# Non-linéarité des fonctions booléennes : applications de la théorie des fonctions booléennes et des codes en cryptographie

Julien Bringer

# THÈSE

présentée devant

## l'Université du Sud Toulon – Var

pour obtenir

le grade de Docteur de l'Université du Sud Toulon – Var
spécialité Mathématiques, Informatique et applications

par

**Julien Bringer**

Groupe de Recherche en Informatique et Mathématiques
École Doctorale
UFR Sciences et Techniques

# Non-linéarité des fonctions booléennes

# Applications de la théorie des fonctions booléennes et des codes en cryptographie

soutenue le 16 novembre 2007 devant la commission d'examen

| | | | | |
|---|---|---|---|---|
| MM. : | François Rodier | DR | (IML & CNRS) | Rapporteurs |
| | Serge Vaudenay | PR | (EPFL) | |
| MM. : | Hervé Chabanne | Dr. | (Sagem Sécurité) | Examinateurs |
| | Henri Gilbert | Dr. | (France Télécom R&D) | |
| | Patrick Solé | DR | (UNSA & CNRS) | |
| | Jacques Wolfmann | PRE | (Univ. Sud Toulon – Var) | |
| M. : | Philippe Langevin | PR | (Univ. Sud Toulon – Var) | Directeur |

*À Karell et Maud*

# Remerciements

C'est avec plaisir que je remercie ici toutes les personnes qui m'ont aidé ou accompagné durant cette thèse.

Tout d'abord Philippe Langevin, mon directeur de thèse, qui a su me faire profiter de ses connaissances scientifiques et répondre à mes différentes interrogations. Je lui suis notamment reconnaissant pour son soutien, et la compréhension dont il a fait preuve depuis mon départ de Toulon pour SAGEM (aujourd'hui Sagem Sécurité). Je remercie également Hervé Chabanne, responsable du Pôle Recherche en Sécurité et Cryptographie de Sagem Sécurité, qui depuis fin 2004 m'accorde sa confiance et partage avec moi son expérience et ses idées. Sans lui cette thèse serait moins riche, puisque les résultats de la deuxième partie ont tous été obtenus avec lui.

Je remercie chaleureusement François Rodier et Serge Vaudenay d'avoir accepté la lourde tâche de rapporteur, de l'intérêt porté à cette thèse, et pour tous leurs conseils éclairés. J'aimerais les remercier de plus, ainsi que Hervé, Henri Gilbert, Patrick Solé et Jacques Wolfmann, de m'avoir fait l'honneur et le plaisir de faire partie de mon jury.

J'ai effectué la majorité de mon travail de doctorant au sein de l'université de Toulon entre 2001 et 2004, et je tiens donc à remercier Jacques pour m'avoir accueilli au sein du GRIM, dont il assurait alors la direction. Je n'oublie pas non plus tous les membres du GRIM, avec lesquels j'ai eu l'occasion de goûter aux joies de l'enseignement, et bien sûr d'étudier des problèmes de recherche variés. Je leur dois de plus l'opportunité d'avoir participé à l'organisation des colloques YACC en 2002 et 2004. J'en profite pour saluer toutes les personnes rencontrées lors de ces trois années à Toulon, notamment les membres du troisième étage ou du département d'informatique, et les doctorants de Toulon, dont Chloé, ou de Marseille, je pense à la fine équipe du CMI, pour les très bons moments passés avec eux. Le séminaire AZURCRYPT co-organisé par Pascal Véron m'a également permis de découvrir un peu plus l'approche industrielle de la cryptographie et de la sécurité. Cela a contribué à mon orientation vers un parcours industriel et je remercie tout particulièrement Yannick Teglia et Béatrice Peirani pour leur aide en ce sens. Un merci très spécial à Pierre-Yvan Liardet, dont le soutien m'a aidé à continuer ma thèse.

Même si cette thèse n'a pas été effectuée dans le cadre direct de mon emploi à la Sagem, une part non négligeable y est liée (les résultats décrits dans la deuxième partie en sont tous issus) et je remercie donc mes collègues et tous ceux, Hervé étant l'un des principaux, qui ont fait que cela soit possible.

Un des plaisirs de la recherche étant de travailler à plusieurs, toute ma reconnaissance va aussi à mes différents co-auteurs pour leurs collaborations, y compris en dehors du cadre de cette thèse. Je pense également à toutes les personnes rencontrées lors d'un déplacement dans une conférence, ou par hasard, avec qui j'ai eu plaisir à travailler, même sans résultats précis à la clé : en particulier merci à Gregor Leander.

Enfin, j'aimerais finir par une dédicace spéciale à ma famille, mes amis et à tous ceux qui ont contribué à leur manière à mon parcours.

# Table des matières

# Introduction

Liée au souci de protection d'informations de toute sorte, la cryptologie – la science du secret – a été très longtemps considérée comme une arme de guerre, et de fait, elle n'est devenue un axe de recherche publique et reconnue que depuis le milieu du XXème siècle. Limitée historiquement à des opérations élémentaires d'un point de vue mathématique pour des raisons de simplicité d'utilisation, le développement des capacités de calcul en informatique et l'explosion des communications numériques ont été de pair avec une extraordinaire évolution de la théorie. Elle est composée de la cryptographie – l'écriture secrète – et de la cryptanalyse – l'analyse de celle-ci. De nos jours, ses applications sont nombreuses (confidentialité, authentification, intégrité, non-répudiation, ...) et l'utilisation de la cryptographie devient de plus en plus courante et variée lorsqu'on traite des données sous forme numérique.

Traditionnellement, les traitements utilisent un secret commun à deux interlocuteurs pour chiffrer et déchiffrer les messages, mais une avancée majeure du domaine est l'invention du concept de cryptographie à clé publique par Diffie et Hellman [DH76] (aujourd'hui appelée cryptographie asymétrique) en 1976 et la publication du premier schéma concret, par Rivest, Shamir et Adleman en 1977 [RSA78]. Depuis, les efforts de formalisation, l'introduction de nouveaux usages, l'amélioration des procédés de cryptanalyse, ou tout simplement l'évolution des connaissances, ont lié la cryptographie à des théories mathématiques et informatiques élaborées et variées (parfois développées pour l'occasion, parfois connues depuis des siècles). La cryptologie est liée à l'algèbre en général (dont la théorie des nombres, la théorie des corps, l'arithmétique), la théorie de la complexité, les probabilités, la théorie de l'information, etc... Réciproquement, elle participe également à l'avancée d'autres sujets ; par exemple les travaux menés à Bletchley Park, pour réussir à retrouver les messages chiffrés par la machine Enigma lors de la seconde guerre mondiale, ont amené Max Newman et Alan Turing à concevoir le premier calculateur totalement électronique, programmable et capable de traiter des opérations arithmétiques binaires.

De nombreux problèmes continuent à être étudiés. En particulier, face à la prédominance de la théorie des nombres dans le domaine de la cryptographie asymétrique, l'application de nouveaux problèmes difficiles est recherchée afin d'assurer une certaine sécurité aux cryptosystèmes associés. Mais les cryptanalyses se développant en parallèle, les alternatives aux problèmes de factorisation et du logarithme discret restent encore peu nombreuses (parmi les plus crédibles, on retiendra le cryptosystème de McEliece [McE78], fondé sur la théorie des codes, qui suivait la publication du système RSA de seulement un an). De plus, de nouveaux usages, tels que l'utilisation de composants matériels avec une très faible capacité de calcul, introduisent de nouvelles contraintes.

Contrairement à la cryptographie asymétrique, la cryptographie symétrique repose très souvent sur la multiplication d'opérations élémentaires, principalement afin d'atteindre des débits élevés de traitements. Ce faisant la sécurité est plus difficile à établir (même parfois imprévisible) et les critères de résistance sont généralement définis suite à la découverte de nouvelles attaques spécifiques. Suivant la théorie de Boole, les objets les plus élémentaires dans ce cadre sont naturellement une combinaison d'opérations logiques, i.e. des fonctions agissant sur des vecteurs binaires et à valeur 0 ou 1 : on parle de fonctions booléennes.

L'étude des fonctions booléennes est liée à de nombreux sujets comme la théorie algébrique des codes, la théorie des séquences, la théorie des designs et, depuis plus récemment, la cryptographie. En fait, les fonctions booléennes sont devenues importantes pour la conception et la sécurité de certains systèmes de chiffrement symétrique : les systèmes de chiffrement par blocs et de chiffrement à flot. Les fonctions utilisées dans ces systèmes doivent avoir des propriétés très particulières pour résister à des attaques spécifiques dont les principales sont la cryptanalyse linéaire, la cryptanalyse différentielle, l'attaque par corrélations et les attaques algébriques. Les attaques mises au point à la fin des années 1980 et au début des années 1990 ont en effet pu être formalisées en mettant en évidence le rôle des fonctions booléennes en cryptographie et les propriétés qu'elles doivent vérifier, dont : haute non-linéarité, équilibre, immunité aux corrélations et degré élevé.

Parmi les nombreux domaines reliés en partie à la cryptologie, les travaux présentés dans cette thèse s'intéressent tout particulièrement à la théorie des codes et à la théorie des fonctions booléennes. Deux objectifs différents sont alors visés. Dans une première partie, l'étude d'un problème ouvert sur les fonctions booléennes est effectuée ; ce problème est notamment lié à un critère cryptographique mais la motivation de son étude est bien plus large. Alors que la deuxième partie, elle, est centrée sur les applications concrètes en cryptographie d'objets provenant de ces 2 théories.

**Partie I : Fonctions booléennes et non-linéarité.** Au centre des propriétés intéressantes des fonctions booléennes utilisées en cryptographie, la haute non-linéarité est un problème historique lié à la théorie des codes, et en particulier au calcul du rayon de recouvrement du code de Reed-Muller d'ordre 1. C'est un des problèmes jugés difficiles dans la théorie des codes et c'est un problème ouvert en général. La non-linéarité d'une fonction booléenne est égale à la distance de la fonction au code de Reed-Muller d'ordre 1. En dimension paire, on connaît la non-linéarité maximale, elle est atteinte pour les fonctions courbes, notion définie par Rothaus en 1976, mais on ne sait pas encore classifier toutes les fonctions courbes ni établir leur nombre exact. En dimension impaire, en général, on ne connaît pas cette valeur maximale. Un des outils principaux pour étudier la non-linéarité d'une fonction booléenne est l'analyse de Fourier, par l'intermédiaire de la transformée de Fourier discrète sur un corps fini et la théorie des caractères. Le problème de déterminer la non-linéarité maximale devient le problème de trouver le rayon spectral, à savoir l'amplitude spectrale minimale. Il est très facile d'obtenir un encadrement du rayon spectral et on connaît également la valeur exacte pour des petites dimensions impaires ($m = 3$, 5 ou 7). Pour $m$ supérieur ou égal à 15, Patterson et Wiedemann ont affiné la borne supérieure en construisant, en 1983, une fonction ayant une amplitude infé-

rieure à celles connues jusque-là. Cette borne est la meilleure connue à ce jour. À noter qu'en 2006, des avancées ont également été obtenues dans les cas $m = 9$, 11 et 13 grâce aux travaux de Kavut, Maitra, Sarkar et Yücel.

La première partie est consacrée à ce problème et nous y étudions la construction de fonctions hautement non-linéaires ayant une structure similaire aux fonctions proposées par Patterson et Wiedemann. Plus précisément, l'idée est de partir d'une construction algébrique de fonctions booléennes invariantes suivant l'action d'un groupe pour laquelle l'estimation de la meilleure non-linéarité possible met alors en évidence le besoin de relâcher certaines contraintes. Pour avoir plus de liberté, nous proposons une généralisation de cette construction, constante suivant l'action d'un groupe sur toutes les orbites excepté sur l'orbite triviale, dans le but d'améliorer la non-linéarité. Le but de cette partie est principalement de montrer pourquoi cette construction est une cible prometteuse pour s'approcher de la conjecture de Patterson-Wiedemann. Cette étude a été partiellement publiée dans [BGL05].

L'étude de l'amplitude spectrale de telles fonctions introduit deux problèmes différents. Le premier point est la valeur de sommes de Gauss : à partir de la théorie des sommes de Gauss et à l'aide de résultats connus, dont ceux établis par Stickelberger, nous essayons d'obtenir des estimations de ces sommes afin de choisir des sous-groupes convenables. D'autre part, un nouveau problème apparaît, à savoir l'estimation de sommes d'exponentielles sur un sous-groupe du groupe multiplicatif d'un corps fini qui, par analogie avec le cas du corps tout entier, peut-être vu comme une généralisation du problème de détermination du rayon spectral (nous parlerons de non-linéarité partielle et de rayon spectral partiel). L'estimation de sommes d'exponentielles est une question très largement traitée dans la littérature mais qui demeure difficile dans les cas généraux ; dans le cas qui nous intéresse, un cas particulier de sommes d'exponentielles incomplètes dans un corps fini, il ne semble pas y avoir de résultats notables à l'heure actuelle. Ce deuxième point pourrait donc donner l'impression que cette construction est artificielle mais ici il n'est pas nécessaire d'obtenir une estimation optimale de ces sommes. Nous étudions donc ce nouveau problème à partir de différentes techniques. Pour faire abstraction des difficultés rencontrées à cardinal fini, nous généralisons les techniques utilisées par Rodier pour l'étude asymptotique de la non-linéarité classique et étudions le comportement asymptotique de la non-linéarité partielle pour des sous-ensembles quelconques. Différents exemples illustrent l'étude pour souligner l'intérêt de la démarche. Nous montrons en particulier un exemple de construction de fonction hautement non-linéaire et établissons des exemples de conditions suffisantes, et réalistes, pour descendre sous la borne de Patterson-Wiedemann.

Cette partie est organisée comme suit. Les principales notions théoriques utilisées, liées à la théorie des caractères, l'analyse de Fourier, et les sommes de Gauss, sont rappelées au chapitre 1 afin de faciliter la lecture et la compréhension. Ce premier chapitre est également l'occasion de faire un survol de la théorie des fonctions booléennes, d'expliquer la problématique de l'étude de leur non-linéarité et de détailler quelques critères cryptographiques importants. Le chapitre 2 décrit les constructions de type Patterson-Wiedemann et la généralisation proposée. Nous montrons en particulier comment réduire le problème de l'étude de la non-linéarité à celle sur un sous-groupe grâce à l'utilisation des sommes de Gauss. Dans le chapitre 3, nous introduisons la notion de non-linéarité partielle et étudions le comportement asymptotique en fonction des ensembles concernés. Les résultats obtenus sont généraux et permettent de prouver que le comportement peut, sous certaines conditions, être proche du cas du corps. Le chapitre 4

permet de conforter et approfondir cette analyse dans le cas d'un sous-groupe multiplicatif en s'intéressant aux résultats applicables en dimension finie. Le chapitre 5 a alors pour but de détailler des applications concrètes de cette construction. Nous décrivons ainsi une fonction de type Patterson-Wiedemann généralisé battant la borne quadratique en dimension 15 et détaillons plusieurs résultats allant dans le sens de la conjecture de Patterson-Wiedemann. Enfin une synthèse des résultats ainsi que des perspectives de recherche sont présentées dans le chapitre 6.

**Partie II : Cryptographie.** Dans la deuxième partie de la thèse, nous présentons différentes applications cryptographiques des fonctions booléennes et de la théorie des codes. Un objectif est de mettre en avant l'intérêt de ces domaines et une partie de la diversité des utilisations possibles.

De plus, ces applications correspondent à des cadres récents d'utilisation de la cryptographie, pour lesquels la cryptographie classique n'offrent pas de solution efficace. Le but est donc d'exploiter les propriétés de certaines fonctions booléennes ou de codes afin d'apporter des solutions pratiques. Deux cadres récents sont étudiés. Le premier concerne les protocoles cryptographiques faisant intervenir des composants – type étiquette électronique – à faibles ressources (capacité de calcul, mémoire, énergie) où un schéma cryptographique classique, asymétrique ou symétrique, est en général trop consommateur pour y être implémenté. Le deuxième est le problème d'authentification de données variables dans le temps, où des primitives cryptographiques tolérant des erreurs sont nécessaires, ce qui conduit naturellement à l'utilisation de codes correcteurs d'erreurs.

Le chapitre 7 est ainsi articulé autour de 3 publications différentes. Le choix qui a été fait est de les insérer telles qu'elles ont été publiées, en ajoutant une présentation du contexte et un résumé de la contribution avant chaque article. Pour aller plus loin, une perspective de poursuite est également précisée à chaque fois. Dans le premier article, nous proposons l'introduction de fonctions puissances afin de prémunir le protocole d'authentification à faibles ressources $\mathrm{HB}^+$ (proposé par Juels et Weis à Crypto'05) contre les attaques par le milieu. Sur un sujet analogue de protection à faible coût, le second article analyse un procédé, qui a été proposé en 2006 par Avoine et Castelluccia, de protection d'équipements ne disposant de presque aucune capacité de calcul, pour lequel nous mettons en évidence le parallèle avec la théorie du canal à jarretière de Wyner pour accroître la sécurité. Enfin, le troisième article présente un travail sur l'utilisation du cryptosystème de McEliece, à base de codes correcteurs d'erreurs, dans des schémas d'authentification de données bruitées, ceci dans le but de restreindre l'accès aux fonctions de vérification.

**Guide de lecture.** Selon le découpage ci-dessus, les deux parties peuvent donc être lues de manière indépendante. En marge des sujets principaux de cette thèse, le lecteur trouvera également en annexe quelques détails et compléments sur la théorie des codes, sur l'étude des fonctions puissances et sur les attaques par corrélations des algorithmes de chiffrement à flot constitués de registres linéaires combinés par une fonction booléenne.

# Première partie

# Fonctions booléennes et non-linéarité

# Chapitre 1

# Définitions et résultats préliminaires

Ce chapitre a pour objet de présenter le cadre de travail des chapitres suivants et de préciser les définitions et les résultats utiles pour leur compréhension. Dans un premier temps, nous rappelons les principaux résultats de la théorie des caractères et de l'analyse de Fourier, principalement dans le cadre des corps finis, sans oublier l'introduction des sommes de Gauss, qui sera un des outils principaux des chapitres 2 et 5. Dans un deuxième temps, la théorie des fonctions booléennes est présentée : quelques notions générales sur les fonctions booléennes sont définies, puis la problématique de l'étude de leur non-linéarité est détaillée. Pour finir, le lien entre les propriétés des fonctions booléennes et les contraintes de construction de schémas cryptographiques est expliqué.

Ce sont des résultats, pour la plupart, classiques, qui ne seront pas démontrés ici ; des références bibliographiques seront précisées régulièrement pour guider le lecteur vers plus de détails et/ou vers des démonstrations.

## 1.1 Théorie des caractères et analyse de Fourier

Un outil important en théorie des nombres, pour la résolution de problèmes liés à des entiers comme à des réels, est le domaine des sommes d'exponentielles. De même, des sommes analogues définies sur un corps fini s'avèrent utiles dans de nombreuses applications. Les sommes d'exponentielles sont des sommes combinant (de manière linéaire ou non) les valeurs d'objets particuliers formant un groupe : les caractères.

Nous présentons tout d'abord les résultats à la base de la théorie des caractères, dont en particulier les relations d'orthogonalité qui font partie des propriétés fondamentales motivant l'utilisation des caractères. Le cas des corps finis où deux structures différentes s'expriment, les caractères additifs et les caractères multiplicatifs, est ensuite détaillé. Nous rappelons après les principes de la transformée de Fourier avec les propriétés découlant de l'application de ces notions. Enfin, nous introduisons les sommes de Gauss, faisant partie des sommes d'exponentielles les plus importantes, qui permettent entre autre de faire le lien entre la structure additive et multiplicative des caractères d'un corps fini. Ces sommes jouent un rôle dans de nombreux domaines (on peut par exemple citer la loi de réciprocité quadratique) mais nous ne donnerons que les quelques résultats qui seront utilisés dans la suite – leur étude générale est en effet un

problème de longue haleine.

### 1.1.1 Caractères des groupes abéliens finis

La suite est en grande partie extraite du très instructif livre de Serre [Ser70].

#### 1.1.1.1 Dualité

Soit $G$ un groupe abélien fini d'ordre $n$, noté multiplicativement.

**Définition 1.1** *On appelle caractère de $G$ tout homomorphisme de $G$ dans le groupe multiplicatif $\mathbb{C}^\times$ des nombres complexes.*

Les caractères de $G$ forment un groupe $\mathrm{Hom}(G, \mathbb{C}^\times)$, que l'on note $\hat{G}$, appelé le *dual* de $G$. En particulier, les caractères sont à valeurs dans le groupe $\mathbb{U}_n$ des racines $n$-ièmes de l'unité.

**Lemme 1.1 (Dedekind)** *L'ensemble des caractères de $G$ est une partie libre de $\mathbb{C}^G$.*

Autrement dit, si $G$ est fini alors les caractères sont $\mathbb{C}$-indépendants. La démonstration (valable même si $G$ n'est pas fini) se fait par contraposée et par récurrence sur le nombre d'homomorphismes intervenant dans une relation de dépendance linéaire.

**Exemple 1.1** *Supposons que $G$ soit cyclique d'ordre $n$, de générateur $s$. Si $\chi : G \to \mathbb{C}^\times$ est un caractère de $G$, l'élément $w = \chi(s)$ vérifie la relation $w^n = 1$, i.e., est une racine $n$-ième de l'unité. Inversement toute racine $n$-ième de l'unité définit un caractère de $G$ au moyen de $s^a \mapsto w^a$. On voit ainsi que l'application $\chi \mapsto \chi(s)$ est un isomorphisme de $\hat{G}$ sur le groupe $\mathbb{U}_n$ des racines $n$-ièmes de l'unité ; en particulier, $\hat{G}$ est cyclique d'ordre $n$.*

$$G \simeq \hat{G} \simeq \mathbb{U}_n.$$

Le résultat fondamental permettant une étude approfondie des caractères est le suivant :

**Proposition 1.1** *Soit $H$ un sous-groupe de $G$. Tout caractère de $H$ peut être prolongé en un caractère de $G$.*

On raisonne par récurrence sur l'indice $(G : H)$ de $H$ dans $G$ pour le démontrer. L'opération de restriction définit un homomorphisme

$$\rho : \hat{G} \to \hat{H}$$

et la proposition 1.1 signifie que $\rho$ est surjectif. D'autre part, le noyau de $\rho$ est formé des caractères de $G$ qui sont triviaux sur $H$ ; il est donc isomorphe au groupe $\widehat{(G/H)}$ dual de $G/H$. D'où la suite exacte $\{1\} \to \widehat{G/H} \to \hat{G} \to \hat{H} \to \{1\}$, i.e.

$$\hat{G}/\widehat{(G/H)} \simeq \hat{H}.$$

Ce qui permet de démontrer facilement la propriété ci-dessous.

**Proposition 1.2** *Le groupe $\widehat{G}$ est un groupe abélien fini de même ordre que G.*

De plus, le bidual de $G$, le dual $\hat{\hat{G}}$ de $\hat{G}$, est isomorphe à $G$ suivant un isomorphisme canonique

$$\begin{array}{ccc} G & \to & \hat{\hat{G}} \\ x & \mapsto & \{\chi \in \hat{G} \mapsto \chi(x)\}. \end{array}$$

Une autre notion importante est l'orthogonal d'une partie.

**Définition 1.2** *Soit P un sous-ensemble non vide de G. Un caractère $\chi \in \hat{G}$ est dit orthogonal à P si sa restriction $\chi_{|P}$ vaut 1. L'ensemble $\{\chi \in \hat{G}, P \subset Ker\ \chi\}$ des caractères orthogonaux à P est un sous-groupe de $\hat{G}$, il est noté $P^{\perp}$.*

*Par dualité, soit $\Pi$ un sous-ensemble non vide de $\hat{G}$. Un élément $x \in G$ est dit orthogonal à $\Pi$ si x a comme image l'unité par tous les éléments de $\Pi$. L'ensemble $\{x \in G | \forall \chi \in \Pi, x \in Ker\ \chi\}$ des éléments orthogonaux à $\Pi$ est également un sous-groupe de G, noté $\Pi^{\perp}$.*

Pour un sous-groupe $H$ de $G$, en notant $\pi$ la projection de $G$ sur $G/H$, on vérifie alors que les morphismes

$$\begin{array}{ccc} H^{\perp} & \to & \widehat{G/H} \\ \chi & \mapsto & \{\overline{\chi} : \pi(x) \mapsto \chi(x)\} \end{array} \quad \text{et} \quad \begin{array}{ccc} \widehat{G/H} & \to & H^{\perp} \\ \chi & \mapsto & \chi \circ \pi \end{array}$$

sont injectifs et donc que $\widehat{G/H}$ et $H^{\perp}$ sont deux groupes isomorphes. Associé à la proposition précédente, on en déduit en particulier que $\#H^{\perp} = (G : H)$ et $(H^{\perp})^{\perp} = H$.

#### 1.1.1.2 Relations d'orthogonalité

Par multiplicativité, on obtient le résultat, connu sous le nom de relation d'orthogonalité, au coeur des principales propriétés des objets utilisant les caractères.

**Théorème 1.1** *Soient n l'ordre de G et $\chi \in \widehat{G}$.*

$$\sum_{x \in G} \chi(x) = \begin{cases} n & si\ \chi = 1, \\ 0 & si\ \chi \neq 1. \end{cases}$$

Le théorème complète le résultat de Dedekind en montrant ainsi que les caractères sont orthogonaux entre eux suivant le produit scalaire hermitien $\langle \chi, \psi \rangle = \sum_{x \in G} \chi(x)\overline{\psi(x)}$.

En appliquant les relations précédentes dans le groupe $\widehat{G}$, on obtient de plus les relations duales :

**Corollaire 1.1** *Soit $x \in G$. On a :*

$$\sum_{\chi \in \widehat{G}} \chi(x) = \begin{cases} n & si\ x = 1, \\ 0 & si\ x \neq 1. \end{cases}$$

Nous ne pourrons résumer en quelques pages les nombreuses applications de la théorie des caractères, mais comme utilisation célèbre de cette théorie, en dehors de l'analyse harmonique, on peut néanmoins citer dans le cadre de caractères modulaires ($G = (\mathbb{Z}/m\mathbb{Z})^\times$) les fonctions $L$ introduite par Dirichlet ($L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$) dont l'utilisation combinée à leur lien avec la fonction zêta de Riemann et à ces relations d'orthogonalité permet de démontrer le théorème sur la progression arithmétique des nombres premiers :

**Théorème 1.2 (Dirichlet)** *Soient $a$ et $m$ des entiers supérieurs ou égaux à 1, premiers entre eux. Il existe une infinité de nombres premiers $p$ tels que $p \equiv a \bmod m$. De plus l'ensemble $\mathcal{P}_a = \{p \in \mathcal{P}, p \equiv a \bmod m\}$ admet une densité analytique qui est $1/\varphi(m)$, où $\varphi$ est l'indicateur d'Euler.*

**Exemple 1.2** *Une autre application classique est l'utilisation de sommes d'exponentielles pour déterminer le nombre de solutions d'une équation. Par exemple si on cherche le nombre de zéros $Z(f)$ de l'équation $f(x) = c$ pour $f : X \to G$ alors en posant $S(f,\chi) = \sum_{x\in X} \chi(f(x))$ pour $\chi \in \hat{G}$, on obtient par orthogonalité*

$$Z(f) = \frac{1}{\#G} \sum_{\chi\in\hat{G}} S(f,\chi)\overline{\chi}(c).$$

### 1.1.2 Caractères des corps finis

Soit $K = \mathbb{F}_q$ un corps fini à $q = p^m$ éléments ($m > 0$ entier et $p$ premier), alors étant donné que $K$ peut être relié à deux structures de groupe abélien fini différentes, deux types de caractères sont associés à $K$ et la distinction est importante.

#### 1.1.2.1 Caractères additifs

**Définition 1.3** *Un caractère additif du corps $K$ est un caractère du groupe additif $(K, +)$.*

La caractéristique de $K$ étant $p$, on part de la trace de $K = \mathbb{F}_q$ sur $\mathbb{F}_p$, soit $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p} : \mathbb{F}_q \to \mathbb{F}_p$, puis on définit

$$\mu_1 : x \in \mathbb{F}_q \quad \mapsto \quad e^{2i\pi\frac{\mathrm{Tr}(x)}{p}}.$$

Par linéarité de la trace, c'est bien un caractère additif de $K$ et il génère l'ensemble des caractères additifs.

**Proposition 1.3** *Pour tout $a \in \mathbb{F}_q$, $\mu_a : x \mapsto \mu_1(ax)$ est un caractère additif de $K$ et tous les caractères additifs sont de ce type.*

Le caractère $\mu_1$ est appelé le caractère additif canonique de $K$ et est généralement noté $\mu_K$. D'autre part, par transitivité de la trace lorsqu'on considère une extension finie de $K$, disons $L$, on a $\mathrm{Tr}_{L/\mathbb{F}_p} = \mathrm{Tr}_{K/\mathbb{F}_p} \circ \mathrm{Tr}_{L/K}$ et ainsi les caractères canoniques sur $K$ et $L$ sont liés par la relation

$$\mu_L = \mu_K \circ \mathrm{Tr}_{L/K}.$$

Cette relation pourra s'avérer utile par la suite.

**Remarque 1.1** *Une autre méthode pour traiter les caractères additifs d'une extension L de K est de représenter L, par isomorphisme, comme le K-espace vectoriel de dimension n = [L : K] et d'utiliser une forme K-bilinéaire, φ, non dégénérée sur $K^n$. Alors l'ensemble des caractères de $K^n$ est*

$$\{\mu_K \circ \varphi(.,a) : x \in K^n \mapsto e^{2i\pi \mathrm{Tr}_{K/\mathbb{F}_p}(\varphi(x,a))/p} \mid a \in K^n\}.$$

*Cette analogie permettra d'exploiter des outils différents au besoin (cf. Définition 1.7). En particulier, on peut choisir de représenter arbitrairement les caractères additifs de K sous la forme $x \mapsto e^{2i\pi \frac{\mathrm{Tr}(ax)}{p}}$ pour x et a vu comme éléments du corps K ou bien du type $y \mapsto e^{2i\pi \frac{b \cdot y}{p}}$ pour y et b vu comme éléments du $\mathbb{F}_p$-espace vectoriel K (où $(y, b) \mapsto y \cdot b$ est un produit scalaire sur $\mathbb{F}_p^m$).*

#### 1.1.2.2 Caractères multiplicatifs

**Définition 1.4** *Un caractère multiplicatif du corps K est un caractère du groupe multiplicatif $K^\times$.*

Le groupe $K^\times$ étant cyclique, d'ordre $q - 1$, ses caractères sont faciles à décrire :

**Proposition 1.4** *Soit α un élément générateur de $K^\times$. Pour tout $j \in \{0, \ldots, q-2\}$, la fonction $\chi_j$ définie par*

$$\forall k \in \{0, \ldots, q-2\}, \chi_j(\alpha^k) = e^{2i\pi \frac{jk}{q-1}}$$

*est un caractère multiplicatif de K et réciproquement tout élément de $\widehat{K^\times}$ s'écrit sous cette forme.*

On remarque que le caractère $\psi_0$ correspond ainsi au caractère trivial.

Les caractères quadratiques sont des cas particuliers des caractères multiplicatifs. Soit $p$ un nombre premier impair. Le caractère quadratique de $\mathbb{F}_q^\times$ est l'unique caractère $\eta_q$ multiplicatif non trivial d'ordre deux. Il est à valeurs dans $\{\pm 1\}$ et prend la valeur 1 en $c \in \mathbb{F}_q^\times$ si et seulement si $c$ est un carré dans $\mathbb{F}_q$. D'après la proposition ci-dessus, on a en fait $\eta_q = \psi_{(q-1)/2}$ et si $q$ est premier (i.e. $q = p$), on retrouve le symbole de Legendre :

$$\eta_p(c) = \left(\frac{c}{p}\right).$$

### 1.1.3 Transformée de Fourier

La théorie des caractères nous amène maintenant sur le chemin de l'analyse de Fourier. Les résultats énoncés ici seront des conséquences plus ou moins directes des relations d'orthogonalité. Soit $G$ un groupe abélien fini, noté additivement, et d'ordre $n$.

**Définition 1.5** *Soient $f : G \to \mathbb{C}$ et $F : \hat{G} \to \mathbb{C}$. La transformée de Fourier de $f$ est l'application $\hat{f} : \hat{G} \to \mathbb{C}$ définie par*

$$\forall \mu \in \hat{G},\ \hat{f}(\mu) = \sum_{x \in G} f(x)\mu(x).$$

*La transformée de Fourier inverse de $F$ est l'application $\check{F} : G \to \mathbb{C}$ définie par*

$$\forall x \in G,\ \check{F}(x) = \frac{1}{n} \sum_{\mu \in \hat{G}} F(\mu)\overline{\mu(x)}.$$

Par isomorphisme de $G$ avec son dual, il est également courant de rencontrer la notation $\hat{f}(a)$ pour la transformée de $f$ en $\mu$ si $\phi(a) = \mu$ pour un isomorphisme $\phi : G \to \hat{G}$.

**Remarque 1.2** *La définition des caractères et par conséquent celle des transformées de Fourier se généralisent à un groupe abélien quelconque. On retrouve alors pour le groupe $(\mathbb{R}, +)$ la définition usuelle de la transformée de Fourier liée aux caractères $t \mapsto e^{itx}$ pour $x \in \mathbb{R}$.*

Ces deux transformations sont clairement $\mathbb{C}$-linéaires et d'après les relations d'orthogonalité (cf. Section 1.1.1.2), on sait qu'elles sont inverses l'une de l'autre.

**Proposition 1.5** *Soient $f : G \to \mathbb{C}$ et $F : \hat{G} \to \mathbb{C}$, alors $\check{\hat{f}} = f$ et $\hat{\check{F}} = F$.*

Cela souligne en particulier le fait que l'ensemble des caractères de $G$ forme une base orthogonale de $\mathbb{C}^G$.

La transformée de Fourier permet de changer l'espace de représentation des fonctions et par là même de transformer une opération en une opération duale. Principalement, les produits de convolution $\star$ et de corrélation $\times$ sur des fonctions deviennent de simples produits sur les transformées :

**Proposition 1.6** *Soient $f, g : G \to \mathbb{C}$ et $\mu \in \hat{G}$,*

$$\widehat{f \times g}(\mu) = \overline{\hat{f}(\mu)}\hat{g}(\mu),$$

$$\widehat{f \star g}(\mu) = \hat{f}(\mu)\hat{g}(\mu),$$

*où $f \star g(y) = \sum_{x \in G} f(x)g(y - x)$ et $f \times g(y) = \sum_{x \in G} \overline{f(x)}g(y + x)$ pour $y \in G$.*

Enfin, rappelons deux formules qui sont à l'origine de nombreux résultats (e.g. en théorie des codes, les identités de MacWilliams [MS77] en sont de bons exemples). Elles nous seront très utiles dans le chapitre suivant.

**Proposition 1.7 (Parseval-Plancherel)** *Soit $H$ un sous-groupe de $G$ et $f : G \to \mathbb{C}$,*

$$\sum_{h \in H} f \times f(h) = \frac{1}{(G : H)} \sum_{\mu \in H^\perp} |\hat{f}(\mu)|^2.$$

*En particulier, on obtient l'identité de Parseval,*

$$\|\hat{f}\|_2^2 = \sum_{\mu \in \hat{G}} |\hat{f}(\mu)|^2 = \#G \sum_{x \in G} |f(x)|^2 = \#G\|f\|_2^2.$$

**Proposition 1.8 (Formule de Poisson)** *Soit $H$ un sous-groupe de $G$ et $f : G \to \mathbb{C}$,*

$$\sum_{h\in H} f(h) = \frac{1}{(G:H)} \sum_{\mu\in H^\perp} \hat{f}(\mu).$$

Formule dont on déduit clairement une généralisation à la sommation sur un coset (un translaté de $H$) :

**Corollaire 1.2** *Soit $H$ un sous-groupe de $G$, $x \in G$ et $f : G \to \mathbb{C}$,*

$$\sum_{h\in x+H} f(h) = \frac{1}{(G:H)} \sum_{\mu\in H^\perp} \hat{f}(\mu)\overline{\mu(x)}.$$

### 1.1.4 Sommes de Gauss

Un cas très intéressant de transformées de Fourier est le cadre des sommes de Gauss. Pour la suite, nous nous restreindrons à $K = \mathbb{F}_q$ un corps fini. Les principaux résultats sont détaillés avec clarté dans [LN88].

**Définition 1.6** *Soient $\mu$ un caractère additif et $\chi$ un caractère multiplicatif de $K$. La somme de Gauss sur $K$ de $\mu$ et $\chi$ est définie par*

$$G_K(\chi,\mu) = \sum_{x\in K^\times} \chi(x)\mu(x).$$

Lorsque $\mu$ est le caractère additif canonique de $K$, i.e. $\mu_K$, pour simplifier on notera

$$G_K(\chi,\mu_K) = \tau_K(\chi).$$

La somme de Gauss $G_K(\chi,\mu)$ est en fait la valeur en $\chi$ de la transformée de Fourier de $\mu_{|K^\times}$ sur $K^\times$ ou encore la valeur en $\mu$ de la transformée de Fourier de $\chi$ sur $K$ (avec la convention $\chi(0) = 0$). C'est ainsi un outil important pour passer de la structure additive à la structure multiplicative de $K$ (et inversement).

**Proposition 1.9** *Si les caractères $\mu \in \hat{K}$ et $\chi \in \widehat{K^\times}$ sont non triviaux alors la somme de Gauss $G_K(\chi,\mu)$ est de module $\sqrt{q}$. Sinon,*

$$G_K(\chi,\mu) = \begin{cases} q-1 & si\ \chi = 1\ et\ \mu = 1, \\ -1 & si\ \chi = 1\ et\ \mu \neq 1, \\ 0 & si\ \chi \neq 1\ et\ \mu = 1. \end{cases}$$

Parmi les résultats essentiels, on dispose du théorème de Davenport-Hasse qui permet de faire le lien entre les sommes de Gauss et celles sur une extension de corps. Soient $\chi$ un caractère multiplicatif et $\mu$ un caractère additif de $K$, soit $L$ une extension finie de $K$, alors les relevés $\chi' = \chi \circ \mathrm{N}_{L/K}$ et $\mu' = \mu \circ \mathrm{Tr}_{L/K}$ sont des caractères de $L$, respectivement multiplicatif et additif.

**Théorème 1.3 (Davenport-Hasse)**

$$-G_L(\chi',\mu') = (-G_K(\chi,\mu))^{[L:K]}$$

Une des preuves possibles est d'utiliser le lien entre la fonction zêta associée aux caractères sur $K$ et celle associée aux caractères sur $L$.

Dans certains cas, on sait déterminer la valeur des sommes de Gauss de manière plus précise que dans la proposition 1.9.

**Théorème 1.4 (Stickelberger)** *Soient $\chi$ un caractère multiplicatif non trivial de $\mathbb{F}_{q^2}$ d'ordre $m|q+1$, alors les sommes de Gauss sont rationnelles et*

$$\tau_{\mathbb{F}_{q^2}}(\chi) = \begin{cases} q & \text{si } m \text{ impair ou } \frac{q+1}{m} \text{ pair}, \\ -q & \text{si } m \text{ pair et } \frac{q+1}{m} \text{ impair}. \end{cases}$$

Ce résultat est l'occasion de penser à un autre cas très connu dont la forme du résultat est ressemblante, bien que distincte : le problème de la détermination des sommes quadratiques de Gauss, étudié par Gauss [Gau07] (à la base de sa théorie de construction de polygones réguliers à la règle et au compas) et qui a initié la théorie des sommes du même nom (le lecteur trouvera la forme générale de son résultat dans [LN88, Theorem 5.15]). Pour en savoir plus sur les périodes de Gauss, je conseille au lecteur intéressé de se tourner vers des ouvrages spécifiques (e.g. [BEW98]).

Mises à part les sommes rationnelles, Langevin explique dans [Lan97] et [Lan99] comment calculer certaines sommes de Gauss quadratiques. Par exemple, si $m$ est une puissance d'un nombre premier $l$ et que $p$ la caractéristique de $K$ engendre les carrés du groupe $(\mathbb{Z}/m\mathbb{Z})^\times$, alors pour $\chi \in \widehat{K^\times}$ d'ordre $m$ (nécessairement quadratique car $p$ engendre un sous-groupe d'indice 2), on sait déterminer des entiers $h$, $a$ et $b$ tels que $\tau_K(\chi) = 2^h(a+b\sqrt{-l})$. L'article [Lan97] explique ce résultat en détail ; nous nous en servirons dans la construction d'un exemple spécifique au chapitre 5. A noter en parallèle que dans [Lan99], il est souligné que la caractérisation en elle-même de l'ordre des caractères pour lesquels les sommes sont quadratiques reste un problème ouvert.

#### 1.1.4.1 Congruences de Stickelberger

Pour finir, un résultat plus général permettant d'estimer la valeur d'une somme de Gauss, dont nous nous servirons plus tard, est appelé congruences de Stickelberger (cf. [Lan90a]). Il est lié au théorème du même nom en théorie algébrique des nombres, qui fournit un idéal annulateur du groupe de classe d'un corps cyclotomique.

Soient $q = p^f$ la puissance d'un nombre premier, $a \in \mathbb{Z}$ et $a = a_0 + a_1^p + \cdots + a_{f-1}p^{f-1}$ sa décomposition $p$-adique modulo $q-1$, on définit alors $s_p(a) = \sum_i a_i$ et $\gamma_p(a) = (-1)^f \prod_{i=0}^{f-1} a_i!$. L'idée est d'étudier la décomposition en idéaux premiers des sommes de Gauss dans une extension cyclotomique. Soient $\mathcal{P}$ un idéal premier au dessus de $p$ dans $\mathbb{Z}[\xi_p, \xi_{q-1}]$ (pour $\xi_n$ une racine primitive $n$-ième de l'unité), $\pi$ un générateur de $\mathcal{P}$ et $\chi_\mathcal{P}$ le caractère multiplicatif associé à $\mathcal{P}$ (connu sous le nom de caractère de Teichmüller, c'est en fait le caractère multiplicatif d'ordre $q-1$ du quotient $K = \mathbb{Z}[\xi_p, \xi_{q-1}]/\mathcal{P}$ qui stabilise la classe de $\xi_{q-1}$ – je renvoie le lecteur vers [Lan99] où plus de détails sont facilement accessibles – c'est un générateur du groupe des caractères multiplicatifs de $\mathbb{F}_q$).

**Théorème 1.5 (Congruences de Stickelberger)**

$$\tau_K(\chi_{\mathcal{P}}^{-a}) \equiv \frac{(-1)^{s_p(a)}\pi^{s_p(a)}}{\gamma_p(a)} \mod \mathcal{P}^{s_p(a)+1}$$

Ce théorème trouve de plus une forme encore plus générale dans la formule de Gross-Koblitz [GK79] qui peut être utilisée pour des développements $p$-adique d'ordre supérieur. Ces deux théorèmes sont très utiles pour l'étude de problèmes de divisibilité, ils permettent par exemple d'étudier le poids de codes cycliques irréductibles (dans l'esprit du théorème de McEliece) comme cela est exploité dans [AL05].

## 1.2 Fonctions booléennes

Ces principaux outils étant rappelés, nous revenons maintenant au sujet principal qui nous occupe ici. Nous nous intéressons à l'étude de fonctions définies sur $\{0,1\}^m$ ($m > 0, m \in \mathbb{N}$) et à valeurs dans $\{0,1\}$, i.e. des fonctions booléennes. L'objet de cette partie est d'introduire la problématique étudiée dans les chapitres suivants, à savoir l'étude de la non-linéarité des fonctions booléennes. Cependant, en présentant la théorie des fonctions booléennes, nous aurons l'occasion d'entrevoir d'autres propriétés ou problèmes intéressants, mais sans nécessairement entrer dans les détails. Le lecteur pourra consulter [Car06a] pour avoir une présentation plus en profondeur de ces sujets.

### 1.2.1 Généralités

L'ensemble $\{0,1\}$ sera très souvent muni de sa structure de corps à 2 éléments, $\mathbb{F}_2$, et par isomorphisme, l'espace de départ pourra être vu soit comme un $\mathbb{F}_2$-espace vectoriel de dimension $m$ soit comme le corps à $2^m$ éléments ; ceci dépendra de la structure adéquate à prendre pour les propriétés étudiées. On notera $\mathcal{BF}(m)$ l'ensemble des fonctions booléennes de $\mathbb{F}_2^m \simeq \mathbb{F}_{2^m}$ dans $\mathbb{F}_2$, alors $\#\mathcal{BF}(m) = 2^{2^m}$ et $(\mathcal{BF}(m), +, \times, .)$ est une $\mathbb{F}_2$-algèbre.

On munit $\mathcal{BF}(m)$ d'une distance $d_H$ en définissant le poids de $f \in \mathcal{BF}(m)$, $w_H(f)$, comme la taille de son support (i.e. le poids binaire du vecteur vérité ou vecteur image de $f$) et on définit $d_H(f,g)$ ($f, g \in \mathcal{BF}(m)$) comme le nombre de différences entre leurs vecteurs image respectifs, i.e. $w_H(f+g)$.

Par isomorphisme, toute fonction $f$ de $\mathcal{BF}(m)$ peut être vue dans $\mathbb{F}_2[X_1,\ldots,X_m]/(X_1^2 - X_1,\ldots,X_m^2 - X_m)$, cette représentation est définie comme la forme algébrique normale (ANF) de $f$. Pour $x = (x_1,\ldots,x_m) \in \mathbb{F}_2^m$ et $u = (u_1,\ldots,u_m) \in \{0,1\}^m$, on désigne par $x^u$ le produit $x_1^{u_1}\cdots x_m^{u_m}$. De plus, pour $u, v \in \{0,1\}^m$, on écrira $v \subset u$ si $v_i \leq u_i$ pour tout $i \in \{1,\ldots,m\}$. On peut alors en dire plus sur l'ANF :

**Lemme 1.2** *Soit $f \in \mathcal{BF}(m)$ alors il existe un unique $2^m$-uplet $(a_u)_{u\in\{0,1\}^m}$ tel que pour tout $x \in \mathbb{F}_2^m$,*

$$f(x) = \sum_{u\in\{0,1\}^m} a_u x^u.$$

*Pour $u \in \{0,1\}^m$, on a $a_u = \sum_{v\subset u} f(v)$.*

Le degré de l'ANF d'une fonction booléenne est appelé le degré algébrique de cette fonction. L'ensemble des fonctions booléennes de degré inférieur ou égal à un entier $r \leq m$ (si $r = 1$ on parle de fonctions affines) est alors le code de Reed-Muller binaire de degré $r$ et d'ordre $m$, noté $RM(r, m)$. L'ANF permet donc de faire le lien entre les fonctions booléennes et la théorie des codes. A noter que le code $RM(1, m)$, ensemble des fonctions affines, est important en cryptographie puisque une fonction hautement non-linéaire se veut en être éloignée le plus possible.

D'autres représentations existent dont la représentation trace d'un polynôme à coefficients dans $\mathbb{F}_{2^m}$ ou la forme numérique normale (NNF). La première représentation permet de travailler avec des valeurs du corps $\mathbb{F}_{2^m}$ et ainsi d'en utiliser plus aisément la structure (voir notamment les travaux de Wolfmann) ; c'est par exemple la représentation associée à l'étude des fonctions monômes, appelées fonctions puissances (cf. Annexe B.1 ou par exemple [Fel05] qui exploite la structure de corps). En cryptographie, l'AES [Nat01] utilise ainsi une fonction puissance (la fonction inverse, pour la nommer) à la base de la construction de ses boîtes S. Concernant la deuxième représentation, contrairement à l'ANF, il s'agit d'une représentation polynomiale à coefficients à valeurs réelles [CG99]. Elle permet de caractériser efficacement plusieurs critères cryptographiques de choix de fonctions booléennes.

### 1.2.2 Transformées de Fourier d'une fonction booléenne

Selon la structure, de corps fini ou d'espace vectoriel de $\mathbb{F}_2^m \simeq \mathbb{F}_{2^m}$, choisie, deux définitions de la transformée de Fourier sont possibles ; elles sont équivalentes.

Du point de vue corps fini, soit $\mu_K$ le caractère additif canonique de $K = \mathbb{F}_{2^m}$, i.e. $\mu_K : x \mapsto (-1)^{\mathrm{Tr}_{K/\mathbb{F}_2}(x)}$ et $\mu_a : x \mapsto \mu_K(ax)$ pour $a \in K$, alors la transformée de Fourier d'une application $h : \mathbb{F}_{2^m} \to \mathbb{C}$ est définie dans ce contexte par

$$\hat{h}(\mu_a) = \sum_{x \in K} h(x)\mu_a(x) = \sum_{x \in K} h(x)(-1)^{\mathrm{Tr}_K(ax)},$$

pour $a \in K$. Pour simplifier, on note souvent $\hat{h}(a)$ au lieu de $\hat{h}(\mu_a)$.

Du point de vue espace vectoriel, les caractères sont obtenus à partir d'un produit scalaire, $\mu_a : x \in \mathbb{F}_2^m \mapsto (-1)^{\langle a,x \rangle} = (-1)^{a \cdot x}$, et la transformée de Fourier de $h : \mathbb{F}_2^m \to \mathbb{C}$ en $a \in \mathbb{F}_2^m$ vaut

$$\hat{\mathrm{h}}(a) = \sum_{x \in \mathbb{F}_2^m} h(x)(-1)^{a \cdot x}.$$

Les deux définitions donneront le même ensemble de valeurs spectrales mais dans des ordres différents. On peut appliquer ces transformations directement à une fonction booléenne $f$, on parle alors de transformée de Fourier de $f$, mais il est d'usage de l'appliquer à la fonction signe de $f$, $f_\chi = (-1)^f$ (c'est une fonction dite binaire, i.e. à valeurs dans $\{-1, 1\}$). Cette transformée est appelée transformée de Walsh ou de Walsh-Hadamard, mais par raccourci de langage on rencontre également le terme transformée de Fourier lorsqu'aucune confusion n'est possible. On obtient ainsi en fonction de la structure utilisée,

**Définition 1.7** *Soit $f \in \mathcal{BF}(m)$. D'un point de vue corps fini, la transformée de Walsh de $f$ est définie par*

$$\forall a \in \mathbb{F}_{2^m}, \hat{f}_\chi(a) = \sum_{x \in \mathbb{F}_{2^m}} (-1)^{f(x) \oplus \mathrm{Tr}_K(ax)}.$$

*D'un point de vue espace vectoriel, elle est définie par*

$$\forall a \in \mathbb{F}_2^m, \widehat{\mathrm{f}_\chi}(a) = \sum_{x \in \mathbb{F}_2^m} (-1)^{f(x) \oplus a \cdot x}.$$

Les propriétés qui en découlent dans la suite seront valables dans les 2 cas, sauf mention du contraire, et pour simplifier nous énoncerons donc ces résultats seulement pour la version "corps fini".

Il est immédiat que les coefficients de Walsh sont tous des entiers, nous verrons ci-dessous qu'ils sont de plus tous pairs (cf. Proposition 1.11). De ces définitions, de nombreuses relations utiles pour étudier les fonctions booléennes sont héritées de la théorie de Fourier – comme par exemple celles détaillées dans la section 1.1.3, dont la relation de Parseval ou la formule de Poisson. L'utilisation de la formule de Poisson permet par exemple de mettre en évidence un lien entre le degré d'une fonction et la divisibilité des ses coefficients de Walsh (cf. [Lan90b]).

**Proposition 1.10** *Soient $f \in \mathcal{BF}(m)$ et $s$ un entier entre 1 et $m$.*
- *Si pour tout $a \in \mathbb{F}_{2^m}$, $2^s | \hat{f}_\chi(a)$ alors* $\deg(f) \leq m - s + 1$.
- *Si pour tout $a \in \mathbb{F}_{2^m}$, $2^s \| \hat{f}_\chi(a)$ (i.e. $2^s | \hat{f}_\chi(a)$ et $2^{s+1} \nmid \hat{f}_\chi(a)$) alors* $\deg(f) \leq m - s$.

La théorie de Fourier est en fait l'un des principaux outils pour l'étude des fonctions booléennes, et spécifiquement de la non-linéarité comme nous allons le voir ci-après.

### 1.2.3 Rayon de recouvrement de $RM(1, m)$ et non-linéarité

La non-linéarité d'une fonction booléenne $f \in \mathcal{BF}(m)$, notée $\mathrm{nl}(f)$, est la distance de Hamming minimale entre $f$ et les fonctions affines de $\mathcal{BF}(m)$, i.e. la distance de $f$ au code de Reed-Muller de degré 1, $RM(1, m)$. Cette distance peut être calculée en utilisant la transformée de Walsh.

**Proposition 1.11** *Soient $f \in \mathcal{BF}(m)$, $a \in \mathbb{F}_{2^m}, b \in \mathbb{F}_2$, et $l_{a,b}$ l'application affine définie par $l_{a,b}(x) = \mathrm{Tr}_K(ax) + b$ pour $x \in \mathbb{F}_{2^m}$, alors*

$$d_H(f, l_{a,b}) = 2^{m-1} - \frac{(-1)^b}{2} \hat{f}_\chi(a).$$

*En particulier, $w_H(f) = 2^{m-1} - \frac{1}{2}\hat{f}_\chi(0)$ et*

$$\mathrm{nl}(f) = d_H(f, RM(1, m)) = 2^{m-1} - \frac{1}{2} \|\hat{f}_\chi\|_\infty,$$

*où*

$$\|g\|_\infty = sup_{x \in E} |g(x)|$$

*représente la norme sup d'une fonction $g$ définie sur un ensemble $E$ et à valeurs dans $\mathbb{C}$.*

De plus, la non-linéarité $\mathrm{nl}(f)$ est évidemment invariante sous l'action des transformations affines sur $f$.

On déduit de la proposition que la distance maximale entre une fonction booléenne et le code $RM(1,m)$ vérifie

$$\rho(m) = \sup_{f\in\mathcal{BF}(m)} \mathrm{nl}(f) = 2^{m-1} - \frac{1}{2} \inf_{f\in\mathcal{BF}(m)} \|\hat{f}_\chi\|_\infty.$$

Cette notion correspondant à trouver le mot de $\mathcal{BF}(m)$ le plus éloigné de $RM(1,m)$ est la détermination du rayon de recouvrement $\rho(m)$ du code $RM(1,m)$. En effet, connaître $\rho(m)$ implique qu'on sait que quelque soit le mot de départ, alors il existe au moins un mot du code à une distance inférieure ou égale à $\rho(m)$. Une connaissance précise de ce paramètre peut permettre d'améliorer les algorithmes de décodage du code. Ceci est donc une raison importante pour étudier de manière approfondie la non-linéarité des fonctions booléennes, d'autant plus qu'à part le rayon de recouvrement[1], les autres paramètres (longueur, dimension, distance minimale) de $RM(1,m)$ sont faciles à déterminer. Une autre motivation majeure depuis quelques années est l'utilisation des fonctions booléennes en cryptographie comme nous l'expliquerons plus loin (cf. Section 1.2.4).

L'égalité ci-dessus entraîne que l'étude du rayon de recouvrement se ramène à la recherche de la plus faible amplitude spectrale quand $f$ parcourt l'ensemble des fonctions booléennes. Par analogie, on parle de rayon spectral, noté $R(m)$ et défini par

$$R(m) = \inf_{f\in\mathcal{BF}(m)} \|\hat{f}_\chi\|_\infty.$$

D'après l'identité de Parseval, on sait que pour $f \in \mathcal{BF}(m)$,

$$\sum_{a\in\mathbb{F}_{2^m}} \hat{f}_\chi(a)^2 = 2^{2m},$$

ce qui implique que $\|\hat{f}_\chi\|_\infty \geq 2^{\frac{m}{2}}$, soit

$$\mathrm{bp}(m) \leq R(m) \tag{1.1}$$

où $\mathrm{bp}(m) = 2^{\frac{m}{2}}$ est la borne dite de Parseval. Elle est valable pour toutes les fonctions booléennes sans restriction. Pour la non-linéarité, cela devient une borne supérieure qui s'écrit $\mathrm{nl}(f) \leq 2^{m-1} - 2^{m/2-1}$. Les fonctions admettant une non-linéarité maximale (donc au plus $2^{m-1} - 2^{m/2-1}$) sont appelées fonctions non-linéaires maximales et les fonctions s'approchant presque au plus près de la borne maximale sont des fonctions hautement non-linéaires. Il est également intéressant de connaître la non-linéarité maximale de certains sous-ensembles de $\mathcal{BF}(m)$ (par exemple pour l'ensemble des fonctions équilibrées, on parle d'étude du rayon spectral *équilibré* – cf. Section 5.2).

On aimerait alors connaître la distribution de la non-linéarité des fonctions booléennes. Les premiers cas sont observables expérimentalement (cf. [Mai91] pour le cas $n = 6$) mais

[1] De manière plus ou moins anecdotique (cela dépendra du lecteur, concerné ou non), le rayon de recouvrement d'un code permet d'optimiser le nombre de grille à remplir pour des paris sportifs [Zan01].

pour $n > 6$, cela devient vite irréalisable ($\#\mathcal{BF}(m) = 2^{2^m}$). En 1998, Olejár et Stanek [OS98] ont démontré que les fonctions booléennes à $m$ variables ont presque toutes une non-linéarité supérieure à $2^{m-1} - c\sqrt{m}2^{m/2-1}$, où $c = \sqrt{2(1+\epsilon)\ln 2}$ avec $\epsilon > 0$ quelconque. Dans [Rod03a] (voir aussi [Rod06]), Rodier a confirmé que cette borne était la meilleure en exhibant une borne inférieure de telle sorte que les fonctions booléennes $f \in \mathcal{BF}(m)$ admettent presque toutes une non-linéarité $\mathrm{nl}(f)$ asymptotiquement proche de $2^{m-1} - 2^{m/2-1}\sqrt{2m\ln 2}$.

Une question qui se pose en particulier est de déterminer si la borne (1.1) est atteinte ou, si non, quelle est la non-linéarité maximum possible. D'après l'identité de Parseval, l'égalité est réalisée dans (1.1) si et seulement si tous les coefficients de Walsh sont de module $2^{m/2}$, ce qui n'est possible que pour $m$ pair, puisque les coefficients sont des entiers. Ce cas a été introduit et étudié par Rothaus [Rot76] et Dillon [Dil74] :

**Proposition 1.12** *Soit m pair. Une fonction $f \in \mathcal{BF}(m)$ est non-linéaire maximale si et seulement si*

$$\forall a \in \mathbb{F}_{2^m},\ |\hat{f}_\chi(a)| = 2^{\frac{m}{2}}.$$

*On dit que $f$ est courbe ; sa non-linéarité est maximale et vérifie* $\mathrm{nl}(f) = 2^{m-1} - 2^{m/2-1}$.

Comme le cas des fonctions quadratiques l'illustrera ci-dessous, quelque soit $m$ pair, il existe toujours des fonctions courbes ; ce qui démontre que pour $m$ pair le rayon de recouvrement de $RM(1,m)$ est $\rho(m) = 2^{m-1} - 2^{m/2-1}$ et son rayon spectral est $R(m) = 2^{m/2}$.

#### 1.2.3.1 Formes quadratiques

Soit $q : K = \mathbb{F}_{2^m} \to \mathbb{F}_2$ une forme quadratique nulle en 0. Il existe une forme bilinéaire symplectique $\phi$, associée à $q$, telle que

$$\forall x, y \in K,\ \ q(x+y) = q(x) + q(y) + \phi(x,y).$$

Par définition le radical (ou noyau) de $q$ est le sous-espace vectoriel $\mathrm{Ker}\, q = \{x \in K / \forall y \in K, \phi(x,y) = 0\}$ de dimension $d = \dim \mathrm{Ker}\, q$. On sait que la co-dimension du noyau, i.e. le rang de la forme quadratique, est toujours paire. La forme sera alors de rang maximal si son noyau est de dimension 0 pour $m$ pair ou 1 pour $m$ impair. De plus, la restriction de $q$ à son noyau est une forme linéaire, soit nulle, soit équilibrée. Si la restriction est non-nulle, on dit qu'un point est un défaut de $q$ si c'est un élément $x$ du noyau tel que $q(x) = 1$ ; on note $\mathrm{def}\,(q)$ l'ensemble des défauts de $q$.

La transformée de Walsh de $q$ en $a \in K$ se calcule facilement. On a

$$\begin{aligned}
\hat{q_\chi}(a)^2 &= \sum_{x,y\in K} \mu_K(q(x) + ax + q(y) + ay) \\
&= \sum_{x,y\in K} \mu_K(q(x+y) + \phi(x,y) + a(x+y)) \\
&= \sum_{x\in K} \mu_K(q(x) + ax) \sum_{y\in K} \mu_K(\phi(x,y)) \\
&= 2^m \sum_{x\in \mathrm{Ker}\, q} \mu_K(q(x) + ax)
\end{aligned}$$

par orthogonalité. On en déduit alors 3 valeurs différentes (toutes atteintes, sauf peut être 0, mais fonction de la valeur de $a$ par rapport à Ker $q^\perp$ et def $q$) :

$$\forall a \in K,\ \hat{q_\chi}(a) \in \{-2^{\frac{m+d}{2}}, 0, 2^{\frac{m+d}{2}}\}.$$

On distingue trois types de formes quadratiques possibles, elliptique (du type $x_0x_1 + x_0 + x_1$), parabolique (du type $x_0x_1 + x_2$) ou hyperbolique (du type $x_0x_1$). Le poids de la forme quadratique dépend par exemple exclusivement du type de la forme et de la dimension de son noyau.

En choisissant une forme quadratique non dégénérée dans le cas pair, on obtient une fonction courbe (i.e. de non-linéarité maximale égale à $2^{m-1} - 2^{m/2-1}$), ce qui est possible pour tout $m$ pair. Dans le cas impair, la dimension minimale du noyau étant de 1, on en déduit la borne quadratique usuelle, à savoir qu'il existe une fonction booléenne de non-linéarité supérieure à $2^{m-1} - 2^{\frac{m+1}{2}-1}$. La borne quadratique générale est donc

$$R(m) \leq \mathrm{bq}(m) = 2^{\lfloor \frac{m+1}{2} \rfloor}.$$

Pour résumé, nous avons les inégalités suivantes.

**Proposition 1.13** *Le rayon de recouvrement de $RM(1,m)$ vérifie*

$$2^{m-1} - 2^{\lfloor \frac{m+1}{2} \rfloor - 1} \leq \rho(m) \leq 2^{m-1} - \frac{1}{2}\lceil 2^{\frac{m}{2}} \rceil \tag{1.2}$$

*et son rayon spectral se situe entre les bornes quadratique et de Parseval*

$$\mathrm{bp}(m) = \lceil 2^{\frac{m}{2}} \rceil \leq R(m) \leq 2^{\lfloor \frac{m+1}{2} \rfloor}. \tag{1.3}$$

Si $m$ est pair, ce sont toutes des égalités, atteintes pour les fonctions courbes. En plus du cas quadratique, Rothaus a démontré dans [Rot76] qu'il en existe de degré $t$ pour tout $2 \leq t \leq m/2$. Par contre, contrairement au cas quadratique, où les fonctions courbes sont toutes connues, il est difficile de caractériser les fonctions courbes d'un degré quelconque. Plus généralement, même si on connaît des constructions génériques de fonctions courbes (dont les classes de Maiorana-McFarland $\mathcal{M}$ [McF73], des *Partial Spreads* $\mathcal{PS}$ [Dil74], et la classe $\mathcal{GPS}$ [Car95], parmi les plus connues), même le nombre total de fonctions courbes reste un problème ouvert pour $m$ un entier pair quelconque (la dernière avancée sur cette question est le calcul d'une estimation pour $m = 8$ [LRVZ06]). Déterminer ce nombre et classifier les fonctions courbes est un des deux principaux challenges liés à l'étude de la non-linéarité des fonctions booléennes ; le deuxième est la détermination de la meilleure non-linéarité possible dans le cas impair.

#### 1.2.3.2 Non-linéarité en dimension impaire

Si $m$ impair, la différence entre les 2 bornes n'est pas très grande pour des petites valeurs : $\mathrm{bq}(m) - \mathrm{bp}(m) < (\sqrt{2} - 1)\sqrt{2^m}$. Pour $m \in \{1, 3, 5, 7\}$, on sait que $R(m)$ est égal à la borne quadratique : Le cas $m = 1$ est trivial ; Pour $m = 3$, de l'encadrement on déduit directement que $R(3) = 2^{\frac{3+1}{2}}$ ; Berlekamp et Welch l'ont démontré en 1972 [BW72] pour $m = 5$ ; Et en 1980, Mykkelveit [Myk80] a apporté la preuve du cas $m = 7$ (voir aussi la preuve de Hou [Hou96]).

Cependant pour $m$ quelconque, l'incertitude peut être très grande et on ne sait pas en dire beaucoup plus en général. Lorsque Mykkelveit montre que $R(7) = 2^{\frac{7+1}{2}}$ dans [Myk80], il conjecture alors que cette égalité avec la borne quadratique reste vraie pour toutes les dimensions.

Dans [PW83], Patterson et Wiedemann trouve un contre-exemple à cette conjecture en dimension 15 et montre du même temps que $R(m) < \mathrm{bq}(m)$ pour tout $m \geq 15$.

**Proposition 1.14 (Patterson-Wiedemann)** *Soient $\alpha$ et $\beta$ des racines primitives respectivement de $\mathbb{F}_{2^3}^{\times}$ et $\mathbb{F}_{2^5}^{\times}$, soit $G$ le sous-groupe de $GL(\mathbb{F}_{2^{15}})$ engendré par les applications*

$$x \mapsto \alpha x,\ x \mapsto \beta x,\ x \mapsto x^2,$$

*alors il existe une application booléenne invariante sous l'action de $G$ d'amplitude spectrale* 216.

Pour y parvenir, ils construisent en fait toutes les fonctions stables sous l'action de $G$, i.e. dont le support correspond à une réunion d'orbites, et obtiennent deux fonctions booléennes d'amplitude 216.

En particulier, cela implique $R(15) \leq 216 = \frac{27}{32}2^{7+1} < \mathrm{bq}(15)$ et plus généralement, l'inégalité $R(m + m') \leq R(m)R(m')$ implique :

**Corollaire 1.3** *Si $m$ est un entier impair supérieur à* 15*,*

$$R(m) \leq \frac{27}{32}\sqrt{2}\sqrt{2^m} = \mathrm{bpw}(m) < \mathrm{bq}(m).$$

Dans la suite, on désignera cette borne comme la borne de Patterson-Wiedemann, notée $\mathrm{bpw}(m)$. En pratique, afin de souligner la différence avec la borne quadratique, toute fonction vérifiant $\|\hat{f}\|_\infty < \mathrm{bq}(m) = 2^{(m+1)/2}$, pour $m \geq 9$, $m$ impair, sera dite *hautement non-linéaire*.

Dans le même article [PW83], même s'ils n'ont pas obtenu une construction générique à la suite du résultat sur $R(15)$, ils proposent de plus une nouvelle conjecture sur le comportement du rayon spectral de $RM(1, m)$, cette fois-ci asymptotiquement équivalent à la borne de Parseval.

**Conjecture 1.1 (Patterson-Wiedemann)** *Pour $m \to \infty$,*

$$R(m) \sim \sqrt{2^m}.$$

Depuis 1983, cette conjecture n'a pu être confirmée ni infirmée. De même, la borne de Patterson-Wiedemann est pour le moment, malgré d'autres expériences (voir par exemple [CCW85]), la meilleure connue puisque qu'aucune construction d'amplitude plus faible n'a été découverte jusqu'ici. D'autre part, même les cas $m \in \{9, 11, 13\}$ ont continué à poser problème longtemps après les travaux de Patterson et Wiedemann : jusqu'en 2006, on ne savait pas si on pouvait battre[2] la borne quadratique. Depuis 2006, on sait néanmoins que $R(9) < \mathrm{bq}(9)$ grâce à l'obtention dans [KMY06] d'une fonction d'amplitude 30<32 (voir également [KMSY06]).

[2]Le terme battre signifie ici prouver l'existence d'une fonction booléenne en $m$ variables d'amplitude strictement inférieure à $\mathrm{bq}(m)$.

Cette construction a été trouvée grâce à l'utilisation d'un algorithme de recherche par optimisation itérative, de manière similaire aux expériences de sélection de fonctions booléennes à partir d'algorithmes génétiques (voir par exemple [MCD98]) ou par recuit simulé (e.g. [CJ00]). Ici l'algorithme est une variation de la descente de gradient (avec une modification adaptée pour sortir des minima locaux), appliqué sur une famille restreinte de fonctions booléennes : les fonctions symétriques par rotations (fonctions dont les valeurs sont inchangées lorsqu'on applique une rotation quelconque sur les variables en entrée). L'espace de recherche réduit (de l'ordre de $2^{\frac{2^m}{m}}$ au lieu de $\#\mathcal{BF}(m) = 2^{2^m}$) combiné à la stratégie employée a permis dans [KMSY06] d'établir le résultat suivant :

**Lemme 1.3** *A équivalence affine près, il existe exactement deux fonctions booléennes symétriques par rotation, en* 9 *variables, de non-linéarité* $2^{9-1} - 2^{\frac{9-1}{2}} + 1 = 241$.

De plus, Kavut *et al.* montrent que cette valeur est la valeur maximum possible pour la non-linéarité des fonctions booléennes symétriques par rotation en 9 variables. Ce résultat ne reste pas vrai dans $\mathcal{BF}(9)$. En effet, en élargissant la classe de recherche, [KY07b] a prouvé très récemment l'existence de plusieurs fonctions de non-linéarité 242 (amplitude 28). Il reste ainsi à déterminer si $\rho(9) = 242,\ 243,$ ou 244.

**Proposition 1.15 ([KY07b])** *Le rayon de recouvrement et le rayon spectral de* $RM(1,9)$ *vérifient*

$$242 \leq \rho(9) \leq 244 \quad et \quad 24 \leq R(9) \leq 28.$$

Toutefois, si on suit la conjecture de [BCP92] sur la parité du rayon de recouvrement de $RM(1,m)$, cela ne laisserait plus que 2 possibilités (242 ou 244).

Une conséquence immédiate de ce résultat pour les dimensions supérieures est que la non-linéarité maximale pour $m = 11$ est au minimum 996 et 4040 pour $m = 13$. Ce qui implique,

**Corollaire 1.4** *Si m est un entier impair,* $R(m) < \mathrm{bq}(m)$ *si et seulement si* $m \geq 9$.

Il est intéressant de préciser que la fonction trouvée dans [KMY06] permet en même temps de prouver l'existence d'un *urcoset* (*orphan* ou classe latérale maximale, cf. [Lan91]) de $RM(1,m)$ de poids impair pour $m = 9$. Ce qui permet de compléter les résultats d'existence établis par Hou pour $m > 11$ (cf. [Hou90]) et de [LZ96] pour $m = 7$.

Bien sûr le problème de la détermination de la non-linéarité maximale existe aussi quand on se limite à des ensembles de fonctions particuliers : par exemple pour les fonctions équilibrées, où on parle de rayon de recouvrement équilibré $\rho_B(m)$ (et de rayon spectral équilibré $RB(m)$), la valeur maximale est inconnue même dans le cas pair (cf. Sec. 5.2.2). Un autre exemple important, expliqué en annexe B.1, sujet de nombreux travaux mais pour lequel de nombreux problèmes restent ouverts, est le cas des fonctions puissances où cette fois-ci l'indétermination se situe sur le cas pair. La non-linéarité généralisée à des fonctions multi-booléennes (ou vectorielles), i.e. à valeurs dans une extension de $\mathbb{F}_2$ est également génératrice de nombreuses études.

Avant de passer au coeur de ce mémoire, nous rappelons ci-dessous quel est l'intérêt de l'étude des fonctions booléennes pour la cryptographie.

### 1.2.4 Fonctions booléennes et cryptographie

L'étude de la non-linéarité est liée au rayon de recouvrement des codes de Reed-Muller, mais ce n'est pas le seul point à y être associé. Les fonctions booléennes jouent un rôle important en cryptographie, en particulier dans les algorithmes de chiffrement à flot ou par blocs, et la non-linéarité fait partie des nombreux critères cryptographiques de choix de fonctions.

Le rôle des fonctions booléennes dans la construction de schéma de chiffrement est l'application concrète des principes de confusion et diffusion. Compte-tenu de ces principes et des nombreuses attaques publiées sur des constructions existantes (e.g. [Sie84]), des critères de résistance ont été définis afin d'être satisfaits au mieux lors de l'élaboration d'un système de chiffrement et du choix des *fonctions booléennes cryptographiques* à utiliser (on pourra consulter l'annexe B.2 pour quelques exemples rapides de principes d'attaques sur du chiffrement à flot). Cependant tous les critères ne vont pas dans le même sens et de nombreux compromis peuvent être nécessaires (indépendamment certains critères peuvent être faciles à étudier, mais les compromis compliquent les choses). Nous faisons maintenant une brève description de ces différents points et quelques-uns de ces compromis.

**Degré algébrique.** Le degré d'une fonction booléenne doit être élevé. Par exemple, pour du chiffrement à flot, dans le cas d'une combinaison de plusieurs registres à décalage linéaires (*Linear Feedback Shift Register* – LFSR) de longueur $L_1, \ldots, L_m$ alors la suite de valeurs retournée par la fonction de combinaison $f$ est modélisable par un LFSR de longueur inférieure à $L = f(L_1, \ldots, L_m)$. Ce LFSR peut en particulier être reconstruit à partir d'au plus $2L$ valeurs via l'algorithme de Berlekamp-Massey. La construction d'une fonction de degré maximale est très simple, mais en respectant d'autres critères les contraintes sont plus fortes. Par exemple, on sait qu'une fonction courbe est de degré au plus $m/2$.

**Équilibre.** Afin que la sortie d'une fonction soit uniformément distribuée pour éviter un biais de la suite produite et donc une dépendance statistique avec l'entrée, les fonctions devront être équilibrées. Une généralisation mise en évidence par [Sie84] est que cela doit rester vrai quand on fixe certaines coordonnées à des valeurs constantes en entrée, dans le cas d'une fonction de combinaison de plusieurs LFSR. Dans le cas contraire, en fixant ces coordonnées on est à même d'observer une corrélation entre les entrées et sorties de la fonction et ceci permet de réduire la complexité d'une recherche exhaustive pour retrouver les initialisations des LFSR. L'attaque proposée par Siegenthaler a été améliorée dans de nombreux articles, pour devenir des *attaques à corrélation rapide* dont le principe repose sur l'utilisation de techniques de décodage d'erreurs (voir par exemple [MS89] et Annexe B.2).

**Définition 1.8** *Une fonction $f \in \mathcal{BF}(m)$ est dite sans corrélation d'ordre $t \geq 1$ si sa distribution de valeurs ne change pas lorsqu'on fixe au plus $t$ entrées. Si de plus $f$ est équilibrée, $f$ est dite résiliente d'ordre $t$.*

La borne de Siegenthaler nous dit alors qu'une fonction sans corrélation d'ordre $t$ est de degré au plus $m - t$. Un résultat établi par Xiao et Massey [XM88] permet d'obtenir la caractérisation en fonction de coefficients de Walsh (point de vue espace vectoriel) :

**Proposition 1.16** *Soit $f \in \mathcal{BF}(m)$ et $t \geq 1$, alors $f$ est sans corrélation d'ordre $t$ si et seulement si $\widehat{f_\chi}(u) = 0$ pour tout $u \in \mathbb{F}_2^m$ de poids compris entre 1 et $t$. De même, $f$ est résiliente d'ordre $t$ si et seulement si $\widehat{f_\chi}(u) = 0$ pour tout $u \in \mathbb{F}_2^m$ de poids compris entre 0 et $t$.*

D'après l'égalité de Parseval, ce résultat entraîne que la non-linéarité maximale possible d'une fonction sans corrélation d'ordre $t$ se dégrade quand $t$ croît.

**Non-linéarité.** La non-linéarité des fonctions booléennes utilisées en cryptographie est également un paramètre important. L'existence d'une approximation affine d'une fonction permet en effet, dans le cas de chiffrements par blocs comme ceux à flot, très souvent de construire des attaques efficaces. La non-linéarité maximale possible dépend également des autres critères cryptographiques pris en compte (e.g. une fonction équilibrée ne peut être courbe).

Notamment, Matsui a exploité le concept de cryptanalyse linéaire en 1993 [Mat93] qui consiste à approcher plusieurs rondes d'un algorithme de chiffrement itératif par blocs (e.g. type schéma de Feistel) par une application linéaire en le message à chiffrer et en la clé. Plus cette approximation se réalise de nombreuses fois, plus la clé se retrouve rapidement. Ceci n'est possible que si les transformations non-linéaires internes (en particulier les boites $S$) sont elles-mêmes approchées par des fonctions linéaires avec une bonne probabilité. Un critère de résistance d'une fonction vectorielle $S : \mathbb{F}_2^m \to \mathbb{F}_2^l$ à une attaque par cryptanalyse linéaire est alors que le paramètre

$$\lambda_S = \left| \max_{\alpha} \max_{\beta \neq 0} \#\{x \in \mathbb{F}_2^m,\ \alpha \cdot x + \beta \cdot S(x) = 0\} - 2^{m-1} \right|$$

soit le plus petit possible. En écrivant $S = (S_1, \ldots, S_l)$, cela est équivalent au fait que toutes les combinaisons linéaires des fonctions booléennes $S_i$ ont la meilleure non-linéarité possible. On pourra se reporter à l'annexe B.1 pour l'étude de ce paramètre sur un cas particulier, les fonctions puissances.

L'importance de ce paramètre pour une fonction de combinaison de registres à décalage est liée en partie aux attaques par corrélation. En effet, on peut démontrer que pour une fonction $f \in \mathcal{BF}(m)$ résiliente d'ordre $t$ alors la fonction de $t+1$ variables la plus proche de $f$ sera une fonction affine, donc la probabilité de succès d'une attaque par corrélation sur $t+1$ variables sera d'autant plus faible que $f$ sera éloignée de la fonction affine correspondante – éloignement qui correspond au minimum à la valeur de non-linéarité de $f$.

De plus, une généralisation des attaques linéaires via des approximations de petit degré $r > 1$ est possible (cf. [KR96]) d'où la nécessité d'étudier la distance minimale de $f \in \mathcal{BF}(m)$ à l'ensemble des fonctions booléennes de degré inférieur ou égal à $r$, i.e. $d_H(f, RM(r,m))$. Cette distance est appelée la non-linéarité d'ordre $r$ de $f$ et est notée $\mathrm{nl}_r(f)$ ; la non-linéarité maximale d'ordre $r$ correspond au rayon de recouvrement du code de Reed-Muller d'ordre $r$ mais contrairement au cas $r = 1$, cette fois-ci le problème dual (détermination du rayon spectral, généralisé) n'est plus aussi naturel, puisqu'il ne s'exprime pas uniquement en fonction des coefficients de Walsh de $f$ (ce qui rend le problème encore plus difficile à étudier, y compris le calcul de $\mathrm{nl}_r$ pour

une fonction donnée). L'étude du profil de non-linéarité, i.e. des valeurs $\mathrm{nl}_r(f)$ pour $r \in \{1, \ldots, m-1\}$, est un critère plus général que la non-linéarité d'ordre 1 puisqu'il permet entre autre de connaître le degré de $f$, mais il n'est pas utile d'y accorder la même importance à tous les ordres ; comme on l'a vu la non-linéarité d'ordre 1 doit être optimale pour éviter les attaques linéaires mais quand l'ordre augmente, la complexité des attaques augmentant, plus de souplesse est possible. De nombreux articles s'intéressent au problème sous un angle cryptographique (par exemple d'un point de vue général [Car06c], ou en le reliant à d'autres critères, comme la résilience [KIY04, BBNP05] ou l'immunité algébrique [Car05, Car06b]) ou d'un point de vue théorie des codes (e.g. [Lan90b, BL03, LRVZ06]).

**Propagation.** L'effet avalanche est une autre propriété importante dans les fonctions cryptographiques (dont les fonctions de hachage et les algorithmes de chiffrement par blocs). L'idée est d'obtenir des modifications de plus en plus importantes au fur et à mesure que les données se propagent dans la structure de l'algorithme, et ainsi d'obtenir des résultats en sortie totalement différents avec très peu de perturbations en entrée ; c'est donc à mettre en parallèle avec le principe de diffusion. Le critère d'avalanche stricte (*Strict Avalanche Criterion – SAC*) [WT85] et sa généralisation, le critère de propagation (*Propagation Criterion – PC*) [PVV+90] ont été introduits afin de mesurer cet effet sur une fonction booléenne.

D'après le concept de cryptanalyse différentielle introduit sur le DES par Biham et Shamir [BS90] qui permet de retrouver de l'information sur la clé d'un système de chiffrement par blocs, en exploitant la connaissance de couples de chiffrés dont les messages clairs sont connus et ont une différence fixée, on sait qu'un tel système ne doit pas avoir de différentielle (différence fixée en entrée donnant une deuxième différence fixée en sortie) dont la probabilité d'apparition est élevée. Cette contrainte a un impact similaire sur le choix des boites $S$ internes, il faut que le paramètre

$$\delta_S = \max_{\beta \in \mathbb{F}_2^l} \max_{\alpha \in \mathbb{F}_2^m - \{0\}} \#\{x \in \mathbb{F}_2^m,\, S(x+\alpha) + S(x) = \beta\}$$

soit le plus faible possible (pour $S$ une fonction vectorielle de $\mathbb{F}_2^m$ dans $\mathbb{F}_2^l$) ; nous reparlerons de ce critère dans un cadre un peu inhabituel au chapitre 7 ou encore dans le cadre plus classique des fonctions puissances en annexe B.1. La meilleure résistance est atteinte lorsque $\delta_S = 2^{m-l}$ et cela implique que les composantes booléennes de $S$ vérifient $\delta_{S_i} = 2^{m-1}$ ($i \in \{1, \ldots, l\}$).

**Définition 1.9** *La fonction $f \in \mathcal{BF}(m)$ satisfait le critère $PC(l)$ de propagation de degré $l$ ($1 \leq l \leq m$) si pour tout $a \in \mathbb{F}_2^m$ de poids $0 < w_H(a) \leq l$, la fonction dérivée de $f$ en $a$, $D_a f : x \mapsto f(x+a) \oplus f(x)$, est équilibrée.*

Cela correspond en fait à l'annulation des coefficients d'auto-corrélation de $f$ en ces même points.

Le cas $l = 1$ correspond au critère $SAC$. On dit que $f$ satisfait le critère $PC(l)$ d'ordre $k$ si toute restriction de $f$ obtenue en fixant $k$ coordonnées en entrée vérifie le critère $PC(l)$. Signalons un autre critère plus fort, le critère – *Extended Propagation Criterion – EPC*($l$)

d'ordre $k$ qui impose que la dérivée en $a \neq 0$ de poids inférieur à $l$ soit $k-$résiliente. On sait prouver par exemple que si $f$ satisfait le critère $PC(l)$ d'ordre $k < m - 2$ alors son degré est inférieur à $m - k - 1$.

Le critère $PC(m)$ correspond à la définition des fonctions parfaitement non-linéaires (PN), équivalent à la notion de fonctions courbes.

**D'autres critères**... De nombreux autres critères existent (non-existence de structure linéaire, corrélation maximale, norme $L^4$ de la transformée de Walsh, fonctions non normales, nombre de termes de l'ANF, rayon de recouvrement généralisé, ...) dont en particulier l'immunité algébrique qui a été introduit récemment suite à de nouvelles attaques, dites algébriques, réalisables sur des systèmes de chiffrements par blocs ou à flot (cf. [Ars05] pour une présentation et une étude détaillée de plusieurs de ces techniques et de l'immunité algébrique).

Le principe général est de retrouver la clé par résolution de systèmes d'équations polynomiales multivariées surdéterminés, via l'utilisation par exemple de bases de Gröbner. En effet, les chiffrés s'expriment la plupart du temps polynomialement sur un corps fini (en général $\mathbb{F}_2$) en la clé et en fonction des variables d'entrée. Ces méthodes sont d'autant plus efficaces que le degré des relations obtenues (ou de toutes relations pouvant en découler) est faible. On définit ainsi l'immunité algébrique de $f$, $AI(f)$, comme le degré minimum de $g \neq 0$ tel que $f.g = 0$ ou $(f \oplus 1).g = 0$. Au niveau contrainte vis-à-vis des autres critères, ce point est relié aux non-linéarités d'ordres supérieurs ou égaux à 1 dans [Car05, Car06b] ou encore à la normalité d'une fonction dans [Can05].

Dans le domaine des codes ou de la cryptographie, la théorie des fonctions booléennes est donc reliée à de nombreuses contraintes très différentes et certaines questions importantes sur ces objets restent toujours difficiles à atteindre (dont la question de la construction d'un grand nombre de fonctions satisfaisant au mieux à tous les critères). Nous allons voir par la suite que même l'une des plus anciennes propriétés étudiées, à savoir la non-linéarité (d'ordre 1), n'est pas encore entièrement maîtrisée et que de nombreuses pistes sont encore à creuser. Dans les chapitres suivants, nous nous intéresseront de près à la conjecture de Patterson-Wiedemann et aux conditions pour battre la borne établie. Nous proposerons en particulier une construction, à partir d'une généralisation de l'exemple de Patterson et Wiedemann, qui est prometteuse pour obtenir des fonctions booléennes (équilibrées ou non) de non-linéarité élevée. Une question qui se pose alors est la possibilité ou non de généraliser cette conjecture sur des sous-ensembles quelconques de définition (des fonctions) et de sommation (dans le calcul de la transformée de Walsh).

# Chapitre 2

# Non-linéarité des fonctions de type Patterson-Wiedemann

En 1983, dans [PW83], Patterson et Wiedemann expliquent comment construire une fonction booléenne en 15 variables de non-linéarité strictement supérieure à $2^{15-1} - 2^{(15-1)/2}$, à partir d'une réunion de cosets du sous-groupe $\mathbb{F}_{2^5}^\times$ dans $\mathbb{F}_{2^{15}}^\times$. Plus précisément, parmi les fonctions invariantes sur ces cosets sous l'action du produit semi-direct du groupe $\mathbb{F}_{2^3}^\times \times \mathbb{F}_{2^5}^\times$ par le groupe des automorphismes de Frobenius, ils obtiennent par recherche exhaustive une fonction de non-linéarité $2^{14} - 2^7 + 20$. Ce qui signifie $R(15) \leq 216 = \frac{27}{32} 2^{7+1}$, d'où pour tout $m \geq 15$,

$$R(m) \leq \frac{27}{32} \sqrt{2}\, \sqrt{2^m}, \tag{2.1}$$

grâce à l'inégalité $R(m + 2k) \leq R(m)R(2k)$.

Ce résultat permit alors de mettre en défaut la conjecture de Mykkelveit de 1980 [Myk80], qui supposait que pour tout $t$, $R(2t+1) = 2^{t+1}$. La valeur $2^{t+1}$ étant ainsi dépassée, la conjecture suivante est alors proposée dans le même article [PW83].

**Conjecture 2.1 (Patterson-Wiedemann (1983))** *Pour $m \to \infty$,*

$$R(m) \sim \sqrt{2^m}.$$

A ce jour, cette conjecture n'a été ni démontrée ni contredite et le fait de battre la borne quadratique demeure toujours un problème intéressant. Comme expliqué dans le chapitre précédent, grâce à [KMY06, KY07b], on sait depuis peu que $R(m) < \mathrm{bq}(m)$ pour tout $m \geq 9$ impair mais la construction ne permet pas encore de conforter cette conjecture.

De plus, depuis 1983 aucune nouvelle fonction permettant d'obtenir un rayon spectral inférieur à la borne (2.1) n'a été trouvée ; dans la suite, nous désignerons cette borne comme étant la borne de Patterson-Wiedemann. La structure particulière de la construction effectuée dans [PW83] a été étudiée, notamment par Langevin et Zanotti dans [LZ01], mais même si la technique se généralise à une dimension quelconque, elle ne donne pas de résultat immédiat. Néanmoins, en cherchant sur $\mathbb{F}_{2^{15}}$, ils obtiennent quatre nouvelles fonctions battant la borne quadratique en parcourant cette fois-ci les fonctions invariantes modulo le sous-groupe

d'ordre 151. Ces fonctions ont une amplitude moins bonne que la fonction de Patterson et Wiedemann : 248, 246, 234 et 232. Les fonctions dites de type Patterson-Wiedemann ont également été étudiées, avec succès, pour la recherche de fonctions équilibrées hautement non-linéaires [SZZ93, MS02, SM07], de fonctions avec de bonnes propriétés d'autocorrélation (e.g. [Mai01, SM07]) ou de fonctions résilientes hautement non-linéaires [SM00].

Comme indiqué ci-dessus, la construction de [PW83] est basée sur une fonction invariante sous l'action d'un sous-groupe $G$ d'ordre 217 de $\mathbb{F}_{2^{15}}^{\times}$ – c'est en quelque sorte une généralisation de l'idée des constructions de Dillon par réunion de sous-espaces dans le cas pair (cf. [Dil74]). Afin de chercher des fonctions d'amplitude spectrale plus faible, l'idée d'explorer la non-linéarité via des sous-groupes multiplicatifs, qui est en partie formalisée dans [LZ01, LZ05], est reprise ici et généralisée dans le but d'en assouplir les contraintes associées. Plus précisément, nous considérons des fonctions constantes sur les classes non-triviales sous l'action d'un sous-groupe. Nous montrons alors comment l'utilisation des sommes de Gauss permet de se restreindre à un sous-problème, l'étude des fonctions définies sur le sous-groupe, quand l'estimation des sommes de Gauss est possible.

Le but principal des chapitres suivants est alors d'explorer des pistes pouvant mener à une confirmation de la conjecture de Patterson-Wiedemann via les constructions introduites ici.

## 2.1 Fonctions de type Patterson-Wiedemann

Cette première section poursuit l'étude effectuée dans [LZ01] afin de définir une famille de fonctions généralisant l'idée de Patterson-Wiedemann et d'analyser les critères de définitions pour améliorer la borne de Patterson-Wiedemann.

Soit $m \in \mathbb{N}^*$ (en général $m$ sera impair) et $L = \mathbb{F}_{2^m}$ le corps à $q = 2^m$ éléments. Soit $G$ un sous-groupe multiplicatif de $L^\times$ d'ordre $N$ et d'indice $v$, $Nv = 2^m - 1$.

On définit $\delta_G : L \to \{0, 1\}$ l'indicatrice de $G$ sur $L$ par $\delta_G(x) = 1$ si $x \in G$, 0 sinon. Soit $\Omega$ le groupe quotient de $L^\times$, $\Omega = L^\times/G$, et $\pi_G : L^\times \to \Omega$ le morphisme surjectif associé. Par la suite on identifiera les éléments de $\Omega$ à des éléments de $L^\times$. $\Omega$ est en fait un sous-groupe cyclique d'ordre $v$ dont le dual s'identifie (cf. Remarque 2.1 ou Chapitre 1) au groupe des caractères multiplicatifs de $L^\times$ orthogonaux à $G$.

**Remarque 2.1** *$G^\perp = \{\chi \in \widehat{L^\times}, \chi \perp G\}$ est isomorphe à $\hat{\Omega} = \widehat{L^\times/G}$ : soit $\chi \in G^\perp$, et $y \in L^\times$, alors il existe $x \in G$ et $\omega \in \Omega$ tels que $y = x\omega$, et comme $\chi_{|G} = 1$, on a $\chi(y) = \chi(x)\chi(\omega) = \chi(\omega)$.*

Une fonction de type Patterson-Wiedemann est une fonction qui est constante sur chaque coset de $G$ (ou classe modulo $G$).

**Définition 2.1** *Soit $f_0 \in \{0, 1\}$, $s : \Omega \to \{\pm 1\}$ un choix de valeurs sur les cosets de $G$, et $f$ la réunion de cosets selon $s$ définie pour $x \in L$ par*

$$f(x) = \begin{cases} (-1)^{f_0} & \text{si } x = 0, \\ \sum_{\omega \in \Omega} s(\omega)\delta_G(\frac{x}{\omega}) & \text{sinon.} \end{cases}$$

*$f$ est une fonction binaire de type Patterson-Wiedemann suivant $G$ et de représentant $s$ sur $\Omega$.*

Dans la suite, nous écrirons *fonction de type PW*. Le groupe $L^\times$ étant cyclique, il existe un unique sous-groupe d'indice $v$ et on dira alors plus simplement que $f$ est une *fonction* $(v, s)$*–PW*. Afin d'alléger les notations, nous fixerons, sauf mention contraire, la valeur en 0 par $f(0) = 1$ (i.e. $f_0 = 0$).

Soit $f$ une telle fonction et $a \in L$, alors pour $\mu : x \mapsto (-1)^{\mathrm{Tr}_L(x)}$ caractère additif canonique de $L$, la transformée de Fourier de $f$ en $a$ est $\hat{f}(a) = \sum_{x\in L} f(x)\mu(ax)$. En particulier,

$$\hat{f}(0) = 1 + \sum_{\omega\in\Omega} s(\omega).N,$$

i.e. $\hat{f}(0) = q - 2\mathrm{w}(s).N = q - 2\mathrm{w}(f)$, où le poids $\mathrm{w}(f)$ correspond ici au nombre de $-1$ dans l'image de $f$. Ainsi pour s'approcher de la borne de Patterson-Wiedemann, il faudra contrôler cette valeur en choisissant par exemple $s$ presque équilibrée et $G$ tel que $N$ soit de l'ordre de $\sqrt{q}$.

En dehors de 0, on va introduire des sommes de Gauss dans l'expression de la transformée de Fourier en $a$ afin de mieux exploiter les propriétés associées à $f$ par définition ; en particulier, le fait que $G$ soit un groupe multiplicatif sera naturellement mieux mis en avant sous cette forme.

**Proposition 2.1** *Si $f$ est une fonction $(v, s)$–PW (i.e. de représentant $s$ sur $\Omega$), alors pour $a \in L^\times$, la transformée de Fourier de $f$ en $a$ vérifie*

$$\hat{f}(a) = 1 + \frac{1}{v}\sum_{\chi\perp G} \tau_L(\chi)\hat{s}(\overline{\chi})\overline{\chi}(a) \tag{2.2}$$

*où $\hat{s}(\chi) = \sum_{\omega\in\Omega} s(\omega)\chi(\omega)$ correspond à la transformée de Fourier multiplicative de $s$ évaluée en le caractère $\chi$.*

*Démonstration.* Soit $a \in L^\times$,

$$\begin{aligned}
\hat{f}(a) &= \sum_{x\in L} f(x)\mu(ax) \\
&= f(0) + \sum_{x\in L^\times}\sum_{\omega\in\Omega} s(\omega)\delta_G(\frac{x}{\omega})\mu(ax) \\
&= 1 + \sum_{\omega'\in\Omega}\sum_{x\in G}\sum_{\omega\in\Omega} s(\omega)\delta_G(\frac{\omega' x}{\omega})\mu(a\omega' x) \\
&= 1 + \sum_{\omega'\in\Omega}\sum_{x\in G} s(\omega')\mu(a\omega' x)
\end{aligned}$$

par définition de $\delta_G$, ce qui implique :

$$\hat{f}(a) = 1 + \sum_{\omega\in\Omega} s(\omega)\sum_{x\in G}\mu(a\omega x). \tag{2.3}$$

D'après la formule de Poisson (Proposition 1.8), on a

$$\sum_{x\in G}\mu(a\omega x) = \frac{1}{[L^\times : G]}\sum_{\chi\perp G}\tau_L(\chi)\overline{\chi}(a\omega)$$

où $\tau_L(\chi) = G_L(\chi, \mu)$ est la somme de Gauss sur $L$ pour le caractère multiplicatif $\chi \in \widehat{L^\times}$ et pour $\mu$ le caractère canonique additif dans $\hat{L}$. L'équation (2.3), pour $a$ non nul, conduit alors au résultat. □

En particulier, la valeur des coefficients de Fourier en $a \neq 0$ ne dépend que du coset contenant $a$, et donc le spectre contient au plus $v + 1$ valeurs différentes. Pour estimer l'amplitude spectrale de $f$, et les valeurs de la transformée de Fourier, le problème réside dans le choix de $s$ et la difficulté est donc désormais de calculer des sommes de sommes de Gauss. Le calcul de sommes de Gauss étant ardu en général, nous essayerons de nous placer dans des cas particuliers où leurs valeurs sont connues, ou bien d'utiliser une méthode indirecte ne nécessitant pas leur calcul.

## 2.2 Simplification des sommes de Gauss

### 2.2.1 Cas hypothétique

On sait que $\tau_L(1) = -1$ et $|\tau_L(\chi)| = \sqrt{q}$ pour tout $\chi \neq 1, \chi \in \widehat{L^\times}$. En général $\tau_L(\chi)$ peut être d'argument quelconque sur le cercle des nombres complexes de module $\sqrt{q}$, mais nous supposons ici être dans le cas "idéal" où $\tau_L(\chi) = \epsilon\sqrt{q}$ pour tout $\chi \neq 1$, $\epsilon \in \{\pm 1\}$ constant (par exemple $\epsilon = 1$).

Bien que ce cas ne se produise que dans des configurations très particulières[1], la simplification des calculs qui en découle permettra de mieux comprendre et d'affiner les critères de construction de $f$.

**Lemme 2.1** *Soient $\epsilon \in \{\pm 1\}$ et $G$ un sous-groupe de $L^\times$ d'indice $v$. Supposons que $\tau_L(\chi) = \epsilon\sqrt{q}$ pour tout $\chi \neq 1, \chi \perp G$ alors, pour $a \in L^\times$, la transformée de Fourier en $a$ d'une fonction $f$ de type $(v, s)$–PW vérifie*

$$\hat{f}(a) = 1 + \epsilon s(\omega_a)\sqrt{q} - \frac{\epsilon\sqrt{q} + 1}{v}\hat{s}(1) \tag{2.4}$$

*avec $\hat{s}(1) = \sum_{\omega\in\Omega} s(\omega) = v - 2\mathrm{w}(s)$ et $\omega_a \in \Omega$ tel que $\omega_a \in a^{-1}G$, i.e. $\omega_a = \pi_G(a^{-1})$.*

*Démonstration.* La transformée de Fourier (2.2) de $f$ en $a \in L^\times$ devient

$$\begin{aligned}\hat{f}(a) &= 1 + \frac{1}{v}\sum_{\chi\perp G}\tau_L(\chi)\hat{s}(\overline{\chi})\overline{\chi}(a) \\ &= 1 + \frac{\epsilon\sqrt{q}}{v}\sum_{\chi\perp G, \chi\neq 1}\hat{s}(\overline{\chi})\overline{\chi}(a) - \frac{1}{v}\hat{s}(1)\end{aligned}$$

[1] Si les sommes de Gauss pour un sous-groupe de $\widehat{L^\times}$ sont toutes pures, on sait d'après [BMW82] qu'elles sont alors toutes rationnelles et que 2 est semi-primitif modulo $2^m - 1$, ce qui implique entre autre $m$ pair (cf. section suivante).

où $\epsilon$ est dans $\{\pm 1\}$. Ainsi,

$$\hat{f}(a) \quad = \quad 1 + \frac{\epsilon\sqrt{q}}{v} \sum_{\chi \perp G} \hat{s}(\overline{\chi})\overline{\chi}(a) - \frac{\epsilon\sqrt{q}+1}{v}\hat{s}(1)$$

où $\hat{s}(1) = \sum_{\omega\in\Omega} s(\omega) = v - 2\mathrm{w}(s)$ dépend uniquement du poids de $s$. De plus,

$$\begin{aligned}\sum_{\chi \perp G} \hat{s}(\overline{\chi})\overline{\chi}(a) &= \sum_{\chi \perp G}\sum_{\omega\in\Omega} s(\omega)\overline{\chi}(\omega)\overline{\chi}(a) \\ &= \sum_{\omega\in\Omega} s(\omega)\sum_{\chi\perp G}\overline{\chi}(a\omega),\end{aligned}$$

où $\sum_{\chi\perp G}\overline{\chi}(a\omega) = \sum_{\chi\perp G}\chi(a\omega) = v\delta_G(a\omega)$ par la relation d'orthogonalité (Proposition 1.1) appliquée au groupe $G^\perp$. D'où le résultat. □

**Corollaire 2.1** *Sous ces conditions, si* $\hat{s}(1) = o(v)$ *alors* $|\hat{f}(a)| \sim \sqrt{q}$ *pour* $a \neq 0$.

**Exemple 2.1** *Comme v impair, on ne peut éliminer le terme parasite dans (2.4) en choisissant* $\hat{s}(1) = 0$*, mais on peut choisir s presque équilibrée, par exemple* $\hat{s}(1) = 1$ *(*$\mathrm{w}(s) = (v-1)/2$*). Dans ce cas, le corollaire s'applique et* $\hat{f}(0) = q - N(v-1) = 1 + N$*. Si* $N \leq \sqrt{q} + o(\sqrt{q})$*, on obtient l'équivalence* $\|\hat{f}\|_\infty \sim \sqrt{q}$.

Cependant, l'hypothèse $\tau_L(\chi) = \epsilon\sqrt{q}$ avec $\epsilon = \pm 1$ pour tout $\chi \in G^\perp, \chi \neq 1$, est fausse en général : on a $\tau_L(\chi) = \sqrt{q}e^{2i\pi\theta_\chi}$ où $\theta_\chi$ peut prendre une valeur quelconque dans $[0, 1[$.

### 2.2.2 Cas idéal : *m* pair

Si $m$ est pair ($m = 2t$ avec $t \geq 1$), il est possible de choisir le groupe $G$ tel que l'hypothèse précédente soit vérifiée. En effet, d'après un théorème de Stickelberger (cf. Théorème 1.4), on sait que les sommes de Gauss sont toutes rationnelles et plus particulièrement que $\tau_L(\chi) = \sqrt{q}$ pour tout $\chi \in \widehat{F^\times_{2^{2t}}} - \{1\}$, si $\chi$ est d'ordre divisant $2^t + 1$.

Soit $G$ le sous-groupe de $L^\times = \mathbb{F}^\times_{2^{2t}}$ d'ordre $N = 2^t - 1$, isomorphe à $\mathbb{F}^\times_{2^t}$. Dans ce cas, $\Omega$ est isomorphe à $\mathbb{F}^\times_{2^{2t}}/\mathbb{F}^\times_{2^t}$, donc l'ordre de $G^\perp$ est égal à $v = (2^{2t}-1)/(2^t-1) = 2^t + 1$. D'où $\tau_L(\chi) = \sqrt{q}$ pour tout $\chi\perp G, \chi \neq 1$. On peut alors obtenir des non-linéarités maximales :

**Corollaire 2.2** *Si* $m = 2t$*, il existe*

- $f$ *une fonction de type PW suivant le sous-groupe* $\mathbb{F}^\times_{2^t}$ *qui soit une fonction courbe, i.e.* $\|\hat{f}\|_\infty = 2^t$,
- $f$ *de type PW suivant* $\mathbb{F}^\times_{2^t}$ *telle que* $\|\hat{f}\|_\infty = 2^t + 2$.

A noter qu'on retrouve en fait ici un résultat de Dillon [Dil74], sur le choix de $2^{t-1}$ cosets de $G$ afin d'obtenir une fonction courbe (fonction de la classe $\mathcal{PS}$).

*Démonstration.* $v$ étant impair, on choisit $s$ tel que $\hat{s}(1) = \xi = \pm 1$, dans ce cas, on a $\hat{f}(0) = q - 2f_0 - N\mathrm{w}(s) = q - 2f_0 - Nv + N\xi = 1 - 2f_0 + \xi(2^t - 1) = \xi 2^t + (\mu(f_0) - \xi)$ et l'équation (2.4) s'écrit pour $a \in L^\times$

$$\begin{aligned}\hat{f}(a) &= \mu(f_0) + s(\omega_a)\sqrt{q} - \frac{2^t+1}{v}\xi \\ &= s(\omega_a)\sqrt{q} + (\mu(f_0) - \xi).\end{aligned}$$

Si on choisit, $f_0 = (1-\xi)/2$, on trouve alors $|\hat{f}(a)| = 2^t = \sqrt{q}$ pour tout $a \in L$, i.e. que la fonction $f$ est une fonction courbe. Si maintenant, on fixe $f_0 = (1+\xi)/2$, alors $\|\hat{f}\|_\infty = 2^t + 2$.□

## 2.3 Passage à la limite pour $m$ impair

Supposons $m = 2t+1$, $t \geq 0$. Dans le cas impair, l'hypothèse précédente est irréaliste, nous pouvons seulement essayer de nous approcher du cadre intéressant où tous les arguments $\theta_\chi$ pour $\chi \neq 1$ sont proches de la même valeur. En fait, nous allons voir que si $m$ est très grand, on peut effectivement obtenir que toutes les sommes de Gauss non triviales sont proches de $\sqrt{2^m}$. Nous verrons une deuxième application de cette idée dans la section 5.1 pour des petites dimensions.

Ici, nous allons utiliser le théorème de densité de Kronecker (cf. [HW79]) appliqué à $(\theta_\chi)_\chi$ et un résultat d'indépendance des sommes de Gauss démontré dans [Rod93] afin d'obtenir un cadre favorable en passant à la limite (de manière similaire à ce qui est fait dans [Rod93]). Rappelons tout d'abord l'énoncé du théorème sur le cercle des complexes de module 1.

**Théorème 2.1 (Kronecker)** *Soient $z_1, \ldots, z_l$ dans $\mathbb{U}$, l'ensemble des nombres complexes de module 1, tels que : pour tout l-uplet d'entiers $(u_1, \ldots, u_l)$ vérifiant $\prod_{1\leq i\leq l} z_i^{u_i} = 1$ alors pour tout $i \in \{1, \ldots, l\}$, $u_i = 0$.*

*Alors les points de coordonnées $(z_1^n, \ldots, z_l^n)$ sont partout denses sur le tore $\mathbb{U}^l$.*

Ensuite, le résultat d'indépendance des sommes de Gauss (modulo l'action du Frobenius et de la conjugaison) démontré dans [Rod93] est le suivant.

### 2.3.1 Indépendance des sommes de Gausss

Soient $k$ un nombre premier impair, $\mathrm{ord}_k(2)$ l'ordre de 2 modulo $k$ (on supposera $\mathrm{ord}_k(2)$ impair), alors il existe $l$ et $c$ des entiers tels que

$$k - 1 = 2\mathrm{ord}_k(2)l \text{ et } ck = 2^{\mathrm{ord}_k(2)} - 1.$$

Soient $b \in (\mathbb{Z}/k\mathbb{Z})^\times$ un élément primitif et $\xi$ racine primitive $2l$-ième de l'unité sur $\mathbb{Q}$. On définit alors $a(X) = \sum_{1\leq v\leq l} a_v X^v$ avec

$$a_v = 2w_2(c.j_v) - \mathrm{ord}_k(2),$$

où $j_\nu$ est l'unique entier de $\{1,\dots,k-1\}$ congru à $b^\nu$ modulo $k$. Soit $\lambda$ un caractère multiplicatif de $\mathbb{F}^\times_{2^{\mathrm{ord}_k(2)}}$ d'ordre $k$, alors d'après un théorème de Stickelberger (Théorème 1.5, voir aussi [Lan90a, Lan99]) $a_\nu$ dépend de la valuation dyadique de sommes de Gauss sur $\mathbb{F}_{2^{\mathrm{ord}_k(2)}}$ :

$$a_\nu = 2\mathrm{v}_2(\tau_{\mathbb{F}_{2^{\mathrm{ord}_k(2)}}}(\lambda^{b^\nu})) - \mathrm{ord}_k(2).$$

Suivant ces hypothèses et définitions, on obtient le résultat suivant.

**Théorème 2.2 (Rodier)** *Si $a(\xi^{2r-1}) \neq 0$ pour tout $r \in \{1,\dots,l\}$, alors pour $(u_\nu)_{1\leq\nu\leq l}$ entiers*

$$\prod_{1\leq\nu\leq l}\left(\frac{\tau_{\mathbb{F}_{2^{\mathrm{ord}_k(2)}}}(\lambda^{b^\nu})}{2^{\mathrm{ord}_k(2)/2}}\right)^{u_\nu} = 1 \text{ si et seulement si } \forall\nu\in\{1,\dots,l\}, u_\nu = 0. \tag{2.5}$$

En d'autres termes, les sommes de Gauss normalisées $\frac{\tau(\lambda)}{\sqrt{2^{\mathrm{ord}_k(2)}}}$ pour $\lambda \in \widehat{\mathbb{F}^\times_{2^{\mathrm{ord}_k(2)}}}$ d'ordre $k$ sont indépendantes modulo les classes cyclotomiques et la conjugaison complexe [Rod93, lemmes 4.2 et 4.3]. A noter que le rôle de $a$ dans la preuve est de représenter le lien entre le produit considéré et la valuation dyadique de ce dernier. Suivant [Rod93], la condition sur $a$ peut être allégée par rapport à ce théorème d'indépendance.

**Corollaire 2.3** *Si une des deux conditions suivantes est vérifiée*
- *$l$ est une puissance de 2*
- *ou $l$ est un nombre premier impair et l'assertion $a_1 = -a_2 = a_3 = -a_4 = \cdots = a_l$ est fausse,*

*alors $a(\xi^{2r-1}) \neq 0$ pour tout $r \in \{1,\dots,l\}$ et l'équivalence (2.5) est vérifiée.*

Par exemple, on peut choisir parmi les nombres de Mersenne premier $k = 2^m-1$ (alors $\mathrm{ord}_k(2) = m$) tel que $l$ est premier ; pour $l = 3$ (et $m = 5$), la deuxième condition ci-dessus est satisfaite.

### 2.3.2 Construction d'une famille de fonctions de type PW selon une suite de sous-groupes projectifs

Soit $m$ tel que $2^m - 1$ soit un nombre de Mersenne premier permettant de satisfaire le théorème précédent (e.g. $m = 5$) et $L_r = \mathbb{F}_{2^{mr}}$ une extension de $K = \mathbb{F}_{2^m}$ de degré $r$ impair. Soit $G_r = L_r^\times/K^\times$ le sous-groupe projectif de $L_r^\times$, soit $f_r$ une fonction de type PW invariante sous l'action de $G_r$ et soit $s_r$ son représentant sur $\Omega_r = L_r^\times/G_r \sim K^\times$ (on choisira $s_r$ presque équilibrée, ici $|\hat{s_r}(1)| = 1$).

Alors, $G_r$ est d'ordre $N_r = (2^{mr}-1)/(2^m-1)$, d'indice $v_r = v = 2^m-1$ et ${G_r}^\perp$ est isomorphe à $\widehat{K^\times}$. Plus précisément, on a

**Proposition 2.2** *Soit $\mathrm{N}_{L_r/K}$ la norme de $L_r$ sur $K$ définie par $\mathrm{N}_{L_r/K}(x) = x.x^{2^m}.x^{2^{2m}}\cdots x^{2^{(r-1)m}}$. L'application du dual de $K^\times$ dans l'orthogonal de $G_r$, qui à un caractère $\chi$ fait correspondre le caractère $\chi^{1/r}\circ\mathrm{N}_{L_r/K}$, est un isomorphisme.*

*Démonstration.* En effet, soit $\psi \in G_r^\perp$, $x \in L_r^\times$. Par définition de $G_r$, il existe $y \in K^\times$ et $g \in G_r$ tels que $x = yg$. Alors par définition de $N_{L_r/K}$ et comme $\psi$ est un morphisme, on a $\psi^{1/r}\circ\mathrm{N}_{L_r/K}(x) = \psi(x.x^{2^m}.x^{2^{2m}}\cdots x^{2^{(r-1)m}})^{1/r} = \psi(y.y^{2^m}\cdots y^{2^{(r-1)m}})^{1/r}\psi(g.g^{2^m}\cdots g^{2^{(r-1)m}})^{1/r}$.

Or $\psi \perp G_r$, d'où $\psi(g.g^{2^m} \cdots g^{2^{(r-1)m}}) = 1$ et

$$\psi^{1/r} \circ \mathrm{N}_{L_r/K}(x) = \psi(y.y^{2^m} \cdots y^{2^{(r-1)m}})^{1/r}.$$

De plus $y \in K^\times = \mathbb{F}_{2^m}^\times$, on en déduit, pour $x \in L_r^\times$

$$\psi^{1/r} \circ \mathrm{N}_{L_r/K}(x) = \psi(y^r)^{1/r} = \psi(y)$$

où $y$ est le projeté de $x$ sur $K^\times$. □

D'autre part, d'après le théorème de Davenport-Hasse (cf. Théorème 1.3), pour $\chi \in \widehat{K^\times}$, comme $r$ est impair, on a

$$\tau_{L_r}(\chi \circ \mathrm{N}_{L_r/K}) = \tau_K(\chi)^r.$$

Ainsi, si $a \in L_r^\times$, on décompose $a$ en $xy$ où $x \in G_r$ et $y \in K^\times$, et l'équation (2.2) s'écrit alors en fonction de

$$\begin{aligned}
\sum_{\chi \perp G_r} \tau_{L_r}(\chi)\hat{s}(\overline{\chi})\overline{\chi}(a) &= \sum_{\chi \perp G_r} \tau_{L_r}(\chi)\hat{s}(\overline{\chi})\overline{\chi}(y) \\
&= \sum_{\chi \in \widehat{K^\times}} \tau_{L_r}(\chi^{1/r} \circ \mathrm{N}_{L_r/K})\hat{s}(\overline{\chi^{1/r} \circ \mathrm{N}_{L_r/K}})\overline{\chi^{1/r} \circ \mathrm{N}_{L_r/K}}(y) \\
&= \sum_{\chi \in \widehat{K^\times}} \tau_{L_r}(\chi^{1/r} \circ \mathrm{N}_{L_r/K})\hat{s}(\overline{\chi})\overline{\chi}(y)
\end{aligned}$$

puisque $y \in K^\times$, $\hat{s}(\psi) = \sum_{\omega \in K^\times} s(\omega)\psi(\omega)$ et que $\alpha \in K^\times$ implique $N_{L_r/K}(\alpha) = \alpha^r$. De plus $r$ étant premier avec l'ordre de $\widehat{K^\times}$ (en effet ce dernier est d'ordre $2^r - 1$, qui vaut 1 modulo $r$ via le petit théorème de Fermat) on obtient par changement de variable sous la somme,

$$\begin{aligned}
\sum_{\chi \perp G_r} \tau_{L_r}(\chi)\hat{s}(\overline{\chi})\overline{\chi}(a) &= \sum_{\chi \in \widehat{K^\times}} \tau_{L_r}(\chi \circ \mathrm{N}_{L_r/K})\hat{s}(\overline{\chi^r})\overline{\chi^r}(y) \\
&= \sum_{\chi \in \widehat{K^\times}} \tau_K(\chi)^r \hat{s}(\overline{\chi^r})\overline{\chi^r}(y).
\end{aligned}$$

On peut remplacer $y$ par $a$ dans la dernière expression en utilisant la convention $\chi(x) = 1$ si $x \in G_r$ et $\chi \in \widehat{K^\times}$. D'où, en remplaçant le représentant $s$ de $f$ par le représentant $s_r$ de $f_r$ sur $\Omega_r$,

**Lemme 2.2** *Pour $a \neq 0$, la transformée de Fourier de la fonction binaire $f_r$, de type $(v, s_r)$–PW, vérifie*

$$\hat{f}_r(a) = f_r(0) + \frac{1}{v} \sum_{\chi \in \widehat{K^\times}} \tau_K(\chi)^r \hat{s}_r(\overline{\chi^r})\overline{\chi^r}(a).$$

On souhaite maintenant appliquer le théorème 2.2. On considère $k = 2^m - 1$, alors c'est bien un nombre premier par choix de $m$, $\mathrm{ord}_k(2)$ est égal à $m$, $c = 1$ et $l = (2^{m-1} - 1)/m$. Par exemple, si on fixe $m = 5$ le théorème s'applique (cf. corollaire ci-dessus et [Rod93, Remarque 3.7]), d'où l'indépendance des sommes de Gauss. Par le théorème de Kronecker et via les classes cyclotomiques et la conjugaison, on en déduit que $((e^{2i\pi\theta_\chi})^r)_{\chi\in\widehat{K^\times}-\{1\}}$ est dense pour $r$ impair dans le cercle des nombres complexes de module 1, où $e^{2i\pi\theta_\chi} = \tau_K(\chi)/\sqrt{2^m}$. Ainsi on peut se ramener asymptotiquement au cas $\tau_K(\chi)^r$ proche de $\sqrt{2^{mr}}$ pour tout $\chi \neq 1, \chi \in \widehat{K^\times}$.

On obtient ainsi une estimation des coefficients de Fourier en dehors de zéro des fonctions de type PW définies suivant les sous-groupes $G_r$ ($r \geq 1$).

**Théorème 2.3** *Si pour tout $r > 0$, $f_r$ est une fonction binaire de type Patterson-Wiedemann invariante sous l'action de $G_r$, sous-groupe d'indice $2^m - 1$ de $L_r^\times = \mathbb{F}_{2^{mr}}^\times$, et de représentant $s_r$ sur $\Omega_r = L_r^\times/G_r$ tel que $|\hat{s_r}(1)| = 1$ alors, pour tout $\epsilon > 0$ et $r_0 \in \mathbb{N}$, il existe $r_1 \geq r_0$ tel que les coefficients de Fourier en dehors de zéro de $f_r$ vérifient*

$$\left|\max_{a\neq 0} \frac{|\hat{f_{r_1}}(a)|}{\sqrt{2^{mr_1}}} - 1\right| \quad \leq \quad (v-1)\epsilon + \frac{1}{\sqrt{2^{mr_1}}} + \frac{1}{v}\frac{\sqrt{2^{mr_1}}+1}{\sqrt{2^{mr_1}}}. \tag{2.6}$$

Il faut remarquer que l'estimation ne s'applique pas en 0. Il serait en fait nécessaire de satisfaire une contrainte supplémentaire pour contrôler le coefficient de Fourier en 0 (cf. Remarque 2.4).

*Démonstration.* Par densité, si $\epsilon > 0$ et $r_0 \in \mathbb{N}$, alors il existe $r_1 \geq r_0$ impair tel que pour tout $\chi \neq 1, \chi \in \widehat{K^\times}$, $\left|\frac{\tau_K(\chi)^{r_1}}{\sqrt{2^{mr_1}}} - 1\right| \leq \epsilon$. Or pour $r$ quelconque, on a pour $a \neq 0$,

$$\begin{aligned}
&\sum_{\chi\in\widehat{K^\times}} \tau_K(\chi)^r \hat{s_r}(\overline{\chi^r})\overline{\chi^r}(a) \\
&= \sum_{\chi\in\widehat{K^\times},\chi\neq 1} (\tau_K(\chi)^r - \sqrt{2^{mr}})\hat{s_r}(\overline{\chi^r})\overline{\chi^r}(a) - \hat{s_r}(1) \\
&\quad + \sqrt{2^{mr}} \sum_{\chi\in\widehat{K^\times},\chi\neq 1} \hat{s_r}(\overline{\chi^r})\overline{\chi^r}(a) \\
&= \sum_{\chi\in\widehat{K^\times},\chi\neq 1} (\tau_K(\chi)^r - \sqrt{2^{mr}})\hat{s_r}(\overline{\chi^r})\overline{\chi^r}(a) - \hat{s_r}(1)(\sqrt{2^{mr}}+1) \\
&\quad + \sqrt{2^{mr}} \sum_{\chi\in\widehat{K^\times}} \hat{s_r}(\overline{\chi^r})\overline{\chi^r}(a).
\end{aligned}$$

Comme $r$ est premier avec l'ordre de $\widehat{K^\times}$, $\sum_{\chi\in\widehat{K^\times}} \hat{s_r}(\overline{\chi^r})\overline{\chi^r}(a) = \sum_{\chi\in\widehat{K^\times}} \hat{s_r}(\overline{\chi})\overline{\chi}(a)$, d'où, de même que pour l'équation (2.4),

$$\sum_{\chi\in\widehat{K^\times}} \hat{s_r}(\overline{\chi^r})\overline{\chi^r}(a) = v \times s(\omega_a)$$

où ici $\omega_a \in K^\times$ est défini tel que $\omega_a \in a^{-1}G_r$.

Pour $r \geq 1$, on définit $l_{f_r}$ sur $L_r^\times$ par $l_{f_r}(a) = f_r(0) + s(\omega_a)\sqrt{2^{mr}} - \hat{s_r}(1)(\sqrt{2^{mr}}+1)/v$ pour $a \in L_r^\times$. On peut maintenant montrer que $\hat{f_{r_1}}(a)$ est proche de $l_{f_{r_1}}(a)$. D'après ce qui précède,

pour $a$ non nul,

$$\begin{aligned}
|\hat{f}_{r_1}(a) - l_{f_{r_1}}(a)| &\leq \frac{1}{v}\left|\sum_{\chi\in\widehat{K^\times},\chi\neq 1}(\tau_K(\chi)^{r_1} - \sqrt{2^{mr_1}})\hat{s_{r_1}}(\overline{\chi^{r_1}})\overline{\chi^{r_1}}(a)\right| \\
&\leq \frac{1}{v}\sum_{\chi\in\widehat{K^\times},\chi\neq 1}|\tau_K(\chi)^{r_1} - \sqrt{2^{mr_1}}||\hat{s_{r_1}}(\overline{\chi^{r_1}})\overline{\chi^{r_1}}(a)| \\
&\leq \frac{\epsilon\sqrt{2^{mr_1}}}{v}\sum_{\chi\in\widehat{K^\times},\chi\neq 1}|\hat{s_{r_1}}(\overline{\chi^{r_1}})| \\
&\leq \epsilon\sqrt{2^{mr_1}}(v-1).
\end{aligned}$$

Ainsi, on peut comparer le coefficient de Fourier de $f_{r_1}$ en $a$ par rapport à la valeur visée,

$$\begin{aligned}
\left|\frac{\hat{f}_{r_1}(a)}{\sqrt{2^{mr_1}}} - s(\omega_a)\right| &\leq \frac{1}{\sqrt{2^{mr_1}}}\left|\hat{f}_{r_1}(a) - l_{f_{r_1}}(a) + f_{r_1}(0) - \hat{s_{r_1}}(1)\frac{\sqrt{2^{mr_1}}+1}{v}\right| \\
&\leq (v-1)\epsilon + \frac{1}{\sqrt{2^{mr_1}}}\left(1 + |\hat{s_{r_1}}(1)|\frac{\sqrt{2^{mr_1}}+1}{v}\right),
\end{aligned}$$

et si on suppose que $\hat{s_{r_1}}(1) = \pm 1$, on obtient

$$\left|\frac{\hat{f}_{r_1}(a)}{\sqrt{2^{mr_1}}} - s(\omega_a)\right| \leq (v-1)\epsilon + \frac{1}{\sqrt{2^{mr_1}}} + \frac{1}{v}\frac{\sqrt{2^{mr_1}}+1}{\sqrt{2^{mr_1}}}.$$

□

Donc $\max_{a\neq 0}\frac{|\hat{f_{r_1}}(a)|}{\sqrt{2^{mr_1}}}$ peut être très proche de 1, il suffit pour cela d'augmenter la valeur de $v$ (et donc de $m$) fixée au départ.

**Remarque 2.2** *En pratique, cela signifie que pour r grand, on sait obtenir* $\max_{a\neq 0}|\hat{f}_r(a)| \simeq (1+\frac{1}{v})\sqrt{2^{mr}}$*, donc il faut* $v \geq 7$ *pour espérer battre la borne de Patterson-Wiedemann (cf. Proposition 1.14, il faut* $1 + 1/v \leq 27\sqrt{2}/32$*).*

Cependant, malgré ce comportement favorable des coefficients de Fourier en dehors de 0, ce résultat n'est pas suffisant puisque dans ce cas la valeur en 0 diverge. En effet, on sait que $\hat{f}_r(0) = 2^{mr} - (1 - f_r(0)) - 2N_r\mathrm{w}(s_r)$, avec $\mathrm{w}(s_r) = (v-1)/2$ pour $\hat{s_r}(1) = 1$, d'où

$$\hat{f}_r(0) = 2^{mr} - 1 + f_r(0) - N_r(v-1) = f_r(0) + N_r \simeq N_r = \frac{2^{mr}-1}{2^m-1}.$$

Ainsi, pour $r$ grand, en dépit des coefficients de Fourier de $f_r$ pouvant être proches de $\sqrt{2^{mr}}$ en $a \neq 0$, l'amplitude spectrale de $f_r$ sera grande, de l'ordre de $2^{r(m-1)}$. L'effet est en fait inverse à celui désiré, comme $v$ est constant et $N_r \gg \sqrt{2^{mr}}$, l'amplitude de $f_r$ s'éloigne d'autant plus de $\sqrt{2^{mr}}$ que $r$ augmente.

**Remarque 2.3** *Dans la preuve précédente, l'application du théorème de Kronecker et l'hypothèse d'indépendance des sommes de Gauss normalisées ne sont pas réellement nécessaires puisque nous les faisons toutes converger vers la même valeur* 1 *sur le cercle unité. En effet, un théorème d'approximation diophantienne dû à Dirichlet nous assure que pour tout* $\eta > 0$*, il existe* $r \in \mathbb{N}^*, r \leq \eta^{-(v-1)}$ *et* $p \in \mathbb{Z}^{v-1}$ *tels que* $|r\theta_\chi - p_\chi| \leq \eta$ *pour tout* $\chi \in \widehat{K^\times} - \{1\}$*. Dans ce cas, les contraintes sont moins fortes et cela conduit à remplacer* $\epsilon$ *par la valeur* $\sqrt{2(1-\cos\eta 2\pi)}$ *dans l'expression (2.6). Le théorème de Kronecker permettant de choisir un autre vecteur limite, cela nous amène à la question suivante.*

**Problème 2.1** *Pour r entier non nul et s une fonction binaire sur* $K^\times$*, existe t-il un vecteur* $(\alpha_\chi)_{\chi\in\widehat{K^\times}}$ *différent du vecteur tout à* 1 *et formé de nombres complexes de modules 1, tel que pour tout* $a \neq 0$*,*

$$\left|\sum_{\chi\in\widehat{K^\times}} \alpha_\chi \hat{s}(\overline{\chi^r})\overline{\chi^r}(a)\right| < v \quad ?$$

Ce problème peut être vu comme un cas particulier du problème de minimisation d'une somme de polynômes trigonométriques. Dans [Kah85], Kahane s'intéresse à ce problème dans le cas aléatoire (cf. [Kah85, Théorème 2]) mais le résultat ne semble pas nous aider ici.

### 2.3.3 Cas général

Suivant le principe du théorème 2.3, on a finalement le résultat général suivant.

**Théorème 2.4** *Soient* $L = \mathbb{F}_q$ *un corps de caractéristique* 2 *et G un sous-groupe d'indice v de* $L^\times$ *tels qu'il existe* $\epsilon \geq 0$ *tel que pour tout* $\chi \in G^\perp - \{1\}$*,* $|\frac{\tau_L(\chi)}{\sqrt{q}} - 1| \leq \epsilon$*. Soit f une fonction de type Patterson-Wiedemann invariante suivant G et de représentant s sur* $\Omega = L^\times/G$ *tel que* $|\hat{s}(1)| = 1$*, alors ses coefficients de Fourier vérifient*

$$\left|\max_{a\neq 0} \frac{|\hat{f}(a)|}{\sqrt{q}} - 1\right| \quad \leq \quad (v-1)\epsilon + \frac{1}{\sqrt{q}} + \frac{1}{v}\frac{\sqrt{q}+1}{\sqrt{q}}$$

*et*

$$|\hat{f}(0)| \leq 1 + \frac{q-1}{v}.$$

**Remarque 2.4** *Pour résumé, une condition suffisante, pour que la construction de type PW soit intéressante par rapport à la borne de Patterson-Wiedemann, est d'avoir*

- $\tau(\chi)$ *proche de* $\sqrt{q}$ *pour tout* $\chi \neq 1, \chi \in G^\perp$*,*
- $\hat{s}(1) = \pm 1$ *et v grand,*
- *le cardinal de G au plus de l'ordre de* $\sqrt{q}$ *(cela permet de contrôler* $\hat{f}(0)$*),*
- *et dans le cas d'un passage à la limite :* $\frac{1}{v} << 1$ *et* $v - 1$ *borné.*

Il faut donc, pour contrôler la valeur en 0, que l'ordre du sous-groupe soit au plus de l'ordre de $\sqrt{q}$. Nous savons que cette contrainte n'est pas réalisée dans le cas du théorème 2.3 mais c'est également une difficulté pour le théorème 2.4. En effet sous cette hypothèse, l'ordre des

caractères orthogonaux à $G$ est très grand et il est alors très difficile de trouver des exemples numériques satisfaisant la condition "les sommes de Gauss sont proches de $\sqrt{q}$" compte tenu de la complexité de leurs calculs (contrairement au cas d'indice petit, cf. Chapitre 5).

Pour lever cette contrainte, nous allons généraliser, dans la section suivante, la construction des fonctions de type PW afin de disposer d'un choix de cosets équilibré (permettant d'avoir $\hat{s}(1) = 0$). Ceci permettra d'éliminer un des termes parasites, et de mieux contrôler le poids de la fonction $f$ construite (et la valeur de $\hat{f}(0)$).

## 2.4 Généralisation à une construction pseudo-équilibrée

Nous reprenons ici les idées introduites dans [Bri04]. Par la suite, $G$ continue à désigner un sous-groupe de $L^\times = \mathbb{F}_q^\times$ ($G \lhd L^\times$), $N$ est le cardinal du sous-groupe et $v$ son indice.

### 2.4.1 Équilibrage sur le sous-groupe

Nous modifions la construction d'une fonction de type PW en ajoutant plus de liberté sur $G$ afin d'équilibrer la fonction et d'en contrôler le poids :

**Définition 2.2** *Soit $f_0 \in \mathbb{F}_2$, $s : \Omega^* \to \{\pm 1\}$ un choix de valeurs sur les cosets de $G$ différents de $G$ ($\Omega^* = \Omega - \{1\}$), $h : G \to \{\pm 1\}$ un choix de valeurs sur $G$ et $f$ la réunion de ces valeurs selon $s$ et $h$ définie pour $x \in L$ par*

$$f(x) = \begin{cases} (-1)^{f_0} & \text{si } x = 0 \\ h(x) & \text{si } x \in G \\ \sum_{\omega \in \Omega^*} s(\omega)\delta_G(\frac{x}{\omega}) & \text{sinon} \end{cases}$$

*Nous dirons que $f$ est une fonction de type Patterson-Wiedemann (PW) généralisé, suivant $G$ et de représentant $(s, h)$, i.e. $s$ sur $\Omega^*$ et $h$ sur $G$.*

Comme dans le cas précédent, une fonction ainsi définie peut être vue comme une configuration d'indice $v$ suivant la séquence $s$ et la section $h$. On notera pour simplifier que $f$ est une fonction $(v, s, h)$*–PW généralisée*. De plus, on fixe encore une fois $f_0 = 0$.

Afin d'obtenir une construction quasi-équilibrée (ou du moins un coefficient de Fourier en 0 de taille raisonnable comparé à la construction de la partie précédente), on choisit $s$ parfaitement équilibrée (ce qui est maintenant possible puisque $v - 1 = \text{Card}\,\Omega^*$ est impair) : Fixons une fois pour toute la condition $\mathrm{w}(s) = (v-1)/2$.

On en déduit tout d'abord la valeur du spectre de Fourier en 0 de $f$ :

$$\begin{aligned} \hat{f}(0) &= 2^m - 2\mathrm{w}(f) = 2^m - 2N\mathrm{w}(s) - 2\mathrm{w}(h) \\ &= Nv + 1 - N(v-1) - 2\mathrm{w}(h). \end{aligned}$$

**Lemme 2.3** *La transformée de Fourier de $f$ en 0 vérifie*

$$\hat{f}(0) = 1 + (N - 2\mathrm{w}(h)). \tag{2.7}$$

En fonction du choix de $h$, $\hat{f}(0)$ sera plus ou moins grand, on privilégiera donc une fonction de poids proche de $N/2$. En particulier si $h$ est de poids égal à $(N+1)/2$ (ou $(N-1)/2$ si $f_0 = 1$) alors $f$ est une fonction équilibrée.

Pour $a \in L^\times$, le calcul de la transformée est similaire à celui effectuée dans la section 2.1 mais avec l'introduction d'une somme supplémentaire suivant les éléments de $G$ et faisant intervenir $h$.

**Proposition 2.3** *Pour $a \in L^\times$ et $f$ une fonction de type PW généralisé définie comme ci-dessus alors*

$$\hat{f}(a) = 1 + \frac{1}{v}\sum_{\chi \perp G} \underbrace{\tau_L(\chi)\hat{s}(\overline{\chi})\overline{\chi}(a)}_{0 \text{ si } \chi=1} + \sum_{x\in G} h(x)\mu(ax)$$

*où $\hat{s}$ est la transformée de Fourier multiplicative de $s$ sur $\Omega^*$.*

*Démonstration.* La démonstration est analogue au cas des fonctions de type PW simple. Soit $a \neq 0$,

$$\begin{aligned}\hat{f}(a) &= \sum_{x\in L} f(x)\mu(ax)\\ &= 1 + \sum_{x\in L^\times - G} f(x)\mu(ax) + \sum_{x\in G} h(x)\mu(ax)\end{aligned}$$

où le terme central peut s'écrire via des sommes de Gauss comme précédemment.

$$\begin{aligned}\sum_{x\in L^\times - G} f(x)\mu(ax) &= \sum_{x\in L^\times - G}\sum_{\omega\in\Omega^*} s(\omega)\delta_G(\frac{x}{\omega})\mu(ax)\\ &= \sum_{y\in G}\sum_{\omega,\omega'\in\Omega^*} s(\omega)\delta_G(\frac{\omega' y}{\omega})\mu(a\omega' y)\\ &= \sum_{y\in G}\sum_{\omega\in\Omega^*} s(\omega)\mu(a\omega y) \text{ (car } \delta_G \text{ est l'indicatrice de } G).\end{aligned}$$

Comme pour l'équation (2.2), on utilise le fait que $\sum_{y\in G}\mu(ay) = \frac{1}{[L^\times:G]}\sum_{\chi\perp G}\tau_L(\chi)\overline{\chi}(a)$ pour obtenir

$$\begin{aligned}\sum_{x\in L^\times - G} f(x)\mu(ax) &= \frac{1}{v}\sum_{\omega\in\Omega^*} s(\omega)\sum_{\chi\perp G}\tau_L(\chi)\overline{\chi}(a\omega)\\ &= \frac{1}{v}\sum_{\chi\perp G}\tau_L(\chi)\hat{s}(\overline{\chi})\overline{\chi}(a),\end{aligned}$$

pour $\hat{s}(\overline{\chi}) = \sum_{\omega\in\Omega^*} s(\omega)\overline{\chi}(\omega)$. De plus, $s$ étant équilibrée, i.e. $\mathrm{w}(s) = (v-1)/2$, on a alors $\hat{s}(1) = \sum_{\omega\in\Omega^*} s(\omega) = (v-1) - 2\mathrm{w}(s) = 0$. □

Ainsi le spectre de Fourier de $f$ peut contenir plus de valeurs : en effet, l'introduction de la fonction $h$ fait que $\hat{f}(a)$ n'est plus déterminé par le coset contenant $a$.

### 2.4.2 Conséquences

**Remarque 2.5** *Pour comparer par rapport à la construction précédente, supposons pour simplifier que* $\tau_L(\chi) = \sqrt{q}$ *(comme dans la section 2.2.1, ce qui n'est possible que dans le cas pair) pour tout* $\chi \neq 1$. *Dans ce cas, on en déduit pour* $a \neq 0$,

$$\begin{aligned}\hat{f}(a) &= 1 + \frac{1}{v}\sum_{\chi \perp G} \tau_L(\chi)\hat{s}(\overline{\chi})\overline{\chi}(a) + \sum_{x\in G} h(x)\mu(ax) \\ &= 1 + \frac{\sqrt{q}}{v}\sum_{\chi \perp G} \hat{s}(\overline{\chi})\overline{\chi}(a) + \sum_{x\in G} h(x)\mu(ax),\end{aligned}$$

*puisque* $\hat{s}(1) = 0$. *Ce qui implique,*

$$\hat{f}(a) = 1 + \frac{\sqrt{q}}{v}\sum_{\omega\in\Omega^*} s(\omega)\sum_{\chi\perp G}\overline{\chi}(\omega a) + \sum_{x\in G} h(x)\mu(ax).$$

*La différence avec le cas précédent se situe ici sur le calcul de* $\sum_{\chi\perp G}\overline{\chi}(\omega a)$ *; si* $a \in G$ *alors cette somme sera toujours nulle puisqu'on s'est restreint à* $\omega \in \Omega^*$. *On a finalement :*

1. *Si* $a \in G$,
$$\hat{f}(a) = 1 + \sum_{x\in G} h(x)\mu(ax).$$
2. *Si* $a \in L^\times - G$,
$$\hat{f}(a) = 1 + s(\omega_a)\sqrt{q} + \sum_{x\in G} h(x)\mu(ax),$$
*où* $\omega_a = \pi_G(a^{-1})$ *(cf. Sec. 2.2.1).*

De manière analogue à la section précédente, si on choisit de définir une suite de fonctions $(f_r)$ de type PW généralisé suivant la suite de sous-groupes projectifs $(G_r)$, de représentants $s_r$ sur $\Omega_r^*$ et $h_r$ sur $G_r$ tels que $\hat{s_r}(1) = 0$, alors on obtient une estimation des coefficients de Fourier en $a$ non nul :

**Proposition 2.4** *Pour tout* $\epsilon > 0$ *et* $r_0 \in \mathbb{N}$, *il existe* $r \geq r_0$ *tel que pour* $a \in L_r^\times - G_r$,

$$\left|\frac{\hat{f}_r(a)}{\sqrt{2^{mr}}} - s_r(\omega_a)\right| \leq \epsilon\frac{(v-1)^2}{v} + \frac{1}{\sqrt{2^{mr}}} + \frac{|\sum_{x\in G_r} h_r(x)\mu(ax)|}{\sqrt{2^{mr}}}, \tag{2.8}$$

*et pour* $a \in G_r$,

$$\left|\frac{\hat{f}_r(a)}{\sqrt{2^{mr}}}\right| \leq \epsilon\frac{(v-1)^2}{v} + \frac{1}{\sqrt{2^{mr}}} + \frac{|\sum_{x\in G_r} h_r(x)\mu(ax)|}{\sqrt{2^{mr}}}.$$

*Démonstration.* C'est une simple adaptation de la preuve du théorème 2.3. □

Par rapport au théorème 2.3, une contrainte sur $v$ est levée : il n'est plus nécessaire a priori d'avoir $1/v$ petit pour que l'expression de droite dans l'inégalité (2.8) soit négligeable. De

plus, comme $\hat{f_r}(0) = 1 + (\#G_r - 2\mathrm{w}(h_r))$, la valeur de l'ordre de $G_r$ n'est plus une contrainte si $h_r$ est quasi-équilibrée sur $G_r$. Mais un nouveau problème se situe alors sur le dernier terme, i.e. sur le choix de $h_r$ pour minimiser la somme $|\sum_{x\in G_r} h_r(x)\mu(ax)|$.

Le théorème 2.4 se généralise de la même manière et nous donne les majorations suivantes.

**Théorème 2.5** *Soient $L = \mathbb{F}_q$ un corps de caractéristique 2 et $G$ un sous-groupe d'indice $v$ de $L^\times$ tels qu'il existe $\epsilon \geq 0$ tel que pour tout $\chi \in G^\perp - \{1\}$, $|\frac{\tau_L(\chi)}{\sqrt{q}} - 1| \leq \epsilon$. Soit $f$ une fonction de type Patterson-Wiedemann généralisé suivant $G$ et de représentant $s$ sur $\Omega^* = L^\times/G - \{1\}$ et $h$ sur $G$ tel que $\hat{s}(1) = 0$, alors les coefficients de Fourier de $f$ vérifient : Pour $a \in L^\times - G$*

$$\left|\frac{|\hat{f}(a)|}{\sqrt{q}} - 1\right| \leq \epsilon\frac{(v-1)^2}{v} + \frac{1}{\sqrt{q}} + \frac{|\sum_{x\in G} h(x)\mu(ax)|}{\sqrt{q}},$$

*pour $a \in G$*

$$\frac{|\hat{f}(a)|}{\sqrt{q}} \leq \epsilon\frac{(v-1)^2}{v} + \frac{1}{\sqrt{q}} + \frac{|\sum_{x\in G} h(x)\mu(ax)|}{\sqrt{q}},$$

*et*

$$|\hat{f}(0)| \leq 1 + \left|\frac{q-1}{v} - 2\mathrm{w}(h)\right|.$$

Il est donc nécessaire d'étudier comment minimiser les sommes faisant intervenir la fonction binaire $h$ qui est définie sur le sous-groupe pour appliquer ce résultat efficacement.

**Remarque 2.6** *Retenons qu'ici une condition suffisante, pour que la construction de type PW généralisé suivant un sous-groupe $G$ soit pertinente vis-à-vis de la borne de Patterson-Wiedemann, est d'avoir*

- *$\tau(\chi)$ proche de $\sqrt{q}$ pour tout $\chi \neq 1, \chi \in G^\perp$,*
- *$|\sum_{x\in G} h(x)\mu(ax)|$ petit par rapport à $\sqrt{q}$ pour $a \notin G, a \neq 0$,*
- *$|\sum_{x\in G} h(x)\mu(ax)|$ au plus de l'ordre de $\sqrt{q}$ pour $a \in G \cup \{0\}$ (en particulier en 0, cela implique que $\hat{f}(0)$ est contrôlé),*
- *et dans le cas d'un passage à la limite : $v - 1$ majoré.*

*Les conditions semblent plus réalistes que dans la remarque 2.4, en particulier le fait que l'ordre du groupe soit supérieur à $\sqrt{q}$ n'est plus un point bloquant au vu de ces contraintes.*

L'écriture suivant les sommes de Gauss introduit donc un nouveau problème : l'estimation de sommes d'exponentielles sur le sous-groupe $G$. Dans la suite, nous allons étudier ce problème de recherche dans le cadre de la transformée de Fourier d'une fonction binaire définie sur un sous-ensemble quelconque (cf. Chapitre 3). Nous approfondissons ensuite le cas particulier d'un sous-groupe multiplicatif (cf. Chapitre 4). L'utilisation de propriétés classiques des sommes de Gauss nous conduira alors à proposer une nouvelle conjecture du type de celle de Patterson et Wiedemann. Enfin, le fait de ne plus avoir de contrainte forte sur l'ordre du sous-groupe nous permettra au chapitre 5 d'étudier les comportements des sommes de Gauss pour des indices petits à travers des exemples numériques.

# Chapitre 3

# Estimation asymptotique de la non-linéarité partielle

Soit $f$ une fonction booléenne à $m$ variables, alors la non-linéarité de $f$ est inférieure à $2^{m-1} - 2^{m/2-1}$ (cf. Section 1.2.3). Cette borne est atteinte lorsque $f$ est une fonction courbe, voir [Rot76], ce qui n'est possible que pour $m$ pair. De manière générale, lorsque $f$ varie, la connaissance de la distribution de valeurs de la non-linéarité est complexe.

On s'attend à ce que la plupart des fonctions booléennes aient une complexité élevée, et ce compte tenu de l'augmentation doublement exponentielle du nombre de fonctions booléennes quand la dimension de l'espace de départ croît. En 1998, Olejár et Stanek [OS98] confirment que cette observation s'applique à la notion de non-linéarité en démontrant alors que les fonctions booléennes à $m$ variables ont presque toutes une non-linéarité supérieure à $2^{m-1} - c\sqrt{m}2^{m/2-1}$, où $c = \sqrt{2(1+\epsilon)\ln 2}$ avec $\epsilon > 0$ quelconque. Avant ce résultat et un résultat analogue établi ensuite par Carlet [Car02], la distribution de valeurs de la non-linéarité était très peu connue, si ce n'est via des observations expérimentales (dans la limite du réalisable, e.g. $n \leq 6$).

Dans [Rod03b, Rod04], Rodier affine ces résultats en exhibant une borne supérieure, $2^{m-1} - 2^{m/2-2}\sqrt{m \ln 2}$, pour la non-linéarité de presque toutes les fonctions. Récemment (cf. [Rod03a] ou encore [Rod06]), il a démontré que les fonctions booléennes $f$ de $\mathbb{F}_{2^m}$ admettent, en fait, presque toutes une non-linéarité $\mathrm{nl}(f)$ asymptotiquement proche de $2^{m-1} - 2^{m/2-1}\sqrt{2m \ln 2}$.

Par analogie, on aimerait savoir si ce résultat est vrai, pour des notions dérivées de la non-linéarité, sur des sous-ensembles quelconques. Dans ce chapitre, nous montrons comment les idées et outils introduits par Rodier peuvent être utilisés dans le but d'étudier une telle généralisation du problème de la non-linéarité à des sous-ensembles de $\mathbb{F}_{2^m}$. Nous mettons en particulier en évidence les différentes contraintes à satisfaire pour obtenir des bornes asymptotiques inférieures et supérieures valables dans ce contexte et similaires au cas du corps. Une des motivations principales de ces travaux est l'application des estimations aux constructions de type Patterson-Wiedemann généralisé étudiées au chapitre 2, où apparaît entre autre le problème des sommes d'exponentielles sur des sous-groupes de $\mathbb{F}_{2^m}^{\times}$.

## 3.1 Introduction

Étudier la non-linéarité d'une fonction booléenne revient à étudier le minimum de sa transformée de Fourier. Un fait intéressant, expliqué dans [Rod03b, Rod04], est le lien entre la non-linéarité d'une fonction booléenne et le problème de minimisation de séries de Fourier à coefficients binaires sur le tore des nombres complexes de module 1 (ou plus généralement pour des polynômes réels aléatoires). Ce problème de la distribution de valeurs de la norme infinie sur $\mathbb{R}$

$$\|x \mapsto \sum_{s=0}^{n} a_{s,n} e^{isx}\|_\infty$$

pour $n$ tendant vers l'infini, a été largement étudié (voir entre autres les articles de Salem et Zygmund [SZ54], de Kahane [Kah85] ou de Halácz [Hal73]). Comme expliqué ci-dessus, Rodier adapte, dans [Rod03a, Rod03b, Rod04, Rod06], ces techniques au contexte de la non-linéarité des fonctions booléennes.

Nous suivons ici les idées de Rodier afin de mener l'étude de la répartition des valeurs de non-linéarité partielle, ou plus exactement des amplitudes spectrales associées (voir Section 3.2.2 pour la définition de cette notion). Ces techniques sont centrées sur une analyse asymptotique (et presque sûrement) de l'amplitude d'une fonction grâce à l'application de résultats connus de théorie des probabilités.

Nous effectuons une description du problème section 3.2, où après avoir introduit les espaces mesurés étudiés, on décrit la notion de rayon spectral d'un sous-ensemble de $\mathbb{F}_{2^m}$ sur un sous-groupe de caractères. Dans la section 3.3, nous étudions les bornes supérieures et inférieures obtenues en généralisant la méthode de Rodier [Rod03b] basée sur les idées de Kahane. Dans la section 3.4, nous analysons ensuite les conditions pour appliquer l'amélioration des estimations asymptotiques proposées dans [Rod03a] via les résultats de [Hal73]. La section 3.5 permet de faire le lien avec les précédents résultats de [Rod03a]. Enfin dans la section 3.6, nous étudions différents cas d'applications des quatre théorèmes principaux 3.2, 3.3, 3.4 et 3.5 à des instances particulières (dont celles introduites au chapitre 2).

## 3.2 Description du problème

### 3.2.1 Suite d'espaces de fonctions binaires

Soit une suite $(G_j)_{j\in\mathbb{N}}$ croissante d'ensembles où $G_j$ est un sous-ensemble de $\mathbb{F}_{q_j}$ ($q_j = 2^{m_j}$) de cardinal $N_j > 0$. Nous allons maintenant considérer des suites de fonctions binaires construites sur ces ensembles afin de pouvoir étudier le passage à la limite.

Soit $j \in \mathbb{N}$. On définit $\Phi_j$ comme étant l'injection canonique de $G_j$ dans $G_{j+1}$. On peut ainsi définir $G_\infty = \lim_{j\to\infty} G_j$ par limite inductive suivant la famille $(\Phi_j)_{j\in\mathbb{N}}$.

Soit $\Omega_j = \{g : G_j \to \{\pm 1\}\}$ l'espace des fonctions binaires (ou exponentielles de fonctions booléennes) définies sur $G_j$. On définit de même $\Omega_\infty$ l'espace des fonctions binaires définies sur $G_\infty$. Pour simplifier, on le désignera également dans cette partie par $\Omega$ ($\Omega = \Omega_\infty$).

**Lemme 3.1** *$\Omega_j$ pour tout $j \in \mathbb{N}$ et $\Omega_\infty$, munis des lois $(+, \times, .)$, sont des $\mathbb{F}_2$-algèbres.*

On définit de plus une application de transition de $\Omega_{j+1}$ à $\Omega_j$, $\pi_j : \Omega_{j+1} \to \Omega_j$, telle que $\pi_j(g) = g_{|G_j}$ pour $g \in \Omega_{j+1}$. En d'autres termes, $\Omega_\infty$ est la limite projective des $\Omega_j$. On considère alors l'application de projection dans $\Omega_j$, $\Pi_j : \Omega_\infty \to \Omega_j$, qui à $g \in \Omega_\infty$ associe sa restriction à $G_j$ :

$$\begin{array}{rccc} \Pi_j : & \Omega_\infty & \to & \Omega_j \\ & g & \mapsto & g_{|G_j}. \end{array}$$

On munit finalement les algèbres $\Omega_j$ de la probabilité uniforme $\mathbb{P}$ et par extension $\Omega_\infty$ d'une probabilité, également notée $\mathbb{P}$, définie par :

$$\forall f \in \Omega_j,\ \mathbb{P}(h \in \Omega_\infty \mid \Pi_j^{-1}(f) = h) = \frac{1}{\#\Omega_j} = \frac{1}{2^{N_j}}.$$

### 3.2.2 Amplitude spectrale et non-linéarité partielles

Par analogie avec l'étude de la non-linéarité des fonctions booléennes (ou de manière équivalente binaires) sur $\mathbb{F}_{2^m}$ où la transformée de Fourier est largement utilisée, on introduit la transformée de Fourier d'une fonction binaire $g$ sur $E$, afin d'étudier maintenant le problème de la non-linéarité généralisé à un sous-ensemble $E$ de $\mathbb{F}_{2^m}$.

**Définition 3.1** *Soit $E$ un sous-ensemble strict de $\mathbb{F}_{2^m}$ et $g : E \to \{\pm 1\}$. La transformée de Fourier sur $E$ de la fonction $g$ est la fonction de l'ensemble des caractères additifs de $\mathbb{F}_{2^m}$, i.e. son dual $\widehat{\mathbb{F}_{2^m}}$, dans $\mathbb{Z}$ définie de la manière suivante*

$$\begin{array}{rccc} \widehat{\mathbb{F}_{2^m}} & \to & \mathbb{Z} \\ \mu & \mapsto & \displaystyle\sum_{x \in E} g(x)\mu(x). \end{array}$$

*La transformée de Fourier sur $E$ de $g$ est notée $\tilde{g}_E$ (ou simplement $\tilde{g}$ s'il n'y a pas d'ambiguïté possible).*

Comme toujours, la transformation de Fourier, qui à $g$ associe sa transformée de Fourier $\tilde{g}$, est inversible de sorte que pour $x \in E$, on a

$$g(x) = \frac{1}{2^m} \sum_{\mu \in \widehat{\mathbb{F}_{2^m}}} \tilde{g}(\mu)\mu(x).$$

**Remarque 3.1** *Cette définition est similaire à la définition classique d'une transformée de Fourier par rapport à un espace de caractères ; la notation $\tilde{g}$, différente de la notation générale d'une transformée de Fourier peut alors surprendre. Ce choix est fait afin de différencier clairement cette définition à celle de la transformée de Fourier d'une fonction booléenne sur le corps en entier. En effet, ici la différence avec la définition usuelle réside dans le fait que la fonction $\tilde{g}_E$ est définie sur l'ensemble des caractères de $\mathbb{F}_{2^m}$ alors que son calcul fait intervenir une somme indexée uniquement par $E$ (contrairement à la définition 1.5).*

Ici, pour un caractère additif $\mu$ de $\mathbb{F}_{q_j}$, la transformée de Fourier de $g \in \Omega_j$ sur $G_j$ est donc définie par $\tilde{g}(\mu) = \sum_{x \in G_j} g(x)\mu(x)$. De plus, par passage à l'inverse, on a pour $x \in G_j$, $g(x) = 1/q_j \sum_{\mu \in \widehat{\mathbb{F}_{q_j}}} \tilde{g}(\mu)\mu(x)$. Cette dernière expression peut être interprétée comme étant

$$g(x) = \int_{\mu \in \widehat{\mathbb{F}_{q_j}}} \tilde{g}(\mu)\mu(x)d\mu.$$

De manière générale, nous utiliserons cette description "continue" pour toutes les sommes pondérées, i.e. si $F$ est une fonction quelconque sur un ensemble $\Upsilon$ fini, on note $\int_{\varphi \in \Upsilon} F(\varphi)d\varphi$ au lieu de $\frac{1}{\text{Card } \Upsilon} \sum_{\varphi \in \Upsilon} F(\varphi)$. Cette écriture "continue" est plus naturelle dans la théorie des probabilités, son utilisation simplifie l'application des propriétés classiques du domaine. D'autre part, elle a également l'avantage de s'étendre à $\Omega_\infty$ où les fonctions ne sont plus à support fini.

Pour un sous-ensemble $E$ de $\mathbb{F}_{2^m}$ et $g$ une fonction binaire définie sur $E$, $g : E \to \{\pm 1\}$, on définit enfin l'amplitude spectrale $\|\tilde{g}\|_\infty$ de $g$ sur $E$ par

$$\|\tilde{g}\|_\infty = \max_{\mu \in \widehat{F_{2^m}}} |\tilde{g}(\mu)| = \max_{\mu \in \widehat{F_{2^m}}} |\sum_{x \in E} g(x)\mu(x)|,$$

qui correspond donc à la plus grande valeur du spectre de $g$ sur $E$. Le rayon spectral de $E$ sur $\mathbb{F}_{2^m}$, noté $R_E(m)$, est alors défini comme la plus petite amplitude spectrale possible :

$$R_E(m) = \min_{g:E \to \{\pm 1\}} \|\tilde{g}\|_\infty = \min_{g:E \to \{\pm 1\}} \max_{\mu \in \widehat{F_{2^m}}} |\sum_{x \in E} g(x)\mu(x)|.$$

Plus généralement, on peut étendre ces définitions à l'analyse spectrale par rapport à un sous-groupe quelconque de $\widehat{\mathbb{F}_{2^m}}$.

**Définition 3.2** *Soient $m \in \mathbb{N}^\times$, $E \subset \mathbb{F}_{2^m}$, $g : E \to \{\pm 1\}$ et $H$ un sous-groupe de $\mathbb{F}_{2^m}$. L'***amplitude spectrale partielle** *de $g$ sur $E$ par rapport à $\hat{H}$ est*

$$\|\tilde{g}\|_{\hat{H}} = \max_{\mu \in \hat{H}} |\tilde{g}(\mu)| = \max_{\mu \in \hat{H}} |\sum_{x \in E} g(x)\mu(x)|.$$

*Le* **rayon spectral partiel** *de $E$ par rapport à $H$ est*

$$R_E(H) = \min_{g:E \to \{\pm 1\}} \|\tilde{g}\|_{\hat{H}}.$$

Par isomorphisme de $H$ avec son dual, et s'il n'y a pas d'ambiguïté, on écrira parfois $\|\tilde{g}\|_H$.

**Remarque 3.2** *Dans la définition précédente, par convention si $E \not\subset H$, pour $\mu \in \hat{H}$ on prolonge le caractère $\mu$ à $E$ en posant $\mu(x) = 1$ si $x \in E$ et $x \notin H$, afin que la transformée de Fourier de $g$ en $\mu$ soit bien définie. En fait, on peut même voir ce prolongement comme l'unique caractère de $\mathbb{F}_{2^m}$ égal à $\mu$ sur $H$ et de valeur $1$ en dehors de $H$.*

Déterminer l'amplitude spectrale partielle de $g$ sur $E$ par rapport à $H$ est équivalent à trouver la distance minimale (sur $E$) de $g$ par rapport à l'ensemble des fonctions affines définies relativement à $H$, $\{x \mapsto \text{Tr}(ax) + b, a \in \mathbb{F}_{2^m}/H^\perp, b \in \mathbb{F}_2\}$ ; d'où le terme de **non-linéarité partielle**.

**Remarque 3.3** *En pratique, on pourrait même étudier l'amplitude spectrale par rapport à un sous-ensemble de caractères quelconques ; une partie des résultats établis dans ce chapitre peuvent se généraliser à ce cadre, mais pour simplifier nous travaillerons avec un ensemble de caractères formant un sous-groupe (ici $\hat{H}$).*

Pour rappel, le résultat démontré par Rodier dans [Rod03a], exprimé en terme d'amplitude spectrale et que nous cherchons à généraliser dans la suite, est le suivant.

**Théorème 3.1 (Rodier)** *Pour presque (selon $\mathbb{P}$) toute fonction $f \in \Omega_\infty$,*

$$\limsup_{m\to\infty} \frac{\|\hat{f}\|_\infty}{\sqrt{2^m \ln 2^m}} = \sqrt{2},$$

*où $\Omega_\infty$ est ici vu comme la limite des espaces de fonctions booléennes définies sur le corps en entier (i.e. $G_j = \mathbb{F}_{q_j}$).*

Soit $(H_j)_j$ une suite de sous-groupes telle que $H_j$ est un sous-groupe de $\mathbb{F}_{q_j}$ d'ordre $h_j \geq 1$ pour tout $j$. Nous voulons principalement étudier la distribution des valeurs $\|\tilde{g}\|_\infty$ pour $g \in \Omega_j$ et le rayon spectral $R_{G_j}(m_j)$, mais nous étudierons en même temps les valeurs $\|\tilde{g}\|_{\hat{H}_j}$ et $R_{G_j}(H_j)$ telles qu'elles sont définies dans la définition précédente. Il est en effet intéressant de savoir à quel point le résultat énoncé dans le théorème 3.1 peut se généraliser.

## 3.3 Techniques de Kahane

Tout d'abord, nous montrons dans cette section que les deux bornes asymptotiques établies dans [Rod03b] se généralisent facilement en modifiant légèrement les preuves, à la condition que les suites $(G_j)$ et $(\hat{H}_j)$ satisfassent quelques critères spécifiques.

### 3.3.1 Borne supérieure

**Lemme 3.2** *Soit $j \in \mathbb{N}$, $\mu \in \hat{H}_j$ et $t \in \mathbb{R}$ alors l'espérance pour $g \in \Omega_j$ de $e^{t\tilde{g}(\mu)}$ vérifie*

$$e^{\frac{t^2 N_j}{2} - t^4 N_j} \leq \varepsilon(e^{t\tilde{g}(\mu)}) \leq e^{\frac{t^2 N_j}{2}}.$$

*Démonstration.* On a

$$\varepsilon(e^{t\tilde{g}(\mu)}) \quad = \quad \varepsilon(\prod_{x\in G_j} e^{tg(x)\mu(x)}) = \prod_{x\in G_j} \varepsilon(e^{tg(x)\mu(x)})$$

puisque les variables aléatoires $g \mapsto e^{tg(x)\mu(x)}$ sont indépendantes pour $x$ parcourant $G_j$. D'autre part, comme à $x$ fixé, $g(x)\mu(x)$ prend la valeur 1 pour la moitié des éléments de $\Omega_j$ et $-1$ sur l'autre moitié, on remarque que $\varepsilon(e^{tg(x)\mu(x)}) = \int_{\Omega_j} e^{tg(x)\mu(x)} d\mathbb{P} = e^t/2 + e^{-t}/2 = \cosh t$. En utilisant les développements en série entière usuels de $e^{t^2/2}$ et $\cosh t$, on obtient aisément

$$1 + t^2/2 \leq \cosh t \leq e^{t^2/2}. \tag{3.1}$$

De plus, $1 + t^2/2 \geq e^{t^2/2-t^4}$, d'où

$$(e^{t^2/2-t^4})^{N_j} \leq \varepsilon(e^{t\tilde{g}(\mu)}) \leq (e^{t^2/2})^{N_j}.$$

□

**Théorème 3.2** *Soient $(G_j)$ une suite croissante de sous-ensembles, $G_j \subset \mathbb{F}_{q_j}$ et de cardinal $N_j$, et $(\hat{H}_j)$ suite de sous-ensembles, $\hat{H}_j \subset \widehat{\mathbb{F}_{q_j}}$ et de cardinal $h_j$. Si la série $\sum \frac{1}{h_j^\alpha}$ est convergente pour tout $\alpha$ strictement positif, alors pour $j$ suffisamment grand et $g \in \Omega_\infty$, on a presque sûrement*

$$\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} \leq \sqrt{2N_j \ln h_j} = \sqrt{2 \times Card\, G_j \times \ln Card\, \hat{H}_j}.$$

Le comportement est donc similaire à celui de la non-linéarité classique, asymptotiquement la plupart des fonctions ont une non-linéarité partielle élevée.

*Démonstration.* Soient $j \in \mathbb{N}$, $t \in \mathbb{R}$ et $g \in \Omega_j$ alors il existe $\mu_0 \in \hat{H}_j$ (dépendant de $g$) tel que $\tilde{g}(\mu_0) = \pm\|\tilde{g}\|_{\hat{H}_j}$, ainsi $e^{t\|\tilde{g}\|_{\hat{H}_j}} \leq e^{t\tilde{g}(\mu_0)} + e^{-t\tilde{g}(\mu_0)}$ et

$$e^{t\|\tilde{g}\|_{\hat{H}_j}} \leq h_j \int_{\hat{H}_j} (e^{t\tilde{g}(\mu)} + e^{-t\tilde{g}(\mu)}) d\mu.$$

L'espérance se calculant ensuite sur $g$ variant dans $\Omega_j$, les intégrales peuvent être interverties et on obtient $\varepsilon(e^{t\|\tilde{g}\|_{\hat{H}_j}}) \leq h_j \int_{\hat{H}_j} (\varepsilon(e^{t\tilde{g}(\mu)}) + \varepsilon(e^{-t\tilde{g}(\mu)})) d\mu$. Le lemme précédent implique alors

$$\varepsilon(e^{t\|\tilde{g}\|_{\hat{H}_j}}) \leq 2h_j e^{\frac{N_j t^2}{2}}.$$

Soit $\alpha > 0$ un réel, en multipliant par $h_j^{-\alpha}$, on en déduit l'inégalité suivante

$$\varepsilon(\exp(t\|\tilde{g}\|_{\hat{H}_j} - \frac{N_j t^2}{2} - (1+\alpha)\ln h_j)) \leq 2h_j^{-\alpha},$$

d'après l'inégalité de Markov (cf. par exemple [Loe77, Rev97] ou [Shi96]) cela entraîne

$$\mathbb{P}(\exp(t\|\tilde{g}\|_{\hat{H}_j} - \frac{N_j t^2}{2} - (1+\alpha)\ln h_j) \geq 1) \leq \varepsilon(\exp(t\|\tilde{g}\|_{\hat{H}_j} - \frac{N_j t^2}{2} - (1+\alpha)\ln h_j)) \leq 2h_j^{-\alpha},$$

i.e. $\mathbb{P}(t\|\tilde{g}\|_{\hat{H}_j} - \frac{N_j t^2}{2} - (1+\alpha)\ln h_j \geq 0) \leq 2h_j^{-\alpha}$. Pour $t = \sqrt{2\frac{(1+\alpha)\ln h_j}{N_j}}$, on obtient

$$\mathbb{P}\left(\|\tilde{g}\|_{\hat{H}_j} \geq \sqrt{2N_j(1+\alpha)\ln h_j}\right) \leq 2h_j^{-\alpha}. \tag{3.2}$$

Soit maintenant $g \in \Omega_\infty$. Si la série $\sum \frac{1}{h_j^\alpha}$ est convergente pour $\alpha$ strictement positif alors la somme $\sum_j \mathbb{P}\left(\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} \geq \sqrt{2N_j(1+\alpha)\ln h_j}\right)$ est finie. Par conséquent, pour $j$ suffisamment grand, on a presque sûrement

$$\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} < \sqrt{2N_j(1+\alpha)\ln h_j},$$

par l'application du lemme de Borel-Cantelli (voir par exemple [KS64] ou [Rev97]). Il faut faire tendre $\alpha$ vers 0 pour conclure. □

Cette borne supérieure est une première généralisation de la borne démontrée par Rodier dans [Rod03b]. La contrainte sur $h_j$ est par exemple satisfaite si $H_j$ est le corps $\mathbb{F}_{q_j}$ tout entier avec $q_j \to \infty$. Elle n'est par contre pas satisfaite si $(H_j)$ est stationnaire. On remarquera que le théorème reste valable si on remplace $\hat{H}_j$ par un sous-ensemble quelconque de caractères (additifs ou multiplicatifs).

Le résultat sera rejoint au chapitre 4 par le théorème 4.1 (ou du moins sa généralisation expliquée dans la remarque 4.2) lorsque la condition $h_j \to \infty$ est satisfaite (condition moins forte que celle du théorème ci-dessus en général, même si elle est proche ; si $H_j$ est un sous-groupe la condition est équivalente sachant que $h_j \mid 2^{m_j}$).

### 3.3.2 Borne inférieure

Nous faisons ici l'hypothèse supplémentaire que $G_j$ est un sous-groupe de $\mathbb{F}_{q_j}$ pour tout $j$.

**Remarque 3.4** *Pour $j$ fixé, on remarquera que $\hat{G}_j$, $\hat{H}_j$ sont des groupes cycliques puisque ce sont des sous-groupes du groupe cyclique $\widehat{\mathbb{F}_{q_j}}$. On définira d'autre part l'extension (ou restriction si $G_j \lhd H_j$) $\mu_{|G_j}$ de $\mu \in \hat{H}_j$ à $G_j$ par défaut comme sa restriction à $G_j \cap H_j$ étendue à $G_j$ tout entier par l'élément unité, i.e $\mu_{|G_j}(x) = \mu(x)$ si $x \in G_j \cap H_j$ et $\mu_{|G_j}(x) = 1$ si $x \in G_j - G_j \cap H_j$. Suivant cette définition, on obtient $\mu_{|G_j} \in \hat{G}_j$. En pratique, on pourra se ramener essentiellement à $G_j \lhd H_j$.*

**Lemme 3.3** *Soit $j \in \mathbb{N}$, $\mu, \nu \in \hat{H}_j$ et $t \in \mathbb{R}$ alors l'espérance pour $g \in \Omega_j$ de $e^{t(\tilde{g}(\mu)+\tilde{g}(\nu))}$ vérifie*

$$\varepsilon(e^{t(\tilde{g}(\mu)+\tilde{g}(\nu))}) \leq \begin{cases} e^{t^2 N_j} \text{ si } \mu \neq \nu \\ e^{2t^2 N_j} \text{ si } \mu = \nu \end{cases}$$

*Démonstration.* Par indépendance des variables aléatoires pour $x \in G_j$,

$$\varepsilon(e^{t(\tilde{g}(\mu)+\tilde{g}(\nu))}) = \prod_{x \in G_j} \varepsilon(e^{tg(x)(\mu(x)+\nu(x))}) = \prod_{x \in G_j} \cosh(t(\mu(x)+\nu(x))).$$

D'après l'équation (3.1) dans la démonstration du lemme 3.2, ce dernier terme est inférieur à $\prod_{x \in G_j} e^{(\mu(x)+\nu(x))^2 t^2/2} = \prod_{x \in G_j} e^{(1+\mu(x)\nu(x))t^2} = e^{N_j t^2} \prod_{x \in G_j} e^{\mu(x)\nu(x)t^2} = e^{N_j t^2} e^{t^2 \sum_{x \in G_j} \mu(x)\nu(x)}$. D'où les résultats suivant les relations d'orthogonalité sur $G_j$ appliquées à $\mu \times \nu$. □

**Remarque 3.5** *On voit apparaître dans cette démonstration l'intérêt de l'hypothèse "$G_j$ sous-groupe de $\mathbb{F}_{q_j}$". En effet, si ce n'était pas le cas, l'ensemble $\{x \in G_j \mapsto \mu(x) \,|\, \mu \in \hat{H}_j\}$ ne serait pas un sous-groupe des caractères de $G_j$ et on ne pourrait pas simplifier la dernière somme par relations d'orthogonalité.*

**Proposition 3.1** *Soient deux réels $\alpha > 0$, $0 < \lambda < 1$, et $j \in \mathbb{N}$, $g \in \Omega_j$, alors*

$$\mathbb{P}\left(\|\tilde{g}\|_{\hat{H}_j} \geq (\frac{\lambda}{2} - \lambda^3 \frac{\ln h_j}{N_j} - \frac{\alpha}{\lambda}) \sqrt{N_j \ln h_j}\right) \geq (1 - 2h_j^{-\alpha})(1 - 2\frac{(\ln h_j)^2}{N_j})(1 - h_j^{\lambda^2 - 1}).$$

*Démonstration.* Soit $t \in \mathbb{R}^+$ et $X_j$ la v.a. $X_j = \int_{\hat{H}_j} e^{t\tilde{g}(\mu)} d\mu$, d'après le lemme 3.2,

$$\varepsilon(X_j) \geq \int_{\hat{H}_j} e^{N_j(t^2/2 - t^4)} d\mu = e^{N_j(t^2/2 - t^4)}.$$

De plus, $\varepsilon(X_j^2) = \int_{\mu,\nu \in \hat{H}_j} \varepsilon(e^{t(\tilde{g}(\mu)+\tilde{g}(\nu))}) d\mu d\nu \leq \int_{\mu \neq \nu} e^{t^2 N_j} d\mu d\nu + \frac{1}{h_j} \int_{\mu \in \hat{H}_j} e^{2N_j t^2} d\mu = \frac{h_j - 1 + e^{N_j t^2}}{h_j} e^{N_j t^2}$ selon le lemme précédent.

D'après [Kah85] (cf. *Inequality II*, §1.6 ; cette inégalité – élémentaire – repose sur l'inégalité de Schwartz), si $X$ est une v.a. dans $L^2$ et $0 < \lambda < 1$ alors

$$\mathbb{P}(X \geq \lambda \varepsilon(X)) \geq (1 - \lambda)^2 \frac{\varepsilon^2(X)}{\varepsilon(X^2)}.$$

Ainsi, pour un réel $\alpha > 0$, on obtient

$$\mathbb{P}(X_j \geq h_j^{-\alpha} \varepsilon(X_j)) \geq (1 - h_j^{-\alpha})^2 \frac{e^{-2N_j t^4}}{(h_j - 1 + e^{N_j t^2})/h_j}.$$

Comme d'autre part, $e^{t\|\tilde{g}\|_{\hat{H}_j}} \geq X_j$, on obtient

$$\mathbb{P}(\|\tilde{g}\|_{\hat{H}_j} \geq N_j \frac{t}{2} - N_j t^3 - \alpha \frac{\ln h_j}{t}) \geq (1 - h_j^{-\alpha})^2 \frac{e^{-2N_j t^4}}{(h_j - 1 + e^{N_j t^2})/h_j}.$$

Si on écrit $t = \lambda \sqrt{\frac{\ln h_j}{N_j}}$ (avec $\lambda > 0$) cela devient

$$\mathbb{P}\left(\|\tilde{g}\|_{\hat{H}_j} \geq (\frac{\lambda}{2} - \lambda^3 \frac{\ln h_j}{N_j} - \frac{\alpha}{\lambda}) \sqrt{N_j \ln h_j}\right) \geq (1 - h_j^{-\alpha})^2 \frac{e^{-2\lambda^4 \frac{(\ln h_j)^2}{N_j}}}{(h_j - 1 + h_j^{\lambda^2})/h_j}.$$

Si $\lambda^2 < 1$, le terme de droite est supérieur à $(1 - h_j^{-\alpha})^2 e^{-2\lambda^4 \frac{(\ln h_j)^2}{N_j}} (1 - h_j^{\lambda^2 - 1})$, avec $e^{-2\lambda^4 \frac{(\ln h_j)^2}{N_j}} \geq 1 - 2\frac{(\ln h_j)^2}{N_j}$. D'où le résultat. □

**Théorème 3.3** *Soit $(G_j)$ (respectivement $(H_j)$) une suite de sous-groupes additifs de $(\mathbb{F}_{q_j})$ de cardinal $N_j$ (resp. $h_j$). Si la série $\sum \frac{1}{h_j^\alpha}$ est convergente pour tout $\alpha$ strictement positif et si la série $\sum \frac{(\ln h_j)^2}{N_j}$ est convergente alors pour $j$ suffisamment grand et $g \in \Omega_\infty$, on a presque sûrement*

$$\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} \geq \frac{1}{2}\sqrt{N_j \ln h_j} + o(\sqrt{N_j \ln h_j}) \sim \frac{1}{2\sqrt{2}}\sqrt{2 \times Card\, G_j \times \ln Card\, \hat{H}_j}.$$

*Démonstration.* Soit $g \in \Omega_\infty$, d'après la proposition précédente

$$\mathbb{P}\left(\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} < (\frac{\lambda}{2} - \lambda^3\frac{\ln h_j}{N_j} - \frac{\alpha}{\lambda})\sqrt{N_j \ln h_j}\right) \leq 1 - (1 - 2h_j^{-\alpha})(1 - 2\frac{(\ln h_j)^2}{N_j})(1 - h_j^{\lambda^2-1}),$$

donc via le lemme de Borel-Cantelli, si la série dont le terme principal est le terme de droite ci-dessus converge, alors pour $j$ suffisamment grand, on a presque sûrement

$$\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} \geq (\frac{\lambda}{2} - \lambda^3\frac{\ln h_j}{N_j} - \frac{\alpha}{\lambda})\sqrt{N_j \ln h_j}.$$

La convergence de la série étant une conséquence de nos hypothèses, on conclut en faisant tendre $\lambda$ vers 1 et $\alpha$ vers 0 (puisque $\ln h_j = o_{j\to\infty}(N_j)$ par convergence de la série de terme principal $\frac{(\ln h_j)^2}{N_j}$). □

Cette borne est donc une généralisation de la borne inférieure expliquée dans [Rod03b], les hypothèses de convergence ne sont pas beaucoup plus fortes que dans le théorème 3.2, alors que la restriction aux sous-groupes additifs est très contraignante. A l'aide de ces deux théorèmes, quand ils s'appliquent, on observe donc asymptotiquement que les non-linéarités partielles sont concentrées dans une même région, de manière analogue au cas de la non-linéarité classique.

## 3.4 Techniques de Halácz

Nous étudions maintenant les possibles améliorations de la borne inférieure obtenue ci-dessus en généralisant les résultats de [Rod03a] qui utilisent des techniques inspirées par les travaux de Halácz [Hal73].

L'outil principal au coeur des démonstrations suivantes est la définition d'une variable aléatoire (cf. Sec. 3.4.1) qui permet de comparer la valeur de l'amplitude spectrale par rapport à deux réels donnés. Toujours à l'aide de résultats classiques de théorie des probabilités, nous étudions alors les conditions pour obtenir la borne asymptotique supérieure visée (dans la section 3.4.2) et la borne asymptotique inférieure (Sec. 3.4.3).

### 3.4.1 Outils

Pour démarrer cette étude, nous adaptons tout d'abord les fonctions auxiliaires utilisées dans [Rod03a] à notre contexte.

Soit $(M_j)_{j\in\mathbb{N}}$, $(\Delta_j)_{j\in\mathbb{N}}$ deux suites de réels strictement positifs quelconques. Soit $\alpha$ une fonction réelle monotone infiniment différentiable sur $[0,1]$ telle que $\alpha(0)=0, \alpha(1)=1$ et telle que les dérivées $p$-ième $\alpha^{(p)}$ s'annulent toutes en 0 et 1. Pour tout $j\in\mathbb{N}$, on construit alors une fonction $u_j$ définie en $x\in\mathbb{R}$ par

$$u_j(x)=\begin{cases} 0 & \text{si} \quad |x|\leq M_j, \\ \alpha(\frac{|x|-M_j}{\Delta_j}) & \text{si} \quad M_j\leq|x|\leq M_j+\Delta_j, \\ 1 & \text{si} \quad |x|\geq M_j+\Delta_j. \end{cases}$$

Ainsi $u_j$ est également infiniment différentiable. Pour $g$ fixé ($g\in\Omega_j$), on définit enfin la variable aléatoire (réelle et positive)

$$\eta_j=\int_{\hat{H}_j} u_j(\tilde{g}(\mu))d\mu=\frac{1}{h_j}\sum_{\hat{H}_j} u_j(\tilde{g}(\mu)).$$

On a ainsi

$$\begin{aligned}\eta_j &= \int_{\hat{H}_j}\int_{\mathbb{R}} \exp(it\tilde{g}(\mu))U_j(t)dtd\mu \\ &= \int_{\mathbb{R}}\int_{\hat{H}_j} \exp(it\tilde{g}(\mu))d\mu U_j(t)dt,\end{aligned}$$

où $U_j$ est la transformée de Fourier réelle de $u_j$. Elle vérifie

$$u_j(x)=\int_{\mathbb{R}}\exp(itx)U_j(t)dt.$$

Les résultats obtenus par la suite sur la non-linéarité sont fortement liés à l'analyse probabiliste de certaines valeurs prises par les variables aléatoires $\eta_j$. En fait, la définition de $u_j$ implique que $\eta_j$ est dépendante de la position de l'amplitude spectrale de $g$ par rapport à $M_j$ et $M_j+\Delta_j$ et nous étudions donc la distribution de valeurs de $\eta_j$ via des méthodes probabilistes adaptées. Cette méthode peut être vue comme une méthode duale de la précédente où $\eta_j$ correspond à une projection de $g$, permettant au final une analyse plus fine.

Afin de faciliter la lecture, les démonstrations techniques sont reportées en annexe C à partir de la page 187.

#### 3.4.1.1 Estimations asymptotiques élémentaires

Nous énonçons ici les principales estimations utilisées par la suite pour obtenir les estimations asymptotiques des différentes probabilités liées à la distribution des fonctions dans $\Omega_\infty$.

**Remarque 3.6** *Pour le résultat suivant, on utilise notamment le fait que la transformée de Fourier dans $\mathbb{R}$ d'une gaussienne $x\mapsto e^{-\frac{x^2}{2}}$ est la gaussienne définie par $t\mapsto\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$. Réciproquement la transformée inverse de la gaussienne $t\mapsto e^{-\frac{t^2}{2}}$ est alors $x\mapsto\sqrt{2\pi}e^{-\frac{x^2}{2}}$.*

**Proposition 3.2** *Pour $j \to \infty$,*

$$\int_{\mathbb{R}} |U_j(t)dt| = O(\frac{M_j}{\Delta_j}),\ si\ \Delta_j = O(M_j), \tag{3.3}$$

*et*

$$\int_{\mathbb{R}} |t^p U_j(t)dt| = O(\frac{1}{\Delta_j^p})\ pour\ p \geq 1,\ si\ \Delta_j \geq 1. \tag{3.4}$$

*De plus pour $c > 0$,*

$$|\int_{\mathbb{R}} e^{-ct^2/2} t^p U_j(t)dt| = O(\frac{1}{\Delta_j^{p-1}\sqrt{c}} e^{-M_j^2/2c})\ pour\ 1 \leq p, \tag{3.5}$$

*et*

$$|\int_{\mathbb{R}} e^{-ct^2/2} U_j(t)dt| = O(\frac{\sqrt{c}}{M_j} e^{-M_j^2/2c}). \tag{3.6}$$

*Démonstration.* La démonstration est détaillée en annexe C, elle reste similaire à celle de Halácz (cf. [Rod03a]), même si les estimations obtenues sont ici légèrement plus générales. □

Dans toute la suite, on supposera $\Delta_j \geq 1$ afin de pouvoir appliquer l'estimation (3.4).

### 3.4.2 Borne supérieure

Soit $j \in \mathbb{N}$ fixé ; dans cette section, $g$ désignera un élément de $\Omega_j$ excepté lors de l'estimation finale où $j$ ne sera plus fixe. Dans cette partie, nous allons montrer que pour des choix adéquats (cf. Théorème 3.4), $M_j + \Delta_j$ peut être vu comme une borne supérieure pour presque toutes les fonctions. La propriété suivante est l'idée à la base de la démonstration, c'est ce qui motive l'introduction de la variable $\eta_j$.

**Proposition 3.3** *Si $\|\tilde{g}\|_{\hat{H}_j} \geq M_j + \Delta_j$ alors $\eta_j \geq 1/h_j$.*

*Démonstration.* Si $\|\tilde{g}\|_{\hat{H}_j} \geq M_j + \Delta_j$, alors il existe $\mu \in \hat{H}_j$ tel que

$$|\tilde{g}(\mu)| \geq M_j + \Delta_j,$$

i.e. tel que $u_j(\tilde{g}(\mu)) = 1$. D'où, le résultat par définition de $\eta_j$. □

Par la suite, nous étudions donc la probabilité d'avoir $\eta_j$ plus grand que $1/h_j$. Afin de majorer cette probabilité, nous allons nous intéresser à l'estimation de l'espérance de $\eta_j$ pour pouvoir utiliser l'inégalité de Markov.

**Lemme 3.4** *Si $t \in \mathbb{R}$, $\mu \in \hat{H}_j$ et $x \in G_j$, alors l'espérance pour $g \in \Omega_j$ de $e^{itg(x)\mu(x)}$ est* $\cos t$.

*Démonstration.* Ce résultat est similaire à celui vu dans la partie 3.3. En effet, à $x$ fixé, $g(x)\mu(x)$ prend la valeur 1 pour la moitié des éléments de $\Omega_j$ et $-1$ sur l'autre moitié. D'où $\varepsilon(e^{itg(x)\mu(x)}) = \int_{\Omega_j} e^{itg(x)\mu(x)} d\mathbb{P} = e^{it}/2 + e^{-it}/2$. □

Par l'intermédiaire de développements limités, on en déduit :

**Lemme 3.5** *Soit $\mu \in \hat{H}_j$ fixé et $t \in \mathbb{R}$, alors $\varepsilon(e^{it\tilde{g}(\mu)}) = (\cos t)^{N_j}$, et pour $|t| \leq 1$ tel que $t \to 0$,*

$$\varepsilon(e^{it\tilde{g}(\mu)}) = e^{-N_j t^2/2} - N_j t^4/12 e^{-N_j t^2/2} + N_j O_{t\to 0}(t^6) + N_j^2 O_{t\to 0}(t^8). \tag{3.7}$$

*Démonstration.* cf. Annexe C. □

Les deux lemmes ci-dessus permettent d'obtenir une approximation asymptotique de l'espérance de $\eta_j$ :

**Lemme 3.6**

$$\varepsilon(\eta_j) = O(\frac{\sqrt{N_j}}{M_j} e^{-M_j^2/2N_j}) + O(\frac{\sqrt{N_j}}{\Delta_j^3} e^{-M_j^2/2N_j}) + O(\frac{N_j}{\Delta_j^6}) + O(\frac{N_j^2}{\Delta_j^8}).$$

*Démonstration.* La démonstration est réalisée en se ramenant à l'application du développement limité du lemme précédent via une adaptation des intervalles d'intégration (cf. Annexe C). □

On peut maintenant faire le lien entre $\varepsilon(\eta_j)$ et la valeur de l'amplitude spectrale d'une fonction $g$ par rapport à $M_j + \Delta_j$.

**Proposition 3.4** *Pour $j$ quelconque,*

$$\mathbb{P}(\|\tilde{g}\|_{\hat{H}_j} \geq M_j + \Delta_j) \leq \mathbb{P}(\eta_j \geq 1/h_j) \leq h_j \varepsilon(\eta_j).$$

*Démonstration.* La preuve découle directement de la proposition 3.3 et de l'inégalité de Markov appliquée dans l'espace de probabilité $\Omega_j$. □

Ce qui implique finalement,

**Théorème 3.4** *Soient $(G_j)$ une suite de sous-ensembles, $G_j \subset \mathbb{F}_{q_j}$ et de cardinal $N_j$, et $(\hat{H}_j)$ suite de sous-ensembles de $\widehat{\mathbb{F}_{q_j}}$ et de cardinal $h_j$.*

*Si les séries $\sum \frac{h_j\sqrt{N_j}}{M_j} e^{-M_j^2/2N_j}$, $\sum \frac{\sqrt{N_j}h_j}{\Delta_j^3} e^{-M_j^2/2N_j}$, $\sum \frac{N_j h_j}{\Delta_j^6}$ et $\sum \frac{N_j^2 h_j}{\Delta_j^8}$ sont convergentes, alors pour $j$ suffisamment grand, et $g \in \Omega_\infty$, on a presque sûrement (au sens de la probabilité $\mathbb{P}$)*

$$\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} < M_j + \Delta_j.$$

*Démonstration.* Cela découle du lemme 3.6 et de l'application du lemme de Borel-Cantelli comme pour le théorème 3.2 puisque sous les hypothèses de convergences des séries énoncées ci-dessus,

$$\sum_j \mathbb{P}(\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} \geq M_j + \Delta_j) \leq \sum_j h_j \varepsilon(\eta_j) < \infty.$$

□

**Remarque 3.7** *Les hypothèses de convergence de ces séries impliquent en particulier que les suites formées des différents termes convergent vers* 0*, i.e. qu'il faut au minimum* $h_j \sqrt{N_j} = o(M_j e^{M_j^2/2N_j})$*,* $h_j \sqrt{N_j} = o(\Delta_j^3 e^{M_j^2/2N_j})$*,* $h_j N_j = o(\Delta_j^3)$ *et* $h_j N_j = o(\Delta_j^6)$*.*

Nous verrons par la suite des exemples d'applications de ce théorème, et en particulier comment obtenir la même borne que dans le théorème 3.2 mais avec des conditions légèrement différentes. En effet, ici les contraintes sur $N_j$ et $h_j$ sont dépendantes des choix de $M_j$ et $\Delta_j$, ce qui laisse une certaine liberté supplémentaire.

### 3.4.3 Borne inférieure

Comme dans la section précédente, on considère pour commencer $j \in \mathbb{N}$ fixé et $g$ dans $\Omega_j$. Nous allons maintenant étudier la probabilité pour $\eta_j$ de prendre la valeur 0, ceci après avoir remarqué que si $\|\tilde{g}\|_{\hat{H}_j} < M_j$ alors $\eta_j = 0$. Nous allons démontrer que cette probabilité est presque sûrement nulle pour $j$ suffisamment grand, de sorte à en déduire une borne inférieure asymptotique sur les amplitudes spectrales.

Par rapport à la section précédente, nous faisons l'hypothèse supplémentaire dans cette partie (comme dans la section 3.3.2) que $G_j$ est un sous-groupe de $\mathbb{F}_{q_j}$ pour tout $j$.

Tout d'abord, d'après l'inégalité de Bienaymé-Tchebychev (cf. par exemple [Loe77, Rev97, Shi96]) on remarque que :

**Proposition 3.5**

$$\mathbb{P}(\eta_j = 0) \leq \mathbb{P}\Big(|\eta_j - \varepsilon(\eta_j)| \geq \varepsilon(\eta_j)\Big) \leq \frac{\varepsilon(\eta_j^2) - \varepsilon^2(\eta_j)}{\varepsilon^2(\eta_j)}.$$

Pour appliquer cette proposition, il nout faut donc connaître une estimation de l'espérance de $\eta_j$ et de $\eta_j^2$, du moins pour $j$ tendant vers l'infini.

#### 3.4.3.1 Espérance de $\eta_j$

L'étude découle directement du corps de la démonstration effectuée pour le lemme 3.6 dans la section sur la borne supérieure.

**Lemme 3.7** *Pour* $j \to \infty$*,*

$$\varepsilon(\eta_j) = \int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t) dt + O_j\Big(\frac{\sqrt{N_j}}{\Delta_j^3} e^{-M_j^2/2N_j}\Big) + O\Big(\frac{N_j}{\Delta_j^6}\Big) + O\Big(\frac{N_j^2}{\Delta_j^8}\Big).$$

*Démonstration.* cf. Annexe C. □

#### 3.4.3.2 Espérance de $\eta_j^2$

Pour le calcul de $\varepsilon(\eta_j^2)$, on commence par séparer $\eta_j^2$ en 2 morceaux,

$$\begin{aligned}\eta_j^2 &= \Big(\int_{\hat{H}_j} u_j(\tilde{g}(\mu))d\mu\Big)^2 \\ &= \int\int_{\mu_{|G_j}=\nu_{|G_j}} u_j(\tilde{g}(\mu))u_j(\tilde{g}(\nu))d\mu d\nu + \int\int_{\mu_{|G_j}\neq\nu_{|G_j}} u_j(\tilde{g}(\mu))u_j(\tilde{g}(\nu))d\mu d\nu.\end{aligned}$$

Nous nous intéressons tout d'abord à la première partie dont voici un résultat concernant son ensemble de sommation.

**Lemme 3.8** *Soit $\mu \in \hat{G}_j$ un générateur du groupe cyclique $\hat{G}_j$, alors*

$$\sigma_j(\mu_a) = Card\,\{\nu \in \hat{H}_j, \nu_{|G_j} = \mu_a\} = \sigma_j$$

*est indépendant de a pour $a \in G_j$ où $\mu_a \in \hat{G}_j$ est défini par $\mu_a(x) = 1$ pour $x \in G_j - G_j \cap H_j$ et $\mu_a(x) = \mu(ax)$ pour $x \in G_j \cap H_j$.*

*Démonstration.* On remarque que $\sigma_j$ est le cardinal du quotient $\hat{H}_j/\widehat{G_j \cap H_j}$ ; cf. Annexe C. □

Il résulte immédiatement de ce lemme une majoration du premier morceau de $\eta_j^2$.

**Lemme 3.9**

$$\int\int_{\mu_{|G_j}=\nu_{|G_j}} u_j(\tilde{g}(\mu))u_j(\tilde{g}(\nu))d\mu d\nu \leq \sigma_j\eta_j/h_j.$$

*Démonstration.* cf. Annexe C. □

Concernant le deuxième morceau, par définition de la transformée de $u_j$, on remarque que

$$\begin{aligned}&\int\int_{\mu_{|G_j}\neq\nu_{|G_j}} u_j(\tilde{g}(\mu))u_j(\tilde{g}(\nu))d\mu d\nu \\ &= \int_{\mu_{|G_j}\neq\nu_{|G_j}}\int_{\mathbb{R}^2} e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}U_j(t_1)U_j(t_2)dt_1dt_2d\mu d\nu.\end{aligned}$$

Pour calculer son espérance, nous nous intéressons tout d'abord à l'espérance du terme interne.

**Lemme 3.10** *Soient $t_1, t_2 \in [-1/2, 1/2]$ et $\mu, \nu \in \hat{H}_j$ avec $\mu_{|G_j} \neq \nu_{|G_j}$, alors*

$$\begin{aligned}\varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}) = e^{-N_j(t_1^2+t_2^2)/2} &- \frac{1}{12}N_j(t_1^4+t_2^4+6t_1^2t_2^2)e^{-N_j(t_1^2+t_2^2)/2} \\ &+N_jO((|t_1|+|t_2|)^6) + N_j^2O((|t_1|+|t_2|)^8).\end{aligned}$$

*Démonstration.* La démonstration (cf. Annexe C) utilise des développements limités comme pour le lemme 3.5 et nécessite, pour simplifier les résultats, l'utilisation de relations d'orthogonalité. On y utilise notamment le fait que $\sum_{x\in G_j}\mu(x)\nu(x)=0$. □

**Remarque 3.8** *Comme dans la partie 3.3.2, on peut voir dans la démonstration, page 192, ce qui motive le choix de la condition "$G_j$ sous-groupe de $\mathbb{F}_{q_j}$". En effet, si ce n'était pas le cas, on ne pourrait pas simplifier les dernières sommes via les relations d'orthogonalités appliquées à $\mu\times\nu$, et des puissances impaires de $t_1$ et $t_2$ à coefficients indéterminés seraient à prendre en compte. Le fait de n'avoir que des puissances paires à coefficients connus et de plus une expression symétrique en $t_1$ et $t_2$ permet d'aboutir à une estimation plus précise dans le lemme suivant.*

On continue le travail d'estimation de la deuxième partie de l'espérance de $\eta_j^2$, en sommant maintenant sur $\mathbb{R}^2$.

**Lemme 3.11** *Soit $\mu,\nu\in\hat{H}_j$ avec $\mu_{|G_j}\neq\nu_{|G_j}$, alors*

$$\begin{aligned}\varepsilon\Big(&\int_{\mathbb{R}^2} e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}U_j(t_1)U_j(t_2)dt_1dt_2\Big)\\ &= \left(\int_{\mathbb{R}} e^{-N_jt^2/2}U_j(t)dt\right)^2+O\Big(\Big(\frac{1}{\Delta_j^2}+\frac{N_j}{\Delta_j^3M_j}\Big)e^{-M_j^2/N_j}\Big)+O\Big(\frac{N_j}{\Delta_j^6}\Big)+O\Big(\frac{N_j^2}{\Delta_j^8}\Big).\end{aligned}$$

*Démonstration.* C'est une application du lemme précédent, sous l'intégrale – cf. Annexe C. □

Enfin, on peut en déduire l'estimation finale de $\varepsilon(\eta_j^2)$.

**Proposition 3.6** *Pour $j\to\infty$,*

$$\varepsilon(\eta_j^2) \leq \frac{\sigma_j}{h_j}\varepsilon(\eta_j)+\left(\int_{\mathbb{R}} e^{-N_jt^2/2}U_j(t)dt\right)^2+O\Big(\Big(\frac{1}{\Delta_j^2}+\frac{N_j}{\Delta_j^3M_j}\Big)e^{-M_j^2/N_j}\Big)+O\Big(\frac{N_j}{\Delta_j^6}\Big)+O\Big(\frac{N_j^2}{\Delta_j^8}\Big).$$

*Démonstration.* D'après les lemmes précédents,

$$\begin{aligned}\varepsilon(\eta_j^2) &= \varepsilon\Big(\int\int_{\mu_{|G_j}=\nu_{|G_j}} u_j(\tilde{g}(\mu))u_j(\tilde{g}(\nu))d\mu d\nu\Big)+\varepsilon\Big(\int\int_{\mu_{|G_j}\neq\nu_{|G_j}} u_j(\tilde{g}(\mu))u_j(\tilde{g}(\nu))d\mu d\nu\Big)\\ &\leq \frac{\sigma_j}{h_j}\varepsilon(\eta_j)+\varepsilon\Big(\int_{\mu_{|G_j}\neq\nu_{|G_j}}\int_{\mathbb{R}^2} e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}U_j(t_1)U_j(t_2)dt_1dt_2d\mu d\nu\Big)\\ &\leq \frac{\sigma_j}{h_j}\varepsilon(\eta_j)+\int_{\mu_{|G_j}\neq\nu_{|G_j}}\left(\int_{\mathbb{R}} e^{-N_jt^2/2}U_j(t)dt\right)^2 d\mu d\nu\\ &\quad+\int_{\mu_{|G_j}\neq\nu_{|G_j}}\left(O\Big(\Big(\frac{1}{\Delta_j^2}+\frac{N_j}{\Delta_j^3M_j}\Big)e^{-M_j^2/N_j}\Big)+O\Big(\frac{N_j}{\Delta_j^6}\Big)+O\Big(\frac{N_j^2}{\Delta_j^8}\Big)\right)d\mu d\nu\\ &\leq \frac{\sigma_j}{h_j}\varepsilon(\eta_j)+\left(\int_{\mathbb{R}} e^{-N_jt^2/2}U_j(t)dt\right)^2+O\Big(\Big(\frac{1}{\Delta_j^2}+\frac{N_j}{\Delta_j^3M_j}\Big)e^{-M_j^2/N_j}\Big)+O\Big(\frac{N_j}{\Delta_j^6}\Big)+O\Big(\frac{N_j^2}{\Delta_j^8}\Big),\end{aligned}$$

où la dernière inégalité est réalisée puisque $\int_{\mu_{|G_j} \neq \nu_{|G_j}} d\mu d\nu \leq 1$. □

On applique maintenant l'inégalité de Tchebychev :

**Proposition 3.7** *Il existe $A > 0$ tel que pour $j$ qui tend vers l'infini,*

$$\begin{aligned}
\mathbb{P}(\eta_j = 0) \quad \leq \quad & \left(\frac{\sigma_j}{h_j}\right) \Bigg/ \left( \sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left( \frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5}) \right) \right) \\
+ \quad & \left( O((\frac{1}{\Delta_j^2} + \frac{N_j}{\Delta_j^3 M_j})e^{-M_j^2/N_j}) + \frac{N_j}{\Delta_j^6}(1 + O(\frac{N_j}{\Delta_j^2})) \right. \\
& + \frac{\sqrt{N_j}N_j e^{-M_j^2/2N_j}}{\Delta_j^6}(O(\frac{1}{\Delta_j^3}) + O(\frac{N_j}{\Delta_j^5}) + O(\frac{N_j}{\Delta_j^2 M_j}) + O(\frac{1}{M_j})) \\
& \left. + \frac{N_j^2}{\Delta_j^{12}}(O(1) + O(\frac{N_j}{\Delta_j^2}) + O(\frac{N_j^2}{\Delta_j^4})) \right) \\
& \Bigg/ \left( \sqrt{\frac{2\pi}{N_j}}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left( \frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5}) \right) \right)^2
\end{aligned}$$

*Démonstration.* C'est une application technique des lemmes précédents, combinée avec l'inégalité de Tchebychev ; cf. Annexe C. □

D'où la borne asymptotique suivante, valide pour presque toutes les fonctions binaires.

**Théorème 3.5** *Soit $(G_j)$ (respectivement $(H_j)$) une suite de sous-groupes additifs de $(\mathbb{F}_{q_j})$ de cardinal $N_j$ (resp. $h_j$). Si la série $\sum \mathbb{P}(\eta_j = 0)$ est convergente, alors pour $j$ grand, on a presque sûrement pour $g \in \Omega_\infty$,*

$$\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} \geq M_j.$$

*Démonstration.* On applique le lemme de Borel-Cantelli de la même manière que dans la démonstration du théorème 3.4. □

Ici encore, la borne inférieure est restreinte au cas où $G_j$ est un groupe additif. Cependant, comparé au théorème 3.3, la contrainte de convergence de la série $\sum \mathbb{P}(\eta_j = 0)$ laisse une liberté intéressante quant au choix de $M_j$ et nous verrons plus loin comment cela permet d'améliorer la borne inférieure.

## 3.5 Cas de la non-linéarité sur $\mathbb{F}_{q_j}$

Comme première application, nous montrons comment retrouver les résultats précédemment démontrés dans [Rod03a] en appliquant les théorèmes 3.4 et 3.5 avec les instances adéquates (à noter que cet exercice est immédiat pour les théorèmes 3.2 et 3.3 et leur lien avec [Rod03b]). Ces 4 théorèmes sont en effet des généralisations des résultats établis par Rodier.

Si on fixe $G_j = H_j = \mathbb{F}_{q_j}$, alors pour $g \in \Omega_j$, on voit que $\tilde{g} = \hat{g}$ et $\|\tilde{g}\|_{\widehat{\mathbb{F}_{q_j}}} = \|\hat{g}\|_\infty$ ; on retrouve ainsi la notion d'amplitude spectrale liée à la non-linéarité de $g$ sur $\mathbb{F}_{q_j}$. Par ce choix, on a $N_j = h_j = q_j$.

Si on pose $\Delta_j = \sqrt{\frac{q_j}{\ln q_j}}$ comme dans [Rod03a], on a alors $\Delta_j \geq 1$, l'estimation (3.4) est donc bien vérifiée. On va ensuite faire un choix pour $M_j$ spécifique à chaque cas (borne supérieure et borne inférieure).

### 3.5.1 Borne supérieure

Pour $M_j = \sqrt{2q_j \ln q_j} + 2\sqrt{\frac{q_j}{\ln q_j}} \ln\ln q_j$, on souhaite utiliser le théorème 3.4, ce qui donnerait pour $j$ suffisamment grand, presque sûrement

$$\|\widehat{f_{|\mathbb{F}_{q_j}}}\|_\infty \quad < \quad \sqrt{2q_j \ln q_j} + 2\sqrt{\frac{q_j}{\ln q_j}} \ln\ln q_j + \sqrt{\frac{q_j}{\ln q_j}},$$

soit :

**Corollaire 3.1** *Pour $j$ grand,*

$$\|\widehat{f_{|\mathbb{F}_{q_j}}}\|_\infty \quad < \quad \sqrt{2q_j \ln q_j} + 3\sqrt{\frac{q_j}{\ln q_j}} \ln\ln q_j \tag{3.8}$$

*pour presque tout $f \in \Omega$.*

*Démonstration.* Pour appliquer ce théorème, il suffit de vérifier que les séries $\sum \frac{h_j\sqrt{N_j}}{M_j} e^{-M_j^2/2N_j}$, $\sum \frac{\sqrt{N_j}h_j}{\Delta_j^3} e^{-M_j^2/2N_j}$, $\sum \frac{N_j h_j}{\Delta_j^6}$ et $\sum \frac{N_j^2 h_j}{\Delta_j^8}$ sont convergentes. Pour commencer, nous avons

$$\begin{aligned} e^{-M_j^2/2N_j} &= e^{-(\ln q_j + \frac{2(\ln\ln q_j)^2}{\ln q_j} + 2\sqrt{2}\ln\ln q_j)} \\ &= e^{-2\frac{(\ln\ln q_j)^2}{\ln q_j}} \frac{1}{q_j(\ln q_j)^{2\sqrt{2}}}. \end{aligned}$$

Or $\frac{(\ln\ln q_j)^2}{\ln q_j} = o(1)$ donc $e^{-M_j^2/2N_j} = O(\frac{1}{q_j(\ln q_j)^{2\sqrt{2}}})$, ce qui implique

1. $\frac{h_j\sqrt{N_j}}{M_j} e^{-M_j^2/2N_j} = O(\frac{1}{(\ln q_j)^{2\sqrt{2}+1/2}})$,
2. et $\frac{\sqrt{N_j}h_j}{\Delta_j^3} e^{-M_j^2/2N_j} = O(\frac{1}{q_j^{5/2}(\ln q_j)^{2\sqrt{2}}})$.

D'autre part, $\frac{N_j h_j}{\Delta_j^6} = \frac{(\ln q_j)^3}{q_j}$ et $\frac{N_j^2 h_j}{\Delta_j^8} = \frac{(\ln q_j)^4}{q_j}$, donc ces 4 expressions sont clairement en $O(\frac{1}{(\ln q_j)^2})$. La somme $\sum_{x\in\mathbb{N}^\times} \frac{1}{x^2}$ étant finie, les hypothèses de convergence de ces 4 séries sont vérifiées et le théorème s'applique. □

### 3.5.2 Borne inférieure

Nous allons utiliser le théorème 3.5 avec

$$M_j = \sqrt{2q_j \ln q_j} - 4\sqrt{\frac{q_j}{\ln q_j}} \ln\ln q_j.$$

**Corollaire 3.2** *Pour $j$ suffisamment grand, on a*

$$\|\widehat{f_{|\mathbb{F}_{q_j}}}\|_\infty \geq \sqrt{2q_j \ln q_j} - 4\sqrt{\frac{q_j}{\ln q_j}} \ln\ln q_j$$

*pour presque tout $f \in \Omega$.*

*Démonstration.* Il faut pour cela que la série $\sum \mathbb{P}(\eta_j = 0)$ converge. Or, on connaît d'après la proposition 3.7 une majoration de $\mathbb{P}(\eta_j = 0)$,

$$\begin{aligned}
&\mathbb{P}(\eta_j = 0)\\
&\leq \left(\frac{\sigma_j}{h_j}\right)\left/\left(\sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)\right)\right.\\
&+ \left(O((\frac{1}{\Delta_j^2} + \frac{N_j}{\Delta_j^3 M_j})e^{-M_j^2/N_j}) + \frac{N_j}{\Delta_j^6}(1 + O(\frac{N_j}{\Delta_j^2}))\right.\\
&\quad + \frac{\sqrt{N_j}N_j e^{-M_j^2/2N_j}}{\Delta_j^6}(O(\frac{1}{\Delta_j^3}) + O(\frac{N_j}{\Delta_j^5}) + O(\frac{N_j}{\Delta_j^2 M_j}) + O(\frac{1}{M_j}))\\
&\quad \left. + \frac{N_j^2}{\Delta_j^{12}}(O(1) + O(\frac{N_j}{\Delta_j^2}) + O(\frac{N_j^2}{\Delta_j^4}))\right)\\
&\quad \left/\left(\sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)\right)^2.\right.
\end{aligned}$$

De manière similaire au cas précédent, on prouve aisément que les relations suivantes sont réalisées[1], $e^{-(M_j+2\Delta_j)^2/2N_j} = \Omega(\frac{(\ln q_j)^{4\sqrt{2}}}{q_j})$ et que $e^{-(M_j)^2/2N_j} = O(\frac{(\ln q_j)^{4\sqrt{2}}}{q_j})$. A noter d'autre part qu'ici $\sigma_j = 1$. Ainsi,

$$\left(\sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)\right) = \Omega(\frac{(\ln q_j)^{4\sqrt{2}-1/2}}{q_j})$$

[1] $a_j = \Omega_{j\to\infty}(b_j)$ signifiant que $a_j$ est minoré à partir d'un certain rang par un multiple non nul de $b_j$, i.e. $b_j = O(a_j)$

et

$$\begin{aligned}
&\mathbb{P}(\eta_j = 0)\\
&\leq O(\frac{1}{(\ln q_j)^{4\sqrt{2}-1/2}}) + \Big(O(\frac{(\ln q_j)^{8\sqrt{2}}}{q_j^2}\frac{\ln q_j}{q_j}) + \frac{(\ln q_j)^3}{q_j^2}(1+O(\ln q_j))\\
&\quad + \frac{(\ln q_j)^{4\sqrt{2}+3}}{q_j^3}\Big(O(\frac{(\ln q_j)^{3/2}}{q_j}) + O(\frac{(\ln q_j)^{5/2}}{q_j}) + O(\sqrt{\ln q_j}) + O(\frac{1}{\sqrt{\ln q_j}})\Big)\\
&\quad + \frac{(\ln q_j)^6}{q_j^4}(O(1)+O(\ln q_j)+O((\ln q_j)^2))\Big)\Big/\left(\frac{(\ln q_j)^{4\sqrt{2}-1/2}}{q_j}\right)^2\\
&\leq O(\frac{1}{(\ln q_j)^{4\sqrt{2}-1/2}}) + O(\frac{(\ln q_j)^2}{q_j}) + O(\frac{1}{(\ln q_j)^{8\sqrt{2}-4}})\\
&\quad + O(\frac{1}{q_j(\ln q_j)^{4\sqrt{2}-9/2}}) + O(\frac{1}{q_j^2(\ln q_j)^{8\sqrt{2}-7}})
\end{aligned}$$

ce qui implique $\mathbb{P}(\eta_j = 0) = O(\frac{1}{(\ln q_j)^{4\sqrt{2}-1/2}})$. Or $4\sqrt{2} - 1/2 > 1$ donc la série $\sum \mathbb{P}(\eta_j = 0)$ est convergente et le théorème 3.5 entraîne que pour $j$ suffisamment grand, on a

$$\|\widehat{f_{|\mathbb{F}_{q_j}}}\|_\infty \geq \sqrt{2q_j \ln q_j} - 4\sqrt{\frac{q_j}{\ln q_j}} \ln\ln q_j$$

pour presque tout $f \in \Omega$. □

A l'aide de ces majoration et minoration, on en déduit (cf. [Rod03a]) que :

**Proposition 3.8** *Pour presque tout $f \in \Omega$,*

$$\lim_{j\to\infty} \frac{\|\widehat{f_{|\mathbb{F}_{q_j}}}\|_\infty}{\sqrt{q_j \ln q_j}} = \sqrt{2}.$$

C'est à dire qu'on retrouve bien le théorème 3.1.

**Remarque 3.9** *La même méthode pourrait vraisemblablement être utilisée pour affiner le résultat en un développement asymptotique d'ordre plus important. Ici, on a immédiatement un premier ordre :* $\sqrt{2}\sqrt{q_j \ln q_j} + O(\frac{\ln\ln q_j}{\ln q_j})$.

## 3.6 Applications à des instances quelconques

Nous appliquons maintenant les versions généralisées des théorèmes 3.4 et 3.5 sur des sous-ensembles stricts de $\mathbb{F}_{q_j}$. Dans un premier temps, nous analysons globalement les principaux cas possibles ; le but étant d'étudier les différences entre les bornes résultantes et celles

attachées aux théorèmes 3.2 et 3.3 (en particulier on observe si une amélioration est possible pour la borne inférieure, comme dans le cas du corps) et les conditions associées à ces bornes sont comparées. Puis nous donnons quelques exemples avant de traiter le cas de sous-groupes projectifs. Ce dernier cas nous intéresse tout particulièrement puisque de tels sous-groupes $G_j$ sont utilisés dans certaines constructions de fonctions de type Patterson-Wiedemann généralisé (cf. Chapitre 2) où un passage à la limite permet d'approcher les sommes de Gauss sur $G_j^\perp$ par $\sqrt{q_j}$ .

### 3.6.1 Estimations sur des sous-ensembles

Ici $G_j$ et $H_j$ sont des sous-ensembles quelconques de $\mathbb{F}_{q_j}$ et ce pour tout $j \in \mathbb{N}$. Pour plus de clarté, adoptons maintenant les notation suivantes : on pose $u_j = \frac{h_j\sqrt{N_j}}{M_j}e^{-M_j^2/2N_j}$, $v_j = \frac{\sqrt{N_j}h_j}{\Delta_j^3}e^{-M_j^2/2N_j}$, $w_j = \frac{N_j h_j}{\Delta_j^6}$ et $x_j = \frac{N_j^2 h_j}{\Delta_j^8}$. Pour appliquer le théorème 3.4, il faut que les séries de termes respectifs $u_j, v_j, w_j$ ou $x_j$ soient convergentes.

On supposera ci-dessous, pour simplifier, que $(N_j)$ est une suite croissante divergente vers l'infini telle que la série $\sum \frac{1}{(\ln N_j)^\alpha}$ converge pour un $\alpha > 0$ (par exemple $\alpha > 1$ convient s'il existe une suite d'entiers $(\kappa_j)_j$ croissante et divergente vers l'infini vérifiant $2^{\kappa_j} = O(N_j)$, de sorte que $\frac{1}{\ln N_j} = O(\frac{1}{\kappa_j})$).

#### 3.6.1.1 Borne supérieure

Soient

$$M_j = \sqrt{2N_j(\ln h_j + \alpha \ln\ln N_j)}$$

et

$$\Delta_j = N_j^{1/4} h_j^{1/6} (\ln N_j)^{\alpha/6},$$

alors $e^{-M_j^2/2N_j} = \frac{1}{h_j(\ln N_j)^\alpha}$. On obtient pour ces choix

$$u_j = \frac{N_j^{1/2} h_j}{M_j h_j (\ln N_j)^\alpha} = \frac{\sqrt{N_j}/M_j}{(\ln N_j)^\alpha}.$$

Or

$$\frac{\sqrt{N_j}}{M_j} = \frac{1}{\sqrt{2(\ln h_j + \alpha \ln\ln N_j)}} = O(1)$$

donc la série $\sum u_j$ converge.

$$v_j = \frac{\sqrt{N_j}h_j}{\Delta_j^3 h_j (\ln N_j)^\alpha} = \frac{1}{N_j^{1/4} h_j^{1/2} (\ln N_j)^{\alpha+\alpha/2}} = O(\frac{1}{(\ln N_j)^{\alpha+\alpha/2}})$$

d'où la convergence de la série de terme $v_j$. De plus,

$$w_j = \frac{N_j h_j}{N_j^{3/2} h_j (\ln N_j)^\alpha} = \frac{1}{\sqrt{N_j}(\ln N_j)^\alpha} = O(\frac{1}{(\ln N_j)^\alpha})$$

et

$$x_j = \frac{N_j^2 h_j}{N_j^2 h_j^{4/3} (\ln N_j)^{4\alpha/3}} = O(\frac{1}{(\ln N_j)^{4\alpha/3}}),$$

donc les séries numériques de termes $w_j$ et $x_j$ convergent. Ainsi, d'après le théorème 3.4 pour $j$ suffisamment grand,

$$\|\widetilde{g_{|G_j}}\|_{\hat{H}_j} \quad < \quad \sqrt{2N_j(\ln h_j + \alpha \ln \ln N_j)} + N_j^{1/4} h_j^{1/6} (\ln N_j)^{\alpha/6} \tag{3.9}$$

pour presque tout $g \in \Omega$.

**Remarque 3.10** *Si $(N_j)$ ne vérifie pas l'hypothèse mais que c'est le cas pour $(h_j)$, i.e. il existe $\alpha > 0$ tel que la série $\sum \frac{1}{(\ln h_j)^\alpha}$ est convergente alors on obtient le même résultat en remplaçant $N_j$ par $h_j$ dans les logarithmes $\ln \ln N_j$ et $(\ln N_j)^{\alpha/6}$.*

L'équation (3.9) permet donc d'obtenir la même borne que dans le théorème 3.2 avec quelques conditions complémentaires (le troisième point ci-dessous est repris directement du théorème pour faciliter la comparaison).

**Proposition 3.9** *Soit $(G_j)$ une suite croissante de sous-ensembles de $\mathbb{F}_{q_j}$ de cardinal $N_j$ et $(H_j)$ de cardinal $h_j$ telles que l'une des conditions suivantes est vérifiée :*

1. *Il existe $\alpha > 0$ tel que la série $\sum \frac{1}{(\ln N_j)^\alpha}$ converge, $\ln \ln N_j = o(\ln h_j)$ et $h_j (\ln N_j)^\alpha = o(N_j^{3/2} (\ln h_j)^3)$.*
2. *Il existe $\alpha > 0$ tel que la série $\sum \frac{1}{(\ln h_j)^\alpha}$ converge et $h_j (\ln h_j)^{\alpha-3} = o(N_j^{3/2})$.*
3. *Pour tout $\alpha > 0$, la série $\sum \frac{1}{h_j^\alpha}$ converge.*

*Alors pour presque tout $g \in \Omega$,*

$$\overline{\lim_{j\to\infty}} \frac{\|\widetilde{g_{|G_j}}\|_{\hat{H}_j}}{\sqrt{N_j(\ln h_j)}} \leq \sqrt{2}.$$

Par rapport au théorème 3.2, cette généralisation de la borne asymptotique supérieure obtenue par Rodier dans [Rod03b, Rod04] est valable pour de nouvelles conditions sur les cardinaux $N_j$ et $h_j$. Cependant, on remarque que seule la première condition est réellement nouvelle puisque la troisième (qui est extraite du théorème 3.2) est automatiquement vérifiée quand la deuxième l'est.

#### 3.6.1.2 Borne inférieure

Nous allons maintenant analyser la borne inférieure pour estimer la marge potentielle par rapport à la borne supérieure. Supposons que $G_j$ soit un sous-groupe de $\mathbb{F}_{q_j}$. Pour appliquer le théorème 3.5, il faut que la série $\sum \mathbb{P}(\eta_j = 0)$ converge et la généralisation est alors plus complexe.

**Cas 1.** Pour commencer, si[2] $h_j = \Theta(N_j)$, par analogie avec la démonstration sur $\mathbb{F}_{q_j}$, on a $\mathbb{P}(\eta_j = 0) = O(\frac{\sigma_j}{(\ln N_j)^{4\sqrt{2}-1/2}})$ pour $M_j = \sqrt{2N_j \ln N_j} - 4\sqrt{\frac{N_j}{\ln N_j}} \ln\ln N_j$ et $\Delta_j = \sqrt{\frac{N_j}{\ln N_j}}$. Si par exemple $(\sigma_j)$ est bornée, alors la série $\sum \mathbb{P}(\eta_j = 0)$ converge si la série $\sum \frac{1}{(\ln N_j)^{4\sqrt{2}-1/2}}$ converge.

**Cas 2.** Si $h_j \neq \Theta(N_j)$, d'après la proposition 3.7,

$$\begin{aligned}
&\mathbb{P}(\eta_j = 0)\\
&\leq \left(\frac{\sigma_j}{h_j}\right)\left/\left(\sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)\right.\right.\\
&+ \left(O((\frac{1}{\Delta_j^2} + \frac{N_j}{\Delta_j^3 M_j})e^{-M_j^2/N_j}) + \frac{N_j}{\Delta_j^6}(1 + O(\frac{N_j}{\Delta_j^2}))\right.\\
&\quad + \frac{\sqrt{N_j}N_j e^{-M_j^2/2N_j}}{\Delta_j^6}(O(\frac{1}{\Delta_j^3}) + O(\frac{N_j}{\Delta_j^5}) + O(\frac{N_j}{\Delta_j^2 M_j}) + O(\frac{1}{M_j}))\\
&\quad \left.+ \frac{N_j^2}{\Delta_j^{12}}(O(1) + O(\frac{N_j}{\Delta_j^2}) + O(\frac{N_j^2}{\Delta_j^4}))\right)\\
&\quad \left.\left/\left(\sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)\right.\right)^2.
\end{aligned}$$

Soit $M_j = \sqrt{2N_j \ln h_j} - 4\Delta_j \ln\ln h_j$ et $\Delta_j = \sqrt{\frac{N_j}{\ln h_j}}$, alors si $h_j$ tend vers l'infini et si les conditions $h_j(\ln h_j)^{7-4\sqrt{2}+1/2} = o(N_j^2)$ et $(\ln h_j)^2 = o(N_j^{1/2})$ sont vérifiées, on obtient en développant

$$\begin{aligned}
\mathbb{P}(\eta_j = 0) \quad = \quad & O(\frac{\sigma_j}{(\ln h_j)^{4\sqrt{2}-1/2}}) + O(\frac{\ln h_j}{N_j}) + O(\frac{h_j^2}{N_j^2(\ln h_j)^{8\sqrt{2}-4}})\\
& + O(\frac{h_j}{N_j^2(\ln h_j)^{4\sqrt{2}-3-1/2}}) + O(\frac{h_j^2}{N_j^4(\ln h_j)^{8\sqrt{2}-2}}).
\end{aligned}$$

Par exemple, $\sum \mathbb{P}(\eta_j = 0)$ converge lorsque $h_j = O(N_j)$, que $(\sigma_j)$ est bornée et qu'il existe une suite d'entiers $(\kappa_j)$ telle que $h_j = \Omega(2^{\kappa_j})$.

Il existe donc au moins deux cas pour lesquels on obtient l'amélioration espérée de la borne du théorème 3.3 :

**Proposition 3.10** *Soit $(G_j)$ une suite croissante de sous-groupes de $\mathbb{F}_{q_j}$ de cardinal $N_j$ et $(H_j)$ sous-groupes d'ordre $h_j$ tels que l'une des conditions suivantes est vérifiée :*

1. *$h_j = \Theta(N_j)$ et la série $\sum \frac{\sigma_j}{(\ln N_j)^{4\sqrt{2}-1/2}}$ est convergente.*
2. *Les séries $\sum \frac{1}{(\ln N_j)^{8\sqrt{2}-4}}$ et $\sum \frac{\sigma_j}{(\ln h_j)^{4\sqrt{2}-1/2}}$ sont convergentes et $h_j = O(N_j)$.*

[2] $a_j = \Theta_{j\to\infty}(b_j)$ signifie que $a_j$ est encadré à partir d'un certain rang par des multiples non nuls de $b_j$, i.e. $a_j = O(b_j)$ et $b_j = O(a_j)$

*Alors pour presque tout $g \in \Omega$,*

$$\underline{\lim}_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_{\hat{H}_j}}{\sqrt{N_j(\ln h_j)}} \geq \sqrt{2}.$$

On notera dans ces cas qu'au prix de contraintes plus fortes que celles du théorème 3.3 ($\sum \frac{(\ln h_j)^2}{N_j}$ et $\sum \frac{1}{h_j^\alpha}$ convergentes pour tout $\alpha > 0$), la borne obtenue devient optimale.

**Autres cas.** Le cas $N_j = o(h_j)$ semble plus difficile à appliquer, une version affaiblie de la borne inférieure précédente est toutefois concevable. Plaçons-nous dans un cas particulier : supposons qu'ils existent une suite $(e_j)$ et $\gamma > 0$ tels que $N_j = \Omega(2^{e_j})$, $h_j \geq N_j^\gamma$ et tels que la série $\sum \frac{h_j}{N_j^{\gamma+1}}$ converge. Pour $\Delta_j = \sqrt{N_j}$ et $M_j = \sqrt{2N_j(\ln h_j - \gamma \ln N_j)}$, on obtient

$$\mathbb{P}(\eta_j) \quad = \quad O(\frac{\sigma_j}{N_j^\gamma}) + O(\frac{1}{N_j}) + O(\frac{h_j^2}{N_j^{2\gamma+2}}) + O(\frac{h_j}{N_j^{\gamma+2}}) + O(\frac{h_j}{N_j^{2\gamma+4}}). \tag{3.10}$$

Par exemple, si la suite $(\sigma_j)$ est majorée, la série converge et le théorème 3.5 s'applique.

**Remarque 3.11** *D'autres exemples d'applications de ces théorèmes sont bien entendu possibles. A noter également que les bornes et les conditions ne sont pas nécessairement optimales et peuvent être affinées par d'autres choix de $M_j$ et $\Delta_j$ adaptés au contexte.*

### 3.6.2 Quelques exemples

Afin d'illustrer ces différents cas nous détaillons quelques exemples particuliers. La plupart des résultats suivants découle directement de l'analyse précédente.

#### 3.6.2.1 Cas $H_j = G_j$

D'après la proposition 3.9, les conditions se réduisent ici à une seule condition suffisante : Si pour tout $\alpha > 0$, la série $\sum \frac{1}{N_j^\alpha}$ est convergente alors pour presque tout $g \in \Omega$,

$$\overline{\lim_{j\to\infty}} \frac{\|\widetilde{g_{|G_j}}\|_{\hat{G}_j}}{\sqrt{N_j(\ln N_j)}} \leq \sqrt{2}.$$

Si d'une part $G_j = H_j$ est un groupe additif pour tout $j$ alors d'après le théorème 3.3, on sait qu'avec la même condition de convergence des séries $\sum \frac{1}{N_j^\alpha}$, pour presque tout $g \in \Omega_\infty$,

$$\underline{\lim}_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_{\hat{G}_j}}{\sqrt{N_j \ln N_j}} \geq \frac{1}{2}.$$

D'autre part, ici $\sigma_j = 1$, et les conditions de la proposition 3.10 se réduisent à $\sum \frac{1}{(\ln N_j)^{4\sqrt{2}-1/2}}$ converge. Or, comme $N_j | q_j$, $N_j$ est une puissance de 2 (qui diverge vers $\infty$ par convergence

des séries de terme principal $\frac{1}{N_j^\alpha}$) d'où $\sum \frac{1}{(\ln N_j)^{4\sqrt{2}-1/2}}$ converge puisque c'est dans ce cas, à une constante multiplicative près, une série extraite de la série de terme principal $(\frac{1}{n})^{4\sqrt{2}-1/2}$ (qui est convergente car $4\sqrt{2} - 1/2 > 1$). On obtient alors un résultat généralisant le résultat obtenu sur le corps $\mathbb{F}_{q_j}$ tout entier, à savoir :

**Proposition 3.11** *Soit $N_j$ l'ordre du sous-groupe $G_j = H_j$ tel que pour tout $\alpha > 0$, la série $\sum \frac{1}{N_j^\alpha}$ est convergente, alors pour presque tout $g \in \Omega_\infty$,*

$$\lim_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_{\hat{G_j}}}{\sqrt{N_j \ln N_j}} = \sqrt{2}.$$

Ce dernier résultat n'est absolument pas surprenant puisque $G_j = H_j$ en tant que sous-groupe additif de $\mathbb{F}_{q_j}$, peut être vu comme un sous-corps lorsque $1 \in G_j$. Seul le cas où $1 \notin G_j$ semble nouveau.

#### 3.6.2.2 Cas $H_j = \mathbb{F}_{q_j}$

On a alors $h_j = q_j$, et la condition du théorème 3.2 est donc immédiatement vérifiée si $q_j \to \infty$. Pour presque tout $g \in \Omega_\infty$,

$$\overline{\lim}_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_{\widehat{\mathbb{F}_{q_j}}}}{\sqrt{N_j \ln q_j}} \leq \sqrt{2}.$$

Si d'autre part $G_j$ est un sous-groupe, alors la condition $\sum \frac{(\ln h_j)^2}{N_j}$ converge est équivalente à $N_j \to \infty$ puisque $N_j | q_j$. Ainsi, si $N_j \to \infty$, on a pour presque tout $g \in \Omega_\infty$,

$$\underline{\lim}_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_{\widehat{\mathbb{F}_{q_j}}}}{\sqrt{N_j \ln q_j}} \geq \frac{1}{2}.$$

Pour améliorer cette borne, certains critères supplémentaires sont nécessaires. Dans la proposition 3.10, les conditions se ramènent à $N_j = \Theta(q_j)$ et on a ainsi le résultat suivant :

**Corollaire 3.3** *Soit $N_j$ l'ordre de $G_j$ un sous-groupe additif de $\mathbb{F}_{q_j}$ tel que $N_j = \Theta(q_j)$ et $N_j \to \infty$, alors pour presque tout $g \in \Omega_\infty$,*

$$\lim_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{N_j \ln q_j}} = \sqrt{2}.$$

Par exemple, pour $G_j = \mathbb{F}_{q_j/2^a}$ ($a$ constant), on a bien $N_j = \Theta(q_j)$ et $\sigma_j = 2^a$ est une constante.

Dans d'autres cas où les conditions de la proposition 3.10 ne sont pas vérifiées, on peut tout de même essayer d'en dire plus que le théorème 3.3. Pour illustrer ce point, supposons $G_j = \mathbb{F}_{2^{2m_j/3}}$ (sous l'hypothèse que $m_j$ est un multiple de 3 et $m_j \to \infty$), cas où $N_j = o(h_j)$

(3ème cas de la section 3.6.1.2), alors en choisissant $1 \geq \gamma > 1/2$ dans (3.10), on assure la convergence de la série $\sum \sigma_j/N_j^\gamma$, puisque ici

$$\frac{\sigma_j}{N_j^\gamma} = \frac{2^{m_j/3}}{2^{2\gamma m_j/3}} = \frac{1}{2^{(2\gamma-1)m_j/3}},$$

où $2\gamma - 1 > 0$. On peut alors montrer que la série $\sum \mathbb{P}(\eta_j = 0)$ converge. Ainsi, en combinant les deux bornes, on obtiendra les inégalités suivantes.

**Proposition 3.12** *Soient $(m_j)$ une suite de multiples de 3, divergente vers l'infini, et pour tout $j$, $G_j = \mathbb{F}_{2^{2m_j/3}}$ sous-groupe de $\mathbb{F}_{q_j}$ ($q_j = 2^{m_j}$). Alors pour presque tout $g \in \Omega_\infty$*

$$\sqrt{\frac{4}{3}} \leq \underline{\lim}_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{N_j \ln q_j}} \leq \overline{\lim}_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{N_j \ln q_j}} \leq \sqrt{2}.$$

Même si la précision est moindre dans ce cas, on aboutit à une estimation relativement satisfaisante par rapport au théorème 3.3.

Compte tenu des comportements asymptotiques proches des bornes inférieures et supérieures obtenues, une conséquence à souligner est que la probabilité de tirer une fonction de faible amplitude spectrale (même restreinte à une fonction définie sur $G$) est faible. Cependant dans certaines circonstances, la borne inférieure obtenue peut être éloignée de la borne supérieure (voire ne pas être connue), comme par exemple dans une "faible" mesure pour le dernier cas ci-dessus ; on peut alors se demander, même si cela paraît peu probable, si la borne supérieure est améliorable.

#### 3.6.2.3 Le cas particulier des sous-groupes projectifs

Nous nous attardons finalement sur le cas où $G_j$ est un sous-groupe multiplicatif projectif de $\mathbb{F}_{q_j}^\times$ avec $H_j = \mathbb{F}_{q_j}$, où donc seuls les théorèmes 3.2 et 3.4 sont applicables.

Le choix de ces sous-groupes est guidé par la construction de fonction de type Patterson-Wiedemann (cf. Chapitre 2) de non-linéarité élevée. En effet, nous avons montré (e.g. Section 2.3) l'intérêt de tels sous-groupes pour contrôler les sommes de Gauss sur les caractères orthogonaux à ces sous-groupes. Le passage à la limite permet de faire converger les sommes de Gauss vers la même valeur ; les calculs s'en trouvant simplifiés, la contrainte porte alors sur le choix des valeurs sur le sous-groupe lui-même (cf. Remarque 2.6).

On a $h_j = q_j$, et on suppose $q_j \to \infty$, donc la série $\sum \frac{1}{h_j^\alpha}$ converge pour tout $\alpha > 0$, d'où via le théorème 3.2 : Pour presque tout $g \in \Omega_\infty$,

$$\overline{\lim}_{j\to\infty} \frac{\|\widetilde{g_{|G_j}}\|_{\widehat{\mathbb{F}_{q_j}}}}{\sqrt{N_j \ln q_j}} \leq \sqrt{2}.$$

On aimerait en déduire, conformément au chapitre 2, que $\|\widetilde{g_{|G_j}}\|_{\widehat{\mathbb{F}_{q_j}}} = o(\sqrt{q_j})$.

**Cas 1 : $(N_j)$ stationnaire.** Soit $t \in \mathbb{N}$ tel que $t | m_j$ pour tout $j \geq j_0$ et $G_j$ défini par

$$G_j = \mathbb{F}_{2^t}^{\times} \simeq \mathbb{F}_{q_j}^{\times} / (\mathbb{F}_{q_j}^{\times} / \mathbb{F}_{2^t}^{\times})$$

pour $j \geq j_0$. On en déduit ainsi que pour presque tout $g \in \Omega$,

$$\overline{\lim_{j \to \infty}} \frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{(2^t - 1) \ln q_j}} \leq \sqrt{2}$$

soit pour $j$ suffisamment grand et pour presque tout $g \in \Omega$,

$$\frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{q_j}} \leq \sqrt{\frac{2^{t+1} \ln q_j}{q_j}} = o(1). \tag{3.11}$$

Cependant, cette estimation n'est pas pertinente puisque qu'on sait déjà par une majoration très grossière que $\|\widetilde{g_{|G_j}}\|_\infty \leq N_j = 2^t - 1$.

**Cas 2 : $(N_j)$ minorée par une suite de puissances de 2 divergente.** Partons d'une suite de sous-groupes projectifs non constante. Soit $(t_j)$ une suite d'entiers tels que $t_j | m_j$ pour tout $j$ et telle que $(m_j - t_j)_j$ diverge vers l'infini. On choisit alors $G_j$ le sous-groupe projectif $\mathbb{F}_{q_j}^{\times} / \mathbb{F}_{2^{t_j}}^{\times}$, de sorte que $N_j \sim_{j \to \infty} 2^{m_j - t_j}$.

On a ainsi pour $j$ suffisamment grand et pour presque tout $g \in \Omega$,

$$\frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{q_j - 1}} \leq \sqrt{\frac{2 \ln q_j}{2^{t_j} - 1}}. \tag{3.12}$$

En particulier, on a presque sûrement

$$\frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{q_j}} = O(\sqrt{\frac{\ln q_j}{2^{t_j}}}),$$

et il est alors facile de choisir des sous-groupes tels que le deuxième terme tende vers 0. Ci-dessous on présente un exemple de construction :

Soient des entiers $(\alpha_j, r_j)_{j \in \mathbb{N}^\times}$ tels que $m_j = \alpha_j . r_j$, $t_j = \alpha_j$, et soient les entiers $N_j = (2^{\alpha_j r_j} - 1)/(2^{\alpha_j} - 1)$ vérifiant

$$N_j \mid N_{j+1}, \tag{3.13}$$

pour tout $j$ et tels que

$$\alpha_j (r_j - 1) \to \infty. \tag{3.14}$$

Soit une suite de groupes $(G_j)_{j \in \mathbb{N}}$ définie par $G_j = \mathbb{F}_{2^{\alpha_j r_j}}^{\times} / \mathbb{F}_{2^{\alpha_j}}^{\times}$. Le choix de la condition (3.13) implique que la suite $(G_j)$ est bien une suite croissante de sous-ensembles de $\mathbb{F}_{q_j}$ où $q_j = 2^{\alpha_j r_j}$ pour $j \geq 0$. Plus précisément :

**Lemme 3.12** *Pour tout $j \in \mathbb{N}$, $G_j$ est cyclique d'ordre $N_j$ et est un sous-groupe distingué de $G_{j+1}$.*

*Démonstration.* Soit $j \in \mathbb{N}$, $G_j$ est un sous-groupe de $\mathbb{F}^{\times}_{2^{\alpha_j r_j}}$, donc cyclique. Il est d'ordre $N_j$ par construction. De plus, $\mathbb{F}_{2^{\alpha_j \alpha_{j+1} r_j r_{j+1}}}$ étant une extension de corps de $\mathbb{F}_{2^{\alpha_j r_j}}$ et de $\mathbb{F}_{2^{\alpha_{j+1} r_{j+1}}}$, $G_j$ et $G_{j+1}$ sont deux sous-groupes cycliques de $\mathbb{F}^{\times}_{2^{\alpha_j \alpha_{j+1} r_j r_{j+1}}}$ tels que l'ordre de $G_j$ divise l'ordre de $G_{j+1}$, d'où $G_j \lhd G_{j+1}$. □

D'autre part, le résultat précédent implique que pour presque tout $g \in \Omega$, et pour $j$ suffisamment grand,

$$\frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{q_j}} = O(\sqrt{\frac{\alpha_j r_j}{2^{\alpha_j}}}). \tag{3.15}$$

## 3.7 Application aux constructions de type Patterson-Wiedemann

Concernant le cadre étudié dans la section 2.3 pour la construction des fonctions de type Patterson-Wiedemann, nous avons immédiatement :

**Proposition 3.13** *Soit $m \in \mathbb{N}$ et $(r_j)$ une suite d'entiers impairs divergente vers l'infini, telle que $r_j | r_{j+1}$ pour $j \geq 0$. On pose $\alpha_j = m$ pour tout $j$. Pour presque tout $g \in \Omega$, et pour $j$ suffisamment grand,*

$$\frac{\|\widetilde{g_{|G_j}}\|_\infty}{\sqrt{2^{mr_j}}} = O(\sqrt{\frac{mr_j}{2^m}}) = O(\sqrt{r_j}).$$

Observons maintenant la conséquence sur l'estimation (2.8) obtenue au chapitre 2 : soit une suite $(f_r)$ de fonctions $(2^m-1, s_r, h_r)$–PW généralisée suivant la suite de sous-groupes projectifs $(G_r)$, alors pour tout $\eta > 0$, il existe une constante $C > 0$ et $r \in \mathbb{N}^*, r \leq \eta^{-(2^m-2)}$ tels que pour $a \in \mathbb{F}^{\times}_{2^{mr}} - G_r$,

$$\begin{aligned} |\frac{\hat{f}_r(a)}{\sqrt{2^{mr}}} - s_r(\omega_a)| &\leq \sqrt{2(1-\cos\eta 2\pi)}\frac{(2^m-2)^2}{2^m-1} + \frac{1}{\sqrt{2^{mr}}} + C\sqrt{\frac{mr}{2^m}} \\ &\leq \sqrt{2(1-\cos\eta 2\pi)}(2^m-2) + \frac{1}{\sqrt{2^m}} + C\sqrt{\frac{m}{2^m(\eta)^{(2^m-2)}}}, \end{aligned}$$

où on utilise ici la variante déduite de l'application du résultat de Dirichlet (cf. Remarque 2.3) pour connaître une majoration de $r$. Pour que l'expression de droite soit petite, il faut choisir $\eta$ très petit dans le premier terme de l'addition, et $m$ grand dans le deuxième terme. Malheureusement, dans ce cas, le troisième terme explose. Réciproquement si on réduit le troisième terme alors le premier ou le deuxième est gros.

En pratique, contrairement à ce qu'on aimerait atteindre dans le chapitre 2, on ne peut donc pas se servir directement de ce résultat pour obtenir une fonction d'amplitude très petite par rapport à $\sqrt{q_j}$. Les deux estimations asymptotiques ("convergence" des sommes de Gauss et non-linéarité sur les sous-groupes) sont trop contradictoires et il faudra en améliorer au

moins une (soit de manière générale si c'est possible, soit en se restreignant à une construction spécifique...).

Comme les théorèmes 3.3 et 3.5 ne s'appliquent pas ici, on aimerait pouvoir répondre au problème suivant afin de connaître la marge par rapport à la borne supérieure.

**Problème 3.1** *Que vaut la borne inférieure asymptotique pour G un sous-groupe de $L^\times$ ?*

Même si le résultat précédent ne permet pas de s'approcher de la borne de Parseval, on en déduit l'existence de familles de fonctions de type Patterson-Wiedemann généralisé qui ont asymptotiquement presque toutes une non-linéarité nettement supérieure à la non-linéarité asymptotique moyenne. Soit $m$ un entier tel que $2^m - 1$ soit un nombre de Mersenne premier et tel que les sommes de Gauss définies sur $K^\times = \mathbb{F}_{2^m}^\times$ vérifient le théorème d'indépendance 2.2. Le résultat est alors le suivant.

**Théorème 3.6** *Pour tout $r \in \mathbb{N}^*$, on considère $G_r^{(m)}$ le sous-groupe d'indice $2^m - 1$ de $\mathbb{F}_{2^{mr}}^\times$, $\Omega_r^{(m)}$ l'ensemble des fonctions binaires définies sur $G_r^{(m)}$ et $\Omega_\infty^{(m)}$ la limite projective de ces ensembles, quand r tend vers l'infini.*

*Pour tout $\epsilon > 0$, il existe $r > 0$ tel que pour presque tout $h \in \Omega_\infty^{(m)}$, si $f$ est une fonction de type Patterson-Wiedemann généralisé suivant le sous-groupe $G_r^{(m)}$ de représentant $h_{|G_r^{(m)}}$ sur $G_r^{(m)}$ et de représentant équilibré sur les cosets non-triviaux de $G_r^{(m)}$. Alors l'amplitude spectrale de $f$ vérifie*

$$\|\hat{f}\|_\infty \leq \left(\epsilon + \sqrt{\frac{2}{2^m - 1}}\right) \sqrt{2^{mr} \ln 2^{mr}}. \tag{3.16}$$

*Démonstration.* D'après le théorème 3.2 (cf. aussi équation (3.12)), pour tout $m \in \mathbb{N}^*$, il existe $r_0 > 0$ tel que pour tout $r \geq r_0$, on ait

$$\|\widetilde{h_{|G_r^{(m)}}}\|_\infty \leq \sqrt{2} \sqrt{\frac{2^{mr} - 1}{2^m - 1} \ln 2^{mr}}$$

pour presque tout $h \in \Omega_\infty^{(m)}$.

D'après la proposition 2.4, pour tout $\epsilon' > 0$, il existe $r \geq r_0$ tel que si $f_r$ est une fonction $(2^m - 1, s_r, h_r)$–PW généralisée de représentant $s_r$ équilibrée sur les cosets non-triviaux de $G_r^{(m)}$ et de représentant $h_r$ sur $G_r^{(m)}$ alors, en notant $\widetilde{h_r}_{|G_r^{(m)}}(a)$ la transformée de Fourier en $a$ de $h_r$ sur $G_r^{(m)}$ :

- Si $a \in \mathbb{F}_{2^{mr}}^\times - G_r^{(m)}$, $\left|\frac{\hat{f}_r(a)}{\sqrt{2^{mr}}}\right| \leq 1 + \epsilon'(2^m - 2) + \frac{1}{\sqrt{2^{mr}}} + \frac{|\widetilde{h_r}_{|G_r^{(m)}}(a)|}{\sqrt{2^{mr}}}$.
- Si $a \in G_r^{(m)}$, $\left|\frac{\hat{f}_r(a)}{\sqrt{2^{mr}}}\right| \leq \epsilon'(2^m - 2) + \frac{1}{\sqrt{2^{mr}}} + \frac{|\widetilde{h_r}_{|G_r^{(m)}}(a)|}{\sqrt{2^{mr}}}$.
- Si $a = 0$, $|\hat{f}_r(0)| \leq 1 + |\#G_r^{(m)} - 2\mathrm{w}(h_r)| = 1 + |\widetilde{h_r}_{|G_r^{(m)}}(0)|$.

Ainsi, à partir de $h \in \Omega_\infty^{(m)}$ si on pose $h_r = h_{|G_r^{(m)}}$ alors pour presque tout $h$, on a pour tout $a \in \mathbb{F}_{2^{mr}}$,

$$\left|\frac{\hat{f}_r(a)}{\sqrt{2^{mr} \ln 2^{mr}}}\right| \leq \frac{1}{\sqrt{\ln 2^{mr}}} \left(1 + \epsilon'(2^m - 2) + \frac{1}{\sqrt{2^{mr}}}\right) + \sqrt{\frac{2}{2^m - 1}}.$$

Pour conclure, il suffit de fixer $\epsilon'$ pour que le terme de droite soit inférieur à $\epsilon$. □

Par exemple, pour $m = 5$ qui vérifie bien la condition d'indépendance des sommes de Gauss du théorème 2.2, on obtient un facteur de l'ordre de $0.25\sqrt{2^{mr}\ln 2^{mr}}$ ; pour $m = 13$, on arrive à un facteur d'environ $0.0157\sqrt{2^{mr}\ln 2^{mr}}$, ce qui est nettement éloigné de la moyenne asymptotique $\sqrt{2}\sqrt{2^{mr}\ln 2^{mr}}$ et nous incite à poursuivre l'étude de ces constructions.

Il est difficile de connaître la valeur maximale (s'il y en a une) possible pour $m$ dans l'inégalité (3.16), compte tenu de la difficulté de vérifier les hypothèses du théorème 2.2 puisque la taille des nombres premiers de Mersenne croît très rapidement. Si on admettait qu'il n'y a pas de valeur maximale, i.e. qu'il existe un nombre infini de $m$ valides[3], alors on pourrait remplacer, dans l'inégalité (3.16), $\epsilon + \sqrt{\frac{2}{2^m-1}}$ par $\epsilon$. Cette hypothèse est cependant très difficile à vérifier.

## 3.8 Conclusion

Nous avons étudié le comportement asymptotique de l'amplitude spectrale partielle (par rapport à un sous-ensemble) des fonction binaires définies sur un sous-ensemble d'un corps fini. Il ressort de cette étude que le comportement peut être sensiblement le même que dans le cas des fonctions booléennes ; des bornes supérieures et inférieures analogues aux bornes connues sont décrites. Cependant, même si les bornes supérieures permettent d'affirmer qu'asymptotiquement presque toutes les fonctions ont une amplitude spectrale relativement faible, les bornes inférieures obtenues ne sont valables que pour des sous-ensembles particuliers. Le problème de connaître la borne inférieure pour un sous-ensemble quelconque pourrait ainsi faire l'objet d'une autre étude.

Il est probable que les même bornes demeurent valides, mais la démonstration sans l'utilisation des relations d'orthogonalités paraît difficile.

D'autre part, ces résultats ne présagent en rien de l'impossibilité de trouver des instances avec des bonnes propriétés. Comme introduit dans [Rod05] via l'étude de la norme $L^4$ de la transformée de Fourier sur le corps tout entier, si on souhaite aller plus loin dans les estimations asymptotiques, on pourra chercher à appliquer d'autres théorèmes de probabilités sur la convergence de variables aléatoires (e.g. loi des grands nombres, grandes déviations). Il s'agit en effet d'essayer de mieux comprendre le comportement des fonctions qui ne se situent pas dans la moyenne.

En particulier, nous avons montré via le théorème 3.6 comment l'étude de la non-linéarité partielle permet de prouver l'existence de familles non-triviales de fonctions de type Patterson-Wiedemann généralisé de non-linéarité asymptotique nettement supérieure à la moyenne.

Finalement, on peut se demander s'il est possible d'obtenir des résultats asymptotiques pour d'autres propriétés que la non-linéarité, et sur des sous-ensembles de fonctions particu-

[3]Cette hypothèse est en particulier dépendante de l'existence d'un nombre infini de nombres de Mersenne premiers.

liers. Comme nous l'avons vu au chapitre 1, de nombreux autres critères sont étudiés dans le cadre de l'utilisation des fonctions booléennes en cryptographie et la connaissance de leur comportement pour $m$ grand est à approfondir.

Comme pour la non-linéarité, on peut également imaginer étudier ces propriétés pour des fonctions définies sur un ensemble de départ quelconque, par exemple pour en déduire d'autres propriétés des fonctions de type Patterson-Wiedemann généralisé.

# Chapitre 4

# Non-linéarité partielle sur un sous-groupe multiplicatif

Après l'analyse asymptotique, nous essayons dans ce chapitre d'obtenir des informations supplémentaires sur la non-linéarité partielle sur un sous-groupe multiplicatif. En particulier, nous cherchons à savoir ce qui est envisageable en dimension finie. Nous examinons également comment satisfaire les contraintes imposées par les constructions de type Patterson-Wiedemann généralisé. En effet, une des problématiques majeures de cette méthode est l'estimation de sommes d'exponentielles sur des sous-groupes : plusieurs approches sont envisagées, dont la détermination d'une borne universelle ou des majorations en moyenne. Nous obtenons notamment une borne supérieure équivalente à celle établie dans le chapitre précédent en suivant une technique employée par [Car03] pour la non-linéarité classique.

## 4.1 Approche du problème

### 4.1.1 Introduction

Nous étudions la non-linéarité partielle sur un sous-groupe $G$ d'ordre $N$ et d'indice $v$ de $L^\times$. Soit $h$ une fonction binaire définie sur $G$, suivant la définition 3.1 la transformée de Fourier en $a \in L$ de $h$ sur $G$ est la somme de caractères

$$\tilde{h}_G(a) = \sum_{x \in G} h(x)\mu(ax).$$

On souhaite alors savoir quelle est la valeur maximale atteinte, i.e. l'amplitude spectrale de $h$ sur $G$ :

$$\|\tilde{h}_G\|_\infty = \max_{a \in L} |\tilde{h}_G(a)|.$$

Plus généralement, on veut connaître la valeur du *rayon spectral d'indice v*, soit le rayon spectral de $G$ par rapport à $L$

$$R_v(m) = R_G(m) = \min_{h:G \to \{\pm 1\}} \|\tilde{h}_G\|_\infty,$$

en fonction de son indice $v$. Comme expliqué dans la section 3.2.2, cette notion revient à rechercher la plus grande non-linéarité possible sur $G$. De plus, sous les conditions décrites à la

fin du chapitre 2, sa valeur a une grande influence sur la non-linéarité sur $L$ d'une fonction de type PW généralisé.

Dans son principe, l'étude de ces différentes valeurs est un cas particulier d'un problème posé par Langevin sur la question de l'existence d'une fonction à support fini définie sur un espace de dimension infinie pour laquelle l'amplitude spectrale est de l'ordre de la racine carrée de la taille du support (cf. [Lan99, Problème 3.1, page 94]). Ici, la question est donc de savoir si la conjecture de Patterson-Wiedemann (cf. Conjecture 1.1) peut être transposée en un résultat réel sur $G$.

**Problème 4.1** *Soit $q$ une puissance de 2 et $G$ sous-groupe de $\mathbb{F}_q^\times$ d'indice $v$, quelles sont les conditions sur $q$ et $v$ pour qu'il existe une fonction binaire $h$ définie sur $G$ telle que*

$$\|\tilde{h}_G\|_\infty \sim \sqrt{\#G} \ ?$$

L'égalité de Parseval, qui s'applique à toute transformée de Fourier, s'écrit ici :

**Lemme 4.1** *Soit $h : G \to \{\pm 1\}$,*

$$\sum_{a \in L} \tilde{h}_G(a)^2 = q \times N.$$

On en déduit l'inégalité de Parseval habituelle, à savoir la borne inférieure

$$R_v(m) \geq \sqrt{N} = \sqrt{\frac{2^m - 1}{v}}. \tag{4.1}$$

Il s'agit alors de savoir à quelle distance de cette borne inférieure se situe $R_v(m)$.

### 4.1.2 Contexte des constructions de type PW

Il faudrait également réussir à comparer intelligemment $R_v(m)$ et $\sqrt{2^m}$. En effet, de manière plus spécifique à la construction de fonctions de type PW généralisé, on aimerait pouvoir se prononcer sur les choix possibles de $G$ et $h$ afin d'obtenir, suivant la remarque 2.6, $\|\tilde{h}_G\|_{G\cup\{0\}} \leq \sqrt{q}$ et $\|\tilde{h}_G\|_{L^\times - G}$ petit par rapport à $\sqrt{q}$ (où $\|\phi\|_E$ désigne la valeur maximale de $|\phi(u)|$ pour $u$ parcourant $E$).

**Remarque 4.1** *Pour illustrer, si on se place dans le cas idéal (mais non parfaitement réalisable pour m impair), comme pour la remarque 2.5 (i.e. $\tau_L(\chi) = \sqrt{q} = \sqrt{2^m}$ pour $\chi \neq 1$), alors pour battre la borne quadratique (respectivement battre la borne de Patterson-Wiedemann) à l'aide d'une fonction de type PW généralisé, il faudrait exhiber une fonction $h : G \to \{\pm 1\}$ telle que*

1. *$|\tilde{h}_G(a)| \leq \kappa \sqrt{2}\sqrt{q} - 1$, pour $a \in G \cup \{0\}$,*
2. *$|\tilde{h}_G(a)| \leq (\kappa\sqrt{2} - 1)\sqrt{q} - 1$, pour $a \in L^\times - G$,*

*où $\kappa$ vaut 1 pour la borne quadratique (resp. $\frac{27}{32}$ pour la borne de Patterson-Wiedemann). De plus pour avancer vers une démonstration de la conjecture de Patterson-Wiedemann, il faudrait que $\kappa$ soit aussi proche de $\frac{1}{\sqrt{2}}$ que possible.*

*À noter que comme $\sqrt{q} \simeq \sqrt{\frac{2^m-1}{v}} \times \sqrt{v}$, plus $v$ est grand plus on a de marge vis-à-vis de la borne de Parseval sur $G$.*

Malgré le fait que la détermination du rayon spectral sur $G$, problème bien plus général que dans le cas du corps tout entier, est certainement plus difficile, on remarque aussi que pour notre finalité (la construction de fonctions de type PW généralisé hautement non-linéaires), la marge de manoeuvre par rapport à la borne de Parseval est beaucoup plus importante. C'est donc la principale motivation de l'utilisation de telles constructions et la raison pour laquelle l'étude de $R_v(m)$ est importante.

En particulier, dans les paragraphes ci-dessous, nous essayons d'établir des comparaisons, non optimales, comme suit : Soit $m \in \mathbb{N}^\times$. Il existe $\lambda$ réel tel que pour tout sous-groupe $G \lhd L^\times$ d'ordre $N$, alors

$$R_v(m) \leq \lambda \sqrt{N}.$$

### 4.1.3 Sommes d'exponentielles et polynômes

Un première façon d'aborder ce problème est de le voir comme un problème de recherche de sommes d'exponentielles quelconques de faible amplitude. Cependant, il est en général difficile de connaître les meilleures constructions et/ou bornes possibles. On peut essayer d'utiliser les bornes classiques telles que les bornes de Hasse-Weil, de Deligne (ou d'autres généralisations cf. [Gil95]) mais elles ne sont pas satisfaisantes en général : par exemple, pour $P(X) \in L[X]$ alors $\widetilde{\mu(P)}_G(a) = \sum_{x\in G} \mu(P(x) + ax) = \frac{1}{v} \sum_{x\in L^\times} \mu(P(x^v) + ax^v)$ et si $P$ est de degré impair $s > 1$, la borne de Hasse-Weil implique

$$\|\widetilde{\mu(P)}_G\|_\infty \leq \frac{1}{v}(sv-1)\sqrt{2^m} + \frac{1}{v} \lesssim s\sqrt{2^m} \simeq s\sqrt{v}\sqrt{N}. \tag{4.2}$$

Ce qui comparé à l'inégalité de Parseval (4.1) paraît très mauvais, dès que $v$ ou $s$ est élevé.

Si on s'intéresse au cas des fonctions puissances, en suivant les travaux de [Gil95], on s'aperçoit que l'on peut améliorer légèrement cette borne (cf. [BGL05]). Nous avons en effet estimé l'amplitude spectrale d'indice $v$ de fonctions monômes du type $f : x \mapsto \gamma x^s$ pour $\gamma \in L$ et un entier $s$. Si $m$ n'est pas premier ($m = lt$), la stratégie consiste à évaluer la somme d'exponentielles sur $K = \mathbb{F}_q$ au lieu de $L$, où $[L : K] = l$ et $q = 2^t$. Nous cherchons alors des instances $(m, l, t, v, s)$ avec $v$ l'indice du sous-groupe $G$ et $s$ un exposant tel que $\|\widetilde{\mu(\gamma x^s)}_G\|_\infty$ est petit pour un choix judicieux de $\gamma \in L$. En pratique, nous déterminons les formes de $s$ et $vs$ pour appliquer les résultats de [Gil95].

Soit $w_q(e)$ la somme des composantes de l'expression $q$-aire d'un entier $e$. Supposons que $w_q(s) \neq w_q(sv)$, et notons $w = \max\{w_q(s), w_q(sv)\}$, et $d \in \{v, sv\}$ l'entier tel que $w = w_q(d)$. Si $d < q$ est impair ou bien si l'expression $q$-aire de $d$ est $d = 1 + kq^j$, alors pour tout entier $k$ pair et $j < (m/l)$, le Théorème 2.1 dans [Gil95] fournit l'estimation suivante

$$\|\widetilde{\mu(P)}_G\|_{L^\times} \leq \frac{1}{v}(w-1)^l\sqrt{2^m} + \frac{1}{v}. \tag{4.3}$$

Numériquement, on est capable de trouver de nombreuses instances $(m, l, t, v, s)$ vérifiant $(w-1)^l < (sv-1)$, i.e. telles que l'inégalité (4.3) soit meilleure que (4.2). Cependant, nous n'avons pas rencontré le cas $(w-1)^l < v$ qui aurait permis de se rapprocher de la borne de Parseval sur $G$. On peut toutefois obtenir un résultat encourageant, comparé au problème 4.1, pour les

sous-groupes d'indice 3 : la proposition suivante entraîne que

$$R_3(m) \lesssim \frac{4}{\sqrt{3}}\sqrt{\frac{2^m-1}{3}} \simeq 2.31\sqrt{\frac{2^m-1}{3}}$$

pour $m \equiv 2 \mod 4$.

**Proposition 4.1** *Soit $m = 2t$ où $t$ est impair, et $P : x \mapsto \gamma x^s$, avec $Tr_{L/K}(\gamma) \neq 0$. Alors le choix de paramètres $(2t, 2, t, 3, (q+1)/3)$ implique*

$$\|\widetilde{\mu(P)}_G\|_\infty \leq \frac{4}{3}\sqrt{2^m} + \frac{1}{3} \tag{4.4}$$

*Démonstration.* Soient $m = 2t$, $v = 3$, $vs = q + 1$. Nous voulons estimer

$$\widetilde{\mu(P)}_G(a) = \frac{1}{v}\sum_{x\in L^\times} \mu(\gamma x^{sv} + ax^v) = \frac{1}{v}\sum_{x\in L^\times} \mu(\gamma x^{q+1} + ax^3).$$

Si $a \neq 0$, $\max\{w_q(3), w_q(q+1)\} = 3$, alors l'inégalité (4.3) donne (4.4) sur $L^\times$. Si $a = 0$, on obtient par le calcul :

$$\widetilde{\mu(P)}_G(0) = \frac{1}{v}\sum_{x\in L^\times} \mu(\gamma x^{q+1}).$$

Soit $\mu_K$ le caractère additif canonique de $K$ et soit $x \in L^\times$,

$$\mu(\gamma x^{q+1}) = \mu_K(\mathrm{Tr}_{L/K}(\gamma x^{q+1})) = \mu_K(x^{q+1}\mathrm{Tr}_{L/K}(\gamma)).$$

On en déduit

$$\widetilde{\mu(P)}_G(0) = \frac{q+1}{v}\sum_{y\in K^\times} \mu_K(y\mathrm{Tr}_{L/K}(\gamma)) = -\frac{q+1}{v}$$

L'inégalité 4.4 est alors satisfaite puisque $|\widetilde{\mu(P)}_G(0)| = \frac{q+1}{3} \leq \frac{4}{3}q + \frac{1}{3}$. □

## 4.2 Estimation en moyenne

Nous utilisons maintenant une méthode plus classique d'étude de la non-linéarité, directement à partir de la définition (distance aux fonctions affines). En moyenne, il est connu que la majorité des fonctions booléennes ont une non-linéarité relativement élevée, et donc qu'elles ont une amplitude spectrale relativement faible. Pour des fonctions binaires sur un sous-groupe $G$ de $L^\times$, nous montrons ci-dessous, en adaptant une idée de [Car03], que ce phénomène est toujours vrai, y compris à petite dimension. Une comparaison avec les résultats du chapitre 3 est également recherchée afin de savoir si ceux-ci sont applicables en dimension finie.

Tout d'abord, rappelons un encadrement connu de sommes de coefficients binomiaux (voir [MS77, page 310]).

**Lemme 4.2** *Soit $N$ un entier positif quelconque et $0 < \lambda < 1/2$. Alors*

$$\frac{2^{NH_2(\lambda)}}{\sqrt{8N\lambda(1-\lambda)}} \leq \sum_{0\leq i\leq \lambda N} \binom{N}{i} \leq 2^{NH_2(\lambda)} < 2^N e^{-2N(1/2-\lambda)^2}$$

*où $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ ($H_2$ est la fonction d'entropie).*

Ce lemme permet de démontrer le résultat suivant :

**Théorème 4.1** *Soit $m > 0$ un entier, $G$ un sous-groupe de $L^\times = \mathbb{F}_{2^m}^\times$ et $N$ l'ordre de $G$ (respectivement $v$ son indice). Soit $c > 0$ un réel strictement positif quelconque tel que $N > 2c^2 m$. Alors, la densité de l'ensemble $\{h : G \to \{\pm 1\}, \|\tilde{h}_G\|_\infty \leq c\sqrt{2Nm}\}$ par rapport à l'ensemble des fonctions binaires définies sur $G$ est supérieure à $1 - 2^{m(1-c^2\log_2(e))}$.*

*De plus, si $c^2\log_2(e) > 1$, alors cette densité tend vers 1 pour $m$ tendant vers l'infini.*

*Démonstration.* Soit $l : L \to \mathbb{F}_2$ une application linéaire et $l_{|G}$ sa restriction à $G$, alors le nombre de fonctions binaires définies sur $G$, $h : G \to \{\pm 1\}$, telle que la distance entre $h$ et $\mu(l_{|G})$ sur $G$ est plus petite que $N/2 - c\sqrt{m}\sqrt{N/2}$, est :

$$A = \sum_{0\leq i\leq N/2 - c\sqrt{m}\sqrt{N/2}} \binom{N}{i}.$$

Grâce au lemme 4.2, on en déduit que $A \leq 2^N e^{-2N(1/2-\lambda)^2}$, où $0 < \lambda = 1/2 - c\sqrt{m}/\sqrt{2N} < 1/2$. D'où,

$$A \leq 2^{N - mc^2\log_2(e)}.$$

Ainsi, le nombre de fonctions $h$ à une distance, sur $G$, inférieure à $N/2 - c\sqrt{m}\sqrt{N/2}$ d'au moins une fonction linéaire quelconque est au plus

$$2^m A \leq 2^{m+N-mc^2\log_2(e)}.$$

Comme $\tilde{h}_G(a) = N - 2d(h, x \mapsto \mu(ax))$, on obtient que la densité de l'ensemble défini dans l'énoncé du théorème parmi l'ensemble de toutes les fonctions binaires sur $G$ est supérieure à

$$1 - 2^{m(1-c^2\log_2(e))}.$$

De plus, si $c^2\log_2(e) > 1$ et si on dispose d'une suite de sous-groupes $(G_m)_m$, telle que pour tout $m$, $G_m$ est un sous-groupe d'ordre $N_m > 2c^2 m$ de $\mathbb{F}_{2^m}^\times$, alors la densité des fonctions binaires définies sur $G_m$ telles que $\|\tilde{h}_{G_m}\|_\infty \leq c\sqrt{2N_m m}$ tend vers 1 quand $m$ croît vers l'infini. □

A noter que si la condition $N > 2c^2 m$ n'est pas réalisée, alors le résultat reste trivialement vrai (mais sans intérêt particulier), puisque dans ce cas $c\sqrt{2Nm} \geq N$.

Ce résultat complète les résultats du chapitre 3 qui établissent des bornes inférieures et supérieures asymptotiques valables pour presque toutes les fonctions (i.e. sauf un sous-ensemble

de mesure négligeable) et pour $G$ un sous-ensemble quelconque. Par rapport à la proposition 3.9, on s'aperçoit que cela correspond au cas limite où on choisit $c^2 \log_2 e$ très proche de 1, et pour lequel la densité de l'ensemble des fonctions d'amplitude inférieure à

$$\sqrt{2c^2 \log_2 e \times N_j \ln h_j}$$

tend vers 1. La proposition 3.9 peut donc être vue comme la version asymptotique et optimale du théorème 4.1. Ici, on peut en déduire différentes applications concrètes.

Pour appliquer le théorème et obtenir une densité non nulle, il faut avoir au minimum $m < mc^2 \log_2 e$ ; or par hypothèse $mc^2 < N/2$, d'où $m < mc^2 \log_2 e < N/2 \times \log_2 e$. Si la condition $m < N/2 \times \log_2 e$ est vérifiée, alors en prenant $c^2 \log_2 e$ très proche de 1, on obtient une des meilleures bornes envisageable via ce théorème, à savoir :

**Corollaire 4.1** *Si $m < N/2 \times \log_2 e$, alors il existe une fonction $h$ sur $G$ telle que*

$$\|\tilde{h}_G\|_\infty \leq \sqrt{2 \ln 2 \times mN} = \sqrt{2N \ln q} < 1.18 \sqrt{mN}.$$

Sous les conditions de la remarque 4.1, pour battre la borne quadratique, il faut $\|\tilde{h}_G\|_\infty \lesssim 0.41 \sqrt{q}$, i.e. qu'ici il faudrait $N \lesssim 0.12q/m$. Une autre application du théorème donnera :

**Corollaire 4.2** *Pour tout $m \geq 3$ et $G \lhd L^\times$ d'ordre $N > 2m$, une majorité de fonctions binaires $h$ définies sur $G$ vérifient la relation*

$$\|\tilde{h}_G\|_\infty \leq \sqrt{2Nm}.$$

*On en déduit que*

$$\sqrt{N} = \sqrt{\frac{2^m - 1}{v}} \leq R_v(m) \leq \sqrt{2m} \sqrt{\frac{2^m - 1}{v}} = \sqrt{2m}\sqrt{N},$$

*et la majorité des fonctions binaires sur $G$ se situe entre ces deux bornes.*

*Démonstration.* Il suffit de remarquer que $2^{m(1-c^2 \log_2(e))} < \frac{1}{2}$ si $m \geq 3$ et $c = 1$. □

En particulier si l'ordre du sous-groupe vérifie $N = o(2^m/m)$, la majorité des fonctions binaires $h : G \to \{\pm 1\}$ satisfont $\|\tilde{h}_G\|_\infty = o(\sqrt{2^m})$. Cette relation correspond à un des critères identifiés dans la section 2.4.2 pour construire des fonctions de type PW généralisé hautement non-linéaires. Cependant cette contrainte sur l'ordre de $G$ n'est pas vérifiée par les sous-groupes projectifs utilisés pour obtenir des sommes de Gauss proches de $\sqrt{q}$ par passage à la limite (on ne peut donc pas appliquer la proposition 2.4 simultanément).

On aimerait trouver des estimations plus précises pour étudier d'autres conditions sur $G$ permettant d'obtenir $R_v(m) = o(\sqrt{2^m})$. Suivant la remarque 4.1, il serait également intéressant de se focaliser sur une minimisation de l'amplitude spectrale d'une fonction $h$ en dehors de $G$.

**Remarque 4.2** *Le théorème 4.1 se généralise à la comparaison avec un sous-ensemble quelconque de fonctions linéaires de cardinal l en remplaçant toute occurrence de m dans l'énoncé par* $\log_2 l$*. On en déduit par exemple :*

– *si* $c, N$ *sont tels que*

$$N > 2c^2 \log_2(N+1) > 2\frac{\log_2(N+1)+1}{\log_2 e}$$

*alors il existe une majorité stricte de fonctions h définies sur G telles que*

$$\|\tilde{h}_G\|_{G\cup\{0\}} \leq c\sqrt{2N\log_2(N+1)},$$

– *si* $c, N$ *sont tels que*

$$N > 2c^2 \log_2(q-1-N) > 2\frac{\log_2(q-1-N)+1}{\log_2 e}$$

*alors il existe une majorité stricte de fonctions h définies sur G telles que*

$$\|\tilde{h}_G\|_{L^\times - G} \leq c\sqrt{2N\log_2(q-1-N)}.$$

*Cela ne permet malheureusement pas d'améliorer substantiellement l'apport des corollaires précédents dans le contexte de la remarque 4.1. En effet la contrainte est plus forte pour la borne sur* $L^\times - G$ *dans la remarque 4.1, mais pour N petit par rapport à q, elle reste très proche du résultat du théorème 4.1 et seule la borne sur* $G \cup \{0\}$ *est nettement améliorée.*

La borne supérieure trouvée dans les divers cas étudiés section 3.6 correspond au cas le plus favorable possible du théorème 4.1 (i.e. $c^2 \log_2 e$ proche de 1 à valeurs supérieures) mais compte tenu des comportements proches des bornes inférieures et supérieures asymptotiques obtenues pour la plupart des cas, il est difficile d'espérer mieux que la borne du théorème 4.1.

Néanmoins un point positif à souligner est le fait que même sans assurer ni une densité proche de 1 ni une majorité, le corollaire 4.1 assure l'existence d'une fonction vérifiant une condition très proche de la borne asymptotique établie section 3.6.2.2 tout en étant directement applicable en dimension finie. Ainsi, son utilisation dans le cadre d'estimation d'amplitudes (comme par exemple celles des constructions du chapitre 2) est entièrement justifiée.

On peut finalement se demander s'il existe un résultat indépendant de $G$ (ou valable pour une famille de sous-groupes et de $m$) :

**Problème 4.2** *Existe t-il un réel* $\lambda$ *tel que pour tout m et tout indice v alors* $R_v(m) \leq \lambda\sqrt{\frac{2^m-1}{v}}$ *?*

## 4.3 Non-linéarité sur le groupe multiplicatif des inversibles d'un corps fini

Lorsqu'on travaille sur le corps tout entier, on sait que l'utilisation d'une forme quadratique permet de répondre aisément au problème ci-dessus, par l'obtention de la borne dite quadratique.

Ici on suppose que $m$ n'est pas un nombre premier, soit un sous-corps $K = \mathbb{F}_{2^k}$ de $L = \mathbb{F}_{2^m}$ et on considère $G$ le sous-groupe de $L^\times$ défini par $G = K^\times$, i.e d'ordre $N = 2^k - 1$ et d'indice $v = \frac{2^m-1}{2^k-1}$. Alors on peut utiliser des fonctions quadratiques pour déterminer une meilleure borne que dans le théorème précédent.

Comme indiqué dans la section 1.2.3.1, on sait qu'il existe une forme quadratique définie sur $K$ telle que $\|\widehat{\mu(q)}\|_K = 2^{\lfloor \frac{k+1}{2} \rfloor}$ (et il est facile de vérifier que cette assertion reste vraie si on regarde sa transformée pour $a \in L$). Or sa transformée de Fourier sur $G$ pour $a \in L$ est fortement liée à sa transformée sur $K$ :

$$\widetilde{\mu(q)}_G(a) = \widehat{\mu(q)}(a) - \mu(q(0)).$$

Ce qui implique que

$$\|\widetilde{\mu(q)}_G\|_\infty = 2^{\lfloor \frac{k+1}{2} \rfloor} + 1.$$

On obtient dans ce cas un résultat bien plus favorable que ceux de la section précédente, à savoir l'équivalent de la borne quadratique sur $G$ :

$$R_v(m) \leq \mathrm{bq}_G(m) = 2^{\lfloor \frac{k+1}{2} \rfloor} + 1,$$

où ici $k = \log_2(N+1)$. Dans le cas pair, cela donne $R_v(m) \leq \sqrt{N+1} + 1$ ; ou $R_v(m) \leq \sqrt{2(N+1)} + 1$ dans le cas $k$ impair (plus précisément on a $R_v(m) \leq R(k) + 1$).

Pour les fonctions de type PW généralisé, il serait intéressant de choisir un tel sous-groupe $G$ puisque si $N$ est petit devant $2^m$, le rayon spectral sur $G$ sera automatiquement petit devant $\sqrt{2^m}$ (ce que l'on cherche à obtenir, cf. Remarque 4.1). Cependant dans les exemples étudiés dans le chapitre 5 (choisis de façon à pouvoir manipuler ou estimer facilement les sommes de Gauss sur $G^\perp$), un tel choix ne sera pas possible.

**Cas des constructions de type PW généralisé.** Si $f$ est une fonction $(v, s, h)$–PW généralisée suivant un sous-groupe $G$ de $L^\times$, de représentants $s$ équilibré sur $\Omega^* = L^\times/G - \{1\}$ et $h$ sur $G$, alors pour $a \in L^\times$ :

$$\hat{f}(a) = 1 + \frac{1}{v} \sum_{\chi \perp G, \chi \neq 1} \tau_L(\chi) \hat{s}(\overline{\chi}) \overline{\chi}(a) + \sum_{x \in G} h(x) \mu(ax).$$

Dans le cas où $G = K^\times$, on en déduit immédiatement le résultat suivant :

**Corollaire 4.3** *Soit $K$ un sous-corps de $L$ et $f$ de type PW généralisé sur $L$ suivant $K^\times$ de représentant $s$ équilibré sur $\Omega^*$ alors pour $a \in L$,*

$$\hat{f}_L(a) = \frac{1}{v} \sum_{\chi \perp K^\times, \chi \neq 1} \tau_L(\chi) \hat{s}(\overline{\chi}) \overline{\chi}(a) + \widehat{f_{|K}}_K(a). \tag{4.5}$$

Ce corollaire permet d'envisager des constructions par récurrence suivant une suite décroissante de sous-corps de $L$, l'équation ci-dessus impliquant en particulier :

**Corollaire 4.4**

$$R(m) \leq R(k) + \frac{2^k - 1}{2^m - 1} \min_{s:\Omega^* \to \{\pm 1\}} \max_{a \in L} \left| \sum_{\chi \perp \mathbb{F}_{2^k}^{\times}, \chi \neq 1} \tau_L(\chi) \hat{s}(\overline{\chi}) \overline{\chi}(a) \right| \tag{4.6}$$

*où* $L = \mathbb{F}_{2^m}$ *et* $K = \mathbb{F}_{2^k}$.

Nous en verrons une application un peu plus loin dans le cas des fonctions équilibrées.

Un résultat provenant d'une construction par récurrence pour un sous-groupe $G$ quelconque semble quant à lui beaucoup plus difficile à obtenir.

## 4.4 Perspectives

Un problème intéressant serait d'éclaircir le lien entre les différents rayons spectraux quand $v$ et $m$ varient. Si $(2^{m'} - 1)/v'$ divise $(2^m - 1)/v$ alors le sous-groupe de $\mathbb{F}_{2^{m'}}^{\times}$ d'indice $v'$ peut être vu comme un sous-groupe du sous-groupe de $\mathbb{F}_{2^m}^{\times}$ d'indice $v$ et il est facile de prouver que $R_v(m) \leq \frac{(2^m-1)/v}{(2^{m'}-1)/v'} R_{v'}(m')$.

Si par exemple $(2^m - 1)/v$ est divisible par $2^k - 1$ pour un certain $k$ impair, cela permet de relier $R_v(m)$ au cas de la section 4.3, puisqu'on a alors $R_v(m) \leq \frac{2^m-1}{v(2^k-1)} R_{v'}(m)$ où $v' = (2^m - 1)/(2^k - 1)$, d'où

$$R_v(m) \lesssim \sqrt{2\frac{2^m - 1}{v(2^k - 1)}} \sqrt{\frac{2^m - 1}{v}}.$$

Mais cette borne ne nous aide pas à déduire un résultat général concernant le problème 4.2.

**Problème 4.3** *Existe-t-il un lien pertinent entre* $R_v(m)$ *et* $R_{v'}(m')$*, en particulier si* $(2^{m'} - 1)/v'$ *divise* $(2^m - 1)/v$ *?*

À noter qu'une réponse partielle à ce problème et qui permettrait d'aboutir à un résultat général du type $R_v(m) = O(\sqrt{2^m/v})$ nous suffirait pour confirmer la conjecture de Patterson-Wiedemann.

# Chapitre 5

# Applications et constructions en dimension finie

Le but de ce dernier chapitre est d'explorer des applications et exemples concrets des constructions du chapitre 2. En général le calcul ou une estimation des sommes de Gauss est difficile, et il est alors très difficile d'obtenir des résultats en une dimension donnée. Comme nous l'avons vu dans les chapitres précédents, les sommes de Gauss sont plus faciles à manipuler quand elles ont la même valeur (où, dans le cas impair, cette situation idéale est obtenue par passage à la limite).

Dans un premier temps, nous expliquons comment obtenir une approximation de ce cas idéal en dimension finie en fonction de la distribution angulaire des sommes de Gauss. Nous montrons ensuite comment utiliser ce résultat et ceux du chapitre 4 pour étudier la construction de fonctions équilibrées de type Patterson-Wiedemann. Enfin, nous nous restreignons à un cas particulier où le calcul de ces sommes est simplifié, ce qui nous permet d'étudier une construction concrète avec de bonnes propriétés. Nous détaillons en particulier la construction d'une fonction hautement non-linéaire battant la borne quadratique.

## 5.1 Somme de Gauss et écart angulaire

Nous nous intéressons ici au cas particulier où les sommes de Gauss non triviales ont presque toutes le même argument $\theta = 0$ ou $\theta = \pi$ : c'est une deuxième généralisation (après celle par passage à la limite) du cas où les sommes de Gauss non triviales ont toutes la même valeur $\sqrt{q}$ ou $-\sqrt{q}$ (cf. Section 2.2.2 pour l'étude de cette situation).

Soit $f$ une fonction binaire de type PW généralisé de représentants $s$ équilibré sur les cosets de $G$ et $h$ sur $G$, alors pour $a \in L^\times$, le coefficient de Fourier de $f$ en $a$ vérifie

$$\hat{f}(a) = 1 + \frac{1}{v} \sum_{\chi \perp G, \chi \neq 1} \tau_L(\chi) \hat{s}(\overline{\chi}) \overline{\chi}(a) + \tilde{h}_G(a).$$

Le contrôle des amplitudes spectrales dépend alors de la largeur du plus petit secteur angulaire (centré en 0 ou en $\pi$) contenant les sommes de Gauss non triviales. Soit $\tau_L(\chi) = \sqrt{q} e^{i\theta_{\tau_L(\chi)}}$ pour $\chi \neq 1$ ($\theta_{\tau_L(\chi)} \in [-\pi, \pi]$), alors on dispose de la majoration suivante.

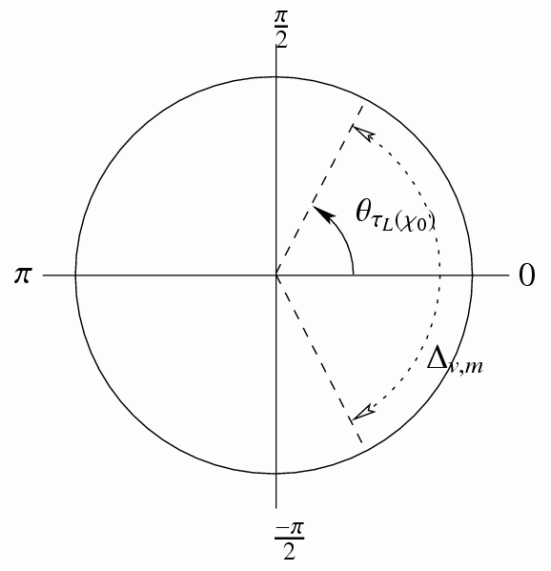


FIG. 5.1 – Exemple d'écart angulaire

**Proposition 5.1** *Soient*

$$\Delta_{v,m} = 2 \times \min\left(\max_{\chi \perp G, \chi \neq 1} |\theta_{\tau_L(\chi)} - 0|, \max_{\chi \perp G, \chi \neq 1} |\theta_{\tau_L(\chi)} - \pi|\right),$$

$$M_{v,m} = \sqrt{2(1 - \cos(\Delta_{v,m}/2))} \in [0, 2].$$

*Soit $f$ une fonction $(v, s, h)$–PW généralisée suivant le sous-groupe $G$ d'indice $v$ de $L^\times$ avec $s$ équilibrée sur $\Omega^* = L^\times/G - \{1\}$ et $h$ son représentant sur $G$. Pour $a \neq 0$, la transformée de Fourier de $f$ vérifie*

$$|\hat{f}(a)| \leq 1 + \frac{\sqrt{q}(v-1)\|\hat{s}\|_\infty}{v} M_{v,m} + \max(\sqrt{q} + \|\tilde{h}_G\|_{L^\times - G}, \|\tilde{h}_G\|_G).$$

*Démonstration.* Pour simplifier, on supposera dans la suite que $\theta = 0$.

On a $\Delta_{v,m} = 2 \max_{\chi \perp G, \chi \neq 1} |\theta_{\tau_L(\chi)}|$, ce qui implique $\max |\frac{\tau_L(\chi)}{\sqrt{q}} - 1| = \sqrt{2(1 - \cos(\Delta_{v,m}/2))} = M_{v,m}$ puisque $|e^{ia} - e^{ib}|^2 = 2(1 - cos(a - b))$. D'où, en se rappelant que $\hat{s}(1) = 0$ et en posant

$\|\hat{s}\|_\infty = \sup_{\chi \perp G} |\hat{s}(\chi)|$, on obtient

$$\begin{aligned}
|\hat{f}(a)| &= \left| 1 + \tilde{h}_G(a) + \frac{\sqrt{q}}{v} \sum_{\chi \perp G, \chi \neq 1} (e^{i\theta_{\tau_L(\chi)}} - 1 + 1)\hat{s}(\overline{\chi})\overline{\chi}(a) \right| \\
&\leq 1 + |\tilde{h}_G(a)| + \frac{M_{v,m}(v-1)\|\hat{s}\|_\infty \sqrt{q}}{v} + \left| \frac{\sqrt{q}}{v} \sum_{\chi \perp G} \hat{s}(\overline{\chi})\overline{\chi}(a) \right| \\
&\leq 1 + |\tilde{h}_G(a)| + \frac{M_{v,m}(v-1)\|\hat{s}\|_\infty \sqrt{q}}{v} + \frac{\sqrt{q}}{v} \left| \sum_{w \in \Omega^\times} s(w) \sum_{\chi \perp G} \overline{\chi}(aw) \right| \\
&\leq 1 + |\tilde{h}_G(a)| + \frac{M_{v,m}(v-1)\|\hat{s}\|_\infty \sqrt{q}}{v} + \sqrt{q}|s(\pi_G(a^{-1}))(1 - \delta_G(a))| \\
&\leq 1 + |\tilde{h}_G(a)| + \frac{M_{v,m}(v-1)\|\hat{s}\|_\infty \sqrt{q}}{v} + \sqrt{q}(1 - \delta_G(a)),
\end{aligned}$$

ce qui nous mène à la majoration recherchée qui dépend directement du plus petit secteur angulaire centré soit en 0, soit en $\pi$, contenant tous les arguments des sommes de Gauss non triviales. □

Plus l'écart angulaire et $\|\hat{s}\|_\infty$ sont faibles, alors (modulo les valeurs dépendantes de $h$) plus on restera proche de $\sqrt{q}$.

Notons $A_{v,m}$ le terme d'erreur $\frac{\sqrt{q}(v-1)\|\hat{s}\|_\infty}{v} M_{v,m}$. Un bon choix pour $s$ est de minimiser $\|\hat{s}\|_\infty$. Selon les conjectures habituelles, on souhaiterait obtenir $\|\hat{s}\|_\infty \approx \sqrt{v}$, et on sait dans tous les cas que $\|\hat{s}\|_\infty < v - 1$[1]. Les $s$ et $G$ intéressants sont ceux pour lesquels, on a $A_{v,m} < \sqrt{q}$, ce qui est tout à fait réalisable au vu des exemples ci-dessous (cf. Table 5.2). Pour obtenir une non-linéarité intéressante, il reste alors à trouver de bonnes fonctions $h$, soit en les construisant (e.g. tentatives numériques à partir de recollement) soit par un résultat d'existence à l'aide d'analyses en moyenne.

Par exemple, pour $m = 3r$, on choisit le sous-groupe $\mathbb{F}_{2^m}^\times / \mathbb{F}_{2^3}^\times$ de $L^\times = \mathbb{F}_{2^m}^\times$, d'indice 7. On calcule alors la taille des secteurs angulaires correspondants aux sommes de Gauss pour différentes valeurs de $r$ afin de trouver les écarts les plus faibles (pour $r \leq 100$, cf. Table 5.1). On voit que les secteurs peuvent être très petits, ce qui donne un intérêt réel à la proposition précédente. La dimension $m = 15$ (i.e. $r = 5$, dimension correspondant au contre exemple de Patterson-Wiedemann [PW83]) apparaît et correspond à un secteur angulaire de 27.05 degrés. La meilleur situation dans la table 5.1 se situe à la dimension $m = 39$ pour laquelle on obtient un secteur très petit d'angle $\Delta_{7,39} \approx 1.67^\circ$(c'est également le premier cas à partir de $m = 15$, d'angle plus petit que celui à $m = 15$). Dans les 3 cas les plus favorables, on calcule le terme d'erreur maximum (ici on majore $\|\hat{s}\|_\infty$ par $v - 1$) dans la table 5.2. En particulier, pour $m = 39$ et $m = 78$, le terme d'erreur $A_{7,3r}$ étant très inférieur à $0,2\sqrt{q}$, il semble possible de descendre en dessous de la borne de Patterson-Wiedemann à condition de trouver une fonction

[1] Pour $v$ grand, on peut raisonnablement espérer choisir $s$ tel que $\|\hat{s}\|_\infty \leq \sqrt{2v \ln v}$ selon le théorème 3.2 et même $\|\hat{s}\|_\infty \simeq \sqrt{v \ln v}$ d'après les résultats de [Kah85, SZ54] sur les polynômes trigonométriques aléatoires à coefficients binaires.

| $r$ | 13 | 26 | 39 | 52 | 65 | 78 | 91 | 96 |
|---|---|---|---|---|---|---|---|---|
| $\Delta_{7,3r}$ | 1.67 | 3.35 | 5.02 | 6.70 | 8.37 | 10.05 | 11.72 | 15.32 |
| $r$ | 83 | 70 | 57 | 31 | 44 | 18 | 5 | 8 |
| $\Delta_{7,3r}$ | 17.00 | 18.67 | 20.35 | 22.02 | 23.70 | 25.37 | 27.05 | 28.72 |

TAB. 5.1 – Taille en degré des secteurs angulaires pour différentes valeurs de $r$

| $r$ | 13 | 26 | 39 |
|---|---|---|---|
| $A_{7,3r}$ | $\lesssim 0.075\sqrt{q}$ | $\lesssim 0.150\sqrt{q}$ | $\lesssim 0.225\sqrt{q}$ |

TAB. 5.2 – Valeur maximale approximative du terme $A_{7,3r}$

$h$ définie sur le sous-groupe d'indice 7 avec une contribution modérée. On remarque que le résultat du chapitre précédent ne sera d'aucune aide pour cela puisque, bien que l'estimation en moyenne du théorème 4.1 s'applique (la condition $2m < N \log_2 e$ étant réalisée), la majoration par $1.18\sqrt{mN}$ n'est pas suffisante (car $N$ est trop proche de $2^m$).

**Exemple 5.1** *Dans le cas a priori le plus favorable, $m = 39$, pour battre la borne de Patterson-Wiedemann avec une fonction $(7, s, h)$–PW généralisée, il suffit de satisfaire aux contraintes suivantes :*

- $\hat{f}(0) = 1 + (\frac{q-1}{7} - 2\mathrm{w}(h)) < \frac{27}{32}\sqrt{2}\sqrt{q}$ *(cf. équation (2.7), Sec. 2.4) ; par exemple, $h$ presque équilibrée sur $G$,*
- $\|\tilde{h}_G\|_{L^\times - G} < (\frac{27}{32}\sqrt{2} - 1 - 0.075)\sqrt{q} - 1$*, soit environ au plus* $0.31\sqrt{N}$*,*
- $\|\tilde{h}_G\|_G < (\frac{27}{32}\sqrt{2} - 0.075)\sqrt{q} - 1$*, soit environ au plus* $2.96\sqrt{N}$*.*

Une telle fonction $h$ n'est pas facile à trouver (en particulier au regard de la deuxième condition), la marge de manoeuvre est cependant intéressante (pour estimer l'écart avec la borne de Parseval sur $G$, on peut remarquer que la somme des carrés des valeurs maximales de $|\tilde{h}_G(a)|$ pour $a \in L$ vaudrait environ $1.33q \times N$, ce qui est plus important que $q \times N$). Mais pour trouver une telle fonction $h$, il est nécessaire de définir une stratégie appropriée, une recherche exhaustive n'étant pas envisageable (il y a environ $2^{2^{36}}$ fonctions binaires différentes définies sur $G$)[2].

Un cas intéressant pour éviter ce problème est de se ramener au contexte de la section 4.3 (où $G$ est le groupe multiplicatif d'un sous-corps), ce qui soulève le point suivant :

**Problème 5.1** *Quel est l'écart angulaire minimal possible pour $m$ entier impair et $v = \frac{2^m-1}{2^k-1}$ où $k|m$ ?*

**Remarque 5.1** *On pourrait chercher à généraliser l'estimation au cas où le secteur angulaire n'est pas centré en 0 (ou $\pi$), i.e. quand il existe 2 secteurs angulaires (dû à la conju-*

[2]D'après l'étude asymptotique du chapitre 3, les chances de succès d'un tirage aléatoire sont probablement minimes pour $m$ grand.

*gaison complexe) centré en $\theta$ et $-\theta$, mais cela implique de savoir estimer les sommes du type $\sum_{\chi\perp G,\chi\neq 1,\arg\tau_L(\chi)>0}\chi(x)$, ce qui demeure difficile.*

**Remarque 5.2** *Dans cette partie, outre l'obligation de trouver des espaces où les sommes de Gauss ont un écart angulaire réduit et celle de construire sur G des fonctions de faible amplitude, le problème de l'estimation de l'amplitude spectrale minimale de la transformée de Fourier multiplicative de s surgit. C'est un problème intéressant à lui seul.*

Pour les différentes dimensions de la table 5.1 (et en général pour un indice $v$ petit), le fait que les sommes de Gauss non triviales s'accumulent au même endroit (arguments proches) devrait permettre de trouver des valeurs de $s$ adéquates pour aboutir à une construction menant à une estimation plus performante. Le nombre très important de possibilités rend cependant la tâche difficile à réaliser "à l'aveugle". Pour contourner ce problème, nous proposons dans la section 5.3 une approche de recherche légèrement différente à partir d'une construction explicite.

**Remarque 5.3** *Le critère d'accumulation des arguments des sommes de Gauss en un même point est en fait équivalent à ce que la somme sur les caractères orthogonaux à G, $\sum_{\chi\perp G}\tau_L(\chi)$, soit de module proche de $(v-1)\sqrt{q}$, puisque*

$$\left|\sum_{\chi\perp G}\tau_L(\chi)\right| = \left|-1+\sqrt{q}\sum_{\chi\neq 1\in G^\perp} e^{i\theta_{\tau_L(\chi)}}\right|.$$

*D'après la formule de Poisson (cf. Proposition 1.8), ceci est donc équivalent à $\left|v\sum_{x\in G}\mu(x)\right|$ proche de $(v-1)\sqrt{q}$. Or $\sum_{x\in G}\mu(x)=\frac{1}{v}\sum_{x\in\mathbb{F}_q}\mu(x^v)$, et le critère est donc que le polynôme $X^v$ s'approche de la borne de Weil, i.e.*

$$\left|\sum_{x\in\mathbb{F}_q}\mu(x^v)\right| \simeq (v-1)\sqrt{q}.$$

*Ce problème, lié à la borne de Carlitz-Uchiyama, a été étudié dans [Rod92] dans le cas du dual d'un code BCH binaire de distance construite $\delta = 9$ : le résultat correspond au cas $v = 7$ et Rodier observe alors qu'on se situe très proche de la borne en dimension 39. Ce résultat est donc à rapprocher de la taille des secteurs angulaires indiqués Table 5.1.*

## 5.2 Fonctions équilibrées hautement non-linéaires

Contrairement à la construction de type PW, la définition d'une fonction de type PW généralisé permet d'obtenir des fonctions équilibrées. Nous étudions ici leur non-linéarité.

Soit $f$ une fonction de type PW généralisé définie par $\mu(f_0)$ en 0, par $s$ équilibrée sur les cosets de $G$ et par la fonction $h$ sur $G$. Pour que $f$ soit équilibrée sur $L$, il faut que le vecteur $\{\mu(f_0),(h(x))_{x\in G}\}$ soit équilibré. Supposant $s$ fixé, un simple dénombrement permet de compter le nombre de possibilités pour $(f_0,h)$ : $1-2h$ doit être de poids $(N+1)/2$ si $f_0=0$ et de poids $(N-1)/2$ si $f_0=1$. Il y a donc au total $\binom{N}{(N+1)/2}+\binom{N}{(N-1)/2}=2\binom{N}{(N-1)/2}$ constructions différentes.

A partir de ce constat et du théorème 4.1, on en déduit le résultat d'existence suivant :

**Théorème 5.1** *Soit $m > 0$ un entier, $G$ un sous-groupe de $L^\times$ d'ordre $N$, d'indice $v$ et $s$ une fonction équilibrée sur $\Omega^* = L^\times/G - \{1\}$. Soit $c > 0$ un réel quelconque tel que $N > 2c^2m$ et $\binom{N}{(N-1)/2} \geq 2^{m(1-c^2\log_2 e)+N-1}$.*

*Si un tel réel c existe, alors il existe une fonction $f$ de type PW généralisé de représentant $s$ sur les cosets non triviaux de $G$ telle que $f$ est équilibrée sur $L$ et vérifiant*

$$\|\tilde{f_{|G}}\|_\infty \leq c\sqrt{2Nm}.$$

*Démonstration.* Étant donnés les valeurs de $f$ en 0 et sur $G$, $\mu(f_0)$ et $h$, il suffit de regarder la densité de l'ensemble des $(f_0, h)$ permettant d'obtenir un résultat équilibré. En effet, la condition $\binom{N}{(N-1)/2} > 2^{m(1-c^2\log_2 e)+N-1}$ s'écrit

$$\frac{\binom{N}{(N-1)/2}}{2^{N-1}} + (1 - 2^{m(1-c^2\log_2 e)}) > 1,$$

donc, d'après le théorème 4.1, la somme de la densité des fonctions $h$ convenables et de la densité des fonctions de l'ensemble $\{h : G \to \{\pm 1\}, \|\tilde{h_G}\|_\infty \leq c\sqrt{2Nm}\}$ parmi l'ensemble des fonctions définies sur $G$ est strictement plus grande que 1. Ainsi, les deux ensembles ne sont pas disjoints, d'où le résultat. □

Ce théorème signifie donc qu'il existe des fonctions binaires définies sur $G$ et équilibrées qui ont elles-aussi une non-linéarité partielle élevée. En outre, plus la différence

$$\binom{N}{(N-1)/2} - 2^{m(1-c^2\log_2 e)+N-1}$$

est grande, plus le nombre de fonctions possibles est grand. De ce théorème, on déduit en particulier :

**Corollaire 5.1** *Si la relation suivante est vérifiée*

$$\frac{N}{2} > \frac{m + N - 1 - \log_2\binom{N}{(N-1)/2}}{\log_2 e}, \tag{5.1}$$

*alors il existe une fonction $f$ du type précédent telle que*

$$\|\tilde{f_{|G}}\|_\infty \leq \sqrt{2 \times \frac{m + N - 1 - \log_2\binom{N}{(N-1)/2}}{\log_2 e} \times N}.$$

*Démonstration.* Il suffit de fixer $c$ dans le théorème précédent avec la plus petite valeur possible telle que $\binom{N}{(N-1)/2} \geq 2^{m(1-c^2\log_2 e)+N-1}$. □

En prenant $G = \mathbb{F}_q^\times$, la condition (5.1) est automatiquement vérifiée et on obtient :

**Corollaire 5.2** *Pour tout $m \geq 3$, il existe une fonction binaire équilibrée $f : \mathbb{F}_{2^m} \to \{\pm 1\}$ telle que*

$$\|\hat{f}\|_\infty \leq 1 + \sqrt{2 \times \frac{m + 2^m - 2 - \log_2 \binom{2^m-1}{2^{m-1}-1}}{\log_2 e} \times (2^m - 1)}.$$

Cependant le résultat n'est pas très intéressant dans ce cas. Pour illustrer, pour $m = 8$, cela donne une borne d'environ $4\sqrt{2^8}$ – ce qui est élevé comparé à l'existant (cf. Table 5.3).

Pour $G$ quelconque, d'après le corollaire 5.1 associé à la proposition 5.1, dont on reprend ici les notations, on en conclut que si l'inégalité (5.1) est satisfaite, alors il existe une fonction $f$ équilibrée de type PW généralisé de représentant $s$ telle que

$$\|\hat{f}\|_\infty \leq 1 + \sqrt{q} + \frac{\sqrt{q}(v-1)\|\hat{s}\|_\infty}{v} M_{v,m} + \sqrt{2 \times \frac{m + N - 1 - \log_2 \binom{N}{(N-1)/2}}{\log_2 e} \times N}, \qquad (5.2)$$

où $M_{v,m} = \sqrt{2(1 - \cos(\Delta_{v,m}/2))}$ et $\Delta_{v,m}$ est l'écart angulaire des sommes de Gauss $\Delta_{v,m} = 2 \times \min(\max_{\chi \perp G, \chi \neq 1} |\theta_{\tau_L(\chi)} - 0|, \max_{\chi \perp G, \chi \neq 1} |\theta_{\tau_L(\chi)} - \pi|)$.

**Remarque 5.4** *Le dernier terme peut être approché par*

$$\sqrt{\frac{2m + \log_2 N}{v \log_2 e}} \times \sqrt{q},$$

*pour $m, N \to \infty$ (en utilisant la formule de Stirling, cf. détails pour $m = 2t$ ci-après), si bien que si $2m + \log_2 N$ est négligeable devant $v$ alors le terme est négligeable devant $\sqrt{q}$. Comme $1 \leq N < q$, alors $2m \leq 2m + \log_2 N \leq 3m$, et il suffit d'avoir $m$ négligeable devant $v$ (c'est par exemple le cas si $v$ est de l'ordre de $\sqrt{q}$).*

*Si on compare ce résultat à celui du corollaire 4.1, ici la borne varie environ entre $1.18\sqrt{mN}$ et $1.44\sqrt{mN}$, donc la différence engendrée par la restriction à des fonctions équilibrées n'est pas très importante.*

Enfin comme dans la section 4.3, on peut regarder le cadre favorable où $G$ est le groupe des inversibles d'un sous-corps de $L$ pour obtenir des bornes plus intéressantes.

**Corollaire 5.3** *Si $G = K^\times$ avec $K = \mathbb{F}_{2^k}$ un sous-corps de $L$, alors d'après l'équation (4.5), si on choisit $f$ équilibrée sur $K$, on en déduit la majoration suivante :*

$$\|\hat{f}\|_\infty \leq \sqrt{q} + \sqrt{q}(1 - \frac{2^k - 1}{q - 1})\|\hat{s}\|_\infty M_{v,m} + RB(k).$$

Pour utiliser une forme quadratique pour construire explicitement $f$, il faudra prendre une forme parabolique. Au mieux, on aura alors $\|\hat{f}_K\|_\infty$ égal à $2^{\frac{k+2}{2}}$ si $k$ est pair et $2^{\frac{k+1}{2}}$ si $k$ est impair (où $k = \log_2(N+1)$).

### 5.2.1 Application aux fonctions équilibrées en dimension paire

Dans le cas pair, comme nous l'avons déjà vu, les sommes de Gauss n'ont pas le même comportement qu'en dimension impaire, et par là même elles peuvent être plus faciles à manipuler. Nous avons vu qu'il est possible d'obtenir des sommes de Gauss rationnelles en considérant un sous-groupe construit à partir du corps pour lequel $L$ est une extension quadratique (voir aussi le théorème 1.4). Il existe également d'autres situations où les sommes de Gauss ont un écart angulaire faible. En pratique, il est alors plus facile de trouver des paramètres permettant de réduire la borne supérieure dans (5.2).

Soit $m = 2t$ un entier pair. Soit $G$ sous-groupe de $L^\times$ d'ordre $N = 2^t - 1$ et d'indice $v = 2^t + 1$ (cf. section 2.2.2). Dans ce cas, $\tau_L(\chi) = \sqrt{q}$ pour tout $\chi \perp G, \chi \neq 1$ et l'écart angulaire $\Delta_{v,m}$ vaut 0, tout comme $M_{v,m}$. L'inégalité (5.2) lorsque l'équation (5.1) est vérifiée (ce qui est le cas à partir de $t = 5$) devient

$$\|\hat{f}\|_\infty \leq 1 + \sqrt{q} + \sqrt{2 \times \frac{2t + 2^t - 2 - \log_2 \binom{2^t-1}{2^{t-1}-1}}{\log_2 e} \times (2^t - 1)}.$$

Ce qui donne

$$\|\hat{f}\|_\infty \leq 1 + \sqrt{q}\left(1 + \sqrt{2\frac{2t + 2^t - 2 - \log_2 \binom{2^t-1}{2^{t-1}-1}}{(2^t+1)\log_2 e}}\right). \tag{5.3}$$

D'un point de vue asymptotique (pour $t \to \infty$), d'après la formule de Stirling

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,$$

on a

$$\binom{2^t - 1}{2^{t-1} - 1} \sim \frac{(2^t - 1)!}{((2^{t-1} - 1)! \times 2^{t-1})^2} \sim \frac{\sqrt{2}2^{2^t}}{\sqrt{\pi 2^t e}}.$$

Le terme d'erreur dans l'inégalité précédente est donc équivalent à

$$\sqrt{\frac{5}{\log_2 e}}\sqrt{\frac{t}{2^t}}$$

pour $t$ qui tend vers l'infini.

**Proposition 5.2** *Pour $m = 2t$, il existe $f : \mathbb{F}_{2^m} \to \{\pm 1\}$ équilibrée telle que*

$$\|\hat{f}\|_\infty \leq 2^t\left(1 + \sqrt{\frac{5}{\log_2 e}}\sqrt{\frac{t}{2^t}} + o\left(\sqrt{\frac{5}{\log_2 e}}\sqrt{\frac{t}{2^t}}\right)\right).$$

Asymptotiquement, nous avons donc un résultat sur le rayon spectral équilibré de $RM(1, 2t)$ :

**Corollaire 5.4**

$$RB(2t) \leq 2^t + \sqrt{\frac{5}{\log_2 e}}\sqrt{t2^t} + o_{t\to\infty}\left(\sqrt{\frac{5}{\log_2 e}}\sqrt{t2^t}\right).$$

Ainsi, $RB(2t)$ a un comportement asymptotique très proche de $R(2t)$.

**Remarque 5.5** *Suivant l'esprit de la remarque 4.2, on peut également dans le cas équilibré distinguer la non-linéarité par rapport à $G \cup \{0\}$ et à $L^\times - G$ pour améliorer très légèrement les bornes obtenues, au prix cependant de contraintes sur N et m plus difficiles à exprimer et à traiter. Une approximation rapide permet de diminuer le* 5 *de l'inégalité précédente en un* 4*.*

### 5.2.2 Comparaison avec l'existant

Comme pour le rayon spectral $R(m)$, la détermination du rayon spectral équilibré $RB(m)$ (ou par équivalence la non-linéarité maximale correspondante $2^{m-1} - \frac{1}{2}RB(m)$) est délicate. On sait que les fonctions courbes ne sont pas équilibrées et la détermination de $RB(m)$ est par conséquent difficile également dans le cas pair. Les bornes de Parseval et quadratique deviennent

$$2^{\frac{m}{2}} + 4 \leq RB(m) \leq 2^{\lceil \frac{m+1}{2} \rceil}$$

dans le cas pair ; et on a pour $m, m' \in \mathbb{N}$

$$RB(m + m') \leq RB(m)R(m').$$

Une méthode générale utilisée pour construire des fonctions équilibrées hautement non-linéaires est alors la technique d'équilibrage de fonctions hautement non-linéaires. Dans le cas impair, le contre-exemple de Patterson-Wiedemann ($m = 15$, amplitude 216) est naturellement la cible principale de ces techniques. Dans [LVZ98] (voir aussi [LZ05]) un équilibrage d'amplitude 244 est obtenu. Puis dans [MS02], ce résultat est amélioré pour obtenir une fonction équilibrée d'amplitude 240. Enfin, toujours à partir de ces fonctions, de nouvelles améliorations ont très récemment été trouvées dans [SM07], pour une amplitude 236, puis dans [KY07a] aboutissant à une amplitude de 232, ce qui implique :

**Proposition 5.3** *Si m est un entier impair supérieur ou égal à* 15 *alors*

$$RB(m) \leq \frac{29}{32} 2^{\frac{m+1}{2}} < 2^{\frac{m+1}{2}}.$$

Suivant la même idée d'équilibrage, dans [Mai07], Maitra explique comment obtenir une fonction équilibrée en 13 variables d'amplitude 124 (battant donc la borne quadratique) à partir de la fonction hautement non-linéaire en 9 variables de non-linéarité 242 obtenue dans [KY07b].

Dans le cas pair, un algorithme constructif proposé par Dobbertin [Dob94] permettant l'équilibrage de fonctions courbes donne le résultat suivant.

**Proposition 5.4** *Soit $m = 2t$ et $f$ une fonction courbe normale, alors il existe $g$ un équilibrage de $f$ tel que*

$$\|\hat{g}\|_\infty \leq 2^t + RB(t).$$

*En écrivant $m = 2^r s$ ($r \geq 1$, $s$ impair), on obtient*

$$RB(m) \leq \sum_{i=1}^{r} 2^{\frac{m}{2^i}} + 2^{\frac{s+1}{2}}.$$

De plus, Dobbertin, dans le même article [Dob94], conjecture que la première inégalité est en fait une égalité.

On voit que cette proposition permet de montrer que $RB(2t) \sim R(2t)$ tout comme le corollaire 5.4, mais la borne de Dobbertin met en évidence une différence asymptotique plus faible. Toutefois, le lien entre les 2 résultats permet de nous conforter dans l'intérêt potentiel de la construction par rapport à un sous-groupe de $L^\times$ pour $m$ impair.

**Utilisation de la construction par récurrence sur le groupe multiplicatif d'un sous-corps.** A partir du corollaire 4.4, on remarque que la construction de fonctions équilibrées de type PW généralisé dans le cas où $G = \mathbb{F}_{2^t}^\times$ et $L = \mathbb{F}_{2^{2t}}$ permet d'obtenir le même résultat (ce n'est pas surprenant, car dans ce cas, la construction correspond réellement à l'équilibrage d'une fonction courbe normale). En effet, le corollaire restreint aux fonctions équilibrées donne ici $RB(2t) \leq RB(t) + \sqrt{2^{2t}}$. Ainsi par récurrence, en distinguant $t$ pair et $t$ impair, on obtient exactement le résultat de Dobbertin.

**Bornes connues.** A noter qu'en pratique, à partir de $m = 8$ (pair ou impair), on ne connaît pas les valeurs exactes prises par $RB(m)$. On reporte dans la table 5.3, pour quelques petites valeurs de $m$, les encadrements connus pour $RB(m)$, dont la plupart sont obtenus à partir des bornes ci-dessus et en utilisant le fait que $RB(m)$ est un multiple de 4 (en tant que coefficient de Walsh d'une fonction équilibrée). On inclut une colonne pour comparer avec les valeurs de $R(m)$ (déduites des bornes et des constructions connues – cf. partie 1.2.3).

| $m$ | $RB(m)$ | $R(m)$ |
|---|---|---|
| 3 | 4 | 4 |
| 4 | 8 | 4 |
| 5 | 8 | 8 |
| 6 | 12 | 8 |
| 7 | 16 | 16 |
| 8 | 20 ou 24 | 16 |
| 9 | $24 \leq RB(m) \leq 32$ | $24 \leq R(m) \leq 28$ |
| 10 | 36 ou 40 | 32 |
| 11 | $48 \leq RB(m) \leq 64$ | $46 \leq R(m) \leq 56$ |
| 12 | 68, 72 ou 76 | 64 |
| 13 | $92 \leq RB(m) \leq 124$ | $92 \leq R(m) \leq 112$ |
| 14 | $132 \leq RB(m) \leq 144$ | 128 |
| 15 | $184 \leq RB(m) \leq 232$ | $182 \leq R(m) \leq 216$ |

TAB. 5.3 – Les premières valeurs de $RB(m)$ et $R(m)$

## 5.3 Construction par recollement quadratique

En plus des études asymptotique (cf. Chapitre 3) et en moyenne (cf. Section 4.2) des bornes envisageables, trouver une construction concrète, à dimension finie, de fonctions permettant d'approcher la borne de Patterson-Wiedemann est également un problème difficile. Nous proposons ici une construction de fonctions de type PW généralisé et expliquons en quoi elle est intéressante. Afin d'obtenir de bonnes propriétés et de choisir au mieux les différents paramètres, nous allons en fait nous placer dans un cas où les sommes de Gauss sont calculables relativement aisément (ce qui n'est pas vrai en général – ici elles seront quadratiques). L'avantage, dans ce contexte, est de choisir (lorsque cela est possible) les fonctions $s$ et $h$ en adéquation avec les valeurs de ces sommes.

Plus concrètement, nous utilisons des résidus quadratiques pour obtenir un exemple de fonction hautement non-linéaire en 15 variables constante sur les cosets non-triviaux du sous-groupe d'indice 7 de $\mathbb{F}_{2^{15}}^{\times}$ et battant alors la borne quadratique.

### 5.3.1 Description

Soit $v > 3$ un nombre premier congru à 3 modulo 4 tel que 2 engendre le groupe des résidus quadratiques modulo $v$. Pour reprendre la terminologie de [Lan96], $(v, 2)$ vérifie les conditions de résiduosité quadratique. Soit $\chi$ un caractère multiplicatif de $L^{\times}$ d'ordre $v$ (si $G$ sous-groupe d'indice $v$, c'est le cas pour tout $\chi \neq 1$ orthogonal à $G$), alors comme $v$ est premier, $\chi$ génère l'ensemble des caractères d'ordre $v$. De plus, d'après [Lan97], on sait sous cette hypothèse que les sommes de Gauss sont quadratiques et qu'il existe des entiers $h$, $A$ et $B$ tels que :

$$\tau_L(\chi) = 2^h(A + B\sqrt{-v}), \qquad 2 \nmid \operatorname{pgcd}(A, B),$$

où $\sqrt{-v}$ représente une racine carrée complexe de $-v$ et $h$ est déterminé en partie via le théorème de Stickelberger (cf. Théorème 1.5) et le nombre de classes du corps quadratique imaginaire $\mathbb{Q}(\sqrt{-v})$. D'autre part, pour $j \in \{1, \ldots, v-1\}$, on connaît également la somme de Gauss en $\chi^j$ :

$$\tau_L(\chi^j) = 2^h(A + \left(\frac{j}{v}\right) B\sqrt{-v}).$$

**Remarque 5.6** *Si on choisit le plus petit corps possible pour disposer de caractères d'ordre $v$, soit $\mathbb{F}_{2^{(v-1)/2}}$, alors dans ce cas si $f$ représente la somme des chiffres de $(2^{(v-1)/2} - 1)/v$, Stickelberger permet d'affirmer que $\tau_L(\chi)/2^f$ est un élément de $\mathbb{Q}(\sqrt{-v})$ de norme $2^l$ où $l = h(-v)$ est le nombre de classes de $\mathbb{Q}(\sqrt{-v})$ : plus précisément $\tau_L(\chi)/2^f = (a + b\sqrt{-v})$ avec $2a, 2b$ entiers et $a^2 + b^2 v = 2^l$. La détermination de $h, A, B$ dépend ainsi de $a, b, f$ et $l$ où $2^{2(h-f)}(A^2 + vB^2) = 2^l$. A noter qu'on a en particulier $2f + l = m = (v-1)/2$. D'après [Lan97], on sait de plus que $A \equiv -2^{1+(v-1)/2-f} \mod v$.*

*Dans le cas général où L est une extension de $\mathbb{F}_{2^{(v-1)/2}}$, le relèvement par Davenport-Hasse permet d'étendre immédiatement ce résultat.*

Soit $\gamma$ une racine primitive de $L$. Supposons que $\chi(\gamma)$ soit la racine principale $\zeta_v$ de l'unité d'ordre $v$ dans $\mathbb{C}$. Alors les éléments $\gamma^0, \gamma^1, \ldots, \gamma^{v-1}$ forment un système de représentants du quotient $\Omega = L^{\times}/G$.

**Définition 5.1** *Soit g la fonction ternaire de L dans* $\{\pm 1, 0\}$ *définie en* $x \in L$ *par*

$$g(x) = \sum_{j=1}^{v-1} \left(\frac{j}{v}\right) \delta_G(\gamma^{-j} x).$$

*g est le recollement quadratique (quadratic residue spread) associé à v et G.*

Elle est en fait construite à partir d'une fonction de type PW généralisé de représentant $s : \gamma^j \mapsto \left(\frac{j}{v}\right)$ sur $\Omega^*$. Par construction, les coefficients de Fourier de la fonction ternaire $g$ s'expriment en fonction du symbole de Legendre.

**Lemme 5.1** *Soit* $0 \leq k \leq 2^m - 2$,

$$\hat{g}(\gamma^k) = 2^h \left( -\left(\frac{k}{v}\right) A - B + vB\delta_0(k) \right). \tag{5.4}$$

*Démonstration.* On a $v\hat{g}(\gamma^k) = \sum_{\psi \perp G} \tau_L(\psi) \hat{s}(\overline{\psi}) \overline{\psi}(\gamma^k)$, soit $v\hat{g}(\gamma^k) = \sum_{j=1}^{v-1} \tau_L(\chi^j) \hat{s}(\overline{\chi}^j) \overline{\chi}^j(\gamma^k)$ où $\chi$ génère $G^\perp$. Or, d'après un résultat de Gauss [Gau07] (voir par exemple [Lan99]),

$$\sum_{i=1}^{v-1} \left(\frac{i}{v}\right) \zeta_v^{is} = \left(\frac{s}{v}\right) \sqrt{-v},$$

ainsi $\hat{s}(\overline{\chi}^j) = \sum_{i=1}^{v-1} \left(\frac{i}{v}\right) \overline{\chi}^j(\gamma^i) = \left(\frac{-j}{v}\right) \sqrt{-v}$. On en déduit

$$\begin{aligned}
v\hat{g}(\gamma^k) &= \sum_{j=1}^{v-1} 2^h \left( A + \left(\frac{j}{v}\right) B \sqrt{-v} \right) . \left(\frac{-j}{v}\right) \sqrt{-v} . \zeta_v^{-jk} \\
&= 2^h \sum_{j=1}^{v-1} \left( A \left(\frac{-j}{v}\right) \sqrt{-v} - v(-1)^{(v-1)/2} B \right) \zeta_v^{-jk} \\
&= 2^h \sum_{j=1}^{v-1} \left( A \left(\frac{-j}{v}\right) \sqrt{-v} + vB \right) \zeta_v^{-jk}, \text{ puisque } v \equiv 3 \mod 4 \\
&= 2^h \left( A \sqrt{-v} \sum_{j=1}^{v-1} \left(\frac{-j}{v}\right) \zeta_v^{-jk} + Bv \sum_{j=1}^{v-1} \zeta_v^{-jk} \right) \\
&= 2^h v \left( -A \left(\frac{k}{v}\right) + B \left( \sum_{j=0}^{v-1} \zeta_v^{jk} - 1 \right) \right).
\end{aligned}$$

□

On prêtera attention au fait que le choix des valeurs sur les cosets en fonction de la résiduosité quadratique nous permet ici de construire un représentant $s : \gamma^j \mapsto \left(\frac{j}{v}\right)$ d'amplitude spectrale exactement $\sqrt{v}$, i.e. optimale.

### 5.3.2 Un exemple de construction hautement non-linéaire

On veut maintenant compléter la construction de $g$ sur $G$ afin d'obtenir une fonction binaire. Fixons $v = 7$, alors $f = 1$ et d'après la remarque 5.6, il existe un caractère non trivial $\chi'$ d'ordre $v$ tel que $\tau_{\mathbb{F}_8}(\chi') = 2^{f-1}(a' + b'\sqrt{-7})$ où $a'^2 + 7b'^2 = 2^{l+2} = 2^{m-2f+2} = 8$ pour $a', b'$ entiers. En fait $a' \equiv -2^{1+(v-1)/2-f} \mod v \equiv -1$ et $b' = \pm 1$. Le corps étant trop petit pour qu'une fonction intéressante soit possible sur $G$, on considère une extension.

Soit $\chi$ un caractère multiplicatif d'ordre 7 dans $\mathbb{F}_{2^{15}}$ tel qu'on peut réaliser $\chi$ par le relevé du caractère non trivial $\chi'$ de $\mathbb{F}_8$ de sorte que (via Davenport-Hasse – cf. Chapitre 1),

$$\tau_{\mathbb{F}_{2^{15}}}(\chi) = \tau_{\mathbb{F}_8}(\chi')^5 = (-1 + \sqrt{-7})^5 = -16(11 + \sqrt{-7})$$

i.e. $h = 4$, $A = -11$ et $B = \pm 1$. Les valeurs prises par la transformée de Fourier ternaire du recollement quadratique sont donc $-160$, $-96$ et 192 (soit une amplitude d'environ $1.06\sqrt{2^{15}}$).

On effectue enfin une recherche exhaustive parmi les monômes $x^\alpha$, pour finalement obtenir que la fonction binaire

$$f(x) = -\delta_0(x) + \mu(x^\alpha)\delta_G(x) + \sum_{j=1}^{v-1} \left(\frac{j}{v}\right) \delta_G(\gamma^{-j}x)$$

a une amplitude spectrale de 232 pour $\alpha = 755$. Le spectre de la fonction $f$ et de la fonction puissance associée sont détaillés dans les tables 5.4 et 5.5. On ne dépasse pas la borne de Patterson-Wiedemann avec $f$, mais la fonction permet de battre la borne quadratique. C'est un exemple de fonction de type PW généralisé hautement non-linéaire.

**Corollaire 5.5** *Pour $m = 15$, il existe une fonction $f$ de type PW généralisé hautement non-linéaire vérifiant $\|\hat{f}\|_\infty < \mathrm{bq}(15)$.*

De plus, ont atteint le meilleur résultat trouvé dans [LZ01] à partir d'une construction de type PW constante sur le coset trivial. Ce qui justifie donc pleinement l'intérêt d'étudier cette construction généralisée. La liberté supplémentaire que nous avons introduite dans cette notion devrait permettre (si on était capable de parcourir toutes les fonctions définies sur $G$) de faire mieux que dans [LZ01].

| Valeur | -216 | -152 | -88 | -24 | 40 | 104 | 168 | 232 |
|---|---|---|---|---|---|---|---|---|
| Multiplicité | 7550 | 6494 | 1208 | 3020 | 755 | 151 | 6795 | 6795 |

TAB. 5.4 – Spectre de $f(x) = -\delta_0(x) + \mu(x^{755})\delta_G(x) + \sum_{j=1}^{v-1} \left(\frac{j}{v}\right) \delta_G(x/\gamma^j)$.

Cette construction effective surclasse les estimations en moyenne vues précédemment. Par exemple, le corollaire 4.1 nous permet d'assurer l'existence d'une fonction définie sur $G$ d'amplitude seulement inférieure à $\sqrt{2\ln 2}\sqrt{mN} = \sqrt{2\ln 2}\sqrt{15 \times \frac{2^{15}-1}{7}}$, soit environ 312, ce qui est moins favorable que la fonction puissance choisie (amplitude 279). D'autre part, le principe même de l'estimation d'une amplitude par inégalité triangulaire ne peut être suffisant puisqu'ici, même en prenant une fonction définie sur $G$ d'amplitude $69 = \lceil \sqrt{N} \rceil$ (i.e. d'amplitude

| Valeur | -279 | -151 | -119 | -55 | -23 |
|---|---|---|---|---|---|
| Multiplicité | 453 | 1 | 755 | 8305 | 6795 |
| Valeur | 9 | 41 | 73 | 137 | 201 |
| Multiplicité | 5738 | 6795 | 755 | 3020 | 151 |

TAB. 5.5 – Spectre de la fonction binaire $\mu(x^{755})\delta_G(x)$.

théorique minimale au plus près de la borne de Parseval), la somme des amplitudes en valeur absolue ne permettrait que d'obtenir une borne supérieure ou égale à 261.

A souligner que la fonction puissance trouvée implique que $R_7(15) \leq 4.08\sqrt{\frac{2^{15}-1}{7}}$.

**Remarque 5.7** *Il serait également intéressant d'essayer d'équilibrer cette construction par résiduosité quadratique. Pour conserver l'idée de la section 5.2, on pourrait alors essayer d'équilibrer uniquement la partie définie sur le sous-groupe d'indice* 7 *sans modifier les valeurs sur les cosets non-triviaux.*

### 5.3.3 Autres constructions

Pour construire d'autres fonctions ternaires de non-linéarité élevée, il semble adéquat de se placer dans un cadre où $B$ est faible, puisque selon le lemme précédent, le spectre est $\left\{2^h(-A-B), 2^h(A-B), 2^h(v-1)B\right\}$ (ou encore $\left\{2^f(-a-b), 2^f(a-b), 2^f(v-1)b\right\}$ dans le corps de base $\mathbb{F}_{2^{(v-1)/2}}$ en reprenant les notations de la remarque 5.6). En effet, plus $v$ est grand, plus il est important au vu de la dernière des 3 valeurs d'avoir B petit ; de plus, si $B$ est petit par rapport à $A$ cela permet de limiter la variation entre la première et la deuxième valeur.

Suivant la même remarque 5.6, dans $\mathbb{F}_{2^{(v-1)/2}}$, on sait que $(2a)^2 + v(2b)^2 = 2^{l+2}$ où $2a, 2b$ sont entiers et $l$ est le nombre de classes du corps quadratique imaginaire $\mathbb{Q}(\sqrt{-v})$ : un nombre de classes tel que $2^{l+2}$ soit proche de $v$ permet donc d'aboutir à $2b$ petit (si $2^{l+2} < 4v$ alors $|2b| < 2$)[3]. La table 5.6 permet de se faire une idée sur les possibilités pour $l \leq 7$ ; on pourra se référer à [Wei] pour plus de détails sur la valeur du nombre de classes $h(-v)$ pour $v$ quelconque.

| Nombre de classes $l = h(-v)$ | (Valeurs de $v$, sommes $f$ des chiffres) |
|---|---|
| 1 | (7, 1) |
| 3 | (23, 4) |
| 5 | (47, 9) ; (79, 17) |
| 7 | (71, 14) |

TAB. 5.6 – Nombre premiers $v$ ($v < 100$) vérifiant la condition de résiduosité pour un nombre de classes inférieur à 7

[3]Trouver tous les indices $v$ vérifiant cette condition est un problème délicat, la détermination du nombre de classes $h(-v)$ pour $v$ quelconque étant déjà un problème difficile. Une piste possible est d'utiliser les estimations obtenues sous l'hypothèse de Riemann généralisée.

Par exemple, dans le cas $v = 23$, on a l'équation $(2a)^2 + (2b)^2 v = 2^{3+2}$ d'où $A = \pm 3$ et $B = \pm 1$ et il existe $\chi$ d'ordre $v$ tel que $\tau_{\mathbb{F}_{2^{11}}}(\chi) = 2^3(A + \sqrt{-23})$ (où on sait que $A \equiv -3 \mod 23$ toujours d'après la remarque). Le spectre ternaire sera donc $\{2^4, -2^5, 2^4 \times 11\}$ ; la borne quadratique pour $m = 11$ étant de valeur 64, nous sommes bien au dessus et les estimations seront donc inutiles : la seule solution est de réellement trouver des valeurs sur le sous-groupe telles que les amplitudes se compensent (mais une recherche exhaustive n'est toujours pas envisageable puisque celui-ci est d'ordre 89). Pour y remédier, on peut se placer dans une extension (comme pour le cas $v = 7$), par exemple de degré 3, alors on obtient des sommes de Gauss de la forme $\tau_{\mathbb{F}_{2^{33}}}(\chi) = 2^9(180 \pm 4\sqrt{-23})$ ; ce qui nous donne le spectre ternaire $\{-2^{12} \times 23, 2^{12} \times 22, 2^{12} \times 11\}$ avec une amplitude de l'ordre de $1.016\sqrt{2^{33}}$. D'après le corollaire 4.1, on sait qu'il existe une fonction définie sur le sous-groupe d'indice $v = 23$ de $\mathbb{F}_{2^{33}}$ d'amplitude inférieure à $\sqrt{2\ln 2}\sqrt{33 \times \frac{2^{33}-1}{23}} \simeq 1.410\sqrt{2^{33}}$, ce qui n'est toujours pas suffisant. Par contre on dispose de plus de marge par rapport à la borne de Parseval sur le sous-groupe car $\sqrt{\frac{2^{33}-1}{23}} \simeq 0.209\sqrt{2^{33}}$.

Via la condition $B$ faible par rapport à $A$, on retrouve finalement ici l'observation faite à la section 5.1 : si les sommes de Gauss ont un petit écart angulaire alors on reste proche de $\sqrt{q}$ sur toute la partie en dehors de $G$ dans la transformée de Fourier d'une fonction PW-généralisée. La construction effective présentée ci-dessus permet d'améliorer fortement la proposition 5.1 (surtout pour $v$ grand – pour l'exemple du paragraphe précédent, la proposition ne fournit qu'une majoration d'environ $3.233\sqrt{2^{33}}$). Dans la table 5.7 nous indiquons (de gauche à droite), pour plusieurs indices $v$ de la table 5.6, la forme des sommes de Gauss sur $G^\perp$, l'extension (de degré inférieur à 15) permettant de s'approcher au mieux de $\sqrt{q}$ sur la partie définie par le recollement quadratique associé à $v$. On y indique également (toujours de g. à d.) l'écart angulaire des sommes de Gauss et la majoration obtenue via la proposition 5.1 (pour souligner l'amélioration obtenue), la borne retournée par le corollaire 4.1 et la valeur de la borne de Parseval sur le sous-groupe d'indice $v$ (pour mesurer la marge de manoeuvre sur $G$).

On y voit que l'amplitude de $g$ est la plus faible pour les écarts angulaires les plus faibles mais la différence entre les 4 cas est très relative comparée aux différentes bornes extraites de la proposition 5.1. Les deux colonnes de droite confirment que l'estimation en moyenne n'est pas intéressante pour $m = (v-1)/2$ et que la marge par rapport à la borne de Parseval sur $G$ est d'autant plus grande que $v$ est grand. Dans les deux dernières lignes, on remarque en particulier que si on peut s'approcher de cette borne alors on sait dépasser la borne de Patterson-Wiedemann. Pour les deux premières lignes, une construction explicite sera nécessaire.

**Proposition 5.5** *Si* $R_{47}(115) < 1.256\sqrt{\frac{2^{115}-1}{47}}$ *ou encore si* $R_{79}(273) < 1.682\sqrt{\frac{2^{273}-1}{79}}$*, alors il existe des fonctions PW-généralisées battant la borne de Patterson-Wiedemann.*

En particulier, la deuxième condition est moins forte que $\sqrt{2}\sqrt{\#G}$ (équivalent numérique de la borne quadratique).

**Problème 5.2** *Même si ces conditions semblent réalistes, il reste à déterminer une majoration précise de* $R_{47}(115)$ *et de* $R_{79}(273)$*.*

| $(v;\ m=\frac{v-1}{2})$ | $\tau_{\mathbb{F}_{2^m}}(\chi)$ | $r$ | $\|\|\hat{g}\|\|_{L^\times/G}$ | $\Delta_{v,mr}$ | $A_{v,mr}+\sqrt{2^{mr}}$ | $\sqrt{2\ln 2}\sqrt{mr\frac{2^{mr}-1}{v}}$ | $\sqrt{\frac{2^{mr}-1}{v}}$ |
|---|---|---|---|---|---|---|---|
| $(7;3)$ | $-1\pm i\sqrt{v}$ | 13 | $\simeq 1.005\sqrt{2^{39}}$ | $\simeq 1.67^\circ$ | $\lesssim 1.033\sqrt{2^{39}}$ | $\simeq 2.779\sqrt{2^{39}}$ | $\simeq 0.378\sqrt{2^{39}}$ |
| $(23;11)$ | $-3\pm i\sqrt{v}$ | 3 | $\simeq 1.016\sqrt{2^{33}}$ | $\simeq 12.17^\circ$ | $\lesssim 1.487\sqrt{2^{33}}$ | $\simeq 1.410\sqrt{2^{33}}$ | $\simeq 0.209\sqrt{2^{33}}$ |
| $(47;23)$ | $-9\pm i\sqrt{v}$ | 5 | $\simeq 1.010\sqrt{2^{115}}$ | $\simeq 12.98^\circ$ | $\lesssim 1.760\sqrt{2^{115}}$ | $\simeq 1.842\sqrt{2^{115}}$ | $\simeq 0.146\sqrt{2^{115}}$ |
| $(79;39)$ | $7\pm i\sqrt{v}$ | 7 | $\simeq 1.004\sqrt{2^{273}}$ | $\simeq 4.88^\circ$ | $\lesssim 1.374\sqrt{2^{273}}$ | $\simeq 2.189\sqrt{2^{273}}$ | $\simeq 0.113\sqrt{2^{273}}$ |

Tab. 5.7 – Meilleure extension (degré $r$ inférieur à 15) en fonction de l'indice $v$

**Remarque 5.8** *L'intérêt de cette approche se situe également dans le calcul des sommes de Gauss. On se passe en effet du calcul de base (i.e. en partant de la définition* $\tau_L(\chi)=\sum_{x\in L^\times}\chi(x)\mu_L(x)$*) puisque les sommes de Gauss sont des éléments du corps quadratique imaginaire* $\mathbb{Q}(\sqrt{-v})$ *de norme connue. On se limite ainsi à la résolution d'une équation de norme, ce qui permet de travailler en très grande dimension (dans la table 5.7, on voit par exemple apparaître une dimension de base de* 79 *et une extension de dimension* 273 *sur* $\mathbb{F}_2$*).*

*Sous Magma, la fonction de résolution de ce type d'équation (implémentation de l'algorithme de Cornacchia – cf. [Coh93]), combinée à la fonction de calcul du nombre de classes, nous a ainsi permis de parcourir et de calculer très rapidement les sommes de Gauss pour tous les v vérifiant la condition de résiduosité quadratique et inférieur à* 99*, et pour toutes les extensions de degré inférieur à* 15 *(on explore alors des dimensions pouvant dépasser* 1000*).*

Finalement, on s'aperçoit qu'il serait intéressant de continuer à étudier cette situation où les sommes de Gauss sont quadratiques pour des indices plus grands. Pour des petits indices, cette dernière construction nous a en effet permis d'obtenir un exemple de fonction de type Patterson-Wiedemann généralisé hautement non-linéaire en dimension 15 et d'établir des conditions réalistes (cf. Proposition 5.5) pour dépasser la borne de Patterson et Wiedemann. Ces résultats confortent, selon nous, la conjecture de Patterson et Wiedemann.

# Chapitre 6

# Synthèse

L'objet principal de cette partie était d'étudier la construction de fonctions hautement non-linéaires s'approchant de la conjecture de Patterson et Wiedemann. Pour cela, nous proposons une généralisation de la construction de type Patterson-Wiedemann (fonctions constantes sur les cosets d'un sous-groupe) en autorisant plus de liberté sur le coset trivial. Nous étudions la non-linéarité de cette construction et introduisons alors les deux problèmes sous-jacents : le calcul de sommes de Gauss et leurs répartitions dans le corps des nombres complexes, ainsi que l'estimation de sommes d'exponentielles incomplètes.

Dans le chapitre 2, nous décrivons ces constructions et les conditions à atteindre pour se rapprocher d'une amplitude spectrale $\sqrt{2^m}$ sur $\mathbb{F}_{2^m}$. Soit $G$ un sous-groupe de $L^\times = \mathbb{F}_{2^m}^\times$ d'indice $v$, supposons :

1. que pour tout $\chi \perp G, \chi \neq 1$, $\tau_L(\chi) \approx \epsilon\sqrt{2^m}$ $(\epsilon \in \{\pm 1\})$ ;
2. et qu'il existe une fonction binaire $h : G \to \{\pm 1\}$ tel que ses coefficients de Fourier sur $G$ vérifient $\|\tilde{h}_G\|_{G\cup\{0\}} \leq \sqrt{2^m} + o(\sqrt{2^m})$ et $\|\tilde{h}_G\|_{L^\times - G} << \sqrt{2^m}$ ;

alors une construction de type Patterson-Wiedemann généralisée selon $G$ et $h$ battrait la borne de Patterson-Wiedemann. On démontre de plus que la première hypothèse est réalisable dans le cas $m$ pair et approchable dans le cas $m$ impair par passage à la limite.

Dans le chapitre 3, nous introduisons la notion générale de non-linéarité partielle, liée à l'étude de la deuxième hypothèse, et montrons que le comportement asymptotique peut être équivalent à celui de la non-linéarité classique. En particulier, la non-linéarité partielle d'une fonction est asymptotiquement presque sûrement élevée. Ce résultat permet notamment de prouver l'existence de fonctions de type Patterson-Wiedemann généralisé de non-linéarité asymptotique plus élevée que la moyenne.

Dans le chapitre 4, nous approfondissons l'étude du rayon spectral partiel $R_v(m)$ pour un sous-groupe multiplicatif d'indice $v$ de $\mathbb{F}_{2^m}^\times$ où nous démontrons entre autre que la borne supérieure asymptotique est applicable en dimension finie. D'après le théorème 4.1, on en déduit par exemple que la deuxième hypothèse ci-dessus est automatiquement vérifiée lorsque $m = o(v)$. Ce résultat souligne l'intérêt d'étudier le comportement de $R_v(m)$ : c'est un problème bien plus général que la conjecture de Patterson-Wiedemann mais pour lequel nous ne cherchons pas nécessairement un résultat optimal.

Dans le chapitre 5, nous décrivons des applications concrètes de cette étude. Nous expliquons comment choisir des sommes de Gauss proches de la même valeur, en observant leur écart angulaire, afin de s'approcher de la première hypothèse. Nous appliquons les résultats à la construction de fonctions équilibrées de type Patterson-Wiedemann. Et nous détaillons une construction particulière, dite par résiduosité quadratique, où le calcul des sommes de Gauss est possible et pour laquelle une fonction hautement non-linéaire, battant la borne quadratique en dimension 15, est obtenue. De plus, nous établissons sur des exemples numériques des conditions sur $R_v(m)$ pour dépasser la borne de Patterson-Wiedemann (cf. Proposition 5.5).

**Perspectives.** Les difficultés rencontrées dans cette partie sont principalement liées au manque de connaissance sur le rayon spectral partiel $R_v(m)$ et au problème d'estimation des sommes de Gauss. Ce travail soulève ainsi de nouvelles interrogations et l'étude de problèmes ouverts dont certains ont été abordés au cours des précédents chapitres ; d'autres sont expliqués ci-après.

Le comportement de $R_v(m)$ doit être mieux compris pour aboutir à d'autres constructions hautement non-linéaires. En particulier, on aimerait savoir précisément à quelle distance de la borne de Parseval nous sommes. Par analogie avec le cas du corps tout entier, et selon les analyses présentées dans cette partie, nous pensons que la conjecture de Patterson-Wiedemann pourrait se généraliser ainsi :

**Conjecture 6.1** *Soit $v$ un entier impair, alors pour $m$ grand tel que $v \mid 2^m - 1$,*

$$R_v(m) \simeq \sqrt{\frac{2^m}{v}}.$$

D'après les exemples de construction étudiés, on peut même préciser que si $R_v(m) = O(\sqrt{\frac{2^m-1}{v}})$ (c'est le cas pour $v = 1$ d'après la borne quadratique), alors il suffirait de trouver des sommes de Gauss proches de $\epsilon\sqrt{2^m}$ pour $\frac{1}{v} = o(1)$ afin de s'approcher de la conjecture de Patterson et Wiedemann.

Un autre problème de recherche intéressant est le suivant. Pour $m$ impair, si on observe l'ordre maximal d'un sous-groupe de $\mathbb{F}_{2^m}^\times$ qui soit inférieur à $\sqrt{2^m}$, on réalise numériquement que celui-ci peut être très proche de $\sqrt{2^m}$. Notons $\delta(m) = \max_{d|2^m-1, d\leq\sqrt{2^m}} \frac{d}{\sqrt{2^m}}$ alors $\delta(m)$ peut prendre une valeur proche[1] de 1 comme l'illustre la table 6.1. On y retrouve en particulier le cas $m = 15$ correspondant au contexte du contre exemple de Patterson-Wiedemann.

Dans ces chapitres, nous avons rencontré plusieurs cas qui soulignent l'intérêt de ce problème. Nous avons étudié le cas pair (cf. Sec. 2.2.2) pour lequel $\delta(m) \simeq 1$ et où le sous-groupe d'ordre $\delta(m)\sqrt{2^m}$ donne des sommes de Gauss rationnelles. Étudier le comportement

[1] En revanche, même si $\limsup_{m\to+\infty,\, m \text{ impair}} \delta(m)$ a de bonnes chances de converger vers 1, il n'existe aucun outil pour envisager une preuve. Des résultats existent toutefois pour les diviseurs d'entiers quelconques (voir entre autre les travaux de Tenenbaum ou de Ford). On peut également essayer de s'accorder un peu plus de marge comme le problème abordé dans [LeB02].

| $m$ | 75 | 55 | 47 | 81 | 39 | 45 | 15 | 63 | 43 | 79 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\delta(m)$ | 0.9872 | 0.9374 | 0.8944 | 0.8770 | 0.8727 | 0.8528 | 0.8342 | 0.7652 | 0.7080 | 0.6982 |

Tab. 6.1 – Les dix plus grandes valeurs de $\delta(m)$ pour $m \leq 85$

des sommes de Gauss dans une situation approchante pour $m$ impair serait également intéressant. Plus simplement pour $m$ impair, en notant $q = 2^m$, nous avons également vu que si les sommes de Gauss (sur un caractère non-trivial orthogonal au sous-groupe) sont proches de $\sqrt{q}$, alors une construction de type PW permet d'atteindre une non-linéarité élevée si l'ordre est inférieur à $\sqrt{q}$ et si $v$ est grand mais de taille raisonnable (dépendante de l'écart des sommes de Gauss par rapport à $\sqrt{q}$). De même, une fonction de type PW généralisé sera hautement non-linéaire si $v$ reste de taille raisonnable et si son représentant sur $G$ a une amplitude petite devant $\sqrt{q}$. Le choix de $v$ et $N$ proches de $\sqrt{q}$ apparaît comme un très bon compromis : en particulier, sous ces conditions le théorème 4.1 est suffisant pour assurer une faible amplitude sur $G$. Le problème de recherche qui se pose est donc le suivant.

**Problème 6.1** *1. Est-ce que* $\limsup_{m\to+\infty,\, m \text{ impair}} \delta(m)$ *est égale à 1 ?*

*2. Soit m impair et* $d(m)$ *le plus grand diviseur de* $2^m - 1$ *inférieur à* $\sqrt{2^m}$*. Que peut-on dire des sommes de Gauss (en particulier de leur argument) sur un sous-groupe d'indice* $d(m)$ *quand* $\delta(m) = \frac{d(m)}{\sqrt{2^m}} \simeq 1$ *?*

Et ceci est lié à un problème de recherche général : l'élaboration d'un algorithme efficace pour le calcul des sommes de Gauss ou de leurs arguments. En effet, pour $m$ grand, il est impossible d'utiliser la formule de base $\sum_{x\in\mathbb{F}_{2^m}^\times} \chi(x)\mu(x)$. Il faudrait par exemple s'intéresser à la formule de Gross-Koblitz [GK79] où on peut envisager d'obtenir une estimation des sommes, dérivée d'un développement dyadique d'ordre fini.

Nous avons porté notre attention sur des situations s'approchant du cas où les sommes de Gauss ont la même valeur et sur l'étude de la non-linéarité sur un sous-groupe d'indice $v$. Toutefois d'autres pistes sont ouvertes. Un point que nous n'avons pas étudié de près concerne un choix de cosets adéquat. Un choix de $s$ tel que l'amplitude de sa transformée de Fourier est de l'ordre de $\sqrt{v}$ semble opportun (c'est en particulier le cas dans la construction de type RQ de la section 5.3). Mais ce point pose alors des problèmes similaires à la conjecture qui nous intéresse. Il serait certainement très intéressant d'étudier une solution où les arguments des coefficients de Fourier de $s$ seraient pris en adéquation avec les valeurs des sommes de Gauss pour combiner efficacement ces sommes même quand elles ne sont pas proches de $\sqrt{q}$. Nous sommes alors de nouveau sur des problématiques s'approchant de l'étude des séries de Fourier aléatoires et l'étude de la norme de polynômes trigonométriques à coefficients de module 1, comme par exemple dans les travaux de [Kah85, SZ54].

# Deuxième partie

# Cryptographie

# Chapitre 7

# Exemples d'application des fonctions booléennes et de la théorie des codes correcteurs d'erreurs en cryptographie

La réalisation de protocoles cryptographiques, domaine de recherche relativement récent, entraîne l'utilisation de nombreux objets et théories mathématiques, principalement dans le domaine des mathématiques discrètes. Les fonctions booléennes et la théorie des codes font ainsi partie des "outils" intéressants dans ce domaine, et leur étude est depuis quelques années fortement liée aux contraintes rencontrées en cryptographie. Ainsi, comme expliqué au chapitre 1, des critères spécifiques ont été définis pour choisir convenablement les fonctions booléennes au coeur d'algorithmes cryptographiques tels que les boites S pour le chiffrement par blocs ou les filtres pour du chiffrement à flot. En ce qui concerne les codes (autres que ceux associés aux fonctions booléennes), liés par définition à la théorie de l'information (pour une finalité en général différente de la cryptographie), plusieurs études ont déjà démontré leur importance en cryptographie : La difficulté du décodage général ou de la recherche d'un mot de code de poids minimum dans un code linéaire quelconque (problème NP-complet [BMvT78]) suscite notamment beaucoup d'intérêt et de nombreuses constructions basées sur ce problème ont été proposées, par exemple les cryptosystèmes à clé publique de McEliece [McE78] et de Niederreiter [Nie86] ou différents schémas d'authentification à partir du problème connexe de *syndrome decoding* (SD) tels que [Ste94] ou [Vér96].

Dans ce chapitre, nous nous intéressons à des applications cryptographiques utilisant la théorie des codes ou des fonctions booléennes et nous illustrons à travers trois exemples différents une partie de la diversité des utilisations possibles en cryptographie. Chaque exemple est associé à une publication pour laquelle une présentation du sujet et un résumé sont effectués avant l'article en lui-même.

Dans un premier temps, deux applications dans le cadre de la sécurisation des échanges entre une étiquette électronique et un récepteur sont détaillées. Le premier article [BCD06b] propose, via l'utilisation d'une fonction booléenne bien choisie, une amélioration d'un proto-

cole d'authentification dynamique afin de résister contre des attaques par le milieu. Le deuxième article [BC06] se restreint à un mode d'authentification passif où l'introduction d'un codage spécifique permet d'augmenter la sécurisation de l'authentification d'une étiquette électronique n'ayant presque aucune capacité de calcul. Enfin, une application dans le cadre de l'authentification à partir de données variables dans le temps est étudiée où l'ajout d'une variation autour du cryptosystème de McEliece permet de restreindre les droits d'accès à la fonction d'authentification afin d'éviter les compromissions des données en jeu (cf. [BCD06a]).

Le lecteur pourra remarquer que chaque article présente une amélioration d'une construction existante grâce à une adaptation plus ou moins directe de résultats connus dans le domaine des fonctions booléennes ou de la théorie des codes, participant ainsi à la mise en valeur de leurs intérêts pour la cryptographie.

## 7.1 RFID et cryptographie

La problématique de sécurisation des échanges faisant intervenir des équipements à faibles ressources nécessite de trouver des solutions appropriées en terme de souplesse et de surcoût. En particulier, le cas de la Radio Frequency IDentification (RFID) a incité de nombreux chercheurs à travailler sur ce problème technique. Cette technologie fait intervenir 3 éléments, des transmetteurs (tags), un récepteur (ou lecteur) et une base de données, les 2 premiers éléments communiquent par ondes radio à courte portée ; pour établir cette communication, le problème de la sécurité se pose alors. Les tags suivent la norme EPC [EPC] qui distinguent 3 classes de composants, la plus adaptée aux primitives cryptographiques usuelles correspond aux cartes à puces sans contact (classe 3) alors que les composants des classes 1 et 2 sont des composants à très faibles ressources avec peu de mémoire et très peu de puissance calculatoire. Ainsi, les tags sont très différents d'une classe à une autre ($<$ 1 kbit à plusieurs Mbits pour des prix très faibles, moins de 0.10 € à plusieurs euros). Les tags de classes 1 ou 2 sont passifs (énergie fournie par le lecteur), ceux de classe 1 se comportent comme des mémoires avec quelques centaines de bits et ceux de classe 2 disposent généralement de plus de mémoire (éventuellement en partie volatile) et d'une petite puissance de calcul. La principale contrainte sur ces tags (et sur leurs ressources) pour leur utilisation étant de conserver des prix très faibles, ce sont alors des cibles difficiles pour des algorithmes cryptographiques usuels, trop lourd à implémenter en général, et par là même une motivation à développer de nouveaux protocoles à faibles ressources. En particulier, un lecteur étant amené à communiquer avec de nombreux tags, chaque tag est associé à un unique identifiant $ID$ et il est nécessaire de trouver des solutions permettant une authentification rapide de ces tags plus sûre que le simple envoi de $ID$ (cf. Fig. 7.1). Le lecteur peut se référer à [Jue06] pour disposer d'une vue d'ensemble de ces problèmes.

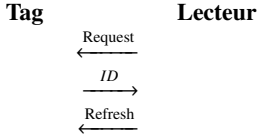


FIG. 7.1 – Exemple d'échange classique et non sécurisé entre un tag et un lecteur

Nous verrons dans la section 7.1.2 une solution susceptible de s'appliquer pour la classe 1 et nous présentons dans la section suivante un protocole pour les composants de la classe 2.

### 7.1.1 HB$^{++}$ : a Lightweight Authentication Protocol Secure against Some Attacks

Dans cet article, nous proposons une amélioration du protocole d'authentification à faibles ressources HB$^{+}$, décrit par Juels et Weis à Crypto'05 [JW05], afin de résister à certaines attaques par le milieu.

En 2001, Hopper et Blum [HB01] ont présenté un protocole d'authentification "léger", dans le but d'être mis en oeuvre directement par des êtres humains, composé de plusieurs itérations successives de la forme suivante – voir Fig. 7.2.

| **Tag** $(x)$ | | **Lecteur** $(x, \eta)$ |
|---|---|---|
| $\nu \in \{0,1 \vert \mathbb{P}(\nu = 1) = \eta\}$ | | |
| | $\xleftarrow{a}$ | Challenge aléatoire $a \in_R \{0,1\}^k$ |
| Calculer $z = a \cdot x \oplus \nu$ | $\xrightarrow{z}$ | Vérifier si $a \cdot x \approx_\eta z$ |

FIG. 7.2 – Une itération du protocole HB

Ici $x$ est un secret de $k$ bits servant à authentifier le tag, celui-ci retourne un résultat correct ou non à chaque itération, et le lecteur l'authentifie au final si la proportion de résultats justes sur l'ensemble des itérations respecte une certaine loi de probabilité. La sécurité du protocole HB ne repose ainsi pas sur une solution classique de chiffrement symétrique mais sur la difficulté du problème calculatoire LPN (Learning Parity with Noise) [BFKL93, BKW00, Hås97].

**Définition 7.1 (problème LPN)** *Soient $A$ une matrice binaire aléatoire de taille $q \times k$, $X$ un vecteur aléatoire de $k$ bits, $\eta$ un paramètre (constant) de bruit, et $\vec{\nu}$ un vecteur aléatoire de $q$ bits tel que $w_H(\vec{\nu}) \leq \eta q$.*

*Étant donnés $A$, $\eta$, et $\vec{z} = AX \oplus \vec{\nu}$, trouver un vecteur de $k$ bits $X'$ tel que $w_H(AX' \oplus \vec{z}) \leq \eta q$.*

En résumé, l'idée est donc pour le tag d'envoyer une version bruitée de $A \cdot x$ pour une matrice aléatoire $A$ au lecteur, selon un biais pré-déterminé, de sorte que ce-dernier puisse vérifier ce biais grâce à la connaissance de $A$ et $x$ mais tel qu'un attaquant interceptant les données sur le canal de communication ne peut déterminer $x$. Les algorithmes de résolution du problème sous-jacent étant de complexité exponentielle (voir par exemple les articles récents [FMI$^{+}$06, LF06]), le protocole obtenu est considéré comme sûr contre des attaquants dits passifs. Le problème LPN peut d'autre part être vu comme le cas moyen du problème SD (Syndrom Decoding) qui est connu comme un problème NP-complet depuis [BMvT78], il est même difficile à approcher (cf. [Hås97]) (le problème SD est à l'origine de l'intérêt de l'utilisation des codes en cryptographie asymétrique, en particulier du cryptosystème de McEliece dont nous reparlerons section 7.2).

Dans [JW05] (voir également [Wei06] pour plus de détails), Juels et Weis proposent un schéma d'authentification à faibles ressources, le protocole HB$^{+}$, construit comme une amélioration du protocole HB afin de résister aux attaques actives. L'idée est principalement de

faire dépendre le calcul de $z$ de 2 données aléatoires, de manière symétrique, une choisie par le lecteur et l'autre choisie par le tag. Cependant, le modèle de sécurité choisi ne permet pas de couvrir toutes les attaques actives. Il ne tient en particulier pas compte de l'information supplémentaire donnée par le résultat de l'authentification (réussite ou non) ; Gilbert *et al.* ont ainsi décrit dans [GRS05] une attaque par le milieu sur $HB^+$ très efficace (linéaire en la taille du secret) en utilisant cette information. Cette attaque exploite la linéarité des opérations en modifiant à chaque itération le challenge envoyé par le lecteur via une translation fixe afin de retrouver au fur et à mesure la valeur du secret.

Notre travail [BCD06b] a alors été motivé par cette attaque, dans un souci d'améliorer le protocole $HB^+$ afin de se prémunir de ces attaques par le milieu tout en conservant le même principe et les mêmes contraintes d'implémentation dans des composants à faibles ressources. Ce nouveau protocole, $HB^{++}$, peut en quelque sorte être vu comme deux exécutions parallèles de $HB^+$ avec des challenges corrélés. Dans ce contexte, nous allons voir comment un des critères utilisés sur les fonctions booléennes pour la définition de primitive de chiffrement symétrique par blocs, critère de sécurité contre la cryptanalyse différentielle, peut être considéré pour casser la linéarité des opérations afin de parer l'attaque par le milieu de [GRS05].

Nous construisons un exemple de fonction booléenne à utiliser dans le protocole en tenant compte des restrictions de ressources, à partir de fonctions puissances de type Gold prises via un découpage du corps $\mathbb{F}_{2^k}$ en petits sous-corps (ces sous-fonctions sont alors quadratiques et facilement implémentables).

D'autre part, afin d'empêcher un attaquant de gagner de l'information d'une authentification à une autre via la connaissance du résultat de plusieurs authentifications, nous introduisons un renouvellement des secrets à chaque utilisation du protocole. Ceci permet de prendre en compte une plus large classe d'attaques par le milieu. Il reste cependant à déterminer s'il est possible de fixer les différents paramètres afin de se prémunir de toutes les attaques par le milieu (on notera que la deuxième attaque décrite dans [BCD06b] est proche d'une attaque générique par le milieu).

Ainsi, on peut également voir le protocole $HB^{++}$ comme un candidat potentiel au problème d'extension du protocole $HB^+$ en un protocole résistant à toutes les attaques actives (attaques par le milieu y compris).

Après l'article, nous détaillons dans la section 7.1.1.1 page 130 l'analyse de sécurité du protocole proposé.

# HB$^{++}$: a Lightweight Authentication Protocol Secure against Some Attacks

Julien Bringer Hervé Chabanne Emmanuelle Dottax
Sagem Défense Sécurité

**Abstract**

At Crypto'05, Juels and Weis introduce HB$^+$, an enhancement of the Hopper and Blum (HB) authentication protocol. This protocol HB$^+$ is proven secure against active attacks, though preserving HB's advantages: mainly, requiring so few resources to run that it can be implemented on an RFID tag. However, in a wider adversarial model, Gilbert, Robshaw and Sibert exhibit a very effective attack against HB$^+$.

We here show how a modification of the HB$^+$ protocol thwarts Gilbert et al's attack. The resulting protocol, HB$^{++}$, remains a good candidate for RFID tags authentication.

**Keywords.** HB$^+$ protocol, active attacks, man-in-the-middle attacks, RFID.

## 1 Introduction

The problems of security and privacy for Radio Frequency Identification (RFID) have recently attracted many technical research.

RFID systems are made of three components: some tags, a reader, and a database which contains information on the tagged objects. Tags (transponders) follow the ISO and EPC [8] standards and communicate with the reader (transceiver) over the air. One main constraint here is that these tags have to be quite inexpensive (the order of magnitude is US cents) and thus they can embed only scarce resources, of which only some part is dedicated to security. Typically, computations are hardwired and some thousands of logic gates are kept for cryptography. This means that tags seem, at first glance, difficult targets for the implementation of classical cryptographic schemes, even if Feldhofer, Dominikus and Wolkerstorfer [9] have described an implementation of the AES algorithm which looks promising. Anyway,

the introduction of new cryptographic schemes, requiring less resources, is today tempting.

In the typical setting, each tag comes with a unique identifier and an adversary should not be able to counterfeit tag responses. Many authentication protocols for RFID tags have been proposed so far (see e.g. in 2003 [18, 26], [12, 13, 16, 24] in 2004, [1, 3, 7, 22] in 2005, see also Juels [17] for a general survey and [2] for fresh references). Notably, at Crypto'05, $HB^+$, a lightweight cryptographic authentication scheme very well suited for low-cost hardware implementation, was introduced by Juels and Weis [19]. It provides a symmetric-key protocol allowing tags to identify themselves on the reader (the reader does not need to know a priori which tags and secrets are involved for the protocol to work). $HB^+$ is presented as an improvement of the HB protocol, which had been introduced in [14]. The security of the HB protocol does not rely on classical symmetric key cryptography solutions, but rather on the hardness of the computational Learning Parity with Noise (LPN) problem [4, 5, 15]. While the HB protocol is made to be secure against passive attacks only, the aim of $HB^+$ is to be resistant to active attacks. A proof of security is provided but at the same time, Gilbert, Robshaw and Sibert [10] describe a man-in-the-middle attack on $HB^+$ not covered by the corresponding security model.

The principal contribution of our work is to improve the $HB^+$ protocol in order to avoid the attacks of [10] and [25], while keeping its design principles and, thus, its advantages. We call $HB^{++}$ our new protocol. In fact, $HB^{++}$ can be seen as running $HB^+$ twice under independent secrets but with correlated challenges. Moreover, the secrets are renewed at each authentication. Two functions are shared by all the tags and readers; one is introduced to link together challenges of the protocol, the other is needed to determine secrets used for an authentication. At the end, the $HB^{++}$ protocol seems to us a good susbtitute for $HB^+$ for RFID tags authentication.

The paper is organised as follows. In Sect. 2, we recall the $HB^+$ protocol. In Sect. 3, we summarize Gilbert et al's attack [10]. In Sect. 4, we introduce a first construction of $HB^{++}$ and show that it is at least as secure as $HB^+$. In Sect. 5, we study its resistance against an extended version of the attack of [10]. In Sect. 6, we describe a more general man-in-the-middle attack by Wagner [25] and show how to thwart this point. We then draw the final design of our protocol with, as last improvement, a renewal of secrets at each authentication. Section 7 concludes.

## 2 The $\mathrm{HB}^+$ protocol

A brief description of one round of the $\mathrm{HB}^+$ protocol is given by Fig. 1 where $a \cdot x$ stands for the scalar product of the binary vectors $a$ and $x$, and $\oplus$ is the exclusive or.

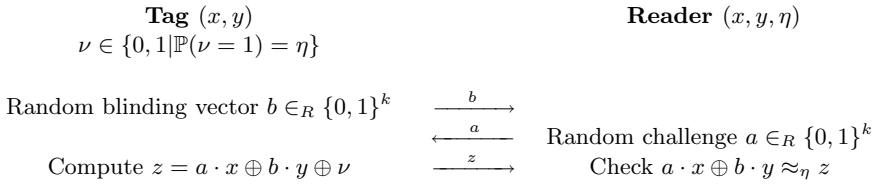


Figure 1: One round of $\mathrm{HB}^+$

The two $k$-bit vectors $x$ and $y$ are secret keys shared by the tag and the reader. Note that an extra noise is added to the response $a \cdot x \oplus b \cdot y$ by the tag, this error bit $\nu$ equals 1 with probability $\eta$.

The $\mathrm{HB}^+$ round described by Fig. 1 is repeated $r$ times and the tag is successfully authenticated if the check fails about $\eta r$ times (this is what is denoted by $\approx_\eta$ in Fig. 1 and in the following).

**Remark 1** *The principal difference between the* $\mathrm{HB}^+$ *and* HB *protocols is the introduction of* $y$ *and* $b$ *in the* $\mathrm{HB}^+$ *protocol in order to avoid active attacks.*

In [19], the authors define a security model, and then show how to reduce an attack on $\mathrm{HB}^+$ to an attack on HB.

The security of the HB protocol is based on the Learning Parity with Noise (LPN) problem. Juels and Weis extend this result in their security model to $\mathrm{HB}^+$ and explain how an attack on $\mathrm{HB}^+$ can be used to solve an instance of the LPN problem (see Sect. 4.1 for an extension to our ideas).

Unfortunately, they do not take into account the extra information given by the result (positive or negative) of the protocol and this is exploited during the attack [10] (see Sect. 3).

## 3 A man-in-the-middle attack against $\mathrm{HB}^+$

In [10], an attack is described against the $\mathrm{HB}^+$ protocol. It is a linear-time man-in-the-middle attack where an adversary located between the reader and the tag is able to modify the challenge at every round. The adversary

chooses a vector $\delta$ in $\{0,1\}^k$ and when a challenge $a$ is sent by the reader, he intercepts the challenge and makes a switch to $a+\delta$ (see Fig. 2). Hence, at the end of the round, the reader will receive $\tilde{z} = (a+\delta)\cdot x \oplus b\cdot y \oplus \nu$ from the tag.

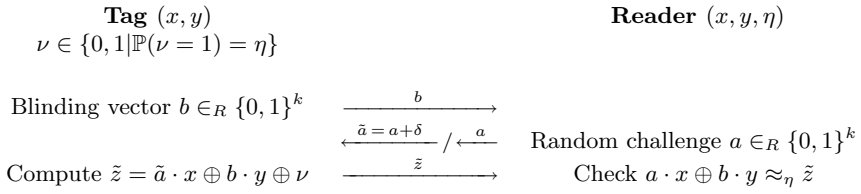


Figure 2: An effective attack against $\mathrm{HB}^+$

This is repeated along all the rounds in order to deduce information from the success or failure of the authentication. Indeed, if the authentication succeeds (resp. fails), we have $\delta \cdot x = 0$ (resp. $\delta \cdot x = 1$) with a high probability. So one can recover $x$ "bit after bit" by varying $\delta$ progressively.

**Remark 2** *Similarly, an adversary can send $b+\delta$ instead of $b$ to the reader, in order to recover $y$.*

# 4 A first attempt

## 4.1 Proposed construction of $\mathrm{HB}^{++}$

The protocol $\mathrm{HB}^{++}$ needs two new secrets $x'$, $y'$ and the introduction of a permutation of the set $\{0,1\}^k$, $f$, designed as described in Sect. 5.

A round of this protocol consists then, for given challenges $(a, b)$, in computing correlated responses thanks to $f$ and for the tag to send these responses together with independent errors $\nu$ and $\nu'$, i.e. the tag sends the responses $z = a\cdot x \oplus b\cdot y \oplus \nu$ and $z' = f(a)\cdot x' \oplus f(b)\cdot y' \oplus \nu'$ (see Fig. 3).

## 4.2 Consistency with the previous security models

The model of active security standing for the $\mathrm{HB}^+$ protocol in [19] can be translated to this first construction.

**Proposition 1** *An adversary who has the capability of breaking a random sequence of challenges-responses of this first attempt of $\mathrm{HB}^{++}$ can successfully attack $\mathrm{HB}^+$.*

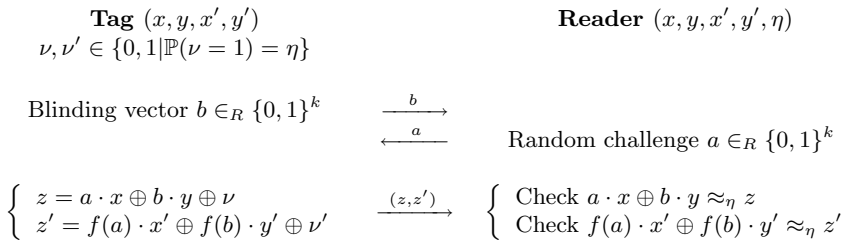


Figure 3: First attempt

*Proof.* Indeed, if an adversary $\mathcal{A}$ obtains a sequence of challenges-responses $\mathcal{S} = \{a_i, b_i, a_i \cdot x \oplus b_i \cdot y \oplus \nu_i\}_{i \in I}$ from successive rounds of the HB$^+$ protocol between a tag $\mathcal{T}_{x,y}$ and a reader $\mathcal{R}$, then, by randomly picking $x'$, $y'$ and a variable $\nu'$ such that $\mathbb{P}(\nu' = 1) = \eta$, he can simulate a sequence of challenges-responses

$$\{a_i, b_i, a_i \cdot x \oplus b_i \cdot y \oplus \nu_i, f(a_i) \cdot x' \oplus f(b_i) \cdot y' \oplus \nu'_i\}_{i \in I}$$

of successive rounds of this first attempt of the HB$^{++}$ protocol between $\mathcal{R}$ and a tag $\mathcal{T}_{x,y,x',y'}$. Thus his ability to cryptanalyse this protocol allows $\mathcal{A}$ to recover the value of $x, y, x'$ and $y'$ given a sufficiently large number of challenges-responses, and so to gain the knowledge of the secrets of the original tag $\mathcal{T}_{x,y}$. □

If $\mathcal{A}$ needs to use an active attack for this last point, the only constraint is to obtain the sequence $\mathcal{S}$ of challenges-responses by applying the same modification on $a$ and $b$ during the rounds of HB$^+$ as if he was trying the attack on the new protocol.

The reduction to the LPN problem, which ensures the security of HB and HB$^+$ against a passive attack, is always true for the new protocol.

Let $\mathrm{wt}_H$ stand for the hamming weight.

**Definition 1 (LPN problem)** *Let $A$ be a random $q \times k$ binary matrix, let $X$ be a random $k$-bit vector, let $\eta$ be a constant noise parameter, and let $\vec{\nu}$ be a random $q$-bit vector such that $\mathrm{wt}_H(\vec{\nu}) \leq \eta q$.*

*Given $A$, $\eta$, and $\vec{z} = AX \oplus \vec{\nu}$, find a $k$-bit vector $X'$ such that $\mathrm{wt}_H(AX' \oplus \vec{z}) \leq \eta q$.*

**Proposition 2** *If a "passive" adversary has the capacity of breaking this first attempt of the HB$^{++}$ protocol with 4 secrets of size $k$, he can also solve a random instance of the LPN problem of size $2k$.*

*Proof.* The adversary $\mathcal{A}$ can recover the secrets given a sufficiently large sequence.

Let $A$ be a random $q \times 2k$ binary matrix, $X$ a random $2k$-bit vector, $\vec{\nu}$ a random $q$-bit vector such that $\mathrm{wt}_H(\vec{\nu}) \leq \eta q$ and $\vec{z} = AX \oplus \vec{\nu}$. $\mathcal{A}$ can construct the $k$-bit vectors $x, y$, $a_i, b_i$ for $i \in \{1, \ldots, q\}$ such that:

$$X = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_i & b_i \\ \vdots & \vdots \\ a_q & b_q \end{pmatrix}$$

The adversary $\mathcal{A}$ can interpret $\vec{z} = (a_i \cdot x \oplus b_i \cdot y \oplus \nu_i)_{i=1\ldots q}$ as responses of the HB$^+$ protocol. As in the HB$^{++}$ protocol the errors $\nu_i$ are independant of the errors $\nu'_i$, by taking random vectors $x', y', \vec{\nu'}$ and by computing $\vec{z'} = (f(a_i) \cdot x' \oplus f(b_i) \cdot y' \oplus \nu'_i)_{i=1\ldots q}$, then $(\vec{z}, \vec{z'})$ can be viewed as responses of the new protocol which allows $\mathcal{A}$ to recover $X = \begin{pmatrix} x \\ y \end{pmatrix}$. □

# 5 Protection against Gilbert et al.'s attack

## 5.1 Choice of $f$

We primarily choose $f$ in order to thwart the attack presented in [10] but $f$ has also to be taken with a low complexity and must not desequilibrate the distribution of scalar products.

As $f$ is taken as a bijection, the last point is always true, the distribution of values does not change:

$$\forall x \in \{0,1\}^k,\ \mathbb{P}(c \in \{c | f(c) \cdot x = 0\}) = \mathbb{P}(c \in \{c | c \cdot x = 0\}).$$

Henceforth, we focus on the first point. In order to avoid the attack [10], $f$ is chosen such that $\Delta_f$ is small with:

$$\Delta_f = \max_{\delta \neq 0, \gamma} |\{a \in \{0,1\}^k | f(a+\delta) + f(a) = \gamma\}|.$$

In fact, this comes to force $f$ to respect only a small number of linear relations, such that ultimately no linear relation holds for all the rounds.

**Definition 2** *Let $f : \mathbb{F}_2^k \to \mathbb{F}_2^k$ be a vectorial boolean function, for $u \neq 0, v \in \mathbb{F}_2^k$, let*

$$\delta_f(u,v) = \big|\{a \in \{0,1\}^k | f(a+u) + f(a) = v\}\big|.$$

**Remark 3** *This $\delta_f(u,v)$ has been introduced in [6] to measure the resistance of an S-box against differential cryptanalysis. We have*

$$\Delta_f = \max_{u\neq 0,v} \delta_f(u,v).$$

*And for instance, the lower the value $\Delta_f$ will be, the more resistant against differential cryptanalysis the function $f$ will be.*

We will see in the sequel how $\Delta_f$ can be used to measure the resistance against the attack of Gilbert et al.

### 5.2 Resistance of the protocol

One can try to extend the attack [10] by corrupting $(a,b)$ with $G(a,b) = (g_1(a,b), g_2(b))$, where $G \neq Id$, and sending $g_2(b)$ to the reader and $g_1(a,b)$ to the tag such that the reader will check if

$$\begin{cases} g_1(a,b)\cdot x \oplus b\cdot y \oplus \nu & \approx_\eta \quad a\cdot x \oplus g_2(b)\cdot y \\ f(g_1(a,b))\cdot x' \oplus f(b)\cdot y' \oplus \nu' & \approx_\eta \quad f(a)\cdot x' \oplus f(g_2(b))\cdot y' \end{cases}$$

i.e. if

$$\begin{cases} (g_1(a,b)+a, b+g_2(b))\cdot(x,y)\oplus\nu & \approx_\eta 0 \\ (f(g_1(a,b))+f(a), f(b)+f(g_2(b)))\cdot(x',y')\oplus\nu' & \approx_\eta 0 \end{cases} \quad (1)$$

Fortunately, an adversary does not know the result of this comparison but only the result of the authentication which depends on the results of all the $r$ rounds of the protocol. So, if one wants to obtain some information on the secrets via this method, (1) has to be independent of $a$ and $b$. We suppose also that an adversary has no knowledge of $x, y, x'$ and $y'$ and so they have to be considered as random vectors. In consequence, to achieve an attack, $\delta_1^{(x,y)}, \delta_2^{(x,y)}, \lambda_1^{(x',y')}$ and $\lambda_2^{(x',y')}$ have to be chosen such that the following equalities stand for all the $r$ rounds:

$$\begin{cases} g_1(a,b) &= a + \delta_1^{(x,y)} \\ g_2(b) &= b + \delta_2^{(x,y)} \\ f(g_1(a,b)) &= f(a) + \lambda_1^{(x',y')} \\ f(g_2(b)) &= f(b) + \lambda_2^{(x',y')} \end{cases}$$

If $\{(a_i, b_i)\}_{i=1..r}$ is the set of all the values used during the $r$ rounds, those equalities induce two linear relations involving $f$: $\forall i \in \{1, \ldots, r\}$,

$$\begin{aligned} f(a_i + \delta_1^{(x,y)}) + f(a_i) &= \lambda_1^{(x',y')}, \\ f(b_i + \delta_2^{(x,y)}) + f(b_i) &= \lambda_2^{(x',y')}. \end{aligned}$$

As $\Delta_f = \max_{\delta \neq 0, \gamma} \left| \{a \in \{0,1\}^k | f(a+\delta) + f(a) = \gamma\} \right|$ is small, these relations are verified during all the rounds only with a small probability. So it is possible to deduce something on the secrets from the success or failure of the authentication only with a small probability $\mathbb{P}$, which verifies:

$$\mathbb{P} \leq \left( \frac{\Delta_f}{2^k} \right)^r.$$

Consequently, the smaller $\Delta_f$ is, the smaller $\mathbb{P}$ is, and we have thus the following criterion for candidate functions $f$:

**Criterion 1** *The security of the* HB$^{++}$ *protocol against generalizations of the active attack described in [10] is ensured whenever the function $f$ satisfies the following property: $\Delta_f$ is small enough such that $\left( \frac{\Delta_f}{2^k} \right)^r$ is negligible.*

An example of construction of function $f$ is given for realistic parameters in Appendix A.1.

# 6 The final design

## 6.1 Another man-in-the-middle attack due to Wagner [25]

This first construction remains sensitive to an attack due to Wagner where the idea is to modify both the challenges sent by the reader and the responses received from the tag along one authentication. For understanding concern, we describe the method on a particular case.

Assume that a large part of the output of function $f$ is independant of a small part of its input, for instance, say that all the bits of $f(a)$ but the first three bits can be computed without the first three bits of $a$. An adversary can then try to find the value of, say, the first three bits of $x$ and $x'$ as follows:

1. he makes a guess for these six bits;
2. for each challenge $a$, he tries to choose $\delta$ such that:

(a) all its bits are 0, except the first three ones that can be 0 or 1,

(b) the same holds for $\delta' = f(a \oplus \delta) \oplus f(a)$, i.e.

$$\delta = (*, *, *, 0, \ldots, 0), \quad \delta' = (*, *, *, 0, \ldots, 0);$$

3. he replaces the reader's challenge $a$ by $a \oplus \delta$ and the tag responses $z, z'$ by $z \oplus \delta \cdot x$ and $z' \oplus \delta' \cdot x'$, respectively;

4. at the end of the protocol, if the adversary has succeeded in constructing such triplets $(a, \delta, \delta')$ along the rounds, he can exploit the result of the authentication. On one hand, if the authentication fails, he chooses another value for the first three bits of $x$ and $x'$; on the other hand, if the authentication succeeds, this increases his confidence in his choice.

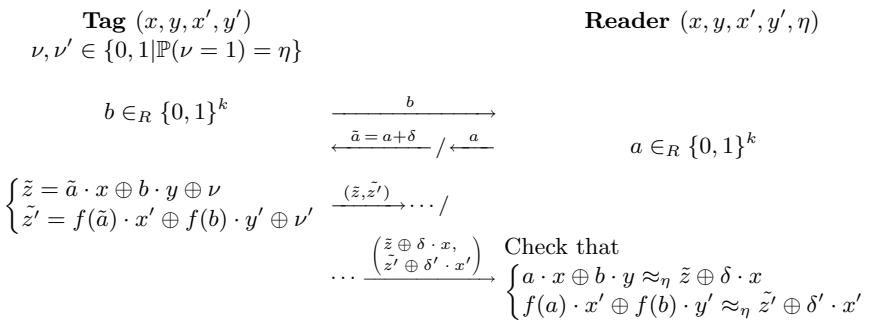


Figure 4: Wagner's attack

Indeed, we see in Fig. 4 that if the adversary has made a good guess the responses he sends to the reader are $\tilde{z} \oplus \delta \cdot x = z$ and $\tilde{z'} \oplus \delta' \cdot x' = z'$.

Actually, the existence of $(a, \delta, \delta')$'s verifying conditions 2a and 2b is likely to occur according to our initial hypothesis on $f$. In this example, if the same holds for the other output bits, after these 6 bits are recovered, the adversary can iteratively extend his knowledge of $x$,$x'$ by one bit at a time using a similar process, until he have learned all of $x$,$x'$. This attack requires a number of iterations that is linear in the key size.

Note that Wagner's attack is closer than Gilbert et al's one to a general man-in-the-middle attack. As the adversary modifies the messages in both ways, he changes challenge $a$ at his will, and the responses $\tilde{z}, \tilde{z'}$ are partly random, since the adversary tries all possible values for some bits of $x, x'$.

To counter this attack, we add in the computations of $f(a)\cdot x'$ and $f(b)\cdot y'$ a rotation which depends on the current round; i.e. we let

$$z' = rot(f(a), \rho) \cdot x' \oplus rot(f(b), \rho) \cdot y' \oplus \nu'$$

where $\rho$ stands for the index of the current round. This way, an adversary has to take into consideration in turn all portions of the secrets during one authentication, and so the attack is not practicable anymore for large $k$.

We address the problem of the security of the protocol across different authentications in the next section.

### 6.2 Description of HB$^{++}$

As already mentionned in Sect. 4.2, we conserve proofs of security for the model of adversary of [19]. Along one authentication, we design a protocol which resists to known man-in-the-middle attacks [10, 25]. Furthermore, among several authentications, we now execute this protocol under renewed secrets.

The HB$^{++}$ protocol can now be fully described. Each tag comes with a unique secret $Z$. At the beginning of each authentication, two challenges are exchanged between the reader and the tag. These challenges are derived under $Z$ with a universal hash function $h$ to obtain $x, x', y$ and $y'$ (see Fig. 5).

| **Tag** $(Z)$ | | **Reader** $(Z, \eta)$ |
|---|---|---|
| $B \in_R \{0,1\}^k$ | $\xrightarrow{B}$ | |
| | $\xleftarrow{A}$ | $A \in_R \{0,1\}^k$ |
| $(x, x', y, y') = h(Z, A, B)$ | | $(x, x', y, y') = h(Z, A, B)$ |

Figure 5: Preliminary stage of one authentication

These keys $x, x', y$ and $y'$ are then used to perform the authentication via $r$ successive rounds. Figure 6 illustrates the round $\rho \in \{1, \ldots, r\}$ of the protocol.

An example of construction of function $h$ is given for realistic parameters in Appendix A.2.

## 7 Conclusion

The main contribution of this paper is to present HB$^{++}$, a new identification protocol which can be used as a replacement of HB$^{+}$ for low-cost pervasive

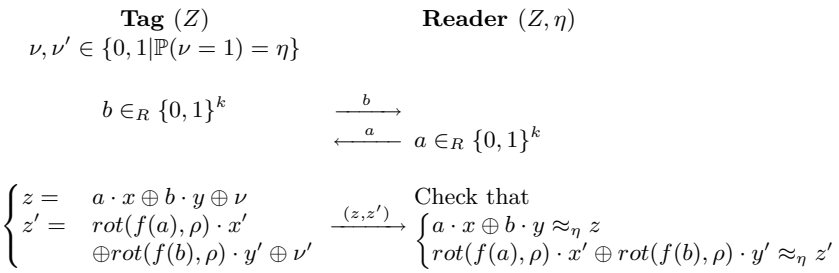


Figure 6: One round of the HB$^{++}$ protocol

computing devices. At the price of making more computations than in HB$^+$, it allows to achieve security in a stronger adversarial model than HB$^+$ as it is resistant to the attacks [10, 25] and at least as secure as HB$^+$ in its adversarial model. This point was left as "an essential line for future work" in [19]. In fact, with HB$^{++}$, we switch from the "detection security model" to a more classical one (i.e. a "prevention-based" model).

The way we improve HB$^+$, i.e. forcing challenges to a specific form, is, to the best of our knowledge, new.

Its formal security reduction, against any man-in-the-middle attack, to a hard problem is left as an open question. Our point is that for attacks considered by Juels and Weis, security proofs continue to hold. For man-in-the-middle attacks – at this time – we only rely on know-how techniques as, for instance, this is the case for the design of block ciphers. Note that here, the adversary is severly constrained in his actions as he has only access to the result of the authentication at the end of the entire protocol.

## Acknowledgment

The authors are quite grateful to David Wagner for his cryptanalysis of a previous version of the protocol (see Sect. 6.1), this helped a lot to improve this paper. They also wish to thank Ari Juels, Jooyoung Lee and Stephen Weis for their comments about HB$^+$ and HB$^{++}$, and the anonymous referees for their useful suggestions which improved the presentation of this paper.

## References

[1] G. Ateniese, J. Camenisch, and B. de Medeiros. Untraceable RFID tags via insubvertible encryption. In *Conference on Computer and Commu-*

*nications Security – CCS'05*. ACM Press, 2005.

[2] G. Avoine. `http://lasecwww.epfl.ch/~gavoine/rfid/`.

[3] G. Avoine and P. Oechslin. A scalable and provably secure hash based RFID protocol. In *International Workshop on Pervasive Computing and Communication Security – PerSec 2005*, pages 110–114. IEEE Computer Society Press, 2005.

[4] A. Blum, M. L. Furst, M. J. Kearns, and R. J. Lipton. Cryptographic primitives based on hard learning problems. In *Advances in Cryptology – CRYPTO'93*, Lecture Notes in Computer Science, pages 278–291. Springer-Verlag, 1993.

[5] A. Blum, A. Kalai, and H. Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. In *STOC 2000*, pages 435–440, 2000.

[6] F. Chabaud and S. Vaudenay. Links between differential and linear cryptanalysis. In A. D. Santis, editor, *Advances in Cryptology – EUROCRYPT'94*, volume 950 of *Lecture Notes in Computer Science*, pages 356–365. Springer-Verlag, 1994.

[7] S. Dominikus, E. Oswald, and M. Feldhofer. Symmetric authentication for RFID systems in practice. Handout of the Ecrypt Workshop on RFID and Lightweight Crypto, July 2005.

[8] EPC. Electronic product code global inc. `http://www.epcglobalinc.org/`.

[9] M. Feldhofer, S. Dominikus, and J. Wolkerstorfer. Strong authentication for RFID systems using the AES algorithm. In M. Joye and J.-J. Quisquater, editors, *Workshop on Cryptographic Hardware and Embedded Systems – CHES 2004*, volume 3156 of *Lecture Notes in Computer Science*, pages 357–370, Boston, Massachusetts, USA, August 2004. IACR, Springer-Verlag.

[10] H. Gilbert, M. Robshaw, and H. Sibert. An active attack against $HB^+$ - a provably secure lightweight authentication protocol. *IEE Electronic Letters*, 41:1169–1170, 2005. See also Cryptology ePrint Archive, Report 2005/237, `http://eprint.iacr.org`.

[11] R. Gold. Maximal recursive sequences with 3-valued crosscorrelation functions. *IEEE Trans. on Inform. Theory*, 14:154–156, 1968.

[12] P. Golle, M. Jakobsson, A. Juels, and P. Syverson. Universal re-encryption for mixnets. In T. Okamoto, editor, *The Cryptographers' Track at the RSA Conference – CT-RSA*, volume 2964 of *Lecture Notes in Computer Science*, pages 163–178. Springer-Verlag, 2004.

[13] D. Henrici and P. Müller. Hash-based enhancement of location privacy for radio-frequency identification devices using varying identifiers. In R. Sandhu and R. Thomas, editors, *International Workshop on Pervasive Computing and Communication Security – PerSec 2004*, pages 149–153. IEEE Computer Society, 2004.

[14] N. J. Hopper and M. Blum. Secure human identification protocols. In C. Boyd, editor, *Advances in Cryptology - ASIACRYPT 2001*, volume 2248 of *Lecture Notes in Computer Science*, pages 52–66. Springer-Verlag, 2001.

[15] J. Håstad. Some optimal inapproximability results. In *STOC 1997*, pages 1–10, 1997.

[16] A. Juels. "yoking-proofs" for RFID tags. In R. Sandhu and R. Thomas, editors, *International Workshop on Pervasive Computing and Communication Security – PerSec 2004*, pages 138–143. IEEE Computer Society, 2004.

[17] A. Juels. RFID security and privacy: A research survey. To appear in the *IEEE Journal on Selected Areas in Communication*, 2006.

[18] A. Juels, R. Rivest, and M. Szydlo. The blocker tag: Selective blocking of RFID tags for consumer privacy. In V. Atluri, editor, *Conference on Computer and Communications Security – ACM CCS*, pages 103–111. ACM Press, 2003.

[19] A. Juels and S. Weis. Authenticating pervasive devices with human protocols. In V. Shoup, editor, *Advances in Cryptology – CRYPTO'05*, volume 3126 of *Lecture Notes in Computer Science*, pages 293–308. Springer-Verlag, 2005.

[20] J.-P. Kaps, K. Yüksel, and B. Sunar. Energy scalable universal hashing. *IEEE Trans. on Computers*, 54:1484–1495, 2005.

[21] J. Katz and J. S. Shin. Parallel and concurrent security of the HB and HB$^+$ protocols. Cryptology ePrint Archive, Report 2005/461, 2005. `http://eprint.iacr.org/`.

[22] D. Molnar, A. Soppera, and D. Wagner. A scalable, delegatable pseudonym protocol enabling ownership transfer of RFID tags. In B. Preneel and S. Tavares, editors, *Selected Areas in Cryptography – SAC 2005*, Lecture Notes in Computer Science. Springer-Verlag, 2005.

[23] K. Nyberg and L. R. Knudsen. Provable security against differential cryptanalysis. In E. F. Brickell, editor, *Advances in Cryptology – CRYPTO'92*, volume 740 of *Lecture Notes in Computer Science*, pages 566–574. Springer-Verlag, 1992.

[24] J. Saito, J.-C. Ryou, and K. Sakurai. Enhancing privacy of universal re-encryption scheme for RFID tags. In L. Jang, M. Guo, G. Gao, and N. Jha, editors, *Embedded and Ubiquitous Computing – EUC 2004*, volume 3207 of *Lecture Notes in Computer Science*, pages 879–890. Springer-Verlag, 2004.

[25] D. Wagner. Private communication, December 2005.

[26] S. Weis, S. Sarma, R. Rivest, and D. Engels. Security and privacy aspects of low-cost radio frequency identification systems. In D. Hutter, G. Müller, W. Stephan, and M. Ullmann, editors, *International Conference on Security in Pervasive Computing – SPC 2003*, volume 2802 of *Lecture Notes in Computer Science*, pages 454–469. Springer-Verlag, 2003.

# A An example of practical settings

## A.1 A practical construction of $f$

### A.1.1 Some theoretic results on APN function

**Proposition 3 ([23])** *Let $f : \mathbb{F}_2^k \to \mathbb{F}_2^k$, then $\Delta_f \geq 2$. In case of equality, $f$ is said to be* Almost Perfect Nonlinear (APN).

**Proposition 4 ([11])** *Let $s = 2^j + 1$, known as a Gold exponent, with $\gcd(k, j) = 1$. If $k$ is odd, the power function $F$ defined as $F : x \mapsto x^s$ over $\mathbb{F}_{2^k}$ is a permutation and APN.*

Let $(\alpha_1, \ldots, \alpha_k)$ be a basis of $\mathbb{F}_{2^k}$ over $\mathbb{F}_2$, and $\varphi : (x_i)_{i=1..k} \in \mathbb{F}_2^k \mapsto \sum_i x_i \alpha_i$ the associated isomorphism. Let $s = 2^j + 1$ be a Gold exponent, $F$

the corresponding power function over $\mathbb{F}_{2^k}$ and $f = \varphi^{-1} \circ F \circ \varphi : \mathbb{F}_2^k \to \mathbb{F}_2^k$. Hence $f$ is a permutation and APN, that is $\Delta_f = 2$. Moreover, it is easy to see that $f$ is a quadratic function. But, even if it is quadratic, $f$ has a large complexity in terms of elementary operations, for a large $k$.

A way to reduce the complexity is to use a composition of functions defined over subspaces of $\mathbb{F}_2^k$. In particular, in the following case, the value $\Delta_f$ is easy to compute.

**Proposition 5** *Let $k = k_1 + k_2$, $f : \mathbb{F}_2^k \to \mathbb{F}_2^k$ defined for $a = (a_1, a_2) \in \mathbb{F}_2^{k_1} \times \mathbb{F}_2^{k_2}$ by $f(a) = (f_1(a_1), f_2(a_2))$ with $f_i : \mathbb{F}_2^{k_i} \to \mathbb{F}_2^{k_i}$. We have*

$$\Delta_f = \max(\Delta_{f_1}\Delta_{f_2}, \Delta_{f_1}2^{k_2}, \Delta_{f_2}2^{k_1}).$$

For example, we can use this construction with a "good" function $g : \mathbb{F}_2^{k_1} \to \mathbb{F}_2^{k_1}$ with low complexity (e.g. a Gold power function over a small field) and define $f : \mathbb{F}_2^k \to \mathbb{F}_2^k$ as $f(a_1, \ldots, a_j) = (g(a_1), \ldots, g(a_j))$ where $k = jk_1$. For well-chosen parameters, this function $f$ satisfies all the conditions to design the HB$^{++}$ protocol: $f$ is a permutation, has a low complexity and $\Delta_f = \Delta_g \times 2^{(j-1)k_1}$ is small compared to $2^k$.

### A.1.2 A example of practical settings

Let $k = 80$, the best known algorithm to solve the relying LPN problem has a computational runtime and needs a number of challenges greater than $2^{35}$ [19] (with $\eta = 1/4$).

Let $k_1 = 5$, $j = 16$ and $(\alpha_1, \ldots, \alpha_{k_1})$ be a basis of $\mathbb{F}_{2^{k_1}}$ over $\mathbb{F}_2$, and $\varphi : \mathbb{F}_2^{k_1} \to \mathbb{F}_{2^{k_1}}, (x_i)_{i=1..k_1} \mapsto \sum_i x_i\alpha_i$ the associated isomorphism.

We construct $f : \mathbb{F}_2^k \to \mathbb{F}_2^k$ thanks to the power function

$$\begin{array}{rccc} g: & \mathbb{F}_{2^{k_1}} & \to & \mathbb{F}_{2^{k_1}} \\ & x & \mapsto & x^3 \end{array}$$

by

$$f(a_1, \ldots, a_j) = (\tilde{g}(a_1), \ldots, \tilde{g}(a_j)),$$

for $a = (a_1, \ldots, a_j) \in (\mathbb{F}_2^{k_1})^j$ and $\tilde{g}(x) = \varphi^{-1} \circ g \circ \varphi(x)$.

As explained above, $g$ is a permutation and $\Delta_g = 2$ ($s = 3$ is a Gold exponent, so $g$ is an APN function). Hence, $f$ is a permutation and

$$\Delta_f = 2^{(j-1)k_1+1} = 2^{k-4}.$$

Thus, the probability for an attack, like the one described in [10], to succeed is lower than $(2^{-4})^r$. For $r \geq 20$, the probability of success is smaller than $2^{-80}$.

One remaining constraint has to be checked: $f$ must have a low complexity.

We set the representation of the field $\mathbb{F}_{2^{k_1}}$ as

$$\mathbb{F}_{2^{k_1}} = \mathbb{F}_2[X]/(P)$$

where $P = X^5 + X^2 + 1$ is an irreducible polynomial over $\mathbb{F}_2$. For $\alpha$ a root of $P$ in $\mathbb{F}_{2^{k_1}}$, let

$$(\alpha_1, \ldots, \alpha_{k_1}) = (1, \alpha, \alpha^2, \alpha^3, \alpha^4)$$

be the canonical basis of $\mathbb{F}_{2^{k_1}}$. For this basis, a description of $\tilde{g} : \mathbb{F}_2^5 \to \mathbb{F}_2^5$ is given below.

$$\begin{aligned}\tilde{g} : (x_0,\, x_1,\, x_2,\, x_3,\, x_4) \longmapsto (&x_0 \oplus x_1x_3 \oplus x_1x_2 \oplus x_2x_3 \oplus x_0x_4,\\ &x_0x_1 \oplus x_2 \oplus x_0x_3 \oplus x_3 \oplus x_3x_4 \oplus x_4,\\ &x_0x_2 \oplus x_0x_1 \oplus x_1x_2 \oplus x_2x_4 \oplus x_0x_4 \oplus x_3x_4 \oplus x_4,\\ &x_1 \oplus x_2 \oplus x_2x_4 \oplus x_2x_3 \oplus x_3 \oplus x_0x_4 \oplus x_4,\\ &x_0x_4 \oplus x_1x_2 \oplus x_0x_2 \oplus x_2x_3 \oplus x_1x_3 \oplus x_3 \oplus x_3x_4 \oplus x_1x_4 \oplus x_2x_4).\end{aligned}$$

The computation of $\tilde{g}$ requires only the evaluation of 10 AND and 29 XOR.

Finally, suppose that each authentication requires 40 rounds. At each round, we rotate $f(a)$ and $f(b)$ by 2 bits. At the end of an authentication, with these rotations, the output of function $f$, $f(a)$ (resp. $f(b)$), is thus related to all 2-bit blocks of $x'$ (resp. $y'$) during the computation of the scalar product.

## A.2 Construction of $h$

We follow [20], choose $h = \mathrm{WH^T}$-16 and, from now, we adopt the notation of this article (here we let $w = 16$).

We have $\mathrm{WH^T}$-16 : $\{0,1\}^{n \times w} \to \{0,1\}^{t \times w}$ and here choose $t = 20$ and $n = 10$.

The key, $Z$, has $n + 2(t-1)$ $w$-bit words, i.e. 768 bits. Let $Z = Z_1, \ldots, Z_{n+2(t-1)}$ where each $Z_i$ is a $w$-bit word.

The output of $\mathrm{WH}^{\mathrm{T}}$-16 is made of $t$ $w$-bit words. And for each of these words, $n$ words $Z_i$ of the key are involved in the computation:

$$\begin{aligned}\mathrm{WH}^{\mathrm{T}}\text{-16}(M) \;=\; & \big(\text{WH-16}(M; Z_1, \ldots, Z_n),\\ & \text{WH-16}(M; Z_3, \ldots, Z_{n+2}),\\ & \ldots,\\ & \text{WH-16}(M; Z_{2t-1}, \ldots, Z_{n+2t-2})\big),\end{aligned}$$

where the function WH-16 needs 460 gates to be implemented and consumes only 2.95 $\mu W$ at 500 kHz.

About the security of this construction, the following result is proven in [20].

**Theorem 1** *The function* $\mathrm{WH}^{\mathrm{T}}$-16 *is universal on equal-length strings with collision probability of* $2^{-wt}$.

To compute new secrets $x, x', y$ and $y'$, challenges exchanged at the beginning of each authentication are 80 bits long and are concatenated together to form the input of $h = \mathrm{WH}^{\mathrm{T}}$-16.

#### 7.1.1.1 Analyse de sécurité du protocole $\mathrm{HB}^{++}$ final

Dans l'article, la preuve qu'une attaque sur le protocole peut se transposer en une attaque sur le protocole $\mathrm{HB}^{+}$ ou sur une instance du problème LPN n'est faite que pour la première version (cf. [BCD06b, section 4]). Nous détaillons ici en quoi cette propriété reste valable pour le protocole final (décrit par [BCD06b, section 6.2, Fig. 5–6]) dans le cas d'une série d'itérations utilisant la même initialisation.

**Proposition 7.1** *Un attaquant $\mathcal{A}$, qui a la faculté de retrouver le secret $Z$ à partir d'un ensemble quelconque de n challenges-réponses du protocole $\mathrm{HB}^{++}$ entre un tag $\mathcal{T}_Z$ et un lecteur, sait attaquer le protocole $\mathrm{HB}^{+}$ avec la même complexité.*

*Démonstration.* Supposons que $\mathcal{A}$ obtienne un ensemble de $n$ challenges-réponses du protocole $\mathrm{HB}^{+}$, $\mathcal{S} = \{a_i, b_i, a_i \cdot x \oplus b_i \cdot y \oplus \nu_i\}_{i\in\{1,\dots,n\}}$, à partir de plusieurs échanges entre un tag $\mathcal{T}_{x,y}$ et un lecteur $\mathcal{R}$. $\mathcal{A}$ choisit alors $x'$, $y'$ au hasard et une variable aléatoire $\nu'$ tel que $\mathbb{P}(\nu' = 1) = \eta$ afin de générer l'ensemble

$$\mathcal{S}' = \left\{a_i,\, b_i,\, a_i \cdot x \oplus b_i \cdot y \oplus \nu_i,\, rot\left(f\left(a_i\right), i \bmod r\right) \cdot x' \oplus rot\left(f\left(b_i\right), i \bmod r\right) \cdot y' \oplus \nu'_i\right\}_{i\in\{1,\dots,n\}},$$

simulant $n$ challenges-réponses successifs du protocole $\mathrm{HB}^{++}$. Soient $A$ et $B$ deux vecteurs binaires quelconques de longueur $k$, alors $\mathcal{S}'$ correspond aux échanges générés par environ $n/r$ authentifications d'un certain tag $\mathcal{T}_Z$ sur un lecteur quelconque, avec des pré-challenges $A$ et $B$ fixes, où $Z$ est tel que $h(Z, A, B) = (x, x', y, y')$ ; comme $h$ est une fonction de hachage universelle, on sait qu'il existe un tel $Z$. Finalement si $\mathcal{A}$ retrouve le secret $Z$, il pourra en déduire $x$ et $y$. □

**Proposition 7.2** *Un attaquant $\mathcal{A}$, qui a la faculté de retrouver le secret $Z$ à partir d'un ensemble quelconque de n challenges-réponses du protocole $\mathrm{HB}^{++}$ entre un tag $\mathcal{T}_Z$ et un lecteur, sait résoudre une instance quelconque du problème LPN associée à une matrice aléatoire de taille $n \times 2k$ et un secret de longueur $2k$.*

*Démonstration.* Soient $A$ une matrice binaire aléatoire de taille $n \times 2k$, $X = (x, y)$ un vecteur aléatoire de $2k$ bits, $\eta$ un paramètre de bruit, et $\vec{\nu}$ un vecteur aléatoire de $n$ bits tel que $w_H(\vec{\nu}) \leq \eta n$ et soit $\vec{z} = AX \oplus \vec{\nu}$. Alors $\mathcal{A}$ construit les vecteurs de $k$ bits $a_i, b_i$ pour $i \in \{1, \dots, n\}$ tels que

$$A = \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix}.$$

L'attaquant $\mathcal{A}$ peut alors interpréter $\vec{z} = (a_i \cdot x \oplus b_i \cdot y \oplus \nu_i)_{i\in\{1,\dots,n\}}$ comme des réponses du protocole $\mathrm{HB}^{+}$ et procéder comme pour la proposition précédente, afin de retrouver $X$. □

La généralisation à un ensemble arbitraire de séries utilisant différentes initialisations est moins naturelle, puisqu'une attaque se réduit alors à une attaque sur plusieurs instances du protocole $HB^+$ ou du problème LPN mais avec des secrets corrélés (obtenus à partir d'un même secret $Z$ via la fonction $h$). La sécurité est difficile à analyser dans ce cas sans hypothèse supplémentaire sur $h$[1], mais il paraît raisonnable de penser que c'est au moins aussi difficile à résoudre que le cas où tous les secrets sont fixes.

Concernant la sécurité contre les attaques par le milieu, l'utilisation de la fonction $f$ et de rotations permet de réduire la probabilité de gagner de l'information à partir du résultat d'une authentification via les attaques de Gilbert *et al.* [GRS05] et de Wagner [Wag05]. Plus exactement, le but est de faire en sorte qu'étant donné un tag $\mathcal{T}$ pré-initialisé avec des vecteurs $(x, x', y, y')$, alors un attaquant qui suit une des techniques [GRS05, Wag05] ne peut deviner un bit quelconque de ces vecteurs, sauf avec une probabilité très faible, en observant une seule tentative d'authentification. D'autre part, sur une seule authentification, il semble difficile d'envisager d'autres types d'attaques par le milieu.

À noter que la rotation introduite dans la version finale du protocole, qui a pour but d'empêcher l'attaque de Wagner, ne réduit en rien la résistance contre [GRS05] étudiée dans [BCD06b, section 5]. En effet, pour obtenir de l'information en une seule authentification, un attaquant qui modifie les challenges envoyés au tag suivant la méthode [GRS05], doit trouver $\delta \neq 0$ et $\gamma$ dans $\{0,1\}^k$ tels que pour tout $i \in \{0, \ldots, r\}$, $rot(f(a_i \oplus \delta), i) = rot(f(a_i), i) \oplus \gamma$ (où $a_i$ correspond au $i$-ième challenge), i.e.

$$f(a_i \oplus \delta) = f(a_i) \oplus rot^{-1}(\gamma, i).$$

Or, à $\delta \neq 0$ et $\gamma$ fixés, la probabilité d'obtenir des challenges vérifiant cette équation reste égale à $\left(\frac{\Delta_f}{2^k}\right)^r$.

Enfin, l'étape préliminaire est à voir comme une mesure de précaution contre d'autres attaques par le milieu : l'utilisation de la fonction $h$ de hachage universelle a uniquement pour rôle de décorréler les différentes authentifications afin que ces attaques ne puissent pas être généralisées à plusieurs authentifications (pour augmenter les chances de succès, par exemple en ne modifiant à chaque authentification qu'un petit nombre d'itérations). Ainsi, aucune résistance à la recherche de collisions ou de pré-images n'est présumée (c'est ce qui permet de choisir une fonction légère à implémenter contrairement à une fonction de hachage cryptographique plus traditionnelle, e.g. SHA-1), le seul point important est que les mêmes secrets temporaires $(x, x', y, y')$ ne soient pas ré-utilisés trop souvent. En particulier, on ne veut pas que les mêmes secrets temporaires soient rencontrés facilement pour des clés $Z$ différentes, afin d'éviter que des tags soient confondus ou usurpés.

[1] Dans le cas d'une implémentation sur RFID, nous préférons l'éviter : plus de contraintes sur la fonction de hachage entraînerait un coût d'implémentation accru.

### 7.1.2 RFID et canal à jarretière

Dans l'article *On the Wiretap Channel Induced by Noisy Tags* [BC06], nous nous intéressons cette fois-ci à une solution de sécurisation des échanges dont les cibles sont principalement les tags les plus élémentaires (classe 1), se comportant tels que des mémoires avec une très faible capacité de calcul.

On se place ici dans le contexte des protocoles combinant des procédés algorithmiques et des principes physiques (j'invite le lecteur à consulter [BCG+06] pour un survol de quelques-unes de ces méthodes) ; l'intérêt de ce type de solutions est de nécessiter beaucoup moins de ressources au niveau du tag qu'une solution purement cryptographique. Des solutions de restrictions d'usage ont tout d'abord été proposées (telles que la désactivation du tag, l'utilisation d'une antenne amovible, de "bouclier" radio, la limitation de distance, ...). Plus récemment, des techniques exploitant le bruit du canal de communication sont également mises en avant. Le fait de tirer parti du bruit du canal n'est pas nouveau, c'est une des bases de la théorie quantique de l'information, mais son application au contexte des RFID l'est. Ainsi, [CF06] propose un protocole d'établissement de clés bien adapté aux tags à faibles ressources en développant un algorithme suivant le principe de distillation–réconciliation–amplification. En quelque sorte, le bruit naturel du canal permet de créer un brouillage contre un attaquant passif.

Afin de protéger l'établissement ou l'échange d'un secret entre un lecteur et un tag en présence d'écoutes éventuelles du canal, Avoine et Castelluccia ont proposé [CA06] à leur tour de créer artificiellement ce brouillage externe (et de le contrôler). Comme décrit par la figure 7.3, le tag envoie seulement une séquence de bits mais en parallèle un faux tag spécifique, appelé *noisy tag* (NT), envoie une séquence de bits pour perturber la séquence originale. Ce *noisy tag* est contrôlé par le lecteur de sorte que ce dernier puisse facilement reconstruire la séquence envoyée par le tag alors qu'un observateur extérieur ne verra que du bruit. Typiquement, le lecteur et NT partagent une clé secrète et le lecteur envoie une graine à NT afin de générer une séquence pseudo-aléatoire qui est ainsi connue du lecteur.

En théorie, seul le lecteur peut soustraire ce bruit mais en pratique, l'implémentation du protocole fait qu'en cas de collision (sur un bit envoyé par le tag et le noisy tag), un attaquant sait quelle est la valeur du bit en question. C'est le cas par exemple lorsque qu'un bit à 1 est représenté par un saut de tension puisque que les amplitudes des deux signaux s'additionnent alors. Dans cet exemple (correspondant au premier protocole proposé dans [CA06]) la probabilité de collision étant d'environ $1/2$, le tag doit donc envoyer $2n$ bits en moyenne pour obtenir $n$ bits secrets.

En modélisant alors ce qu'observe un attaquant sur le canal par un canal avec effacements, nous proposons dans [BC06] d'utiliser la théorie du canal à jarretière (ou *wiretap channel*) développée par Wyner [Wyn75], pour améliorer le protocole et tendre vers une sécurité presque parfaite via un codage adapté. L'idée est de transmettre un élément d'un coset d'un code linéaire binaire au lieu du secret directement, de sorte que si le nombre d'effacements est suffisant alors on ne peut déterminer quel est le secret de départ. Un exemple de construction à partir de codes LDPC (*Low Density Parity Check*, cf. [Gal63]) réguliers est proposé pour lequel la sécurité est presque sûrement atteinte quand la longueur croît vers l'infini.

Cependant, au delà de la mise en évidence de ce parallèle *noisy tags – wiretap channel* et de son exploitation, il serait intéressant de compléter l'analyse par une étude de la sécurité

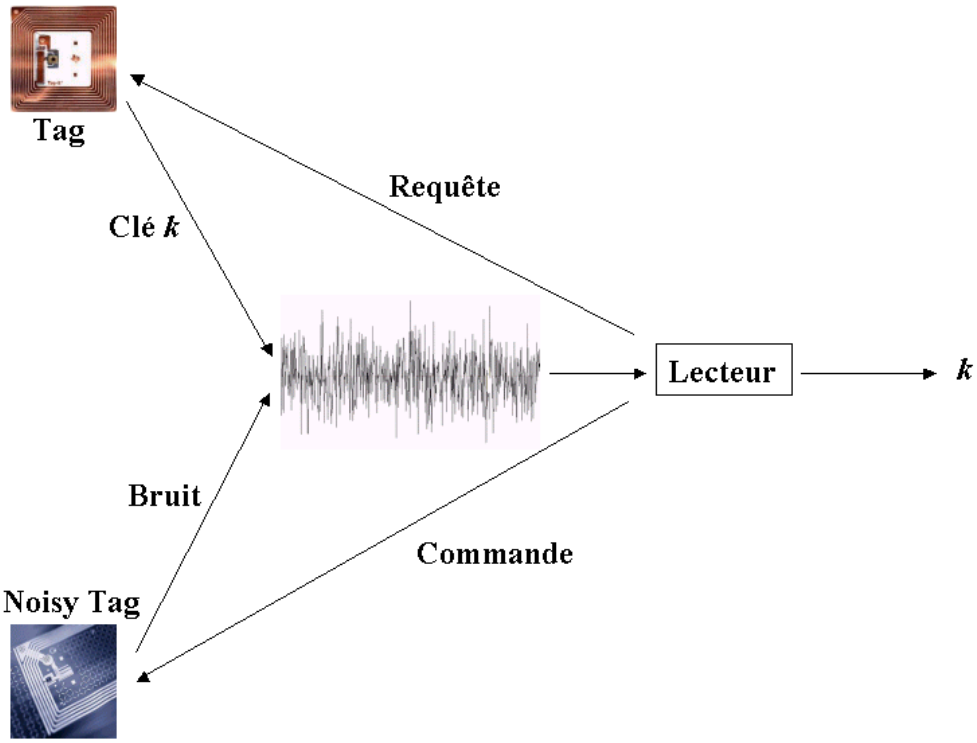


FIG. 7.3 – Protocole Noisy Tags

réellement atteinte à longueur finie avant de valider le choix d'un code en particulier. De plus, il serait intéressant d'étudier la sécurisation du protocole face à des attaques actives de manière analogue au protocole HB (cf. Sec. 7.1.1).

# On the Wiretap Channel Induced by Noisy Tags

Julien Bringer Hervé Chabanne
Sagem Défense Sécurité

**Abstract**

At CARDIS'06, Castelluccia and Avoine introduce noisy tags to allow key exchange between an RFID tag and a reader. We here show that their protocol leads to a well-known information problem: the wiretap channel. We then make use of works by Thangaraj *et al.* on the case where the main channel is noiseless and where there are only erasures on the wiretapper's channel to improve previous results on noisy tags. In particular, we show how one can achieve, in a practical manner, perfect secrecy for key exchange in this noisy tags context.

**Keywords.** RFID, wiretap channel, noisy tags, LDPC codes.

## 1 Introduction

Securely pairing RFID tags to a reader is a particularly hard challenge due to the cost constrainsts which push to always reduce resources inside RFID tags. A very attractive solution was proposed by Castelluccia and Avoine in [2] as in their protocol the tag behaves like a memory, i.e. only sends a sequence of bits to the reader. Note that, doing so, the protocol is naturally more resistant to side channel attacks, which begin to appear for RFIDs [6].

What makes the simplicity of the protocol by Castelluccia and Avoine possible is the introduction on the reader side of a particular RFID called the noisy tag which allows to add perturbations in the communications between the RFID tag and its reader. This extra noise is controlled by the reader but remains unknown to eavesdroppers.

Following Castelluccia and Avoine, we here improve their protocol by introducing new coding scheme of the informations sent by the RFID tag, leading to perfect secrecy.

In Sect. 2, we recall how noisy tags are used by Castelluccia and Avoine and show that this leads to a well-known problem: a particular wiretap channel. In Sect. 3, we recall results on the classical wiretap channel problem

and advances due to Thangaraj *et al.* on the very case of noisy tags. In Sect. 4, we give some examples to illustrate that perfect secrecy is achievable in a practical manner. Section 5 concludes.

For references on RFID security, we invite the reader to check online references at `http://lasecwww.epfl.ch/~gavoine/rfid`. See also [4] for a recent survey.

## 2 Noisy Tags

Quoting Castelluccia and Avoine, the idea of noisy tags comes from a key exchange scheme developped at Bells Telephone Labs during WWII.

Here each reader comes equipped with a special RFID tag: the Noisy Tag (NT). And both the RFID Tag (T) which wants to establish a key with the reader and the Noisy Tag emit bits, simultaneously in order to hide the bits sent by T. Typically, the sequence of bits issued from the Noisy Tag is a pseudo-random sequence of bits, controlled by its associated reader.

We suppose (cf. [2]) that the bit '1' is implemented by a pulse of $x$ mV and that the bit '0' corresponds to a pulse of 0 mV.

- From an attacker's point of view:
  - when T and NT have sent a different bit, he observes $x$ mV in the air and can not distinguish, between T and NT, which one sent the bit '0' and which one sent the '1' : this corresponds to an erasure on the wiretap channel,
  - when both T and NT have sent the same bit, he certainly knows which bit it was as he gets 0 mV, if both T and NT have sent '0', or $2x$ mV if T and NT have sent '1'.
- From the reader's view point, as it knows in advance the sequence of bits produced by NT, at the end, it can retrieve the bits emitted by NT from those which are received and thus obtains the ones issued by T.

This situation which corresponds to the Bit-Based Protocol, Version 1 of [2], can be described as follows (see Fig. 1):

1. there is a noiseless main channel between T and the reader,
2. the attacker gets the information from a Binary Erasure Channel where an erasure has one chance over two to happen.

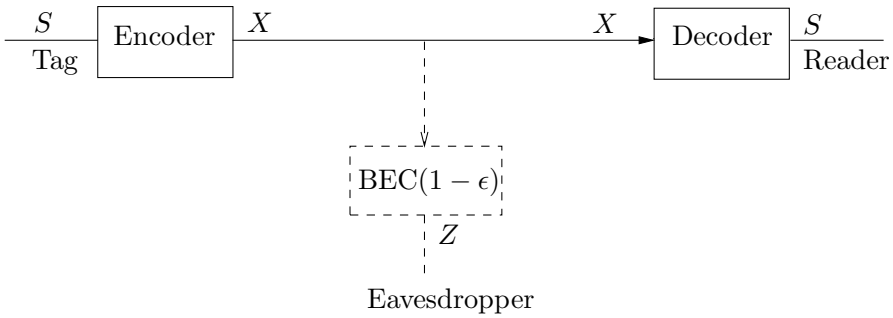


Figure 1: Wiretap channel with erasure

In Figure 1, BEC$(1-\epsilon)$ stands for Binary Erasure Channel with a probalility of $1-\epsilon$ to have an erasure.

**Remark 1** *Castelluccia and Avoine introduce three different protocols in [2]. In the last one, to obtain a security level of* $2^{80}$*, an RFID tag T has to send more than one Kbits to establish an* 80*-bit long key with a reader. The first two ones require that the RFID tag T sends roughly 2 times the key length, but in these protocols, some informations are leaked to the eavesdropper.*

Note that we do not address here (and in [2] neither) the problem of exchanging the same key several times. I.e. we do not consider different executions of the protocol between an RFID tag and readers for establishing the same key across time. We may only consider that the attacker has access to a wiretap modelized as a Binary Erasure Channel BEC$(1-1/2)$ if he observes one execution of the protocol, BEC$(1/4)$ if he eavesdrops two executions and so on. The last protocol of [2] - Code-Based Protocol - suffers also from this kind of situation as with only two observations of the same key exchange, an attacker knows exactly which key is emitted by the RFID tag T (even with several noisy tags). Some solution may be envisaged to alleviate this problem. For instance, we may think at a renewal of key material inside RFID tag T after each successful execution of the protocol [5].

## 3 The Erasure Wiretapper's Channel

The scenario of wiretap channels with erasure, including the one associated to the noisy tags model depicted in Fig. 1, has been studied by Thangaraj *et*

*al.* [10].

The wiretap channel problem was first introduced by Wyner [11] in 1975. Classicaly, to transmit $k$-bit messages, a binary linear code $C$ of length $n$ is chosen and each message is associated to a chosen coset of $C$. More precisely, let

- $n = k + l$,
- $G = (g_1, \ldots, g_l)$ be a generator matrix $l \times n$ of $C$,
- $h_1, \ldots, h_k$ be $k$ linearly independent vectors from $\{0,1\}^n$, not in $C$,
- $v = (v_1, \ldots, v_l)$ be a uniformly random $l$-bit vector,

then a message $s = (s_1, \ldots, s_k)$ in $\{0,1\}^k$ is encoded as

$$x = s_1 h_1 + s_2 h_2 + \cdots + s_k h_k + v_1 g_1 + v_2 g_2 + \cdots + v_l g_l. \qquad (1)$$

We will only consider here the case where $(g_1, \ldots, g_l, h_1, \ldots, h_k)$ span the entire space vector $\{0,1\}^n$. Note that with such a construction, the information rate of the communication channel will be $R = k/n$ (i.e. 1 minus the information rate of $C$).

Given $z$ received by an eavesdropper, if a coset of C contains at least one vector that agrees with $z \in \{0,1\}^n$ in the unerased positions, we say that the coset is consistent with $z$. Let $N(C, z)$ be the total number of cosets of $C$ consistent with $z$. For a code $C$ of length $n$ and dimension $l = n - k$, the maximum possible value for $N(C, z)$ is $2^k$. If the maximum value is reached, i.e. $N(C, z) = 2^k$, we say that $z$ is secured by $C$ and that perfect secrecy is achieved. Indeed, in this case, for messages from a random variable $S$, we have $\mathrm{H}(S|Z = z) = k = \mathrm{H}(S)$, i.e. that $z$ does not reveal anything on $S$.

Now, we have a nice characterization:

**Theorem 1 ([7, 10])** *Let an $[n, n-k]$ code $C$ have a generator matrix $G = \begin{pmatrix} a_1 \cdots a_n \end{pmatrix}$ where $a_i$ is the $i$-th column of $G$. Consider an instance of the eavesdropper's observation $z \in \{0,1,?\}^n$ with $\mu$ unerased positions given by $\{i : z_i \neq ?\} = \{i_1, i_2, \ldots, i_\mu\}$.*

*Then, $z$ is secured by $C$ if and only if the matrix $G_\mu = \begin{pmatrix} a_{i_1} a_{i_2} \cdots a_{i_\mu} \end{pmatrix}$ has maximal rank:* $\operatorname{Rank} G_\mu = \mu$.

*Proof. [For the purpose of completeness, we here give the proof from [10].]* If $G_\mu$ has rank $\mu$, the code $C$ has all $2^\mu$ possible $\mu$-tuples in the $\mu$ unerased

positions. So each coset of $C$ also has all $2^\mu$ possible $\mu$-tuples in the revealed positions. Hence $N(C, z) = 2^k$.

If $G_\mu$ has rank less than $\mu$, the code $C$ does not have all $\mu$-tuples in the $\mu$ unerased positions. So there exists at least one coset that does not contain a given $\mu$-tuple in the $\mu$ unerased positions, this implies $N(C, z) < 2^k$. □

The main purpose of [10] is to introduce codes that approach secrecy capacity (i.e. the largest $k/n$ for which the objectives of secure and reliable communication is achievable) over some wiretap channels. In particular, for a wiretap channel with a noiseless main channel and a binary erasure channel as the wiretapper's channel (as in Fig. 1, see previous section), the authors of [10] exhibit constructions which allow to reach linear-time decodable codes.

One construction of [10] relies on duals of Low-Density Parity-Check (LDPC) codes (see [3] for details on LDPC codes). They choose an LDPC code $[n, k]$ and use the dual (or equivalently a parity check matrix), which is an $[n, n-k]$ code, as the code $C$.

LDPC codes are linear codes obtained from sparse bipartite graphs. Consider a bipartite graph with $n$ left (message or variable) nodes and $r$ right (check) nodes. Then forms the binary matrix in which the entry $(i, j)$ is 1 if and only if the $i$-th check node is connected to the $j$-th message node in the graph via an edge (see Fig. 2 for an example). This "adjacency" matrix is used as a parity check matrix for an LDPC code.

In [8], the threshold $\alpha^*$ of an LDPC code is introduced. This threshold serves us - following [10] - to bind LDPC codes to the wiretap channel via Theorem 1:

**Theorem 2 ([10])** *Let $H$ be a parity-check matrix of an LDPC code with threshold $\alpha^*$. Then a submatrix formed by selecting columns of $H$ independently with probability $\epsilon$ will have full column rank for $\epsilon < \alpha^*$ for large $k$ with high probalility.*

In other words, a code $C$ which is the dual of an LDPC code has a great chance to be a good candidate for encoding messages over our noisy tag channel provided that $\epsilon < \alpha^*$.

With such a code, the exchange is then the following (see Fig. 1). To reveal its secret $s$ to a reader, the tag sends its encoding value $x$ in the air and the reader retrieves the value of the secret from the string $x$ sent by the tag simply by decoding it.

We therefore must know how much it costs. Now recovering $s$ from $x$ in (1) is an $O(n^2)$ operation (as there are no errors on the main channel

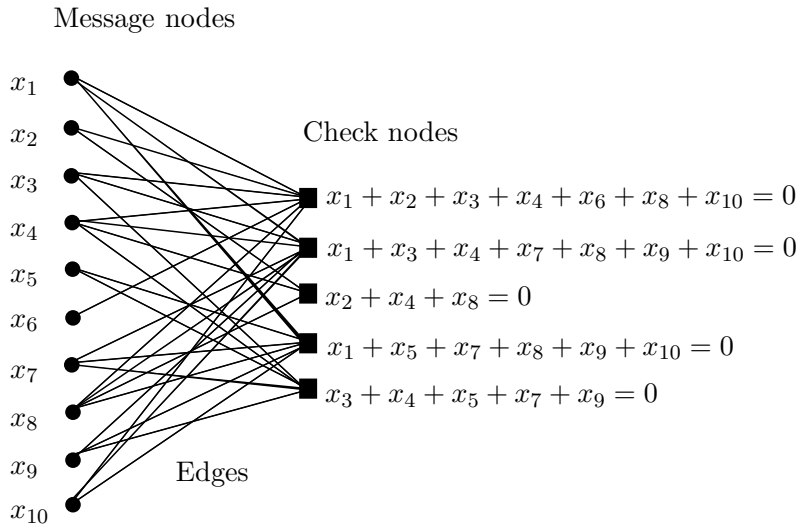


Figure 2: A bipartite graph of an irregular LDPC code

because the noisy tag is controlled by the reader), the computation of $s$ is essentially a syndrome computation for $C$. So to recover $s$, it is sufficient to multiply $x$ with a matrix. Moreover, this could be improved: [10] gives also designs for linear or almost-linear decodable secrecy codes for the system shown in Fig. 1. Finally, this decoding stage is easily performed by the reader.

# 4 A Practical Example

In practice, we limit ourselves to a particular class of LDPC codes, the regular codes, where the threshold (in the Binary Erasure Channel) is easily computable. An LDPC code with an underlying bipartite graph for which each message node is connected to exactly $j$ check nodes and each check node involves $k$ message nodes is called a $(j,k)$-regular LDPC code. It is shown in [1] that we have the following result.

**Lemma 1** *Let $C$ be a $(j,k)$-regular LDPC code and $\sigma$ the unique positive real root of $P(X) = \big((j-1)(k-1)-1\big)X^{k-2} - \sum_{i=0}^{k-3} X^i$. Then the threshold $\alpha^*$ of $C$ for the BEC is*

$$\alpha^* = \frac{1-\sigma}{(1-\sigma^{k-1})^{j-1}}.$$

In particular, according to [1], a $(3,4)$-regular LDPC code has a threshold $\alpha^*$ of 0.647426. This is reasonably greater than the 0.5 imposed by the wiretapper's channel in the noisy tag model and so it may lead to perfect secrecy. Moreover, the information rate of a $(j,k)$-regular code is $1-\frac{j}{k}$, so for $C$ the dual of this code, the information rate on the main channel is $1-3/4=1/4$ and one can transmit a 80-bit key long by sending only 320 bits to the reader, which is really lower than in the third protocol proposed in [2]. For instance, one can find constructions of $(3,4)$-regular LDPC codes in [9].

Many other constructions of regular LDPC codes with a threshold greater than $1/2$ are available, and thus the choice would be easily made with respect to the local constraints. For instance, it could be valuable to increase the information rate if the number of bits has to be lower. For example, a $(4,6)$-regular code and a $(3,5)$-regular code have both a threshold around 0.5. As the first one gives a rate of $1/3$ and the second one a rate of $2/5$, it allows to send only 240 bits in the first case and 200 bits in the second case, for a secret of 80 bits. At last, it could also be necessary to increase the threshold if the probabily of erasures (from the attacker's point of view) decreases.

**Remark 2** *For the last point, we can also use several noisy tags at the same time to avoid the probability to be too low. Another "trick" is possible: as the reader also knows when a bit is leaked over the wiretapper's channel, then it can have a special command to tell to suspend its communication with the RFID tag $T$ when too many bits have escaped to an eavesdropper.*

Actually, a good thing is that for the RFID tag this coding scheme does not change the implementation but only the length of the vector to store, as the idea is to run the encoding algorithm outside the tag and thus to write directly the encoded value $x$ (and in option the key $s$, if it is used after the execution of this protocol) in the tag, in order to keep its behaviour like a memory. Hence, the encoding complexity is not a constraint for the tag.

**Reusability.** We focused on the case where the string $x$ is sent only once, but to consider the impact of repeated observation by an adversary, we have seen mainly three ideas which are summarized below. Of course, to improve efficiency, these solutions should be combined. Note that, as mentioned in Sect. 2, it is possible to simulate repetitions by taking a BEC with a lower probability of erasure.

- First, as in [2], we can increase the number of noisy tags in parallel in order to stay in the same model, for instance to keep a probability of erasure greater than 1/2. With $Q$ noisy tags, an attacker would need about $2^{Q-1}$ repetitions to observe the channel with a probability lower than 1/2 (i.e. to be in a BEC$(1-\epsilon)$ more favourable with $\epsilon > 1/2$).
- A second way is to use a dual code of an LDPC code with an increased threshold in order to retain the result of theorem 2. A possible drawback is that it would decrease the capacity and thus would have a marked effect on the efficiency of the scheme if the adversary can observe many attempts.
- Another point is to update the tag's key after each successful authentication. This option looks more reasonable than the two previous ones, but then it becomes important to create an authenticated channel from the reader to the tag in order to communicate success. Otherwise, the tag could suffer, for instance, a denial-of-service attack in which the attacker communicate success several times to the tag leading to a state de-synchronization between tag and reader.

## 5 Conclusion

We have shown how, following Thangaraj *et al.*, the use of dual of LDPC codes may allow to achieve perfect secrecy for the noisy tags at low cost. In fact, we have given a solution, where the eavesdropper gets no information on the key from a theoretical point of view, with a cost which compares favourably in terms of communication with the protocols by Castelluccia and Avoine. To be fair, one should notice that it requires more computations on the reader side: at most $O(n^2)$ operations, which seems not critical to us.

The security provided here concerns the passive attackers, a remaining point is to analyse if this wiretap model is sufficient for providing perfect secrecy against active attacks, where the adversary can impersonate either tag or reader in many communication attempts. The above-mentioned problem of dealing with multiple observations is a part of the security analysis to do in order to address all practical security issues. The next step in this direction is to focus ourselves on a well-chosen code and to check its behaviour.

## Acknowledgments

The authors wish to thank the anonymous referees for their useful suggestions which helped to improve the presentation of this paper.

## References

[1] Louay Bazzi, Thomas J. Richardson, and Rüdiger L. Urbanke. Exact thresholds and optimal codes for the binary-symmetric channel and gallager's decoding algorithm A. *IEEE Transactions on Information Theory*, 50(9):2010–2021, 2004.

[2] Claude Castelluccia and Gildas Avoine. Noisy tags: A pretty good key exchange protocol for RFID tags. In Josep Domingo-Ferrer, Joachim Posegga, and Daniel Schreckling, editors, *International Conference on Smart Card Research and Advanced Applications – CARDIS*, volume 3928 of *Lecture Notes in Computer Science*, pages 289–299, Tarragona, Spain, April 2006. IFIP, Springer-Verlag.

[3] R. G. Gallager. *Low-Density Parity Check Codes*. Number 21 in Research monograph series. MIT Press, Cambridge, MA, 1963.

[4] Simson L. Garfinkel, Ari Juels, and Ravikanth Pappu. RFID Privacy: An Overview of Problems and Proposed Solutions. *IEEE Security & Privacy*, 3(3):34–43, 2005.

[5] David Molnar and David Wagner. Privacy and security in library RFID: Issues, practices, and architectures. In Birgit Pfitzmann and Peng Liu, editors, *Conference on Computer and Communications Security – ACM CCS*, pages 210–219, Washington, DC, USA, October 2004. ACM, ACM Press.

[6] Yossi Oren and Adi Shamir. Power analysis of RFID tags, 2006. `http://www.wisdom.weizmann.ac.il/~yossio/rfid/`.

[7] L. H. Ozarow and A. D. Wyner. Wire–tap channel II. *The Bell System Technical Journal*, 63(10):2135–2157, 1984.

[8] Thomas J. Richardson and Rüdiger L. Urbanke. Efficient encoding of low-density parity-check codes. *IEEE Transactions on Information Theory*, 47(2):638–656, 2001.

[9] Roxana Smarandache and Pascal O. Vontobel. On regular quasi-cyclic LDPC codes from binomials. In *Proceedings of IEEE International Symposium on Information Theory*, page 274, Chicago, IL, USA, June 27 – July 2 2004.

[10] Andrew Thangaraj, Souvik Dihidar, A. Robert Calderbank, Steven W. McLaughlin, and Jean-Marc Merolla. On the application of LDPC codes to a novel wiretap channel inspired by quantum key distribution. arXiv.org, Report cs.IT/0411003, 2004. `http://arxiv.org/abs/cs.IT/0411003`.

[11] A.D. Wyner. The wire–tap channel. *The Bell System Technical Journal*, 54(8):1355–1387, 1975.

## 7.2 Authentification de données bruitées et cryptosystème de McEliece

Dans l'article [BCD06a] *A Fuzzy Sketch with Trapdoor*, nous construisons un schéma de type *secure sketch* disposant d'une trappe grâce à une utilisation détournée du cryptosystème de McEliece.

À des fins d'authentification, 3 facteurs se distinguent généralement, chacun avec des propriétés très spécifiques :

- ce que l'on possède (comme une carte mémoire ou une carte à puce) permettant de stocker des clés cryptographiques longues et de bonne qualité mais parfois trop coûteux ou pouvant manquer d'ergonomie,
- ce que l'on sait (tel un mot de passe), facile à utiliser mais souvent trop court (ou difficile à retenir),
- et ce que l'on est (ici la biométrie), très facile à utiliser, permanent mais la mesure d'une donnée est soumise à des variations dans le temps et certaines considérations de respect de la vie privée en limitent parfois son usage.

Le premier facteur s'intègre très bien dans les protocoles cryptographiques, et la question de l'intégration adéquate des deux autres facteurs se pose. Par exemple, il est facile de comparer directement deux versions provenant d'une même source (i.e. avec quelques erreurs les différenciant) telles qu'une liste de réponses à des questions particulières, un très long mot passe ou une donnée biométrique, mais les protocoles cryptographiques classiques ne tolérant pas les erreurs, d'autres solutions sont requises pour aboutir à un mode d'authentification sécurisé.

Un des axes de recherche principaux sur ce sujet est alors la construction de primitives cryptographiques absorbant les erreurs grâce à l'utilisation de codes correcteurs d'erreurs. L'introduction par Davida *et al.* [DFM98] de l'idée d'utiliser de la correction dans le cadre de la biométrie, la présentation de deux schémas d'authentification dans le cadre de la biométrie [JW99] puis pour des réponses à une liste de questions [FJ01] sont les précurseurs de ces recherches. Depuis, de nombreux travaux (e.g. [JS06, LT03, DRS04, DS05, DKRS06]) ont alors formalisé ces concepts, examiné diverses variations et analysé leur sécurité, par exemple vis-à-vis de problèmes de ré-utilisation [Boy04] ou contre des attaques actives [BDK+05].

Afin de réconcilier les erreurs, une primitive importante, appelée *secure sketch* (ou *fuzzy sketch*), est définie (cf. Fig. 7.4). Elle est basée sur deux fonctions, une fonction dite de *sketch* Fsk permettant d'extraire une donnée d'aide à partir de la donnée mesurée et une fonction dite de correction Cor pouvant corriger, grâce à la donnée d'aide, une deuxième donnée mesurée. La sécurité est généralement étudiée du point de vue de la théorie de l'information, en particulier la perte d'entropie lors de la publication de la donnée d'aide doit être limitée et il existe alors un compromis avec la capacité de correction finale, d'autant plus que l'on souhaite que seule une donnée proche doit permettre de retrouver la donnée de départ.

Pour plus de liberté, les données impliquées ne sont pas considérées comme directement dans un code mais souvent comme des translatés de mots de codes aléatoires où la donnée d'aide est alors la translation en elle-même (suivant la construction de Juels et Wattenberg [JW99]). Dans le cas d'un code linéaire binaire comme pour l'exemple Fig. 7.5, la capacité de correction de la fonction de correction est la capacité de correction du code, et la perte

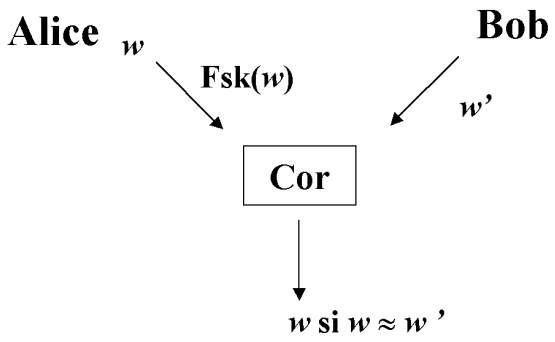


FIG. 7.4 – Principe de secure sketch

d'entropie engendrée par la donnée d'aide est liée à la redondance du code.

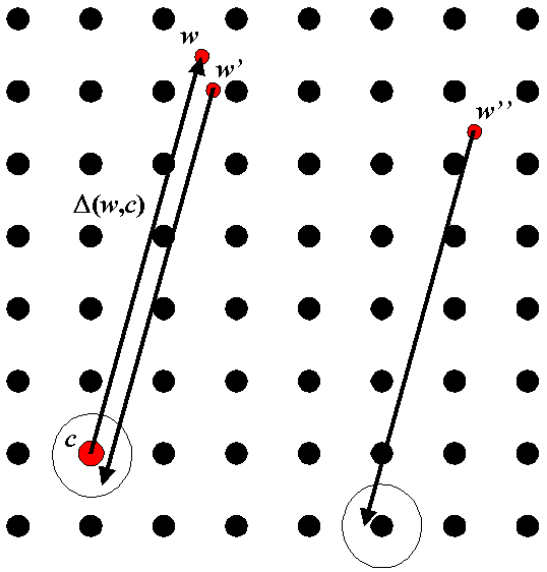


FIG. 7.5 – Un exemple de construction d'un secure sketch

Suivant ce modèle de sécurité, on peut se demander ce qu'il faut faire quand l'entropie des données sources est trop faible ou ce qu'il se passe quand ce sont des données publiques. Dans ce dernier cas, si les *sketches* sont stockés dans une base de données, un attaquant pourrait tester l'appartenance d'une donnée à la base en testant le lien aux *sketches* via la fonction de correction. Dans [BCD06a], nous adressons ces problèmes en proposant, à partir de la construction de [JW99], d'ajouter une sécurité calculatoire au schéma. Afin de travailler avec des données quelconques, l'idée principale est de restreindre l'accès à la fonction de correction.

Pour cela, nous modifions en partie le schéma afin d'y introduire une instance détournée du cryptosystème de McEliece [McE78] de sorte que seul le possesseur de la clé privée (i.e. du

code de Goppa secret sous-jacent) puisse décoder les données altérées. Nous devons pour cela assurer une erreur artificielle additionnelle afin de situer le problème du décodage dans la zone où il s'avère réellement difficile pour un attaquant (voir Fig. 7.6) ; ce qui a pour conséquence une perte de capacité de correction. Ensuite, nous étudions la sécurité, la capacité de correction et les paramètres de ce *secure sketch* à trappe en se basant sur la complexité des meilleurs attaques connues sur le cryptosystème de McEliece.

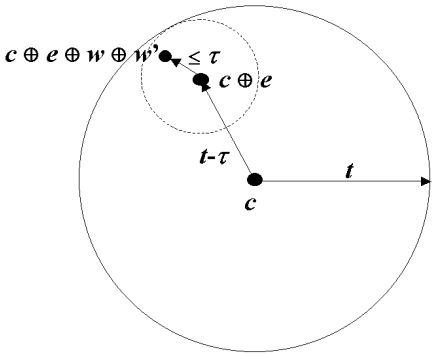


FIG. 7.6 – Décodage au-delà des possibilités d'un attaquant

Cet article a également pour but, par une utilisation nouvelle du cryptosystème de McEliece, de promouvoir l'intérêt de la cryptographie à clé publique basée sur la théorie des codes. Finalement, il serait intéressant de tester l'applicabilité de l'idée à d'autres primitives tolérantes aux erreurs.

### 7.2.1 Exemple d'application

L'intérêt de cette solution par rapport à une couche de chiffrement en supplément d'un *secure sketch* est qu'ici il n'est pas nécessaire de "déchiffrer" le *sketch* (la donnée d'aide) avant d'appliquer la fonction de correction (il faut donc simultanément la clé et une donnée proche de la donnée originale pour récupérer cette dernière). Pour expliquer cet avantage, nous détaillons un exemple d'utilisation pratique ci-dessous.

On considère un protocole d'authentification biométrique du type client-serveur, où les communications entre le client et le serveur sont protégées en intégrité et en confidentialité, avec un client correspondant à un capteur biométrique $C$ et un serveur formé d'une base de données biométriques $\mathcal{B}$ et d'un module de comparaison (matching) $\mathcal{M}$. Le fonctionnement est le suivant :

- Un utilisateur $U_i$ présente son identité $i$ et une donnée biométrique $w'_i$ qui est alors capturée par $C$.
- Le capteur envoie la donnée $w'_i$ à $\mathcal{M}$ et l'identité $i$ à la base $\mathcal{B}$.
- Celle-ci transmet alors la donnée stockée de référence $z_i$ à $\mathcal{M}$ qui peut ainsi décider si l'authentification est valide à partir des données $w'_i$ et $z_i$.

On suppose de plus que $\mathcal{M}$ nécessite l'utilisation de secrets spécifiques pour fonctionner et que le capteur $C$ s'assure que les données biométriques proviennent d'une personne vivante

(principe de détection du vivant dans un système biométrique [MMJP03]) ; si bien que $\mathcal{M}$ n'utilise les secrets nécessaires à son fonctionnement que si les données biométriques sont des données validées par $C$. En pratique, cela implique que $\mathcal{M}$ ne peut pas être utilisé sans passer par $C$, i.e. que le système ne fonctionne que dans le cas d'une tentative d'authentification d'une personne vivante.

On distingue deux situations. Dans le premier protocole, on utilise un *secure sketch* classique défini par une fonction publique de *sketch* Fsk et une fonction publique de correction Cor. On dispose également d'un cryptosystème asymétrique défini par une fonction de génération de clés, une fonction de chiffrement et une fonction de déchiffrement, respectivement $(\mathcal{K}, \mathcal{E}, \mathcal{D})$. Le module $\mathcal{M}$ possède un couple de clés $(pk, sk)$ et garde $sk$ secret. On utilise la couche de chiffrement supplémentaire comme suit.

1. La phase d'enrôlement consiste à stocker dans la base $\mathcal{B}$ la valeur de référence

$$z_i = \mathcal{E}(\mathrm{Fsk}(w_i); pk)$$

   à partir d'une donnée biométrique de référence $w_i$ pour l'utilisateur $U_i$.

2. La phase de vérification opérée par $\mathcal{M}$ est de déchiffrer $\mathrm{Fsk}(w_i) = \mathcal{D}(z_i; sk)$ puis d'appliquer la fonction de correction en calculant $\mathrm{Cor}(\mathrm{Fsk}(w_i), w_i')$. L'authentification sera un succès si la valeur récupérée correspond à $w_i$ (cette dernière vérification peut par exemple être effectuée en stockant le haché du mot de code de référence pour le comparer avec le mot de code après correction – cf. [JW99]).

Dans le deuxième protocole, on utilise uniquement le *secure sketch* à trappe tel qu'il est défini dans [BCD06a] via la fonction publique de *sketch* $\mathrm{Fsk}_{\mathrm{trap}}$ et la fonction privée de correction $\mathrm{Cor}_{\mathrm{trap}}$. Le module $\mathcal{M}$ est le seul à connaître cette fonction $\mathrm{Cor}_{\mathrm{trap}}$.

1. La phase d'enrôlement consiste à stocker dans la base $\mathcal{B}$ la valeur de référence

$$z_i = \mathrm{Fsk}_{\mathrm{trap}}(w_i)$$

   à partir d'une donnée biométrique de référence $w_i$ pour l'utilisateur $U_i$.

2. La phase de vérification au sein de $\mathcal{M}$ est d'appliquer la fonction de correction secrète en calculant $\mathrm{Cor}_{\mathrm{trap}}(z_i, w_i')$ et l'authentification est réussie si la valeur récupérée correspond à $w_i$.

Dans les deux cas, la confidentialité des clés possédées uniquement par $\mathcal{M}$ (i.e. $sk$ ou $\mathrm{Cor}_{\mathrm{trap}}$) permet d'augmenter la sécurité du stockage des données biométriques de référence dans la base $\mathcal{B}$. En effet, sans l'utilisation du secret, il sera en pratique très difficile de retrouver la donnée biométrique de départ $w_i$ à partir de $z_i$ même en connaissant une donnée proche $w_i'$.

Or les données biométriques étant des données publiques, il ne faut pas non plus que des tests d'appartenance à la base $\mathcal{B}$ ou des recoupements avec d'autres bases soient possibles. Comme précisé dans le paragraphe ci-dessus, si on ne dispose pas des clés possédées par $\mathcal{M}$, ce n'est pas possible en ne connaissant que les données stockées $z_i$. Par contre, seul le deuxième

protocole permet de se protéger efficacement contre un attaquant $\mathcal{A}$ qui peut observer une partie des traitements. Si l'attaquant $\mathcal{A}$ peut obtenir (e.g. en espionnant le serveur) le résultat des opérations de déchiffrement et de correction effectuées par $\mathcal{M}$, alors la différence entre les deux protocoles vient de l'opération de déchiffrement du premier protocole qui est réalisée quelque soit la valeur de la donnée biométrique d'authentification $w'_i$.

Supposons que $\mathcal{A}$ ait capturé une donnée biométrique $w$ appartenant à un individu $U$ pour lequel $\mathcal{A}$ souhaite savoir s'il est enregistré dans la base $\mathcal{B}$.

- Dans les deux cas, ne disposant pas des clés de $\mathcal{M}$, il ne peut pas effectuer les traitements à sa place.
- De plus, ne pouvant pas présenter la donnée biométrique capturée $w$ directement au capteur $C$ puisqu'elle n'est alors plus attachée à la personne vivante $U$, il ne peut pas demander directement à $\mathcal{M}$ de comparer $w$ avec une donnée de référence $z_i$ ($i$ quelconque).
- Pour récupérer des informations liées à $z_i$, il peut par contre essayer de s'authentifier en prétendant être l'utilisateur $i$ tout en présentant sa propre donnée biométrique au capteur (qui est donc bien attachée à une personne vivante). L'authentification échouera certainement, donc dans les deux protocoles l'étape de correction ne donnera pas de résultat exploitable pour $\mathcal{A}$, cependant dans le premier protocole, cela lui permet d'obtenir le déchiffrement de $z_i$, à savoir $\mathrm{Fsk}(w_i)$ qu'il pourra alors comparer *hors-ligne* à la donnée $w$.

Cela montre bien l'avantage d'avoir une fonction de correction privée par rapport à une simple couche de chiffrement supplémentaire qui nécessite de déchiffrer avant correction sans savoir à l'avance si les données traitées sont légitimement proches ou non.

# A Fuzzy Sketch with Trapdoor*

Julien Bringer[1] Hervé Chabanne[1] Quoc Dung Do[2]

[1]Sagem Défense Sécurité,
[2]Ecole Polytechnique.

**Abstract**

In 1999, Juels and Wattenberg introduce an effective construction of Fuzzy Sketch, i.e. a way of handling errors into string verification. This allows them to consider data varying in time, such as, for instance, answers to a list of subjective questions. To this end, they utilize an Error Correcting Code.

We here show how to embed a trapdoor into Fuzzy Sketches, reducing to authorized people the ability to correct errors and thus to verify the fuzzy equality to the Fuzzy Sketch.

**Keywords.** Cryptosystem of McEliece, Fuzzy Sketch.

## 1 Introduction

Handling errors into verification comes from [8]. In [14], Juels and Wattenberg give an effective construction of Fuzzy Sketches for the Hamming space. I.e. denoting $Fsk(w)$ the Fuzzy Sketch of $w$, from $Fsk(w)$ and $w \oplus e$ where $e$ symbolized some errors, one can recover $w$. Later on, [13] describes a Fuzzy Vault which can be interpreted as a Fuzzy Sketch for the set difference metric. Finally, [9] sets a general formal framework and defines several techniques to be used for those distances as well as for the edit distance. We will here only consider binary Hamming space.

Fuzzy Sketches are often associated to biometric information but some other data can also be considered. For instance, [11] proposes a fault tolerant scheme allowing very long passwords and some errors.

Previous works [3, 4] make the assumption that they are dealing with secret information. Here, we do not take this hypothesis as granted and consider rather that we are working with public fuzzy data. In our setting, one can, for instance, store fuzzy sketches inside a database to enforce control

*This work was exclusively supported by Sagem Défense Sécurité.

access to some assets. We here stress that letting an attacker the correcting capability of the underlying error correcting code may lead to awkward situations. For example, if he has access to the database, he will be able to check the membership of a given person, which could be unacceptable for privacy reasons.

We show that a slight modification of the Fuzzy Sketch of Juels and Wattenberg allows to embed a trapdoor in this construction restricting the ability to check the fuzzy equality. Starting from the cryptosystem of McEliece [16], we exploit the fact that there is today a gap between the errors which can be corrected, for a given security parameter, by an attacker and the error correction capacity of those who have knowledge of the trapdoor and of the specialized decoding algorithms for the underlying Goppa codes. In fact, in twenty years, the work factor of the best known attack of the original McEliece cryptosystem has decreased from $2^{81}$ [16] to $2^{49}$ [12] but the complexities of the attacks are still exponential.

## 2 Preliminaries and notations

### 2.1 Coding theory

Let $w$ be a word of $\{0,1\}^n$, the *Hamming weight* $\mathrm{wt}_H(w)$ of $w$ is the number of coordinates of $w$ which are non-zero. The *Hamming distance* over $\{0,1\}^n$ is the canonical metric distance defined as the number of differences between two words. Hence, for two words $w_1, w_2$, the Hamming distance $\mathrm{d}_H(w_1, w_2)$ is equivalent to the weight of $w_1 \oplus w_2$ for a binary alphabet. For some integer $n$, the set $\{0,1\}^n$ equipped with the Hamming distance $\mathrm{d}_H$ is an *Hamming space* $\mathcal{H}$.

An *error correcting code* in $\mathcal{H}$ is a subset $C$ of $\mathcal{H}$, the elements of $C$ are called codewords, the *minimum distance* $d_{min}$ of $C$ is the smallest distance between two distinct codewords. This means that one can detect up to $d_{min} - 1$ errors in a codeword. The *capacity of correction* $t$ of $C$ is the radius of the largest ball for which for any $w \in \mathcal{H}$ there is at most one codeword in the ball of radius $t$ centered on $w$. In fact for the Hamming distance $\mathrm{d}_H$,

$$t = \lfloor (d_{min} - 1)/2 \rfloor.$$

As $\{0,1\}$ can be equipped with the structure of the finite field $\mathbb{F}_2$, $\mathcal{H}$ is also a vector space $\mathbb{F}_2^n$ and we define on it a linear code $C$ as a vector subspace of $\mathcal{H}$. If $k$ is the dimension of the vector space $C$ over $\mathbb{F}_2$, we denote $C$ as a

$[n, k, d_{min}]$ code over $\mathbb{F}_2$. Due to linearity property, the minimum distance $d_{min}$ is then equal to the minimum weight of the codewords $w \in C - \{0\}$. Moreover, $C$ can be described by a matrix $G$, call a *generator matrix* of $C$, with $k$ lines, $n$ rows and a rank $k$, such that the lines of $G$ is a basis of $C$: we obtain

$$C = \{mG \mid m \in \mathbb{F}_2^k\}.$$

**Remark 1** *For a random linear code given by its generator matrix, we then easily encode the codewords but the decoding process is not a computationally efficient function in general. We will see further how this fact is used in the cryptosystem of McEliece.*

### 2.2 Fuzzy Sketches

Let $X$ and $Y$ be two random discrete variables. We recall the definition of the *entropy* of $X$,

$$\mathbf{H}(X) = \mathbf{E}_{x \leftarrow X}(-\log_2 \mathbb{P}(X = x)),$$

and the definition of the *conditional entropy* of $X$ given $Y$,

$$\begin{aligned} \mathbf{H}(X \mid Y) &= \mathbf{E}_{y \leftarrow Y} \mathbf{H}(X \mid Y = y) \\ &= \mathbf{E}_{x,y \leftarrow X,Y}(-\log_2 \mathbb{P}(X = x \mid Y = y)). \end{aligned}$$

We also give the *min-entropy* of $X$ defined as

$$\mathbf{H}_\infty(X) = -\log_2 \max_x \mathbb{P}(X = x)$$

and the *average min-entropy* for $X$ given $Y$ defined as

$$\begin{aligned} \overline{\mathbf{H}}_\infty(X \mid Y) &= -\log_2 \mathbf{E}_{y \leftarrow Y}(2^{-\mathbf{H}_\infty(X|Y=y)}) \\ &= -\log_2 \mathbf{E}_{y \leftarrow Y}(\max_x \mathbb{P}(X = x \mid Y = y)), \end{aligned}$$

which are more relevant for cryptographic use. Namely, the *average min-entropy* of $X$ given $Y$ is the logarithm of the average probability of the most likely value of $X$ given $Y$. This notion is useful to measure the difference from uniform distributions. For example, if $Y$ has values in $\{0,1\}^n$ then,

$$\overline{\mathbf{H}}_\infty(X \mid Y) \geq \mathbf{H}_\infty(X) - n.$$

This notion allows us to introduce the definition of fuzzy sketches:

**Definition 1** *A $(\mathcal{H}, m, m', t)$-fuzzy sketch is a pair of functions $(Fsk, Cor)$ where:*

- *$Fsk$ is a (typically randomized thanks to a randomizer $x$) sketching function that on input $w \in \mathcal{H}$ outputs a sketch or redundancy data $P \in \{0,1\}^*$, such that for all random variable $W$ over $\mathcal{H}$ with min-entropy $\mathbf{H}_\infty(W) \geq m$, the average min-entropy of $W$ given $Fsk(W)$ is at least $m'$. That is,*

$$\overline{\mathbf{H}}_\infty(W \mid Fsk(W)) \geq m'.$$

- *$Cor$ is a correction function which allows to recover $w$ from its sketch and any vector close to $w$: given a word $w' \in \mathcal{H}$ and a sketch $P$, it outputs a word $w'' \in \mathcal{H}$, such that for any $P = Fsk(w)$ and $\mathrm{d}_H(w, w') \leq t$, it holds that $w'' = w$.*

In practice, the sketching and correction functions, $Fsk$ and $Cor$, have to be efficiently computable, that is they run in polynomial time in $n$. Here, we denote the parameter $t$ as the *correction capacity of the fuzzy sketch.*

The following construction has been considered by Juels and Wattenberg:

**Claim 1 (Code-offset construction [14])** *Given $C$ a binary linear code of length $n$, dimension $k$ and correction capacity $t$, the Fuzzy Sketch of Juels and Wattenberg is given by*

$$Fsk_{\mathrm{JW}}(w; x) = w \oplus c(x)$$

*where*

- *$x$, the randomizer of $Fsk_{\mathrm{JW}}$, is a random vector of length $k$,*
- *$c(x)$ is taken at random from $C$.*

*This yields a $(\mathcal{H}, m, m + k - n, t)$-fuzzy sketch.*

**Remark 2** *The value $m - m' = n - k$ represents the entropy loss of the fuzzy sketch. When $m$ is not big enough, the amount $m + k - n$ may not be sufficient to ensure the security of the sketch from an information theory point of view. We here show how to strengthen this construction in order to add a "computational security" which stands even when $m$ is small (see also Remark 5).*

# 3 Our construction

## 3.1 Cryptosystem of McEliece

We place ourselves into a hard instance of the McEliece cryptosystem, i.e. let $d = c \oplus e$ where

- $c$ is a word taken at random from a concealed Goppa code,
- $e$ stands for errors of sufficient weight,

such that there is no computational effective way, for a given security parameter $\Sigma$, to recover $c$ from $d$ without the knowledge of the trap. We would want then to determine the capacity to handle additional errors using the original Goppa code.

**Claim 2 (Goppa codes)** *Corresponding to each irreducible polynomial of degree $t$ over $GF(2^m)$, there exists a binary irreducible Goppa code of length $n = 2^m$, dimension $k \geq n - tm$, capable of correcting any pattern of $t$ or fewer errors. The decoding of these codes can be done in $O(nt)$ operations.*

The cryptosystem of McEliece is an asymmetric cryptosystem defined as follows:

**Claim 3 (Cryptosystem of McEliece [16])**
*We consider the two codes defined by the matrices $G$ and $G_{\text{pub}}$ such that*

- *$G$ is a $k \times n$ generator matrix of a Goppa code,*
- *$G_{\text{pub}} = SGP$ where $S$ is a $k \times k$ random invertible dense matrix and $P$ a random $n \times n$ permutation matrix.*

*In the McEliece cryptosystem, $G_{\text{pub}}$ is a public data and the matrices $G$, $S$ and $P$ are kept secret, they constitute the underlying trapdoor.*

*Let $d = c \oplus e$ where $c = xG_{\text{pub}}$ is a codeword with $x$ a random binary vector of length $k$, and $e$ stands for some errors ($\text{wt}_H(e) \leq t$), then, given the properties of the cryptosystem, we have these facts:*

- *Knowledge of the trapdoor allows to recover $c$ from $d$ in polynomial running time using the decoding algorithm of the underlying Goppa code,*

- *without the trapdoor, a basic attack is to perform*

$$o(1)\binom{n}{k}/\binom{n-\mathrm{wt}_H(e)}{k}$$

*guesses to find $k$ columns of $G_{\mathrm{pub}}$ with no errors enabling the recovery of $c$. For each guess, $k^3$ binary operations must be performed.*

Improvements of the basic attack are numerous, see [1, 2, 15, 17, 5, 12], leading to a complexity $C_{\mathrm{alg}} = N_{\mathrm{alg}} \times C_{\mathrm{iter}}$ in the basic algorithm of [12] where

$$N_{\mathrm{alg}} = \left( \sum_{j=1}^{M} (-1)^{j+1} \binom{M}{j} \left( \sum_{i=0}^{p} \frac{\binom{k}{i}\binom{n-k-l}{r-i}}{\binom{n}{r}} \right)^{j} \right)^{-1}, \tag{1}$$

$$C_{\mathrm{iter}} = \frac{nk}{2}(k+M) + Ml + \sum_{i=1}^{p} \binom{k}{i} il + \frac{\sum_{j=0} p\binom{k}{j} Mn(j+1)}{2^l}. \tag{2}$$

This complexity is obtained for an $[n, k]$ linear code, $M$ received codewords with at most $r$ errors (below the correction capacity) and for two other algorithm parameters $p$, $l$. For example, some optimal complexities with some modifications and well chosen parameters $M$, $p$ and $l$ are given below [12].

| $M$ | $C_{\mathrm{alg}}$ for $n = 1024$, $k = 524$, $t = 50$ |
|---|---|
| $2^{15}$ | $2^{56.2}$ |
| $2^{30}$ | $2^{50.2}$ |
| $2^{40}$ | $2^{49.0}$ |

Table 1: Complexity of the attack [12] for different number $M$ of received words.

Usually, instances of McEliece's cryptosystem are made such that the error weight is taken equal to $t$, at the limit of the correction capacity of the underlying Goppa code. However, if it is not the case, one can note that the complexity of known attacks is a growing function of this error weight. For a given security parameter, and considering an increasing length $n$, there is some place for correcting additional errors to $e$.

The following table gives for $\Sigma = 80$, i.e. a work factor greater than $2^{\Sigma}$, the corresponding minimum weight of errors for the cryptosystem of

McEliece according to the best known attack [12] with $M = 2^{20}$. In other words and for example, for $n = 2048$, $k = 1024$, when $\mathrm{wt}_H(e) \geq 78$, an attacker must perform more than $2^{80}$ operations to decode. For $n \leq 1024$, for all parameters $n$ and $k$, the attacker gets a work factor smaller than $2^{80}$.

| $k/n$ | $n = 2048$ | $n = 4096$ | $n = 8192$ |
|---|---|---|---|
| 0.2 | – | 235 | 232 |
| 0.4 | 105 | 104 | 101 |
| 0.5 | 78 | 77 | 74 |
| 0.6 | 60 | 59 | 56 |
| 0.8 | 35 | 34 | 33 |

Table 2: Errors minimum weight for a work factor greater than $2^{80}$

Note that when $n$ increases, for a given $k/n$, $\mathrm{wt}_H(e)$ slowly decreases.

### 3.2 Fuzzy sketch with a trapdoor

We use the fuzzy sketch construction of Juels and Wattenberg (see Claim 1) in order to introduce a trapdoor. We point out that utilizing $Fsk_{\mathrm{JW}}$ with a trapdoor is indeed possible:

**Definition 2 (Fuzzy Sketch with Trapdoor)** *Let $G_{\mathrm{pub}}$ be as in Claim 3 and $C$ the corresponding code of dimension $k$, then a Fuzzy Sketch with trapdoor is given by*

$$Fsk^{\Sigma}_{\mathrm{trap}}(w; x) = w \oplus c(x) \oplus e(x),$$

*where*

- *$x$ is a random vector of length $k$,*
- *$c(x) = xG_{\mathrm{pub}}$ is a codeword of $C$,*
- *$\Sigma$ is a security parameter,*
- *$e(x)$ stands for a well chosen error of weight $\mathrm{wt}_H(e) = \varpi$ with $\varpi$ computed from $n$, $k$ and $\Sigma$.*

In the sequel, we consider that $e(x)$ is taken in a pseudo random way leaded by $x$ from the set of words of length $n$ and weight $\varpi$.

A correction function is allowed for the possessor of the trapdoor as from $w' = w \oplus f$ and $Fsk^{\Sigma}_{\text{trap}}(w; x)$, we get

$$w' \oplus Fsk^{\Sigma}_{\text{trap}}(w; x) = c(x) \oplus e(x) \oplus f,$$

which allows to compute $c(x)$ whenever $e(x) \oplus f$ has a Hamming weight smaller than the decoding capacity of the underlying Goppa code. From $c(x)$, we get $x$ and $e(x)$ and finally $w$. Thus, the same randomizer $x$ is used for computing $c$ and $e$.

We will denote $Cor_{\text{trap}}$ the corresponding procedure. Note that this function $Cor_{\text{trap}}$ is efficient only for those who know the trapdoor. For the others, our $Fsk^{\Sigma}_{\text{trap}}$ is made such that no correction function operates in less than $2^{\Sigma}$ steps, with an overwhelming probability. Indeed, the weight of errors $e(x) \oplus f$ needs to be very small to be corrected without the knowledge of the trapdoor.

We have

$$\varpi - \text{wt}_H(f) \leq \text{wt}_H(e(x) \oplus f) \leq \varpi + \text{wt}_H(f).$$

Let $\tau = t - \varpi$. For given $n$, $k$ and $\Sigma$, we compute $\varpi$ such that $\varpi - \tau$ errors leads to a work factor for an attacker greater than $2^{\Sigma}$ (as we have done for $\Sigma = 80$ during Section 3.1), see Table 3.

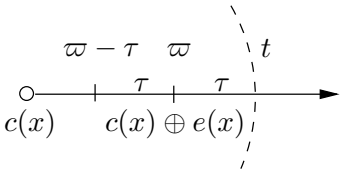


Figure 1: The choice of $\varpi$ and $\tau$ according to the correction capacity $t$.

**Remark 3** *The error $e(x)$ is chosen such that*

$$\text{wt}_H(e(x) \oplus w) \geq \varpi - \tau$$

*in order to avoid the fuzzy sketch to be invertible without a vector $w'$ close to $w$. Whenever this inequality does not hold, we simply choose a new randomizer $x$.*

With our construction, the possessor of the trapdoor can use the correction function $Cor_{\text{trap}}$ for an error of weight $\text{wt}_H(e(x) \oplus f) \leq t$, as for instance

it always happens whenever $\mathrm{wt}_H(f) \leq \tau$. Letting its capacity correction be $\tau$, our Fuzzy Sketch with trapdoor has nearly the same behaviour in terms of min-entropy than those of code-offset construction (see Claim 1).

**Remark 4** *Note that even the possessor of the trapdoor can not recover $w$ from $Fsk^{\Sigma}_{\mathrm{trap}}(w; x)$ without $w'$ close to $w$, if $w$ has not a weight too small.*

**Remark 5** *In [9, Lemma 3.1 (Fuzzy Extractors from Sketches)], Fuzzy Extractors are introduced to alleviate the anticipated poor behaviour of $w$ and then $Fsk(w)$ in terms of min-entropy. Our construction has the same goal. Nevertheless, if needed, it is still possible to apply Fuzzy Extractor to our scheme because the involved technique of [9] is quite general and needs only to operate to have Fuzzy Sketches as defined in Definition 1.*

### 3.3 Correction capacity

We here place ourselves in a specific security setting, as $Fsk^{\Sigma}_{\mathrm{trap}}$ is designed to resist to known attacks of McEliece up to a given security parameter. Using the best cryptanalysis algorithm known to date [12], we compute $\tau$ for a security parameter $\Sigma = 80$:

| $k/n$ | $\tau$ for $n = 2048$ | $\tau$ for $n = 4096$ | $\tau$ for $n = 8192$ |
|---|---|---|---|
| 0.2 | – | 19 | 136 |
| 0.4 | 3 | 50 | 139 |
| 0.5 | 7 | 47 | 121 |
| 0.6 | 7 | 39 | 98 |
| 0.8 | 1 | 17 | 47 |

Table 3: $\tau$ for the security parameter $\Sigma = 80$

Note that for a given $k/n$, $\tau$ grows together with $n$.

In Section 3.2, we have seen how the capacity correction of our fuzzy sketch depends on the capacity correction of the underlying code and on the security parameter $\Sigma$. Indeed, the capacity correction of this fuzzy sketch with trapdoor is $\tau$ and it has been computed according to $\Sigma$ (see Table 3).

However, not only the errors $f$ of weight $\tau$ can be handled, i.e. much more than $\sum_{i=0}^{\tau} \binom{n}{i}$ errors can be corrected. Actually, if $f$ has a weight

$$\tau + 1 \leq \mathrm{wt}_H(f) \leq 2\varpi + \tau$$

and has enough intersection with $e$, than it can be corrected. Hence the total number of errors $f$, for which the correction function works correctly, increases to

$$\sum_{i=0}^{\tau} \binom{n}{i} + \sum_{i=1}^{2\varpi} \binom{\varpi}{\lceil i/2 \rceil} \binom{n - \lceil i/2 \rceil}{\tau + \lfloor i/2 \rfloor}. \tag{3}$$

| $k/n$ | $n = 2048$ | $n = 4096$ | $n = 8192$ |
|---|---|---|---|
| 0.2 | – | 6.7% | 6.1% |
| 0.4 | 5.4% | 5% | 4.6% |
| 0.5 | 4.5% | 4.1% | 3.8% |
| 0.6 | 3.6% | 3.3% | 3.1% |
| 0.8 | 1.8% | 1.7% | 1.5% |

Table 4: Equivalent rate of number of errors $f$ correctly processed

As an example, for $n = 4096$, $R = k/n = 0.2$, the designed capacity correction $\tau = 19$ of our fuzzy sketch with trapdoor is about 0.5% of $n$, but according to (3), the correction function can correct the same number of errors $f$ as a function with a capacity correction 273, which is nearly 6.7% of $n$. So, $Cor_{\text{trap}}$ succeeds in managing correctly a large amount of errors. See Table 4 for other values obtained via (3).

# 4 Applications

To illustrate our scheme, we briefly describe here two applications with especially the wish to protect personal data.

Assume that a central authority holds a database containing "traditional" personal data used for authentication purpose, and that some access point are linked to this database in order to check users. Each point of access comes equipped with $G_{\text{pub}}$. Whenever someone requires an asset, the corresponding Fuzzy Sketch with trapdoor is computed. Depending on context, the access point sends the sketch at once or sends regularly a list containing several sketches to the authority. Hence, using the original personal data in the central database, the authority can apply the correction function with the trapdoor and so comparing with its reference, the authority can take the appropriate decision. This application is intended to protect the processing and the transmission to the database of personal data.

Now, suppose we want to avoid an attacker to have access to personal data in the database or to learn the membership of a given person directly from the database. We then construct the central database, not with personal data, but with the corresponding fuzzy sketches. Hence without the trapdoor, an attacker would not succeed in using the correction function $Cor_{\text{trap}}$ (or an equivalent one), and so would not be able to link a personal data to a fuzzy sketch stored in the database. However, the authority, which possesses the knowledge of the trapdoor, can compute $Cor_{\text{trap}}$ when an access point sends it data.

## 5 Conclusion

We show how to include a trapdoor into the Fuzzy Sketches of Juels and Wattenberg.

This renewal in the utilization of the cryptosystem of McEliece can also be viewed as a way of encrypting fuzzy data. And we hope that this will incite to retain more attention on public-key cryptosystem based on error-correcting codes. In particular, the correction capacity of – what we call – McEliece channel (errors that can be added to a hard instance of the cryptosystem of McEliece and corrected) has to be improved.

We give two examples of use of our ideas in section 4 and hope that our work might serve as inspiration for future work in this area, leading to new applications.

## Acknowledgment

The authors wish to thank the anonymous referees for their useful suggestions which improved the presentation of this paper.

## References

[1] C.M. Adams and H. Meijer, *Security-Related Comments Regarding McEliece's Public-Key Cryptosystem*, Advances in cryptology – CRYPTO 1987, pp. 224–228.

[2] C.M. Adams and H. Meijer, *Security-related comments regarding McEliece's public-key cryptosystem*, IEEE Transactions on Information Theory, vol. 35(2), 1989, pp. 454–455.

[3] X. Boyen, *Reusable cryptographic fuzzy extractors*, ACM Conference on Computer and Communications Security 2004, pp. 82–91.

[4] X. Boyen, Y. Dodis, J. Katz, R. Ostrovsky and A. Smith, *Secure Remote Authentication Using Biometric Data*, Advances in cryptology – EUROCRYPT 2005, pp. 147–163.

[5] A. Canteaut and F. Chabaud, *A New Algorithm for Finding Minimum-Weight Words in a Linear Code: Application to McEliece's Cryptosystem and to Narrow-Sense BCH Codes of Length 511*, IEEE Transactions on Information Theory, vol. 44(1), 1998, pp. 367–378.

[6] G. Cohen and G. Zémor, *The wire-tap channel applied to biometrics*, ISITA2004, Parma, Italy, October 10-13, 2004.

[7] G. Cohen and G. Zémor, *Generalized coset schemes for the wire-tap channel: application to biometric*, In IEEE International Symp. on Information Theory, 2004.

[8] G. Davida, Y. Frankel, and B. Matt, *On enabling secure applications through offline biometric identification*, In Proc. IEEE Symp. on Security and Privacy, 1998, pp. 148–157.

[9] Y. Dodis, L. Reyzin and A. Smith, *Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data*, Advances in cryptology – EUROCRYPT 2004, pp. 523–540.

[10] I. Dumer, D. Micciancio and M. Sudan, *Hardness of approximating the minimum distance of a linear code*, IEEE Transactions on Information Theory, vol. 49(1), 2003, pp. 22–37.

[11] N. Frykholm and A. Juels, *Error-tolerant password recovery*, ACM Conference on Computer and Communications Security 2001, pp. 1–9.

[12] T. Johansson and F. Jönsson, *On the complexity of some cryptographic problems based on the general decoding problem*, IEEE Transactions on Information Theory, vol. 48(10), 2002, pp. 2669–2678.

[13] A. Juels and M. Sudan, *A Fuzzy Vault Scheme*, In IEEE International Symp. on Information Theory, 2002.

[14] A. Juels and M. Wattenberg, *A Fuzzy Commitment Scheme*, ACM Conference on Computer and Communications Security 1999, pp. 28–36.

[15] P. J. Lee and E. F. Brickell, *An Observation on the Security of McEliece's Public-Key Cryptosystem*, Advances in cryptology – EUROCRYPT 1988, pp. 275–280.

[16] R.J. McEliece, *A public-key cryptosystem based on algebraic coding theory*, JPL DSN Progress Report, 1978, pp. 114–116.

[17] J. van Tilburg, *On the McEliece Public-Key Cryptosystem*, Advances in cryptology – CRYPTO 1988, pp. 119–131.

# Bibliographie


[AL05] Yves Aubry and Philippe Langevin. On the weights of irreducible cyclic codes. In *Proceedings of International Workshop on Coding and Cryptography (WCC 2005)*, 2005.

[Ars05] Gwénolé Ars. *Applications des bases de Gröbner à la cryptographie*. PhD thesis, Université de Rennes I, 2005.

[BBNP05] Yuri Borissov, An Braeken, Svetla Nikova, and Bart Preneel. On the covering radii of binary Reed-Muller codes in the set of resilient boolean functions. *IEEE Transactions on Information Theory*, 51(3) :1182–1189, 2005.

[BC06] Julien Bringer and Hervé Chabanne. On the wiretap channel induced by noisy tags. In *Third European Workshop on Security and Privacy in Ad hoc and Sensor Networks – ESAS*, 2006.

[BCD06a] Julien Bringer, Hervé Chabanne, and Quoc Dung Do. A fuzzy sketch with trapdoor. *IEEE Transactions on Information Theory*, 52(5) :2266–2269, 2006. See also Cryptology ePrint Archive, Report 2005/331, `http://eprint.iacr.org/`.

[BCD06b] Julien Bringer, Hervé Chabanne, and Emmanuelle Dottax. $HB^{++}$ : a lightweight authentication protocol secure against some attacks. In *IEEE International Conference on Pervasive Services, Second International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing (SecPerU 2006), 29 June 2006, Lyon, France*, pages 28–33, 2006.

[BCG+06] Neil Bird, Claudine Conrado, Jorge Guajardo, Stefan Maubach, Geert-Jan Schrijen, Boris Skoric, A.M.H. Tombeur, Peter Thueringer, and Pim Tuyls. ALGSICS — Combining Physics and Cryptography to Solve RFID. In *1st Benelux Workshop on Information and System Security – WISSec*, 2006.

[BCL06] Lilya Budaghyan, Claude Carlet, and Gregor Leander. A class of quadratic APN binomials inequivalent to power functions. Cryptology ePrint Archive, Report 2006/445, 2006. `http://eprint.iacr.org/`.

[BCP92] Richard A. Brualdi, Ning Cai, and Vera S. Pless. Orphan structure of the first-order Reed-Muller codes. *Discrete Math.*, 102(3) :239–247, 1992.

[BDK+05] Xavier Boyen, Yevgeniy Dodis, Jonathan Katz, Rafail Ostrovsky, and Adam Smith. Secure remote authentication using biometric data. In Ronald Cramer, editor, *EUROCRYPT*, volume 3494 of *Lecture Notes in Computer Science*, pages 147–163. Springer, 2005.

[BEW98] B.C. Berndt, R.J. Evans, , and K.S. Williams. *Gauss and Jacobi Sums*. Wiley-Interscience, N.Y., 1998.

[BFKL93] Avrim Blum, Merrick L. Furst, Michael J. Kearns, and Richard J. Lipton. Cryptographic primitives based on hard learning problems. In *Advances in Cryptology – CRYPTO'93*, Lecture Notes in Computer Science, pages 278–291. Springer-Verlag, 1993.

[BGL05] Julien Bringer, Valérie Gillot, and Philippe Langevin. Exponential sums and boolean functions. In J-F. Michon, P. Valarcher, and J-B. Yunès, editors, *Boolean Functions : Cryptography and Applications*, BFCA'05, pages 77–87, Rouen, 2005.

[BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting coding and decoding : Turbo-codes. In *Proc. 1993 IEEE International Conference on Communications, Geneva, Switzerland*, pages 1064–1070, 1993.

[BKW00] Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. In *STOC 2000*, pages 435–440, 2000.

[BL03] Eric Brier and Philippe Langevin. Classification of boolean cubic forms in nine variables. In *IEEE Information Theory Workshop*, pages 179–182, 2003.

[BMvT78] Elwyn R. Berlekamp, Robert J. McEliece, and Henk C. van Tilborg. On the inherent intractability of certain coding problems. *IEEE Transactions on Information Theory*, 24(3) :384–386, 1978.

[BMW82] L.D. Baumert, W.H. Mills, and R.L. Ward. Uniform cyclotomy. *Journal of Number Theory*, 14 :67–82, 1982.

[Boy04] Xavier Boyen. Reusable cryptographic fuzzy extractors. In Vijayalakshmi Atluri, Birgit Pfitzmann, and Patrick Drew McDaniel, editors, *ACM Conference on Computer and Communications Security*, pages 82–91. ACM, 2004.

[Bri01] Julien Bringer. Codes cycliques sur $\mathbb{Z}_4$. Mémoire de D.E.A, Université d'Aix-Marseille II, 2001.

[Bri04] Julien Bringer. Nonlinearity of some Patterson-Wiedemann type functions. In *Yet "Another" Conference on Cryptography, YACC'04*, Porquerolles, France, 2004.

[BS90] Eli Biham and Adi Shamir. Differential cryptanalysis of DES-like cryptosystems. In Alfred Menezes and Scott A. Vanstone, editors, *CRYPTO*, volume 537 of *Lecture Notes in Computer Science*, pages 2–21. Springer, 1990.

[BW72] E.R. Berlekamp and L.R. Welch. Weight distributions of the cosets of the (32 ; 6) Reed-Muller code. *IEEE Transactions on Information Theory*, 18(1) :203–207, 1972.

[CA06] Claude Castelluccia and Gildas Avoine. Noisy tags : A pretty good key exchange protocol for RFID tags. In Josep Domingo-Ferrer, Joachim Posegga, and Daniel Schreckling, editors, *International Conference on Smart Card Research and Advanced Applications – CARDIS*, volume 3928 of *Lecture Notes in Computer Science*, pages 289–299, Tarragona, Spain, April 2006. IFIP, Springer-Verlag.

[Can05] Anne Canteaut. Open problems related to algebraic attacks on stream ciphers. In *Proceedings of International Workshop on Coding and Cryptography (WCC 2005)*, 2005.

[Car95] Claude Carlet. Generalized partial spreads. *IEEE Transactions on Information Theory*, 41(5) :1482–1487, 1995.

[Car02] Claude Carlet. On cryptographic complexity of Boolean functions. In G.L. Mullen, H. Stichtenoth, and H. Tapia-Recillas, editors, *Sixth Conference on Finite Fields with Applications to Coding Theory, Cryptography and Related Areas*, pages 53–69. Springer, 2002.

[Car03] Claude Carlet. On the algebraic thickness and non-normality of boolean function. In *IEEE Information Theory Workshop*, pages 147–150, Paris, France, 2003.

[Car05] Claude Carlet. A lower bound on the higher order nonlinearity of algebraic immune functions. Cryptology ePrint Archive, Report 2005/469, 2005. `http://eprint.iacr.org/`.

[Car06a] Claude Carlet. *Boolean Functions for Cryptography and Error Correcting Codes*. In "Boolean methods and models" published by Cambridge University Press (Peter Hammer et Yves Crama editors), 2006.

[Car06b] Claude Carlet. On the higher order nonlinearities of algebraic immune functions. In Dwork [Dwo06], pages 584–601.

[Car06c] Claude Carlet. Recursive lower bounds on the nonlinearity profile of boolean functions and their applications. Cryptology ePrint Archive, Report 2006/459, 2006. `http://eprint.iacr.org/`.

[CCD00] Anne Canteaut, Pascale Charpin, and Hans Dobbertin. Weight divisibility of cyclic codes, highly nonlinear functions on $f_2$, and crosscorrelation of maximum-length sequences. *SIAM J. Discret. Math.*, 13(1) :105–138, 2000.

[CCW85] J. Constantin, Bertrand Courteau, and Jacques Wolfmann. Numerical experiments related to the covering radius of some first order Reed-Muller codes. In Jacques Calmet, editor, *AAECC*, volume 229 of *Lecture Notes in Computer Science*, pages 69–75. Springer, 1985.

[CF00] Anne Canteaut and Eric Filiol. Ciphertext only reconstruction of stream ciphers based on combination generators. In Bruce Schneier, editor, *FSE*, volume 1978 of *Lecture Notes in Computer Science*, pages 165–180. Springer, 2000.

[CF06] Hervé Chabanne and Guillaume Fumaroli. Noisy cryptographic protocols for low-cost RFID tags. *IEEE Transactions on Information Theory*, 52(8) :3562–3566, 2006.

[CG99] Claude Carlet and Philippe Guillot. A new representation of boolean functions. In *Proceedings of AAECC'13, LNCS 1719*, pages 94–103, 1999.

[CJ00] John A. Clark and Jeremy L. Jacob. Two stage optimisation in the design of boolean functions. In *5th Australian Conference on Security and Information Privacy (ACSIP)*, 2000.

[Coh93] Henri Cohen. *A course in computational algebraic number theory*, volume 138 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1993.

[CT00] Anne Canteaut and Michaël Trabbia. Improved fast correlation attacks using parity-check equations of weight 4 and 5. In *EUROCRYPT*, pages 573–588, 2000.

[CV94] Florent Chabaud and Serge Vaudenay. Links between differential and linear cryptanalysis. In Alfredo De Santis, editor, *Advances in Cryptology – EUROCRYPT'94*, volume 950 of *Lecture Notes in Computer Science*, pages 356–365. Springer-Verlag, 1994.

[DFM98] George I. Davida, Yair Frankel, and Brian J. Matt. On enabling secure applications through offline biometric identification. *In Proc. IEEE Symp. on Security and Privacy*, 1998.

[DH76] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on Information Society*, 22(6) :644–654, november 1976.

[Dil74] J.F. Dillon. Elementary hadamard difference sets. Ph. D. Thesis, University of Maryland, 1974.

[DKRS06] Yevgeniy Dodis, Jonathan Katz, Leonid Reyzin, and Adam Smith. Robust fuzzy extractors and authenticated key agreement from close secrets. In Dwork [Dwo06], pages 232–250.

[Dob94] Hans Dobbertin. Construction of bent functions and balanced boolean functions with high nonlinearity. In *Fast Software Encryption, LNCS*, volume 1008, pages 61–74. Springer-Verlag, 1994.

[Dob99a] Hans Dobbertin. Almost perfect nonlinear power functions on $GF(2^n)$ : The Niho case. *Inf. Comput.*, 151(1-2) :57–72, 1999.

[Dob99b] Hans Dobbertin. Almost perfect nonlinear power functions on $GF(2^n)$ : The Welch case. *IEEE Transactions on Information Theory*, 45(4) :1271–1275, 1999.

[Dob99c] Hans Dobbertin. Another proof of Kasami's theorem. *Des. Codes Cryptography*, 17(1-3) :177–180, 1999.

[Dob01] Hans Dobbertin. Almost perfectly nonlinear power functions on $GF(2^n)$ : a new case for $n$ divisible by 5. In *Proceedings of the Fifth Conference on Finite Fields and Applications Fq5*, pages 113–121, 2001.

[DRS04] Yevgeniy Dodis, Leonid Reyzin, and Adam Smith. Fuzzy extractors : How to generate strong keys from biometrics and other noisy data. In Christian Cachin and Jan Camenisch, editors, *EUROCRYPT*, volume 3027 of *Lecture Notes in Computer Science*, pages 523–540. Springer, 2004.

[DS05] Yevgeniy Dodis and Adam Smith. Correcting errors without leaking partial information. In *STOC '05 : Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 654–663, New York, NY, USA, 2005. ACM Press.

[Dwo06] Cynthia Dwork, editor. *Advances in Cryptology - CRYPTO 2006, 26th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2006, Proceedings*, volume 4117 of *Lecture Notes in Computer Science*. Springer, 2006.

[EPC] EPC. Electronic product code global inc. `http://www.epcglobalinc.org/`.

[Fel05] Patrick Felke. *A Systematic Approach with the Multi-Variate Method over Finite Fields of Odd Characteristic*. PhD thesis, Ruhr Universität Bochum, 2005.

[FJ01] Niklas Frykholm and Ari Juels. Error-tolerant password recovery. In *ACM Conference on Computer and Communications Security*, pages 1–9, 2001.

[FMI+06] Marc P. C. Fossorier, Miodrag J. Mihaljevic, Hideki Imai, Yang Cui, and Kanta Matsuura. An algorithm for solving the LPN problem and its application to security evaluation of the HB protocols for RFID authentication. In Rana Barua and Tanja Lange, editors, *INDOCRYPT*, volume 4329 of *Lecture Notes in Computer Science*, pages 48–62. Springer, 2006. See also Cryptology ePrint Archive, Report 2006/197, `http://eprint.iacr.org/`.

[Gal63] R. G. Gallager. *Low-Density Parity Check Codes*. Number 21 in Research monograph series. MIT Press, Cambridge, MA, 1963.

[Gau07] Carl Friedrich Gauss. *Disquisitiones Arithmeticae*, chapter Sectio Septima De Aquationibus Circuli Sectiones Definientibus. 1807.

[Gil95] Valérie Gillot. Bounds for exponential sums over finite fields. *Finite Fields and Their Applications*, 1 :421–436, 1995.

[GK79] Benedict H. Gross and Neal Koblitz. Gauss sums and the $\mathfrak{p}$-adic $\gamma$-function. *The Annals of Mathematics, 2nd Ser.*, 109(3) :569–581, May 1979.

[Gol68] R. Gold. Maximal recursive sequences with 3-valued crosscorrelation functions. *IEEE Trans. on Inform. Theory*, 14 :154–156, 1968.

[GRS05] Henri Gilbert, Matt Robshaw, and Hervé Sibert. An active attack against HB$^+$ – a provably secure lightweight authentication protocol. *IEE Electronic Letters*, 41 :1169–1170, 2005. See also Cryptology ePrint Archive, Report 2005/237, `http://eprint.iacr.org`.

[Hal73] G. Halász. On a result of Salem and Zygmund concerning random polynomials. *Studia Sci. Math. Hungar.*, 8 :369–377, 1973.

[HB01] Nicholas J. Hopper and Manuel Blum. Secure human identification protocols. In Colin Boyd, editor, *Advances in Cryptology - ASIACRYPT 2001*, volume 2248 of *Lecture Notes in Computer Science*, pages 52–66. Springer-Verlag, 2001.

[HKC+94] A. Roger Hammons Jr., P. Vijay Kumar, A. Robert Calderbank, Neil J. A. Sloane, and Patrick Solé. The $Z_4$-linearity of Kerdock, Preparata, Goethals, and related codes. *IEEE Transactions on Information Theory*, 40(2) :301–319, 1994.

[Hou90] Xiang Dong Hou. *Covering radius and error correcting codes*. PhD thesis, University of Illinois, Chicago, US, 1990.

[Hou96] Xiang Dong Hou. Covering radius of the Reed-Muller code (1, 7) - a simpler proof. *J. Comb. Theory, Ser. A*, 74(2) :337–341, 1996.

[Hås97] Johan Håstad. Some optimal inapproximability results. In *STOC 1997*, pages 1–10, 1997.

[HW79] G. H. Hardy and E. M. Wright. An introduction to the theory of numbers. Oxford Science Publications, Oxford, fifth edition, 1979.

[HX01] H. D. L. Hollman and Q. Xiang. A proof of the Welch and Niho conjectures on cross-correlations of binary m-sequences. In *Finite Fields Appl. 7*, pages 253–286, 2001.

[JJ99] Thomas Johansson and Fredrik Jönsson. Fast correlation attacks based on turbo code techniques. In Michael J. Wiener, editor, *CRYPTO*, volume 1666 of *Lecture Notes in Computer Science*, pages 181–197. Springer, 1999.

[JS06] Ari Juels and Madhu Sudan. A fuzzy vault scheme. *Des. Codes Cryptography*, 38(2) :237–257, 2006. Also in IEEE International Symp. on Information Theory 2002.

[Jue06] Ari Juels. RFID security and privacy : A research survey. To appear in the *IEEE Journal on Selected Areas in Communication*, 2006.

[JW99] Ari Juels and Martin Wattenberg. A fuzzy commitment scheme. In *ACM Conference on Computer and Communications Security*, pages 28–36, 1999.

[JW05] Ari Juels and Stephen Weis. Authenticating pervasive devices with human protocols. In Victor Shoup, editor, *Advances in Cryptology – CRYPTO'05*, volume 3126 of *Lecture Notes in Computer Science*, pages 293–308. Springer-Verlag, 2005.

[Kah85] J.-P. Kahane. *Some random series of functions*, volume 5 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge – New York, 2nd edition, 1985.

[Kas71] Tadao Kasami. The weight enumerators for several classes of subcodes of the 2nd order binary Reed-Muller codes. *Information and Control*, 18(4) :369–394, 1971.

[Ker72] A. M. Kerdock. A class of low-rate nonlinear binary codes. *Information and Control*, 20(2) :182–187, 1972.

[KIY04] Kaoru Kurosawa, Tetsu Iwata, and Takayuki Yoshiwara. New covering radius of Reed-Muller codes for t-resilient functions. *IEEE Transactions on Information Theory*, 20(3) :468–475, 2004.

[KMSY06] Selcuk Kavut, Subhamoy Maitra, Sumanta Sarkar, and Melek D. Yucel. Enumeration of 9-variable rotation symmetric boolean functions having nonlinearity > 240. Cryptology ePrint Archive, Report 2006/249, 2006. `http://eprint.iacr.org/`.

[KMY06] Selçuk Kavut, Subhamoy Maitra, and Melek D. Yücel. There exist boolean functions on $n$ (odd) variables having nonlinearity $> 2^{n-1} - 2^{\frac{n-1}{2}}$ if and only if $n > 7$. Cryptology ePrint Archive, Report 2006/181, 2006. `http://eprint.iacr.org/`.

[KR96] Lars R. Knudsen and Matthew J. B. Robshaw. Non-linear approximations in linear cryptoanalysis. In *EUROCRYPT*, pages 224–236, 1996.

[KS64] S. Kochen and C. Stone. A note on the Borel–Cantelli lemma. *Illinois J. Math*, 8 :248–251, 1964.

[KY07a] Selcuk Kavut and Melek Diker Yucel. Balanced 15–variable boolean functions with nonlinearity 16268. Cryptology ePrint Archive, Report 2007/321, 2007. `http://eprint.iacr.org/`.

[KY07b] Selcuk Kavut and Melek Diker Yucel. Generalized rotation symmetric and dihedral symmetric boolean functions – 9 variable boolean functions with nonlinearity 242. Cryptology ePrint Archive, Report 2007/308, 2007. `http://eprint.iacr.org/`.

[Lan90a] S. Lang. *Cyclotomic Fields I and II*, volume 121 of *Graduate Texts in Mathematics*. Springer-Verlag, 1990.

[Lan90b] Philippe Langevin. Covering radius of $RM(1,9)$ in $RM(3,9)$. In Gérard D. Cohen and Pascale Charpin, editors, *EUROCODE*, volume 514 of *Lecture Notes in Computer Science*, pages 51–59. Springer, 1990.

[Lan91] Philippe Langevin. On the orphans and covering radius of the Reed-Muller codes. In *9th International Symposium on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, AAECC 91*, pages 234–240, 1991.

[Lan96] Philippe Langevin. A new class of two weight codes. In *Finite Fields Conference Fq3*, pages 181–187, Glasgow, Scotland, 1996.

[Lan97] Philippe Langevin. Calcul de certaines sommes de gauss. *Journal of Number Theory*, 63(1) :59–64, march 1997.

[Lan99] Philippe Langevin. *Les sommes de caractères et la formule de Poisson dans la théorie des codes, séquences et fonctions booléennes.* Université de Toulon-Var, 1999. Mémoire d'habilitation à diriger des recherches.

[LeB02] Marc LeBrun. Sequence A067742 – Number of middle divisors of n, i.e. divisors in the half-open interval $[\sqrt{n/2}, \sqrt{2n})$. The On-Line Encyclopedia of Integer Sequences, `http://www.research.att.com/~njas/sequences/A067742`, 2002.

[LF06] Éric Levieil and Pierre-Alain Fouque. An improved LPN algorithm. In Roberto De Prisco and Moti Yung, editors, *SCN*, volume 4116 of *Lecture Notes in Computer Science*, pages 348–359. Springer, 2006.

[LN88] Rudolf Lidl and Harald Niederreiter. *Introduction to Finite Fields and their Applications*. Cambridge University Press, 1988.

[Loe77] M. Loeve. *Probability theory*. Springer, 1977.

[LRVZ06] Philippe Langevin, Patrice Rabizonni, Pascal Véron, and Jean-Pierre Zanotti. On the number of bent functions with 8 variables. In J-F. Michon, P. Valarcher, and J-B. Yunès, editors, *Boolean Functions : Cryptography and Applications*, BFCA'06, Rouen, 2006.

[LT03] Jean-Paul M. G. Linnartz and Pim Tuyls. New shielding functions to enhance privacy and prevent misuse of biometric templates. In Josef Kittler and Mark S. Nixon, editors, *AVBPA*, volume 2688 of *Lecture Notes in Computer Science*, pages 393–402. Springer, 2003.

[LVZ98] Philippe Langevin, Pascal Veron, and Jean-Pierre Zanotti. Fonctions booléennes équilibrées (II). Technical Report, GRIM–SCSSI, 1998.

[LW90] Gilles Lachaud and Jacques Wolfmann. The weights of the orthogonals of the extended quadratic binary goppa codes. *IEEE Transactions on Information Theory*, 36(3) :686–, 1990.

[LZ96] Philippe Langevin and Jean-Pierre Zanotti. Fonctions booléennes équilibrées (I). Technical Report, GRIM–SCSSI, 1996.

[LZ01] Philippe Langevin and Jean-Pierre Zanotti. A note on the counter–example of Patterson–Wiedemann. In *Proceedings of the Sixth International Conference on Finite Fields and Applications – Fq6*, pages 214–219, Oaxaca, Mexico, 2001.

[LZ05] Philippe Langevin and Jean-Pierre Zanotti. Nonlinearity of some invariant boolean functions. *Des. Codes Cryptography*, 36(2) :131–146, 2005.

[Mai91] James A. Maiorana. A classification of the cosets of the Reed–Muller code $R(1, 6)$. *Mathematics of Computation*, 57(195) :403–414, 1991.

[Mai01] Subhamoy Maitra. Highly nonlinear balanced boolean functions with very good autocorrelation property. In *Workshop on Coding and Cryptography, Electronic Notes in Discrete Math., Elsevier*, 2001.

[Mai07] Subhamoy Maitra. Balanced boolean function on 13-variables having nonlinearity strictly greater than the bent concatenation bound. Cryptology ePrint Archive, Report 2007/309, 2007. `http://eprint.iacr.org/`.

[Mat93] Mitsuru Matsui. Linear cryptoanalysis method for DES cipher. In *EUROCRYPT*, pages 386–397, 1993.

[MCD98] William Millan, Andrew Clark, and Ed Dawson. Heuristic design of cryptographically strong balanced boolean functions. In *Advances in Cryptology EUROCRYPT 98, Springer Verlag LNCS 1403*, pages 489–499, 1998.

[McE78] R.J. McEliece. A public-key cryptosystem based on algebraic coding theory. JPL DSN Progress Report, 1978.

[McF73] Robert L. McFarland. A family of difference sets in non-cyclic groups. *J. Comb. Theory, Ser. A*, 15(1) :1–10, 1973.

[MMJP03] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer-Verlag New York, 2003.

[MS77] F. MacWilliams and N. Sloane. The theory of error-correcting codes. North-Holland, 1977.

[MS89] Willi Meier and Othmar Staffelbach. Fast correlation attacks on certain stream ciphers. *J. Cryptology*, 1(3) :159–176, 1989.

[MS02] Subhamoy Maitra and Palash Sarkar. Modifications of patterson-wiedemann functions for cryptographic applications. *IEEE Transactions on Information Theory*, 48(1) :278–284, 2002.

[Mul54] D.E. Muller. Application of boolean algebra to switching circuit design and to error detection. *IEEE Trans. on Electronic Computers*, 3 :6–12, 1954.

[Myk80] J.J. Mykkelveit. The covering radius of the $(128, 8)$ Reed-Muller code is 56. *IEEE Transactions on Information Theory*, 26 :359–362, 1980.

[NAS72] NASA. Mariner 9 mission. `http://nssdc.gsfc.nasa.gov/imgcat/html/mission_page/MR_Mariner_9_page1.html`, 1972.

[Nat01] National Institute of Standards, U.S. Department of Commerce. *FIPS 197 : Advanced encryption standard*, November 2001.

[Nie86] H. Niederreiter. Knapsack-type cryptosystems and algebraic coding theory. *Problems Control Inform. Theory*, 15(2) :159–166, 1986.

[OS98] D. Olejár and M. Stanek. On cryptographic properties of random boolean functions. *Journal of Universal Computer Science*, 4(8) :705–717, 1998.

[Pre68] Franco P. Preparata. A class of optimum nonlinear double-error-correcting codes. *Information and Control*, 13(4) :378–400, 1968.

[PVV+90] Bart Preneel, Werner Van Leekwijck, Luc Van Linden, René Govaerts, and Joos Vandewalle. Propagation characteristics of boolean functions. In *EUROCRYPT*, pages 161–173, 1990.

[PW83] N. Patterson and D. Wiedemann. The covering radius of the $(2^{15}, 16)$ Reed - Muller code is at least 16276. *IEEE Transactions on Information Theory*, 29(3) :354–356, 1983.

[Ree54] I.S. Reed. A class of multiple-error-correcting codes and their decoding scheme. *IEEE Trans. on Information Theory*, 4 :38–42, 1954.

[Rev97] D. Revuz. *Probabilités*. Hermann, 1997.

[Rod92] François Rodier. On the spectra of the duals of binary bch codes of designed distance $\delta = 9$. *IEEE Transactions on Information Theory*, 38(2) :478–479, 1992.

[Rod93] François Rodier. Minorations de certaines sommes exponentielles (II). In R. Pellikaan, M. Perret, and S.G. Vladut, editors, *Arithmetic, Geometry and Coding Theory, AGCT-93*, pages 185–198, 1993.

[Rod03a] François Rodier. Asymptotic nonlinearity of Boolean functions. *preprint IML n. 2003-10*, 2003.

[Rod03b] François Rodier. On the nonlinearity of boolean functions. In D. Augot, P. Charpin, and G. Kabatianski, editors, *WCC2003, Workshop on coding and cryptography*, pages 397–405. INRIA, 2003.

[Rod04] François Rodier. Sur la non-linéarité des fonctions booléennes. *Acta Arithmetica*, 115 :1–22, 2004.

[Rod05] François Rodier. Asymptotic distribution for the nonlinearity of boolean functions. In J.-F. Michon, P. Valarcher, and J.-B. Yunes, editors, *First Workshop on Boolean Functions : Cryptography and Applications*, pages 49–64, 2005.

[Rod06] François Rodier. Asymptotic Nonlinearity of Boolean Functions. *Des. Codes Cryptography*, 40(1) :59–70, 2006.

[Rot76] O. S. Rothaus. On Bent Functions. *J. Comb. Theory, Ser. A*, 20(3) :300–305, 1976.

[RS60] I. S. Reed and G. Solomon. Polynomial codes over certain finite fields. *SIAM Journal on Applied Mathematics*, 8(2) :300–304, June 1960.

[RSA78] Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2) :120–126, 1978. See also report MIT/LCS/TM-82, 1977.

[Ser70] Jean-Pierre Serre. *Cours d'arithmétique*. PUF, 1970.

[Sha48] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27 :379–423, 623–, july, october 1948. `http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html`.

[Shi96] A. N. Shiryaev. *Probability*. Springer, New York, 1996.

[Sid71] V.M. Sidelnikov. On mutual correlation of sequences. *Soviet Math Dolk*, 12(1) :197–201, 1971.

[Sie84] Thomas Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory*, 30(5) :776–, 1984.

[SM00] Palash Sarkar and Subhamoy Maitra. Construction of nonlinear boolean functions with important cryptographic properties. In *EUROCRYPT*, pages 485–506, 2000.

[SM07] Sumanta Sarkar and Subhamoy Maitra. Idempotents in the neighbourhood of Patterson–Wiedemann functions having walsh spectra zeros. *WCC 2007, International Workshop on Coding and Cryptography*, à paraître, 2007.

[Ste94] Jacques Stern. A new identification scheme based on syndrome decoding. In *CRYPTO '93 : Proceedings of the 13th annual international cryptology conference on Advances in cryptology*, pages 13–21, New York, NY, USA, 1994. Springer-Verlag New York, Inc.

[SZ54] R. Salem and A. Zygmund. Some properties of trigonometric series whose terms have random signs. *Acta Mathematica*, 91 :245–301, 1954.

[SZZ93] Jennifer Seberry, Xian-Mo Zhang, and Yuliang Zheng. Nonlinearly balanced boolean functions and their propagation characteristics (extended abstract). In Douglas R. Stinson, editor, *CRYPTO*, volume 773 of *Lecture Notes in Computer Science*, pages 49–60. Springer, 1993.

[Vér96] Pascal Véron. Improved identification schemes based on error-correcting codes. *Appl. Algebra Eng. Commun. Comput.*, 8(1) :57–69, 1996.

[Wag05] David Wagner. Communication personnelle sur $HB^{++}$, décembre 2005.

[Wei] Eric Weisstein. Class number. From MathWorld–A Wolfram Web Resource. http ://mathworld.wolfram.com/ClassNumber.html.

[Wei06] Stephen A. Weis. New foundations for efficient authentication, commutative cryptography, and private disjointness testing. MIT Computer Science Ph.D. Thesis, 2006. `http://saweis.net/`.

[WT85] A. F. Webster and Stafford E. Tavares. On the design of S-boxes. In Hugh C. Williams, editor, *CRYPTO*, volume 218 of *Lecture Notes in Computer Science*, pages 523–534. Springer, 1985.

[Wyn75] A.D. Wyner. The wire–tap channel. *The Bell System Technical Journal*, 54(8) :1355–1387, 1975.

[XM88] Guo-Zhen Xiao and James L. Massey. A spectral characterization of correlation-immune combining functions. *IEEE Transactions on Information Theory*, 34(3) :569–, 1988.

[Zan01] Jean-Pierre Zanotti. Football pool problem. Séminaire GRIM – 13 novembre, 2001.

# Annexe A

# Quelques notions de théorie des codes

Le principe général d'un code correcteur d'erreurs est de remplacer un message de longueur $k$ sur un alphabet $\mathcal{A}$ par un mot de code de longueur $n > k$ permettant, grâce à une certaine redondance, de corriger en partie les erreurs apparaissant lors de la transmission du mot de code. Un tel mot de code est pris dans un sous-ensemble $C$, appelé code, de $\mathcal{A}^n$ ; $n$ est la longueur de $C$.

Lorsque $\mathcal{A}^n$ est muni d'une distance (typiquement la distance de Hamming, i.e. le nombre de coordonnées différentes entre 2 vecteurs), le minimum des distances des mots de $C$ pris 2 à 2 est appelé la distance minimale, notée $d_{min}$, du code $C$. Le code permet ainsi de corriger jusqu'à $t = \lfloor \frac{d_{min}-1}{2} \rfloor$ erreurs quelconques puisque si on considère la boule de Hamming centrée autour du message reçu et de rayon $t$ alors par définition de $d_{min}$, il ne peut y avoir qu'un seul mot de $C$ dans cette boule. Plus $d_{min}$ est grand, plus le code permet de corriger d'erreurs (du moins s'il dispose d'un algorithme de décodage exploitable).

Lorsque $\mathcal{A}^n$ est un espace vectoriel (en général on considère un corps fini $\mathcal{A} = \mathbb{F}_q$), et que $C$ en est un sous-espace vectoriel alors $C$ est un code linéaire ; si sa dimension sur $\mathcal{A}$ est $k$, on précise que $C$ est un code $[n, k, d_{min}]_{\mathcal{A}}$. Un tel code peut être décrit par une matrice génératrice $G$, matrice à coefficients dans $\mathcal{A}$ de taille $k \times n$ et de rang $k$ tel que $C = \{mG, m \in \mathcal{A}^k\}$, i.e. que les lignes de $G$ forment un base du sous-espace vectoriel $C$.

Soit $m \in \mathbb{N}^*$, $n = \#\mathcal{A}^m$ et $p_1, \ldots, p_n$ les éléments de $\mathcal{A}^m$, alors le code de Reed-Muller d'ordre $r$ en $m$ variables sur $\mathcal{A}$ est défini à une équivalence près par

$$RM_{\mathcal{A}}(r, m) = \{(f(p_1), \ldots, f(p_n)) \mid f \in \mathcal{A}[X_1, \ldots, X_m], \deg f \leq r\}.$$

Le cas qui nous intéresse au chapitre 1 est le cas $\mathcal{A} = \mathbb{F}_2$, on parle de code de Reed-Muller binaire, noté simplement $RM(r, m)$ ; c'est en fait la configuration dans laquelle ils ont été introduits par Reed [Ree54] et Muller [Mul54]. Un code de Reed-Muller binaire d'ordre $r$ est donc l'ensemble des vecteurs formés des valeurs prises par les fonctions booléennes à $m$ variables de degré au plus $r$. Le code $RM(r, m)$ est alors un code binaire $[2^m, \sum_{i=0}^{r} \binom{m}{i}, 2^{m-r}]_2$.

L'intérêt de ce type de code vient d'une propriété très particulière sur les poids de ses mots de codes : Soit $f \in \mathcal{BF}(m)$ de degré inférieur à $r$, alors $w_H(f) \geq 2^{m-r}$. Cependant, mis à part les cas où $r \leq 2$ (et $r \geq m-3$ par l'identité de MacWilliams [MS77]), on ne connaît pas exactement sa distribution de poids (pour $m$ quelconque).

Historiquement, le code $RM(1,5)$ a été utilisé par les sondes Mariner lancées par la NASA entre 1969 et 1973 pour assurer la transmission des premières photos (en niveau de gris) de Mars [NAS72]. Il contient 64 mots de longueur 32 et permet de corriger $\lfloor \frac{16-1}{2} \rfloor = 7$ erreurs.

La théorie des codes est un sujet de recherche très vaste et ses applications, notamment depuis l'explosion des communications filaires et radios, touchent de très nombreux domaines. Comme autres exemples de référence, on pense aux codes en blocs de Reed-Solomon [RS60] utilisés depuis plusieurs années par le disque compact (CD) pour le protéger des rayures ou encore aux Turbo-Codes introduits par [BGT93] qui sont utilisés dans de nombreux protocoles de communications développés récemment (e.g. l'UMTS). L'attrait pour ces derniers étant dû à la faculté de s'approcher au plus près des limites du théorème de Shannon [Sha48] (limite théorique du rendement maximum de transmission sur un canal bruité).

Quelques problématiques principales dans la recherche de bons codes sont : l'optimisation des paramètres du code, dont le rendement (la dimension divisée par la longueur) en fonction du taux de correction, i.e. connaître et atteindre les limites possibles ; la recherche d'algorithme de décodage efficace ; le décodage au-delà de la capacité de correction théorique...

Pour terminer sur un sujet plus proche des fonctions booléennes, le cas du code de Kerdock [Ker72] est intéressant à citer. Le code de Kerdock $\mathcal{K}_m$ est un code binaire non-linéaire de longueur $m$, de taille $2^{2m}$ et de distance minimale $2^{m-1} - 2^{m/2-1}$ (dans l'ensemble des codes de même longueur et de même taille, cette distance est optimale ; de plus, on sait qu'un code linéaire de mêmes paramètres ne peut exister). C'est en fait un sous-ensemble de $RM(2,m)$ formé d'une réunion de cosets $f + RM(1,m)$ de $RM(1,m)$ (dont le coset trivial) tel que pour tout $f, g \in \mathcal{K}_m$, $f+g$ est courbe. Actuellement tous les codes connus avec les mêmes paramètres sont équivalents au code de Kerdock et le problème d'en construire un qui ne soit pas inclus dans $RM(2,m)$ est un problème ouvert, dont la résolution irait certainement de pair avec une meilleure compréhension de la classe des fonctions courbes.

Les codes de Kerdock ont également intéressé pendant de nombreuses années la communauté pour un autre problème, aujourd'hui résolu. Ce sont des codes non-linéaires mais dont les propriétés sont très proches de codes linéaires. En effet, la distribution des distances entre mots de codes est la même que la distribution de ses poids ; et avec le code de Preparata [Pre68], leurs polynômes énumérateur de poids (i.e. le représentant de leur distribution de poids) vérifient une relation analogue à l'identité de MacWilliams (valable normalement pour un code linéaire et son dual). Cette particularité a trouvé son explication dans [HKC$^+$94], où les auteurs démontrent que ce sont effectivement des codes formellement duaux mais en tant qu'images, par l'application de Gray, de codes linéaires cycliques duaux, sur l'anneau $\mathbb{Z}_4 = \mathbb{Z}/4\mathbb{Z}$ muni de la distance de Lee. Cette découverte développa entre autre l'intérêt pour les codes cycliques définis sur des anneaux finis (je tiens [Bri01] à disposition pour une présentation de quelques résultats sur $\mathbb{Z}_4$).

# Annexe B

# Compléments sur les fonctions booléennes

## B.1 Digression sur les fonctions puissances

Dans la suite, on parle brièvement du cas des fonctions puissances, afin d'illustrer le fait que même dans les cas a priori les plus "simples", l'étude des propriétés des fonctions booléennes n'est pas a fortiori facile. D'autre part, nous mettons en évidence ici que les exposants des fonctions puissances *Almost Bent* connues peuvent être finalement classés en deux catégories de par le degré des fonctions associées.

Plusieurs problèmes sont liés à l'étude des fonctions trace d'un polynôme et en particulier d'un monôme (fonctions booléennes, corrélation de séquences, codes cycliques à 2 zéros...) et plus récemment la cryptographie : Ce type de fonctions est utilisé dans les schémas de chiffrement par blocs. Pour évaluer l'importance du choix des fonctions booléennes afin d'obtenir un système de chiffrement par blocs solide, Chabaud et Vaudenay [CV94] ont introduit, en 1994, deux paramètres importants pour mesurer la résistance aux attaques par cryptanalyse linéaire et par cryptanalyse différentielle.

Soit $m$ un entier et $L = \mathbb{F}_{2^m}$ un corps fini de caractéristique 2. Comme on s'intéresse à des fonctions traciques, on peut considérer les fonctions à valeurs dans $L$ du type, $f : L \to L$.

La résistance à la cryptanalyse linéaire est mesurée par $\lambda_f = \sup_{a\in L, b\in L^\times} \lambda_f(a,b)$, où

$$\lambda_f(a,b) = |\widehat{\mathrm{Tr}(bf)}(a)| = 2\left|\#\{x \in L, \mathrm{Tr}(ax + bf(x)) = 0\} - 2^{m-1}\right|.$$

Si $f$ est une permutation, $\lambda_f = \|\widehat{\mathrm{Tr}(f)}\|_\infty$.

La résistance à la cryptanalyse différentielle est mesurée par $\delta_f = \max_{a\neq 0, b\in L} \delta_f(a,b)$, où

$$\delta_f(a,b) = \#\{x \in L, f(x+a) + f(x) = b\}.$$

Plus $\delta_f$ et $\lambda_f$ sont faibles, plus ces attaques sont difficiles.

**Remarque B.1** *Ce sont des invariants sous l'action du groupe* GL($L$) *des permutations linéaires.*

**Proposition B.1** *On a*

$$\delta_f \geq 2.$$

*En cas d'égalité, on parle de fonction Almost Perfect Nonlinear (APN). D'après Sidelnikov [Sid71], on a d'autre part*

$$\lambda_f \geq 2^{(m+1)/2}.$$

*En cas d'égalité, le spectre de $f$ est composé d'exactement 3 valeurs : 0, $-2^{(m+1)/2}$ et $+2^{(m+1)/2}$. Dans ce cas la fonction $f$ est dite Almost Bent (AB).*

*Démonstration.* Résultat de Sidelnikov via $\sum_{a,b} \widehat{\mathrm{Tr}(bf)}(a)^4 = 2^{2m} \sum_{a,b} \delta_f(a,b)^2$. □

On en déduit que les fonctions AB n'existent que dans le cas impair. De plus les deux notions sont liées puisqu'on peut montrer que toute fonction AB est APN.

A priori le cas le plus facile à traiter est celui des fonctions puissances $x \mapsto \mathrm{Tr}(x^s)$, il a fait l'objet de nombreux travaux mais plusieurs questions restent à élucider. Il est établi que parmi les fonctions puissances, il existe des fonctions AB et des fonctions APN non AB, dans le cas impair. Et dans le cas pair il existe des fonctions APN et la meilleur amplitude spectrale connue est $\lambda = 2^{m/2+1}$.

De manière similaire à l'étude de la non-linéarité, on cherche à connaître la valeur minimale $\Lambda(m) = \min_{0 \leq s < 2^m} \lambda_{x \mapsto x^s}$ : on sait alors que

$$\Lambda(m) = 2^{(m+1)/2}$$

si $m$ est impair et

$$2^{(m+1)/2} < \Lambda(m) \leq 2^{m/2+1}$$

si $m$ est pair.

**Conjecture B.1 (Welch)** $\Lambda(m) = 2^{m/2+1}$ *si $m$ est pair.*

Le problème est donc de déterminer la valeur de $\Lambda(m)$ et de savoir quelles sont les puissances qui permettent d'atteindre ce minimum.

**Remarque B.2** *$\delta_f$ et $\lambda_f$ sont invariants sous l'action des permutations linéaires et par passage à l'inverse (dans le cas $f$ inversible). Entre autre, pour les fonctions puissances (en agissant sur la puissance $s$), ils sont stables sur les classes cyclotomiques et par passage à l'inverse (si $s$ inversible).*

$$\lambda_{x^s} = \lambda_{x^{2s}} = \lambda_{x^{s^{-1}}}$$

$$\delta_{x^s} = \delta_{x^{2s}} = \delta_{x^{s^{-1}}}$$

*On considère ainsi les exposants par classes d'équivalences.*

### B.1.1 Cas $m$ impair

**Proposition B.2** – *Si $x \mapsto x^s$ APN, alors $s$ est inversible modulo $2^m - 1$.*
– *$x^s$ est AB ssi $x^s$ est APN et $\frac{m+1}{2}$-divisible*[1].

Ce critère de divisibilité est utilisé pour démontrer le caractère AB, via le théorème de McEliece sur le poids des codes cycliques ou directement via un théorème de congruence de Stickelberger.

A équivalence près les fonctions puissances AB connues ont les exposants indiqués dans la table B.1. Nous avons quelques fonctions APN supplémentaires (cf. Table B.2).

| Type | $s$ | |
|---|---|---|
| Gold (quadratique) [Gol68] | $2^i + 1$ | $i \wedge m = 1, 1 \leq i \leq m/2$ |
| Kasami [Kas71] | $2^{2i} - 2^i + 1$ | $i \wedge m = 1, 1 \leq i \leq m/2$ |
| Welch | $2^{(m-1)/2} + 3$ | |
| Niho | $2^{2r} + 2^r - 1$ | si $t$ pair, $r = t/2$<br>si $t$ impair, $r = (3t + 1)/2$<br>avec $1 \leq r \leq m = 2t + 1$ |

Tab. B.1 – Fonctions puissances $x^s$ AB connues

| Type | $s$ | |
|---|---|---|
| Inverse | $s = 2^m - 2$ | |
| Dobbertin [Dob01] | $s = 2^{4i} + 2^{3i} + 2^{2i} + 2^i - 1$ | $m = 5i$ |

Tab. B.2 – Fonctions puissances $x^s$ APN (non AB) connues

Il est alors conjecturé qu'on les connaît tous. A noter que dans le cas pair, le même type de conjecture – modulo quelques différences dans les exposants – existe. On notera également que dans le cas pair, la fonction puissance inverse, qui n'est alors plus APN mais dont la non-linéarité est la meilleure connue (cf. [LW90]), est la fonction qui a été choisie pour construire les boites S de l'AES.

**Remarque B.3** *Jusqu'à il y a peu, les seules fonctions AB et APN connues étaient des fonctions puissances (à équivalence près) ; or ceci a changé récemment et des contre-exemples généraux sont en cours d'étude : le lecteur pourra par exemple consulter [BCL06] et les références s'y trouvant pour s'informer sur ces dernières évolutions.*

Revenons aux fonctions puissances AB : La preuve dans le cas de l'exposant de Gold se fait via la théorie des formes quadratiques. Pour l'exposant de Kasami, la première preuve provient de Kasami en 1978, et Dobbertin a fournit plus tard une preuve plus simple via la théorie

[1] I.e. dont les coefficients de Walsh sont divisibles par $2^{\frac{m+1}{2}}$.

des formes quadratiques [Dob99c]. Pour les exposants de Welch et Niho, le caractère APN a d'abord été démontré (voir par exemple [Dob99a] et [Dob99b]) puis la condition de divisibilité a été établie (pour Welch via le théorème de McEliece sur la divisibilité des poids d'un code par Canteaut *et al.* (cf. [CCD00]) ; pour Niho via un algorithme d'addition avec reste modulo $2^m - 1$ par Hollman-Xiang [HX01]). Dans les deux cas, la preuve est relativement technique comparée à celles pour les exposants de Gold et de Kasami.

Il est intéressant de regarder le degré de ces fonctions booléennes.

**Proposition B.3** *On a* $\deg \mathrm{Tr}(x^s) \leq w(s)$*, avec égalité si la fonction n'est pas constante.*

*Démonstration.* $x = \sum x_i b_i$, $s = \sum_1^w 2^{r_i}$, $\mathrm{Tr}(x^s) = \sum_{i_1,\dots,i_w} x_{i_1} \cdots x_{i_w} \mathrm{Tr}(b_{i_1}^{2^{r_1}} \cdots b_{i_w}^{2^{r_w}})$. □

Ainsi Gold est de degré au plus 2 (d'où la preuve du caractère AB via la théorie des formes quadratiques) et Welch est de degré au plus 3. Pour Kasami et Niho, les degrés sont plus élevés a priori (et pas nécessairement constant, dans le cas de Kasami). En fait, dans ces 2 cas, on peut se ramener à des fonctions qui ont des degrés qui demeurent petits. Revenons sur les preuves des exposants de Gold et Kasami pour mieux le comprendre :

**Exposant de Gold.** Soit $f : x \mapsto x^{2^i+1}$, avec $i \wedge m = 1$. Soit $a \in L$, $\hat{f}(a) = \sum(-1)^{Q_a(x)}$ où $Q_a(x) = \mathrm{Tr}(x^{2^i+1} + ax)$. $Q_a$ est une forme quadratique et alors $\hat{Q}_a(0) = \pm 2^{(m+k)/2}$ ou 0.

Considérons la forme symplectique associée, $\Phi(x, y) = \mathrm{Tr}(x^{2^i}y + xy^{2^i})$ alors $\mathrm{Ker}\, Q = \{x, x^{2^i} + x^{-2^i} = 0\} = \{0\} \cup \{x, x^{2^{2i}} = 1\} = \mathbb{F}_2$ car $2i \wedge m = 1$. D'où $k = 1$ et $|\hat{f}(a)| \in \{0, 2^{(m+1)/2}\}$.

Ainsi, l'exposant de Gold est AB.

**Exposant de Kasami.** Soit $r \wedge m = 1$. L'idée de la démonstration de Dobbertin est d'utiliser le fait que $(2^{2r} - 2^r + 1)(2^r + 1) = 2^{3r} + 1$, pour $2^r + 1$ premier avec $2^m - 1$. Alors $\hat{f}(a) = \sum_{x \in L} \mu(x^s + a) = \sum_{x \in L} \mu(x^{2^{3r}+1} + ax^{2^r+1})$ via le changement de variable $x \mapsto x^{2^r+1}$ ; i.e. qu'on se ramène à un problème de formes quadratiques. L'étude du noyau (plus technique que le cas de Gold) suffit ensuite à établir le résultat.

Finalement, dans les 2 cas, les exposants de Gold et de Kasami se ramènent aux formes quadratiques. De son côté, l'exposant de Welch correspond clairement à une cubique. Et il reste à traiter le cas de l'exposant de Niho.

**Exposant de Niho.** Tout d'abord remarquons que les conditions dans la table B.1 sont équivalentes à choisir $r$ tel que $4r \equiv -1 \mod m$, cela permet de traiter les deux cas simultanément. Soit $4r \equiv -1 \mod m$ et $s_N = 2^{2r} + 2^r - 1$ l'exposant de Niho associé. Alors une astuce similaire à celle de Dobbertin pour l'exposant de Kasami s'applique :

$$s_N(2^{2r+1} + 2^r + 1) \equiv 3.2^{3r} \tag{B.1}$$

On peut ainsi se ramener à l'étude d'une cubique (comme dans le cas de Welch) via $x \mapsto x^{2^{2r+1}+2^r+1}$ (c'est bien une permutation, il est facile de montrer que l'exposant est premier avec $2^m - 1$) puisqu'alors ce changement de variables implique $\hat{f}(a) = \sum \mu(x^3 + ax^{2^{2r+1}+2^r+1})$. (Pour $t$ égal à 3 ou 5, c'est même quadratique.)

Le fait que l'exposant de Welch soit de degré 3 et le point ci-dessus nous incite donc à poser la question suivante par analogie avec le degré 2.

**Problème B.1** *Peut-on développer et utiliser une théorie sur les formes cubiques permettant de trouver une nouvelle preuve (directe, sans passer par la divisibilité et le caractère APN) pour Welch et Niho, et plus généralement permettant d'explorer la non-linéarité (en particulier de chercher des fonctions AB) des formes cubiques associées à des fonctions puissances ?*

Il est également intéressant de souligner que l'équation (B.1), qui lie l'exposant de Niho à une cubique tout comme l'exposant de Welch, permet de plus de (re-)démontrer que ces exposants ne sont pas équivalents (via $s \mapsto 2s$ et $s \mapsto s^{-1}$). Pour cela, il suffit de montrer que $w(s_W) \notin \{w(s_N), w(s_N^{-1})\}$. Il est évident que le poids de Welch et le poids de Niho sont distincts. Il faut montrer que le poids de l'inverse de $s_N$ est différent de 3.

On a $s_N.\alpha \equiv 3.2^{3r}$ où $\alpha = 2^{2r+1} + 2^r + 1$. Comme $m$ est impair, 3 est inversible modulo $2^m - 1$ d'où $s_N.\alpha.2^{-3r}.3^{-1} \equiv 1$. Alors, $w(s_N^{-1}) = w(\alpha.3^{-1})$. Or $-\frac{2^m-2}{3}.3 \equiv 1$, donc $w(s_N^{-1}) = w(-\alpha\frac{2^{m-1}-1}{3})$. Il suffit de montrer que $w(s_W) = 3 \neq w(-\alpha\frac{2^{m-1}-1}{3})$.

*Idée* :

$$2^{m-1} - 1 = \sum_{0}^{m-2} 2^i = 3 + 32^2 + 32^4 + \cdots + 32^{m-3},$$

$$\frac{2^{m-1}-1}{3} = \sum_{i \text{ pair}, i=0}^{m-3} 2^i,$$

$$-\frac{2^{m-1}-1}{3} = 2^{m-1} + \sum_{i \text{ impair}, i=1}^{m-2} 2^i \mod 2^m - 1.$$

D'où après multiplications par $\alpha$ et quelques calculs, on obtient 4 cas possibles :

$$w(s_N^{-1}) - 1 = \begin{cases} 3t/4 & \text{si} \quad t \equiv 0(4) \\ (3t+3)/4 & \text{si} \quad t \equiv 3(4) \\ (3t+1)/4 & \text{si} \quad t \equiv 2(4) \\ (3t+2)/4 & \text{si} \quad t \equiv 1(4) \end{cases},$$

où $t = (m-1)/2$.

Si $t \geq 3$, alors $w(s_N^{-1}) > 3$, et les exposants de Welch et Niho ne sont pas équivalents.

## B.2 Attaques de fonctions booléennes sur du chiffrement à flot : quelques exemples

**Principe du chiffrement à flot.** Le Chiffrement de Vernam est inconditionnellement sûr : Soit un message clair $m$ de longueur $N$

$$m = (m_0, \ldots, m_{N-1}) \in \mathbb{F}_2^N,$$

et une suite binaire $s$ de même longueur

$$s = (s_0, \ldots, s_{N-1}) \in \mathbb{F}_2^N.$$

Le chiffré $c$ est obtenu en ajoutant bit à bit $m$ et $s$ :

$$c = m \oplus s = (m_0 + s_0, \ldots, m_{N-1} + s_{N-1}).$$

Cependant dans la pratique $s$ n'est pas entièrement aléatoire mais est obtenue à l'aide d'un générateur pseudo-aléatoire. Le but des attaques est alors soit de retrouver la clé $s$ afin de déchiffrer le message, soit de retrouver la structure du générateur pseudo-aléatoire dans le cas où celle-ci est inconnue.

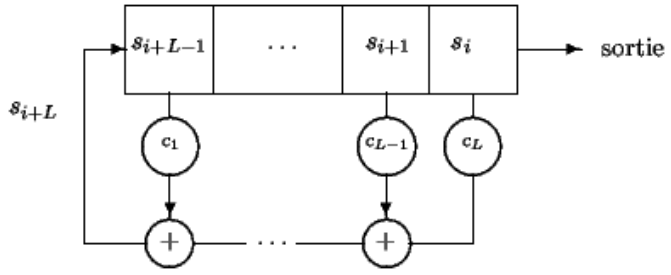


FIG. B.1 – Un LFSR

**Registre à décalage à rétroaction linéaire.** Un registre à décalage à rétroaction linéaire (ou LFSR – Linear Feedback Shift Register) est le générateur d'une séquence définie par une récurrence du type

$$\forall n \geq L, s_n = \sum_{i=1}^{L} c_i s_{n-i},$$

où le polynôme associé $P(X) = 1 + c_1 X + c_2 X^2 + \cdots + c_L X^L \in \mathbb{F}_2[X]$ est appelé polynôme de rétroaction du registre. On sait alors que si $c_L \neq 0$, la période de la séquence générée vérifie $T \leq 2^L - 1$. De plus, la série génératrice s'écrit sous la forme $S(X) = \sum_{n\geq 0} s_n X^n = Q(X)/P(X)$ avec $Q \in \mathbb{F}_2[X]$, $\deg Q < \deg P$ et il existe un unique polynôme minimal de rétroaction pour $(s_n)$ : $S(X) = Q_0/P_0$, où $\deg Q_0 < \deg P_0$ et $\operatorname{pgcd}(P_0, Q_0) = 1$. On appelle la complexité linéaire de la suite le degré de ce polynôme, $\Lambda(s) = \deg P_0$. Si $P_0$ est primitif, $T = 2^{\Lambda(s)} - 1$.

Ce type de constructions n'est pas suffisant puisqu'il existe une attaque à clairs connus très efficace : $2L$ bits consécutifs suffisent pour retrouver l'initialisation.

**Algorithme de Berlekamp - Massey.** L'algorithme permet de retrouver le polynôme minimal.

**Proposition B.4** *Soit $(s_n)_{n\geq 0}$ une suite binaire à récurrence linéaire. A partir de $2\Lambda(s)$ bits consécutifs de la suite, on sait déterminer le plus petit LFSR qui engendre la suite $(s_n)$.*

Ce qui impose de prendre un registre de complexité linéaire très élevée pour résister à cet algorithme. Une méthode alternative pour augmenter la complexité linéaire de la suite produite par un générateur est de prendre une combinaison de plusieurs registres via une fonction booléenne $f \in \mathcal{BF}(n)$. Dans ce cas, il y a $\prod_{i=1}^{n} 2^{L_i}$ initialisations possibles. Il est alors clair que $f$

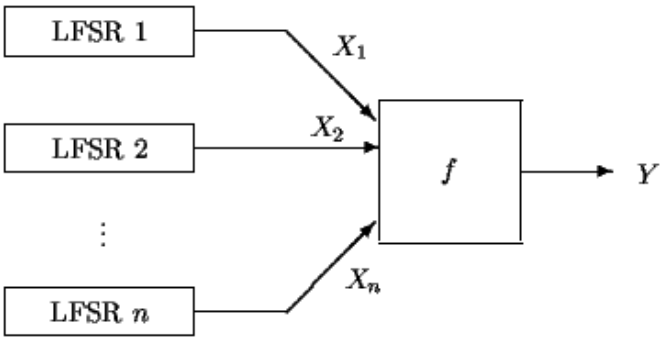


FIG. B.2 – Combinaison de LFSR par une fonction booléenne

doit être équilibrée ($\mathbb{P}[f(X_1, \ldots, X_n) = 1] = \mathbb{P}[f(X_1, \ldots, X_n) = 0] = 1/2$). De plus $f$ doit être de degré élevé puisque :

**Proposition B.5** *Si les polynômes de rétroaction sont primitifs, premiers entre eux deux à deux, de degrés respectifs $L_1, \ldots, L_n$, alors la complexité linéaire de la suite produite par la combinaison est $L = f(L_1, \ldots, L_n) \in \mathbb{Z}$.*

Sous ces conditions, l'algorithme de Berlekamp-Massey devient difficile à appliquer mais une autre attaque, par corrélation, est envisageable.

**Principes de l'attaque par corrélation.** On suppose connaître les paramètres du générateur aléatoire :

- $P_1, \ldots, P_n \in \mathbb{F}_2[X]$ de degré $L_1, \ldots, L_n$ les polynômes de rétroaction des $n$ LFSR.
- $f$ la fonction booléenne de combinaison.

Et soit $N$ le nombre de bits consécutifs connus de la suite chiffrée. L'idée est alors de retrouver l'initialisation du générateur via la détection d'un biais statistique.

**Attaque de Siegenthaler (1984) [Sie84].** C'est une attaque du type "Diviser pour régner" où l'initialisation d'un registre est déterminé indépendamment des autres. Siegenthaler a remarqué l'existence d'une corrélation entre la sortie de $f$ et ses entrées (i.e. les sorties des LFSR). On peut la voir comme une attaque à chiffré seul sous l'hypothèse que le clair est produit par une source binaire dont la sortie vaut 0 avec une probabilité $p_0$ connue.

**Proposition B.6** *Soit $f$ une fonction booléenne à $n$ variables, et $(X_i)_{i=1\ldots n}$ $n$ variables aléatoires indépendantes et uniformément distribuées (v.a.i.u.d) dans $\mathbb{F}_2$. Pour $1 \leq i \leq n$, on note*

$$q_i = \mathbb{P}[f(X_1, \ldots, X_n) = X_i].$$

*Soient $(c_j)_{1\ldots N}$ N bits du texte chiffré. Alors la corrélation $\alpha_i$ entre le chiffré c et la sortie $x^i$ du i-ième LFSR, définie par*

$$\alpha_i = \sum_{j=1}^{N} (1 - 2(c_j + x_j^i))$$

*est une variable aléatoire gaussienne de moyenne $N(2p_i - 1)$ et de variance $4Np_i(1 - p_i)$ où $p_i = 1 - p_0 - q_i + 2p_0q_i$.*

*De plus, la corrélation $\alpha_0$ entre le chiffré et une suite aléatoire x indépendante de $x^1, \ldots, x^n$ est une variable aléatoire de moyenne 0 et de variance N.*

Ainsi si $p_i \neq 1/2$, il suffit d'essayer toutes les initialisations du $i$-ième LFSR et d'observer la corrélation. Il faudra donc $\sum_{i=1}^{n} 2^{L_i}$ essais au maximum pour retrouver toutes les initialisations. Une généralisation en fixant plusieurs entrées à la fois est possible et pour y résister, $f$ doit alors être sans corrélation d'ordre élevé.

**Définition B.1** *Soit $f$ une fonction booléenne à n variables, $1 \leq t \leq n$. $f$ est dite sans corrélation d'ordre t si sa distribution de valeurs ne change pas lorsque qu'on fixe au plus t valeurs.*

**Remarque B.4** *Une autre généralisation de l'attaque est l'étude de $k > t$ LFSR combinés par une fonction booléenne g. On montre alors que la plus forte corrélation est obtenue pour une fonction affine. Ce qui implique que la non-linéarité de $f$ doit également être élevée.*

**Attaques par corrélation rapide.** On assimile ici la recherche de l'initialisation d'un LFSR fixé à un problème de décodage. C'est une attaque à clair connu (mais adaptable à du chiffré seul en fonction de la connaissance du canal) où les modifications induites par les autres LFSR et la fonction de combinaison sont modélisés suivant un canal binaire symétrique (BSC).

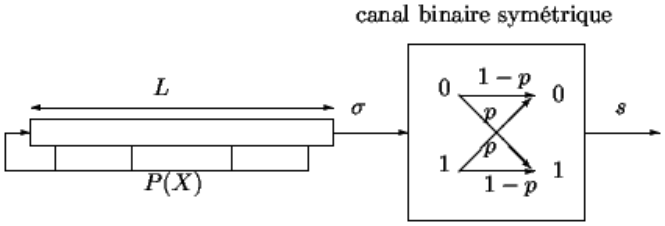


FIG. B.3 – Modélisation du problème par un BSC

Étant donnés $P$ le polynôme de rétroaction minimal, $L$ la longueur, et $\sigma$ la suite produite par le LFSR visé (ou par une combinaison), et $p$ une probabilité d'erreur choisie en fonction du générateur ($p < 1/2$, d'autant plus élevée que la corrélation est faible), on souhaite retrouver $L$ bits consécutifs de $\sigma$ à partir de $N$ bits de la suite chiffrante $s$.

On sait que $(\sigma_0, \ldots, \sigma_{N-1})$ appartient à un code $C$ linéaire, de longueur $N$ et de dimension $L$, défini par $P$, il faut donc trouver un algorithme de décodage efficace pour $C$, avec $N$ le moins grand possible. Le problème est alors de savoir quelles propriétés de $C$ sont exploitables. Parmi les techniques employées, on peut citer :

1. Meier et Staffelbach [MS89] : Un ensemble d'équations de contrôle pour $C$ est obtenu en considérant les polynômes $P^{2^i}$ de degré inférieur à $N$. Pour $P$ creux, et $p$ assez petit, ils utilisent alors la méthode de décodage itératif dû à Gallager pour le décodage de codes LDPC (cf. [Gal63]).
2. Johansson et Jönsson [JJ99] : $(s_n)_{n<N}$ est vu comme un mot perturbé d'un code convolutif ou d'un turbo-code (l'avantage est que $P$ peut être de poids quelconque, et $p$ plus élevé).
3. Canteaut et Trabbia [CT00] : Amélioration de [MS89] via la construction d'équations de contrôle de poids $d$ petit à partir des multiples de $P$ (pour $P$ de poids quelconque) avant d'appliquer l'algorithme de Gallager.

Pour 1 et 2, le nombre $N$ de bits nécessaires est exponentiel en $L/2$ ; pour 3, il est exponentiel en $L/(d-1)$.

**Principes de l'attaque [CT00].** Il y a deux phases, la construction des équations de poids faible en phase de pré-calcul puis la recherche de l'initialisation par une méthode itérative en phase de décodage.

Phase de pré-calcul :

Soit $d \geq 3$. On cherche toutes les équations faisant intervenir $d$ termes d'un mot $(x_0, \ldots, x_N)$ de $C$ sous la forme

$$x_n + \sum_{j=1}^{d-1} x_{i_j} = 0,$$

i.e., tous les polynômes $Q(X)P(X)$ de poids $d$ et de degré au plus $N$. Il suffit de se limiter aux polynômes de terme constant non nul :

$$1 + X^{i_1} + X^{i_2} + \cdots + X^{i_{d-1}}.$$

**Algorithme B.1** *1. Calculer tous les restes $q_i(X) \equiv X^i \mod P$ pour $1 \leq i \leq N$. Définir un tableau $T$ par*

$$\forall\, 0 \leq a < 2^L, T[a] = \{i, q_i(2) = a\}.$$

*2. Soit $\{i_1, \ldots, i_{d-2}\} \subset \{1, \ldots, N-1\}$,*
- *Calculer $A(X) = 1 + q_{i_1}(X) + \cdots + q_{i_{d-2}}(X)$,*
- *Pour $j \in T[A(2)]$, $1 + X^{i_1} + X^{i_2} + \cdots + X^{i_{d-2}} + X^j$ est un multiple de $P$ de poids $d$.*

Le nombre d'opérations est de l'ordre de $\binom{N-1}{d-2} \sim \frac{N^{d-2}}{(d-2)!}$ et le nombre d'équations générées est donné via le résultat suivant.

Si $P$ est un polynôme primitif sur $\mathbb{F}_2$, de degré $L$, alors le cardinal $m(d)$ de $E_d$, l'ensemble des polynômes $Q$ multiples de $P$ de degré inférieur à $N$, de poids $d$ et de terme constant égal à 1, est $m(d) \simeq \frac{N^{d-1}}{(d-1)!2^L}$. Par exemple, pour $L = 40$, $m(4) \simeq 9700$ pour $N = 4.10^5$ ou encore $m(5) \simeq 400$ pour $N = 10^4$.

Phase de décodage :

Via les équations de parité obtenues et l'algorithme de décodage (par maximum de vraisemblance), il est maintenant possible de retrouver le mot de code $(\sigma_n)_{n<N}$ à partir du mot reçu $(s_n)_{n<N}$.

**Définition B.2** *Soit X une v.a. binaire, la vraisemblance de X est définie par :*

$$L(X) = \log \frac{\mathbb{P}[X = 0]}{\mathbb{P}[X = 1]}$$

Le signe de $L(X)$ donne la valeur de $X$ la plus probable et sa valeur absolue correspond à la fiabilité de cette décision.

Au départ, tous les bits ont la même fiabilité assignée : $\forall\, i \in \{0, \ldots, N-1\}, L(\sigma_i) = (-1)^{s_i} \log \frac{1-p}{p}$. Soit $i \in \{0, \ldots, N-1\}$, on écrit les équations de $E_d$ devant être vérifiées par $\sigma_i$ sous la forme : $\sigma_i = t_1, \ldots, \sigma_i = t_j$. De sorte que la vraisemblance de $\sigma_i$ sachant que ces équations sont vraies est $L(\sigma_i) + L(t_1) + \cdots + L(t_j)$. Pour calculer $L(t_k)$, on utilise alors une formule d'approximation :

$$L(\sum X_l) = \Big(\prod \operatorname{sgn}(L(X_l))\Big) \min |L(X_l)|.$$

**Algorithme B.2**

1. *Initialisation : Pour* $i = 0 \ldots N-1$, $L(i) = \log \frac{1-p}{p}$.
2. *Jusqu'à "convergence", itérer :*
   - *Pour* $i = 0 \ldots N-1$,
     $L'(i) := (-1)^{s_i} L(i)$.
     *Pour chaque équation de contrôle "*$x_i = \sum_{j \in J_i} x_j$*", faire*
     $L'(i) := L'(i) + \Big(\prod_{j \in J_i} (-1)^{s_j}\Big) \min_{j \in J_i} |L(j)|$.
   - *Pour* $i = 0 \ldots N-1$,
     $s_i := \operatorname{sgn}(L'(i))$,
     $L(i) := |L'(i)|$.

A chaque itération, on débute avec une estimation de $\sigma$ et une mesure de la fiabilité de cette décision. A l'aide des équations de contrôle, on en déduit alors de nouvelles valeurs et de nouvelles fiabilités. La convergence de l'algorithme correspond à l'obtention d'une fiabilité jugée suffisamment grande, par exemple quand $(s_i)_{i<N}$ n'est plus modifiée par de futures itérations. Le nombre d'équations nécessaire à la convergence est dépendant du poids $d$ choisi. Par exemple, pour $p = 0.4$ et $L = 21$, il faudra $N = 16800$ pour $d = 3$, $N = 2200$ pour $d = 4$ et $N = 1100$ pour $d = 5$. Le comportement de l'algorithme est également variable en fonction de $p$.

L'attaque est finalement assez performante mais le pré-calcul devient très important pour $L$ grand et $p$ proche de $1/2$.

Pour finir, il est intéressant de signaler qu'il est également possible de mener une attaque lorsqu'on ne connaît pas les paramètres internes du générateur. En effet, dans [CF00] il est montré via une attaque à chiffré seul dans le cas où la structure est inconnue comment retrouver les polynômes de rétroaction des LFSR, puis la fonction de combinaison.

# Annexe C

# Démonstrations des résultats techniques du chapitre 3

**Démonstration de la proposition 3.2.** La transformée $U_j$ de Fourier de $u_j$ peut s'écrire, pour $t \in \mathbb{R}$

$$U_j(t) = \frac{1}{2\pi} \int_{\mathbb{R}} u_j(x) e^{-itx} dx = \delta(t) + v(t)$$

où $\delta$ est la distribution de Dirac en zéro et la fonction $v$ est définie par

$$v(t) = \frac{1}{2\pi} \int_{\mathbb{R}} (u(x) - 1) e^{-itx} dx.$$

Par construction de $u_j$, on a

$$\begin{aligned} |v(t)| &\leq \frac{1}{2\pi} \int_{|x| \leq M_j + \Delta_j} |u_j(x) - 1| dx \\ &\leq \frac{1}{\pi}(M_j + \Delta_j). \end{aligned}$$

Et par propriété de la transformée de Fourier de la dérivée, pour $p \geq 0$, on sait que

$$\frac{1}{2\pi} \int_{\mathbb{R}} u_j^{(p)}(x) e^{-ixt} dx = (it)^p U_j(t).$$

De plus comme $U_j = \delta + v$, on obtient $2\pi(it)^p v(t) = \int_{\mathbb{R}} u_j^{(p)}(x) e^{-ixt} dx + 2\pi(it)^p \delta(t)$. Pour $p \geq 1$, comme $(it)^p \delta(t) = 0$ quelque soit $t$, on obtient

$$\begin{aligned} |2\pi t^p v(t)| &\leq \int_{\mathbb{R}} |u_j^{(p)}(x)| dx \leq 2\Delta_j \max_{\mathbb{R}} |u_j^{(p)}(x)| \\ &\leq \frac{2 \max_{\mathbb{R}} |\alpha^{(p)}(x)|}{\Delta_j^{p-1}} = O(\frac{1}{\Delta_j^{p-1}}). \end{aligned}$$

Pour démontrer (3.3) et (3.4), nous écrivons l'intégrale en 2 parties en fonction de la valeur de $|t|$ par rapport à $1/\Delta_j$, puis on utilise les 2 majorations précédentes,

$$\begin{aligned}
\int_{\mathbb{R}} |U_j(t)dt| &\leq 1 + \int_{\mathbb{R}} |v(t)|dt \\
&\leq 1 + \Big| \int_{|t|\leq 1/\Delta_j} |v(t)|dt \Big| + \Big| \int_{|t|>1/\Delta_j} |v(t)|dt \Big| \\
&\leq 1 + \frac{2(M_j+\Delta_j)}{\pi\Delta_j} + \Big| \int_{|t|>1/\Delta_j} \frac{|t^2 v(t)|}{t^2} dt \Big| \\
&\leq 1 + \frac{2(M_j+\Delta_j)}{\pi\Delta_j} + \frac{\max_{\mathbb{R}} |\alpha^{(2)}(x)|}{\pi\Delta_j} \Big| \int_{|t|>1/\Delta_j} 1/t^2 dt \Big| \\
&\leq 1 + \frac{2(M_j+\Delta_j)}{\pi\Delta_j} + \frac{2\max_{\mathbb{R}} |\alpha^{(2)}(x)|}{\pi} \\
&= O(\frac{M_j}{\Delta_j}),
\end{aligned}$$

et pour $p > 0$,

$$\begin{aligned}
\int_{\mathbb{R}} |t^p U_j(t)dt| &\leq \Big| \int_{\mathbb{R}} t^p |v(t)|dt \Big| \\
&\leq \Big| \int_{|t|\leq 1/\Delta_j} t^p |v(t)|dt \Big| + \Big| \int_{|t|>1/\Delta_j} t^p |v(t)|dt \Big| \\
&\leq \frac{2\max_{\mathbb{R}} |\alpha^{(p)}(x)|}{\pi\Delta_j \times \Delta_j^{p-1}} + \Big| \int_{|t|>1/\Delta_j} t^{p+2} |v(t)|/t^2 dt \Big| \\
&\leq \frac{2\max_{\mathbb{R}} |\alpha^{(p)}(x)|}{\pi\Delta_j^p} + \frac{\max_{\mathbb{R}} |\alpha^{(p+2)}(x)|}{\pi\Delta_j^{p+2}} \times 2 \Big| \int_{1/\Delta_j}^{\infty} 1/t^2 dt \Big| \\
&= O(\frac{1}{\Delta_j^p}) + O(\frac{1}{\Delta_j^{p+1}}) = O(\frac{1}{\Delta_j^p}).
\end{aligned}$$

La dernière égalité étant vraie seulement si $\Delta_j \geq 1$ à partir d'un certain rang.

Pour l'estimation (3.5) avec $p \geq 1$, on remarque que la transformée de Fourier inverse de $t \mapsto t^p U_j(t)$ est $x \mapsto i^{-p} u^{(p)}(x)$ et que $t \mapsto e^{-ct^2/2}$ correspond à la transformée de Fourier de $x \mapsto \sqrt{2\pi/c} e^{-x^2/2c}$. On applique alors l'égalité de Parseval,

$$\begin{aligned}
\Big| \int_{\mathbb{R}} e^{-ct^2/2} \times t^p U_j(t)dt \Big| &= \sqrt{2\pi/c} \Big| \int_{\mathbb{R}} e^{-x^2/2c} u^{(p)}(x)dx \Big| \\
&= \sqrt{2\pi/c} \Big| \int_{M_j+\Delta_j>|x|>M_j} e^{-x^2/2c} u^{(p)}(x)dx \Big| \\
&= O(\frac{1}{\Delta_j^p \sqrt{c}} \int_{M_j+\Delta_j>|x|>M_j} e^{-x^2/2c} dx) = O(\frac{1}{\Delta_j^{p-1}\sqrt{c}} e^{-M_j^2/2c}).
\end{aligned}$$

De même, pour $p = 0$,

$$\begin{aligned}\Big|\int_{\mathbb{R}} e^{-ct^2/2}U_j(t)dt\Big| &= \sqrt{2\pi/c}\Big|\int_{\mathbb{R}} e^{-x^2/2c}u_j(x)dx\Big| = \sqrt{2\pi/c}\Big|\int_{|x|>M_j} e^{-x^2/2c}u_j(x)dx\Big| \\ &= O(\frac{1}{\sqrt{c}}\int_{|x|>M_j} e^{-x^2/2c}dx) = O(\frac{\sqrt{c}}{M_j}e^{-M_j^2/2c})\end{aligned}$$

où la dernière relation est obtenue via une majoration de la dernière intégrale par une intégrale plus facile à calculer :

$$\int_{|x|>M_j} e^{-x^2/2c}dx \le \int_{|x|>M_j} \frac{|x|}{M_j}e^{-x^2/2c}dx.$$

□

**Démonstration du lemme 3.5.** De manière similaire au lemme 3.2, on a

$$\varepsilon(e^{it\tilde{g}(\mu)}) = \varepsilon(\prod_{x\in G_j} e^{itg(x)\mu(x)}) = \prod_{x\in G_j}\varepsilon(e^{itg(x)\mu(x)})$$

puisque les variables aléatoires $g \mapsto e^{itg(x)\mu(x)}$ sont indépendantes pour $x$ parcourant $G_j$. D'après le lemme 3.4, on a alors

$$\varepsilon(e^{it\tilde{g}(\mu)}) = \prod_{x\in G_j}\cos t = (\cos t)^{N_j} = e^{N_j\ln(\cos t)}.$$

D'autre part,

$$e^{-a} = e^{-b} + (b-a)e^{-b} + O((b-a)^2) \tag{C.1}$$

pour $a, b \ge 0$ tels que $b - a \to 0$ et

$$\ln\cos t = -t^2/2 - t^4/12 + O(t^6) \tag{C.2}$$

pour $|t| \le 1$ tel que $t \to 0$. On en déduit

$$\begin{aligned}&e^{N_j\ln(\cos t)} \\ &= e^{-N_jt^2/2} - N_jt^4/12e^{-N_jt^2/2} + N_jO_{t\to 0,|t|\le 1}(t^6) + N_j^2O_{t\to 0,|t|\le 1}(t^8/144).\end{aligned}$$

□

**Démonstration du lemme 3.6.** L'espérance étant calculée sur $g$ variant dans $\Omega_j$, les intégrales peuvent être interverties

$$\begin{aligned}\varepsilon(\eta_j) &= \int_{\mathbb{R}}\int_{\hat{H}_j}\varepsilon(e^{it\tilde{g}(\mu)})U_j(t)d\mu dt \\ &= \int_{\mathbb{R}-[-1;1]}\int_{\hat{H}_j}\varepsilon(e^{it\tilde{g}(\mu)})U_j(t)d\mu dt + \int_{[-1;1]}\int_{\hat{H}_j}\varepsilon(e^{it\tilde{g}(\mu)})U_j(t)d\mu dt.\end{aligned}$$

Nous allons tout d'abord appliquer le résultat du lemme 3.5 sur la deuxième partie de cette expression. Comme la première partie de l'estimation (3.7) est indépendante du caractère $\mu$ et que $\int_{\hat{H}_j} d\mu = 1$ , on obtient

$$\begin{aligned}
&\int_{[-1;1]}\int_{\hat{H}_j} \varepsilon(e^{it\tilde{g}(\mu)})U_j(t)d\mu dt \\
&= \int_{[-1;1]}\int_{\hat{H}_j} e^{-N_jt^2/2}(1-\frac{N_jt^4}{12})U_j(t)d\mu dt \\
&\quad + \int_{[-1;1]}\int_{\hat{H}_j} (\varepsilon(e^{it\tilde{g}(\mu)}) - e^{-N_jt^2/2}(1-\frac{N_jt^4}{12}))U_j(t)d\mu dt \\
&= \int_{[-1;1]} e^{-N_jt^2/2}(1-\frac{N_jt^4}{12})U_j(t)dt \\
&\quad + \int_{[-1;1]}\int_{\hat{H}_j} (\varepsilon(e^{it\tilde{g}(\mu)}) - e^{-N_jt^2/2}(1-\frac{N_jt^4}{12}))U_j(t)d\mu dt \\
&= \int_{\mathbb{R}} e^{-N_jt^2/2}(1-\frac{N_jt^4}{12})U_j(t)dt + A_j,
\end{aligned}$$

où

$$\begin{aligned}
A_j &= \int_{[-1;1]}\int_{\hat{H}_j} (\varepsilon(e^{it\tilde{g}(\mu)}) - e^{-N_jt^2/2}(1-\frac{N_jt^4}{12}))U_j(t)d\mu dt \\
&\quad - \int_{\mathbb{R}-[-1;1]} e^{-N_jt^2/2}(1-\frac{N_jt^4}{12})U_j(t)dt.
\end{aligned}$$

D'après (3.7), il existe $\lambda \geq 0$ et $\kappa \in ]0,1[$ tels que pour $|t| \leq \kappa$,

$$|\varepsilon(e^{it\tilde{g}(\mu)}) - e^{-N_jt^2/2}(1-\frac{N_jt^4}{12})| \leq \lambda(N_jt^6 + N_j^2t^8).$$

Ainsi,

$$\begin{aligned}
&\Big|\int_{[-1;1]}\int_{\hat{H}_j} (\varepsilon(e^{it\tilde{g}(\mu)}) - e^{-N_jt^2/2}(1-\frac{N_jt^4}{12}))U_j(t)d\mu dt\Big| \\
&\leq \int_{[-\kappa,\kappa]} \lambda(N_jt^6 + N_j^2t^8)|U_j(t)|dt \\
&\quad + \Big|\int_{[-1;1]-[-\kappa,\kappa]} ((\cos t)^{N_j} - e^{-N_jt^2/2}(1-\frac{N_jt^4}{12}))U_j(t)dt\Big| \\
&\leq \int_{[-\kappa,\kappa]} \lambda(N_jt^6 + N_j^2t^8)|U_j(t)|dt + \int_{[-1;1]-[-\kappa,\kappa]} (2+\frac{N_jt^4}{12})|U_j(t)|dt \\
&\leq \int_{[-\kappa,\kappa]} \lambda(N_jt^6 + N_j^2t^8)|U_j(t)|dt + O_{j\to\infty}(\int_{[-1;1]-[-\kappa,\kappa]} N_jt^6|U_j(t)|dt) \quad (*) \\
&\leq O_{j\to\infty}(\int_{\mathbb{R}} (N_jt^6 + N_j^2t^8)|U_j(t)|dt),
\end{aligned}$$

l'inégalité $(*)$ étant vérifiée car $N_j \geq 1$ et il existe $C_\kappa, C_\kappa^{'} > 0$ tels que $t^4 \leq C_\kappa t^6$ et $2 \leq C_\kappa^{'} t^6$ pour $|t| \geq \kappa$. De même, pour $|t| \geq 1$, il existe $C_\kappa^{''}$ tel que $|1 - \frac{N_j t^4}{12}| \leq C_\kappa^{''} N_j t^6$, d'où

$$\Big| \int_{\mathbb{R}-[-1;1]} e^{-N_j t^2/2}(1 - \frac{N_j t^4}{12}) U_j(t) dt \Big| = O_{j\to\infty}(\int_{\mathbb{R}} N_j t^6 |U_j(t)| dt).$$

Ainsi on obtient $A_j = O_{j\to\infty}(\int_{\mathbb{R}} (N_j t^6 + N_j^2 t^8)|U_j(t)|dt)$.

D'autre part, on voit que cela implique que l'intégrale en dehors de $[-1, 1]$ dans le calcul de $\varepsilon(\eta_j)$ est dominé par ce terme d'erreur :

$$\begin{aligned}
\Big| \int_{\mathbb{R}-[-1;1]} \int_{\hat{H}_j} \varepsilon(e^{it\tilde{g}(\mu)}) U_j(t) d\mu dt \Big| &= \Big| \int_{\mathbb{R}-[-1;1]} \int_{\hat{H}_j} (\cos t)^{N_j} U_j(t) d\mu dt \Big| \\
&= \Big| \int_{\mathbb{R}-[-1;1]} (\cos t)^{N_j} U_j(t) dt \Big| \\
&\leq \int_{\mathbb{R}-[-1;1]} |U_j(t)| dt \\
&\leq O(A_j)
\end{aligned}$$

Finalement, cela donne

$$\varepsilon(\eta_j) = \int_{\mathbb{R}} e^{-N_j t^2/2}(1 - \frac{N_j t^4}{12}) U_j(t) dt + O_{j\to\infty}(\int_{\mathbb{R}} (N_j t^6 + N_j^2 t^8)|U_j(t)|dt). \tag{C.3}$$

D'après les estimations (3.4), (3.5) et (3.6), on en déduit alors que

$$\varepsilon(\eta_j) = O(\frac{\sqrt{N_j}}{M_j} e^{-M_j^2/2N_j}) + O(\frac{\sqrt{N_j}}{\Delta_j^3} e^{-M_j^2/2N_j}) + O(\frac{N_j}{\Delta_j^6}) + O(\frac{N_j^2}{\Delta_j^8}).$$

□

**Démonstration du lemme 3.7.** D'après (C.3),

$$\varepsilon(\eta_j) = \int_{\mathbb{R}} e^{-N_j t^2/2}(1 - \frac{N_j t^4}{12}) U_j(t) dt + O_{j\to\infty}(\int_{\mathbb{R}} (N_j t^6 + N_j^2 t^8)|U_j(t)|dt.$$

En appliquant les estimations (3.4) et (3.5), on obtient donc que

$$\varepsilon(\eta_j) = \int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t) dt + O(\frac{\sqrt{N_j}}{\Delta_j^3} e^{-M_j^2/2N_j}) + O(\frac{N_j}{\Delta_j^6}) + O(\frac{N_j^2}{\Delta_j^8}).$$

□

**Démonstration du lemme 3.8.** Soit $\nu \in \hat{H}_j$ et $a \in G_j$ tel que $\mu_a(x) = 1$ pour $x \in G_j - G_j \cap H_j$. Il est clair que $\nu_{|G_j} = \mu_a$ si et seulement si $\nu_{|G_j \cap H_j} = (\mu_a)_{|G_j \cap H_j}$ dans $\widehat{G_j \cap H_j}$. Ainsi $\sigma_j(\mu_a) =$ Card $(\hat{H}_j / \widehat{G_j \cap H_j})$, i.e. que c'est une constante sur $\hat{G}_j$. On a $\sigma_j =$ Card $H_j$/Card $G_j \cap H_j$. □

**Démonstration du lemme 3.9.** En effet, comme $\tilde{g}(\nu) = \tilde{g}(\mu)$ si $\mu_{|G_j} = \nu_{|G_j}$,

$$\begin{aligned}
\int\int_{\mu_{|G_j}=\nu_{|G_j}} u_j(\tilde{g}(\mu))u_j(\tilde{g}(\nu))d\mu d\nu &= \int\int_{\mu_{|G_j}=\nu_{|G_j}} u_j(\tilde{g}(\mu))^2 d\mu d\nu \\
&= \int_{\hat{H}_j} u_j(\tilde{g}(\mu))^2 (\int_{\nu_{|G_j}=\mu_{|G_j}} d\nu) d\mu \\
&= \sigma_j/h_j \int_{\hat{H}_j} u_j(\tilde{g}(\mu))^2 d\mu \\
&\leq \sigma_j \eta_j / h_j,
\end{aligned}$$

où la dernière inégalité résulte du fait que $u_j$ est à valeurs dans $[0, 1]$. □

**Démonstration du lemme 3.10.** Par indépendance des variables aléatoires sur $\Omega_j$ définies par $g \mapsto e^{ig(x)(t_1\mu(x)+t_2\nu(x))}$ pour $x$ parcourant $G_j$, on a

$$\varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}) = \prod_{x\in G_j} \varepsilon(e^{ig(x)(t_1\mu(x)+t_2\nu(x))}),$$

où

$$\begin{aligned}
\varepsilon(e^{ig(x)(t_1\mu(x)+t_2\nu(x))}) &= \int_{\Omega_j} e^{ig(x)(t_1\mu(x)+t_2\nu(x))} d\mathbb{P} \\
&= \frac{e^{i(t_1\mu(x)+t_2\nu(x))}}{2} + \frac{e^{-i(t_1\mu(x)+t_2\nu(x))}}{2}
\end{aligned}$$

puisque, à $x$ fixé, $g(x)$ prend la valeur 1 pour la moitié des $g \in \Omega_j$ et $-1$ pour l'autre moitié. Donc

$$\begin{aligned}
\varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}) &= \prod_{x\in G_j} \cos(t_1\mu(x) + t_2\nu(x)) \\
&= \exp(\sum_{x\in G_j} \ln\cos(t_1\mu(x) + t_2\nu(x))).
\end{aligned}$$

D'après les hypothèses, $|t_1\mu(x) + t_2\nu(x)| \leq 1$ quelque soit $x$ ; à l'aide des développements

limités (C.1) et (C.2), on trouve alors l'estimation suivante :

$$\begin{aligned}
&\varepsilon\left(\exp\left(it_1\tilde{g}\left(\mu\right)+it_2\tilde{g}\left(\nu\right)\right)\right)\\
&= \exp\left(\sum_{x\in G_j}\left(-\frac{(t_1\mu(x)+t_2\nu(x))^2}{2}-\frac{(t_1\mu(x)+t_2\nu(x))^4}{12}+O\left((t_1\mu(x)+t_2\nu(x))^6\right)\right)\right)\\
&= \exp\left(-\sum_{x\in G_j}(t_1\mu(x)+t_2\nu(x))^2/2\right)+\sum_{x\in G_j}-\frac{((t_1\mu(x)+t_2\nu(x))^4}{12}\\
&\quad +O\left((t_1\mu(x)+t_2\nu(x))^6\right)\times\exp\left(-\sum_{x\in G_j}(t_1\mu(x)+t_2\nu(x))^2/2\right)\\
&\quad +O\left(\left(\sum_{x\in G_j}-\frac{(t_1\mu(x)+t_2\nu(x))^4}{12}+O\left((t_1\mu(x)+t_2\nu(x))^6\right)\right)^2\right)\\
&= \exp\left(-\sum_{x\in G_j}(t_1\mu(x)+t_2\nu(x))^2/2\right)\\
&\quad -\sum_{x\in G_j}\frac{(t_1\mu(x)+t_2\nu(x))^4}{12}\exp\left(-\sum_{x\in G_j}(t_1\mu(x)+t_2\nu(x))^2/2\right)\\
&\quad +N_jO((|t_1|+|t_2|)^6)+N_j^2O((|t_1|+|t_2|)^8).
\end{aligned}$$

Via les relations d'orthogonalité des caractères (cf. Proposition 1.1), on calcule aisément

$$\begin{aligned}
&\sum_{x\in G_j}(t_1\mu(x)+t_2\nu(x))^2\\
&= \left(\sum_{x\in G_j}t_1^2\mu(x)^2+t_2^2\sum_{x\in G_j}\nu(x)^2+2t_1t_2\sum_{x\in G_j}\mu(x)\nu(x)\right)\\
&= N_j(t_1^2+t_2^2),
\end{aligned}$$

où $\sum_{x\in G_j}\mu(x)\nu(x)=0$ car $\mu_{|G_j}\neq\nu_{|G_j}$ dans $\hat{G}_j$. De même, on a

$$\sum_{x\in G_j}(t_1\mu(x)+t_2\nu(x))^4=N_j(t_1^4+t_2^4+6t_1^2t_2^2).$$

Ainsi,

$$\begin{aligned}
\varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}) = e^{-N_j(t_1^2+t_2^2)/2}-\frac{1}{12}N_j(t_1^4+t_2^4+6t_1^2t_2^2)e^{-N_j(t_1^2+t_2^2)/2}\\
+N_jO((|t_1|+|t_2|)^6)+N_j^2O((|t_1|+|t_2|)^8).
\end{aligned}$$

□

**Démonstration du lemme 3.11.** Pour commencer, séparons l'intégrale en 2 parties,

$$\begin{aligned}
&\varepsilon\Big( \int_{\mathbb{R}^2} e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)} U_j(t_1)U_j(t_2)dt_1dt_2\Big)\\
&= \int_{\mathbb{R}^2} \varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)})U_j(t_1)U_j(t_2)dt_1dt_2\\
&= \int_{[-1/2,1/2]^2} \varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)})U_j(t_1)U_j(t_2)dt_1dt_2\\
&\quad + \int_{\mathbb{R}^2-[-1/2,1/2]^2} \varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)})U_j(t_1)U_j(t_2)dt_1dt_2.
\end{aligned}$$

On souhaite appliquer le lemme 3.10, on écrit donc

$$\begin{aligned}
&\int_{[-1/2,1/2]^2} \varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)})U_j(t_1)U_j(t_2)dt_1dt_2\\
&= \Big( \int_{[-1/2,1/2]^2} e^{-N_j(t_1^2+t_2^2)/2} U_j(t_1)U_j(t_2)dt_1dt_2\\
&\quad - \int_{[-1/2,1/2]^2} \frac{1}{12}N_j(t_1^4+t_2^4+6t_1^2t_2^2)e^{-N_j(t_1^2+t_2^2)/2}U_j(t_1)U_j(t_2)dt_1dt_2\Big)\\
&\quad + \Big( \int_{[-1/2,1/2]^2} \varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)})U_j(t_1)U_j(t_2)dt_1dt_2\\
&\quad - \int_{[-1/2,1/2]^2} e^{-N_j(t_1^2+t_2^2)/2}U_j(t_1)U_j(t_2)dt_1dt_2\\
&\quad + \int_{[-1/2,1/2]^2} \frac{1}{12}N_j(t_1^4+t_2^4+6t_1^2t_2^2)e^{-N_j(t_1^2+t_2^2)/2}U_j(t_1)U_j(t_2)dt_1dt_2\Big)\\
&= A_j + \Big( \int_{\mathbb{R}^2} e^{-N_j(t_1^2+t_2^2)/2} U_j(t_1)U_j(t_2)dt_1dt_2\\
&\quad - \int_{\mathbb{R}^2} \frac{1}{12}N_j(t_1^4+t_2^4+6t_1^2t_2^2)e^{-N_j(t_1^2+t_2^2)/2}U_j(t_1)U_j(t_2)dt_1dt_2\Big),
\end{aligned}$$

où

$$\begin{aligned}
&A_j\\
&= \int_{[-1/2,1/2]^2} \varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)})U_j(t_1)U_j(t_2)dt_1dt_2\\
&\quad - \int_{[-1/2,1/2]^2} e^{-N_j(t_1^2+t_2^2)/2}(1-\frac{1}{12}N_j(t_1^4+t_2^4+6t_1^2t_2^2))U_j(t_1)U_j(t_2)dt_1dt_2\\
&\quad - \int_{\mathbb{R}^2-[-1/2,1/2]^2} e^{-N_j(t_1^2+t_2^2)/2}U_j(t_1)U_j(t_2)dt_1dt_2\\
&\quad + \int_{\mathbb{R}^2-[-1/2,1/2]^2} \frac{1}{12}N_j(t_1^4+t_2^4+6t_1^2t_2^2)e^{-N_j(t_1^2+t_2^2)/2}U_j(t_1)U_j(t_2)dt_1dt_2.
\end{aligned}$$

Or d'après le lemme 3.10, il existe $\lambda \geq 0$ et $\kappa \in ]0, 1/2[$ tels que pour $|t_1|, |t_2| \leq \kappa$, alors

$$\begin{aligned} & |\varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}) - e^{-N_j(t_1^2+t_2^2)/2}(1 - \frac{1}{12}N_j(t_1^4 + t_2^4 + 6t_1^2t_2^2))| \\ \leq \quad & \lambda(N_j(|t_1| + |t_2|)^6 + N_j^2(|t_1| + |t_2|)^8). \end{aligned}$$

D'où

$$\begin{aligned} & \Big| \int_{[-1/2,1/2]^2} \varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)})U_j(t_1)U_j(t_2)dt_1dt_2 \\ & \quad - \int_{[-1/2,1/2]^2} e^{-N_j(t_1^2+t_2^2)/2}(1 - \frac{1}{12}N_j(t_1^4 + t_2^4 + 6t_1^2t_2^2))U_j(t_1)U_j(t_2)dt_1dt_2\Big| \\ \leq \quad & \int_{[-\kappa,\kappa]^2} \lambda(N_j(|t_1| + |t_2|)^6 + N_j^2(|t_1| + |t_2|)^8)|U_j(t_1)||U_j(t_2)|dt_1dt_2 \\ & \quad + \Big| \int_{[-1/2,1/2]^2-[-\kappa,\kappa]^2} \prod_{x\in G_j} cos(t_1\mu(x) + t_2\nu(x))U_j(t_1)U_j(t_2)dt_1dt_2 \\ & \quad - \int_{[-1/2,1/2]^2-[-\kappa,\kappa]^2} e^{-N_j(t_1^2+t_2^2)/2}(1 - \frac{N_j(t_1^4 + t_2^4 + 6t_1^2t_2^2)}{12})U_j(t_1)U_j(t_2)dt_1dt_2\Big| \\ \leq \quad & \int_{[-\kappa,\kappa]^2} \lambda(N_j(|t_1| + |t_2|)^6 + N_j^2(|t_1| + |t_2|)^8)|U_j(t_1)||U_j(t_2)|dt_1dt_2 \\ & \quad + \int_{[-1/2,1/2]^2-[-\kappa,\kappa]^2} 2 + \frac{N_j}{12}(t_1^4 + t_2^4 + 6t_1^2t_2^2)|U_j(t_1)||U_j(t_2)|dt_1dt_2 \\ \leq \quad & \int_{[-\kappa,\kappa]^2} \lambda(N_j(|t_1| + |t_2|)^6 + N_j^2(|t_1| + |t_2|)^8)|U_j(t_1)||U_j(t_2)|dt_1dt_2 \\ & \quad + O_{j\to\infty}(\int_{[-1/2,1/2]^2-[-\kappa,\kappa]^2} N_j(|t_1| + |t_2|)^6|U_j(t_1)||U_j(t_2)|dt_1dt_2) \qquad \text{(C.4)} \\ \leq \quad & O_{j\to\infty}(\int_{\mathbb{R}^2} (N_j(|t_1| + |t_2|)^6 + N_j^2(|t_1| + |t_2|)^8)|U_j(t_1)||U_j(t_2)|dt_1dt_2). \end{aligned}$$

L'inégalité (C.4) étant vérifiée puisque $N_j \geq 1$ et considérant l'ensemble $[-1/2, 1/2]^2 - [-\kappa, \kappa]^2$, il existe $C_\kappa, C_\kappa^{'} > 0$ tels que $2 \leq C_\kappa^{'}(|t_1| + |t_2|)^6$ et $t_1^4 + t_2^4 + 6t_1^2t_2^2 \leq C_\kappa(|t_1| + |t_2|)^6$. De même en dehors de $[-1/2, 1/2]^2$, il existe $C_\kappa^{''} > 0$ tel que $|(1 - \frac{1}{12}N_j(t_1^4 + t_2^4 + 6t_1^2t_2^2))| \leq C_\kappa^{''}N_j(|t_1| + |t_2|)^6$, donc

$$\begin{aligned} & \int_{\mathbb{R}^2-[-1/2,1/2]^2} e^{-N_j(t_1^2+t_2^2)/2}U_j(t_1)U_j(t_2)dt_1dt_2 \\ & \quad - \int_{\mathbb{R}^2-[-1/2,1/2]^2} \frac{1}{12}N_j(t_1^4 + t_2^4 + 6t_1^2t_2^2)e^{-N_j(t_1^2+t_2^2)/2}U_j(t_1)U_j(t_2)dt_1dt_2 \\ = \quad & O_{j\to\infty}(\int_{\mathbb{R}^2-[-1/2,1/2]^2} N_j(|t_1| + |t_2|)^6|U_j(t_1)||U_j(t_2)|dt_1dt_2), \end{aligned}$$

ce qui implique finalement que

$$A_j = O_{j\to\infty}(\int_{\mathbb{R}^2} (N_j(|t_1| + |t_2|)^6 + N_j^2(|t_1| + |t_2|)^8)|U_j(t_1)||U_j(t_2)|dt_1dt_2).$$

Par les même arguments,

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^2-[-1/2,1/2]^2} \varepsilon(e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)})U_j(t_1)U_j(t_2)dt_1dt_2\Big| \\
&= \Big|\int_{\mathbb{R}^2-[-1/2,1/2]^2} \prod_{x\in G_j}\cos(t_1\mu(x)+t_2\nu(x))U_j(t_1)U_j(t_2)dt_1dt_2\Big| \\
&\leq \int_{\mathbb{R}^2-[-1/2,1/2]^2} |U_j(t_1)||U_j(t_2)|dt_1dt_2 \\
&= O_{j\to\infty}(\int_{\mathbb{R}^2}(N_j(|t_1|+|t_2|)^6|U_j(t_1)||U_j(t_2)|dt_1dt_2).
\end{aligned}$$

Ce qui revient à négliger les intégrales hors du pavé $[-1/2,1/2]^2$ par rapport aux termes d'erreurs. Finalement, on a

$$\begin{aligned}
&\varepsilon(\int_{\mathbb{R}^2} e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}U_j(t_1)U_j(t_2)dt_1dt_2) \\
&= \int_{\mathbb{R}^2} e^{-N_j(t_1^2+t_2^2)/2}(1-\frac{1}{12}N_j(t_1^4+t_2^4+6t_1^2t_2^2))U_j(t_1)U_j(t_2)dt_1dt_2 \\
&\quad +O_{j\to\infty}(\int_{\mathbb{R}^2}(N_j(|t_1|+|t_2|)^6+N_j^2(|t_1|+|t_2|)^8)|U_j(t_1)||U_j(t_2)|dt_1dt_2).
\end{aligned}$$

Via (3.4), (3.5) et (3.6), on en déduit enfin que

$$\begin{aligned}
&\varepsilon(\int_{\mathbb{R}^2} e^{it_1\tilde{g}(\mu)+it_2\tilde{g}(\nu)}U_j(t_1)U_j(t_2)dt_1dt_2) \\
&= \left(\int_{\mathbb{R}} e^{-N_jt^2/2}U_j(t)dt\right)^2 \\
&\quad -\frac{N_j}{12}\left(2\int_{\mathbb{R}} t_1^4e^{-N_jt_1^2/2}U_j(t_1)dt_1\int_{\mathbb{R}} e^{-N_jt_2^2/2}U_j(t_2)dt_2\right) \\
&\quad -\frac{N_j}{12}\left(6\left(\int_{\mathbb{R}} t^2e^{-N_jt^2/2}U_j(t)dt\right)^2\right) \\
&\quad +O_{j\to\infty}(\int_{\mathbb{R}^2}(N_j(|t_1|+|t_2|)^6+N_j^2(|t_1|+|t_2|)^8)|U_j(t_1)||U_j(t_2)|dt_1dt_2) \\
&= \left(\int_{\mathbb{R}} e^{-N_jt^2/2}U_j(t)dt\right)^2+O((\frac{1}{\Delta_j^2}+\frac{N_j}{\Delta_j^3M_j})e^{-M_j^2/N_j}) \\
&\quad +O_{j\to\infty}(\int_{\mathbb{R}^2}(N_j(|t_1|+|t_2|)^6+N_j^2(|t_1|+|t_2|)^8)|U_j(t_1)||U_j(t_2)|dt_1dt_2) \\
&= \left(\int_{\mathbb{R}} e^{-N_jt^2/2}U_j(t)dt\right)^2+O((\frac{1}{\Delta_j^2}+\frac{N_j}{\Delta_j^3M_j})e^{-M_j^2/N_j})+O(\frac{N_j}{\Delta_j^6})+O(\frac{N_j^2}{\Delta_j^8}).
\end{aligned}$$

□

**Démonstration de la proposition 3.7.** D'après les propositions 3.5 et 3.6 et le lemme 3.7, on a

$$\begin{aligned}
&\mathbb{P}(\eta_j = 0)\\
&\leq \frac{\varepsilon(\eta_j^2) - \varepsilon^2(\eta_j)}{\varepsilon^2(\eta_j)}\\
&\leq \frac{1}{\varepsilon^2(\eta_j)}\Big(\frac{\sigma_j}{h_j}\varepsilon(\eta_j) + \left(\int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t)dt\right)^2 + O((\frac{1}{\Delta_j^2} + \frac{N_j}{\Delta_j^3 M_j})e^{-M_j^2/N_j})\\
&\quad + O(\frac{N_j}{\Delta_j^6}) + O(\frac{N_j^2}{\Delta_j^8})\\
&\quad - (\int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t)dt + O(\frac{\sqrt{N_j}}{\Delta_j^3} e^{-M_j^2/2N_j}) + O(\frac{N_j}{\Delta_j^6}) + O(\frac{N_j^2}{\Delta_j^8}))^2\Big).
\end{aligned}$$

En développant cette expression et en utilisant (3.6), on obtient

$$\begin{aligned}
&\mathbb{P}(\eta_j = 0)\\
&\leq \frac{1}{\varepsilon^2(\eta_j)}\Big(\frac{\sigma_j}{h_j}\varepsilon(\eta_j) + O((\frac{1}{\Delta_j^2} + \frac{N_j}{\Delta_j^3 M_j})e^{-M_j^2/N_j}) + \frac{N_j}{\Delta_j^6}(1 + O(\frac{N_j}{\Delta_j^2}))\\
&\quad + \frac{\sqrt{N_j} N_j e^{-M_j^2/2N_j}}{\Delta_j^6}(O(\frac{1}{\Delta_j^3}) + O(\frac{N_j}{\Delta_j^5}) + O(\frac{N_j}{\Delta_j^2 M_j}) + O(\frac{1}{M_j}))\\
&\quad + \frac{N_j^2}{\Delta_j^{12}}(O(1) + O(\frac{N_j}{\Delta_j^2}) + O(\frac{N_j^2}{\Delta_j^4}))\Big).
\end{aligned}$$

Toujours à partir du lemme 3.7, on cherche maintenant à obtenir une minoration de $\varepsilon(\eta_j)$. On sait que

$$\varepsilon(\eta_j) = \int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t)dt + O(\frac{\sqrt{N_j}}{\Delta_j^3} e^{-M_j^2/2N_j}) + O(\frac{N_j}{\Delta_j^6}) + O(\frac{N_j^2}{\Delta_j^8})$$

donc il existe $j_0 \geq 0$ et $A > 0$ tel que

$$|\varepsilon(\eta_j) - \int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t)dt| \leq A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)$$

pour $j \geq j_0$. Ce qui implique

$$\varepsilon(\eta_j) \geq \int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t)dt - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right). \tag{C.5}$$

On applique alors la version bilinéaire de l'identité de Parseval

$$\int f\hat{g} = \int \hat{f} g$$

à $\int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t)dt = \int_{\mathbb{R}} f_j(t)U_j(t)dt$, où $f_j : t \mapsto e^{-N_j t^2/2}$ est une gaussienne de transformée de Fourier $\hat{f}_j : x \mapsto \sqrt{2\pi/N_j} e^{-x^2/2N_j}$ et $U_j$ est la transformée de Fourier de $u_j$. Ainsi, en utilisant le fait que $u_j$ est à valeurs positives et vaut 1 entre $M_j + \Delta_j$ et $M_j + 2\Delta_j$,

$$\begin{aligned}\int_{\mathbb{R}} e^{-N_j t^2/2} U_j(t)dt &= \sqrt{2\pi/N_j} \int_{\mathbb{R}} e^{-x^2/2N_j} u_j(x)dx \\ &\geq \sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j}.\end{aligned}$$

D'où, pour $j \geq j_0$,

$$\varepsilon(\eta_j) \geq \sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)$$

et

$$\begin{aligned}\mathbb{P}(\eta_j = 0) \leq\ & \left(\frac{\sigma_j}{h_j}\right)\left/\left(\sqrt{2\pi/N_j}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)\right.\right. \\ &+ \left(O((\frac{1}{\Delta_j^2} + \frac{N_j}{\Delta_j^3 M_j})e^{-M_j^2/N_j}) + \frac{N_j}{\Delta_j^6}(1 + O(\frac{N_j}{\Delta_j^2}))\right. \\ &\quad + \frac{\sqrt{N_j}N_j e^{-M_j^2/2N_j}}{\Delta_j^6}(O(\frac{1}{\Delta_j^3}) + O(\frac{N_j}{\Delta_j^5}) + O(\frac{N_j}{\Delta_j^2 M_j}) + O(\frac{1}{M_j})) \\ &\quad \left.+ \frac{N_j^2}{\Delta_j^{12}}(O(1) + O(\frac{N_j}{\Delta_j^2}) + O(\frac{N_j^2}{\Delta_j^4}))\right) \\ &\quad \left.\left/\left(\sqrt{\frac{2\pi}{N_j}}\Delta_j e^{-(M_j+2\Delta_j)^2/2N_j} - A\left(\frac{\sqrt{N_j}}{\Delta_j^3}(e^{-M_j^2/2N_j} + \frac{\sqrt{N_j}}{\Delta_j^3} + \frac{N_j\sqrt{N_j}}{\Delta_j^5})\right)\right)\right.^2.\end{aligned}$$

□

# Quelques notations

| | |
|---|---|
| $\hat{G}$ | Groupe dual (ensemble des caractères) du groupe $G$ |
| $H^{\perp}$ | Sous-groupe des caractères orthogonaux à $H$ |
| $\mu_K$ | Caractère additif canonique du corps $K$ |
| $\mathrm{Tr}_K$ | Trace de $K$ sur son corps de base |
| $\hat{f}(\mu)$ | Transformée de Fourier de $f : G \to \mathbb{C}$ en $\mu \in \hat{G}$ |
| $G_K(\chi, \mu)$ | Somme de Gauss sur $K$ en les caractères additif $\mu$ et multiplicatif $\chi$ |
| $\tau_K(\chi)$ | Somme de Gauss sur $K$ défini en $\mu_K$ |
| $\mathcal{BF}(m)$ | Ensemble des fonctions booléennes en $m$ variables |
| $d_H$ | Distance de Hamming |
| $w_H$ | Poids de Hamming |
| $RM(r, m)$ | Code de Reed-Muller d'ordre $m$ et de degré $r$ |
| $\hat{f}_\chi$ | Transformée de Walsh de $f \in \mathcal{BF}(m)$ – version corps fini |
| $\hat{\mathrm{f}}_\chi$ | Transformée de Walsh de $f \in \mathcal{BF}(m)$ – version espace vectoriel |
| $\mathrm{nl}(f)$ | Non-linéarité de $f \in \mathcal{BF}(m)$ |
| $\|g\|_\infty$ | Norme sup de $g$ sur son ensemble de définition |
| $\|g\|_E$ | Norme sup restreinte à l'ensemble $E$ de $g$ |
| $\rho(m)$ | Rayon de recouvrement du code $RM(1, m)$ |
| $R(m)$ | Rayon spectral de $RM(1, m)$ |
| $RB(m)$ | Rayon spectral équilibré de $RM(1, m)$ |
| $\mathrm{bp}(m)$ | Borne de Parseval |
| $\mathrm{bq}(m)$ | Borne quadratique |
| $\mathrm{bpw}(m)$ | Borne de Patterson-Wiedemann |
| $(v, s)$–PW | Fonction binaire de type Patterson-Wiedemann constante par rapport au sous-groupe d'indice $v$ pour un choix de coset $s$ |
| $\mathrm{w}(f)$ | Poids de Hamming de $1 - 2f$ pour $f$ une fonction binaire |
| $w_2(x)$ | Poids de la décomposition binaire de l'entier $x$ |
| $(v, s, h)$–PW généralisée | Fonction binaire de type Patterson-Widemann généralisé par rapport au sous-groupe $G$ d'indice $v$ pour un choix de coset $s$ et de représentant $h$ sur $G$ |
| $\tilde{h}_G$ | Transformée de Fourier partielle de $h$ sur $G$ |

| | |
|---|---|
| $R_v(m)$ | Rayon spectral d'indice $v$ |
| $R_E(m)$ | Rayon spectral généralisé sur l'ensemble $E$ |
| $R_E(H)$ | Rayon spectral généralisé sur l'ensemble $E$ par rapport aux caractères de $\hat{H}$ |
| $\hat{f}_K$ | Transformée de Fourier restreinte à $K$ de $f$ |
| $\Delta_{v,m}$ | Ecart angulaire des sommes de Gauss sur les caractères d'ordre $v$ dans le corps à $2^m$ éléments |
| $(\frac{k}{p})$ | Symbole de Legendre pour $p$ premier et $k \in \mathbb{Z}_p$ |
| $\varepsilon(X)$ | Espérance de la variable aléatoire $X$ |
| $\mathbb{P}(X = x)$ | Probabilité d'un événement $x$ |

# Table des figures

# Liste des tableaux

# Résumé

Cette thèse s'articule principalement autour de la théorie des codes et des fonctions booléennes liés à la cryptographie. Deux axes de recherche sont suivis : dans une première partie, nous nous concentrons sur la non-linéarité des fonctions booléennes, alors que la deuxième partie présente des applications concrètes, en cryptographie, d'objets provenant de ces théories.

Motivé par la conjecture de Patterson et Wiedemann, nous proposons une généralisation de la construction par réunions d'orbites suivant l'action d'un groupe, pour laquelle la minimisation de l'amplitude spectrale se ramène alors à deux sous-problèmes que nous étudions : l'estimation de sommes de Gauss et l'estimation de sommes d'exponentielles incomplètes. Plusieurs conditions et pistes de résolution de la conjecture sont alors détaillées. Ce travail nous permet de construire asymptotiquement des fonctions de non-linéarité plus élevée que la moyenne et nous obtenons de plus, suivant ce principe, un exemple de recollement quadratique hautement non-linéaire s'approchant de la borne de Patterson et Wiedemann en dimension 15.

Dans la deuxième partie, nous portons tout d'abord notre attention sur des protocoles cryptographiques dits à faibles ressources. Des fonctions booléennes résistantes à la cryptanalyse différentielle sont utilisées afin de protéger le protocole $HB^+$ d'une attaque par le milieu. À partir d'un deuxième protocole basé sur un principe de bruitage, nous effectuons un parallèle avec la théorie du canal à jarretière de Wyner, ce qui permet d'accroître la sécurité. D'autre part, dans le cadre de l'authentification de données variables dans le temps, une variation autour du cryptosystème de McEliece est détaillée afin de contrôler l'accès aux fonctions de vérification.



**Mots clés.** Fonctions booléennes, non-linéarité, conjecture de Patterson et Wiedemann, sommes de Gauss, cryptographie, RFID, codes correcteurs, cryptosystème de McEliece, secure sketch.



# Abstract

This thesis mainly focuses on coding theory and boolean functions which are connected with cryptography. Two research's trends are followed: the first part is dedicated to the nonlinearity of boolean functions whereas the second one shows cryptographic applications of objects coming from these theory.

Interested in Patterson and Wiedemann's conjecture, we propose to generalize their construction based on union of orbits under the action of a group from which the determination of the minimum spectral magnitude reduces to two sub-problems that we intensely study: the evaluation of Gauss sums and the estimation of some partial exponential sums. Several conditions and ideas which may help to confirm the conjecture are thus detailed. This work allows us to construct functions with a nonlinearity greater than the asymptotic mean. Moreover, thanks to this technique, we obtain an example of a quadratic spread which is highly nonlinear and close to the Patterson and Wiedemann bound in 15 variables.

In the second part, we first turn our attention to the field of lightweight cryptographic protocols. Boolean functions with specific resistance to differential cryptanalysis are introduced in the $HB^+$ protocol to strengthen it against man-in-the-middle attacks. Starting from a second protocol using the idea of background noise, we exploit the link with the wiretap channel theory of Wyner to show how to increase the security. At last we deal with authentication of noisy data and secure sketches in which a modification of the McEliece's cryptosystem is explained in order to restrict the access to checking functions.



**Keywords.** Boolean functions, nonlinearity, Patterson and Wiedemann's conjecture, Gauss sums, cryptography, RFID, correcting codes, McEliece's cryptosystem, secure sketch.