Algebraic Immunities of functions over finite fields

Gwénolé Ars, Jean-Charles Faugère

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Algebraic Immunities of functions over finite fields

Gwénolé Ars —

Jean-Charles Faugère

N° 5532

Mars 2005

Thème SYM

ISSN 0249-6399 ISRN INRIA/RR--5532--FR+ENG

INRIA
ROCQUENCOURT

Algebraic Immunities of functions over finite fields

Gwénolé Ars [*†] ,
Jean-Charles Faugère[‡§]

Thème SYM — Systèmes symboliques
Projets SALSA

Rapport de recherche n° 5532 — Mars 2005 — 17 pages

Abstract: A general mathematical definition for a function from $GF(q)^n$ to $GF(q)^m$ to resist to cryptanalytic attacks is developed. It generalize the definition of Algebraic Immunity for Stream Cipher to any finite field and also Block Cipher. This algebraic immunity correspond to equations with low leading term according a monomial ordering. We give properties of this Algebraic Immunity and also compute explicit and asymptotic bounds. We extended the definitions of Algebraic Immunity to functions with memory but they depend on the number of consecutive outputs we look at. We show that all the results obtained for memoryless function give similarly results on memory functions by a change of variables. And then, we prove that, for a memory function f with memory size l and only one output, if there is no relation which not depend on memory for l consecutive output, than we can construct a polynomial that generate all relations without memories. We apply this theorem to the summation generator and compute explicitly the Algebraic Immunity.

Key-words: Algebraic Attacks, Stream Ciphers, Bloc Ciphers, Algebraic Immunity

* LIP6 DGA/UPMC/Université Rennes 1
† Projet SALSA
‡ LIP6/LORIA CNRS/UPMC/INRIA
§ Shared foot note

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Immunité algébrique des fonctions sur les corps finis

Résumé : une definition mathématique générale de la résistance d'une fonction définie sur $GF(q)^n$ dans $GF(q)^m$ à la cryptanalyse algébrique est développé. Elle généralise la définition d"'Algbraic Immunity" des chiffrements par flot à tout corps fini et aussi au chiffrement par blocs. Cette immunité algébrique correspond aux équations de plus petit terme de tête selon un certain ordre monomial. Nous donnons des propriétés de cette immunité algébrique et nous calculons aussi des bornes explicites et asymptotiques. Nous étendons la définition d'immunité aux fonctions avec mémoire mais elle dépend du nombre de sorties consécutives régardées. Nous montrons que tous les résultats obtenues sur les fonctions sans mémoire induisent des résultats similaires par simple changement linéaire des valeurs n et m. Enfin, nous montrons que pour une fonction avec une mémoire de taille l et telle que m=1, si nous n'avons pas de relation ne dépendant pas de la mémoire pour l sorties consécutives, alors nous pouvons construire un polynôme qui engendre toutes les relations sans mémoire de la fonction pour un plus grand nombre de sorties. Nous utilisons cette propriété pour calculer explicitementl'immunité algébrique de "summation generator"

Mots-clés : attaque algébrique, chiffrement par flot, chiffrement par blocs, immunité algébrique, base de Gröbner

1 Introduction

Algebraic attacks are among the most efficient attacks for public key cryptosystems, block ciphers and stream ciphers. They try to recover a secret key by solving a system of algebraic equations. Algebraic attacks were first applied to Matsumoto-Imai Public Key Scheme in [13] by Jacques Patarin. Algebraic attacks were also applied to block ciphers in [7], where the complexity for attacking AES and Serpent was evaluated.

For Stream Cipher, the main cryptographic criteria used for boolean function had previously been a high algebraic degree to counter Berlekamp-Massey algorithm. In [6], it is demonstrated that low degree relations exist. These relations simplify the Algebraic attacks. So a significant step is to find function resistant against these attacks. In [12], the notion of Algebraic Immunity for boolean function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ is introduced.

Some relations with low degree also exist in Block Cipher, and to construct the polynomial system defined by the Block Cipher, we use this equation. We can also extend the algebraic Immunity to Block Ciphers.

The first objective of this paper is to generalize the notion of Algebraic Immunity to function over any finite fields $\mathbb{F}_q$ and we give two definitions Algebraic Immunity for Stream Cipher and Algebraic immunity for Block Cipher denoted respectively $AI_S(f)$ and $AI_B(f)$.

First, we show that these two notions are linked to Gröbner basis for a specific order on monomials, the DRL order for $AI_B(f)$ and the Elimination order for $AI_S(f)$. We prove that the definition of a function over finite fields give immediately a Gröbner basis for a lexicographic order. Having a Gröbner basis help us to find properties on the ideal generated by this basis. These properties give bounds on the notion of Algebraic Immunities. This bound are the power of the first coefficient of the following series which is negative:

$$\begin{array}{ll} AI_B(f) & \frac{q^n}{1-t} - \frac{(1-t^q)^{n+m}}{(1-t)^{n+m+1}} \\ AI_S(f) & \frac{q^{n-m}}{1-t} - \frac{(1-t^q)^n}{(1-t)^{n+1}} \end{array}$$

with f a function from $\mathbb{F}_q^n$ to $\mathbb{F}_q^m$.

From these series, we give explicit bounds on Algebraic Immunities. Then we show asymptotic bound on n for $AI_B(f)$ and $m = n$. We remark that the influence of the field $\mathbb{F}_q$ on $AI_B(f)$ and $m = n$ depends only on the $\sqrt{q}$, whereas, the dependence is on q for $AI_S(f)$ and $m = 1$.

The second part is to extend these notions for functions with memories. We notice that Algebraic Immunities can be extended to these functions but depend on the number of consecutive outputs we look at. We show that all the results obtained for memoryless function give similarly results on memory functions by a change of variables. And then, we prove that, for a memory function f with memory size ℓ and only one output, if there is no relation which not depend on memory for ℓ consecutive output, than we can construct a polynomial that generate all relations without memories. We can apply the theorem to the summation generator and we compute explicitly $AI_S(f)$ for some value of n, it correspond to n for $n \leq 9$.

Section 2 presents the definition of Algebraic immunities for memoryless functions and some properties. Section 3 devoted to the computation of explicit and asymptotic bounds of Algebraic Immunities. And section 4 generalize these notions to function with memory.

2 Definitions of Algebraic Immunities

2.1 Basic Notations and Definitions

Let $\mathbb{F}_q[\mathbf{X},\mathbf{Z}] = \mathbb{F}_q[x_1,\ldots,x_n,z_1,\ldots,z_m]$ be a polynomial ring with variables $x_1,\ldots,x_n,z_1,\ldots,z_m$ over a finite field $\mathbb{F}_q$ with cardinal q.

For a monomial $\mathbf{X}^\alpha\mathbf{Z}^\beta = x_1^{\alpha_1}\cdots x_n^{\alpha_n} z_1^{\beta_1}\cdots z_m^{\beta_m}$, $|(\alpha,\beta)| := \sum_{i=1}^n \alpha_i + \sum_{j=1}^m \beta_j$ is called the *total degree* of this monomial, denoted $\deg(\mathbf{X}^\alpha\mathbf{Z}^\beta)$ and $|\alpha| := \sum_{i=1}^n \alpha_i + \sum_{j=1}^m \beta_j$ is called the *partial degree* of this monomial, denoted $\deg(\mathbf{X}^\alpha\mathbf{Z}^\beta,\mathbf{X})$. In the following, the set of all monomials in variables $x_1,\ldots,x_n,z_1,\ldots,z_m$ is denoted by $M(\mathbf{X},\mathbf{Z})$, or simply by M. In the theory of Gröbner bases, we need to consider a *monomial ordering* (cf. [8]).

Two of such ordering is the *degree reverse lexicographical order* (DRL) and the *elimination order* defined as follows:

Definition 2.1 *For $(\alpha,\beta) = (\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_m)$ and $(\alpha',\beta') \in \mathbb{N}^{n+m}$, We say*

- $\mathbf{X}^\alpha\mathbf{Z}^\beta \succ_{DRL} \mathbf{X}^{\alpha'}\mathbf{Z}^{\beta'}$ *if $|(\alpha,\beta)| > |(\alpha',\beta')|$, or $|(\alpha,\beta)| = |(\alpha',\beta')|$ and the right-most nonzero entry of the vector $(\alpha,\beta)-(\alpha',\beta') \in \mathbb{Z}^{n+m}$ is negative.*

  $\succ_{DRL}$ *defined the DRL order of variable* $[\mathbf{X},\mathbf{Z}]$.

- $\mathbf{X}^\alpha\mathbf{Z}^\beta \succ_{Elim} \mathbf{X}^{\alpha'}\mathbf{Z}^{\beta'}$ *if $\mathbf{X}^\alpha \succ_{DRL} \mathbf{X}^{\alpha'}$, or $\mathbf{X}^\alpha =_{DRL} \mathbf{X}^{\alpha'}$ and $\mathbf{Z}^\beta \succ_{DRL} \mathbf{Z}^{\beta'}$.*

  $\succ_{Elim}$ *defined the Elimination order of variable* $[\mathbf{X}],[\mathbf{Z}]$.

There are many other monomial orderings.

A nonzero polynomial g in $k[\mathbf{X}]$ is written as $g = \sum_{\alpha,\beta} c_{\alpha,\beta}\mathbf{X}^\alpha\mathbf{Z}^\beta$, $c_{\alpha,\beta} \neq 0$. We use the following notations:

$T(g) = \{c_{\alpha,\beta}\mathbf{X}^\alpha\mathbf{Z}^\beta \mid c_{\alpha,\beta} \neq 0\}$: the set of *terms* of g and $M(g) = \mathbf{X}^\alpha\mathbf{Z}^\beta \mid c_{\alpha,\beta} \neq 0\}$: the set of *monomials* of g

We denote the *leading term*, the *leading coefficient* and the *leading term* with respect an order $\prec$, by $\mathrm{LM}(g)$, $\mathrm{LC}(g)$ and $\mathrm{LT}(g)$ respectively. (For each definition, see [8].)

The ideal in $\mathbb{F}_q[\mathbf{X},\mathbf{Z}]$ generated by a subset F is denoted by $\langle F\rangle$.

Under the above notation, a *Gröbner basis* is defined as follows.

Definition 2.2 *Let M be the set of all monomial of $\mathbb{F}_q[\mathbf{X},\mathbf{Z}]$ with a fixed ordering. A finite subset $G = \{g_1,\ldots,g_m\}$ of an ideal $\mathcal{I}$ is called a Gröbner basis if*

$$\langle \mathrm{LT}(g_1),\ldots,\mathrm{LT}(g_m)\rangle = \langle \mathrm{LT}(\mathcal{I})\rangle.$$

For a given ideal $\mathcal{I}$, its Gröbner basis is not unique. But the *reduced Gröbner basis*, which is defined as follows, is uniquely determined.

Definition 2.3 *A Gröbner basis $G = \{f_1, \dots f_m\}$ of an ideal $\mathcal{I}$ is called reduced Gröbner basis if for all i, $\mathrm{LC}(f_i) = 1$ and any monomial of f_i is not divisible by any element of $\mathrm{LM}(G \backslash \{f_i\})$.*

Proposition 2.1 *Let I be an ideal of $\mathbb{F}_q[x_1, \dots, x_n]$ and $k \in \{1, \dots, n\}$. Assume G a Gröbner basis of I for the Elimination order $[x_1, \dots, x_k]$, $[x_{k+1}, \dots, x_n]$ (or the Lexicographical order). Then $G \cap \mathbb{F}_q[x_{k+1}, \dots, x_n]$ is a Gröbner basis of $I \cap \mathbb{F}_q[x_{k+1}, \dots, x_n]$.*

So if we want to find a polynomial depending on severals variables $x_{k+1}, \dots, x_n$, we just need to compute a Gröbner basis with one of these orders.

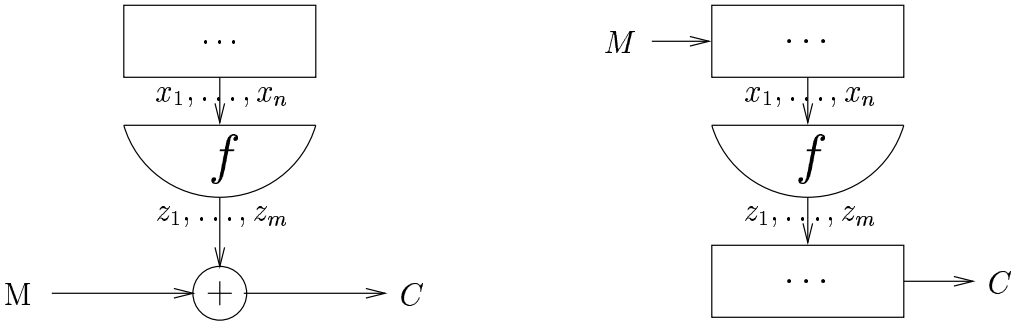
2.2 Algebraic Immunities

Let consider a function $f : \begin{cases} \mathbb{F}_q^n \to \mathbb{F}_q^m \\ \mathbf{X} \mapsto \mathbf{Z} \end{cases}$, denoted as $z_1 = f_1(\mathbf{X}), \dots, z_m = f_m(\mathbf{X})$.

Our objective is to study algebraic equations induced by the graph of f which is to solutions on the field $\mathbb{F}_q$ and not on $\overline{\mathbb{F}_q}$, the algebraic closure of $\mathbb{F}_q$. To restrict the study of equation to $\mathbb{F}_q$, we add field equations $x_i^q - x_i$, denoted as the set $\mathbf{X}^q - \mathbf{X}$ and consider elements of the ideal

$$\mathcal{I} = \langle z_1 - f_1(\mathbf{X}), \dots, z_m - f_m(\mathbf{X}), \mathbf{X}^q - \mathbf{X} \rangle.$$

In symmetric cryptography, there is two important applications: Stream Cipher and Block Cipher.



Stream Cipher Block Cipher

On Stream Cipher, we use functions to filter elements $(x_1, \dots, x_n)$. So when we do an attack on this structure, we supposed known elements $(z_1, \dots, z_m)$ and try to find relations on $(x_1, \dots, x_n)$. Whereas, in Block Cipher, non-linear functions are used inside the process to create the coded message. A possible way to write algebraic equations is to introduced intermediate variables as N. Courtois have done in [7].

Now we will present a definition of Algebraic Immunities for both cases.

Definition 2.4 *Let consider a function* $f : \mathbb{F}_q^n \to \mathbb{F}_q^m$ *and* $\mathcal{I} = \langle z_1 - f_1(\mathbf{X}), \ldots, z_m - f_m(\mathbf{X}), \mathbf{X}^q - \mathbf{X} \rangle$..

We define as:

- *Algebraic Immunity of Block Cipher,*

$$AI_B(f) = min\{\deg(P),\ P \in \mathcal{I}\}.$$

- *Algebraic Immunity of Stream Cipher,*

$$AI_S(f) = min\{\deg(P, \mathbf{X}),\ P \in \mathcal{I}\}.$$

As an ideal with field equations is radical, there is an equivalence between $\mathcal{I}$ and the set of solution of $\mathcal{I}$ which is the graph of f. So $AI(f)$ is a generalization of Algebraic Immunity defined in article [12] and first introduced in article [6].

2.3 Properties of Algebraic Immunities

This definition doesn't give us a way to have Algebraic Immunities, we can find them with Gröbner basis according some orders on monomials. We This is resume in the next theorem.

Theorem 2.1 *Let consider a function* $f : \mathbb{F}_q^n \to \mathbb{F}_q^m$.

- *A reduced Gröbner basis of* $\mathcal{I}$ *for a DRL order* $[\mathbf{X}, \mathbf{Z}]$ *contains a linear basis of polynomials* P *of* $\mathcal{I}$ *such that* $AI_B(f) = \deg(P)$.
- *A reduced Gröbner basis of* $\mathcal{I}$ *for a elimination order on* $[\mathbf{X}], [\mathbf{Z}]$ *contains a linear basis of polynomials* P *of* $\mathcal{I}$ *such that* $AI_S(f) = \deg(P, \mathbf{X})$.

Proof First we have noticed that for an DRL order $[\mathbf{X}, \mathbf{Z}]$ and any polynomial g, $\deg(g) = \deg(LM(g))$ and for an Elimination order $[\mathbf{X}, \mathbf{Z}]$ and any polynomial h, $\deg(h, \mathbf{X}) = \deg(LM(h), \mathbf{X})$.

Moreover to reduce of a polynomial g by another one h, we need that $LT(g) \succ LT(h)$.

As all polynomial of $\mathcal{I}$ are reduced to zero by a Gröbner basis, these both remarks prove the theorem. If P is a polynomial of $\mathcal{I}$ such that $\deg(P) = AI_B(f)$, resp. $\deg(P, \mathbf{X}) = AI_S(f)$, then P is reduced by the Gröbner basis G for a DRL order, respect. the Elimination order, so there is $g \in G$, such that $LT(g) \prec LT(P)$. From the first remark, we prove that G contains a linear generated family satisfying the condition of the theorem.

Then the definition of reduced Gröbner basis imply that the linear generated family is a linearly independent family. □

This theorem give us a way to compute the Algebraic Immunities. Find a Gröbner basis of this ideal $\mathcal{I}$ have a bad theoretical complexity but is very efficient in practice. There is other methods to find $AI_S(f)$ presented in [1, 6, 12].

Moreover, we have several properties as the comparison between this both notions.

Proposition 2.2 *Let consider* $f : \mathbb{F}_q^n \mapsto F_q^m$.
Then $AI_S(f) \leq AI_B(f)$.

Proof Let consider P a polynomial of $\mathcal{I}$ so that $\deg(P) = AI_B(f)$.
We have $\deg(P, \mathbf{X}) \leq \deg(P)$. Thus $AI_S(f) \leq \deg(P, \mathbf{X}) \leq \deg(P) = AI_B(f)$. □

In the article [4], we want to find algebraic relation g on the graph of f satisfying $\deg(g, \mathbf{X}) = AI_S(f)$ so that g can be written as $g(\mathbf{X}, \mathbf{Z}) = g_1(\mathbf{X}) + g_2(\mathbf{X}, \mathbf{Z})$ with $\deg(g_1, \mathbf{X}) = \deg(g_1) > \deg(g_2, \mathbf{X})$. In fact, we can give with this two Algebraic Immunities a condition of existence of these relations.

Proposition 2.3 *Let consider* $f : \mathbb{F}_q^n \mapsto F_q^m$.
There exists $g \in \mathcal{I}$ *satisfying* $\deg(g, \mathbf{X}) = AI_S(f)$ *so that* g *can be written as* $g(\mathbf{X}, \mathbf{Z}) = g_1(\mathbf{X}) + g_2(\mathbf{X}, \mathbf{Z})$ *with* $\deg(g_1, \mathbf{X}) = \deg(g_1) > \deg(g_2, \mathbf{X})$.
If $AI_S(f) = AI_B(f)$.
Moreover if $f : \mathbb{F}_2^n \mapsto F_2$, *then this condition is a necessary and sufficient condition.*

Proof If $AI_S(f) = AI_B(f)$, let g a polynomial so that $\deg(g) = AI_B(f)$, g can be written as $g_1(\mathbf{X}) + g_2(\mathbf{X}, \mathbf{Z})$ so that g_2 have no monomial which depends only of $\mathbf{X}$. This means that $\deg(g_2) > \deg(g_2, \mathbf{X})$.
We have $\deg(g) = max(\deg(g_1), \deg(g_2)) = \deg(g, \mathbf{X})$ then $\deg(g_1) = \deg(g, \mathbf{X}) \geq \deg(g_2) > \deg(g_2, \mathbf{X})$.
Thus g satisfies the condition of the proposition.
For the case of $f : \mathbb{F}_2^n \mapsto F_2$, a polynomial $g \in \mathcal{I}$ satisfying $\deg(g, \mathbf{X}) = AI_S(f)$ can be written as $g(\mathbf{X}, z_1) = g_1(\mathbf{X}) + z_1 g_2(\mathbf{X})$. As $\deg(g_2, \mathbf{X}) < \deg(g_1)$, $\deg(g) = \deg(g_1) \leq \deg(g, \mathbf{X})$. □

As we have found different properties of these algebraic immunities, we can give bound of their value.

3 Bound on the value of Algebraic Immunities

In this section, we give bounds on these Algebraic Immunities.

3.1 Properties of the ideal $\mathcal{I}$

Let consider f a function of $\mathbb{F}_q^n$ on $\mathbb{F}_q^m$. We want to find polynomials with degree $AI_B(f)$, respect $AI_S(f)$ according $\mathbf{X}$, in the ideal $\mathcal{I}$ generated by $z_1 - f_1(\mathbf{X}), \ldots, z_m - f_m(\mathbf{X}), \mathbf{X}^q - \mathbf{X}$ and especially, we consider the $\mathbb{F}_q[X_1, \ldots, X_n, Z_1, \ldots, Z_m]/\mathcal{I}$, denoted by $\mathcal{A}$.

As $\mathcal{I}$ is a zero dimensional ideal, we know that the ring $\mathcal{A}$ is a linear vector space with finite dimension. This subsection give this dimension.

G is a Gröbner basis of $\mathcal{I}$ for a lexicographic order $z_1 \succ \cdots \succ z_m \succ x_1 \succ \cdots \succ x_n]$ of $\mathcal{I}$.
Thus $\mathcal{A}$ is a linear vector space with $\mathbf{X}^\alpha \mid \alpha \in \mathbb{F}_q^n\}$ as a linear basis.
Then $\mathcal{A}$ is a vector space of dimension q^n.

We deduce that the image of q^n+1 monomials in $\mathcal{A}$ is a linearly dependent family, then there is a linear relation in this family and this relation corresponds to a polynomial of $\mathcal{I}$.

Thus considering the q^n+1 first monomials for the DRL order for $AI_B(f)$ and the Elimination order for $AI_S(f)$ is a sufficient condition to have a relation and find the degree. As the algebraic Immunity correspond to a degree, we need to count the monomials there is in $\mathbb{F}_q[\mathbf{X},\mathbf{Z}]/\langle \mathbf{X}^q-\mathbf{X}, \mathbf{Z}^q-\mathbf{Z}\rangle$ for the degree d. Let denote M_p^d the number of monomial in $\mathbb{F}_q[y_1,\ldots,y_p]/\langle y_i^q-y_i$ with degree d. Then the number of monomial m_ℓ satisfying $deg(m_\ell)=d$ in $\mathbb{F}_q[\mathbf{X},\mathbf{Z}]/\langle \mathbf{X}^q-\mathbf{X}, \mathbf{Z}^q-\mathbf{Z}\rangle$ is M_{n+m}^d and the number of monomial m'_ℓ satisfying $deg(m'_\ell,\mathbf{X})=d$ in $\mathbb{F}_q[\mathbf{X},\mathbf{Z}]/\langle \mathbf{X}^q-\mathbf{X}, \mathbf{Z}^q-\mathbf{Z}\rangle$ is $q^m M_n^d$.

We can traduced this condition as a series, a bound of the Algebraic Immunity will be the first degree of the series with a negative or zero coefficient. These series are :

$AI_B(f)$	$\frac{q^n}{1-t}-\frac{(1-t^q)^{n+m}}{(1-t)^{n+m+1}}$
$AI_S(f)$	$\frac{q^{n-m}}{1-t}-\frac{(1-t^q)^{n}}{(1-t)^{n+1}}$

We notice that the difference between the both Algebraic Immunities is only the change of variable :

$$\begin{array}{ccc} AI_B(f) & & AI_S(f) \\ n & \longleftrightarrow & n-m \\ n+m & \longleftrightarrow & n \end{array}$$

3.2 Explicit bounds

In this section, we give explicit bounds on the Algebraic Immunities which is only bounds on the first degree of the series with a negative or zero coefficient. Then we compare these bound to the computed degree of the series for given value of n,m and q. A first bound is given by minoring m by 1:

Proposition 3.1 *Let consider* $f:\mathbb{F}_q^n\to\mathbb{F}_q^m$.
We have

- $AI_B(f)\leq \begin{cases} \lfloor \frac{(n+1)(q-1)}{2} \rfloor & \text{if } q>2 \\ \lceil \frac{n+1}{2} \rceil^2 & \text{if } q=2 \end{cases}$.
- $AI_S(f)\leq \begin{cases} \lfloor \frac{n(q-1)}{2} \rfloor & \text{if } q>2 \\ \lceil \frac{n}{2} \rceil^2 & \text{if } q=2 \end{cases}$.

Proof We can prove it for $AI_B(f)$, the change of variable will give an equivalent bound to $AI_S(f)$.

As M_{n+m}^d denotes the number of monomials with degree d. We have $q^{n+m}=\sum_{i=0}^{(q-1)(n+m)} M_{n+m}^i$. Moreover $M_{n+m}^d=M_{n+m}^{(q-1)(n+m)-d}$.

Thus for $m=1$, $q^{n+1}\leq 2\sum_{i=0}^{\lfloor \frac{(q-1)(n+1)}{2} \rfloor} M_{n+1}^i$.

As $M_{n+m}^d\geq M_{n+1}^d$, then $\sum_{i=0}^{\lfloor \frac{(q-1)(n+1)}{2} \rfloor} M_{n+m}^i\geq \frac{1}{2}q^{n+1}$.

- If $q > 2$, then $\sum_{i=0}^{\lfloor \frac{(q-1)(n+1)}{2} \rfloor} M_{n+m}^i > q^n$.

  Thus there is a polynomial in $\mathcal{I}$ with degree lower or equal to $\lfloor \frac{(q-1)(n+1)}{2} \rfloor$.

- If $q = 2$, we do not have a strict inequality as $2^{n+1} = 2\sum_{i=0}^{\lfloor \frac{n+1}{2} \rfloor} M_{n+1}^i$ for $n+1$ odd.

  But we are sure that $2\sum_{i=0}^{\lceil \frac{n+1}{2} \rceil} M_{n+1}^i > 2^{n+1}$.
  Then $AI_B(f) \leq \lceil \frac{n+1}{2} \rceil$.

□

For the case of $q = 2$, the bound on $AI_S(f)$ is the bound given in the articles [6, 9].

This proposition does not take into account the number of outputs of f. It is optimal for only one output for f. But in Block Cipher, we use in general inversible function, so $m = n$. We need to give other bounds only for $AI_B(f)$. The next bounds can be extended to $AI_S(f)$ by the change of variable given in the previous section.

Proposition 3.2 *Let consider $f : \mathbb{F}_q^n \to \mathbb{F}_q^m$.*
Then
If $\binom{m+n+q-1}{q-1} \geq q^n$,

$$AI_B(f) \leq \frac{\sqrt{(n+m)^2 + 4\left((n+m)!q^n\right)^{\frac{2}{n+m}} - 2(m+n) + 1} - (n+m+1)}{2}$$

And if $\binom{m+n+q-1}{q-1} < q^n$, then

$$AI_B(f) \leq \frac{\sqrt{(n+m-1)^2 - 8(n+m+1)\Omega\Gamma + 4\Omega(\Gamma^2 + m + n)} - (n+m+1) + 2\Omega\Gamma}{2(1-\Omega)}$$

where $\Gamma = (-q + \frac{n+m-1}{2}) + \left((n+m-1)!\right)^{\frac{1}{n+m}} q^{\frac{n}{n+m}})$ and
$\Omega = \frac{1}{4} 2^{\frac{2}{n+m}} (n+m)^{\frac{2}{n+m}}$

Proof This proof is technical and quite long, we only explain in this article how to find it and refer to [2] for complete proof.

First we lower bound the number of monomials with degree lower than d by $\binom{n+m+d}{n+m} - (n+m)\binom{n+m+d-q}{n+m}$. And we find two case, $d < q$, this means $\binom{m+n+q-1}{q-1} \geq q^n$, and $d > q$.

With bounds on this binomials, we can found a lower bound as a polynomial in d with degree $n+m$. A sufficient condition to find the degree d is to have this lower bound higher than q^n.

Using the Hölder inequality, give us that a quadratic polynomial in d must be positive and then the result of the proposition. □

This bound is not useful in this form. But it can give us an asymptotic estimation in n of $AI_B(f)$, for $m = n$.

Corollary 3.1 *Let consider* $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$.

$$\frac{1}{n} AI_B(f) \leq \begin{cases} \frac{5}{6} + o_{n\mapsto+\infty}(1) & \text{if } q \leq 7 \\ \frac{2}{3}\left(\frac{4\sqrt{q}}{e} - \frac{11}{4}\right) + o_{n\mapsto+\infty}(1) & \text{if } q \geq 8 \end{cases}$$

As we can see, this bound is bad for lower value of q, it is worth than the bound given by proposition 3.1. Then we give a better bound for $q = 2$.

We can compare these explicit bounds for Algebraic Immunity for Bloc Cipher, for exemple. For fixed values of q, n and m, we determinate a bound directely from the serie. We have compare this bound withbounds of propositions 3.2 and 3.1 for $q = 16$ and $n = m$.

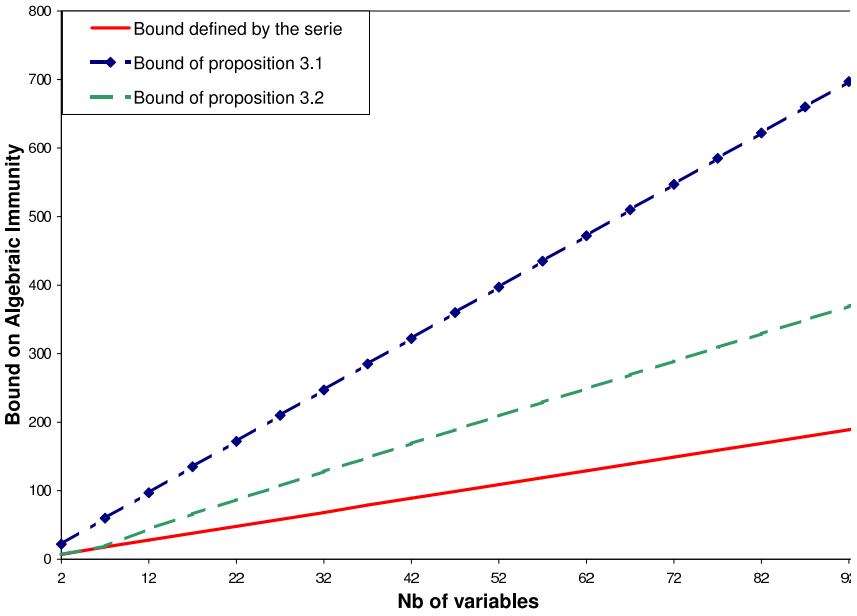


Figure 1: Comparison of bounds according the nb of variables

Figure 1 give us the difference according the value of n, with $m = n$. We see the linear behavior but the asymptotic constant $\frac{1}{n} AI_B(f)$ found is not good.

These bounds depend on the value of q, figure 2 show that the bound of proposition 3.2 is bad for small field.

Now, we give a better bound on $\mathbb{F}_2$.

Proposition 3.3 *Let consider* $f : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$.

$$AI_B(f) \leq \frac{1 - \sqrt{1 - 4\Gamma}}{2}(n + m).$$

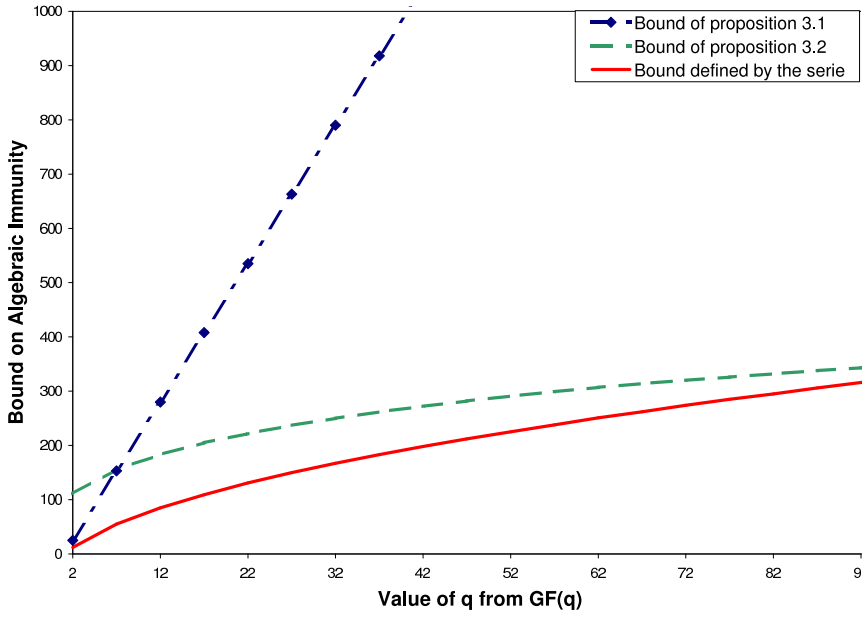


Figure 2: Comparison of bounds according cardinal of $\mathbb{F}_q$

where $\Gamma = \frac{1}{8(n+m+1)\ln 2}\left(\sqrt{(\frac{1}{12(n+m)+1} - \frac{1}{2}\ln 2\pi(n+m) - n\ln 2)^2 + \frac{4(n+m+1)\ln 2}{3(n+m)}}\right.$
$\left. -(\frac{1}{12(n+m)+1} - \frac{1}{2}\ln 2\pi - \frac{1}{2}\ln(n+m) - n\ln 2)\right).$

Proof As for proposition 3.2, we do only explain how we found it and refer to [2] for complete proof.

First, we bound the number of monomial with degree lower than d by $\binom{n+m}{d}$ and using the double inequality of H. Robbins on $n!$, we have :

$$\ln\binom{k}{d} > -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln k - \left((k\lambda + \frac{1}{2})\ln\lambda + (k(1-\lambda)+\frac{1}{2})\ln(1-\lambda)\right) + \frac{1}{12k+1} - \frac{1}{12}\frac{1}{\lambda(1-\lambda)}$$

with $d = \lambda k$.

By studying the variation of the right term as a function in $\lambda \in [\frac{1}{k}; \frac{1}{2}]$, we find a lower bound of $\ln(\binom{k}{d})$. A sufficient condition to find a bound of d is that this lower bound is higher than q^n.

This give us a quadratic polynomial in $\lambda(1-\lambda)$, where $d = \lambda(n+m)$, which must be positive. And then we find a minimum value for λ and thus the result of the proposition. □

Corollary 3.2 *Let consider* $f : \mathbb{F}_2^n \mapsto \mathbb{F}_2^n$.

$$\frac{1}{n} AI_B(f) \leq \frac{1}{8} + o_{n \mapsto +\infty}(1).$$

We compare all the bounds together.

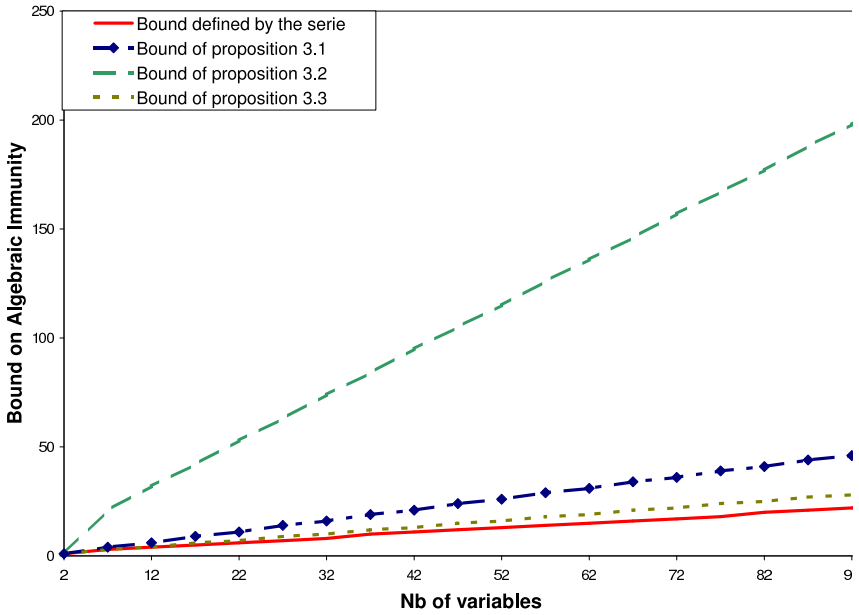


Figure 3: Comparison of bounds according the nb of variables

The bound on $\mathbb{F}_2$ is closed the bound defined by the serie.

4 Functions with memory

For Stream Cipher, use functions with memories give good cryptographic criteria. This idea is used in several stream cipher as $E0$, we find in Bluetooth.

A function with memory is a function that the output depend on the input and also a memory defined by previous inputs of size ℓ.

F. Armknecht and M. Krause have proved in article [10] that for a boolean function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ and considering several consecutive outputs, there is a relation between the inputs and the outputs which does not depend of the memories. The article [5, 9] have developed this study for functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$.

In this section, we look at the Algebraic Immunity for Stream Cipher of a function f, this algebraic immunity is defined according the number of consecutive outputs studied.

4.1 Definition

Let consider a function f with a memory of size ℓ at a moment t.

For this moment t, the memory is denoted by $\mathbf{C}^{(t)} = (c_1^{(t)}, \ldots, c_\ell^{(t)})$, the input by $\mathbf{X}^{(t)} = (x_1^{(t)} \ldots x_n^{(t)})$ and the output by $\mathbf{Z}^{(t)} = (z_1^{(t)} \ldots z_m^{(t)}$.

The function f can be written at moment t as :

$$f : \begin{cases} z_1^{(t)} &= f_1(\mathbf{X}^{(t)}, \mathbf{C}^{(t)}) \\ & \vdots \\ z_m^{(t)} &= f_m(\mathbf{X}^{(t)}, \mathbf{C}^{(t)}) \\ c_1^{(t+1)} &= P_1(\mathbf{X}^{(t)}, \mathbf{C}^{(t)}) \\ & \vdots \\ c_\ell^{(t+1)} &= P_\ell(\mathbf{X}^{(t)}, \mathbf{C}^{(t)}) \end{cases}$$

We note $\mathcal{I}_M$ the ideal generated by $\begin{cases} z_j^{(t+k)} - f_j(\mathbf{X}^{(t+k)}, \mathbf{C}^{(t+k)}) \\ c_i^{(t+k+1)} - P_i(\mathbf{X}^{(t+k)}, \mathbf{C}^{(t+k)}) \end{cases}$ for $j \in \{1, \ldots, m\}$, $i \in \{1, \ldots, \ell\}$ and $k \in \{0, \ldots, M-1\}$ and field equations $\left(\mathbf{X}^{(t+k)}\right)^q - \mathbf{X}^{(t+k)}$, $k \in \{0, \ldots, M-1\}$ and $\left(\mathbf{C}^{(t)}\right)^q - \mathbf{C}^{(t)}$.

Theorem 4.1 *Let consider f a function from $\mathbb{F}_q^n$ to $\mathbb{F}_q^m$ with memory of size ℓ and M consecutive outputs of f.*

If $M \geq \lceil \frac{\ell+1}{m} \rceil$,

Then there exists $P \in \mathbb{F}_q[\overline{X_1^{(t)}}, \ldots, \overline{X_n^{(t+M)}}, \overline{Z_1^{(t)}}, \ldots, \overline{Z_m^{(t+M)}}]$, $P \neq 0$ such that $P \in \mathcal{I}_M$.

Proof Considering the family that generate the ideal $\mathcal{I}_M$.

This family is a Gröbner basis of $\mathcal{I}_M$ according the lexicographic order $mathbf Z^{(t+M)} \succ \cdots \succ \mathbf{Z}^{(t)} \succ \mathbf{C}^{(t+M)} \succ \cdots \succ \mathbf{C}^{(t+1)} \succ \mathbf{C}^{(t)} \succ \mathbf{X}^{(t+M)} \succ \cdots \succ \mathbf{X}^{(t+M)}$. And the ideal $\mathcal{I}_M$ is a zero dimensional ideal.

Thus $\mathcal{A}_m = \mathbb{F}_q[\mathbf{X}^{(t)}, \ldots, \mathbf{C}_\ell^{(t+M)}]/\mathcal{I}_M$ defined a vector space of dimension $q^{n\,M+\ell}$.

We have $q^{(n+m)M}$ distinct monomials in $\mathbb{F}_q[\mathbf{X}^{(t)}, \ldots, \mathbf{X}^{(t+M)}, \mathbf{Z}^{(t)}, \ldots, \mathbf{Z}^{(t+M)}]/\langle \left(\mathbf{X}^{(t+k)}\right)^q - \mathbf{X}^{(t+k)}, \left(\mathbf{Z}^{(t+k)}\right)^q - \mathbf{Z}^{(t+k)}, k \in \{0, \ldots, M-1\} \rangle$.

We deduced that if $q^{(n+m)M} \geq q^{n\,M+\ell}$, then the set of the image of $q^{(n+m)M}$ distinct monomials in $\mathcal{A}_M$ is a linearly dependent family.

Thus for $M > \frac{\ell}{m}$, there exists $P \in \mathcal{I}_M$, $P \neq 0$. The first integer higher than $\frac{\ell}{m}$ is $\lceil \frac{\ell+1}{m} \rceil$. □

With this theorem, we can adapt the definition of Algebraic Immunity to functions with memories.

Definition 4.1 *Let consider f a function from $\mathbb{F}_q^n$ to $\mathbb{F}_q^m$ with memory of size ℓ and M consecutive outputs of f.*

Let denote $\mathcal{J}_M := \mathcal{I}_M \cap \mathbb{F}_q[\mathbf{X}^{(t)}, \ldots, \mathbf{X}^{(t+M)}, \mathbf{Z}^{(t)}, \ldots, \mathbf{Z}^{(t+M)}]$.

We defined the Algebraic Immunities according M outputs

$$AI_S(f, M) := \min_{P \in \mathcal{J}_M, P \neq 0} (\deg(P, \mathbf{X})),$$

$$AI_B(f, M) := \min_{P \in \mathcal{J}_M, P \neq 0} (\deg(P)).$$

4.2 properties

We have simple properties on the Algebraic Immunity.

Proposition 4.1 *Let consider f a function from $\mathbb{F}_q^n$ to $\mathbb{F}_q^m$ with memory of size ℓ and M consecutive outputs of f.*

For all $k \in \mathbb{N}$, $AI_S(f, M+k) \leq AI_S(f, M)$ and $AI_B(f, M+k) \leq AI_B(f, M)$.

Proof Let be $k \in \mathbb{N}$. If $P \in \mathcal{J}_M$ then $P \in \mathcal{J}_{M+k}$. □

As we prove in theorem 4.1, $\mathcal{A}_M$ is a vector space with dimension $q^{n\,M+\ell}$. So we can deduce bounds as in the previous section.

A bound of the Algebraic Immunity will be the first degree of the series with a negative or zero coefficient. These series are :

$AI_B(f, M)$	$\frac{q^{n\,M+\ell}}{1-t} - \frac{(1-t^q)^{(n+m)M}}{(1-t)^{(n+m)M+1}}$
$AI_S(f)$	$\frac{q^{(n-m)M+\ell}}{1-t} - \frac{(1-t^q)^{n\,M}}{(1-t)^{n\,M+1}}$

We notice that there is very few differences with Algebraic Immunity for a function without memories. Moreover, the results of Algebraic Immunity for a function without memories give results for $AI_S(f, M)$ and $AI_B(f, M)$ by the simple change of variable :

$$\begin{array}{ccc} AI_B(f) & & AI_B(f, M) \\ n & \longleftrightarrow & n\,M + \ell \\ n+m & \longleftrightarrow & (n+m)M \end{array} \qquad \begin{array}{ccc} AI_S(f) & & AI_S(f, M) \\ n-m & \longleftrightarrow & (n-m)M + \ell \\ n & \longleftrightarrow & n\,M \end{array}$$

All the bounds given in subsection 3.2 give bounds for $AI_B(f, M)$ and $AI_S(f, M)$ by using the change of variable.

With a large number of variables, the computation of a Gröbner basis can be difficult. Knowing generators of $\mathcal{J}_M$ introduced in definition 4.1 simplifies the computation. The following theorem answer partly to this question.

Theorem 4.2 *Let consider f a function from $\mathbb{F}_q^n$ to $\mathbb{F}_q$ with memory of size ℓ and M consecutive outputs of f.*

Let consider $\mathcal{J}_M$ introduced in definition 4.1.

If $\mathcal{J}_\ell = \{0\}$ *in* $\mathbb{F}_q[\mathbf{X}^{(t)}, \ldots, \mathbf{X}^{(t+\ell)}, z_1^{(t)}, \ldots, z_1^{(t+\ell)}]/\langle \left(\mathbf{X}^{(t+k)}\right)^q - \mathbf{X}^{(t+k)}, \left(z_1^{(t+k)}\right)^q - z_1^{(t+k)}, k \in \{0, \ldots, M-1\}\rangle$

then $\exists P \in \mathbb{F}_q[\overline{X_1^{(t)}}, \ldots, \overline{X_n^{(t+\ell+1)}}, \overline{Z_1^{(t)}}, \ldots, \overline{Z_1^{(t+\ell)}}]$ *so that*

$$z_1^{(t+\ell+1)} = P(\mathbf{X}^{(t)}, \ldots, \mathbf{X}^{(t+\ell+1)}, z_1^{(t)}, \ldots, z_1^{(t+\ell)})$$

And for $M \geq \ell + 1$, $\mathcal{J}_M$ *is generated by*

$$z_1^{(t+i)} - P(\mathbf{X}^{(t+i-\ell-1)}, \ldots, \mathbf{X}^{(t+i)}, z_1^{(t+i-\ell-1)}, \ldots, z_1^{(t+i-1)})$$

for all $i \in \{\ell+1, \ldots, M\}$ *and field equations on variables.*

Proof With notations of the proof of theorem 4.1, we know that $\mathcal{A}_\ell$ is a vector space of dimension $q^{\ell(n+1)}$.

As $\mathbb{F}_q[\mathbf{X}^{(t)}, \ldots, \mathbf{X}^{(t+\ell)}, z_1^{(t)}, \ldots, z_1^{(t+\ell)}]/\langle \left(\mathbf{X}^{(t+k)}\right)^q - \mathbf{X}^{(t+k)}, \left(z_1^{(t+k)}\right)^q - z_1^{(t+k)}, k \in \{0, \ldots, M-1\}\rangle$ have $q^{\ell(n+1)}$ distinct monomials and $\mathcal{J}_\ell = \{O\}$, the image of this monomials in $\mathcal{A}_\ell$ is a linear basis of $\mathcal{A}_\ell$.

So we can express all the memories $c_j^{(t+i)}$ as a polynomial in $\mathbf{X}^{(t+k)}$ and $z_1^{(t+k)}$, $k \in \{0, \ldots, \ell\}$, for all $i \in \{0,\ dots, \ell\}$ and $j\{1,\ dots, \ell\}$.

Then $z_1^{(t+l+1)}$ can be expressed as a polynomial in $\mathbf{X}^{(t+k)}$ and $z_1^{(t+k')}$, $k \in \{0, \ldots, \ell+1\}$, $k' \in \{0, \ldots, \ell+1\}$, for all $i \in \{0,\ dots, \ell\}$ and $j\{1,\ dots, \ell\}$. As same for memory $\mathbf{C}^{t+\ell+1}$.

With iteration to M, we construct polynomials that are a Gröbner basis for the lexicographic order $\mathbf{X}^{(t)} \prec \cdots \prec \mathbf{X}^{(t+M)} \prec z_1^{(t)} \prec \ldots z_1^{(t+M)} \prec \mathbf{C}^{(t)} \prec \cdots \prec \mathbf{X}^{(t+M+1)}$.

Then, from proposition 2.1, the polynomial not depending of memories is a Gröbner basis of $\mathcal{J}_M$ for the deduced lexicographic order. □

An application of this theorem is the summation generator. In this Stream Cipher, the filtering function f have n inputs $\mathbf{X} = (x_1^{(t)}, \ldots, x_n^{(t)})$ in $\mathbb{F}_2$, a memory $C^{(t)} \in \mathbb{Z}/2^\ell\mathbb{Z}$ with $\ell = \lceil \log_2 n \rceil$ the size of the memory and one output $z \in \mathbb{F}_2$ for a moment t. The definition is given by this relation :

$$z^{(t)} = x_1^{(t)} \oplus \cdots \oplus x_n^{(t)} \oplus C^{(t)} \qquad C^{(t+1)} = \left\lfloor \frac{x_1^{(t)} + \cdots + x_n^{(t)} \oplus C^{(t)}}{2} \right\rfloor$$

with $\oplus$, the sum on $\mathbb{F}_2$ and $+$ the sum in the ring $\mathbb{Z}/2^\ell\mathbb{Z}$.

In the article [11], they construct a polynomial P so that $z^{(t+\ell+1)} = P(\mathbf{X}^{(t)}, \ldots, \mathbf{X}^{(t+\ell+1)}, z^{(t)}, \ldots, z^{(t+\ell)})$. In fact in this article, they have proved the hypothesis of theorem 4.2. So, we can compute with this relation the exact value of $AI_S(f, \ell)$ and compare with the bound given by article [11]:

n	2	3	4	5	6	7	8	9
Bound of [11]	2	3	4	6	6	7	8	12
$AI_S(f, \ell)$	2	3	4	5	6	7	8	9

As we can notice, it seems that the $AI_S(f, \ell)$ is equal to n.

5 conclusion

This article generalize the Algebraic Immunity to all finite fields and also for Block Cipher. All these notions are link to Gröbner basis with a specific order : the DRL order for Algebraic Immunity in Block Cipher and the Elimination order for Algebraic Immunity in Stream Cipher.

As the definition of a function f directly give us a Gröbner basis for a lexicographic order, we prove properties of the ideal. Moreover we give explicit and asymptotic bounds on the Algebraic Immunity.

We extend this notion to function with memories over any finite fields and we give a theorem that help computing the relations implies by these Algebraic Immunity.

References

[1] F. Armknecht On the Existence of low-degree Equations for Algebraic Attacks In SASC Ecrypt workshop. Avvailable at eprint.iacr.org/2004/185/.

[2] G. Ars Applications des bases de Gröbner en cryptographie Phd Thesis in University of Rennes 1, 2005.

[3] C. Carlet Improving the algebraic immunity of resilient and nonlinear functions and constructing bent functions. On eprint.iacr.org/2004/276/.

[4] N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. In LNCS, editor, *Crypto 2003*, volume 2729, pages 177–194, 2003.

[5] N. Courtois. Algebraic Attacks on Combiners with Memory and Several Outputs. In ICISC 2004, LNCS, Springer.

[6] N. Courtois and W. Meier. Algebraic Attacks on Stream Ciphers with Linear Feedback. In Eurocrypt 2003, LNCS 2656, pp. 345-359, Springer.

[7] N. Courtois and J. Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. In *Asiacrypt*, volume 2501 of *LNCS*, pages 267–287, 2002.

[8] David A. Cox, John B. Little, and Don O'Shea. *Ideals, Varieties, and Algorithms : An introduction to computational algebraic geometry and commutative algebra*. Springer-Verlag, NY, 2nd edition, 1996. 536 pages.

[9] J. Golic Vectorial Boolean functions and induced algebraic equations. On eprint.iacr.org/2004/225/.

[10] M. Krause and F. Armknecht. Algrebraic Attacks on Combiners with Memory. In Crypto 2003, LNCS 2729, pp 162-176, Springer.

[11] D. H. Lee, J. Kim, J. Hong, J. W. Han and D. Moon. Algebraic Attacks on Summation Generators. In FSE 2004, LNCS, Springer, 2004.

[12] W. Meier, E. Pasalic and C. Carlet. Algebraic Attacks and Decomposition of Boolean Functions. In Eurocrypt 2004, pp. 474-491, LNCS 3027, Springer, 2004.

[13] J. Patarin, Cryptanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88, Crypto'95, Springer, LNCS 963, pp. 248-261, 1995.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399