The Inverse S-box, Non-linear Polynomial Relations and Cryptanalysis of Block Ciphers[⋆]

Nicolas T. Courtois
Axalto Cryptographic Research & Advanced Security,
36-38 rue de la Princesse, BP 45, 78430 Louveciennes Cedex, France
http://www.nicolascourtois.net
courtois@minrank.org

Abstract. This paper is motivated by the design of AES. We consider a broader question of cryptanalysis of block ciphers having very good non-linearity and diffusion. Can we expect anyway, to attacks such ciphers, clearly designed to render hopeless the main classical attacks ? Recently a lot of attention have been drawn to the existence of multivariate algebraic relations for AES (and other) S-boxes. Then, if the XSL-type algebraic attacks on block ciphers [11] are shown to work well, the answer would be positive. In this paper we show that the answer is certainly positive for many other constructions of ciphers. This is not due to an algebraic attack, but to new types of generalised linear cryptanalysis, highly-nonlinear in flavour. We present several constructions of somewhat special practical block ciphers, seemingly satisfying all the design criteria of AES and using similar S-boxes, and yet being extremely weak. They can be generalised, and evolve into general attacks that can be applied - potentially - to any block cipher.

Key Words: block ciphers, AES, Rijndael, interpolation attack on block ciphers, fractional transformations, homographic functions, multivariate equations, Feistel ciphers, generalised linear cryptanalysis, bi-linear cryptanalysis.

1 Introduction

AES (Rijndael) [16, 17] is a rather accomplished realisation of certain philosophy that culminates two decades of research in the design of modern block ciphers. It has important security margins and at present attacking full Rijndael is very ambitious. The research on AES focuses rather on better understanding its security by following two paths. First approach analyses the security of reduced-round versions of AES against known attacks. Another line of research is to attack, instead of Rijndael, its design principles. Though the outcome of this approach will in most cases give results not being directly applicable to AES, it remains extremely interesting. This is because Rijndael pushes some of these design principles such as high non-linearity or good diffusion to their theoretical limits, thus giving us the opportunity and motivation to explore these limits, and uncover possible pitfalls (are they serious or not).

For example the resistance of the Inverse function in $GF(2^n)$ to linear, differential and higher-order differential attacks is exceptional and close to optimality, see [4]. On page 6 of [18], the designers of AES say: *"[...] The disadvantage of these boxes is that they have a simple description in $GF(2^m)$, which is also the*

[⋆] Work supported by the French Ministry of Research RNRT Project "X-CRYPT".

field in which the diffusion layer is linear. This may create uneasy feelings, but we are not aware of any vulnerability caused by this property. For the time being we challenge cryptanalysts to demonstrate any vulnerability caused by this property. Should such a vulnerability exist, one can always replace the Sboxes by Sboxes with similar properties, that are not algebraic over $GF(2^m)$. [...]"

Unfortunately an important vulnerability of the inverse S-box does exist. It follows the line of research that has already been around for some time now. Historically the idea goes back to cryptanalysis of some rather esoteric public key schemes by Patarin [31], greatly improved by Courtois [8, 14], and followed without proper acknowledgment by Faugère and Joux [26]. The seminal idea (due to Patarin) is to study the security of a cipher component not in terms of Boolean/algebraic functions, but in terms of Boolean/algebraic **relations** that involve both inputs and output bits. This idea is very powerful, and in the last two years, it has led to a sudden collapse of several important families of stream ciphers, as demonstrated by Courtois, Meier et al in [12, 13, 1, 9, 10] and numerous other recent papers. But does it matter at all for block ciphers ?

An early warning has been issued by Jakobsen at Crypto'98 [24]. He proposes attacks on block ciphers based on univariate (and tentatively also multivariate) polynomial approximations, and already speaks explicitly of using (probabilistic) algebraic relations. Jakobsen clearly makes his point showing that to obtain secure ciphers *"[...] it is not enough that round functions have high Boolean complexity. Likewise, good properties against differential and linear attacks are no guarantee either. In fact, many almost prefect non-linear functions should be avoided exactly because they are too simple algebraically [...]".* Yet Jakobsen did not propose neither really surprising nor really devastating attacks, and so far his results are rather seen as a very special case of Generalised Linear Cryptanalysis (GLC) that breaks badly some very special ciphers and has no implication whatsoever for all the other ciphers.

Each Rijndael S-box, though very complex when regarded as a function, can be characterised in several ways by algebraic relations, cf. [11, 29], being true with very high probability, usually 1. When the XL attack was first introduced by Courtois, Klimov, Patarin and Shamir [6], it became sensible to combine these ideas and to write the problem of recovering the AES key as solving a system of multivariate quadratic equations. This seems, at first sight, rather extensively stupid, as obviously we are facing an NP-complete problem, and any other cipher can be attacked in a similar way. Even though for AES the system is somewhat over-determined, and even with an optimistic evaluation of XL, there is clearly no hope to get an attack faster than 2^{300}. Yet with time, this idea appeared less and less stupid. Courtois and Pieprzyk proposed a method called XSL [11] that allows to substantially lower the (still naive) complexity estimation of an algebraic attack, by adapting the basic idea of XL to the sparsity and the specific structure of these equations. Then Murphy and Robshaw followed [29] with an (in theory) equivalent version of the same approach, writing quadratic equations over $GF(256)$ instead of $GF(2)$, yet yielding more sparsity, and giving hopes for even faster attacks. Thus a rather outrageous idea appeared: this version of XSL attack appears (in first, very naive estimations) to have a potential to recover an AES key in less than 2^{128} AES computations, given only one single known plaintext. So far the real feasibility of such attacks is far from being clear.

This paper is also exploiting, very loosely speaking the same, vulnerabilities of the inverse function in $GF(2^n)$ function. As Jakobsen in [24] we will work on multivariate/univariate approximations and relations. Our goal is not to propose attacks on AES, it is more educational: we wish to demonstrate that there are ciphers that have very high-nonlinearity and exceptionally good diffusion but can be broken in practice even for a very large number of rounds. We will make extensive use of the inverse function in $GF(2^n)$ and some of its linear equivalents, both as a component for building highly insecure ciphers, and (quite surprisingly) as the algebraic structure that can be exploited in attacks.

In this paper we follow the following methodology: first we develop some results about composing functions with various operations. Then construct weak contrived ciphers incorporating the inverse in $GF(2^n)$, thus looking secure and satisfying all the known design criteria. Finally from here we develop general families of attacks applicable to (more or less) any cipher. The paper is composed of two rather independent parts. In the first part we study, improve and propose attacks that are based on univariate algebraic equations/relations. These are applied to propose weak Substitution-Permutation Network ciphers (SPN), and will lead to describing a new very general class of attacks on such ciphers. In the second part we are rather concerned with Feistel ciphers and we will study attacks based on bi-variate and multivariate equations.

Part I - Insecure Substitution-Permutation Networks

2 Whitening Ciphers and Known Weak Constructions

We define a subclass of Substitution-Permutation Network (SPN) ciphers that we call Whitening Ciphers. There are ciphers in which we alternatively XOR the state with some derived key $K_i \in GF(2^n)$, and some function $F : GF(2^n) \to GF(2^n)$ that does not depend on the key and in most cases (but not always) will be bijective and identical for every round.

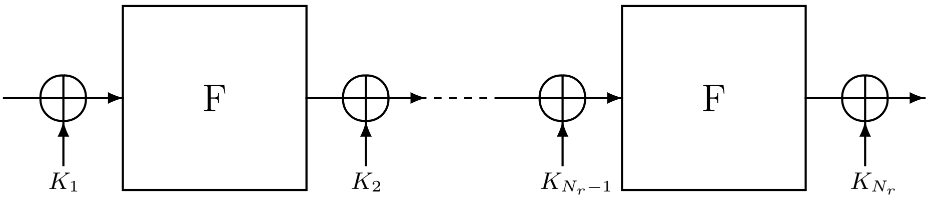


Fig. 1. Whitening Ciphers with identical round function F

Serpent [3] and Rijndael [16] are perfect examples of whitening ciphers. Many other ciphers, for example DES and other Feistel ciphers that use XOR to combine consecutive keys, can still be seen as whitening ciphers. In this case F is not bijective, a different F is used in the first and the last round, the intermediate data are redundant, and the key schedule is weak - some bits are always 0.

We start with some simple examples of insecure whitening ciphers.

2.1 When F is of Low Degree

For example $n = 128$, F is a polynomial of degree 3, and the number of rounds is $N_r = 16$. Then the whole encryption function is a fixed key-dependent polynomial of degree $D = 3^{16} \approx 2^{25}$. This is not a big degree.

The resulting attack is called an *instance deduction* (cf. [20]) - it does not recover the key but a partial equivalent of it, for example a formula that allows to decrypt/encrypt a certain fraction of messages. Naively the full polynomial can be recovered by Gaussian reduction given D known plaintexts. Alternatively, Jakobsen and Knudsen show on page 5 in [22]. that it can be done much faster in time essentially D, given D chosen plaintexts, In [25] it was claimed that it should be done in time D even with D *known* plaintexts, but no proof is given and the result does not seem to be true. At any rate, when $D = 3^{16} \approx 2^{25}$, at least the chosen plaintext attack can be handled in less than 1 second on a PC.

Remark: In practice, if we want to build such a cipher, an additional problem will be to have F that is bijective. When n is odd, $GCD(3, 2^n - 1) = 1$ and $X \mapsto X^3$ can be used. Other solutions are known: Dickson polynomials. For example when $2^n \equiv 2 \bmod 5$ or $2^n \equiv 3 \bmod 5$, the polynomial $X^5 + X^3 + 1$ is a permutation, see Section 7 of [27].

2.2 When F is Approximated by a Function of Low Degree

Then, following Jakobsen, the cipher is still insecure. When the whole cipher can be approximated with a polynomial of degree D^{N_r} true with probability ε^{N_r}, then following [24], we may apply Sudan's algorithm to recover this equation, with a complexity of these attacks being a low degree polynomial in $\frac{D^{N_r}}{\varepsilon^{2N_r}}$.

Remark: These attacks by polynomial interpolation can be declined in two versions. It can be exploited as a known plaintext attack, and in this case it is a special case of Generalised Linear Cryptanalysis (GLC). It can also be exploited in a chosen plaintext attack, and in this case it becomes a special case of Higher-order Differential Attacks, see [22]. The second variant is less frequently applicable in practice, but the complexity to recover the polynomial should be lower (at least in the non-noisy case, as discussed above). Moreover for many practical ciphers, there is no need at all for recovering the polynomial even in the noisy case (!). We can use a differential of some order that makes the polynomial vanish, to detect the noisy polynomial approximation without recovering it, and build a distinguisher on $N_r - 1$ rounds that should allow key recovery for the full cipher by guessing some relevant key bits in the last round.

What's Next. So far we only have serious, but marginal attacks that operate only on contrived (weak) ciphers. Later we will propose more general constructions of insecure ciphers, that can simultaneously incorporate several different components. At some point we will we obtain a completely general attack that applies (potentially) to any cipher, contains the Jakobsen attacks described above, contains the linear cryptanalysis, and - importantly - does break the barrier of linear/low degree approximations.

3 Insecure Ciphers Based on the Rijndael S-box

Some functions became favorite components for designing block ciphers because they are anything but linear/low degree approximations. Let $F = Inv$ be the full-size inverse function with the usual $0 \mapsto 0$:

$$Inv(X) = \begin{cases} X^{-1} & \text{in } GF(2^n) \text{ if } X \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

For now, it is full size, and the representation of the finite field is assumed to be known (later we will consider hidden fields of unknown size).

We call an Almost-Invariant, any property that is invariant but remains true with a slowly decreasing probability. We have the following theorem:

Theorem 3.0.1 (Almost-Invariant Polynomial Relation Attack). For any cipher $X \mapsto Y = E_K(X)$ that composes in any order

(a) N_r applications of Inv in $GF(2^n)$,
(b) any number of XORs with different subkeys or constants,
(c) any number of multiplications by a subkey or a constant, must be $\neq 0$,

there exist $(\alpha, \beta, \gamma, \delta) \in GF(2^n)^4$ such that:

$$\mathbb{P}_{X \in GF(2^n)}\left[\alpha XY + \beta X + \gamma Y + \delta = 0 \mid Y = E_K(X)\right] \geq \left(1 - \frac{1}{2^n}\right)^{N_r} \geq \left(1 - \frac{N_r}{2^n}\right)$$

Proof. The proof is done by induction. We assume that for a cipher with N_r rounds, there is an equation of the form

$$\alpha XY + \beta X + \gamma Y + \delta = 0$$

that holds for the fraction $\left(1 - \frac{1}{2^n}\right)^{N_r}$ of all $(X, Y) = (X, E_K(X))$. The cases of adding one addition or one multiplication are obvious, the equation still exists. What if we add an Inv ? Let $T = \begin{cases} Y^{-1} & \text{if } Y \neq 0 \\ 0 & \text{otherwise} \end{cases}$.

We multiply the equation above by T:

$$\alpha TXY + \beta TX + \gamma TY + \delta T = 0$$

We have $TY = 1$ for all X except when $Y = 0$. The resulting equation will be true with probability that gets multiplied by $\left(1 - \frac{1}{2^n}\right)$.

$$\beta TX + \alpha X + \delta T + \gamma = 0$$

The probability decreases. (This equation will probably not be true when X is such that $Y = 0$, and $T = 0$, except if $\alpha X + \gamma = 0$ which is in general not true.) This ends the proof. □

Remark: When N_r is small, a special case of this theorem exists in another, different form, see equation (6) on page 11 of Jakobsen and Knudsen paper [22] (also appearing as (9) in [25]). Yet as it stands, the result of [22] is false and below we give a generalised and corrected version of it. It can also be seen as another, equivalent (but seemingly better) formulation of Theorem 3.0.1:

Theorem 3.0.2 (Homographic Approximation Version of Thm. 3.0.1). For any cipher $X \mapsto Y = E_K(X)$ that composes in any order

(a) N_r applications of *Inv* in $GF(2^n)$,
(b) any number of XORs with different subkeys or constants,
(c) any number of multiplications by a subkey or a constant, must be $\neq 0$,
there exist $(\alpha, \beta, \gamma, \delta) \in GF(2^n)^4$ such that:

$$\mathbb{P}_{X \in GF(2^n)}\left[Y = \frac{\alpha X + \beta}{\gamma X + \delta} \mid Y = E_K(X)\right] \geq \left(1 - \frac{1}{2^n}\right)^{N_r} \geq \left(1 - \frac{N_r}{2^n}\right)$$

Proof. The proof is obvious and even easier than for Theorem 3.0.1 above.

Comparison to [22]. First of all, our result is more general compared to (6) in [22] that does not include the "multiplying by a constant" case (c). Moreover, there are two flaws in the result of [22]: the authors assume that $\alpha \neq 0$ and they neglect the singularity problem of *Inv* claiming that the approximation would be true all the time. Both these omissions are not serious when the number of round is small $N_r \ll 2^n$ and in this case the result (6) of [22] is true with good probability - for ciphers combining only elements of type (a) and (b). Otherwise, the correct result is our Theorem 3.0.2 or Theorem 3.0.1.

4 Composition and Approximation Properties of *Inv*

As we will see now, the theory of Rijndael *Inv* is rich and non-trivial.

4.1 Homographic Functions

In mathematics the functions of the form $X \mapsto \frac{\alpha X + \beta}{\gamma X + \delta}$ are called *homographic functions* or *linear fractional transformations* or *Möbius transformations*, see [35]. It is well known that they can be represented by 2×2 matrices $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. The composition of these functions is equivalent to multiplying their matrices (!).

A cross-ratio of 4 pairwise different points $R(t, u, v, w) = \frac{t-u}{t-w} / \frac{v-u}{v-w}$ is known to be an invariant for such transformations The cross-ratio can therefore be used in cryptanalysis as suggested by Vaudenay and Aoki, see Section 2.4. of [2]. However, for cryptanalysis of compositions such as in Theorems 3.0.2 and 3.0.1, with *Inv* version of the inverse, it is again an almost-invariant, not an invariant.

4.2 What is the Difference Between *Inv* and the Inverse ?

More precisely, it will be invariant as long as we do not encounter any singularity in which 0 is mapped to 0. Thus we have:

1. $R(a + k, b + k, c + k, d + k) = R(a, b, c, d)$,
2. $R(\mu a, \mu b, \mu c, \mu d) = R(a, b, c, d)$ for $\mu \neq 0$,
3. $R(1/a, 1/b, 1/c, 1/d) = R(a, b, c, d)$ for non-zero elements,
4. $R(1/a, 1/b, 1/c, 0) = a/c \cdot R(a, b, c, 0)$.

This causes a discontinuity in the invariant.

We see that, the function *Inv* of Rijndael is **not** strictly speaking a homographic function. It is equal to a function of the form $X \mapsto \frac{\alpha X + \beta}{\gamma X + \delta}$ except in one point, when 0 is mapped to 0. It is possible to see that this "completion" is the reason why in Theorems 3.0.2 and 3.0.1. the probability **does** decrease with the number of rounds. Functions that we get by composition are less and less homographic and this is why we talk about *homographic approximation.*

This "completion" with $0 \mapsto 0$ has important and non-trivial properties. There are three ways of defining the inverse function for a finite field:

1. We can have a bijection on 255 elements $GF(256)^* \rightarrow GF(256)^*$, thus avoiding the pole. This is acceptable as long as inversion alone is concerned, but in general makes it impossible composing with other homographic functions to form a group: they would have poles at different places, and functions with a different domain can not easily be composed.
2. We can have a bijection on 257 elements $\overline{GF(256)} \rightarrow \overline{GF(256)}$ with $\overline{GF(256)} = GF(256) \cup \{\infty\}$. For example the inverse will be defined as:
$$\overline{Inv}(X) = \begin{cases} X^{-1} & \text{if } X \notin \{0, \infty\} \\ 0 \mapsto \infty \\ \infty \mapsto 0 \end{cases} .$$
This is an eminently interesting version. We can compose this function with other homographic functions defined as follows:
$$H_{a,b,c,d}(X) \overset{def}{=} \begin{cases} \frac{aX+b}{cX+d} & \text{if } X \notin \{-\frac{d}{c}, \infty\} \\ -\frac{d}{c} \mapsto \infty \\ \infty \mapsto \frac{a}{c} \end{cases} \qquad \text{with } \; det\begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0.$$
When composing such functions we still get homographic invertible transformations. They form a group.
3. We can have a bijection on 256 elements $Inv : GF(256) \rightarrow GF(256)$ that is used in Rijndael. It is important to note that Inv can be seen as as restriction to $GF(256)$ of $\overline{Inv} \circ \tau$, with τ being a function that swaps ∞ and 0, leaving all the other points unchanged. This "swap" that occurs in Inv when we put $0 \mapsto 0$ explains why when we compose Inv with key additions, the composition function will be homographic with decreasing probability as shown by Theorems 3.0.2 and 3.0.1.

4.3 Big and Small Groups, Group Approximation, AES and DES

The homographic function $1/X$ over $\overline{GF(256)}$ composes well with constant/key additions and with constant/key multiplications to form a quite small group. What is the group generated when we replace $1/X$ par Inv ? With good probability, we still get homographic functions, but the approximation probability decreases with the number of rounds. The answer is non-trivial and rather surprising. The group generated is the group of all permutations. Moreover, the result remains true also when we do not use multiplications and compose only Inv with additions of constants. We have the following result:

Theorem 4.3.1 (The Group Generated by *Inv* and XORs).
The group generated by composing Inv and constant/key additions is exactly the group of all permutations of $GF(2^n)$.

The proof is given in Appendix A.

Consequences for Block Cipher Cryptanalysis

This fact is very closely related to the famous question whether DES is a group or not. If DES were a group it would have serious consequences on the security of DES, because it means that triple DES would not really be more secure than

DES. Luckily it was concluded that DES is not a group [5] (for AES see [34]). However, as we will explain now, the question of such attacks on triple DES remains widely open. Assume that, as for Theorem 4.3.1 compared to Theorem 3.0.2, DES is not a group, but yet each 64-bit permutation obtained with DES can be seen as equal with some probability, to an element of some group G. Then even if this probability were quite small, (e.g. 2^{-10}), we would obtain an attack on triple DES as follows. We try to guess a presumably 56-bit (maybe less) information that characterise the group element $g \in G$ expected to approximate our triple DES instance. Given a pair (P, C) produced with the triple-DES, the property $C = g(P)$ will be satisfied with probability 2^{-30}. A random permutation will satisfy it with a negligible probability of about 2^{-64}. Thus the right g can be detected in practice. We get a practical distinguishing attack requiring about 2^{56} computations and $\mathcal{O}(2^{30})$ known plaintexts.

We conclude that it is not enough to show that DES or AES is not a group. One should design block ciphers in such a way that the encryption operations (or single round operations) should not belong to a small group, and should not have "good" approximations by elements of some group that is not too large.

5 More Constructions of Insecure SPN Ciphers

With Theorem 3.0.1, we (already) get a very interesting result. We compose *Inv*'s, multiplication by some (key-dependent or not) constants and XORs by other constants. We get a family of block ciphers such that:

1. It is stable by composition of a few ciphers (but not for a big number).
2. Since they apply the inverse to the whole state, they are strongly resistant to linear, differential and higher-order differential attacks, see [4].
3. The security of these ciphers (and also for more general ones we will propose later) does **not** grow exponentially with the number of rounds:
   (a) When the number of rounds is exponential, up to about 2^n, for example $N_r = 2^n$, these ciphers can be easily broken given $\mathcal{O}(1)$ known plaintexts. If our goal is only to distinguish the cipher from random, we can use the cross-ratio (cf. Section 4.2 or [2]). For any 4-tuple of known plaintexts it should be invariant with good probability. Otherwise, we use the equation of Theorem 3.0.1 that is true with large probability being at least $(1-1/2^n)^{2^n} \approx 1/e$. Any subset of four known plaintext allows to recover the equation with good probability, that is then checked: should remain valid for an important fraction of other plaintexts. We get an instance deduction attack that recovers the $(\alpha, \beta, \gamma, \delta) \in GF(2^n)^4$ and uses them to encrypt/decrypt any message with good probability.
   (b) However, when the number of rounds is very large, some of these ciphers are **provably secure without any assumption**. This is an immediate corollary of Theorem 4.3.1.

5.1 Combining with Polynomial Equations

We will combine our class of insecure ciphers with the Jakobsen attack of [24]: we allow also components that are polynomials of small degree.

Theorem 5.1.1 (Higher-Degree Homographic Approximation Attack). For any cipher $X \mapsto Y = E_K(X)$ that mixes:

(a) N_r applications of Inv in $GF(2^n)$,
(b) any number of XORs with different subkeys or constants,
(c) any number of multiplications by a subkey or a constant, must be $\neq 0$,
(d) small number of rounds that are small degree polynomials with the total product of their degrees being D.
(e) all these combined with noise, or equivalently we assume that all the components of the cipher are not (a-d) but equal to such with some probability ε_i, and with total combined approximation probability being $\varepsilon = \prod \varepsilon_i$.

Then there exist two polynomials $P(X)$ and $Q(X)$ of degree D such that:

$$\mathbb{P}_{X \in GF(2^n)} \left[Y = \frac{P(X)}{Q(X)} \mid Y = E_K(X) \right] \geq \varepsilon \left(1 - \frac{1}{2^n}\right)^{N_r} \geq \varepsilon \left(1 - \frac{N_r}{2^n}\right)$$

Resulting Attack: The existence of such polynomial relations can be efficiently checked with a bivariate version of the Sudan's Algorithm, see [24].

Proof of Theorem 5.1.1. Again the proof is done by induction and we need to verify step by step that all transformations preserve the property, with degree of the polynomials increasing multiplicatively in the case (d). The cases (a) and (c) are completely trivial. For (b) we write: $P(X)/Q(X)+C = (P(X)+C \cdot Q(X))/Q(X)$. For (d) we observe that if we apply a polynomial of degree D_1 $A = a_0 \ldots a_{D_1} X^{D_1}$ to a fraction $(P(X)/Q(X))$ of two polynomials of degree D_2 the result can be written as:

$$\frac{a_0 Q(X)^{D_1} + a_1 P(X)^1 Q(X)^{D_1 - 1} + \ldots a_{D_1} P(X)^{D_1}}{Q(X)^{D_1}}.$$

Clearly a fraction of two polynomials of degree $D_1 D_2$. This ends the proof. □

5.2 Further Extension

Our class of insecure ciphers (and our attack) can be extended by using automorphisms of the finite field $GF(2^n)$. This allows to encompass more linear equivalents of the inverse function (so far we used only $X \mapsto a/X$). It also allows to include polynomials that are of high degree without increasing the final D.

Theorem 5.2.1 (Extended Higher-Degree Homographic Attack). For any cipher $X \mapsto Y = E_K(X)$ that mixes:

(a) N_r applications of Inv in $GF(2^n)$,
(b) any number of XORs with different subkeys or constants,
(c) any number of multiplications by a subkey or a constant, must be $\neq 0$,
(d) small number of rounds that are small degree polynomials with the total product of their degrees being D,
(e) any number of squares in $GF(2^n)$,
(f) all these combined with noise, or equivalently we assume that all the components of the cipher are not (a-d) but equal to such with some probability ε_i, and with total combined approximation probability being $\varepsilon = \prod \varepsilon_i$.

Then there exist $r \in \mathbb{N}$ and two polynomials $P(X)$ and $Q(X)$ of degree D such that:

$$\mathbb{P}_{X \in GF(2^n)}\left[Y = \frac{P(X^{2^r})}{Q(X^{2^r})} \mid Y = E_K(X)\right] \geq \varepsilon\left(1 - \frac{1}{2^n}\right)^{N_r} \geq \varepsilon\left(1 - \frac{N_r}{2^n}\right)$$

Proof. The proof is nearly the same. We observe that the Frobenius automorphism (e) commutes with all the other operations (a-d), except it replaces constants/subkeys by a different constant, and polynomials by a different polynomial. Therefore we can safely put all squares at the beginning and we get the result from Theorem 5.1.1. □

Resulting Attack: It is still possible to check if such equation exists, we guess $r \in \{0, n-1\}$ and proceed with a version of Sudan's Algorithm cf. [24].

6 New General Attack on Whitening Ciphers

Now we will introduce a new very general attack that can be applied to potentially any whitening cipher (and even to other ciphers). The idea is as follows: consider 128-bit whitening cipher. It is very unlikely that it has any kind of equational property such as in Theorem 5.2.1. However if we select some, say 4 bits in one round, and a different set of 4 bits in the next round input, and choose some special representation of the field $GF(2^4)$, approximations of the form $\frac{P(X^{2^r})}{Q(X^{2^r})}$ such as as in Theorem 5.2.1 may indeed exist.

Summary of the General Attack. We resume here all the different choices that the attacker should explore to find the best attack of this type.

1. Choose the size of the field, for example $m = 4$.
2. Select some input and output m-bit masks for one round that can be the same (invariant attacks) or different (much more possibilities). These masks can be subsets of bits, can be linear selection functions and can even by arbitrary non-linear functions $GF(2)^n \to GF(2)^m$.
3. Masks should be selected in such a way that a bias exists from the information theoretical point of view: the output mask seen as a parameterised function of the inputs should not be uniform.
4. Select a representation of the two finite fields.
5. Choose parameters (D, ε), guess r, and find the polynomial equation of Theorem 5.2.1 by a version of Sudan's algorithm, cf. [24].
6. By combining connecting approximations of this type, exactly as in linear cryptanalysis, we obtain attacks for an arbitrary number of rounds.

6.1 Our New Attack - Summary of What We Get

1. The possibilities offered by this attack are very large and given a cipher it is hard to know if it can be efficiently applied.
2. It is a special case of GLC. (Looking for attacks that are excessively general doesn't make sense, as we will not be able to explore them and to see if they do apply to a particular cipher).
3. It is not excessively general. Given a particular cipher such as AES, it is hard but probably still possible to find all interesting ways of applying it.

4. It contains linear cryptanalysis. (When $m = 1$. Masks are multivariate functions $GF(2^n) \rightarrow GF(2^1)$, approximation by homographic functions in $GF(2^1)$ amounts to using only the identity function in $GF(2^1)$.).
5. It contains all the attacks of Jakobsen from [24]. It goes beyond: it works locally instead of globally, and does no longer require the components to have low degree approximations.
6. Though both previous attacks will be easily prevented if only we use one big Inv function inside the cipher, the new attack can tolerate an arbitrary number of inverses (of the same size m). Thus it allows to cryptanalyse many ciphers that have very high non-linearity and resist to classical attacks.
7. Since it is a generalised linear attack, it is possible to show by using the Fourier Transform, that if the selection of the m bits is done in a linear way, and if m is small, this attack cannot be "much" faster than the best linear attack on the same cipher. It can however be strictly better (and it is obvious to construct examples by embedding noisy polynomials in a round function). In the case of non-linear selection functions, **or** if m is bigger (e.g. ≥ 32) these attacks can be much faster than any other known attack, as it was already the case in our simple example, cf. point 3a in Section 5.
8. With this attack it is obvious and easy to construct many quite complex ciphers such that their complexity does not grow exponentially with the number of rounds. All we need is to embed in each round an arbitrary combination of the components of Theorem 5.2.1, with the input and the output being hidden. This embedded approximation can be systematically (i.e. for every round) highly non-linear and can also be systematically non-polynomial.

Part II - Constructing Insecure Feistel Ciphers

7 Weak Feistel Ciphers Based on the Inverse S-box

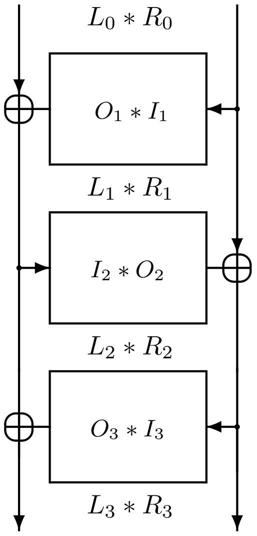
In this section we will exploit a special case of Bi-Linear Cryptanalysis (BLC) [7], being itself a special case of Generalised Linear Cryptanalysis (GLC) [21].We do not really need to understand the whole BLC [7], and only recall the basic principles of BLC when needed. This paper can be read independently without knowing BLC. BLC allows to construct ciphers that look secure w.r.t. the state of the art in cryptanalysis, yet there are extremely weak. For example, we consider a Feistel cipher in which the round function is given by:

$$f_i(X) = K_i \cdot Inv(X) \qquad \text{in } GF(2^n),$$

with $K_i \in GF(2^n)$ being the partial key. We will show that this cipher is insecure. Our notations are as follows: We consider a Feistel cipher with N_r rounds. Let $I_i \in GF(2^n)$ and $O_i \in GF(2^n)$ denote respectively the input and the output of the i-th round function $i = 1..N_r$, let (L_0, R_0) be the input and (L_{N_r}, R_{N_r}) be the output. (Note: in this paper we use "untwisted" version of the Feistel schemes, as on the right-hand figure, page 254 in [28]. Thus the meaning of L and R is as on Fig. 2 and differs from several other papers).

We have then (see Fig. 2) or [7]):

$$L_{N_r} \cdot R_{N_r} \oplus L_0 \cdot R_0 = \sum_{i=1}^{\lceil N_r/2 \rceil} O_{2i-1} \cdot I_{2i-1} \;\oplus\; \sum_{i=1}^{\lfloor N_r/2 \rfloor} I_{2i} \cdot O_{2i} = \sum_{i=1}^{N_r} I_i \cdot O_i$$



$$= \cancel{L_1 * R_1} - L_0 * R_0$$

$$= \cancel{\cancel{L_2 * R_2}} - \cancel{L_1 * R_1}$$

$$= L_3 * R_3 - \cancel{\cancel{L_2 * R_2}}$$

Fig. 2. The principle of Bi-linear Cryptanalysis over $GF(2^n)$ for a Feistel cipher

For every round i, by definition of the round function, we have:

$$I_i \cdot O_i = K_i \text{ with probability } \left(1 - \frac{1}{2^n}\right)$$

And thus we get the following I/O sum for the whole cipher:

$$L_{N_r} \cdot R_{N_r} \oplus L_0 \cdot R_0 = \sum_{i=1}^{N_r} K_i \text{ with probability } \left(1 - \frac{1}{2^n}\right)^{N_r}$$

We do not know $\sum K_i$ but it can be recovered from one known plaintext. Then, this equation allows to distinguish our cipher from a random permutation given 2 plaintexts, for $N_r \ll 2^n$ and with negligible error probability of about 2^{-n}. Even when the number of rounds is $N_r = 2^n$ our characteristic is true with probability $(1-1/2^n)^{2^n} \approx 1/e$ and thus the cipher can be still distinguished from random with good error probability, given about $\mathcal{O}(e^2)$, i.e. a small constant number of plaintexts. (Our formula allows also to recover the plaintext if a half of it can be guessed in a dictionary attack.)

We get a cipher with the following properties:

1. It is based on the inverse in $GF(2^k)$.
2. It mixes two group operations: addition and multiplication in $GF(2^n)$.
3. It has very good diffusion and avalanche properties.
4. It is composed of very "good" Boolean functions (cf. [4]) and thus resists to all common attacks on block ciphers including LC and DC.
5. Yet the security of it does **not** grow exponentially with the number of rounds. It is easy to break in practice even for 2^n rounds, given about 10 plaintexts.

7.1 More Weak Feistel Ciphers Based on the Inverse S-box

What we described here is a very special case of the general family of insecure Feistel ciphers specified in [7]. It is possible to see that in general we have:

Theorem 7.1.1 (General Construction of Weak Feistel Ciphers). If the round function is such that there is a symmetric quadratic multivariate relation over $GF(2)$ (or any other field) between the input and output bits, then the cipher is insecure and can be distinguished from a random permutation given a small constant number of plaintexts.

We omit the proof. (This result is new, but rather obvious if we read [7].)

Further Extensions. Obviously the equations can be probabilistic. Also the representation of the field may be secret, however here, unlike in Section 6, if this representation is linear over $GF(2)$, it does not help: bi-linear equations over $GF(2^k)$ always give bi-linear equations over $GF(2)$. Then, given a cipher it becomes a hard problem to see if it is weak w.r.t. such attacks, even if we are aware of Theorem 7.1.1. Moreover weak ciphers do not limit to these specified by this theorem. In general the multivariate bi-linear relations can also include some linear parts that, when combined for a whole cipher, can be recovered from linear cryptanalysis. Below we such example, that really does not look as a weak cipher, yet it can be broken easily even for one thousand rounds.

Example. We consider a 64-bit Feistel cipher with 32-bit constant parameter $(c, c') \in GF(2^{16})^2$ expanded key being $(K_1, K_1', \ldots, K_{N_r}, K_{N_r}'$ and in which the round function is $(x, x') \mapsto (y, y')$ with $x, x', y, y' \in GF(2^{16})$ defined as:

$$\begin{cases} y = \frac{x+x'+K_i}{x+c} \text{ in } GF(2^{16}) \\ y' = \frac{x+x'+K'_i}{x'+c'} \text{ in } GF(2^{16}) \end{cases}$$

This cipher looks very secure, yet we have the following symmetric relation with additional linear terms, true with probability about $1 - 2^{-15}$:

$$xy + x'y' = K_i + K'_i + yc + y'c' \text{ in } GF(2^{16})$$

This implies that for the whole cipher, if we denote the left part of the input by (L_0, L'_0) and so on respectively, we have in $GF(2^{16})$:

$$L_{N_r} \cdot R_{N_r} \oplus L'_{N_r} \cdot R'_{N_r} \oplus L_0 \cdot R_0 \oplus L'_0 \cdot R'_0 = \sum_{i=1 \ldots N_r} (K_i \oplus K'_i) \oplus \sum_{i=1 \ldots N_r} (c \cdot O_i \oplus c' \cdot O'_i) =$$

$$= \sum_{i=1 \ldots N_r} (K_i \oplus K'_i) \oplus c \cdot (L_0 \oplus L_{N_r} \oplus R_0 \oplus R_{N_r}) \oplus c' \cdot (L'_0 \oplus L'_{N_r} \oplus R'_0 \oplus R'_{N_r}).$$

This equation holds with probability about $(1 - 2^{-15})^{N_r}$ and allows to distinguish the cipher from a random permutation given a few plaintexts, whatever is the number of rounds, for up to about 2^{15} rounds. Another very weak cipher based on inverse in $GF(2^n)$.

Extensions of this particular construction. As in [7] it is possible to see that if $G(x, x')$ is a component (that can be key-dependent) such that fixed some linear combination of outputs of G is biased, then we can replace K_i by $K_i + G(x, x')$ and K'_i by $K'_i + G(x, x')$ in our definition of the round function, and the cipher will still be weak. We can also replace $x + x'$ by an arbitrary function of two variables (the same in both parts).

8 Generalised Feistel Ciphers

We can use similar tricks to ciphers similar to SHA or Skipjack, and give many other constructions of insecure ciphers based on the inverse in $GF(2^n)$. We give here an example. We will build a 64-bit block cipher. We divide our 64-bit state in 4 parts a, b, c, d and each round is as follows:

$$\begin{cases} b \leftarrow a \\ c \leftarrow b \\ d \leftarrow c \\ a \leftarrow d + K_i \cdot Inv(a + b + c) \end{cases}$$

Again it looks very good, mixes all kind of operations... and is very weak. It can be broken, not exactly by bi-linear cryptanalysis (BLC), but by a more general attack that can be called **Multi-Linear Cryptanalysis (MLC)**. For example, we consider the following expression in $GF(2^{16})$:

$$ab + ac + ad + bc + bd + cd$$

It is symmetric by any permutation of the 4 parts. After one round of encryption the same expression becomes:

$$ab + ac + bc + [d + K_i \cdot Inv(a + b + c)](a + b + c)$$

The difference between the two previous expressions is:

$$K_i \cdot Inv(a + b + c) \cdot (a + b + c)$$

which is equal to K_i with probability close to 1. Again, we can sum up these differences over the whole cipher and this allows to break our cipher given $\mathcal{O}(1)$ plaintexts for a large number of rounds up to about 2^{16}.

8.1 Higher Degree Multi-Linear Cryptanalysis

MLC is not limited to quadratic equations. It is easy to show the following result:

Theorem 8.1.1 (General Principle of Multi-Linear Cryptanalysis). Consider a cipher with d parts in $GF(2^n)$ in which each round transforms $(a_1, \ldots, a_d) \mapsto (a'_1, \ldots, a'_d) = (a_d \oplus F_K(a_1, \ldots, a_{d-1}), a_1, \ldots, a_{d-1})$ with an arbitrary function F. Let $P(a_1, \ldots, a_d)$ be an arbitrary d-linear function in $GF(2^n)$. Then for each round we have:

$$P(a_1, \ldots, a_d) - P(a'_2, \ldots, a'_d, a'_1) = P(a_1, a_2, \ldots, a_{d-1}, F_K(a_1, \ldots, a_{d-1}))$$

With this, we can construct multi-linear characteristics for an arbitrary number of rounds of a Feistel cipher. they will be composed of a common (the same for every round) d-linear part and some $(d-1)$-linear expressions that connect one round to another.

Simple Example: We leave the reader the pleasure to find that the following two ciphers are very easy to break by the new MLC attack:

$$\begin{cases} b \leftarrow a \\ c \leftarrow b \\ d \leftarrow c \\ a \leftarrow d + K_i \cdot Inv(abc) \end{cases} \qquad \begin{cases} b \leftarrow a \\ c \leftarrow b \\ d \leftarrow c \\ a \leftarrow d + K_i \cdot Inv(ab + bc + ac) \end{cases} .$$

Extensions: There are many extensions and generalisations possible. We expect that several real-life ciphers such as SHACAL or Skipjack should have interesting attacks of this type. However the number of possible MLC attacks is quite big. and systematic exploration of these attacks will not be obvious to achieve.

9 Conclusion

Proposing insecure ciphers with highly non-linear components may look as an exercise with no definite purpose. However each time we do so, we can usually formulate a general class of attacks that can potentially be applied to (more or less) any cipher. In this paper we introduced several new types of Generalised Linear Cryptanalysis. The universe of such attacks is unfortunately excessively rich, and remains largely unexplored. We show their interest by constructing various insecure ciphers. Their specific form is not trivial and is determined by the high level structure of the cipher. Locally, they exploit the existence if some non-linear multivariate relations that come from the inverse S-box. This demonstrates that such S-boxes can be dangerous and lead to devastating attacks.

In order to prevent such attacks, we advocate, following [11], to use S-boxes that have no such simple polynomial relations. In particular, for software encryption, we can afford to use reasonably large random S-boxes. This should prevent all known attacks on block ciphers: linear/differential cryptanalysis with generalisations, all kinds of attacks described in this paper, and also any kind of global algebraic attack such as XSL [11].

Acknowledgments: We would like to thank Louis Goubin and Lars Knudsen for helpful remarks and encouragements.

References

1. Frederik Armknecht, Matthias Krause: *Algebraic Atacks on Combiners with Memory,* Crypto 2003, LNCS 2729, pp. 162-176, Springer.
2. Kazuaro Aoki and Serge Vaudenay: *On the Use of GF-Inversion as a Cryptographic Primitive. SAC 2003, LNCS 3006, pp. 234-247, Springer 2004.*
3. *Ross Anderson, Eli Biham and Lars Knudsen: Serpent: A Proposal for the Advanced Encryption Standard.*
4. Anne Canteaut, Marion Videau: *Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis,* Eurocrypt 2002, LNCS 2332, Springer.
5. K.W. Campbell, M.J. Wiener: *Proof that DES is not a group.* Crypto'92, LNCS 740, pp. 512-520, Springer-Verlag, New York, 1993.
6. Nicolas Courtois, Adi Shamir, Jacques Patarin, Alexander Klimov, *Efficient Algorithms for solving Overdefined Systems of Multivariate Polynomial Equations*, In Advances in Cryptology, Eurocrypt'2000, LNCS 1807, Springer, pp. 392-407.
7. Nicolas Courtois: *Feistel Schemes and Bi-Linear Cryptanalysis,* in Crypto 2004, LNCS 3152, pp. 23-40, Springer, 2004.
8. Nicolas Courtois: *The security of Hidden Field Equations (HFE)*; Cryptographers' Track Rsa Conference 2001, LNCS 2020, Springer, pp. 266-281.
9. Nicolas Courtois: *Algebraic Attacks on Combiners with Memory and Several Outputs,* ICISC 2004, LNCS, to appear in Springer in early 2005. Extended version available on `http://eprint.iacr.org/2003/125/`.
10. Nicolas Courtois: *Fast Algebraic Attacks on Stream Ciphers with Linear Feedback,* Crypto 2003, LNCS 2729, pp: 177-194, Springer.
11. Nicolas Courtois and Josef Pieprzyk, *Cryptanalysis of Block Ciphers with Overdefined Systems of Equations,* Asiacrypt 2002, LNCS 2501, pp.267-287, Springer, a preprint with a different version of the attack is available at `http://eprint.iacr.org/2002/044/`.
12. Nicolas Courtois: *Higher Order Correlation Attacks, XL algorithm and Cryptanalysis of Toyocrypt,* ICISC 2002, LNCS 2587, pp. 182-199, Springer.
13. Nicolas Courtois and Willi Meier: *Algebraic Attacks on Stream Ciphers with Linear Feedback,* Eurocrypt 2003, Warsaw, Poland, LNCS 2656, pp. 345-359, Springer. An extended version is available at `http://www.minrank.org/toyolili.pdf`
14. Nicolas Courtois, Magnus Daum and Patrick Felke: *On the Security of HFE, HFEv- and Quartz,* PKC 2003, LNCS 2567, Springer, pp. 337-350. The extended version can be found at `http://eprint.iacr.org/2002/138/`.
15. Nicolas Courtois: *The Inverse S-box and Two Paradoxes of Whitening,* Long extended version of the Crypto 2004 rump session presentation, *Whitening The AES S-box,* Available at `http://www.minrank.org/invglc_rump_c04.zip`.
16. Joan Daemen, Vincent Rijmen: *AES proposal: Rijndael,* `http://csrc.nist.gov/encryption/aes/rijndael/Rijndael.pdf`
17. Joan Daemen, Vincent Rijmen: *The Design of Rijndael. AES - The Advanced Encryption Standard,* Springer-Verlag, Berlin 2002. ISBN 3-540-42580-2.
18. Joan Daemen, Vincent Rijmen, Bart Preneel, Anton Bosselaers, Erik De Win: *The Cipher SHARK,* FSE 1996, Springer.
19. Niels Ferguson, Richard Schroeppel and Doug Whiting: *A simple algebraic representation of Rijndael,* SAC 2001, page 103, LNCS 2259, Springer.
20. Lars Knudsen: *Block Ciphers - Analysis, Design and Applications,* PhD thesis, Aarhus University, Denmark, 1994.

21. C. Harpes, G. Kramer, and J. Massey: *A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-up Lemma,* Eurocrypt'95, LNCS 921, Springer, pp. 24-38. `http://www.isi.ee.ethz.ch/~harpes/GLClong.ps`
22. Thomas Jakobsen and Lars Knudsen: *Attacks on Block Ciphers of Low Algebraic Degree,* Journal of Cryptology 14(3): 197-210 (2001).
23. Thomas Jakobsen: *Higher-Order Cryptanalysis of Block Ciphers.* Ph.D. thesis, Dept. of Math., Technical University of Denmark, 1999.
24. Thomas Jakobsen: *Cryptanalysis of Block Ciphers with Probabilistic Non-Linear Relations of Low Degree,* Crypto 98, LNCS 1462, Springer, pp. 212-222, 1998.
25. Thomas Jakobsen, Lars R. Knudsen: *The Interpolation Attack on Block Ciphers,* FSE 97, LNCS 1267, Springer, pp. 28-40, 1997.
26. Antoine Joux, Jean-Charles Faugère: *Algebraic Cryptanalysis of Hidden Field Equation (HFE) Cryptosystems Using Gröbner Bases*, Crypto 2003, LNCS 2729, pp. 44-60, Springer, 2003.
27. R. Lidl, H. Niederreiter: *Finite Fields,* Encyclopedia of Mathematics and its applications, Volume 20, Cambridge University Press.
28. Alfred J. Menezes, Paul C. van Oorschot, Scott A. Vanstone: *Handbook of Applied Cryptography*; CRC Press, 1996.
29. S. Murphy, M. Robshaw: *Essential Algebraic Structure within the AES,* Crypto 2002, Springer.
30. Kaisa Nyberg: *Differentially Uniform Mappings for Cryptography,* Eurocrypt'93, LNCS 765, Springer, pp. 55-64.
31. Jacques Patarin: *Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): two new families of Asymm. Algorithms*, Eurocrypt'96, Springer, pp. 33-48.
32. Claude Elwood Shannon: *Communication theory of secrecy systems,* Bell System Technical Journal 28 (1949).
33. Wen-Ai Jackson and S. Murphy: *Projective Aspects of the AES Inversion,* Designs, Codes and Cryptography, Vol. 43, pp 167-179, 2007. Available also from `http://www.ma.rhul.ac.uk/static/techrep/2006/RHUL-MA-2006-4.pdf`.
34. Ralph Wernsdorf: *The Round Functions of RIJNDAEL Generate the Alternating Group,* Avialble from `Ralph.Wernsdorf@SIT.rohde-schwarz.com`
35. The Wikipedia entry "Möbius transformation", freely available at `http://en.wikipedia.org/wiki/Mobius_group`

A Proof of Theorem 4.3.1

In this section we give the proof of Theorem 4.3.1:

Theorem 4.3.1 (The Group Generated by *Inv* and XORs). The group generated by composing *Inv* and key additions is exactly the group of all permutations of $GF(2^n)$.

Proof: Let $K = GF(2^n)$ and let $a \in K - \{0\}$. The following equality holds for all $X \in \overline{K}$ and any ring/field of characteristic 2:

$$a^2 X = \frac{1}{1/a + \frac{1}{a + \frac{1}{1/a+X}}}$$

We propose to replace the inverse in $\overline{K}$ by the Rijndael inverse. We get the following function defined for $X \in K$ (a is a constant non-zero parameter):

$$A_a(X) = Inv\,(1/a \oplus Inv\,(a \oplus Inv\,(1/a \oplus X)))$$

Again, since Inv and $X \mapsto 1/X$ are almost always equal, this function must be equal to a^2X with overwhelming probability. By inspection we verify that:

$$\begin{cases} A_a(X) = a^2 X \quad \text{for} \quad X \notin \{0, 1/a\} \\ A_a(1/a) = 0 \\ A_a(0) = a \end{cases}$$

Example: when $a = 1$, we get a function $B = A_1$ that is equal to identity except that it swaps two points 0 and 1:

$$\begin{cases} A_1(X) = X \quad \text{for} \quad X \notin \{0, 1\} \\ A_1(1) = 0 \\ A_1(0) = 1 \end{cases}$$

We will construct more functions that exchange points. Let $a \notin \{0, 1\}$. We define:

$$C_a(X) = A_a(A_1(A_{\frac{1}{a}}(X)))$$

By inspection we verify that this function exchanges exactly two points a and a^2:

$$\begin{cases} C_a(X) = X \quad \text{for} \quad X \notin \{a, a^2\} \\ C_a(a) = a^2 \\ C_a(a^2) = a \end{cases}$$

For the next step, we observe that for any field $K = GF(2^n)$ the square is a permutation and $\sqrt{a}$ is well defined. We define the following function:

$$D_a(X) = A_{Inv(\sqrt{a})}(C_a(A_{\sqrt{a}}(X)))$$

By inspection we verify that for $a \notin \{0, 1\}$ this function exchanges exactly two points 1 and a:

$$\begin{cases} D_a(X) = X \quad \text{for} \quad X \notin \{1, a\} \\ D_a(1) = a \\ D_a(a) = 1 \end{cases}$$

We verified that it remains true also for $a = 0$ (otherwise we could use $B = A_1$ to exchange 0 and 1). Thus we can exchange 1 and any other point. Finally we can exchange any couple of points (a, b) as follows:

$$\begin{cases} E_{ab} = D_a \circ D_b \circ D_a \quad \text{for} \quad a \neq 1, b \neq 1 \\ E_{1a} = E_{a1} = D_a \quad \text{for} \quad a \neq 1. \end{cases}$$

The transformations E_{ab} generate the group of all permutations. □

B The Whitening Paradox

In this part we will describe a paradox that is a straightforward consequence of the present work. This part do not appear in the Springer version of this paper, only in the later extended version. It has also been outlined at the rump session of Crypto 2004 see [15].

Introduction. The starting question is the following: we built a cipher with Inv, XOR with some key, Inv etc. How many rounds are necessary to make it secure ? If we look at our proof of Theorem 4.3.1 above, we see that at most 45 rounds are necessary to exchange 2 elements a and b. Thus at most $45 \cdot 2^n$ rounds are necessary to obtain an arbitrary permutation of $GF(2^n)$. (And it cannot be less than $1/n \cdot log_2(2^n!) \approx 2^n$.) Thus with $45 \cdot 2^n$ rounds we achieve the best possible, information-theoretic, security: the system can be made to become a random permutation with uniform probability distribution.

Here starts the paradox: When we look at Theorem 3.0.1, our cipher should be a homographic function with probability that is a constant:

$$\left(1 - \frac{1}{2^n}\right)^{45 \cdot 2^n} \approx e^{-45}$$

This result assumes that all the whitening keys are random and independent, and then the fraction of about e^{-45} inputs does never encounter a singularity of Inv during all the $45 \cdot 2^n$ rounds. For this fraction there is a homographic approximation that can be recovered by Gaussian reduction from 3 plaintexts for which it holds. To find such 3 plaintexts by exhaustive search we need on average about $e^{3 \cdot 45}$ tries, which is still a constant. (We assume here that n is sufficiently large).

In the construction however, $45 \cdot 2^n$ rounds are sufficient to cover up the whole input space with points that will always encounter a singularity at some round, and there is no homographic approximation. The encryption function can be just anything (an arbitrary permutation of $GF(2^n)$).

Due to this construction, we call the Inv S-box a “cryptographic black hole”: it can absorb astronomical quantities of whitening and remain pitch black, (i.e. insecure). However if we do the whitening in a specific way, we can succeed to achieve “lily-white”: the highest level of security possible for a block cipher.

We have a cipher which is extremely insecure for random of pseudo-random choice of internal key material, but is extremely secure for a special choice of key material. We call it “whitening paradox”.