplications, les constructions proposées utilisent des fonctions non-linéaires, mais dont la représentation algébrique reste très simple sur un corps fini de grande taille, comme un polynôme creux de $\mathbb{F}_q[X]$. L'utilisation de transformations ayant une représentation simple sous forme univariée est une des contraintes principales déterminant la conception de ces primitives, et il est alors nécessaire d'analyser les faiblesses introduites par la simplicité de cette structure.

1.3 Le chiffrement par bloc MiMC

Le but du stage était plus précisément d'établir une première évaluation de la sécurité du chiffrement MiMC proposé en 2016 [AGR+16]. Il s'agit d'un cas extrême reposant sur l'application, un grand nombre de fois, de la fonction $x \mapsto x^3$ sur un corps fini de taille 2^m, où m est de l'ordre de 128, et de l'addition d'une clé secrète.

MiMC-n/n est un système de chiffrement construit avec des blocs de taille n et une clé k de taille n. Le nombre d'itérations suggéré est de $r = \lceil n \log_3 2 \rceil$ et chaque tour pour le chiffrement consiste à appliquer la fonction suivante $F_i : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}, x \mapsto F(x \oplus k \oplus c_i) = (x \oplus k \oplus c_i)^3$, où les $c_i \in \mathbb{F}_{2^n}$ sont des constantes aléatoires sauf $c_0 = c_r = 0$ (voir figure 1.1). Le déchiffrement est obtenu en inversant l'ordre des constantes de tour et en utilisant la fonction : $F^{-1}(x) = x^s$ où $s = (2^{n+1} - 1)/3$.

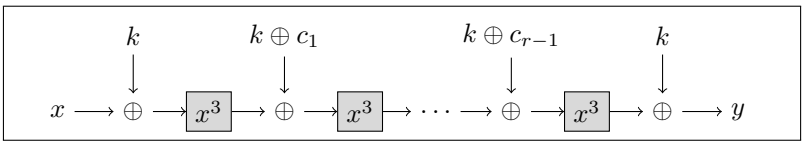

FIGURE 1.1 – Chiffrement MiMC avec r tours

On peut ainsi voir le résultat après r tours comme un polynôme univarié sur $\mathbb{F}_{2^n}$. Le degré algébrique (voir chapitre 2) est alors le degré maximum des n polynômes à n variables décrivant la fonction sur $\mathbb{F}_{2^n}$, ou encore le maximum des poids de Hamming des exposants présents dans la forme développée du polynôme univarié. Dans le cas de MiMC, le degré algébrique maximal est donc $n - 1$ car les opérations sont réalisées dans $\mathbb{F}_{2^n}$.

Le chiffrement par bloc MiMC a été conçu pour résister à diverses techniques de cryptanalyse. Le nombre d'itérations, r, a notamment été choisi pour que le degré de la fonction après r tours atteigne le degré attendu pour une permutation aléatoire dont les opérations sont effectuées sur le même alphabet. Toutefois, le principe de conception algébrique du système suscite une inquiétude naturelle quant à la sécurité contre les techniques

de cryptanalyse algébrique. Plusieurs attaques algébriques possibles sont présentées dans [AGR+16] avec une analyse de la résistance de MiMC contre ces attaques.

Nous nous intéresserons plus particulièrement aux attaques sur le degré algébrique puisqu'il est notamment possible d'exploiter le fait que le degré multivarié du chiffrement reste inférieur à sa valeur maximale. Par exemple [EGL+20] propose une généralisation de la cryptanalyse différentielle d'ordre supérieur en utilisant notamment un distingueur à somme nulle qui exploite le fait que le degré algébrique de MiMC croît lentement.

L'objectif de ce travail est ainsi d'exploiter la simplicité de cette structure, ce qui nécessite d'analyser l'évolution du degré et de la densité des polynômes multivariés représentant MiMC et son inverse avec le nombre d'itérations. Dans la suite, nous analyserons donc l'évolution de ce degré multivarié en nous plaçant dans le cas où la clé est $k = 0$.

Dans un premier temps, nous introduirons quelques résultats d'algèbre dans le chapitre 2 qui nous permettront par la suite d'étudier l'évolution du degré décrivant la transformation. Nous nous attacherons alors à l'analyse de l'évolution du degré algébrique des polynômes représentant MiMC, dans le chapitre 3. L'élément essentiel de ce chapitre est la justification de la présence de paliers, jusqu'à présent inconnus, dans l'évolution du degré, détaillée dans la section 3.3, où nous présenterons notamment une généralisation du palier observé entre les deux premiers tours de MiMC. Enfin, le chapitre 4 sera dédié à l'étude du degré algébrique de la transformation inverse, pour laquelle plusieurs pistes seront proposées pour réaliser cette étude.

1.4 Cadre du stage

Ce travail a été effectué lors de mon stage de fin d'étude qui s'est déroulé au centre INRIA dans le 12e arrondissement de Paris. L'INRIA - Institut National de Recherche en Informatique et en Automatique - est organisé en équipes-projets rassemblant des chercheurs avec des compétences complémentaires autour d'un projet scientifique commun. Intégrée au sein de l'équipe-projet COSMIQ, j'ai été encadrée par Anne Canteaut, directrice de recherche, et Léo Perrin, chargé de recherche. Le projet COSMIQ fait suite au projet SECRET depuis le 1er janvier 2020. Les axes de recherche de l'équipe sont : la cryptologie symétrique, la cryptologie fondée sur les codes et l'information quantique. Le sujet du stage, s'inscrit donc plus particulièrement dans la première de ces thématiques puisqu'il porte sur la cryptanalyse de primitives symétriques.

Mon stage a débuté le 9 mars 2020, mais la crise sanitaire à laquelle nous avons été confrontés m'a contrainte à rejoindre mon domicile familial dès le 16 mars afin de poursuivre le stage en télétravail. Dans ces conditions particulières de confinement, j'ai notamment rencontré quelques problèmes informatiques, car la vétusté et la puissance de calcul réduite de mon ordinateur personnel augmentaient considérablement le temps d'exécution des programmes. Par ailleurs, après seulement une semaine de présence au centre, je n'avais pas encore eu le temps de me familiariser avec le sujet et mon environnement de travail. J'ai donc eu quelques difficultés pour m'approprier le sujet au cours des premières semaines de ce stage.

Chapitre 2

Quelques résultats d'algèbre

Dans ce chapitre, nous introduisons certaines définitions et propriétés algébriques qui nous seront utiles afin d'étudier l'évolution du degré algébrique des polynômes décrivant MiMC et son inverse.

2.1 Degré d'une fonction vectorielle

Définition 2.1. *Une fonction booléenne à n variables est une fonction de $\mathbb{F}_2^n$ dans $\mathbb{F}_2 = \{0, 1\}$.*

Les coordonnées d'une fonction vectorielle F de $\mathbb{F}_2^n$ dans $\mathbb{F}_2^m$ sont les m fonctions booléennes F_i, $1 \leq i \leq m$ telles que $F(x) = (F_1(x), \ldots, F_m(x))$ pour tout x. Le degré algébrique de F est alors défini par le degré algébrique de ses coordonnées, comme suit :

Définition 2.2. *Soit f une fonction de $\mathbb{F}_2^n$ dans $\mathbb{F}_2$. Cette fonction peut s'écrire de manière unique comme un polynôme multivarié dans $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 - x_1), \ldots, (x_n^2 - x_n)$, appelé sa forme normale algébrique (ANF) :*

$$f(x_1, \ldots, x_n) = \sum_{u=(u_1,\ldots,u_n)\in\mathbb{F}_2^n} a_u \prod_{i=1}^{n} x_i^{u_i} \ .$$

Définition 2.3. *Sous les conditions précédentes, le degré algébrique de f est alors défini par :*

$$\deg f = \max\{wt(u) : u \in \mathbb{F}_2^n, a_u \neq 0\} \ ,$$

où wt représente le poids de Hamming.

Pour une fonction $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$, le degré algébrique de F correspond au degré maximum des m polynômes à n variables décrivant la fonction sur l'espace vectoriel $\mathbb{F}_2^m$.

Par ailleurs, chaque fonction vectorielle $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ peut être vue comme un polynôme univarié défini sur $\mathbb{F}_{2^n}$, car $\mathbb{F}_{2^n}$ peut être identifié à un espace vectoriel de dimension n sur $\mathbb{F}_2$. Pour chaque fonction $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, il existe alors une unique représentation polynomiale univariée sur $\mathbb{F}_{2^n}$ de degré au plus $2^n - 1$:

$$F(x) = \sum_{i=0}^{2^n-1} b_i x^i, \ b_i \in \mathbb{F}_{2^n} \ .$$

Dans ce cas, on peut donc faire le lien entre le degré algébrique et le degré univarié. En effet, le degré algébrique de F correspond au maximum des poids de Hamming des exposants [1] apparaissant dans la forme développée du polynôme :

$$\deg(F) = \max\{wt(i), \ 0 \leq i < 2^n, \text{ et } \ b_i \neq 0\} \ .$$

1. le poids de Hamming d'un entier est celui de sa décomposition binaire

2.2 Degré maximal de la composition d'une fonction puissance

Considérons une fonction puissance x^d sur $\mathbb{F}_{2^n}$, et une fonction de chiffrement F consistant à appliquer cette fonction puissance de manière itérative : $F = F_r \circ \ldots \circ F_0$ où pour tout i, $0 \leq i \leq r$, $F_i : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, $x \mapsto (x \oplus c_i)^d$, et les $c_i \in \mathbb{F}_{2^n}$ sont des constantes aléatoires sauf $c_0 = c_r = 0$. Il est alors possible de déterminer de manière récursive les monômes présents à chaque tour de chiffrement. La première constante étant nulle, on a en particulier la relation suivante.

Proposition 2.4. *Soit $\mathcal{M}_r$ la liste des exposants des monômes susceptibles d'apparaître dans le polynôme univarié $\mathcal{P}_r$ après r tours, on a :*

$$\mathcal{M}_r = \{dj \bmod (2^n - 1) \text{ où } j \preceq i\,^2,\ i \in \mathcal{M}_{r-1}\} .$$

Démonstration. En effet, si on a :

$$\mathcal{P}_{r-1}(x) = \sum_{i \in \mathcal{M}_{r-1}} \alpha_i x^i \quad \text{alors} \quad \mathcal{P}_r(x) = \mathcal{P}'_{r-1}(x^d + c_1) = \sum_{i \in \mathcal{M}_{r-1}} \alpha'_i (x^d + c_1)^i ,$$

où $\mathcal{P}'_{r-1}$ correspond au polynôme $\mathcal{P}_{r-1}$ après décalage des constantes (chaque c_k apparaissant dans les α_i est remplacé par c_{k+1} dans les α'_i).

Or

$$(x^d + c_1)^i = \prod_{\ell \in I} (x^d + c_1)^{2^\ell} = \prod_{\ell \in I} (x^{d2^\ell} + c_1^{2^\ell}) \quad \text{où} \quad I = Supp(i) .$$

Ainsi, les termes obtenus après développement sont de la forme :

$$\left(x^{d \sum_{\ell \in J} 2^\ell}\right) \left(c_1^{\sum_{\ell \in I \setminus J} 2^\ell}\right) \quad \text{où} \quad J \subseteq I .$$

Finalement, les exposants présents après r tours sont : $\mathcal{M}_r = \{(dj) \bmod (2^n - 1) \text{ où } j \preceq i, i \in \mathcal{M}_{r-1}\}$. □

Le degré après r tours de la fonction est alors le poids maximal des éléments de $\mathcal{M}_r$.

Par ailleurs, notons que si la première constante n'était pas nulle, le degré algébrique serait identique. En effet, pour toute fonction polynomiale $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, $x \mapsto F(x)$, le degré de $F(x + c)$ où c est une constante est le même que le degré de $F(x)$. Si $F(x) = \sum_i \alpha_i x^i$, alors $\deg(F(x)) = \max_i wt(i)$ et

$$F(x + c) = \sum_i \alpha_i (x + c)^i = \sum_i \alpha_i \prod_{\ell \in I} (x + c)^{2^\ell} = \sum_i \alpha_i \prod_{\ell \in I} (x^{2^\ell} + c^{2^\ell}) = \sum_i \alpha_i \sum_{J \subseteq I} x^{\sum_{\ell \in J} 2^\ell} c^{\sum_{\ell \in I \setminus J} 2^\ell} \quad \text{où } I = Supp(i) .$$

Par conséquent, on a bien : $\deg(F(x + c)) = \max_{J \subseteq Supp(i)} wt(\sum_{\ell \in J} 2^\ell) = \max_i wt(i) = \deg(F(x))$.

2.3 Degré de F^{-1} lorsque F est une permutation

Définition 2.5. *Une permutation polynomiale est un polynôme qui agit comme une permutation c'est-à-dire comme une bijection de l'ensemble dans lui-même.*

En particulier, on a le résultat suivant :

Proposition 2.6. *Tout monôme x^d est une permutation dans $\mathbb{F}_{2^n}$ si et seulement si* $\text{pgcd}(d, 2^n - 1) = 1$.

On en déduit alors que le degré d'extension n du corps dans lequel sont effectuées les opérations pour le design de MiMC doit être impair pour que $x \mapsto x^3$ soit une permutation dans $\mathbb{F}_{2^n}$.

Lorsque $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ est une permutation, il existe un lien entre son degré et celui de son inverse (voir [BC13] pour de plus amples détails).

2. pour des entiers i et j, $j \preceq i$ signifie $Supp(j) \subseteq Supp(i)$ où $Supp(i)$ est le nombre de 1 dans la décomposition binaire de i.

Définition 2.7. *Soit $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$. Pour tout entier k, $1 \leq k \leq m$, on note $\delta_k(F)$ le degré algébrique maximal du produit d'au plus k coordonnées de F.*

$$\delta_k(F) = \max_{K \subset \{1,\dots m\}, |K| \leq k} \deg \left(\prod_{i \in K} F_i \right) .$$

Théorème 2.8 (Théorème 3.1 de [BC13])**.** *Soit $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ une permutation. Pour tous entiers k et l tels que : $1 \leq k, l \leq n$ on a :*

$$\delta_l(F^{-1}) < n - k \Leftrightarrow \delta_k(F) < n - l .$$

Si $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ est une permutation et G une fonction de $\mathbb{F}_2^n$ dans $\mathbb{F}_2^n$, alors on a : $\deg G \circ F \leq \delta_{\deg G}(F)$ et $\delta_l(F) \leq l \deg F$. On en déduit également le corollaire 2.9 que nous utiliserons au chapitre 4 :

Corollaire 2.9 (Corollaire 3.3 de [BC13])**.** *Si $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ est une permutation, alors*

$$\deg(F^{-1}) = n - 1 \iff \deg(F) = n - 1 .$$

2.4 Fonctions AB et APN

Intéressons-nous à présent à des fonctions ayant des propriétés particulières en cryptanalyse différentielle et linéaire. Les fonctions presque courbes (AB) sont les fonctions de $\mathbb{F}_2^n$ dans $\mathbb{F}_2^n$ qui garantissent la meilleure résistance possible à la cryptanalyse linéaire. Ce sont également des fonctions presque parfaitement non-linéaires (APN), c'est-à-dire qu'elles assurent une résistance optimale à la cryptanalyse différentielle. L'inverse n'étant pas vrai, sauf dans le cas des fonctions quadratiques, une condition nécessaire et suffisante pour qu'une fonction APN soit AB est proposée dans [CCD99]. Cette dernière fait intervenir le poids de Hamming d'entiers modulo $2^n - 1$ et nous sera donc utile par la suite.

Soit $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, alors pour tout entier $a, b \in \mathbb{F}_{2^n}$, on définit :

$$\delta_F(a,b) = \#\{x \in \mathbb{F}_2^n, F(x+a) + F(x) = b\} \quad \text{et} \quad \lambda_F(a,b) = |\#\{x \in \mathbb{F}_2^n, ax + bF(x) = 0\} - 2^{n-1}| .$$

Définition 2.10. *Une fonction $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ est presque parfaitement non-linéaire (APN) si*

$$\delta_F = \max_{a \neq 0} \max_b \delta_F(a,b) = 2 .$$

Et une fonction $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ est presque courbe (AB) si

$$\lambda_F = \max_a \max_{b \neq 0} \lambda_F(a,b) = 2^{\frac{n-1}{2}} .$$

Proposition 2.11 (Corollaire 3 de [CCD99])**.** *Soit $n = 2k+1$, et $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}, x \mapsto x^s$ une fonction puissance sans composante affine, alors F est AB si et seulement si F est APN et*

$$\forall u,\ 1 \leq u \leq 2^n - 1,\ wt(us \bmod 2^n - 1) \leq k + wt(u) .$$

En particulier $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}, x \mapsto x^3$ étant quadratique, elle est AB, et par conséquent son inverse $F^{-1} : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}, x \mapsto x^s$ avec $s = (2^{n+1} - 1)/3$ l'est aussi.

Chapitre 3

Étude du degré algébrique de MiMC

Concentrons-nous dans un premier temps sur l'évolution du degré algébrique de la fonction de chiffrement. Afin de faire des simulations de l'évolution du degré sur des versions réduites, j'ai réalisé plusieurs implémentations en Sage et en C. Les calculs étant très coûteux, j'ai notamment utilisé la librairie gmp pour travailler avec de plus grands entiers. Sur la figure 3.1, nous pouvons alors remarquer que le degré algébrique évolue sous forme de paliers. Il s'agit donc de déterminer les tours auxquels apparaissent des paliers.

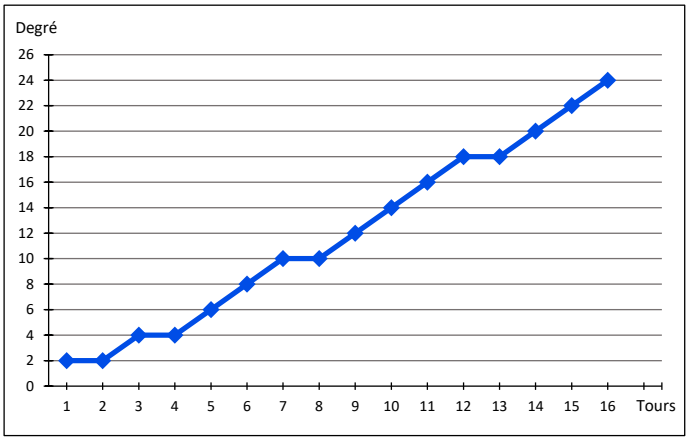

FIGURE 3.1 – Évolution du degré algébrique de la fonction de chiffrement

Dans une première approche, nous étudierons plus particulièrement les tours r tels que $3^r < 2^n - 1$, et nous verrons alors qu'il existe une borne sur le degré dans la section 3.1. Ensuite, dans la section 3.2 nous analyserons plus en détail les monômes présents dans le polynôme univarié décrivant la transformation. Cela nous permettra dans la section 3.3, de nous attacher au point essentiel de ce chapitre, à savoir la justification des paliers dans l'évolution du degré, dont le résultat principal est la proposition 3.6. Toutefois, on peut se demander si le choix de constantes particulières, pourrait faire baisser le degré à certains tours. La section 3.4 sera alors consacrée à l'étude des coefficients des monômes présents dans le polynôme. Finalement, nous ferons quelques remarques pour les tours r tels que $3^r \geq 2^n - 1$, dans la section 3.5.

3.1 Borne sur le degré

Tout d'abord, nous pouvons remarquer que la relation de récurrence de la proposition 2.4 nous permet de définir une borne maximale sur le degré. Dans le cas général, tant que le degré univarié du polynôme n'excède pas $2^n - 1$, le degré algébrique est au plus $\lfloor \log_2(3^r) \rfloor = \lfloor r \log_2 3 \rfloor$. Néanmoins, cette borne peut être améliorée, car les exposants apparaissant sont uniquement des multiples de 3. Or, si $\lfloor r \log_2 3 \rfloor$ est impair, il ne peut y avoir d'exposant atteignant ce degré, car un entier de la forme $2^{2k+1} - 1$ n'est pas divisible par 3. La borne devient donc : $2 \times \lfloor (r \log_2 3)/2 \rfloor$, et le degré maximal est ainsi toujours pair. En observant que

$\lfloor (r \log_2 3)/2 \rfloor \leq \lfloor ((r-1)\log_2 3)/2 \rfloor + 1$, on déduit donc que, entre deux tours consécutifs, le degré augmente de 2 ou reste stable (on parlera alors de palier).

Par ailleurs, on peut également minorer le degré algébrique. En effet, si le degré univarié est inférieur à $2^n - 1$, il y a nécessairement le monôme x^{3^r}. De plus, le coefficient devant ce monôme étant toujours de 1, il est indépendant du choix des constantes et donc ne s'annule pas.

La figure 3.2 permet ainsi de comparer le degré observé avec ces deux bornes, dans le cas particulier où le degré d'extension est $n = 25$. On remarque alors que le degré observé semble très proche de la borne maximale.

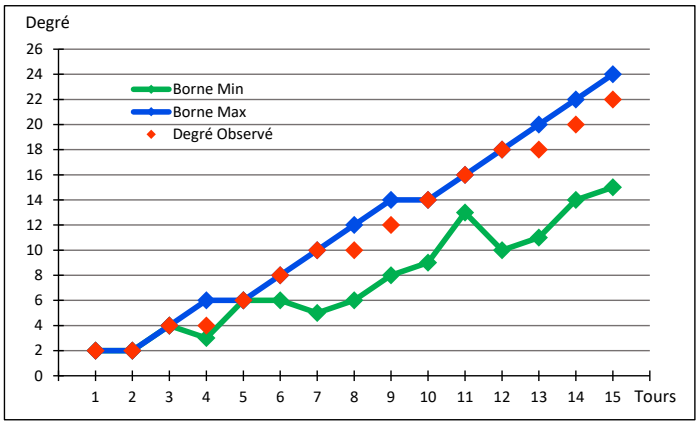

FIGURE 3.2 – Comparaison du degré observé avec les bornes (pour $n = 25$)

3.2 Quelques monômes particuliers

Afin d'étudier plus en détail l'évolution du degré algébrique, il est nécessaire d'observer la forme des exposants des monômes apparaissant dans le polynôme univarié représentant la fonction après r tours. On souhaite ainsi savoir si les exposants absents (sous-section 3.2.1), ou présents (sous-section 3.2.2), à chaque tour ont une forme particulière, permettant alors de les identifier sans avoir à construire chacun des polynômes des tours précédents.

Nous regarderons notamment l'évolution de la parité de $\lfloor r \log_2 3 \rfloor$ pour les puissances successives de 3. Pour cela, notons P les tours tels que $\lfloor r \log_2 3 \rfloor$ est pair et I sinon. On a alors le lemme 3.1.

Lemme 3.1. *Il n'est pas possible d'obtenir des séquences de la forme P-I-P ou I-P-I (c'est-à-dire telles que* $\lfloor r \log_2 3 \rfloor = \lfloor (r-2)\log_2 3 \rfloor + 2$*), ou de la forme P-P-P-P ou I-I-I-I (c'est-à-dire telles que* $\lfloor r \log_2 3 \rfloor = \lfloor (r-3)\log_2 3 \rfloor + 6$*).*

Démonstration. Tout d'abord, constatons que $\lfloor (r-1)\log_2 3 \rfloor + 1 \leq \lfloor r \log_2 3 \rfloor \leq \lfloor (r-1)\log_2 3 \rfloor + 2$. Plus précisément, $\log_2 3 \approx 1,585$ donc $2\log_2 3 > 3$, et on a $\lfloor r \log_2 3 \rfloor \geq \lfloor (r-2)\log_2 3 \rfloor + 3 > \lfloor (r-2)\log_2 3 \rfloor + 2$. De même, $3\log_2 3 < 5$ donc on a $\lfloor r \log_2 3 \rfloor \leq \lfloor (r-3)\log_2 3 \rfloor + 5 < \lfloor (r-3)\log_2 3 \rfloor + 6$. □

3.2.1 Absence de monômes

Premièrement, j'ai pu constater que certains exposants n'apparaissaient jamais dans le polynôme univarié décrivant la transformation. En particulier, aucun monôme n'a d'exposant de la forme $2^{2k} - 1$, et plus généralement on a le lemme 3.2.

Lemme 3.2. *Soit $\mathcal{P}_r$ le polynôme univarié décrivant la transformation après r tours, alors pour tout monôme x^j de $\mathcal{P}_r$, on a $j \not\equiv 5, 7 \mod 8$.*

Démonstration. Montrons-le par récurrence :

— Au tour 3, $\mathcal{M}_3 = \{0, 3, 6, 9, 12, 18, 24, 27\}$ donc les monômes d'exposant $15 \equiv 7 \mod 8$ et $21 \equiv 5 \mod 8$ sont absents.

— Supposons qu'aucun monôme n'a d'exposant congru à 5 ou 7 modulo 8 au tour r. Soit $i \in \mathcal{M}_r$, par hypothèse : $3|i$ mais $i \not\equiv 5,7 \mod 8$. Donc $\forall j \in J_1 = \{j \preceq i \in \mathcal{M}_r\}$, on a $j \not\equiv 5,7 \mod 8$.
Soit $J_2 = \{j,\ j \leq 3^r,\ j \equiv 5,7 \mod 8\}$, on a $J_1 \cap J_2 = \emptyset$. Comme $J_2 = \{5+8k,\ k \in \mathbb{N}\} \cup \{7+8k,\ k \in \mathbb{N}\}$, on a $3 \times J_2 = \{3 \times j,\ j \in J_2\} = \{15+24k,\ k \in \mathbb{N}\} \cup \{21+24k,\ k \in \mathbb{N}\}$.
De plus, $\mathcal{M}_{r+1} = \{3j \bmod (2^n-1) \text{ où } j \in J_1\}$, donc $(3 \times J_2) \cap \mathcal{M}_{r+1} = \emptyset$, et $\mathcal{M}_{r+1}$ ne contient pas les exposants $15+24k$ et $21+24k$ avec $k \in \mathbb{N}$.
Enfin, si $j = 5+8k$ est tel que $3|(5+8k)$, alors $3|(2k+2)$, et on peut écrire $k = 3k'+2$. Donc $j = 21+24k'$.
De même, si $j = 7+8k$ est tel que $3|7+8k$, alors $\exists k',\ j = 15+24k'$.
Finalement, $\forall i \in \mathcal{M}_{r+1}$, $3|i$ et $i \not\equiv 5,7 \mod 8$.

□

L'absence de ces monômes joue un rôle primordial dans la démonstration de la proposition 3.6, dans la sous-section 3.3.2, c'est-à-dire pour expliquer l'évolution du degré algébrique lorsque le degré univarié du polynôme n'excède pas $2^n - 1$.

Par ailleurs, afin de justifier l'évolution du degré dans la sous-section 3.3.2, nous aurons besoin de voir que $2^{2k+1} - 5 > 3^r$ si $\lfloor r \log_2 3 \rfloor = 2k$. Plus généralement, il semble que cette inégalité soit vérifiée pour chaque tour à partir du quatrième, comme le suggère la conjecture 3.3.

Conjecture 3.3. *Soit $k = \lfloor r \log_2 3 \rfloor$, alors pour tout $r > 4$, on a : $2^{k+1} - 5 > 3^r$.*

Lemme 3.4. *Pour tout r, $4 < r \leq 43491450$ et $k = \lfloor r \log_2 3 \rfloor$, on a : $2^{k+1} - 5 > 3^r$.*

Démonstration. Cette preuve s'appuie sur diverses pistes que j'ai explorées pour tenter de montrer que $2^{\lfloor r \log_2 3 \rfloor + 1} - 5 > 3^r$, pour tout $r > 4$.

D'une part, si $\lfloor r \log_2 3 \rfloor = 2k$, $3^r \notin \{2^{2k+1}-4, 2^{2k+1}-3, 2^{2k+1}-2, 2^{2k+1}-1\}$ donc il suffit de voir que $3^r \neq 2^{2k+1} - 5$. En effet, si $3^r \geq 2^{2k+1}$ on aurait $\lfloor r \log_2 3 \rfloor = 2k+1$, de plus $3 \nmid 2^{2k+1}-4, 2^{2k+1}-3, 2^{2k+1}-1$, et $3^r \neq 2^{2k+1} - 2$ car $2 \nmid 3^r$. Afin de prouver que $3^r \neq 2^{2k+1} - 5$, j'ai notamment démontré par récurrence que 3^r peut s'écrire pour $a \in \mathbb{N}$:

$$\begin{array}{rcl|rcl}
1+2^5a & \text{si} & r \equiv 0 \mod 8 & 1+2^4+2^5a & \text{si} & r \equiv 4 \mod 8 \\
1+2+2^5a & \text{si} & r \equiv 1 \mod 8 & 1+2+2^4+2^5a & \text{si} & r \equiv 5 \mod 8 \\
1+2^3+2^5a & \text{si} & r \equiv 2 \mod 8 & 1+2^3+2^4+2^5a & \text{si} & r \equiv 6 \mod 8 \\
1+2+2^3+2^4+2^5a & \text{si} & r \equiv 3 \mod 8 & 1+2+2^3+2^5a & \text{si} & r \equiv 7 \mod 8
\end{array}$$

Donc $\forall r,\ r \not\equiv 3 \mod 8$ on a $3^r \neq 2^{2k+1} - 5$. Il reste alors à voir que les puissances de 3 s'écrivant $1+2+2^3+2^4+2^5a$ ne sont pas de la forme $2^{2k+1} - 5$. Cela correspond aux tours : $11 + 8j$, $j \in \mathbb{N}$ (car $27 = 3^3 = 2^5 - 5$).

D'autre part, on peut aussi montrer $3^r < 2^k - 5$ où $k = \lfloor r \log_2 3 \rfloor + 1$ en utilisant une récurrence forte, c'est-à-dire en supposant l'inégalité vérifiée pour tous les tours jusqu'au tour $r-1$. Premièrement, on peut constater que le résultat est vrai sur les premiers tours. Distinguons ensuite plusieurs cas en fonction de la parité de $\lfloor r \log_2 3 \rfloor$ et des tours précédents :

1. Cas I : si $\lfloor r \log_2 3 \rfloor = 2k+1$, il suffit de voir que $3^r \notin \{2^{2k+2}-5, 2^{2k+2}-4, 2^{2k+2}-3, 2^{2k+2}-2, 2^{2k+2}-1\}$. En effet, si $3^r \geq 2^{2k+2}$ on aurait eu $\lfloor r \log_2 3 \rfloor = 2k+2$. De plus $3 \nmid 2^{2k+2}-5, 2^{2k+2}-3, 2^{2k+2}-2$, et $3^r \neq 2^{2k+2} - 4$ car $4 \nmid 3^r$. Enfin, $3^r \neq 2^{2k+2} - 1$ car $2^{2k+2} - 1 \equiv 7 \mod 8$, ce qui est impossible d'après la forme des puissances de 3. Donc $3^r < 2^{2k+2} - 5$.
2. Cas P-P : lorsque $\lfloor r \log_2 3 \rfloor = 2k$ et $\lfloor (r-1) \log_2 3 \rfloor = 2k-2$, on a par hypothèse $3^{r-1} < 2^{2k-1} - 5$ donc $3^r < 3 \times (2^{2k-1} - 5) = 2^{2k} + 2^{2k-1} - 15 < 2^{2k+1} - 5$.
3. Cas I-I-I-P : si $\lfloor r \log_2 3 \rfloor = 2k$, $\lfloor (r-1) \log_2 3 \rfloor = 2k-1$, $\lfloor (r-2) \log_2 3 \rfloor = 2k-3$ et $\lfloor (r-3) \log_2 3 \rfloor = 2k-5$, on a $3^{r-3} < 2^{2k-4} - 5$ donc $3^r < 27 \times (2^{2k-4} - 5) < 2^{2k+1} - 5$ (car $27 = 2^5 - 5$).
4. Cas P-P-P-I-I-P : lorsque $\lfloor r \log_2 3 \rfloor = 2k$, $\lfloor (r-1) \log_2 3 \rfloor = 2k-1$, $\lfloor (r-2) \log_2 3 \rfloor = 2k-3$, $\lfloor (r-3) \log_2 3 \rfloor = 2k-4$, $\lfloor (r-4) \log_2 3 \rfloor = 2k-6$ et $\lfloor (r-5) \log_2 3 \rfloor = 2k-8$ on a $3^{r-5} < 2^{2k-7} - 5$ donc $3^r < 243 \times (2^{2k-7} - 5)$. Or $243 = 2^8 - 13$ d'où $3^r < 2^{2k+1} - 13 \times 2^{2k-7} - 5 \times 2^8 + 65 < 2^{2k+1} - 5$.
5. Cas I-P-P-I-I-P : si $\lfloor r \log_2 3 \rfloor = 2k$, $\lfloor (r-1) \log_2 3 \rfloor = 2k-1$, $\lfloor (r-2) \log_2 3 \rfloor = 2k-3$, $\lfloor (r-3) \log_2 3 \rfloor = 2k-4$, $\lfloor (r-4) \log_2 3 \rfloor = 2k-6$ et $\lfloor (r-5) \log_2 3 \rfloor = 2k-7$, l'exposant semble a priori également supérieur à 3^r mais les inégalités observées sur les tours précédents ne permettent pas de le démontrer de façon générale.

On a ainsi montré que pour les tours r tels que $r \equiv 0, 1, 2, 4, 5, 6, 7 \mod 8$ on a bien $2^{\lfloor \log_2 3^r \rfloor + 1} - 5 \neq 3^r$. De plus, en regardant la parité de $\lfloor \log_2 3^r \rfloor$ on a montré que $2^{\lfloor \log_2 3^r \rfloor + 1} - 5 > 3^r$ sauf éventuellement pour la suite I-P-P-I-I-P, c'est-à-dire $\lfloor \log_2 3^r \rfloor$ pair - $\lfloor \log_2 3^{r-1} \rfloor$ impair, ... etc. En effet, d'après le lemme 3.1, seules les séquences traitées ci-dessus sont possibles. Il reste donc à voir ce qui se passe dans le cas I-P-P-I-I-P tel que $r \equiv 3 \mod 8$ c'est-à-dire tel que 3^r s'écrit $1 + 2 + 2^3 + 2^4 + 2^5 a$.

Soit $\mathcal{R} = \{123, 147, 523, 547, 947, 971, \ldots\}$ l'ensemble des tours r vérifiant ces deux conditions. Si on avait $2^{\lfloor \log_2 3^r \rfloor + 1} - 5 = 3^r$ alors on aurait $\forall i \leq r,\ 3^i | 2^{\lfloor \log_2 3^r \rfloor + 1} - 5$. J'ai donc observé algorithmiquement la divisibilité de $2^{\lfloor \log_2 3^r \rfloor + 1} - 5$, pour $r \in \mathcal{R}$, par des puissances de 3 (jusqu'à $3^{11} = 177147$ pour des raisons de temps de calcul). On remarque alors que pour tout $r \in \mathcal{R}$ tel que $r \leq 43491450$, $3^{11} \nmid (2^{\lfloor \log_2 3^r \rfloor + 1} - 5)$ donc dans ce cas $2^{\lfloor \log_2 3^r \rfloor + 1} - 5 \neq 3^r$. Ensuite, $43491451 \in \mathcal{R}$ tel que $\lfloor \log_2 3^{43491451} \rfloor = 68932318$ mais $3^{11} | 2^{68932319} - 5$. Il faudrait alors pouvoir utiliser des algorithmes plus performants pour observer la divisibilité de $2^{68932319} - 5$ par de plus grandes puissances de 3 afin de s'assurer que $2^{68932319} - 5 \neq 3^{43491451}$. En conclusion, pour tout $r \leq 43491450$ on a bien $2^{\lfloor \log_2 3^r \rfloor + 1} - 5 \neq 3^r$ (ce qui correspond à $n \leq 68932317$, et $68932317 \sim 2^{26} = 67108864$). □

3.2.2 Présence de monômes

Afin de justifier le degré algébrique à chaque tour, il est également nécessaire d'étudier les monômes ayant un degré de poids maximal. J'ai effectivement pu remarquer, pour les tours 5 à 14, que les exposants de poids maximal ont une forme particulière en fonction de la parité de $\lfloor r \log_2 3 \rfloor$. Soit $\mathcal{N}_r$ l'ensemble des exposants donnant le degré algébrique au tour r :

— si $\lfloor r \log_2 3 \rfloor = 2k + 1$ et $\lfloor (r-1) \log_2 3 \rfloor = 2k$, le degré algébrique est donné par les 4 exposants :

$$\mathcal{N}_r = \{2^{2k+1} - 5, 2^{2k+1} - 2, 2^{2k+2} - 2^{2k-1} - 5, 2^{2k+2} - 2^{2k-1} - 2\}$$

(le tour 5 semble être une exception).

— si $\lfloor r \log_2 3 \rfloor = 2k+1$ et $\lfloor (r-1) \log_2 3 \rfloor = 2k-1$, le degré algébrique est donné par 2 exposants uniquement :

$$\mathcal{N}_r = \{2^{2k+1} - 5, 2^{2k+1} - 2\} .$$

— si $\lfloor r \log_2 3 \rfloor = 2k$ et $\lfloor (r-1) \log_2 3 \rfloor = 2k - 1$, le degré algébrique est donné par les $4k - 3$ exposants :

$$\mathcal{N}_r = \{2^{2k-1} - 5, 2^{2k-1} - 2, 2^{2k} - 7, 2^{2k} - 4\} \cup \{2^{2k} - 2^j - 5, 2^{2k} - 2^j - 2 \text{ où } j = 2i + 1, 1 \leq i \leq k - 2\}$$
$$\cup \{2^{2k+1} - 2^{2k-1} - 2^j - 5, 2^{2k+1} - 2^{2k-1} - 2^j - 2 \text{ où } j = 2i, 2 \leq i \leq k - 1\} \cup \{2^{2k+1} - 2^{2k-1} - 6\} .$$

— si $\lfloor r \log_2 3 \rfloor = 2k$ et $\lfloor (r-1) \log_2 3 \rfloor = 2k - 2$ le degré algébrique est donné par $2k$ exposants :

$$\mathcal{N}_r = \{2^{2k-1} - 5, 2^{2k-1} - 2, 2^{2k} - 7, 2^{2k} - 4\} \cup \{2^{2k} - 2^j - 5, 2^{2k} - 2^j - 2 \text{ où } j = 2i + 1, 1 \leq i \leq k - 2\} .$$

En revanche, il n'est pas possible de déterminer expérimentalement les monômes de degré de poids maximal lorsque $\lfloor r \log_2 3 \rfloor = 2k$, $\lfloor (r-1) \log_2 3 \rfloor = 2k - 2$ et $\lfloor (r-2) \log_2 3 \rfloor = 2k - 4$. En effet, la première apparition de ce motif est pour $r = 19$ c'est-à-dire qu'il faudrait au minimum pouvoir réaliser les opérations dans $\mathbb{F}_{2^{31}}$. On peut néanmoins supposer que les monômes d'exposant $2^{2k-1} - 5, 2^{2k-1} - 2, 2^{2k} - 7, 2^{2k} - 4$ sont présents comme cela semble être le cas pour les tours tels que $\lfloor r \log_2 3 \rfloor = 2k$.

Une démonstration non achevée de ces observations est proposée en annexe A.1.

3.3 Justification des paliers

Cette section est consacrée à la justification des paliers observés sur la figure 3.1. Pour cela, nous traiterons en premier lieu les quatre premiers tours qui sont vérifiables à la main dans la sous-section 3.3.1, puis dans la sous-section 3.3.2, nous tenterons d'établir une preuve de l'évolution du degré algébrique pour les tours suivants.

3.3.1 Cas particuliers des quatre premiers tours

Palier entre les tours 1 et 2

Entre le premier et le deuxième tour, le degré univarié passe de 3 à 9, alors que le degré algébrique reste quadratique. En effet, on rappelle que $c_0 = 0$, donc au deuxième tour, on a :

$$(x^3 + c_1)^3 = x^9 + c_1 x^6 + c_1^2 x^3 + c_1^3 .$$

On en déduit que le degré algébrique est bien quadratique : deg = max$\{wt(i),$ pour $i \in \{0,3,6,9\}\} = 2$.

Il est également possible d'utiliser la proposition 2.4 donnant la liste des exposants présents. Comme $\mathcal{M}_1 = \{3\}$, on a alors :

$$\mathcal{M}_2 = \{3j \bmod (2^n-1) \text{ où } j \preceq 3\} = \{3j \bmod (2^n-1) \text{ où } j \in \{0,1,2,3\}\} = \{0,3,6,9\} \ .$$

Par ailleurs, on peut remarquer qu'en choisissant la première constante aléatoire, cela n'a pas de conséquence sur le palier observé pour le degré car 7 est le seul exposant de degré 3 inférieur à 9, et il n'est pas présent :

$$((x+c_0)^3+c_1)^3 = x^9 + c_0x^8 + c_1x^6 + c_0^2c_1x^4 + c_1^2x^3 + (c_0^4c_1 + c_0c_1^2)x^2 + (c_0^8 + c_0^2c_1^2)x + (c_0^3+c_1)^3 \ .$$

Plus généralement, l'ajout d'une constante ne change pas le degré algébrique, comme nous l'avons vu précédemment dans la section 2.2.

On peut également constater une conservation du degré algébrique entre les tours 1 et 2 du chiffrement MiMC utilisant d'autres permutations.

Proposition 3.5. *Soit $F : \mathbb{F}_2^n \to \mathbb{F}_2^n, x \mapsto x^d$ où $d = 2^k - 1$. Alors, si $d^2 < 2^n - 1$, on a :*

$$\deg((x^d+c)^d) = \deg(x^d) \quad \textit{où } c \textit{ est une constante, et } \deg \textit{ correspond au degré algébrique}$$

Démonstration. En utilisant le fait que

$$\mathcal{M}_2 = \{dj \bmod (2^n-1) \text{ où } j \preceq d\} = \{dj \bmod (2^n-1) \text{ où } j \in [\![0,d]\!]\} \ ,$$

cela revient à montrer que $wt(dj) \leq wt(d)$ pour tout entier $j \leq d$. Montrons, plus précisément, que lorsque $d = 2^k - 1$, $wt(dj) = wt(d) = k$ pour tout $j \in [\![1,d]\!]$.

Soit $j \preceq 2^k - 1$ tel que $j = \sum_{l=1}^{n} 2^{i_l}$, avec $n \in \{1,\dots,k-1\}$ et $k-1 \geq i_1 \geq \dots \geq i_n \geq 0$. Alors

$$jd = (2^{i_1} + \dots + 2^{i_n})(2^k-1) = \sum_{l=i_n}^{i_{n-1}-1} 2^l + \sum_{l=i_{n-1}+1}^{i_{n-2}-1} 2^l + \dots + \sum_{l=i_1}^{k+i_n-1} 2^l + \sum_{l=k+i_{n-1}}^{k+i_1} 2^l \ .$$

On a bien $wt(jd) = k$. □

Par conséquent, en utilisant MiMC avec toute permutation x^d vérifiant la proposition 3.5 (donc en particulier avec x^3) on observera ainsi un palier entre les deux premiers tours.

Tours 2 à 3

On sait que $\mathcal{M}_3 = \{3j \bmod (2^n-1) \text{ où } j \preceq i, i \in M_2\}$ avec $\mathcal{M}_2 = \{0,3,6,9\}$, donc on a forcément $wt(j) \leq 2$, ce qui implique $wt(3j) \leq 4$. De plus, $27 \in \mathcal{M}_3$ et $wt(27) = 4$ donc le degré entre le deuxième et le troisième tour passe de 2 à 4. On peut également le vérifier en regardant le poids de chacun des exposants de $\mathcal{M}_3$: deg = max$\{wt(i),$ pour $i \in \{0,3,6,9,12,18,24,27\}\} = 4$.

Palier entre les tours 3 et 4

Premièrement, on a $\mathcal{M}_3 = \{0,3,6,9,12,18,24,27\}$ avec : $\forall i \in \mathcal{M}_3 \backslash \{27\}$, $wt(i) = 2$ et $wt(27) = 4$. De plus, comme $\mathcal{M}_4 = \{3j \bmod (2^n-1) \text{ où } j \preceq i, i \in \mathcal{M}_3\}$ et $wt(3) = 2$, alors si $wt(j) \leq 2$, on a nécessairement $wt(3j) \leq 4$. Cette condition est toujours vérifiée sauf dans le cas $i = 27$. Or on a :

$$j \preceq 27 \iff j \preceq \sum_{k=0,1,3,4} 2^k \iff j \in \left\{ \sum_{k=0,1,3,4} \varepsilon_k 2^k, \varepsilon_k \in \{0,1\} \right\} \ .$$

Donc

$$3j = 3 \times \sum_{k=0,1,3,4} \varepsilon_k 2^k = \varepsilon_0 + (\varepsilon_0+\varepsilon_1)2 + \varepsilon_1 2^2 + \varepsilon_3 2^3 + (\varepsilon_3+\varepsilon_4)2^4 + \varepsilon_4 2^5 \ .$$

Regardons les cas où $wt(j) > 2$, c'est-à-dire, lorsqu'il y a strictement plus que deux éléments ε_k tels que $\varepsilon_k = 1$:

$$3j \in \{1+2^5, 1+2^3+2^4+2^5, 1+2^3+2^6, 2+2^2+2^3+2^4, 1+2^4+2^6\} \ .$$

Dans tous les cas, on observe que $wt(3j) \leq 4$. On a alors un nouveau palier, car le degré algébrique au quatrième tour est deg = max$\{wt(i),$ pour $i \in \mathcal{M}_4\} = 4$. Plus précisément, $\mathcal{N}_4 = \{27,30,51,54,57,75,78\}$ représente les exposants de poids maximal au quatrième tour.

3.3.2 Tours suivants

Le problème est de pouvoir déterminer les tours présentant un palier, et notamment trouver une généralisation de la démonstration utilisée pour les paliers entre les tours 1 et 2 puis entre les tours 3 et 4.

Premièrement, j'ai pu constater un lien entre les indices de tours r de la forme $\lceil n/\log_2 3\rceil$ (pour n impair) et les tours présentant des paliers. En effet, lorsque $\lceil (n+2)/\log_2 3\rceil - \lceil n/\log_2 3\rceil = 2$, il y a un palier entre les tours $r = \lceil n/\log_2 3\rceil$ et $r+1$ comme nous pouvons le constater dans les tableaux 3.1 et 3.2 :

n	7	9	11	13	15	17	19	21	23
r	5	6	7	9	10	11	12	14	15

TABLE 3.1 – Indices de tours de la forme $r = \lceil n/\log_2 3\rceil$

r	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
d	2	2	4	4	6	8	10	10	12	14	16	18	18	20	22	24

TABLE 3.2 – Tours présentant des paliers

Cette construction implique par exemple que quel que soit le degré de l'extension, il n'y a jamais de palier sur les deux derniers tours, et donc le degré maximal ne peut être atteint qu'au dernier tour.

Partant de cette hypothèse, j'ai également pu remarquer qu'à partir du quatrième tour, le degré algébrique semblait évoluer en fonction de $\lfloor r\log_2(3)\rfloor$, comme le montre le tableau 3.3 :

r	4	5	6	7	8	9	10	11	12	13	14
$\lfloor r\log_2(3)\rfloor$	6	7	9	11	12	14	15	17	19	20	22
d	4	6	8	10	10	12	14	16	18	18	20

TABLE 3.3 – Évolution du degré en fonction de $\lfloor r\log_2(3)\rfloor$

Proposition 3.6. *Soit r un entier qui satisfait la conjecture 3.3. Alors, le degré algébrique après r tours de MiMC est $2\times\lceil\lfloor r\log_2 3\rfloor/2-1\rceil$.*

Démonstration. Cette preuve s'appuie sur certains résultats observés et exposés dans les parties précédentes, mais qui n'ont pas encore été entièrement démontrés.

Montrons que si $\lfloor r\log_2 3\rfloor = 2k+1$ alors le degré algébrique observé au tour r est $2k$:

— Le degré n'est pas $2k+1$ car $3\nmid 2^{2k+1}-1$ et les exposants $2^{2k+2}-2^j-1$ avec $0\le j\le 2k-1$ sont soit non-divisibles par 3, soit non-présents. En effet, pour $j\neq 0,2$, les monômes d'exposants $2^{2k+2}-2^j-1\equiv 5,7 \mod 8$ sont absents (voir sous-section 3.2.1). Et pour $j=0$ ou 2, 3 ne divise ni $2^{2k+2}-5$, ni $2^{2k+2}-2$.

— On a par exemple le monôme d'exposant : $2^{2k+1}-5$ de poids $2k$ (voir début de preuve en partie 3.2.2)

Montrons que si $\lfloor r\log_2 3\rfloor = 2k$ alors le degré algébrique observé au tour r est $2k-2$:

— Le degré n'est pas $2k$ car $2^{2k}-1\equiv 7 \mod 8$ n'est pas présent et les exposants $2^{2k+1}-2^j-1$ avec $0\le j\le 2k-1$ sont soit non-divisibles par 3, soit non-présents. En effet, pour $j\neq 0,2$, les monômes d'exposants $2^{2k+1}-2^j-1\equiv 5,7 \mod 8$ sont absents (voir sous-section 3.2.1). Pour $j=0$: on a $3^r\neq 2^{2k+1}-2$ car $2\nmid 3^r$ donc $2^{2k+1}-2>3^r$ (sinon on aurait eu $\lfloor r\log_2 3\rfloor = 2k+1$). De même pour $j=2$, $2^{2k+1}-5>3^r$ (vérifié tant que $r\le 43491450$ voir sous-section 3.2.1).

— Le degré n'est pas $2k-1$ car $3\nmid 2^{2k-1}-1$ et les exposants $2^{2k}-2^j-1$ avec $0\le j\le 2k-1$ et $2^{2k+1}-2^j-2^i-1$ avec $1\le j\le 2k-1$ et $0\le i\le j-1$ sont soit non-divisibles par 3, soit non-présents. En effet, pour $j\neq 0,2$, les monômes d'exposants $2^{2k}-2^j-1\equiv 5,7 \mod 8$ sont absents (voir sous-section 3.2.1). De plus, pour $j=0$ ou 2, 3 ne divise ni $2^{2k}-5$, ni $2^{2k}-2$. Par ailleurs, si $j\ge 3$, alors $2^{2k+1}-2^j-2^i-1\equiv 5,7 \mod 8$ pour $i\ge 3$ et pour $i=1$. Il suffit donc d'étudier les cas $j=2$ ou $i\in\{0,2\}$ pour couvrir les cas restants. Pour $j=2$ ($i=0,1$), 3 ne divise ni $2^{2k+1}-7$, ni $2^{2k+1}-6$. De même pour $i=0$ ou 2, 3 ne divise ni $2^{2k+1}-2^j-5$, ni $2^{2k+1}-2^j-2$.

— On a notamment le monôme d'exposant : $2^{2k-1}-5$, de poids $2k-2$ (voir début de preuve en sous-section 3.2.2).

Par conséquent, le degré algébrique après r tours de MiMC est bien : $2 \times \lceil \lfloor r \log_2 3 \rfloor / 2 - 1 \rceil = 2k$ si $\lfloor r \log_2 3 \rfloor = 2k+1$ et $2 \times \lceil \lfloor r \log_2 3 \rfloor / 2 - 1 \rceil = 2k-2$ si $\lfloor r \log_2 3 \rfloor = 2k$. □

De cette expression, on peut alors en déduire que la parité de $\lfloor r \log_2 3 \rfloor$ influe sur les tours présentant des paliers et ceux atteignant la borne. En effet, le degré observé atteint la borne pour $2 \times \lceil \lfloor r \log_2 3 \rfloor / 2 - 1 \rceil = 2 \times \lfloor (r \log_2 3)/2 \rfloor$, c'est-à-dire lorsque $\lfloor r \log_2 3 \rfloor$ est impair.

De plus, le degré algébrique observé est conservé entre les tours r et $r+1$ lorsque $2 \times \lceil \lfloor (r+1) \log_2 3 \rfloor / 2 - 1 \rceil = 2 \times \lceil \lfloor r \log_2 3 \rfloor / 2 - 1 \rceil$. Il y a donc un palier dans l'évolution du degré lorsque $\lfloor r \log_2 3 \rfloor$ est impair et $\lfloor (r+1) \log_2 3 \rfloor$ est pair. On peut noter qu'à l'inverse, des paliers sont présents pour la borne lorsque $\lfloor r \log_2 3 \rfloor$ est pair et $\lfloor (r+1) \log_2 3 \rfloor$ est impair.

3.4 Forme des coefficients et lien avec le degré de MiMC utilisant x^9

3.4.1 Premières observations

Une autre piste de réflexion est d'étudier la forme des coefficients des monômes ayant un exposant de poids maximal. En effet, la liste des exposants des monômes donnée par $\mathcal{M}_r$ n'est pas minimale, car les coefficients de certains monômes peuvent s'annuler pour certains choix de constantes et faire chuter le degré à certains tours.

En réutilisant le raisonnement de la section 2.2 :

$$\mathcal{P}_r(x) = \sum_{i \in \mathcal{M}_{r-1}} \alpha'_i (x^d + c_1)^i = \sum_{i \in \mathcal{M}_{r-1}} \alpha'_i \sum_{J \subseteq I = Supp(i)} \left(x^{d \sum_{\ell \in J} 2^\ell} \right) \left(c_1^{\sum_{\ell \in I \setminus J} 2^\ell} \right) ,$$

on peut alors déterminer de manière itérative les coefficients des monômes à chaque tour. On rappelle que chaque c_k apparaissant dans α_i est remplacé par c_{k+1} dans α'_i.

Lemme 3.7. *Soit $\mathcal{C}_r$ l'ensemble des coefficients des monômes, non-distincts, présents au tour r on a :*

$$\mathcal{C}_r = \left\{ \alpha'_i \left(c_1^{\sum_{\ell \in I \setminus J} 2^\ell} \right), i \in \mathcal{M}_{r-1}, J \subseteq Supp(i) \right\} . \tag{3.1}$$

Dans l'expression de $\mathcal{P}_r$, présentée ci-dessus, les monômes $x^{d \sum_{\ell \in J} 2^\ell}$ ne sont pas distincts, c'est pourquoi dans la forme du polynôme telle que les monômes sont tous distincts, les coefficients des monômes correspondent à une somme d'éléments de $\mathcal{C}_r$, susceptible d'être nulle.

Il est également intéressant d'observer l'évolution du degré algébrique de x^9 par rapport à x^3. En effet, la transformation décrivant MiMC avec x^9 est équivalente à la transformation décrivant MiMC avec x^3, avec une constante sur deux qui est nulle. Sur la figure 3.3, nous pouvons ainsi constater que le degré algébrique au tour r pour x^9 n'est pas toujours le même que le degré algébrique au tour $2r$ pour x^3.

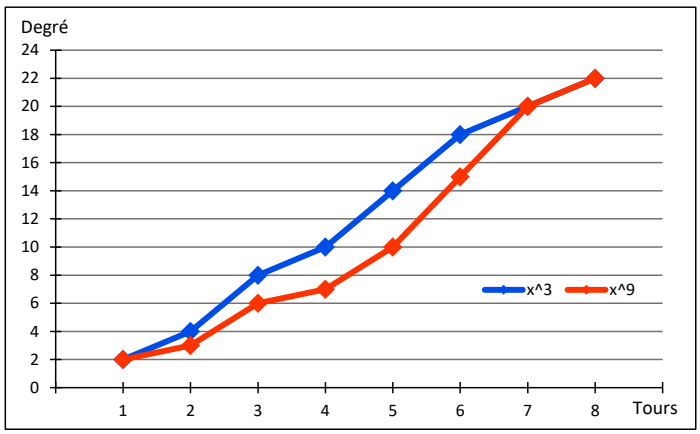

FIGURE 3.3 – Comparaison du degré algébrique aux tours r de MiMC avec x^9 et aux tours $2r$ avec x^3 ($n = 23$)

L'étude de x^9 permet par exemple de remarquer que certains coefficients s'annulent lorsque les constantes c_i pour i impair sont nulles. En effet, si on considère le polynôme représentant MiMC avec x^3, après 4 tours le degré algébrique observé est de 4, or après 2 tours de chiffrement MiMC avec x^9, le degré algébrique est 3. Il est donc, par exemple, possible de faire baisser le degré de 4 à 3 au tour 4 car les coefficients des monômes d'exposant de poids maximal dépendent uniquement des constantes d'indices impairs :

$$27 : c_1^{18} + c_3^2 \qquad 30 : c_1^{17} \qquad 51 : c_1^{10} \qquad 54 : c_1^9 + c_3 \qquad 57 : c_1^8 \qquad 75 : c_1^2 \qquad 78 : c_1$$

De même, le degré algébrique passe de 8 à 6 au tour 6 car les coefficients des monômes d'exposant de poids maximal sont : $c_1^2 c_3^8$ pour 507 et $c_1^{64} c_3 + c_1 c_3^8$ pour 510. De plus, il n'y a qu'un seul exposant de degré 7 : 702 et le coefficient devant le monôme correspondant est c_3. Plus généralement, les coefficients des monômes d'exposants non-divisibles par 9 peuvent se factoriser par une combinaison linéaire de constantes d'indices impairs.

3.4.2 Quelques monômes particuliers et lien avec x^3

Nous avions déjà montré dans la sous-section 3.2.1 que pour x^3, les monômes d'exposants congrus à 5 et 7 modulo 8 n'étaient jamais présents, pour x^9 on a le lemme 3.8.

Lemme 3.8. *Si $\mathcal{Q}_r$ est le polynôme univarié décrivant MiMC avec la fonction x^9 après r tours, alors pour tout monôme x^j de $\mathcal{Q}_r$, on a $j \mod 8 \in \{0, 1\}$.*

Démonstration. Montrons par récurrence qu'aucun exposant n'est de la forme $\{k+8n,\ n \in \mathbb{N}, k = 2, 3, 4, 5, 6, 7\}$
- Au tour 2, $\mathcal{M}_2 = \{0, 9, 72, 81\}$ donc tous les exposants satisfont $j \mod 8 \in \{0, 1\}$.
- Supposons qu'il n'y a aucun monôme d'exposant $\{k + 8n,\ n \in \mathbb{N}, k = 2, 3, 4, 5, 6, 7\}$ au tour r. Soit $i \in \mathcal{M}_r$ par hypothèse, $9|i$ et $i \notin \{k + 8n,\ n \in \mathbb{N}, k = 2, 3, 4, 5, 6, 7\}$.
  Soit $J_1 = \{j \preceq i \in \mathcal{M}_r\}$, on sait que $\forall j \in J_1$, $j \notin J_2 = \{j = k + 8n,\ n \in \mathbb{N}, k = 2, 3, 4, 5, 6, 7\}$, et on veut donc montrer que $\mathcal{M}_{r+1} = \{9j \bmod (2^n - 1) \text{ où } j \in J_1\}$ ne comporte aucun exposant appartenant à J_2.
  On a $J_1 \cap J_2 = \emptyset$ et $9 \times J_2 = \{9 \times j,\ j \in J_2\} = \{k + 72n,\ n \in \mathbb{N}, k = 18, 27, 36, 45, 54, 63\}$.
  Donc $(9 \times J_2) \cap \mathcal{M}_{r+1} = \emptyset$.
  De plus, tous les exposants de J_2 et multiples de 9, s'écrivent $k+72n$ avec $n \in \mathbb{N}$ et $k = 18, 27, 36, 45, 54, 63$.
  Finalement, $\forall i \in \mathcal{M}_{r+1}$, $9|i$ et $i \notin \{k + 8n,\ n \in \mathbb{N}, k = 2, 3, 4, 5, 6, 7\}$.

□

Dans la sous-section 3.2.2, nous avons listé les monômes ayant les exposants de poids maximal à chaque tour pour x^3. Or tous ces monômes ont des degrés congrus à 1,2,3,4,6 modulo 8 donc d'après le lemme 3.8, parmis ces monômes, les seuls susceptibles d'être présents dans le polynôme décrivant la transformation sont ceux dont l'exposant est congru à 1 modulo 8, c'est-à-dire les monômes dont l'exposant est de la forme $2^{2k} - 7$. Ces derniers ne font partie des exposants de poids maximal que sur les tours tels que $\lfloor \log_2 3^r \rfloor = 2k$, on en déduit que le degré du polynôme après r tours de MiMC avec x^9 ne coïncide jamais avec le degré du polynôme après $2r$ tours de MiMC avec x^3 lorsque $\lfloor \log_2 9^r \rfloor$ est impair.

De plus, notons que le monôme d'exposant $2^{2k} - 7$ n'est présent dans MiMC avec x^9 que s'il est divisible par 9, c'est-à-dire si $k \equiv 2 \mod 3$. En effet, $2^{2k} - 7 \equiv 0, 3, 6 \mod 9$ pour respectivement $k \equiv 2, 0, 1 \mod 3$, car d'une part, $2^4 - 7 = 9 \equiv 0 \mod 9$, $2^6 - 7 = 57 \equiv 3 \mod 9$ et $2^8 - 7 = 249 \equiv 6 \mod 9$. Et d'autre part, si $2^{2k} - 7 \equiv 0, 3, 6 \mod 9$, alors on a respectivement $2^{2(k+3)} - 7 = 2^6 \times 2^{2k} - 7 \equiv 0, 3, 6 \mod 9$.

Par conséquent, les degrés sont susceptibles de coïncider lorsque $\lfloor \log_2 9^r \rfloor = 2k$ tel que $k \equiv 2 \mod 3$. En m'appuyant sur la relation de récurrence (3.1) et les observations sur les constantes d'indices impairs, j'ai donc tenté d'exprimer les coefficients des monômes en fonction des constantes, mais les résultats que j'ai pu obtenir ne m'ont pas permis d'expliquer la forme des coefficients dans le cas général.

3.5 Lorsque $3^r > 2^n - 1$

Une fois le degré maximal atteint, on souhaiterait déterminer si le degré algébrique peut chuter ou s'il reste à sa valeur maximale. La proposition 2.4 nous donne la liste des monômes susceptibles d'apparaître dans le polynôme. On constate qu'après quelques tours supplémentaires, tous les monômes sont susceptibles d'être présents, mais cela n'assure pas le fait que le degré reste à son maximum. Pour des constantes aléatoires, il ne semble pas y avoir de chute de degré (vérifié pour $n = 11, 13, 15, 17$) mais ce dernier pourrait baisser pour certains choix de constantes comme nous avons pu le constater dans la section 3.4.

Chapitre 4

Étude du degré algébrique de la transformation inverse

J'ai aussi pu observer certaines particularités de la fonction de déchiffrement. Tout d'abord, le degré croît rapidement jusqu'à $n-2$, puis il y a un important palier (qui augmente avec la taille du corps) avant le dernier tour, car d'après le corollaire 2.9 le degré maximal n'est atteint que lorsqu'il l'est pour le chiffrement. Sur la figure 4.1 on peut également constater un palier entre le premier et le deuxième tour, comme pour la fonction de chiffrement.

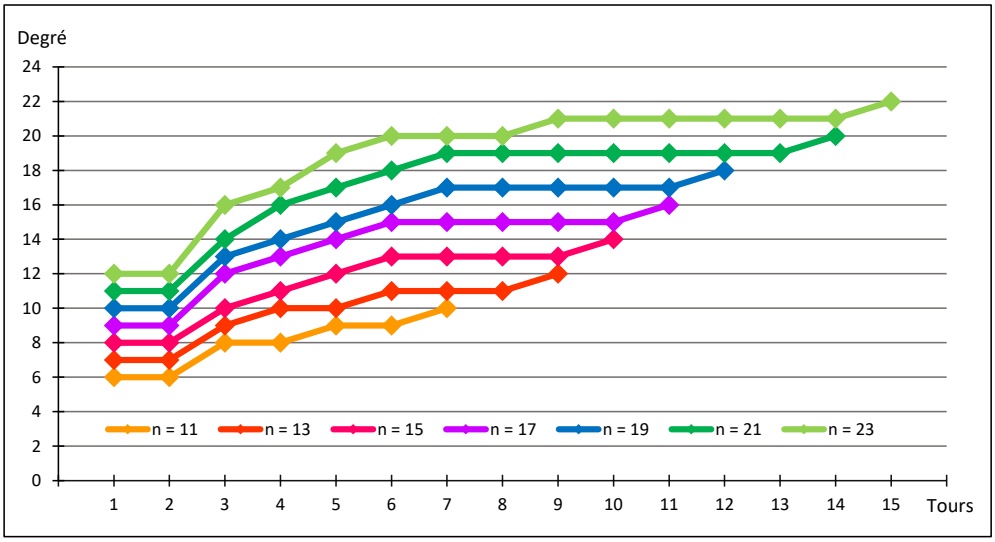

FIGURE 4.1 – Évolution du degré algébrique de la fonction de déchiffrement pour différentes valeurs de n

Dans un premier temps, nous montrerons le palier entre les deux premiers tours dans la section 4.1. Pour cela, nous aurons notamment besoin de démontrer le lemme 4.1 donnant le poids des exposants js en fonction du poids de j. Nous étudierons ensuite, dans la section 4.2, diverses pistes pour tenter de justifier l'évolution du degré sur les tours suivants.

4.1 Palier entre les tours 1 et 2

Cette section est dédiée à la démonstration du palier entre les deux premiers tours pour la transformation inverse. Premièrement, remarquons que $\mathcal{M}_1 = \{s\}$, donc le degré algébrique au premier tour est $wt(s) = (n+1)/2$. On utilise de nouveau la proposition 2.4 avec cette fois $d = s = (2^{n+1}-1)/3$, on a alors :

$$\mathcal{M}_2 = \{sj \bmod (2^n-1) \text{ où } j \preceq s\} = \{(2^{n+1}-1)/3 \times j \bmod (2^n-1) \text{ où } j \preceq (2^{n+1}-1)/3\} .$$

Or

$$j \preceq (2^{n+1}-1)/3 \Longleftrightarrow j \preceq \sum_{k=0}^{(n-1)/2} 2^{2k} \Longleftrightarrow j \in \left\{ \sum_{k=0}^{(n-1)/2} \varepsilon_k 2^{2k}, \varepsilon_k \in \{0,1\} \right\} .$$

Donc les exposants des monômes présents au deuxième tour sont :

$$\begin{aligned} sj &= \varepsilon_0 2^0 + (\varepsilon_0+\varepsilon_1)2^2 + \ldots + \left(\sum_{k=0}^{(n-1)/2} \varepsilon_k\right) 2^{n-1} + \left(\sum_{k=1}^{(n-1)/2} \varepsilon_k\right) 2^{n+1} + \ldots + \varepsilon_{(n-1)/2} 2^{2n-2} \\ &\equiv \varepsilon_0 2^0 + \left(\sum_{k=1}^{(n-1)/2} \varepsilon_k\right) 2^1 + (\varepsilon_0+\varepsilon_1)2^2 + \ldots + \varepsilon_{(n-1)/2} 2^{n-2} + \left(\sum_{k=0}^{(n-1)/2} \varepsilon_k\right) 2^{n-1} \mod 2^n - 1 . \end{aligned}$$

Afin de montrer la présence d'un palier, il convient de s'assurer que le degré algébrique au tour 2, est également $(n+1)/2$, c'est-à-dire que tous les éléments de $\mathcal{M}_2$ sont tels que $wt(sj \bmod (2^n-1)) \leq (n+1)/2$.

Tout d'abord, on constate qu'il suffit de regarder le poids d'un élément pour chaque coset cyclotomique, car si s est une permutation, on a :

$$\forall u, 1 \leq u \leq 2^n-1, (2^i u \bmod (2^n-1))s \bmod (2^n-1) = 2^i (us \bmod (2^n-1)) \bmod (2^n-1) . \tag{4.1}$$

Regardons le poids de js en fonction du poids de j :

— si $wt(j) = 0$: il est évident que $wt(sj) = 0$.

— si $wt(j) = 1$: comme tous les exposants $j = 2^i$, $i \in \{0, 2, \ldots, n-1\}$ appartiennent au même coset cyclotomique, il suffit de regarder le résultat obtenu pour l'un d'eux. Or si $j = 1$, $wt(sj) = wt(s) = \frac{n+1}{2}$. Donc $\forall j$ tel que $wt(j) = 1$, on a $wt(sj) = \frac{n+1}{2}$.

Lemme 4.1. *Soit $j \preceq s$. Alors pour tout j tel que $wt(j) \geq 2$, on a :*

$$wt(js \bmod 2^n - 1) \in \begin{cases} [wt(j)-1, (n-1)/2] \text{ si } wt(j) \equiv 2 \mod 3 \\ [wt(j), (n+1)/2] \text{ sinon} . \end{cases}$$

Démonstration. Dans ce raisonnement, nous utiliserons notamment que pour tous entiers pairs $i_1 < i_2 < i_3$:

$$s(2^{i_1} + 2^{i_2} + 2^{i_3}) = (2^{i_1} + 2^{i_2} + 2^{i_3})/3 = 2^{i_1} + \sum_{l=i_1/2}^{(i_2-2)/2} 2^{2l+1} + \sum_{l=i_2/2}^{(i_3-2)/2} 2^{2l} \mod 2^n - 1 . \tag{4.2}$$

car

$$\begin{aligned} 3 \times \left(2^{i_1} + \sum_{l=i_1/2}^{(i_2-2)/2} 2^{2l+1} + \sum_{l=i_2/2}^{(i_3-2)/2} 2^{2l}\right) &= 2^{i_1} + \sum_{l=i_1/2}^{(i_2-2)/2} 2^{2l+1} + \sum_{l=i_2/2}^{(i_3-2)/2} 2^{2l} + 2^{i_1+1} + \sum_{l=i_1+2/2}^{i_2/2} 2^{2l} + \sum_{l=i_2/2}^{(i_3-2)/2} 2^{2l+1} \\ &= 2^{i_1} + 2^{i_2} + 2^{i_3} . \end{aligned}$$

Par ailleurs, considérons un entier j tel que $wt(j) \equiv 2 \mod 3$, avec $wt(j) = 2+3k$, et $j = 1 + \sum_{m=1}^{3k+1} 2^{i_m}$, où les i_m sont pairs car $j \preceq s$, et $2 \leq i_1 < \ldots < i_{3k+1} \leq n-1$, alors montrons que

$$\begin{aligned} sj = 2^{i_1} + \sum_{l=i_1/2}^{(i_2-2)/2} 2^{2l+1} + \sum_{m=1}^{k-1} \left(\sum_{l=i_{3m-1}/2}^{(i_{3m}-2)/2} 2^{2l} + 2^{i_{3m+1}} + \sum_{l=i_{3m+1}/2}^{(i_{3m+2}-2)/2} 2^{2l+1} \right) \\ + \sum_{l=i_{3k-1}/2}^{(i_{3k}-2)/2} 2^{2l} + 2^{i_{3k+1}} + \sum_{l=i_{3k+1}/2}^{(n-3)/2} 2^{2l+1} \mod 2^n - 1 , \end{aligned} \tag{4.3}$$

et donc $wt(sj \bmod 2^n - 1) = (n + 2k + 1 - i_1 + \sum_{m=1}^{k} i_{3m} - i_{3m+1})/2$.

— pour $wt(j)=2$, si $j=1+2^{i_1}$, $i_1\in\{2,\dots,n-1\}$, alors on a :

$$sj=2^{i_1}+\sum_{k=i_1/2}^{(n-3)/2}2^{2k+1}\mod 2^n-1\quad\text{donc}\quad wt(sj \bmod 2^n-1)=(n-i_1+1)/2\ .$$

— soit j_0 tel que $wt(j_0)=2+3k$ et $j_0=1+\sum_{m=1}^{3k+1}2^{i_m}$, alors par hypothèse, sj_0 vérifie (4.3). Alors si $wt(j)=2+3(k+1)$ tel $j=j_0+2^{i_{3k+2}}+2^{i_{3k+3}}+2^{i_{3k+4}}$, on a :

$$\begin{aligned}sj&=sj_0+s(2^{i_{3k+1}}+2^{i_{3k+2}}+2^{i_{3(k+1)}})\\&=2^{i_1}+\sum_{l=i_1/2}^{(i_2-2)/2}2^{2l+1}+\sum_{m=1}^{k-1}\left(\sum_{l=i_{3m-1}/2}^{(i_{3m}-2)/2}2^{2l}+2^{i_{3m+1}}+\sum_{l=i_{3m+1}/2}^{(i_{3m+2}-2)/2}2^{2l+1}\right)+\sum_{l=i_{3k-1}/2}^{(i_{3k}-2)/2}2^{2l}\\&+2^{i_{3k+1}}+\sum_{l=i_{3k+1}/2}^{(n-3)/2}2^{2l+1}+\left(2^{i_{3k+1}}+\sum_{l=i_{3k+1}/2}^{(i_{3k+2}-2)/2}2^{2l+1}+\sum_{l=i_{3k+2}/2}^{(i_{3k+3}-2)/2}2^{2l}\right)\mod 2^n-1\\&=2^{i_1}+\sum_{l=i_1/2}^{(i_2-2)/2}2^{2l+1}+\sum_{m=1}^{k}\left(\sum_{l=i_{3m-1}/2}^{(i_{3m}-2)/2}2^{2l}+2^{i_{3m+1}}+\sum_{l=i_{3m+1}/2}^{(i_{3m+2}-2)/2}2^{2l+1}\right)+\sum_{l=i_{3k+2}/2}^{(i_{3k+3}-2)/2}2^{2l}\\&+2^{i_{3(k+1)+1}}+\sum_{l=i_{3k+4}/2}^{(n-3)/2}2^{2l+1}\mod 2^n-1\ ,\end{aligned}$$

et dans ce cas, le poids de sj est :

$$\begin{aligned}wt(sj \bmod 2^n-1)&=1+\left(\frac{i_2-2-i_1}{2}+1\right)+\sum_{m=1}^{k}\left(\left(\frac{i_{3m}-2-i_{3m-1}}{2}+1\right)+1+\left(\frac{i_{3m+2}-2-i_{3m+1}}{2}+1\right)\right)\\&+\left(\frac{i_{3k+3}-2-i_{3k+2}}{2}+1\right)+1+\left(\frac{n-3-i_{3k+4}}{2}+1\right)\\&=\left(n+2k+3-i_1+\sum_{m=1}^{k+1}i_{3m}-i_{3m+1}\right)/2\ .\end{aligned}$$

Donc si $wt(j)\equiv 2 \mod 3$, on a bien $wt(sj \bmod 2^n-1)=(n+2k+1-i_1+\sum_{m=1}^{k}i_{3m}-i_{3m+1})/2\in[wt(j)-1,(n-1)/2]$, car, par définition des i_m, on a $i_1+2(m-1)\le i_m\le n-1-2(3k+1-m)$.

Si $wt(j)\equiv 0 \mod 3$, tel que $wt(j)=3k$, et $j=1+\sum_{m=1}^{3k-1}2^{i_m}$, où les i_m sont pairs, et $2\le i_1<\ldots<i_{3k-1}\le n-1$, alors, en utilisant le même raisonnement que précédemment, montrons que

$$sj=1+\sum_{l=0}^{(i_1-2)/2}2^{2l+1}+\sum_{l=i_1/2}^{(i_2-2)/2}2^{2l}+\sum_{m=1}^{k-1}\left(2^{i_{3m}}+\sum_{l=i_{3m}/2}^{(i_{3m+1}-2)/2}2^{2l+1}+\sum_{l=i_{3m+1}/2}^{(i_{3m+2}-2)/2}2^{2l}\right)\mod 2^n-1\ ,\quad(4.4)$$

et donc $wt(sj \bmod 2^n-1)=(2k+i_2+\sum_{m=1}^{k-1}-i_{3m}+i_{3m+2})/2$.

— pour $wt(j)=3$, si $j=1+2^{i_1}+2^{i_2}$, $i_1\in\{2,\dots,n-3\}$, $i_2\in\{i_1+2,\dots,n-1\}$, on a :

$$sj=1+\sum_{l=0}^{(i_1-2)/2}2^{2l+1}+\sum_{l=i_1/2}^{(i_2-2)/2}2^{2l}\mod 2^n-1\quad\text{donc}\quad wt(sj \bmod 2^n-1)=(i_2+2)/2\ .$$

— soit j_0 tel que $wt(j_0)=3k$ et $j_0=1+\sum_{m=1}^{3k-1}2^{i_m}$ alors par hypothèse, sj_0 vérifie (4.4). Alors si $wt(j)=3(k+1)$ tel que $j=j_0+2^{i_{3k}}+2^{i_{3k+1}}+2^{i_{3k+2}}$, on a :

$$sj=1+\sum_{l=0}^{(i_1-2)/2}2^{2l+1}+\sum_{l=i_1/2}^{(i_2-2)/2}2^{2l}+\sum_{m=1}^{k}\left(2^{i_{3m}}+\sum_{l=i_{3m}/2}^{(i_{3m+1}-2)/2}2^{2l+1}+\sum_{l=i_{3m+1}/2}^{(i_{3m+2}-2)/2}2^{2l}\right)\mod 2^n-1\ .$$

Donc si $wt(j) \equiv 0 \mod 3$, on a bien $wt(sj \bmod 2^n-1) = (2k+i_2+\sum_{m=1}^{k-1} -i_{3m}+i_{3m+2})/2 \in [wt(j),(n+1)/2]$.

Si $wt(j) \equiv 1 \mod 3$, tel que $wt(j) = 1+3k$, et $j = 1+\sum_{m=1}^{3k} 2^{i_m}$, où les i_m sont pairs, et $2 \leq i_1 < \ldots < i_{3k} \leq n-1$, alors montrons que

$$sj = 1 + \sum_{l=1}^{(i_1-2)/2} 2^{2l} + \sum_{m=1}^{k-1} \left(2^{i_{3m-1}} + \sum_{l=i_{3m-1}/2}^{(i_{3m}-2)/2} 2^{2l+1} + \sum_{l=i_{3m}}^{(i_{3m+1}-2)/2} 2^{2l} \right) + 2^{i_{3k-1}} + \sum_{l=i_{3k-1}/2}^{(i_{3k}-2)/2} 2^{2l+1} + \sum_{l=i_{3k}}^{(n-1)/2} 2^{2l} \mod 2^n-1 \,, \quad (4.5)$$

et donc $wt(sj \bmod 2^n-1) = (n+2k+1+\sum_{m=1}^{k} i_{3m-2}-i_{3m-1})/2$.

— pour $wt(j)=4$, si $j = 1+2^{i_1}+2^{i_2}+2^{i_3}$, $i_1 \in \{2,\ldots,n-5\}$, $i_2 \in \{i_1+2,\ldots,n-3\}$, $i_3 \in \{i_2+2,\ldots,n-1\}$:

$$sj = 1+\sum_{l=1}^{(i_1-2)/2} 2^{2l}+2^{i_2}+\sum_{l=i_2/2}^{(i_3-2)/2} 2^{2l+1}+\sum_{l=i_3/2}^{(n-1)/2} 2^{2l} \mod 2^n-1 \quad \text{donc} \quad wt(sj \bmod 2^n-1) = (n+i_1-i_2+3)/2 \,.$$

— soit j_0 tel que $wt(j_0)=3k$ et $j_0 = 1+\sum_{m=1}^{3k} 2^{i_m}$ alors par hypothèse, sj_0 vérifie (4.5). Alors si $wt(j) = 1+3(k+1)$ tel que $j = j_0 + 2^{i_{3k+1}} + 2^{i_{3k+2}} + 2^{i_{3(k+1)}}$, on a :

$$sj = 1 + \sum_{l=1}^{(i_1-2)/2} 2^{2l} + \sum_{m=1}^{k} \left(2^{i_{3m-1}} + \sum_{l=i_{3m-1}/2}^{(i_{3m}-2)/2} 2^{2l+1} + \sum_{l=i_{3m}}^{(i_{3m+1}-2)/2} 2^{2l} \right) + 2^{i_{3k+2}} + \sum_{l=i_{3k+2}/2}^{(i_{3k+3}-2)/2} 2^{2l+1} + \sum_{l=i_{3k+3}}^{(n-1)/2} 2^{2l} \mod 2^n-1 \,.$$

Donc si $wt(j) \equiv 1 \mod 3$, on a $wt(sj \bmod 2^n-1) = (n+2k+1+\sum_{m=1}^{k} i_{3m-2}-i_{3m-1})/2 \in [wt(j),(n+1)/2]$. □

Par conséquent le degré algébrique au deuxième tour est également $(n+1)/2$, et il y a ainsi un palier entre les deux premiers tours pour le déchiffrement.

Dans la section 3.3, nous avions notamment pu constater la présence d'un palier entre les deux premiers tours pour les permutations x^d pour $d = 2^k-1$. En revanche, il n'y a pas nécessairement de palier entre les deux premiers tours pour l'inverse de ces permutations. Par exemple, dans $\mathbb{F}_{2^{11}}$, $15 = 2^4-1$ et il y a donc un palier entre les tours 1 et 2 pour le chiffrement, mais pour $15^{-1} = 273$, le degré algébrique au premier tour est $wt(273) = 3$, et au deuxième tour, il est de 5 car on a notamment $wt(273 \times 273 \mod 2^n-1) = 5$.

4.2 Diverses pistes pour les tours suivants

Le degré de la fonction inverse étant beaucoup plus élevé que celui de x^3, étudier l'évolution du degré du déchiffrement au fil des itérations est beaucoup plus difficile que pour le chiffrement. Dans cette partie, différentes approches sont analysées pour tenter d'expliquer l'évolution du degré algébrique pour les tours suivants.

Nous verrons notamment dans la sous-section 4.2.1 comment les propriétés des fonctions presque courbes peuvent appuyer l'étude du degré de la transformation inverse. Dans la sous-section 4.2.2, nous examinerons ensuite l'intérêt de l'étude du degré du chiffrement pour expliquer le comportement de la transformation inverse. Enfin, nous proposerons une autre approche consistant à exprimer le poids des exposants js en fonction du poids des exposants j, dans la sous-section 4.2.3.

4.2.1 Particularité des fonctions AB

Une des caractéristiques des fonctions puissances presque courbes est l'existence d'une borne sur le poids des exposants js où $0 \leq j \leq 2^n-1$, modulo 2^n-1. En particulier, au premier tour, on a $\mathcal{M}_1 = \{s\}$, par conséquent pour les tours suivants on aura notamment les monômes d'exposant $j \preceq s$. Or, la proposition 2.11 nous donne :

$\forall j \preceq s$, $wt(js \bmod (2^n-1)) \leq k + wt(j) \leq k + wt(s) = \frac{n-1}{2} + \frac{n+1}{2} = n$, mais cela ne donne pas de nouvelles informations, car les opérations sont effectuées modulo $2^n - 1$.

En revanche, cela peut par exemple permettre de diminuer le nombre de monômes à étudier. En effet, à partir de la connaissance du degré au tour $r-1$, on peut alors affiner le nombre de monômes x^{js} à étudier pour déterminer le degré au tour r.

Par exemple au deuxième tour, le degré est de $\frac{n+1}{2}$ et il semble qu'au troisième tour, le degré soit $\frac{n+1}{2} + \lfloor \frac{n+1}{6} \rfloor$ d'après le tableau 4.1 :

n	7	9	11	13	15	17	19	21	23	25	27
degré au tour 3	5	6	8	9	10	12	13	14	16	17	18

TABLE 4.1 – Degré algébrique de la transformation inverse au troisième tour

Donc, pour $wt(j) \leq \lfloor \frac{n+1}{6} \rfloor + 1$, la relation précédente, nous donne bien

$$wt(js \bmod (2^n-1)) \leq \frac{n-1}{2} + wt(j) \leq \frac{n+1}{2} + \left\lfloor \frac{n+1}{6} \right\rfloor .$$

Il s'agit alors de regarder les éléments j tels que $\lceil \frac{n+1}{6} \rceil + 1 \leq wt(j) \leq \frac{n+1}{2}$. De plus d'après la relation (4.1) il suffit de regarder uniquement un élément par coset cyclotomique.

Cette borne ne semble néanmoins pas assez fine pour donner des informations précises sur l'évolution du degré algébrique de la transformation inverse.

4.2.2 Influence du degré de la fonction de chiffrement

Une autre piste de réflexion est d'étudier l'influence du degré de la fonction de chiffrement, sur le degré de la fonction de déchiffrement. Le théorème 2.8 nous donne effectivement un lien entre le degré de la fonction et le degré de son inverse. D'une part, l'étude précédemment réalisée pour le chiffrement nous donne des informations sur $\delta_k(F)$ où k suit l'évolution du degré de $F : x \mapsto x^3$. On rappelle que δ_k correspond au degré maximal du produit d'au plus k coordonnées de F. Pour les premiers tours, on obtient par exemple :

$$\begin{aligned}
(x^3+c_1)^3 &= G_1 \circ F(x) \quad \text{où} \quad G_1 = (x+c_1)^3 \quad \text{donc} \quad \deg G_1 \circ F = 2 \leq \delta_{\deg G_1}(F) = \delta_2(F) \\
((x^3+c_1)^3+c_2)^3 &= G_2 \circ F(x) \quad \text{où} \quad G_2 = (G_1+c_2)^3 \quad \text{donc} \quad \deg G_2 \circ F = 4 \leq \delta_{\deg G_2}(F) = \delta_2(F) \\
(((x^3+c_1)^3+c_2)^3+c_3)^3 &= G_3 \circ F(x) \quad \text{où} \quad G_3 = (G_2+c_3)^3 \quad \text{donc} \quad \deg G_3 \circ F = 4 \leq \delta_{\deg G_3}(F) = \delta_4(F) \\
((((x^3+c_1)^3+c_2)^3+c_3)^3+c_4)^3 &= G_4 \circ F(x) \quad \text{où} \quad G_4 = (G_3+c_4)^3 \quad \text{donc} \quad \deg G_4 \circ F = 6 \leq \delta_{\deg G_4}(F) = \delta_4(F)
\end{aligned}$$

L'évolution du degré du chiffrement, étudiée au chapitre 3, nous donne ainsi les degrés des $G_i \circ F$. En effet, le polynôme $G_i \circ F(x)$ correspond exactement au polynôme décrivant le tour $i+1$ de MiMC. On remarque par ailleurs que le degré des G_i correspond au degré des $G_{i-1} \circ F$, car comme nous l'avons vu dans la section 2.2 le degré algébrique de $F(x+c)$ et $F(x)$ est identique pour toute fonction polynomiale F. Cela nous donne donc une borne inférieure sur les $\delta_k(F)$. De plus, on a également une borne supérieure triviale : $\delta_k(F) \leq k \deg F$. On obtient alors le tableau 4.2.

k	1	2	4	6	8	10	...	K
$\delta_k(F)$	2	4	$\in [6,8]$	$\in [8,12]$	$\in [10,16]$	$\in [12,20]$	...	$\in [K+2, 2K]$

TABLE 4.2 – Degré maximal du produit d'au plus k coordonnées de F

D'autre part, pour le déchiffrement, on a :

$$\begin{aligned}
x^s &= F^{-1}(x) \\
(x^s+c_{r-1})^s &= H_2 \circ F^{-1}(x) \quad \text{où } H_2 = (x+c_{r-1})^s \quad \text{donc} \quad \deg H_2 \circ F^{-1} = \tfrac{n+1}{2} \leq \delta_{\deg H_2}(F^{-1}) = \delta_{\frac{n+1}{2}}(F^{-1}) \\
((x^s+c_{r-1})^s+c_{r-2})^s &= H_3 \circ F^{-1}(x) \quad \text{où } H_3 = (H_2+c_{r-2})^s \quad \text{donc} \quad \deg H_3 \circ F^{-1} \leq \delta_{\deg H_3}(F^{-1}) = \delta_{\frac{n+1}{2}}(F^{-1})
\end{aligned}$$

On peut donc, par exemple, obtenir une borne pour le degré au troisième tour. En effet, on a :

$$\delta_k(F) < 2k+1 \Leftrightarrow \delta_{n-(2k+1)}(F^{-1}) < n-k \qquad \text{et} \qquad \delta_k(F) \geq k+2 \Leftrightarrow \delta_{n-(k+2)}(F^{-1}) \geq n-k$$

Or comme le degré de F^{-1} est $(n+1)/2$, la plus petite valeur de l telle que $\delta_l(F^{-1})$ soit utile est $l = (n+1)/2$. En effet, on peut voir dans les lignes précédentes que pour le déchiffrement, les $\delta_l(F^{-1})$ qui interviennent sont tels que $l \geq (n+1)/2$ car l suit l'évolution du degré du déchiffrement. Pour cette valeur de l, k vérifie alors : $k \leq (n-3)/4$ pour $\delta_{n-(2k+1)}(F^{-1}) < n-k$ et $k \leq (n-5)/2$ pour $\delta_{n-(k+2)}(F^{-1}) \geq n-k$. Pour le troisième tour en particulier, on s'intéresse aux k tels que $n-(2k+1) = (n+1)/2$. De plus, k suit l'évolution du degré algébrique du chiffrement, donc k est pair. On obtient alors une borne pour les valeurs de n telles que $n \equiv 3 \mod 8$. Soit d le degré algébrique au troisième tour, on a par exemple :

— pour $n = 11$: $d \leq 8$ (le degré observé est 8)
— pour $n = 19$: $d \leq 14$ (le degré observé est 13)
— pour $n = 27$: $d \leq 20$ (le degré observé est 18)

L'étude réalisée pour le degré de la fonction de chiffrement nous donne ainsi une idée du degré pour la fonction de déchiffrement, mais une nouvelle fois, cela ne nous donne pas son évolution précise.

4.2.3 $wt(js)$ en fonction de $wt(j)$

Dans cette sous-section, nous tentons, comme dans la section 4.1, d'établir un lien entre le poids des exposants js et le poids des exposants j.

Tout d'abord, on note que pour $wt(j) = 0, 1, n-1$, on a respectivement $wt(js) = 0, \frac{n+1}{2}, \frac{n-1}{2}$. Les cas $wt(j) = 0, 1$ ont été traités dans la démonstration du palier entre les tours 1 et 2. De plus, si $wt(j) = n-1$ alors $j = 2^n - 2^i - 1$ où $0 \leq i \leq n-1$. Or $(2^n - 2^i - 1)s = -2^i s \mod (2^n-1)$, donc on a :

$$(2^n - 2^i - 1)s = \begin{cases} -\sum_{l=i/2}^{(n+i-1)/2} 2^{2l} = \sum_{l=0}^{(i-2)/2} 2^{2l} + \sum_{l=i/2}^{(n-3)/2} 2^{2l+1} \mod (2^n-1) \quad \text{si} \quad i \equiv 0 \mod 2 \\ -\sum_{l=(i-1)/2}^{(n+i-2)/2} 2^{2l+1} = \sum_{l=0}^{(i-3)/2} 2^{2l+1} + \sum_{l=(i+1)/2}^{(n-1)/2} 2^{2l} \mod (2^n-1) \quad \text{si} \quad i \equiv 1 \mod 2 \end{cases}$$

Dans tous les cas, on a bien $wt(js \mod 2^n-1) = wt((2^n - 2^i - 1)s \mod 2^n - 1) = \frac{n-1}{2}$.

Par ailleurs, en observant le poids de js en fonction du poids de j, j'ai pu remarquer :

Conjecture 4.2. *Si $2 \leq j \leq 2^n - 2$, alors :*

$$wt(js \bmod 2^n - 1) \in \begin{cases} [k, (n+2k-3)/2] & si \ \ wt(j) = 2k \\ [k+2, (n+2k+1)/2] & si \ \ wt(j) = 2k+1 \ . \end{cases}$$

Cela affinerait donc la borne caractérisant les fonctions presque courbes qui était $wt(js \bmod (2^n-1)) \leq \frac{n-1}{2} + wt(j)$. Un début de démonstration est proposé en annexe B.2.

L'ensemble des résultats présentés dans ce chapitre ne permettent pas de déterminer de manière exacte le degré algébrique au tour r mais ils peuvent néanmoins permettre de donner une approximation du degré (ou une borne) connaissant le degré au tour $r-1$, ou l'évolution du degré pour la fonction de chiffrement.

Chapitre 5

Conclusion

Pendant ce stage, j'ai ainsi pu réaliser une première évaluation de la sécurité du chiffrement MiMC en analysant l'évolution du degré algébrique des polynômes représentants MiMC et son inverse. J'ai notamment pu constater des comportements anormaux, en particulier la présence de paliers dans l'évolution du degré de la fonction de chiffrement. Les simulations que j'ai effectuées semblaient, en effet, montrer que le degré évoluait en fonction du nombre de tours en $2 \times \lceil \lfloor r \log_2 3 \rfloor / 2 - 1 \rceil$. J'ai alors tenté d'établir une preuve de ce comportement, en essayant de déterminer les exposants de poids maximal. Toutefois, certains cas particuliers ne m'ont pas permis de conclure et d'assurer le résultat pour le cas général.

Par ailleurs, j'ai également commencé l'analyse du degré de la transformation inverse. J'ai pu démontrer la présence d'un palier entre les deux premiers tours, mais le comportement pour les tours suivants reste à déterminer. En effet, j'ai identifié diverses pistes et exploré certaines d'entre elles mathématiquement, mais cela ne m'a pas encore permis de déterminer de manière précise l'évolution du degré.

Ayant obtenu une bourse pour poursuivre en thèse sur un sujet plus large, j'aurai ainsi l'opportunité de pouvoir expliquer plus en détails ces comportements. L'objectif de la thèse est, en effet, d'élargir le travail déjà effectué avec le chiffrement MiMC, et plus généralement, de comprendre comment les spécificités des nombreuses primitives employant des fonctions de faible degré sur un corps fini, influent sur leur sécurité.

Il existe un grand nombre de primitives, dont plusieurs variantes de MiMC [AGP+19, GLR+20, AAB+19], c'est pourquoi, l'analyse de la structure algébrique univariée de MiMC, commencée durant ce stage, sera une des pistes de recherche pour étudier la résistance de ces algorithmes à des attaques classiques. Une autre piste qui pourra être explorée sera de rechercher de nouvelles méthodes d'attaques, exploitant la forme particulière du polynôme univarié décrivant la transformation.

Par ailleurs, dans le travail réalisé au cours de ce stage seules les transformations dans des corps finis de la forme $\mathbb{F}_{2^n}$ ont été considérées, mais il est également possible de réaliser les opérations dans des corps premiers. Il existe effectivement des variantes de MiMC définies sur des corps premiers et les outils de cryptanalyse pour les primitives définies sur ce type de corps restent entièrement à découvrir.

Comme nous l'avons vu en première partie de ce rapport, il est important de pouvoir analyser la sécurité de ces primitives récemment proposées, car elles sont notamment en cours de déploiement dans la Blockchain. Il s'agit ainsi de pouvoir apporter des réponses rapides et précises aux acteurs industriels travaillant dans ce domaine.

Bibliographie

[AAB+19] Abdelrahaman Aly, Tomer Ashur, Eli Ben-Sasson, Siemen Dhooghe, and Alan Szepieniec. Design of symmetric-key primitives for advanced cryptographic protocols. Cryptology ePrint Archive, Report 2019/426, 2019. https://eprint.iacr.org/2019/426.

[AGP+19] Martin R. Albrecht, Lorenzo Grassi, Léo Perrin, Sebastian Ramacher, Christian Rechberger, Dragos Rotaru, Arnab Roy, and Markus Schofnegger. Feistel structures for MPC, and more. In Kazue Sako, Steve Schneider, and Peter Y. A. Ryan, editors, *ESORICS 2019, Part II*, volume 11736 of *LNCS*, pages 151–171. Springer, Heidelberg, September 2019.

[AGR+16] Martin R. Albrecht, Lorenzo Grassi, Christian Rechberger, Arnab Roy, and Tyge Tiessen. MiMC : Efficient encryption and cryptographic hashing with minimal multiplicative complexity. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *ASIACRYPT 2016, Part I*, volume 10031 of *LNCS*, pages 191–219. Springer, Heidelberg, December 2016.

[BBHR18] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Scalable, transparent, and post-quantum secure computational integrity. Cryptology ePrint Archive, Report 2018/046, 2018. https://eprint.iacr.org/2018/046.

[BC13] Christina Boura and Anne Canteaut. On the Influence of the Algebraic Degree of F^{-1} on the Algebraic Degree of $G \circ F$. *IEEE Trans. Inf. Theory*, 59(1) :691–702, 2013.

[BCG+13] Eli Ben-Sasson, Alessandro Chiesa, Daniel Genkin, Eran Tromer, and Madars Virza. SNARKs for C : Verifying program executions succinctly and in zero knowledge. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part II*, volume 8043 of *LNCS*, pages 90–108. Springer, Heidelberg, August 2013.

[CCD99] Anne Canteaut, Pascale Charpin, and Hans Dobbertin. A New Characterization of Almost Bent Functions. In *Fast Software Encryption, 6th International Workshop, FSE '99, Rome, Italy, March 24-26, 1999, Proceedings*, volume 1636 of *Lecture Notes in Computer Science*, pages 186–200. Springer, 1999.

[DR02] Joan Daemen and Vincent Rijmen. *The Design of Rijndael : AES - The Advanced Encryption Standard*. Springer Verlag, Berlin, Heidelberg, New York, 2002.

[EGL+20] Maria Eichlseder, Lorenzo Grassi, Reinhard Lüftenegger, Morten Øygarden, Christian Rechberger, Markus Schofnegger, and Qingju Wang. An Algebraic Attack on Ciphers with Low-Degree Round Functions : Application to Full MiMC. Cryptology ePrint Archive, Report 2020/182, 2020. https://eprint.iacr.org/2020/182.

[GLR+20] Lorenzo Grassi, Reinhard Lüftenegger, Christian Rechberger, Dragos Rotaru, and Markus Schofnegger. On a generalization of substitution-permutation networks : The HADES design strategy. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020, Part II*, volume 12106 of *LNCS*, pages 674–704. Springer, Heidelberg, May 2020.

Annexe A

Compléments sur le degré algébrique de MiMC

A.1 Exposants de poids maximal

Cette partie est consacrée à la justification de la présence des exposants de poids maximal pour chaque tour r. Pour cela, on utilisera une récurrence forte en supposant la présence des exposants cités dans la sous-section 3.2.2 pour chaque tour, jusqu'au tour $r-1$. Tout d'abord, on peut effectivement vérifier sur les premiers tours la présence de ces exposants. Ensuite pour montrer l'hérédité, on distinguera 6 cas en fonction de la parité de $\lfloor r \log_2 3 \rfloor$ et des tours précédents :

1. Cas I-P : lorsque $\lfloor r \log_2 3 \rfloor = 2k$ et $\lfloor (r-1) \log_2 3 \rfloor = 2k-1$, on a $2^{2k-1}-5, 2^{2k-1}-2 \in \mathcal{N}_r$ car ils sont présents au tour $r-1$. De plus, $2^{2k}-7, 2^{2k}-4 \in \mathcal{N}_r$ car

$$(2^{2k}-7)/3 = 1+2+\sum_{i=2}^{k-1} 2^{2i} \preceq 1+2+\sum_{i=3}^{2k-2} 2^i = 2^{2k-1}-5 \quad \text{et} \quad (2^{2k}-4)/3 = \sum_{i=1}^{k-1} 2^{2i} \preceq \sum_{i=1}^{2k-2} 2^i = 2^{2k-1}-2 \ ,$$

et comme nous venons de le voir $2^{2k-1}-5, 2^{2k-1}-2 \in \mathcal{N}_{r-1}$.
On a aussi la présence des monômes d'exposant $2^{2k}-2^j-5$ et $2^{2k}-2^j-2$ pour $j = 2i+1$ où $1 \leq i \leq k-2$ et des monômes d'exposant $2^{2k+1}-2^{2k-1}-2^j-5, 2^{2k+1}-2^{2k-1}-2^j-2$ pour $j = 2i$ où $2 \leq i \leq k-1$ car :

$$\begin{aligned}
(2^{2k}-2^j-5)/3 &= 1+\sum_{i=1}^{(j-3)/2} 2^{2i+1} + \sum_{i=(j+1)/2}^{k-1} 2^{2i} && \preceq 2^{2k-1}-5 \ , \\
(2^{2k}-2^j-2)/3 &= \sum_{i=0}^{(j-3)/2} 2^{2i+1} + \sum_{i=(j+1)/2}^{k-1} 2^{2i} && \preceq 2^{2k-1}-5, 2^{2k-1}-2 \ , \\
(2^{2k+1}-2^{2k-1}-2^j-5)/3 &= 1+\sum_{i=1}^{(j-2)/2} 2^{2i+1} + 2^j + \sum_{i=j+1}^{2k-2} 2^i && \preceq 2^{2k-1}-5 \ , \\
(2^{2k+1}-2^{2k-1}-2^j-2)/3 &= \sum_{i=0}^{(j-2)/2} 2^{2i+1} + 2^j + \sum_{i=j+1}^{2k-2} 2^i && \preceq 2^{2k-1}-5, 2^{2k-1}-2 \ .
\end{aligned}$$

Enfin, $2^{2k+1}-2^{2k-1}-2^2-2 = 3 \times (2^{2k-1}-2) \in \mathcal{N}_r$.

2. Cas P-I : si $\lfloor r \log_2 3 \rfloor = 2k+1$ et $\lfloor (r-1) \log_2 3 \rfloor = 2k$, alors $2^{2k+1}-5 \in \mathcal{N}_r$ car

$$(2^{2k+1}-5)/3 = 1+\sum_{i=1}^{k-1} 2^{2i+1} \preceq 1+\sum_{i=3}^{2k-1} 2^i = 2^{2k}-7 \in \mathcal{N}_{r-1} \ .$$

De plus, $2^{2k+1}-2 \in \mathcal{N}_r$ car $2^{2k-2}-7 \in \mathcal{N}_{r-2}$:

$$(2^{2k+1}-2)/3 = \sum_{i=0}^{k-1} 2^{2i+1} \preceq 1+2+2^3+\left(\sum_{i=5}^{2k-3} 2^i\right)+2^{2k-1} = (2^{2k-2}-7) \times 3 \ .$$

D'autres monômes ont pu être observés dans la sous-section 3.2.2 mais leur présence reste à déterminer.

3. Cas P-I-I : lorsque $\lfloor r\log_2 3\rfloor = 2k+1$, $\lfloor (r-1)\log_2 3\rfloor = 2k-1$ et $\lfloor (r-2)\log_2 3\rfloor = 2k-2$, on a $2^{2k+1}-5, 2^{2k+1}-2 \in \mathcal{N}_r$ car $2^{2k-2}-7 \in \mathcal{N}_{r-2}$:

$$(2^{2k+1}-5)/3 = 1+\sum_{i=1}^{k-1} 2^{2i+1} \preceq (2^{2k-2}-7)\times 3 \quad \text{et} \quad (2^{2k+1}-2)/3 = \sum_{i=0}^{k-1} 2^{2i+1} \preceq (2^{2k-2}-7)\times 3 \ .$$

4. Cas I-P-P : si $\lfloor r\log_2 3\rfloor = 2k$, $\lfloor (r-1)\log_2 3\rfloor = 2k-2$ et $\lfloor (r-2)\log_2 3\rfloor = 2k-3$, on a $2^{2k-1}-5 \in \mathcal{N}_r$ car

$$(2^{2k-1}-5)/3 = 1+\sum_{i=1}^{k-2} 2^{2i+1} \preceq 1+\sum_{i=3}^{2k-3} 2^i = 2^{2k-2}-7 \in \mathcal{N}_{r-1} \ .$$

De plus, $2^{2k}-7 \in \mathcal{N}_r$ car $2^{2k-3}-5 \in \mathcal{N}_{r-2}$. En effet, on a :

$$(2^{2k}-7)/3 = 1+2+\sum_{i=2}^{k-1} 2^{2i} \preceq 1+2+2^4+\left(\sum_{i=6}^{2k-4} 2^i\right)+2^{2k-2} \ .$$

$$\text{Or} \qquad \left(1+2+2^4+\left(\sum_{i=6}^{2k-4} 2^i\right)+2^{2k-2}\right)/3 = 1+\sum_{i=4}^{2k-4} 2^i \preceq 1+2+\sum_{i=3}^{2k-4} 2^i = 2^{2k-3}-5 \ .$$

On a aussi $2^{2k-1}-2, 2^{2k}-4 \in \mathcal{N}_r$ car $2^{2k-3}-2 \in \mathcal{N}_{r-2}$:

$$(2^{2k-1}-2)/3 = \sum_{i=0}^{k-2} 2^{2i+1} \preceq 2+\sum_{i=3}^{2k-5} 2^i + 2^{2k-3} \quad \text{et} \quad (2^{2k}-4)/3 = \sum_{i=1}^{k-1} 2^{2i} \preceq \sum_{i=2}^{2k-6} 2^i + 2^{2k-4} + 2^{2k-2} \ .$$

Or

$$\left(2+\sum_{i=3}^{2k-5} 2^i + 2^{2k-3}\right)/3 = \sum_{i=1}^{2k-5} 2^i \preceq (2^{2k-3}-2) \text{ et } \left(\sum_{i=2}^{2k-6} 2^i + 2^{2k-4} + 2^{2k-2}\right)/3 = \sum_{i=1}^{k-3} 2^{2i}+2^{2k-5}+2^{2k-4} \preceq (2^{2k-3}-2) \ .$$

Par ailleurs, on a les monômes d'exposant $2^{2k}-2^j-5$ et $2^{2k}-2^j-2$ pour $j = 2i+1$ où $2 \le i \le k-2$:

$$(2^{2k}-2^j-5)/3 = 1+\sum_{i=1}^{(j-3)/2} 2^{2i+1} + \sum_{i=(j+1)/2}^{k-1} 2^{2i} \quad \preceq 2^{2k-1}-2^{2k-3}-2^l-5, \quad \text{où } l \in [4, j-1] \text{ pair} \ ,$$

$$(2^{2k}-2^j-2)/3 = \sum_{i=0}^{(j-3)/2} 2^{2i+1} + \sum_{i=(j+1)/2}^{k-1} 2^{2i} \qquad \preceq 2^{2k-1}-2^{2k-3}-2^l-5, 2^{2k-1}-2^{2k-3}-2^l-2 \ ,$$

car $\{2^{2k-1}-2^{2k-3}-2^l-5, 2^{2k-1}-2^{2k-3}-2^l-2 \text{ où } l = 2i, 2 \le i \le k-1\} \subset \mathcal{N}_{r-1}$.
Enfin, on a $2^{2k}-2^3-5, 2^{2k}-2^3-2 \in \mathcal{N}_r$ car

$$(2^{2k}-2^3-2)/3 = 2+\sum_{i=2}^{k-1} 2^{2i} \preceq 2^{2k-1}-2^{2k-3}-2^2-2 \in \mathcal{N}_{r-1} \ ,$$

$$\text{et} \qquad (2^{2k}-2^3-5)/3 = 1+\sum_{i=2}^{k-1} 2^{2i} \quad \preceq 1+\sum_{i=4}^{2k-4} 2^i + 2^{2k-2} \qquad \text{or} \left(1+\sum_{i=4}^{2k-4} 2^i + 2^{2k-2}\right)/3 = 2^{2k-3}-5 \ .$$

5. Cas P-P-P : si $\lfloor r\log_2 3\rfloor = 2k$, $\lfloor (r-1)\log_2 3\rfloor = 2k-2$ et $\lfloor (r-2)\log_2 3\rfloor = 2k-4$, on a $2^{2k-1}-5 \in \mathcal{N}_r$:

$$(2^{2k-1}-5)/3 = 1+\sum_{i=1}^{k-2} 2^{2i+1} \preceq 1+\sum_{i=3}^{2k-3} 2^i = 2^{2k-2}-7 \in \mathcal{N}_{r-1} \ .$$

De plus, on a aussi $2^{2k-1}-2$:

$$(2^{2k-1}-2)/3 = \sum_{i=0}^{k-2} 2^{2i+1} \preceq 2+2^3+\sum_{i=5}^{2k-9} 2^i + 2^{2k-7}+2^{2k-5}+2^{2k-3}\ .$$

Or

$$\left(2+2^3+\sum_{i=5}^{2k-9} 2^i + 2^{2k-7}+2^{2k-5}+2^{2k-3}\right)/3 = 2+2^2+2^3+\sum_{i=3}^{k-5} 2^{2i} + 2^{2k-7}+2^{2k-6}+2^{2k-5}$$
$$\preceq 2+2^2+2^3+\sum_{i=5}^{2k-9} 2^i + 2^{2k-7}+2^{2k-6}+2^{2k-5}\ .$$

Puis

$$\left(2+2^2+2^3+\sum_{i=5}^{2k-9} 2^i + 2^{2k-7}+2^{2k-6}+2^{2k-5}\right)/3 = 2+\sum_{i=3}^{2k-9} 2^i + 2^{2k-6} \preceq 2^{2k-5}-2, 2^{2k-5}-5\ ,$$

où $2^{2k-5}-2, 2^{2k-5}-5 \in \mathcal{N}_{r-3}$ (car on a nécessairement $\lfloor (r-3)\log_2(3)\rfloor = 2k-5$).

Pour $2^{2k}-4$ et $2^{2k}-7$, il n'est pas possible de se ramener aux exposants de poids maximal des tours précédents, car $(2^{2k}-4)/3^i$ et $(2^{2k}-7)/3^i$ sont supérieurs aux exposants de poids maximal au tour $r-i$. Par ailleurs, la présence d'autres monômes est également à étudier, mais il est difficile de déterminer la forme possible des exposants de ces monômes. En effet, la première apparition d'une séquence du type P-P-P est pour $r=19$, ce que je n'ai pas pu observer expérimentalement pour des raisons de coûts de calculs.

6. Cas I-I-I : lorsque $\lfloor r\log_2 3\rfloor = 2k+1$, $\lfloor (r-1)\log_2 3\rfloor = 2k-1$ et $\lfloor (r-2)\log_2 3\rfloor = 2k-3$, il faudrait montrer la présence de $2^{2k+1}-5$ et $2^{2k+1}-2$, d'après les résultats observés dans la section 3.2.2. Cependant, il n'est de nouveau pas possible de se ramener aux exposants de poids maximal des tours précédents, car $(2^{2k+1}-5)/3^i$ et $(2^{2k+1}-2)/3^i$ sont supérieurs aux exposants de poids maximal au tour $r-i$.
   Toutefois, j'ai pu noter qu'aux tours 7 et 12, il était possible de retrouver le monôme d'exposant $2^{2k+1}-5$ à partir de $3^{r-3} \in \mathcal{N}_{r-3}$, mais cela nécessite de connaître le développement binaire des puissances de 3.

A.2 Exposants proches de 3^r

Afin de justifier la présence de certains exposants de poids maximal au tour r, il serait également intéressant d'étudier la présence des monômes proches du monôme donnant le degré univarié aux tours précédents, c'est-à-dire, proches de $x^{3^{r-i}}$, sur les tours $r-i$.

Il est par exemple possible de déterminer la présence, ou non, des monômes x^j tels que $j \in [3^r-60, 3^r]$ (voir tableau A.1). En effet, grâce à la proposition 2.4 et à l'expression de 3^r sous la forme $\{1+2^4+2^5a,\ 1+2+2^4+2^5a,\ 1+2^3+2^4+2^5a,\ 1+2+2^3+2^5a,\ 1+2^5a,\ 1+2+2^5a,\ 1+2^3+2^5a,\ 1+2+2^3+2^4+2^5a,\ \text{avec } a\in\mathbb{N}\}$, on peut montrer par récurrence que certains monômes proches de x^{3^r} sont nécessairement présents ou qu'à l'inverse ils sont forcément absents.

3^r-i	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60
$r\equiv 0[8]$	x	x	x	x					x	x	x	x				x			x	x	
$r\equiv 1[8]$	x	x		x		x	x	x	x	x		x		x	x	x	x	x			
$r\equiv 2[8]$	x	x	x	x						x		x				x		x	x		
$r\equiv 3[8]$	x	x		x		x	x	x	x	x											
$r\equiv 4[8]$	x	x	x	x					x	x	x	x				x	x	x	x	x	
$r\equiv 5[8]$	x	x		x		x	x	x	x	x		x		x	x	x	x	x			
$r\equiv 6[8]$	x	x	x	x						x		x				x	x	x	x		
$r\equiv 7[8]$	x	x		x		x	x	x	x	x							x	x		x	

TABLE A.1 – Exposants proches de 3^r

En particulier, on remarque que les monômes d'exposants $3^r - 12$, $3^r - 36$ et $3^r - 60$ ne sont jamais présents. Et plus généralement : pour tout $K \equiv 12 \mod 24$, le monôme x^{3^r-K} n'est pas présent. En effet, on sait que les monômes dont les exposants sont congrus à 5 et 7 modulo 8 ne sont jamais présents, or $3^r \equiv 1 \mod 8$ si $r \equiv 0 \mod 2$ et $3^r \equiv 3 \mod 8$ si $r \equiv 1 \mod 2$, donc $3^r - K \equiv 5, 7 \mod 8$ quel que soit r lorsque $K \equiv 12 \mod 24$. De même, les monômes x^{3^r-K} ne sont pas présents : si $r \equiv 0 \mod 2$ et $K \equiv 18 \mod 24$, ou si $r \equiv 1 \mod 2$ et $K \equiv 6 \mod 24$.

À l'inverse les monômes d'exposants $3^r - 3^i$ sont toujours présents car $3^r - 3^i = 3^i(3^{r-i} - 1)$ or au tour $r - i$ on a le monôme d'exposant 3^{r-i} et comme on a toujours $3^{r-i} - 1 \preceq 3^{r-i}$ alors on a l'exposant $3(3^{r-i} - 1)$ au tour $r - i + 1$, et donc l'exposant $3^2(3^{r-i} - 1)$ au tour $r - i + 2$, ...etc.. et finalement, on a bien l'exposant $3^i(3^{r-i} - 1)$ au tour r.

La proposition 2.4 nous donne par ailleurs le résultat suivant :

$$[\, j \in \mathcal{M}_r \text{ tel que } j \geq 3^r - (3^i - 1) \text{ et } 3^{i-1} | j \,] \iff [\, (j/3^{i-1}) \preceq 3^{r-i+1}] \,,$$

qui nous permet également de justifier la présence des monômes proches de x^{3^r}. Cette relation est notamment utile pour comprendre comment ont été construits les monômes. Par exemple, au troisième tour, on a $\mathcal{M}_3 = \{0, 3, 6, 9, 12, 18, 24, 27\}$, où $27 = 3 \times 9$ a été construit à partir de l'exposant 9 présent au tour précédent, alors que $24 = 3 \times 8$ a été construit à partir de 8 qui n'était pas présent au tour précédent, mais $8 \preceq 9$.

Cette autre approche nous permet par exemple de constater la densité des polynômes univariés décrivant MiMC. Toutefois, afin de pouvoir obtenir des informations sur le poids de ces exposants qui pourrait permettre de justifier la présence d'exposants d'une certaine forme (voir section A.1), il faudrait notamment pouvoir déterminer le développement binaire des puissances de 3.

Annexe B

Compléments sur le degré algébrique du déchiffrement

B.1 Degré observé

Le tableau B.1 donne le degré algébrique du déchiffrement, observé expérimentalement pour certaines valeurs de n :

r	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$n=7$	4	4	5	5	6										
$n=9$	5	5	6	7	7	8									
$n=11$	6	6	8	8	9	9	10								
$n=13$	7	7	9	10	10	11	11	11	12						
$n=15$	8	8	10	11	12	13	13	13	13	14					
$n=17$	9	9	12	13	14	15	15	15	15	15	16				
$n=19$	10	10	13	14	15	16	17	17	17	17	17	18			
$n=21$	11	11	14	16	17	18	19	19	19	19	19	19	19	20	
$n=23$	12	12	16	17	19	20	20	20	21	21	21	21	21	21	22

TABLE B.1 – Degré algébrique de la transformation inverse en fontion de n

Soit d_i le degré algébrique au tour i. On a déjà montré que $d_1 = d_2 = (n+1)/2$. Pour les tours suivants, le tableau suggère une évolution du degré algébrique de la façon suivante :

$$d_3 \approx \left\lfloor \frac{2n+2}{3} \right\rfloor = \frac{n+1}{2} + \left\lfloor \frac{n+1}{6} \right\rfloor \qquad d_6 \approx \left\lfloor \frac{14n+5}{16} \right\rfloor = \frac{n+1}{2} + \left\lfloor \frac{6n-3}{16} \right\rfloor$$

$$d_4 \approx \left\lfloor \frac{3n+1}{4} \right\rfloor = \frac{n+1}{2} + \left\lfloor \frac{n-1}{4} \right\rfloor \qquad d_7 \approx \left\lfloor \frac{8n+3}{9} \right\rfloor = \frac{n+1}{2} + \left\lfloor \frac{7n-3}{18} \right\rfloor$$

$$d_5 \approx \left\lfloor \frac{13n+6}{16} \right\rfloor = \frac{n+1}{2} + \left\lfloor \frac{5n-2}{16} \right\rfloor \qquad d_8 \approx \left\lfloor \frac{8n+3}{9} \right\rfloor = \frac{n+1}{2} + \left\lfloor \frac{7n-3}{18} \right\rfloor$$

Ces conjectures ne sont basées que sur quelques observations et les valeurs ne semblent pas suivre une suite logique, mais on peut supposer que pour toute autre valeur de n, le degré algébrique pour ces tours suivra approximativement ces valeurs.

B.2 Etude de $wt(js)$ en fonction de $wt(j)$

Dans cette partie, nous étudions quelques pistes pour démontrer la conjecture 4.2, présentée dans la sous-section 4.2.3, à savoir que pour $2 \leq j \leq 2^n - 2$,

$$wt(js \bmod 2^n - 1) \in \begin{cases} [k, (n+2k-3)/2] & \text{si } \ wt(j) = 2k \\ [k+2, (n+2k+1)/2] & \text{si } \ wt(j) = 2k+1 \ . \end{cases}$$

En m'inspirant de la démonstration du palier entre les tours 1 et 2, présentée dans la section 4.1, j'ai cherché à utiliser un raisonnement par récurrence à partir des cas $wt(j) = 2$ et $wt(j) = 3$.

Premièrement, notons que :

— si $wt(j) = 2$ tel que $j = 1 + 2^{i_1}$, $i_1 \in \{1, \dots, n-1\}$ on a bien $wt(sj) \in [1, (n-1)/2]$ car :

— si i_1 est pair, on a :

$$sj = 2^{i_1} + \sum_{l=i_1/2}^{(n-3)/2} 2^{2l+1} \mod 2^n - 1 \quad \text{donc} \quad wt(sj) = (n - i_1 + 1)/2 \in [1, (n-1)/2] \ ,$$

— et si i_1 est impair, on a :

$$sj = 1 + \sum_{l=0}^{(i_1-3)/2} 2^{2l+1} \mod 2^n - 1 \quad \text{donc} \quad wt(sj) = (i_1 + 1)/2 \in [1, (n-1)/2] \ .$$

— si $wt(j) = 3$, où $j = 1 + 2^{i_1} + 2^{i_2}$, $i_1 \in \{1, \dots, n-2\}$, $i_2 \in \{i_1 + 1, \dots, n-1\}$ on a $wt(sj) \in [3, (n+3)/2]$:

— si i_1 et i_2 sont pairs, alors :

$$sj = 1 + \sum_{l=0}^{(i_1-2)/2} 2^{2l+1} + \sum_{l=i_1/2}^{(i_2-2)/2} 2^{2l} \mod 2^n - 1 \quad \text{donc} \quad wt(sj) = (i_2 + 2)/2 \in [3, (n+1)/2] \ ,$$

— si i_1 est pair et i_2 impair, alors :

$$sj = \sum_{l=0}^{(i_1-2)/2} 2^{2l} + 2^{i_2} + \sum_{l=(i_2+1)/2}^{(n-1)/2} 2^{2l} \mod 2^n - 1 \quad \text{donc} \quad wt(sj) = (n + i_1 - i_2 + 2)/2 \in [3, (n+3)/2] \ ,$$

— si i_1 est impair et i_2 pair, alors :

$$sj = \sum_{l=0}^{(i_1-1)/2} 2^{2l} + \sum_{l=(i_1-1)/2}^{(i_2-2)/2} 2^{2l+1} + \sum_{l=i_2/2}^{(n-1)/2} 2^{2l} \mod 2^n - 1 \quad \text{donc} \quad wt(sj) = (n+3)/2 \ ,$$

— et si i_1 et i_2 sont impairs, alors :

$$sj = 2^{i_1} + \sum_{l=(i_1+1)/2}^{(i_2-1)/2} 2^{2l} + \sum_{l=(i_2-1)/2}^{(n-3)/2} 2^{2l+1} \mod 2^n - 1 \quad \text{donc} \quad wt(sj) = (n - i_1 + 2)/2 \in [3, (n+1)/2] \ .$$

On peut alors se demander si lorsque $wt(j) \equiv 0 \mod 2$, il existe une relation de récurrence permettant de se ramener au cas $wt(j) = 2$. De même, si $wt(j) \equiv 1 \mod 2$, on souhaite se ramener au cas $wt(j) = 3$.

Toutefois, en observant les expressions de js lorsque $wt(j) = 4$ ou $wt(j) = 5$, il semble difficile de trouver une formule unifiée, car le poids de js dépend de la parité des exposants des puissances de 2 représentant j.

En effet, si $wt(j) = 4$, tel que $j = 1 + 2^{i_1} + 2^{i_2} + 2^{i_3}$, $i_1 \in \{1, \dots, n-3\}$, $i_2 \in \{i_1 + 1, \dots, n-2\}$, $i_3 \in \{i_2 + 1, \dots, n-1\}$, alors le tableau B.2 nous donne les poids de sj en fonction de la parité de i_1, i_2, i_3.

De même, lorsque $wt(j) = 5$, tel que $j = 2^0 + 2^{i_1} + 2^{i_2} + 2^{i_3} + 2^{i_4}$, $i_1 \in \{1, \dots, n-4\}$, $i_2 \in \{i_1 + 1, \dots, n-3\}$, $i_3 \in \{i_2 + 1, \dots, n-2\}$, $i_4 \in \{i_3 + 1, \dots, n-1\}$, le tableau B.3 indique le poids de sj en fonction de la parité des exposants i_1, i_2, i_3, i_4.

$i_1 \mod 2$	$i_2 \mod 2$	$i_3 \mod 2$	$wt(sj \bmod 2^n - 1)$
0	0	0	$(n + i_1 - i_2 + 3)/2$
0	0	1	$(n - i_1 + 3)/2$
0	1	0	$(n - i_1 + i_2 - i_3 + 2)/2$
0	1	1	$(i_3 + 3)/2$
1	0	0	$(n - i_1 + i_2 - i_3 + 2)/2$
1	0	1	$(i_1 - i_2 + i_3 + 2)/2$
1	1	0	$(i_1 - i_2 + i_3 + 2)/2$
1	1	1	$(n + i_2 - i_3 + 3)/2$

TABLE B.2 – $wt(js)$ pour $wt(j) = 4$

$i_1 \mod 2$	$i_2 \mod 2$	$i_3 \mod 2$	$i_4 \mod 2$	$wt(sj \bmod 2^n - 1)$
0	0	0	0	$(n - i_1 + i_3 - i_4 + 3)/2$
0	0	0	1	$(i_2 - i_3 + i_4 + 3)/2$
0	0	1	0	$(i_2 - i_3 + i_4 + 3)/2$
0	0	1	1	$(n + i_1 - i_2 + i_3 - i_4 + 3)/2$
0	1	0	0	$(i_4 + 4)/2$
0	1	0	1	$(n + i_1 - i_2 + i_3 - i_4 + 3)/2$
0	1	1	0	$(n + i_1 - i_2 + 4)/2$
0	1	1	1	$(n - i_1 + i_2 - i_3 + 3)/2$
1	0	0	0	$(i_1 - i_2 + i_4 + 3)/2$
1	0	0	1	$(n + i_3 - i_4 + 4)/2$
1	0	1	0	$(n + 5)/2$
1	0	1	1	$(n - i_1 + i_2 - i_3 + 3)/2$
1	1	0	0	$(n + i_2 - i_3 + 4)/2$
1	1	0	1	$(n - i_1 + 4)/2$
1	1	1	0	$(n - i_1 + i_3 - i_4 + 3)/2$
1	1	1	1	$(i_1 - i_2 + i_4 + 3)/2$

TABLE B.3 – $wt(js)$ pour $wt(j) = 5$