Solving Binary Linear Equation Systems over the Rationals and Binaries

Benedikt Driessen and Christof Paar

Horst-Goertz Institute for IT Security,
Ruhr-University Bochum, Germany
{benedikt.driessen,christof.paar}@rub.de

Abstract. This paper presents intermediate results of our investigations into the potential of analog hardware for the purpose of solving linear equation (LES) systems which are of quadratic form and binary. Based on the assumption that we can efficiently solve binary LES over the rationals with sufficient precision, we present a generic method to map a rational solution to a solution which solves the equation system over $\mathbb{F}_2$. We show that, in order to perform this mapping, we only need to look at two bits of the binary expansion of each of the elements of the rational solution vector.

1 Introduction

Solving binary linear equation systems (LES) of the form $A\boldsymbol{x} = \boldsymbol{b}$ with $A \in \mathbb{F}_2^{n\times n}, \boldsymbol{b}, \boldsymbol{x} \in \mathbb{F}_2^n$ and n unknowns is a common problem and appears in numerous research and technical disciplines. In the field of cryptography, a special form of this problem arises when attacking stream ciphers. Certain attacks, such as attacks on A5/1 and A5/2 in the extremely wide-spread GSM-standard [PFS00,BBK03,GNR08,Gol97,PS00] require solving of a very large number of LES over $\mathbb{F}_2$.

These LES can be solved with the help of Gaussian elimination, which can easily be implemented in soft- and hardware. However, for some implementations this approach is unsatisfying in practice due to its cubic complexity. In spirit of unconventional cryptanalytical computing devices such as TWIRL [Sha99,LS00] and TWINKLE [ST03] we have explored the potential of a hypothetical device which solves a particular LES over $\mathbb{Q}$.

Based on the assumption that we are given a solution to $A\boldsymbol{u} = \boldsymbol{b}$ with $A \in \mathbb{F}_2^{n\times n}, \boldsymbol{b} \in \mathbb{F}_2^n$ and $\boldsymbol{u} \in \mathbb{Q}$, we have developed a method to convert the rational solution which allows us to solve quadratic LES over $\mathbb{F}_2$. We explicitly stress that the presented method is generic and not dependent on the conceptualized device, which is not in the focus of this paper.

2 Background

In the following we will shortly elaborate on what inspired our overall work and how this resulted in this paper. The computing device TWINKLE (and TWIRL) is used for the sieving step of the Number Field Sieve (NFS) [LLMP93] and – in combination with "classical hardware" – assumed to be able to factor 512-bit RSA moduli. Although the device has never been built, it is supposedly possible to do so for $5000.

The idea of the hypothetical devices TWIRL and TWINKLE is to shift the most expensive part of the NFS method from digital hardware into the analog domain. More specifically, finding appropriate smooth numbers for the sieving step is done with the help of an array of LEDs and a light sensor. In TWINKLE, the LEDs, which are switched on and off in a specific manner, are emitting light proportional to the logarithms of successive primes numbers. The sensor operates as instantaneous adder of these intensities and signals when a certain threshold has been reached. The LEDs switched on in this moment, together with the product of their associated primes, indicate a smooth number.

The advantage of TWINKLE/TWIRL is that they exploit a physical property to offload computation, which inspired us to experiment with an electrical device we call the Analog Solver. The core idea of our device is to use a network of switched Operational AMPlifiers (OPAMPs) to solve binary linear equation systems. The basic principle of our circuit is given by the observation that OPAMPs can be used as inverting adders for input voltages u_i, i.e.,

$$u_{\texttt{out}} = -R_{\texttt{add}} \left(\frac{u_1}{R_1} + \frac{u_2}{R_2} + \frac{u_3}{R_3} + \cdots + \frac{u_n}{R_n} \right). \tag{1}$$

See Figure 1 for the corresponding circuit. By choosing all resistors to be equal and inverting the polarity of the input voltages, we can simplify Equation 1 to the following form,

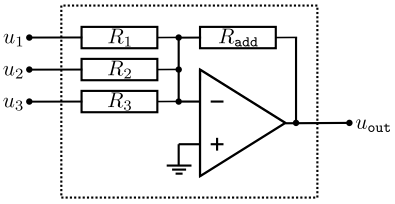

Fig. 1. Inverting adder with single OPAMP

$$u_{\texttt{out}} = u_1 + u_2 + u_3 + \cdots + u_n$$

which is a simple addition of the input voltages. Based on this idea, we can construct a circuit of n OPAMPs with a switched feedback network, which computes a solution to the equation system

$$A\boldsymbol{u} = -\boldsymbol{b}U_{\texttt{in}}, \tag{2}$$

where $U_{\texttt{in}}$ is the input voltage of the circuit.

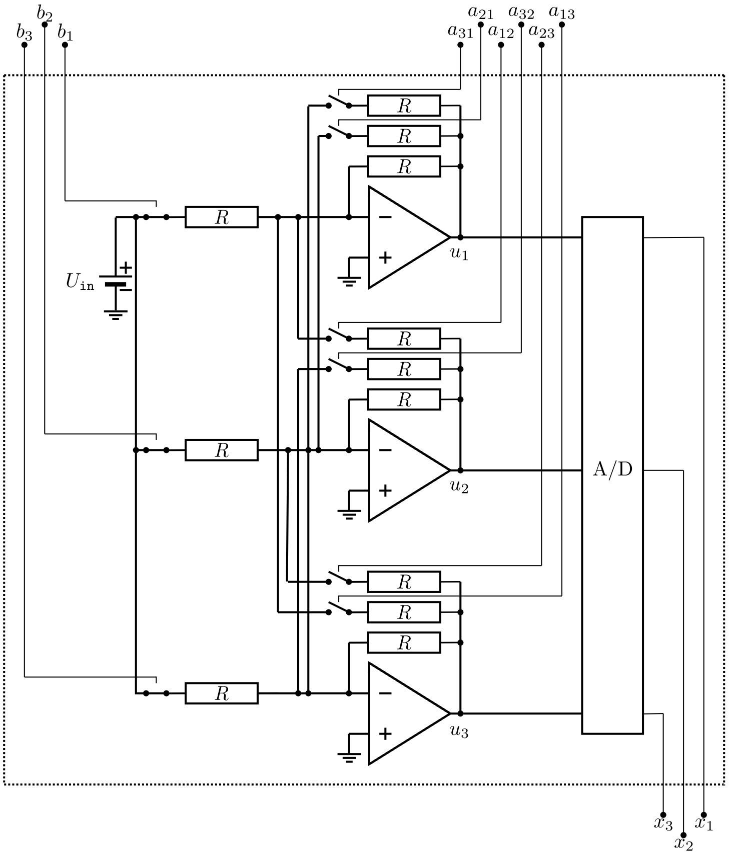

Fig. 2. Basic principle of constructing the Analog Solver (for three unknowns)

In this network, feedback loops between the OPAMPs are closed according to coefficients of an equation system which is to be solved – the solution to the equation system can be measured (after some oscillation) as voltage output of the OPAMPs. An example circuit for $n = 3$ is shown in Figure 2. The setting of the switches is determined by the corresponding binary coefficients of A and $\boldsymbol{b}$ (an open switch represents the binary value 0, and a closed one the value 1). Looking at each of the OPAMPs separately, we can write down equations for the expected output voltage of each OPAMP:

$$\begin{aligned} u_1 &= -b_1 U_{\texttt{in}} - a_{1,2}u_2 - a_{1,3}u_3 \\ u_2 &= -a_{2,1}u_1 - b_2 U_{\texttt{in}} - a_{2,3}u_3 \\ u_3 &= -a_{3,1}u_1 - a_{3,2}u_2 - b_3 U_{\texttt{in}} \end{aligned}$$

By re-arranging each equation accordingly, we easily see that the circuit represents a LES of quadratic form where A has a non-zero diagonal, i.e.,

$$\begin{aligned} u_1 + a_{1,2}u_2 + a_{1,3}u_3 &= -b_1 U_{\mathtt{in}} \\ a_{2,1}u_1 + u_2 + a_{2,3}u_3 &= -b_2 U_{\mathtt{in}} \\ a_{3,1}u_1 + a_{3,2}u_2 + u_3 &= -b_3 U_{\mathtt{in}} \end{aligned}$$

The example circuit solves a particular[1] equation system with three unknowns and binary coefficients which is exactly what we stated in Equation 2. When the circuit has converged after all relevant switches have been set, the output voltages of the OPAMPs will approximate the solution of the LES, i.e.,

$$\boldsymbol{u} = -U_{\mathtt{in}}(A^{-1}\boldsymbol{b}).$$

Without delving further into the specifics of this device (and our attempts to control "unwanted" oscillation and find easy to compute criteria for predicting these behaviors) we know that the device might eventually solve binary equation systems over the rationals. Based on this premise, we have developed a way to convert a rational solution to a solution over $\mathbb{F}_2$, which will be presented now.

3 Deriving a solution over $\mathbb{F}_2$

In this section we present a method to "interpret" a rational solution $\boldsymbol{u} = A^{-1}\boldsymbol{b}$ with $A \in \mathbb{F}_2^{n\times n}, \boldsymbol{b} \in \mathbb{F}_2^n$ and $\boldsymbol{u} \in \mathbb{Q}^n$ in order to find a binary solution for the same equation system. The rational solution $\boldsymbol{u}$ can be obtained either as voltages measured in our hypothetical device, or by other means. The only thing that matters is that the solution is "sufficiently" precise, a requirement that will become clear in Section 4. Furthermore, we assume that the LES is uniquely solvable over $\mathbb{F}_2$, therefore the determinant of A must be $|A| = 1$ over $\mathbb{F}_2$, and thus $|A| \equiv 1 \bmod 2$ when computing the determinant over the rationals. The latter fact is important and will be used extensively in the remainder of this section.

Now we will describe how we can convert the rational solution to a binary vector $\boldsymbol{x} \in \mathbb{F}_2^n$ which satisfies $A\boldsymbol{x} = \boldsymbol{b}$ over $\mathbb{F}_2$. We will do this with the help of three lemmata which will lead to the conversion method.

Lemma 1. *If the LES $A\boldsymbol{u} = \boldsymbol{b}$, $A \in \mathbb{F}_2^{n\times n}, \boldsymbol{b} \in \mathbb{F}_2^n$ is uniquely solvable over $\mathbb{F}_2$ and $\mathbb{Q}$, then computing the solution over $\mathbb{Q}$ yields a vector $\boldsymbol{u} \in \mathbb{Q}_\triangledown^n$ with*

$$\mathbb{Q}_\triangledown = \left\{ \frac{p}{q} : p, q \in \mathbb{Z}, q \equiv 1 \bmod 2 \right\}.$$

[1] The configuration of the switches indicates that the LES solved by the *Analog Solver* shown in the figure is $I\boldsymbol{u} = (1,1,1)^T$, where I is the 3×3 identity matrix.

Proof. Due to Cramer's Rule we know that a rational solution $\boldsymbol{u} = A^{-1}\boldsymbol{b}$ can be computed by the use of determinants, i.e.,

$$\boldsymbol{u} = \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_n \end{pmatrix} = \frac{1}{|A|} \begin{pmatrix} |A_1| \\ |A_2| \\ \vdots \\ |A_n| \end{pmatrix}$$

where $|A| \in \mathbb{Z}\backslash\{0\}$ denotes the determinant of the binary matrix A and $|A_i| \in \mathbb{Z}$ denotes the determinant of A where the i-th column has been replaced by $\boldsymbol{b}$. Given the condition of unique solvability of the equation system over $\mathbb{F}_2$, computing $|A|$ over $\mathbb{F}_2$ cannot be zero. This must also hold when computing $|A|$ over $\mathbb{Z}$ and applying the modulo operator only as last step – since both procedures must yield the same result.

Therefore $|A| \equiv 1 \bmod 2$ holds and thus all potential rational solutions $\boldsymbol{u}$ must be in the set $\mathbb{Q}_{\nabla}^n$.

Considering Lemma 1 and with the help of the ring homomorphism $\varphi(\cdot)$ which is defined as

$$\varphi : \mathbb{Q}_{\nabla} \mapsto \mathbb{F}_2 \quad \text{with} \quad \varphi\left(\frac{p}{q}\right) = \frac{p \bmod 2}{q \bmod 2} = p \bmod 2$$

we could directly deduce a solution $\boldsymbol{x} \in \mathbb{F}_2^n$ for a particular solution $\boldsymbol{u} \in \mathbb{Q}_{\nabla}^n$ by simply applying the $\varphi(\cdot)$-operator to the vector of quotients component-wise, i.e.,

$$\boldsymbol{x} = \varphi(\boldsymbol{u}) = \begin{pmatrix} \varphi(|A_1|) \\ \varphi(|A_2|) \\ \vdots \\ \varphi(|A_n|) \end{pmatrix} \equiv \begin{pmatrix} |A_1| \bmod 2 \\ |A_2| \bmod 2 \\ \vdots \\ |A_n \bmod 2 \end{pmatrix}.$$

However, when measuring the voltage output of our Analog Solver, we do only have fixed-point number representations of the solution vector $\boldsymbol{u}$ and no information about its quotients. Therefore no direct modulo-reduction is possible and we have to find another way to compute $\varphi(\boldsymbol{u})$.

Let us now consider how to convert $u_i \in \mathbb{Q}_{\nabla}$, which is the i-th element of $\boldsymbol{u}$ given as rational number in base-2, to a representation $x_i \in \mathbb{F}_2$. Let u_i be given in the following form which we will call the binary expansion of u_i:

$$u_i = c.d_0d_1d_2d_3\cdots, \quad c = \sum_{i=0}^{l-1} 2^i c_i \in \mathbb{N}, \quad c_i, d_i \in \{0, 1\}. \tag{3}$$

If for some fixed $k \in \mathbb{N}$ and $\forall i \in \mathbb{N}\backslash\{0\}$ we have

$$d_{ik} = d_0, d_{ik+1} = d_1, \cdots, d_{(i+1)k-1} = d_{k-1},$$

we call the representation above the purely periodic binary expansion of u_i and can re-write Equation 3 to

$$u_i = c.\overline{d_0 d_1 d_2 d_3 \cdots d_{k-1}},$$

as there are no non-periodic digit-patterns after the decimal point.

Lemma 2. *If the LES $A\boldsymbol{u} = \boldsymbol{b}$, $A \in \mathbb{F}_2^{n\times n}, \boldsymbol{b} \in \mathbb{F}_2^n$ is uniquely solvable over $\mathbb{F}_2$ and $\mathbb{Q}$, then the binary expansion of any element u_i of the rational solution $\boldsymbol{u} \in \mathbb{Q}_\triangledown^n$ is purely periodic.*

Proof. This proof follows the argumentation found in [YP04]. Suppose we have $u_i = p/q \in \mathbb{Q}_\triangledown$ where $p = |A_i| = cq + r_0$ and $q = |A|$ is odd. Since q is odd, it holds that $q \mid (2^k - 1)$ for some k and hence

$$q = \frac{2^k - 1}{l}, \quad k, l \in \mathbb{N}\backslash\{0\}.$$

Since

$$0 \leq \frac{r_0}{q} \leq 1 \quad \text{and} \quad 0 \leq (2^k - 1)\frac{r_0}{q} \leq 2^k - 1 \quad \text{with} \quad (2^k - 1)\frac{r_0}{q} = lr_0$$

where $d = lr_0$ is an integer we can write

$$d = \sum_{i=1}^{k} 2^{k-i} d_{i-1} = d_0 d_1 d_2 \cdots d_{k-1}, \quad d_i \in \{0, 1\}$$

as a bit-string with k bits. Since

$$d = (2^k - 1)\frac{r_0}{q} \quad \Leftrightarrow \quad \frac{r_0}{q} = 2^{-k} d + 2^{-k}\frac{r_0}{q} \quad \text{and} \quad 2^{-k} d = 0.d_0 d_1 d_2 \cdots d_{k-1}$$

we easily see that the quotient of r_0 and q exhibits a recursive behavior

$$\frac{r_0}{q} = 0.d_0 d_1 d_2 \cdots d_{k-1} + 2^{-k}\frac{r_0}{q} = 0.d_0 d_1 d_2 \cdots d_{k-1} d_0 d_1 d_2 \cdots d_{k-1} + 2^{-2k}\frac{r_0}{q},$$

and therefore the binary expansion of

$$u_i = \frac{|A_i|}{|A|}, \quad u_i \in \mathbb{Q}_\triangledown$$

is purely periodic.

If we actually measure a voltage representation of the solution for a given LES over the rationals and the results are given as purely periodic binary expansions of the form

$$c_{l-1} c_{l-2} \cdots c_0 . d_0 d_1 d_2 d_3 \cdots d_{k-1} d_0 d_1 d_2 d_3 \cdots d_{k-1} \cdots,$$

we "only" need to recover the c_0 and d_k bit of all u_i in order to interpret $\boldsymbol{u} = A^{-1}\boldsymbol{b}$ as a solution over $\mathbb{F}_2^n$. Given c_0 and d_k of a particular u_i we have an alternative way to compute $\varphi(u_i) = \varphi(p/q) \equiv p \bmod 2$, which will be discussed now.

Lemma 3. *Given $A\boldsymbol{u} = \boldsymbol{b}$, $A \in \mathbb{F}_2^{n\times n}, \boldsymbol{b} \in \mathbb{F}_2^n$, a rational solution $\boldsymbol{u} \in \mathbb{Q}_\triangledown^n$ and also a binary solution $\boldsymbol{x} \in \mathbb{F}_2$, the following holds for any pair of elements u_i, x_i, where*

$$c_{l-1}c_{l-2}\cdots c_0.\overline{d_0d_1d_2d_3\cdots d_{k-1}}, \quad c_i, d_i \in \{0,1\}, \quad l,k \in \mathbb{N}$$

is the purely periodic binary expansion of u_i:

$$x_i = \varphi(u_i) = \varphi(p/q) \equiv p \bmod 2 = c_0 \oplus d_{k-1}.$$

Proof. Suppose we have $u_i \in \mathbb{Q}_\triangledown^n$ in purely periodic form with

$$u_i = c_{l-1}\cdots c_0.\overline{d_0\cdots d_{k-1}}$$

which we convert to a quotient via

$$\begin{aligned} 2^k u_i &= c_{l-1}\cdots c_0 d_0\cdots d_{k-1}.\overline{d_0\cdots d_{k-1}} \\ \Leftrightarrow 2^k u_i - u_i &= c_{l-1}\cdots c_0 d_0\cdots d_{k-1} - c_{l-1}\cdots c_0 \\ \Leftrightarrow u_i &= \frac{c_{l-1}\cdots c_0 d_0\cdots d_{k-1} - c_{l-1}\cdots c_0}{2^k-1}. \end{aligned}$$

Now we know that

$$u_i = \frac{|A_i|}{|A|} = \frac{c_{l-1}\cdots c_0 d_0\cdots d_{k-1} - c_{l-1}\cdots c_0}{2^k-1}$$

where both denominators are odd and therefore

$$\varphi(u_i) \equiv |A_i| \bmod 2 = c_{l-1}\cdots c_0 d_0\cdots d_{k-1} - c_{l-1}\cdots c_0 \bmod 2 = c_0 \oplus d_{k-1}$$

holds.

There is also a special case, where deducing $\boldsymbol{x}$ is easy: In case $A \in \mathbb{F}_2^{n\times n}$ is in upper triangular form, we know that the determinant is the product of its diagonal elements. When the corresponding equation system is uniquely solvable over $\mathbb{F}_2$, it holds that

$$|A| = 1.$$

Therefore, for any $u_i \in \mathbb{Z}$ we can compute $\varphi(u_i)$ by looking at the least significant bit of u_i, which is an integer, i.e.,

$$x_i = \varphi(|A_i|) \equiv u_i \bmod 2.$$

4 Discussion

While our method is certainly interesting, there are some limitations. First of all, computing over the rationals with sufficient precision must be more desirable than direct and digital computation over $\mathbb{F}_2$. We were motivated to assume this by the prospect of a device, which promises to compute rational solutions very fast. However, the clear limitation, given a device which is indeed really fast, is the requirement of sufficient output precision – a notion which has been mentioned a few times already. "Sufficient" precision allows us to obtain the last bit of the binary expansion of each element of the rational vector $\boldsymbol{u}$. Our argumentation in Section 3 implies that the smaller the determinant of A (over the rationals), the shorter the length of the binary expansion and hence the lower the required precision of a physical device.

Since we ultimately want to speed up cryptographic attacks and to examine the feasibility of our approach, we have experimentally generated some of the binary linear equation systems which are used to execute Golic's attack [Gol97] against A5/1. For his attack, Golic needs to generate and solve more than 2^{40} equation systems of size $n = 64$ over $\mathbb{F}_2$. We have selected a random subset[2] of the generated equation systems and counted which determinants occur how often. The result of our examination is depicted in Figure 3 in a graph that shows the percentage of determinants (y-axis) which is lower than a specific value (x-axis). Observing the graph quite surprisingly[3] reveals that nearly 50% of all matrices have a determinant below 300, while 90% are still below 3000.

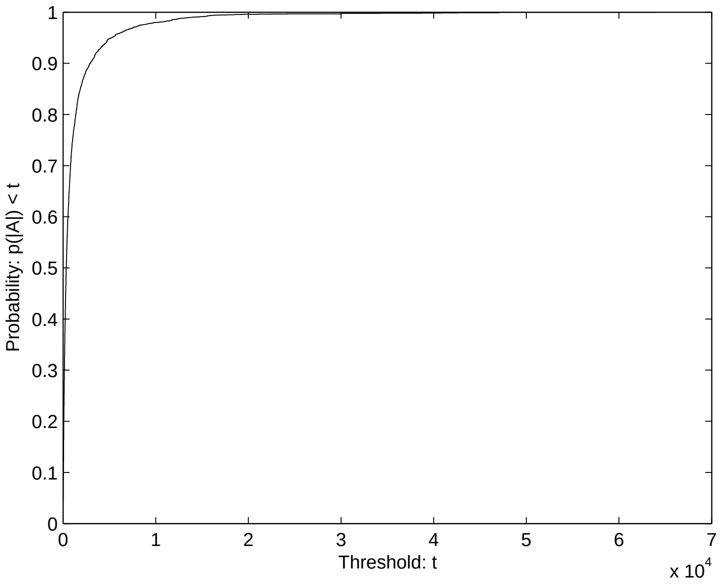

Fig. 3. Probability that the determinant in Golic's A5/1 attack is below a threshold t

[2] Of the more than 2^{40} equation systems 10 000 were chosen.

[3] We assume that this result is due to the inherent structure of the generated matrices.

We deduce that there are indeed existing and practically relevant instances of binary linear equation systems in which the arising determinants are small.

5 Conclusion

We have presented a surprising method that allows us to convert a solution of a particular, binary and quadratic LES over $\mathbb{Q}$ to a solution over $\mathbb{F}_2$. Our investigations were motivated by the idea of a hardware, which is only sketched in this paper. We emphasize that our method is generic in that it can be used in any other scenario where deriving a solution over $\mathbb{Q}$ is more desirable than direct computation over $\mathbb{F}_2$, given that the rational solution is sufficiently precise. We have discussed latter requirement and shown that linear equation systems with a low determinant also occur in cryptanalytic attacks.

References

[BBK03] Elad Barkan, Eli Biham, and Nathan Keller. Instant ciphertext-only cryptanalysis of GSM encrypted communication. In *Advances in Cryptology - CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 600–616. Springer Berlin / Heidelberg, 2003.

[GNR08] Timo Gendrullis, Martin Novotný, and Andy Rupp. A real-world attack breaking A5/1 within hours. In *Proceedings of the 10th international workshop on Cryptographic Hardware and Embedded Systems*, CHES '08, pages 266–282, Berlin, Heidelberg, 2008. Springer-Verlag.

[Gol97] Jovan Dj. Golic. Cryptanalysis of alleged A5 stream cipher. In *Proceedings of the 16th annual international conference on Theory and application of cryptographic techniques*, EUROCRYPT'97, pages 239–255, Berlin, Heidelberg, 1997. Springer-Verlag.

[LLMP93] A. Lenstra, H. Lenstra, M. Manasse, and J. Pollard. The number field sieve. *The development of the number field sieve*, pages 11–42, 1993.

[LS00] Arjen K. Lenstra and Adi Shamir. Analysis and optimization of the TWINKLE factoring device. In *Proceedings of the 19th international conference on Theory and application of cryptographic techniques*, EUROCRYPT'00, pages 35–52, Berlin, Heidelberg, 2000. Springer-Verlag.

[PFS00] Slobodan Petrovic and Amparo Fuster-Sabater. Cryptanalysis of the A5/2 Algorithm. Technical report, 2000. `http://eprint.iacr.org/`.

[PS00] Thomas Pornin and Jacques Stern. Software-hardware trade-offs: Application to A5/1 cryptanalysis. In *Proceedings of the 2nd international workshop on Cryptographic Hardware and Embedded Systems*, CHES '00, pages 318–327, Berlin, Heidelberg, 2000. Springer-Verlag.

[Sha99] Adi Shamir. Factoring large numbers with the TWINKLE device (extended abstract). In *Proceedings of the First International Workshop on Cryptographic Hardware and Embedded Systems*, CHES '99, pages 2–12, London, UK, 1999. Springer-Verlag.

[ST03] Adi Shamir and Eran Tromer. Factoring large numbers with the TWIRL device. In *Advances in Cryptology - CRYPTO 2003*, Lecture Notes in Computer Science, pages 1–26. Springer Berlin / Heidelberg, 2003.

[YP04] Thomas Yeung and Eric Poon. Binary decimal numbers and decimal numbers other than base ten. In *Proceedings of the International Conference on The Future of Mathematics Education*, 2004.