AOC

Why the Degree of Regularity Alone is Bad for Estimating Security – a Counter Example to Common Arguments

by Jan Ferdinand Sauer May 18, 2021 0 Comment

Cryptographic primitives designed to be algebraically simple – AOCs – might be particularly vulnerable to algebraic attacks. One of the most threatening attack vectors in this category is the Gröbner basis analysis. For a cipher or hash function to be considered secure, the Gröbner basis for any polynomial system derivable from the primitive needs to be intractable to compute.

Unfortunately, the complexity of computing a Gröbner basis for a specific polynomial system is generally not known before the computation is completed. However, some complexity bounds exist. One of the most prominently used bounds is based on a polynomial system's degree of regularity.

Generally, computing the degree of regularity for a polynomial system is as hard as computing the Gröbner basis itself. Luckily, for an "average" regular determined system, the degree of regularity equals the *Macaulay bound*. That is, for $\mathcal{F} = \{f_0, \ldots, f_{s-1}\} \subseteq \mathbb{F}[x_0, \ldots, x_{n-1}]$ we have $d_{\text{reg}} = 1 + \sum_{i=0}^{s-1} \deg(f_i) - 1$.

How Current AOCs Argue Resistance to Gröbner Basis Analysis

The **Poseidon** [6] paper mentions the Macaulay bound, and implicitly assumes that the polynomial system arising from Poseidon is a regular sequence. My own experiments indicate that this assumption is *false*. Similarly, **GMiMC** [1] uses the Macaulay bound and assumes the regularity of the system implicitly. My own experiments indicate that this assumption is also *false*. The authors of **Ciminion** [4] explicitly assume the derived system to be regular, but mistakenly describe this to be "the best adversarial scenario" where in fact the opposite is true. Furthermore, my own experiments indicate that the polynomial sequence is *not* regular. For **Rescue** [2], the authors perform Gröbner basis attacks on round-reduced variants, showing that the system arising from Rescue is *not* regular. They then extrapolate the observed degrees to estimate the degree of regularity for the full-round primitive.

In summary, two approaches can be observed: (1) assume regularity of the system, then use the Macaulay bound to compute the degree of regularity, or (2) extrapolate the degree of regularity from round-reduced variants. Both approaches then use the degree of regularity to estimate the complexity for computing the Gröbner basis. This is generally done by looking at the complexity bound of the most efficient Gröbner basis algorithm, F_5. This bound is

$$O\left(\binom{n + d_{\text{reg}}}{n}^{\omega}\right)$$

where n is the number of variables in the polynomial ring [3].

But: this is an upper bound. We need a *lower* bound.

The Degree of Regularity does not Suffice

I'll make a series of increasingly complex and decreasingly pathological examples why the degree of regularity derived from the Macaulay bound does not suffice to accurately estimate the *concrete* complexity of computing a Gröbner basis. The ideals of all the systems below are of dimension 0, meaning that the respective sets of common solutions are non-empty and contain finitely many elements. This accurately reflects the properties of polynomial systems modeling a cryptographic primitive.

The system is already a Gröbner basis

Let's say we want to compute the Gröbner basis for $\mathcal{F}_{\text{gb}} = \{x^7, y^7, z^7\} \subseteq \mathbb{F}[x, y, z]$. We quickly see that $\mathcal{F}$ is a regular sequence, and determine that the degree of regularity is $d_{\text{reg}} = 1 + \sum_{i=0}^{2} 7 - 1 = 19$. Consequently, or so the roughly sketched argument above goes, a Gröbner basis algorithm like F_4 or F_5 should have to perform computations on polynomials of up to degree 19 before being able to output a Gröbner basis.

However, $\mathcal{F}_{\text{gb}}$ is already a Gröbner basis – no computation at all is required!

The system can be split up

Deriving a polynomial system from a cryptographic primitive rarely gives you a Gröbner basis – although there are exceptions, like GMiMC. Instead, let's look at the following polynomial system.

$$\mathcal{F}_{\text{indep}} = \left\{ \begin{matrix} u^2vw + u^2, & x^2yz + x^2, \\ uv^2w + v^2 + 1, & xy^2z + y^2 + 1, \\ uvw^2 + w^2, & xyz^2 + z^2 \end{matrix} \right\} \subseteq \mathbb{F}[u, v, w, x, y, z].$$

The polynomials containing variables u, v and w are completely independent from the polynomials where x, y, and z make an appearance. For the Macaulay bound, this fact is irrelevant. Since $\mathcal{F}_{\text{indep}}$ is a regular sequence, we might derive $d_{\text{reg}} = 1 + \sum_{i=0}^{5} 4\text{–}1 = 19$.

However, the F_4 implementations of magma and FGb as well as the python implementation of F_5 all compute on polynomials of only degree 5 and lower before finding the Gröbner basis – they are not fooled by this attempt to artificially increase the complexity.

The system is not very "involved"

When deriving a polynomial system from a (single) cryptographic primitive, a partition in the set of polynomials like above is unlikely to appear – intuitively, that would lead to weak diffusion. Let's change the system a little, then.

$$\mathcal{F}_{\text{invlv}} = \left\{ \begin{matrix} u^2vw + u^2, & x^2yz + x^2, \\ uv^2w + v^2 + 1, & xy^2z + y^2 + 1, \\ uvw^2 + w^2, & u^4 + z^4 \end{matrix} \right\} \subseteq \mathbb{F}[u, v, w, x, y, z].$$

The sets $\mathcal{F}_{\text{indep}}$ and $\mathcal{F}_{\text{invlv}}$ differ in one polynomial, and this polynomial $(u^4 + z^4) = f_{\text{link}}$ links the two independent subsets of $\mathcal{F}_{\text{indep}}$. I didn't derive the system from any concrete primitive, but a polynomial like f_{link} might express how to move from one round to the next in a cipher.

The Macaulay bound for $\mathcal{F}_{\text{invlv}}$ does not change from the bound for $\mathcal{F}_{\text{indep}}$ since f_{link} is of the same degree as the polynomial it replaced. Also, $\mathcal{F}_{\text{invlv}}$ is still a regular sequence, so we still have $d_{\text{reg}} = 19$.

You might have guessed it by now: the highest polynomials appearing during a Gröbner basis computation for $\mathcal{F}_{\text{invlv}}$ is not 19. Magma's F_4 reports a maximum degree of 6, FGb only reaches degree 5, and so does python-F_5.

While I don't fully understand why this happens, vectors of origin give some hints. Briefly, v_i is a vector of origin for Gröbner basis element g_i if $\mathcal{F}_{\text{invlv}} \cdot v_i = g_i$. Below are the vectors of origin for $\mathcal{F}_{\text{invlv}}$, where any big polynomial is replaced by $\bullet$ to ease reading.

$$\begin{matrix} (\bullet, \bullet, 0, 0, 0, 0), \\ (\bullet, \bullet, 0, 0, 0, 0), \\ (0, 0, \bullet, 3, 0, 0), \\ (0, 0, \bullet, \bullet, 0, 0), \\ (0, 0, \bullet, \bullet, 1, 0), \\ (0, 0, \bullet, \bullet, \bullet, 1) \end{matrix}$$

A zero in position i in a vector of origin means that f_i was unnecessary for computing the Gröbner basis element. Above vectors of origin have a lot of zeros – in fact, even though all polynomials are linked to one another in some (potentially indirect) way, there seems to be a partition.

I describe polynomial systems for which the Gröbner bases' elements can be computed from a few input polynomials at a time as having low "involvement." As of yet, there is no mathematically rigourous way to define this notion, but above example should give a rough intuition. My observations indicate that low involvement means low complexity for computing a Gröbner basis.

Note. Above counter-examples do not disprove the equality of the degree of regularity and the Macaulay bound for *generic* polynomial systems – they only show that regularity of the sequence is not a sufficient requirement.

Existing Lower Bounds

The main message of this post is that we need (tight-ish) *lower*, not upper, bounds for estimating the complexity of a Gröbner basis computation in order to accurately asses the security of cryptographic primitives against this vector of attack. Unfortunately, the scientific literature currently has little to offer in this regard.

Hyun [5] exclusively deals with field $\mathbb{Q}$, while we are interested in finite fields, and Möller & Mora [7] look at ideals of positive dimension, while we are only interested in zero-dimensional ideals. Furthermore, all given bounds are *existential* while we need a *constructive* bound.

In summary, current strategies for arguing that some Arithmetization Oriented Primitive is resistant against Gröbner basis attacks make too many unbacked assumptions, often implicitly. The tools to make these arguments rigorously don't currently exist. Or in other words: "look at me still talking when there's science to do."

References

1. Albrecht, M.R., Grassi, L., Perrin, L., Ramacher, S., Rechberger, C., Rotaru, D.,Roy, A., Schofnegger, M.: Feistel Structures for MPC, and More. In: ESORICS. pp.151–171. Springer (2019)
2. Aly, A., Ashur, T., Ben-Sasson, E., Dhooghe, S., Szepieniec, A.: Design of Symmetric Primitives for Advanced Cryptographic Protocols. IACR ToSC 2020(3), 1–45(2020)
3. Bardet, M., Faugère, J.C., Salvy, B.: On the complexity of the F5 Gröbner basis algorithm. Journal of Symbolic Computation 70, 49–70 (2015)
4. Dobraunig, C.E., Grassi, L., Guinet, A., Kuijsters, D.: Ciminion: Symmetric Encryption Based on Toffoli-Gates over Large Finite Fields. In: Eurocrypt 2021 (2021)
5. Huynh, Dung T.: A superexponential lower bound for Gröbner bases and Church-Rosser commutative Thue systems. Information and Control, 68(1-3):196–206 (1986)
6. Grassi, L., Khovratovich, D., Rechberger, C., Roy, A., Schofnegger, M.: Poseidon: A New Hash Function for Zero-Knowledge Proof Systems. In: USENIX Security. USENIXAssociation (2020)
7. Möller, H. M., and Mora, F.: Upper and lower bounds for the degree of Gröbner bases. In: International Symposium on Symbolic and Algebraic Manipulation, pages 172–183. Springer (1984)

Jan Ferdinand Sauer

Website: https://asdm.gmbh

Leave a Reply

Logged in as Jan Ferdinand Sauer Logout

Comment

Submit