
# Algebraic Attacks on Block Ciphers

Martin Albrecht (M.R.Albrecht@rhul.ac.uk)

Information Security Group,
Royal Holloway, University of London

October 28, 2008

# Outline

# Notation


- $P = \mathbb{K}[x_0, \ldots, x_{n-1}]$; $\mathbb{K}$ a field
- $I$ is an ideal $\subset P$. $f_0, f_1 \in I \rightarrow f_0 + f_1 \in I$; $f_2 \in P \rightarrow f_2 \cdot f_0 \in I$
- $\langle f_0, \ldots, f_{m-1} \rangle$ is the ideal spanned by $f_0, \ldots, f_{m-1}$.
- $V(f_0, \ldots, f_{m-1}) = \{(a_0, \ldots, a_{n-1}) \in \overline{\mathbb{K}}^n : f_i(a_0, \ldots, a_{n-1}) = 0 \text{ for all } 0 \leq i < m\}$.
- $V(I)$ is the variety of $I$. Especially, if $I = \langle f_0, \ldots, f_{m-1} \rangle$, then $V(I) = V(f_0, \ldots, f_{m-1})$ (Hilbert's Basis Theorem)
- In cryptanalysis often $\mathbb{K} = \mathbb{F}_2$ and we're looking for solutions in $\mathbb{K}$ not $\overline{\mathbb{K}}$, so we add $x_i^2 + x_i$ to the ideal basis.

# The $\mathcal{MQ}$ Problem


Given a set $\{f_0, \ldots, f_{m-1}\}$, compute a "simpler" set $\{g_0, \ldots, g_{M-1}\}$, such that

$$\langle f_0, \ldots, f_{m-1} \rangle = \langle g_0, \ldots, g_{M-1} \rangle.$$

and consequently

$$V(f_0, \ldots, f_{m-1}) = V(g_0, \ldots, g_{M-1}).$$

**AES** (128-bit, 10 rounds) 8000 equations in 1600 variables over $\mathbb{F}_2$ [CP02] or 5248 equations in 3968 variables over $\mathbb{F}_{2^8}$ [MR02].

**Present** (80-bit, 31 rounds) 34742 equations in 8140 variables over $\mathbb{F}_2$ [AC08].

# Monomial Orderings


Monomials can be ordered in different ways, i.e. there is no canonical ordering.
Denote $\mathbf{x}^a = \prod x_i^{a_i}$ for $0 \leq i < n$.
Important term odering examples are:

**lex** $\mathbf{x}^a < \mathbf{x}^b \Leftrightarrow \exists\ 0 \leq i < n : a_0 = b_0, \ldots, a_{i-1} = b_{i-1}, a_i < b_i$

**degrevlex** Let $deg(\mathbf{x}^a) = a_0 + \cdots + a_{n-1}$, then
$\mathbf{x}^a < \mathbf{x}^b \Leftrightarrow deg(\mathbf{x}^a) < deg(\mathbf{x}^b)$ or $deg(\mathbf{x}^a) = deg(\mathbf{x}^b)$ and
$\exists\ 0 \leq i < n : a_{n-1} = b_{n-1}, \ldots, a_{i+1} = b_{i+1}, a_i > b_i$.

Leading monomials etc. are always considered with respect to some monomial ordering.
An monomial ordering is called admissable if it respects multiplication.

# (Reduced) Gröbner Bases


## Definition (Gröbner Basis)

Fix a monomial order. A finite subset $G = \{g_0, \ldots, g_{m-1}\}$ of an ideal $I$ is said to be a **Gröbner basis** or standard basis if

$$\langle LT(g_0), \ldots, LT(g_{m-1})\rangle = \langle LT(I)\rangle.$$

## Definition (Reduced Gröbner Basis)

A **reduced Gröbner basis** for a polynomial ideal $I$ is a Gröbner basis for $G$ such that:

1. $LC(f) = 1$ for all $f \in G$;
2. For all $f \in G$, no monomial of $f$ lies in $\langle LT(G - \{f\})\rangle$ .

# Example


Consider: $k_{0,0} + k_{1,2}$, $k_{0,2} + k_{1,1}$, $k_{0,1} + k_{1,0}$, $1 + k_{1,2} + z_{1,2}$, $k_{1,1} + z_{1,1}$, $k_{1,0} + z_{1,0}$, $z_{1,1} + y_{1,2} + y_{1,1}$, $z_{12} + y_{10}$, $z_{10} + y_{11} + y_{10}$, $k_{02} + x_{12}$, $1 + k_{01} + x_{11}$, $k_{00} + x_{10}$, $y_{12} + y_{10}y_{11} + x_{10}$, $y_{12} + y_{11} + y_{11}y_{12} + y_{10} + x_{12} + x_{10}$, $y_{12} + y_{10} + x_{12} + x_{12}y_{11} + x_{10}$, $1 + y_{12} + y_{11} + y_{10}y_{12} + x_{11} + x_{10}$, $1 + y_{12} + y_{11} + y_{10} + x_{11} + x_{11}y_{10} + x_{10}$, $1 + y_{12} + y_{11} + y_{10} + x_{11} + x_{11}x_{12} + x_{10}$, $y_{12} + y_{10} + x_{12}y_{10} + x_{10}y_{12}$, $1 + y_{10} + x_{12} + x_{11} + x_{11}y_{12} + x_{10}y_{12}$, $1 + y_{11} + x_{12}y_{12} + x_{11} + x_{10} + x_{10}y_{12}$, $x_{11}y_{11} + x_{10} + x_{10}y_{12}$, $y_{11} + y_{10} + x_{12} + x_{10}y_{11}$, $1 + y_{11} + x_{11} + x_{10}y_{10}$, $1 + y_{11} + x_{11} + x_{10}x_{12}$, $1 + y_{10} + x_{12} + x_{11} + x_{10} + x_{10}x_{11}$

The reduced Gröbner basis w.r.t. $>_{lex}$:

$k_{0,2}$, $1 + k_{0,1}$, $1 + k_{0,0}$, $1 + k_{1,2}$, $k_{1,1}$, $1 + k_{1,0}$, $z_{1,2}$, $z_{1,1}$, $1 + z_{1,0}$, $1 + y_{1,2}$, $1 + y_{1,1}$, $y_{1,0}$, $x_{1,2}$, $x_{1,1}$, $1 + x_{1,0}$

# More Notation


- An ideal is zero-dimensional if $V(I)$ is finite.
- The radical of $I$ denoted by $\sqrt{I}$, is the set $\{f : f^e \in I \text{ for some integer } e \geq 1\}$.
- A perfect field is a field of characteristic $p$ where every element has a $p - th$ root or the characteristic is zero.
- An elimination ideal is defined as: Given $I = \langle f_0, \ldots, f_{m-1} \rangle \subset \mathbb{K}[x_0, \ldots, x_{n-1}]$, the $l$-th elimination ideal $I_l$ is the ideal of $\mathbb{K}[x_{l+1}, \ldots, x_{n-1}]$ defined by $I_l = I \cap \mathbb{K}[x_{l+1}, \ldots, x_{n-1}]$.

# The Shape Lemma


## Theorem (The Shape Lemma)

*Let $\mathbb{K}$ be a perfect field, let $I \subset P$ be a zero-dimensional radical ideal. Let $g_{n-1} \in \mathbb{K}[x_{n-1}]$ be the monic generator of the elimination ideal $I \cap \mathbb{K}[x_{n-1}]$, and let $d = deg(g_{n-1})$. Then the following statements are true:*

1. *The reduced Gröbner basis of the ideal $I$ with respect to the lexicographic ordering $x_0 > \cdots > x_{n-1}$ is of the form*

$$\{x_0 - g_0, \ldots, x_{n-2} - g_{n-2}, g_{n-1}\},$$

   *where $g_0, \ldots, g_{n-2} \in \mathbb{K}[x_{n-1}]$;*
2. *The polynomial $g_{n-1}$ has $d$ distinct zeros $a_0, \ldots, a_{d-1} \in k$, and the set of zeros of $I$ is $\{(g_0(a_i), \ldots, g_{n-2}(a_i), a_i) : i = 0, \ldots, d-1\}$.*

# Shape Lemma Example


Consider $P = \mathbb{F}_7[a, b, c, d]$, $<_{lex}$ and Cyclic-3:

$$I = \langle a + 2b + 2c + 2d + 6, a^2 - a + 2b^2 + 2c^2 + 2d^2, \\ 2ab + 2bc - b + 2cd, 2ac + b^2 + 2bd - c \rangle$$

$\mathbb{F}_7$ is perfect, $I$ ideal is zero-dimensional and radical. The reduced Gröbner basis is:

$$gb = (a + 5d^6 + d^5 + 5d^4 + 3d^3 + 3d^2 + 5d + 6, \\ b + 4d^6 + 2d^4 + d^3 + 2d^2 + 4d, \\ c + 4d^6 + 3d^5 - d^4 + d^3 + 5d, \\ d^7 + 3d^6 - d^5 + 3d^4 + d^3 - d^2 + 3d)$$

$g_{n-1}$ factors as $d \cdot (d + 2) \cdot (d + 3) \cdot (d^4 + 5d^3 + 3d^2 + 4)$.

# Enforcing the Shape Lemma

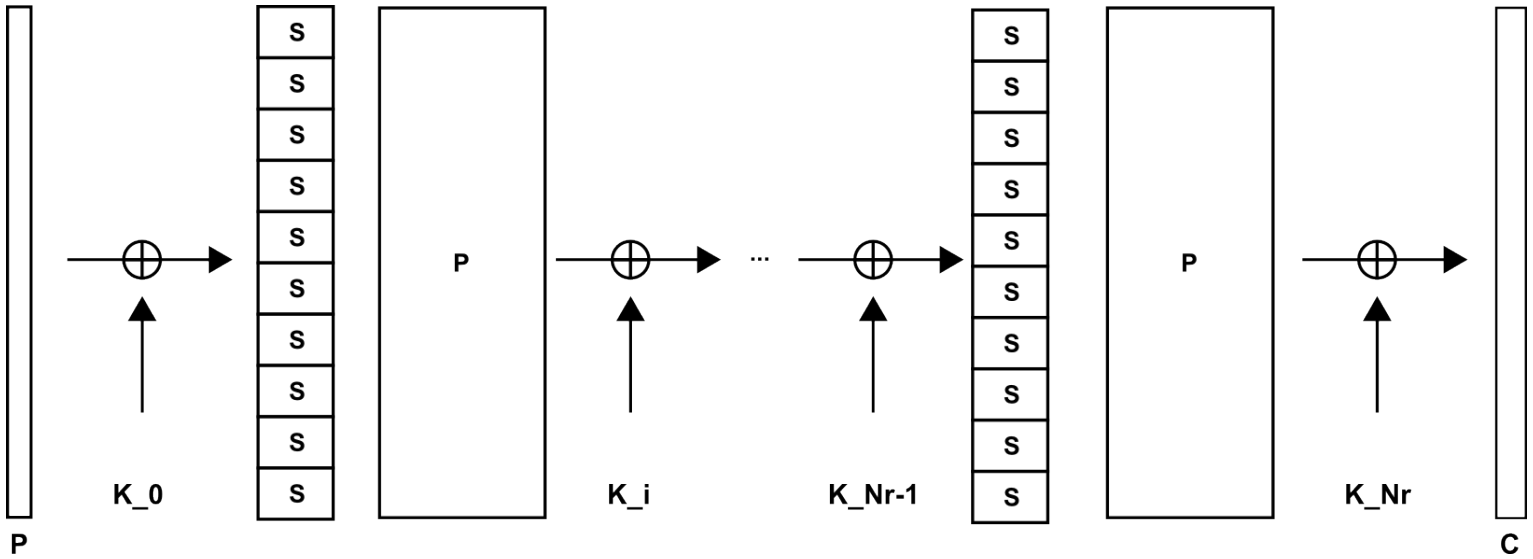

To bring an ideal over $\mathbb{F}_p$ in the form such that the shape lemma applies, add the "field polynomials" ($\{x_i^p - x_i\}$ for every $0 \leq i < n$) to the ideal. This makes sure, that:

- solutions from the algebraic closure are excluded as $x_i^p - x_i$ factors completely over $\mathbb{F}_p$,
- the ideal is zero-dimensional (implied by above statement),
- the ideal is a radical ideal as $\mathrm{GCD}(\frac{d(x_i^p - x_i)}{dx_i}, x_i^p - x_i) = 1$ (Seidenberg's Lemma).

So we can solve systems of equations over $\mathbb{F}_p$ using Gröbner bases.

# Outline

# SP-Networks


$P$ is linear, $S$ is the only non-linear component, $K_i$ are subkeys derived from the user supplied key $K$.

# S-Box Equations I


Consider the S-Box (permutation)

$$[7, 6, 0, 4, 2, 5, 1, 3]$$

as an example.

Construct the matrix on the right and perform fraction-free Gaussian elimination on it (fitting a linear model).

$$\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & x_0 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & x_1 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & x_2 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & y_0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & y_1 \\
1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & y_2 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & x_0x_1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & x_0x_2 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & x_0y_0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & x_0y_1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & x_0y_2 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & x_1x_2 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & x_1y_0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & x_1y_1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & x_1y_2 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & x_2y_0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & x_2y_1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & x_2y_2 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & y_0y_1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & y_0y_2 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & y_1y_2
\end{pmatrix}$$

# S-Box Equations II


$$\left(\begin{array}{cccccccc|l}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_0y_0 + x_1 + x_2 + y_0 + y_1 + 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & x_0y_0 + x_0 + x_1 + y_2 + 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & x_0y_0 + x_0 + y_0 + 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & x_0y_0 + x_0 + x_2 + y_1 + y_2 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & x_0y_0 + x_0 + x_1 + x_2 + y_0 + y_1 + y_2 + 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & x_0y_0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & x_0y_0 + x_2 + y_0 + y_2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & x_0y_0 + x_1 + y_1 + 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_0x_2 + x_1 + y_1 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_0x_1 + x_1 + x_2 + y_0 + y_1 + y_2 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_0y_1 + x_0 + x_2 + y_0 + y_2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_0y_0 + x_0y_2 + x_1 + x_2 + y_0 + y_1 + y_2 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_1x_2 + x_0 + x_1 + x_2 + y_2 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_0y_0 + x_1y_0 + x_0 + x_2 + y_1 + y_2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_0y_0 + x_1y_1 + x_1 + y_1 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_1y_2 + x_1 + x_2 + y_0 + y_1 + y_2 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_0y_0 + x_2y_0 + x_1 + x_2 + y_1 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_2y_1 + x_0 + y_1 + y_2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{x_2y_2 + x_1 + y_1 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{y_0y_1 + x_0 + x_2 + y_0 + y_1 + y_2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{y_0y_2 + x_1 + x_2 + y_0 + y_1 + 1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{y_1y_2 + x_2 + y_0}
\end{array}\right)$$

# Cipher Equations


- Define subkey variables for the subkey bits used in each round.
- Define "state" variables for S-Box input and output bits.
- Diffusion layer is linear in these variables and the subkey variables.
- Define equations for key schedule analogously.

There are many ways the above can be done. How it is done most effectively, is an open research problem.

# Outline

# Gröbner Basis Construction


Recall the definition of a Gröbner basis. It is a set $G$ of polynomials $g_0, \ldots g_{m-1}$ such that:

$$\langle LT(g_0), \ldots, LT(g_{m-1}) \rangle = \langle LT(I) \rangle.$$

Now, try to create elements in $\langle LT(I) \rangle$ and not in $\langle LT(g_0), \ldots, LT(g_{m-1}) \rangle$. If you **find** such an element **add it** to the basis. If such an element provably cannot be constructed $G$ is a Gröbner basis. This procedure **terminates** as the ideals of leading terms created this way are strictly increasing and such a sequence **must stabilise** eventually due to the **Ascending Chain Condition**. At this point

$$\langle LT(g_0), \ldots, LT(g_{m-1}) \rangle = \langle LT(I) \rangle$$

.

# S-Polynomials


Bruno Buchberger [Buc65] showed that every cancellation of leading terms may be accounted to **S-polynomials**.

## Definition (S-Polynomial)

Let $f, g \in \mathbb{K}[x_1, \ldots, x_n]$ be polynomials $\neq 0$ and define $x^\gamma = \mathrm{LCM}(\mathrm{LM}(f), \mathrm{LM}(g))$. Then the S-polynomial of $f$ and $g$ is defined as

$$S(f,g) = \frac{x^\gamma}{\mathrm{LT}(f)} \cdot f - \frac{x^\gamma}{\mathrm{LT}(g)} \cdot g.$$

# Polynomial Reduction


The S-polynomial $h$ of $g_i, g_j$ is not in $\langle \mathrm{LT}(g_i), \mathrm{LT}(g_j) \rangle$ but it is in $< \mathrm{LT}(I) >$. It may be in $G_r = \langle \{g_k | k \neq i, j\} \rangle$.

## Definition

Let $G = \{g_0, \ldots, g_{m-1}\} \subset P$ . Given a polynomial $h \in P$ , we say that $h$ reduces to zero modulo G, written

$$h \xrightarrow[G]{} 0,$$

if $h$ can be written in the form

$$h = a_0 g_0 + \cdots + a_{m-1} g_{m-1},$$

such that whenever $a_i g_i \neq 0$, we have $h \geq a_i g_i$.

# Buchberger's Algorithm


```
def buchberger(F):
    G = set(F)
    B = set([(g1,g2) for g1 in G for g2 in G if g1!=g2])

    while B!=set():
        g1,g2 = select(B)
        B.remove( (g1,g2) )

        h = spol(g1,g2).reduce(G)
        if h != 0: #reductions to zero are useless!
            B = B.union( [(g,h) for g in G] )
            G.add( h )
    return G
```

# Performance Considerations I


- the intermediate basis grows pretty quickly;
- the major bottleneck is reduction modulo $G$;
- need strategy which S-polynomial to construct in which order;
- need criteria which S-polynomial reduces to zero to avoid reduction in that case;
- computing with respect to **lex** takes much longer than e.g. **degrevlex** (cf. [FGLM93]);
- runtime is double exponential in the worst case. solving polynomial equation systems in $\mathrm{NP}$-hard on average;
- we often don't know a priori the actual running time of Buchberger's algorithm applied to a given ideal basis.

# Performance Considerations II


Situation could be better for algebraic attacks, because

- the ideals are zero-dimensional (one solution),
- the systems are often sparse yet overdefined,
- we are working over simple fields like $\mathbb{F}_2$ and
- the systems are highly structured.

Several improvements and specialisations to Buchberger's algorithm exist, for example

- the Gebauer-Möller installation [GM88],
- the $F_4$ [Fau99] and $F_5$ [Fau02] algorithms by Jean-Charles Faugère and
- the SlimGB [Bri05] algorithm by Michael Brickenstein.

# Is This Any Good?


This is an emerging topic, but

- HFE and other multivariate schemes were broken using Gröbner bases, [FJ03],
- some stream cipher constructions were broken using algebraic techniques, [CM03],
- while the direct approach against block ciphers usually fails after a few rounds [CMR06, CB07, BD07]
- we can combine other cryptographic techniques with algebra with promising results [AC08, CBW08, FP08].

# Questions?


**Thank You!**

# References I



Martin Albrecht and Carlos Cid.

Algebraic Techniques in Differential Cryptanalysis.
Cryptology ePrint Archive, Report 2008/177, 2008.
Available at http://eprint.iacr.org/2008/177.pdf.

Michael Brickenstein and Alexander Dreyer.

PolyBoRi: A framework for Gröbner basis computations with Boolean polynomials.
In *Electronic Proceedings of MEGA 2007*, 2007.
Available at http://www.ricam.oeaw.ac.at/mega2007/electronic/26.pdf.

Michael Brickenstein.

Slimgb: Gröbner Bases with Slim Polynomials.
In *Reports On Computer Algebra 35*. Centre for Computer Algebra, University of Kaiserslautern, 2005.
Available at: http://www.mathematik.uni-kl.de/~zca/Reports_on_ca/35/paper_35_full.ps.gz.

Bruno Buchberger.

*Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal*.
PhD thesis, Universität Innsbruck, 1965.

Nicolas T. Courtois and Gregory V. Bard.

Algebraic Cryptanalysis of the Data Encryption Standard.
In Steven D. Galbraith, editor, *Cryptography and Coding – 11th IMA International Conference*, volume 4887 of *Lecture Notes in Computer Science*, pages 152–169, Berlin Heidelberg New York, 2007. Springer Verlag.
Available at http://eprint.iacr.org/2006/402.

Nicolas T. Courtois, Gregory V. Bard, and David Wagner.

Algebraic and slide attacks on KeeLoq.
In *Proceedings of FSE 2008*, 2008.

# References II



Nicolas T. Courtois and Willi Meier.

Algebraic attacks on stream ciphers with linear feedback.
In *Advances in Cryptology — EUROCRYPT 2003*, volume 2656 of *Lecture Notes in Computer Science*, pages 345–359. Springer Verlag, 2003.

Carlos Cid, Sean Murphy, and Matthew Robshaw.

*Algebraic Aspects of the Advanced Encryption Standard*.
Springer Verlag, 2006.

Nicolas T. Courtois and Josef Pieprzyk.

Cryptanalysis of Block Ciphers with Overdefined Systems of Equations.
In *Advances in Cryptology — ASIACRYPT 2002*, volume 2501 of *Lecture Notes in Computer Science*, pages 267–287. Springer Verlag, 2002.

Jean-Charles Faugère.

A New Efficient algorithm for Computing Gröbner Basis (F4), 1999.
Available at `http://modular.ucsd.edu/129-05/refs/faugere_f4.pdf`.

Jean-Charles Faugère.

A New Efficient Algorithm for Computing Gröbner Bases without Reduction to Zero (F5).
In *Proceedings of ISSAC*, pages 75–83. ACM Press, 2002.

Jean-Charles Faugère, Patrizia M. Gianni, Daniel Lazard, and Teo Mora.

Efficient Computation of Zero-Dimensional Gröbner Bases by Change of Ordering.
In *Journal of Symbolic Computation 16*, pages 329–344. Academic Press, 1993.

Jean-Charles Faugère and Antoine Joux.

Algebraic cryptanalysis of hidden field equation (HFE) cryptosystems using Gröbner bases.
In *Advances in Cryptology - CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*. Springer-Verlag, 2003.

# References III



Jean-Charles Faugère and Ludovic Perret.

Algebraic cryptanalysis of curry and flurry using correlated messages.
Cryptology ePrint Archive, Report 2008/402, 2008.
available at `http://eprint.iacr.org/2008/402.pdf`.

R. Gebauer and H.M. Möller.

On an Installation of Buchberger's Algorithm.
In *Journal of Symbolic Computation 6 (2 and 3)*, pages 275–286. Academic Press, 1988.

Sean Murphy and Matthew Robshaw.

Essential Algebraic Structure Within the AES.
In M. Yung, editor, *Advances in Cryptology — CRYPTO 2002*, volume 2442 of *Lecture Notes in Computer Science*, pages 1–16. Springer Verlag, 2002.
Available at `http://www.isg.rhul.ac.uk/~mrobshaw/rijndael/aes-crypto.pdf`.