Some applications of coding theory in cryptography

Some Applications of Coding Theory in Cryptography

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Doumen, Jeroen M.

Some applications of coding theory in cryptography / by Jeroen M. Doumen. – Eindhoven : Technische Universiteit Eindhoven, 2003.
Proefschrift. – ISBN 90-386-0702-4
NUR 919
Subject headings : cryptology / coding theory / prime numbers
2000 Mathematics Subject Classification : 94A60, 11T71, 11A41

Printed by Eindhoven University Press.
Cover by JWL Producties.
Kindly supported by STW.

Some Applications of Coding Theory in Cryptography

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
Rector Magnificus, prof.dr. R.A. van Santen, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op 6 juni 2003 om 16.00 uur

door

Jeroen Mathias Doumen

geboren te Warstein, Duitsland.

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr.ir. H.C.A. van Tilborg

en

prof.dr. A.K. Lenstra

Contents

Preface

Nowadays, many people claim we live in the so-called information age. Clearly, the rise of the internet (among others) has made information available to people on an unprecedented scale and in a magnitude never seen before so widely. This can and has been compared to the introduction of the printing press in the Middle Ages. With its advent, the massive distribution of books and ideas became possible, and the printing press certainly has played a significant role since its invention. Even while it is still much too early to tell, the rise of the internet seems to be of a similar scale - for the first time in history, everyone is able to publish his own words.

However, such new flows of information need new technologies to expedite them. Of course, the basic networks along which the information flows have to be built. But there are other key issues here: one should be able to rely on the received information, in the sense that it is received correctly, even when the underlying network it is transmitted over is imperfect and thus prone to errors. Theoretical work on this began in the late 1940's with work of Shannon [Sha48] and Hamming [Ham50]. This has grown to a new branch of mathematics, called coding theory.

Also, there are many cases in which (a form of) confidentiality is required. An obvious example would be sending a love letter to one's secret lover, or sending other sensitive information in some digital form. But there are also other concerns: for instance, one could want to be sure of whether a certain (electronic) letter actually comes from the mentioned author. In daily life, the author can achieve this by writing his signature on the letter. But how can one do that in an email? Another, more mundane example is getting money from an ATM. Before handing you money, the bank wants to be sure that there is enough money in your account. On the other hand, you would like to be the only person able to withdraw money from your account. Again, going to a bank teller and using handwritten signatures has been the solution for centuries. But this is very difficult, if not impossible for an automated machine, and so other, intrinsically digital methods must be adopted. The tools of choice here, collectively called cryptography, used to protect national secrets. An excellent work on this history is given in Kahn's *The Codebreakers* [Kah67]. From the second world war onward, this rapidly became less of an art form and more and more a serious branch of mathematics.

Both coding theory and cryptography have been already proven to be essential in our information age. While they may seem to achieve opposite goals at first sight, they share much more than that. This thesis aims to reveal at least part of that

relation: how coding theory can be applied in cryptography. In the first chapter, a more detailed introduction to the objectives of cryptography will be given. Also, a short description of the basics of coding theory will be given there.

In Chapter 2 attacks on the McEliece public–key cryptosystem are introduced in which a malicious sender, or an adaptive eavesdropper, has a method available to find out whether a ciphertext decrypts properly or not. From this information she can then extract the plaintext that was encrypted. In this chapter it is shown that the McEliece public–key cryptosystem is indeed susceptible to these kinds of attacks and a detailed algorithm for such an attack is given and analyzed. Thus care should be taken when implementing this scheme, and possible countermeasures are discussed which thwart this attack.

Chapter 3 deals with the security of digital signature schemes based on error–correcting codes. Several attacks against the Xinmei scheme, the first such scheme, are surveyed and reasons for the failure of the Xinmei scheme are given. Another weakness is found in another such scheme, proposed by Alabbadi and Wicker, which leads to an attacker being able to forge signatures at will. Further analysis shows that this new weakness also applies to the original Xinmei scheme.

Then, in Chapter 4, work of a more theoretical nature will be discussed. In this chapter two families of numbers are introduced which can efficiently be tested for primality. These families naturally extend the Mersenne numbers to the area of elliptic curves. The first few primes in these families are also presented and compared to the generalized Wagstaff conjecture. However, results from this chapter will turn out to be useful in the last chapter.

Lastly, Chapter 5 will employ algebraic geometry to produce pseudorandom sequences. Some known constructions to produce pseudorandom sequences with the aid of elliptic curves will be generalized there. Both additive and multiplicative characters on elliptic curves will be used for this purpose. Finally, the use of linear recurrencies on elliptic curves will be studied.

Chapter 1

Preliminaries and notation

1.1 Cryptography

The aim of cryptography is to provide secure transmission of messages, in the sense that two or more persons can communicate in a way that guarantees that the desired subset of the following four primitives is met:

(i). *Confidentiality.* This primitive is usually perceived to be the main focus of cryptography, providing a way such that the information can only be viewed by those people authorized to see it.

(ii). *Data integrity.* This service will provide a means to check if the transmitted information was altered in any way, including but not limited to things like insertion, deletion and substitution of messages.

(iii). *Authentication.* This service will establish some identity pertaining to the message. Thus, this primitive can (among others) be used to guarantee the identity of the sender, guarantee the identity of the receiver, or guarantee the time the message was sent.

(iv). *Non-repudiation.* This serves to prevent someone from denying previous commitments. It is needed in cases where disputes might have to be resolved, for instance in E-commerce.

While cryptography is often thought of as a tool to provide confidentiality, the other three primitives are actually much more important in daily life.

In order to build cryptographic protocols supplying one or more of the above primitives, some building blocks are needed. For instance, one often uses a *one-way function,* of which the values should be easy to compute, but the inverse should be impossible to compute for most values. In practice, one is often content with a function for which it is computationally infeasible to compute inverses from. When a one-way function is defined on arbitrary inputs (i.e. on bitstrings of arbitrary length), it will be called a *hash function.* If a one-way function can be (efficiently) inverted given some additional information, it is called a *trapdoor one-way function.*

In a cryptographic protocol, the users are often called Alice and Bob (instead of A and B). An eavesdropper or adversary will be denoted by Eve. This terminology will be used throughout this thesis. The initiator of the protocol will be called Alice (so usually she will be the sender), and the intended recipient will be called Bob.

Cryptographic protocols include such things as encryption schemes, also called cryptosystems, which aim to achieve confidentiality. A description of such a scheme, called the McEliece cryptosystem, which is based on coding theory, will be given in Section 2.2. Other well-known cryptosystems are, among others, DES [MOV97, Section 7.4], AES [DR98], RSA [RSA78] (which can be used for digital signatures as well) and Diffie-Hellman [DH82], which is used for key agreement. Other examples of cryptographic protocols include digital signature schemes, which try to establish authentication and data integrity of a certain message. A history of signature schemes based on error–correcting codes will be given is Section 3.1. Others include RSA [RSA78] and DSA [MOV97, Section 11.5]. For a more complete list of cryptographic protocols, as well as descriptions of those only mentioned here, see [MOV97].

1.2 Coding Theory

The aim of coding theory is to provide secure transmission of messages, in the sense that (up to a certain number of) errors that occurred during the transmission can be corrected. However, for this capability a price must be paid, in the form of redundancy of the transmitted data. In this thesis only linear codes will be considered.

First the alphabet $\mathbb{F}_q$ is chosen. In practice, this usually is the field of binary numbers, but any prime power q is allowed. Let $\mathbb{F}_q^n$ denote a n–dimensional vector space over the finite field $\mathbb{F}_q$. A linear $[n, k]$ code $\mathcal{C}$ is a k–dimensional linear subspace of $\mathbb{F}_q^n$. The elements of $\mathcal{C}$ are similarly called *codewords.* A *generator matrix* G for $\mathcal{C}$ is a $k \times n$ (q–ary) matrix whose rows span $\mathcal{C}$. This means that each codeword $\mathbf{c}$ can be written as $\mathbf{c} = \mathbf{m}G$ (in $\mathcal{C}$) with $\mathbf{m} \in \mathbb{F}_q^k$. One can formulate this by saying that the message vector $\mathbf{m}$ is encoded in the codeword $\mathbf{c}$. The quantity $n - k$ is the *redundancy* of $\mathcal{C}$. It gives the number of excess symbols in $\mathbf{c}$, compared to the message vector $\mathbf{m}$.

Now $\mathbf{c}$ is sent over an (unreliable) channel and certain errors may be inflicted on $\mathbf{c}$: the received vector is $\mathbf{y} = \mathbf{c} + \mathbf{e}$ where $\mathbf{e}$ is a so called *error vector.* Let the Hamming weight $w_H(\mathbf{x})$ of a vector $\mathbf{x}$ simply count the number of non–zero coordinates of $\mathbf{x}$. If the weight of $\mathbf{e}$ is not too large, the received vector $\mathbf{y}$ coincides on many coordinates with $\mathbf{c}$ and $\mathbf{c}$ can be recovered.

With a code $\mathcal{C}$ one can associate its *dual code* $\mathcal{C}^\perp$, which is the $(n-k)$–dimensional subspace orthogonal to $\mathcal{C}$. In other words, the dual code consists of all vectors of $\mathbb{F}_q^n$ that are orthogonal to all codewords of $\mathcal{C}$. A generator matrix H of the dual code $\mathcal{C}^\perp$ is also called a *parity check matrix* of $\mathcal{C}$, since it has the property that $cH^T = 0$ for each codeword $c \in \mathcal{C}$ and vice versa.

The *(Hamming) distance* $d_H(\mathbf{x}, \mathbf{y})$ between vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as the number of coordinates where $\mathbf{x}$ and $\mathbf{y}$ differ. Note that $d_H(\mathbf{x}, \mathbf{y}) = w_H(\mathbf{x} - \mathbf{y})$. The *minimum distance* d of a code $\mathcal{C}$ is defined as the minimum Hamming distance between different codewords in $\mathcal{C}$. Since $\mathcal{C}$ is linear, an equivalent definition would be that d is the minimum non–zero weight of a codeword in $\mathcal{C}$. A $[n, k]$ code with minimum distance d will be denoted as a $[n, k, d]$ code. In general, determining the minimum distance d of a certain code (given e.g. by a generator matrix) is not an easy problem. However, some bounds on the minimum distance are known, one of which is Singleton's bound $d \leq n - k + 1$. Let A_w denote the number of codewords in $\mathcal{C}$ of Hamming weight w. The numbers $A_0, A_1, \ldots, A_n$ are called the *weight distribution* of $\mathcal{C}$.

The number $t = \lfloor \frac{d-1}{2} \rfloor$ is called the *error–correcting capability* of $\mathcal{C}$. It follows from the triangle inequality that for each element $\mathbf{y}$ in $\mathbb{F}_q^n$ there can be at most one element $\mathbf{c}$ in $\mathcal{C}$ at Hamming distance $\leq t$ to it. So, in principle, one can correct up to t errors inflicted to an element in $\mathcal{C}$ by finding the nearest point in $\mathcal{C}$. However, in practice the process of determining the nearest point (called *decoding*) is often very complex. To illustrate, the problem for general linear codes on deciding on whether there exists a point in $\mathcal{C}$ at a given distance of a given point $x \in \mathbb{F}_q^n$ is known to be in the class NP–complete [BMT78]. Fortunately there are certain classes of linear codes where decoding can be done quite effectively. As an example of this, Goppa codes shall be defined and described in the next subsection. More information about coding theory can for instance be found in [MS77; Til93a].

1.2.1 Goppa codes

In this subsection, a short introduction to Goppa codes will be given. For a more detailed description, as well as for proofs of most (unproven) statements given below, see [MS77, Section 12.3].

First, choose a Goppa parameter r, which will determine both the dimension and the minimum distance of the code. Let $G(x)$ be a polynomial of degree r over the finite field $\mathbb{F}_{q^m}$ and let $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$ contain n distinct members of $\mathbb{F}_{q^m}$ (n will be the length of the codewords) such that $G(\gamma_i) \neq 0$ for all $1 \leq i \leq n$. In practice, the choice $\Gamma = \mathbb{F}_{q^m}$, together with an irreducible polynomial $G(x)$ is often made. The *Goppa code* $\mathcal{C}_\Gamma$ generated by the Goppa polynomial $G(x)$ consists of all words $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in \mathbb{F}_q^n$ satisfying

$$\sum_{i=1}^{n} \frac{c_i}{x - \gamma_i} \equiv 0 \pmod{G(x)}.$$

Observe that the inverse of each polynomial $x - \gamma_i$ exists modulo $G(x)$, as $G(\gamma_i) \neq 0$ for all $1 \leq i \leq n$. The code $\mathcal{C}_\Gamma$ is linear and of dimension $k \geq n - mr$. Its minimum distance d satisfies $d \geq r + 1$. Moreover, if the characteristic of $\mathbb{F}_q$ is 2 and $G(x)$ does not have multiple zeros, then it even satisfies $d \geq 2r + 1$. The usual form of

the parity check matrix of $\mathcal{C}_\Gamma$ is given by

$$H_\Gamma = \begin{pmatrix} G(\gamma_1)^{-1} & \cdots & G(\gamma_n)^{-1} \\ \gamma_1 G(\gamma_1)^{-1} & \cdots & \gamma_n G(\gamma_n)^{-1} \\ \gamma_1^2 G(\gamma_1)^{-1} & \cdots & \gamma_n^2 G(\gamma_n)^{-1} \\ \vdots & & \vdots \\ \gamma_1^{r-1} G(\gamma_1)^{-1} & \cdots & \gamma_n^{r-1} G(\gamma_n)^{-1} \end{pmatrix}. \tag{1.1}$$

From this form, a parity check matrix over $\mathbb{F}_q$ can be obtained by writing each entry as the corresponding column vector of length m from $\mathbb{F}_q$.

An important feature of Goppa codes is the existence of an efficient decoding algorithm $\mathcal{A}_{t'}$ for any t' less than or equal to the designed error–correcting capability t. In practice, one therefore corrects up to the designed error–correcting capability.

Decoding can be done as follows: suppose that the vector $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ is received. The *syndrome* polynomial $S_\mathbf{y}(x)$ of $\mathbf{y}$ is defined by

$$S_\mathbf{y}(x) = \sum_{i=0}^{r-1} x^i (H_\Gamma \cdot \mathbf{y}^T)_{i+1} = \sum_{i=0}^{r-1} x^i \sum_{j=1}^{n} \gamma_j^i \mathbf{y}_j G(\gamma_j)^{-1},$$

or equivalently

$$S_\mathbf{y}(x) \equiv \sum_{i=1}^{n} \frac{\mathbf{y}_i}{x - \gamma_i} \pmod{G(x)}.$$

The syndrome polynomial $S_\mathbf{y}(x)$ is zero if and only if $\mathbf{y} \in \mathcal{C}_\Gamma$. Now write $\mathbf{y}$ as

$$\mathbf{y} = \mathbf{c} + \mathbf{e}, \text{ where } \mathbf{c} \in \mathcal{C}_\Gamma \text{ and } w_H(\mathbf{e}) \leq t. \tag{1.2}$$

Let E be the set of non–zero coordinates of $\mathbf{e} = (e_1, e_2, \ldots, e_n)$. Then the *error–locator polynomial* $\sigma(x)$ and the *error evaluator polynomial* $\omega(x)$ are defined by

$$\sigma(x) = \prod_{i \in E} (x - \gamma_i)$$

and

$$\omega(x) = \sum_{i \in E} e_i \prod_{j \in E \setminus \{i\}} (x - \gamma_j).$$

Then $\deg \omega(x) < \deg \sigma(x) \leq t$, $\gcd(\sigma(x), \omega(x)) = 1$ and the following, so–called *key equation* holds:

$$S_\mathbf{y}(x)\sigma(x) \equiv \omega(x) \pmod{G(x)}. \tag{1.3}$$

In order to decode the received word $\mathbf{y}$, this equation has to be solved. Of course, many solutions exist for this equation, and the one with the least degree of $\sigma(x)$ should be found. Such a solution certainly exists and is unique, since it was assumed that at most t errors occurred.

One way to do this is by using Euclid's extended algorithm. On input of two starting values $r_{-1}(x)$ and $r_0(x)$, this algorithm can be used in step i to calculate U_i, V_i and r_i satisfying

$$r_i(x) = (-1)^i \left(U_i(x)r_0(x) - V_i(x)r_{-1}(x)\right) \equiv (-1)^i U_i(x)r_0(x) \pmod{r_{-1}(x)}.$$

Now, starting with the values $r_{-1}(x) = G(x)$ and $r_0(x) = S_{\mathbf{y}}(x)$, proceed until one finds a $r_l(x)$ satisfying $\deg r_l(x) \leq \frac{1}{2}r - 1$. Then $\sigma(x)$ and $\omega(x)$ can be written as

$$\sigma(x) = \frac{U_k(x)}{U_k(0)} \tag{1.4}$$

and

$$\omega(x) = (-1)^k U_k(0)U_k(x). \tag{1.5}$$

These polynomials are proven to be the correct ones in [MS77]:

Proposition 1.2.1 ([MS77, Chapter 12, Theorem 16]) *The polynomials $\sigma(x)$ and $\omega(x)$ given by Equations (1.4) and (1.5) are the unique solution to the Key Equation (1.3) with $\sigma(0) = 1$, $\deg\sigma(x) \leq \frac{1}{2}r$, $\deg\omega(x) \leq \frac{1}{2}r - 1$ and $\deg\sigma(x)$ as small as possible.*

Clearly, once $\sigma(x)$ and $\omega(x)$ are determined, one can easily determine $\mathbf{c}$ and $\mathbf{e}$. The error locations set E is completely determined by the roots of $\sigma(x)$. Further, for each i in E one can compute the error value e_i (the i–th coordinate of $\mathbf{e}$) from the relation $e_i = \frac{\omega(\gamma_i)}{\sigma(\gamma_i)}$. Finally, the original codeword $\mathbf{c}$ can be computed as $\mathbf{c} = \mathbf{y} - \mathbf{e}$. Note that for this algorithm to work, the parity check matrix must be in the usual form (1.1), since the key equation only holds in that case. Also, the order of coordinates $\{\gamma_i\}$ in $\mathbb{F}_{q^m}$ must be known, since otherwise the error vector $\mathbf{e}$ could not be reconstructed from the error set E. Also note that this algorithm will correct up to $\lfloor r/2 \rfloor$ errors, regardless of the characteristic of $\mathbb{F}_q$.

In order to correct up to r errors if the code is binary (i.e. if the characteristic of $\mathbb{F}_q$ is 2) and if $G(x)$ has no multiple zeroes, one has to slightly adapt the above algorithm. First note that in this case the Key Equation (1.3) can be rewritten as

$$S(x)\sigma(x) \equiv \sigma'(x) \pmod{G(x)}.$$

Splitting of the squares and the non-squares in $\sigma(x)$, one can write $\sigma(x) = \alpha^2(x) + x\beta^2(x)$, where $\deg\beta(x) < \deg\alpha(x) \leq r/2$. Thus the key equation can be further rewritten as

$$\beta^2(x)\left(xS(x) + 1\right) \equiv S(x)\alpha^2(x) \pmod{G(x)}.$$

Multiplying this equation by the inverse[1] of $S(x)$, one sees that $\beta^2(x)T(x) \equiv \alpha^2(x) \pmod{G(x)}$ for some polynomial $T(x)$. Since the characteristic is 2, the Frobenius

[1] If this inverse does not exist, $G(x)$ is not irreducible. Then one can work modulo the irreducible factors of $G(x)$ and apply the Chinese Remainder Theorem. This is possible since $G(x)$ has no multiple zeroes.

map $y \mapsto y^2$ is an automorphism, and thus is it possible and well-defined to take the square root of an element of the field[1] $\mathbb{F}_{q^m}[x]/(G(x))$. So there exists a unique polynomial $R(x)$ satisfying $T(x) = R^2(x)$. Thus the key equation becomes

$$R(x)\beta(x) \equiv \alpha(x) \pmod{G(x)}.$$

On this form, Euclid's extended algorithm can be applied as described earlier.

1.2.2 The Maximal Error Property

Now the following property of a decoding algorithm $\mathcal{A}_t$ shall be investigated. This property will play a crucial role in Chapter 2.

Property 1.2.2 (Maximal Error Property) *On input of a vector $\mathbf{y} \in \mathbb{F}_q^n$, the decoding algorithm $\mathcal{A}_t$ will return a codeword $\mathbf{c}$ in $\mathcal{C}$ at distance $\leq t$ to $\mathbf{y}$ if such a codeword exists. Otherwise, it will return an error–message.*

Note that this property in fact states that the decoding algorithm $\mathcal{A}_t$ will not try to correct $t+1$ errors. This is not possible in general if $t = \lfloor \frac{d-1}{2} \rfloor$. However, even then it might be possible to correct $t+1$ errors in a few isolated cases.

Proposition 1.2.3 states a property of the two main algorithms to determine the pair $\sigma(x)$, $\omega(x)$ satisfying the Key Equation (1.3). One of the main algorithms is Euclid's algorithm which was described in the previous subsection. The other is the Berlekamp–Massey algorithm [MS77, Section 9.6], which tries to the pair $\{\sigma(x), \omega(x)\}$ by solving a set of simultaneous linear equations, namely the generalized Newton identities [MS77, Theorem 24, Chapter 8] It is important to note that the Berlekamp–Massey algorithm is successful if and only if Euclid's algorithm is; they also lead to the same result. This is irrespective of whether $S(x)$ is an actual syndrome polynomial $S_{\mathbf{y}}(x)$. The behavior of the Berlekamp–Massey algorithm and Euclid's algorithm with "bad" input, i.e. when inputting a syndrome polynomial of a vector $\mathbf{y}$ at distance more than t from the code, is the same.

These two decoding algorithms for Goppa codes have the maximal error property:

Proposition 1.2.3 *Let $\mathcal{C}$ be a Goppa code with designed error–correcting capability t (so $r = t$ in the binary case, and $r = 2t$ otherwise), and let $\mathcal{A}_t$ be either Euclid's extended algorithm or the Berlekamp–Massey algorithm for decoding $\mathcal{C}$. Then $\mathcal{A}_t$ has the maximal error property.*

Proof: First, suppose the characteristic of $\mathbb{F}_q$ is not equal to 2, suppose that $\mathbf{y} \in \mathbb{F}_q^n$ is the input to A_t and suppose that A_t outputs a vector $\mathbf{v}$ in $\mathbb{F}_q^n$ (so no error–message is returned). As the implementation is not assumed to explicitly check that the output is actually a codeword, $\mathbf{v}$ does not need to be a codeword. Certainly, if $\mathbf{y}$ is of the form (1.2) then $\mathcal{A}_t$ will output $\mathbf{c}$. However, to prove the proposition the converse has to be shown, i.e. that equality holds in (1.2) with $\mathbf{c} = \mathbf{v}$.

To this end, the decoding process is analyzed. First the decoding process tries – using either Euclid's or Berlekamp–Massey's algorithm – to find two polynomials

$\sigma_{\mathbf{y}}(x)$ and $\omega_{\mathbf{y}}(x)$ of the correct degrees and satisfying the Key Equation (1.3). If this fails, an error–message is assumed to be returned. On the other hand, if $\sigma_{\mathbf{y}}(x)$ and $\omega_{\mathbf{y}}(x)$ are determined, the decoding process determines the presumed set of error locations $E' = \{1 \leq i \leq n | \sigma_{\mathbf{y}}(\gamma_i) = 0\}$. If E' has cardinality strictly less than $\deg(\sigma_{\mathbf{y}}(x))$, an error–message is assumed to be given. Next, the decoding process determines a vector $\mathbf{e}'$ of weight at most t by

$$e'_i = \begin{cases} \omega_{\mathbf{y}}(\gamma_i)/\sigma_{\mathbf{y}}(\gamma_i) & \text{if } i \in E', \\ 0 & \text{otherwise.} \end{cases}$$

Note that all e'_i with $i \in E'$ are necessarily non–zero, as otherwise the $\sigma_{\mathbf{y}}$ and $\omega_{\mathbf{y}}$ would not be relatively prime, and thus the degree of $\sigma_{\mathbf{y}}$ would not be minimal. Finally, the decoding process determines $\mathbf{v} = \mathbf{y} - \mathbf{e}'$, which is returned by A_t.

Now to see that $\mathbf{v}$ is a member of $\mathcal{C}$, we consider the polynomials $\sigma_{\mathbf{e}'}(x)$ and $\omega_{\mathbf{e}'}(x)$ (corresponding to $\mathbf{e}'$ written as $\mathbf{e}' = \mathbf{0} + \mathbf{e}'$), i.e. $S_{\mathbf{e}'}(x)\sigma_{\mathbf{e}'}(x) \equiv \omega_{\mathbf{e}'}(x) \pmod{G(x)}$. First the observation is made that the pair $(\sigma_{\mathbf{e}'}(x), \omega_{\mathbf{e}'}(x))$ is equal to the pair $(\sigma_{\mathbf{y}}(x), \omega_{\mathbf{y}}(x))$. That $\sigma_{\mathbf{y}}(x) = \sigma_{\mathbf{e}'}(x)$ is trivially true, because the error vectors in $\mathbf{y} = \mathbf{v} + \mathbf{e}'$ and $\mathbf{e}' = \mathbf{0} + \mathbf{e}'$ are equal. That $\omega_{\mathbf{y}}(x) = \omega_{\mathbf{e}'}(x)$ follows from the fact that both are polynomials of degree $\leq \deg(\sigma_{\mathbf{y}}(x)) - 1$ that coincide on $\deg(\sigma_{\mathbf{y}}(x))$ distinct points.

So

$$S_{\mathbf{y}}(x)\sigma_{\mathbf{y}}(x) \equiv \omega_{\mathbf{y}}(x) = \omega_{\mathbf{e}'}(x) \equiv S_{\mathbf{e}'}(x)\sigma_{\mathbf{e}'}(x) \equiv S_{\mathbf{e}'}(x)\sigma_{\mathbf{y}}(x) \pmod{G(x)},$$

that is

$$(S_{\mathbf{y}}(x) - S_{\mathbf{e}'}(x))\sigma_{\mathbf{y}}(x) \equiv 0 \pmod{G(x)}.$$

The polynomials $G(x)$ and $\sigma_{\mathbf{y}}(x)$ are relatively prime, because a common factor would also divide $\omega_{\mathbf{y}}(x)$ by the Key Equation (1.3) and this contradicts the fact that the degree of $\sigma_{\mathbf{y}}(x)$ was minimal. Hence, it follows that $S_{\mathbf{v}}(x) = S_{\mathbf{y}}(x) - S_{\mathbf{e}'}(x) = 0$, i.e. $\mathbf{v} \in C$. This concludes the proof that $\mathbf{y}$ is of the form as described in equation (1.2). It also concludes the proof of the proposition in the non–binary case.

The proof that the output $\mathbf{v}$ of $\mathcal{A}_t$ is of the form of Equation (1.2), with $\mathbf{v} \in \mathcal{C}$, is similar to the non–binary case. □

Chapter 2

Adaptive chosen ciphertext attacks on the McEliece cryptosystem

Summary. Attacks are introduced in which a malicious sender or an adaptive eavesdropper Eve has an oracle which allows her to find out whether a ciphertext does, or does not, decrypt properly. From this information Eve can extract the plaintext that was encrypted. In this chapter it is shown that the McEliece public–key cryptosystem is susceptible to these kinds of attacks. This chapter is based on joint work with E. Verheul and H.C.A. van Tilborg [VDT02].

2.1 Introduction

In the last decade, several forms of attacks have been published where some of the inputs of an encryption system with a secret fixed key are adaptively chosen. By letting each new input (either plaintext or ciphertext) depend on the previous outputs and by looking at certain aspects of the resulting output at each step, additional secret information of the cryptosystem (for example the fixed key) may be determined. Among the studied aspects of the output are:

- the differences in the output when the differences in the plain inputs are known, see [BS93], [Mat93];
- some statistical or number–theoretic aspects of the output of the cryptosystem when errors are inflicted to the cryptosystem itself (e.g. by radiation), see for instance [BS97], [BDL97];
- so-called side-channel attacks, where the generation of the output is studied instead of the output itself. For instance, one can study the execution time when the precise complexity of the underlying cryptographic process is known [Koc96], or the power consumption of the cryptographic algorithm [KJJ99].

In this chapter, we will look at a different setting. Here the attacker Eve has access (for instance by interception) to one or more encrypted messages (called

ciphertexts) sent by Alice to a receiver Bob. Eve's aim is to recover the plaintext(s) of those messages.

Also suppose that Eve has access to an oracle that can tell her whether a ciphertext deciphers correctly or not. Oracles similar to this one were studied by Goldwasser, Micali and Tong in [GMT82]. Note that this oracle is weaker then the one that is usually supposed to be at Eve's disposal: Eve only gains knowledge whether or not the ciphertext is valid, but she is not given the decrypted ciphertext.

In practice, such an oracle might be easy to obtain: Eve may have access to Bob's decryption device or Eve might be an active eavesdropper. Another realistic possibility is that Bob's decryption device is in fact automated, and will send an automatic reply if the decryption somehow went wrong, asking for a retransmission. This reply can then be intercepted and used by Eve.

This setting was used in [Ble98] to attack protocols based on the RSA encryption standard PKCS #1. In this chapter, an efficient attack against the McEliece cryptosystem will be presented. The general idea of this attack is based on the following components:

(i). Eve alters the ciphertext slightly in such a way that there is a reasonable probability that the message still deciphers correctly and submits the altered message to her oracle.

(ii). Knowledge on whether the altered ciphertext deciphers correctly or not reveals new information and opens interesting new possibilities for adapting the ciphertext.

Eve will continue to alter messages in this way, until she has retrieved enough secret information. It is very likely that Eve will have to send a considerable number of altered messages.

The aim of this attack is to recover the plaintext of a given ciphertext.

In this chapter, attacks on the McEliece [McE78] public–key cryptosystem will be discussed. It is thus assumed that Eve has a validly encrypted McEliece message for Bob which she can alter and for which she is able to find out (e.g. by using her oracle) if the altered message remains a validly encrypted McEliece message.

The outline of this chapter is as follows: first the McEliece public–key cryptosystem will be described in Section 2.2. Section 2.3 is the main part of this chapter; there an effective message-recovery attack on the McEliece public–key cryptosystem will be described based on the maximal error correcting property of the two widely used decoding algorithms (See Property 1.2.2). In Section 2.4 some countermeasures against the described attack are considered.

Related Work

The attack described here differs from the one in [Ber97] where it is assumed that the (original) sender, instead of an eavesdropper, sends the same message more than once using different random error vectors. Also, an independent description of an algorithm, similar to ours, has appeared in [HGS99]. However, their algorithm

is significantly less efficient. Also, their conclusion that the McEliece cryptosystem should not be used because of this attack is unsupported - in Section (2.4) effective countermeasures will be discussed.

2.2 The McEliece Public–Key Cryptosystem

In 1978, McEliece [McE78] proposed a public–key cryptosystem based on the general difficulty of decoding. Consider a generator matrix G, generating a q–ary code $\mathcal{C}$ with parameters $[n, k, d]$, which is constructed by a user Bob. Let $0 \leq t \leq e = \lfloor (d-1)/2 \rfloor$ and let $\mathcal{A}_t$ be an effective decoding algorithm for $\mathcal{C}$ that can correct at most t errors.

Now, to use this in a cryptographic setting, Bob generates a random, invertible, q–ary, $k \times k$ matrix S and a random permutation matrix P of size $n \times n$. The public key of Bob is $G' = SGP$ together with the value of t. The matrices, S, G, P are kept secret. The idea is that G', although it generates a codespace $\mathcal{C}'$ which is equivalent to $\mathcal{C}$, behaves like a "random" generator matrix for which the decoding problem is hard.

Now suppose that another user, say Alice, wants to encrypt a message $\mathbf{m} \in \mathbb{F}_q^k$ for Bob. To this end she generates a random error vector $\mathbf{e}$ of weight $w(\mathbf{e}) \leq t$ and forms:

$$\mathbf{r} = \mathbf{m}G' + \mathbf{e} \quad (= \mathbf{m}SGP + \mathbf{e}). \tag{2.1}$$

Note that in some variants the weight of the error vector e is always exactly equal to t. On delivery, Bob calculates

$$\mathbf{r}P^{-1} = (\mathbf{m}S)G + \mathbf{e}P^{-1}.$$

As $\mathbf{e}P^{-1}$ has the same weight as $\mathbf{e}$, Bob can determine $\mathbf{m}S$ (and $\mathbf{e}P^{-1}$) from $\mathbf{r}P^{-1}$ by means of his effective decoding algorithm $\mathcal{A}_t$. Since S is a invertible matrix, Bob can easily determine $\mathbf{m}$, for instance by the method of Gaussian elimination.

More in particular, in the original scheme McEliece proposed to use binary (i.e. $q = 2$), irreducible Goppa codes, with $n = 1024$, $k \approx 524$ and $t = 50$. There exist many (different) codes of these parameters, they are easy to generate (randomly) and efficient decoding algorithms for them are easy to find. McEliece's construction can be extended to larger classes of codes (for instance non–binary Goppa codes). No details will be given here as that is not necessary for the attack; it suffices to mention the following bounds on the security–related parameters of the system.

Assumption 2.2.1 (The security of McEliece cryptosystem) *The following observations can be made on the parameters of a McEliece public–key cryptosystem:*

Sec–1. $k \approx n/2 \geq 512$*: this makes "syndrome decoding" as well as an exhaustive search for finding the nearest codeword to the received word infeasible;*

Sec–2. $50 \leq t \leq 100$*: this makes all kinds of techniques that are based on guessing/finding k (almost) error free coordinates less time consuming than the methods in Sec–1, but still infeasible (see [BKT99; Dum96; McE78]).*

The maximal error property (Property 1.2.2) states that the decoding algorithm $\mathcal{A}_t$, on input $\mathbf{r}$, never returns an element $\mathbf{c} \in C$ at distance more than t from $\mathbf{r}$. It is important to realize that if too many transmission errors have occurred, the received vector $\mathbf{r}$ may be at distance $\leq t$ from another codeword than the transmitted one. In this case $\mathcal{A}_t$ will not return an error message. The probability that this occurs, will be of importance in the analysis of the attack and will be discussed in Section 2.3. Recall (see Proposition 1.2.3) that the two relevant decoding algorithms for Goppa codes have the maximal error property.

2.3 An adaptive chosen ciphertext attack

Now the attack on the McEliece cryptosystem will be described.

Algorithm 2.3.1 [Adaptive chosen ciphertext attack on McEliece]
Assume that the decoding algorithm $\mathcal{A}_t$ used in a McEliece cryptosystem has the maximal error property. Let $\mathbf{r}$ be the ciphertext sent by Alice and intercepted by Eve (it is of the form $\mathbf{r}_A = \mathbf{m}G' + \mathbf{e}_A$). Then Eve does the following:

Step 1. Increase the number of errors made by Alice to exactly t.

In order to increase the number of errors to the maximum, Eve repeatedly changes a random coordinate arbitrarily (though each coordinate is selected at most once) and sends the resulting codeword to Bob until an error message is returned, i.e. the message is not accepted as a valid McEliece ciphertext. Once this occurs, Eve knows that this message contains exactly one error too much, and thus that the previous message she sent to Bob has the maximum number of errors. She now goes on to Step 2 with this message $\mathbf{r}'$.

Step 2. Determine enough error–free coordinates.

Once Eve knows she has a message $\mathbf{r}'$ with exactly t errors, she can start probing a random coordinate (different from all preceding choices, including those made in Step 1) by changing this arbitrarily in $\mathbf{r}'$, and sending the mutated message to Bob. If an error message is returned, this coordinate was error–free. Once enough error–free coordinates are determined, Eve can determine the plaintext in Step 3.

Step 3. Determine the plaintext.

Once Eve knows enough error–free coordinates, she can solve the matrix equation $\mathbf{r}' = \mathbf{m}G'$ for the plaintext $\mathbf{m}$ by using Gaussian elimination on the columns corresponding with the (known) error–free coordinates.

Before analyzing this algorithm, we introduce the following notion.

Definition 2.3.2 *The weight distribution* $\{A_w\}$ *of an* $[n, k, d]$ *code* $\mathcal{C}$ *will be called* approximately binomial *if (in the context of the problem here) the weight distribution may be approximated as follows:*

$$A_w \approx \frac{\binom{n}{w}(q-1)^w}{q^{n-k}}, \ d \leq w \leq n. \tag{2.2}$$

Note that in this case,

$$\sum_{w=0}^{n} A_w \approx \sum_{w=0}^{n} \frac{\binom{n}{w}(q-1)^w}{q^{n-k}} = \frac{(1+(q-1))^n}{q^{n-k}} = q^k = |C|,$$

as it should be. Certainly the weight distribution of $\mathbb{F}_q^n$ itself is binomial (in this case the approximation in (2.2) is actually an equality). We are not familiar with any result on how well the weight distribution of Goppa codes can be approximated by the binomial distribution, but based on [KFL85; KL95] and [MS77, Section 9.10] it seems very reasonable to make that assumption here.

For simplicity it will also be assumed that the minimum distance of the used code is odd, i.e. $d = 2t + 1$.

Theorem 2.3.3 *With the notation of Section 2.2, let the McEliece–like cryptosystem be based on a* q*–ary code* $\mathcal{C}$ *with an approximately binomial weight distribution, for instance a Goppa code. Also assume it uses a decoding algorithm that has the maximal error property. Then Algorithm 2.3.1 is an adaptive chosen ciphertext attack on* $\mathcal{C}$ *returning the plaintext* $\mathbf{m}$*.*

Before proving Theorem 2.3.3 two lemmas are needed. First recall that the binary entropy function $h(x)$ is defined on the interval $[0, 1]$ by $h(0) = h(1) = 0$ and $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ if $0 < x < 1$.

Lemma 2.3.4 *In the notation of Section 2.2, let* $\mathbf{e} \in \mathbb{F}_q^n$ *be an 'error–vector' of weight* $t + 1$*. Then the following holds:*

i) *There is at most one vector* $\mathbf{f} \in \mathbb{F}_q^n$ *of weight* $\leq t$ *such that* $\mathbf{e} + \mathbf{f} \in \mathcal{C}$*. Also, if such an* $\mathbf{f}$ *exists, then the weight of* $\mathbf{f}$ *is exactly equal to* t*, the supports (the sets of non–zero coordinates) of* $\mathbf{e}$ *and* $\mathbf{f}$ *are disjoint and* $d = 2t + 1$*.*

ii) *If* $\mathbf{e}$ *is chosen uniformly random, then the probability* P *that a vector* $\mathbf{f}$ *of weight at most* t *exists such that* $\mathbf{e} + \mathbf{f} \in \mathcal{C}$ *is given by*

$$P = \frac{A_{2t+1}\binom{2t+1}{t+1}}{\binom{n}{t+1}(q-1)^{t+1}}. \tag{2.3}$$

iii) *If the weight distribution of* $\mathcal{C}$ *is approximately binomial, then*

$$P \approx \frac{\binom{n-(t+1)}{t}(q-1)^t}{q^{n-k}} \leq \frac{2^{(n-(t+1))h\left(\frac{t}{n-(t+1)}\right)}(q-1)^t}{q^{n-k}},$$

where h *is the binary entropy function.*

iv) *Assume that the weight distribution of the Goppa codes is indeed approximately binomial. If a binary Goppa code is used in a McEliece cryptosystem, the above probability P is negligible. To be more precise, when the parameters originally proposed by McEliece in [McE78] are used, we have $P \leq 2^{-215}$. When the improved parameters as mentioned in Assumptions 2.2.1 are used, we have that $P \leq 2^{-54}$.*

Proof:

i) Suppose that two distinct candidates for $\mathbf{f}$ – as mentioned in the first part of the lemma – exist, say $\mathbf{f}$ and $\mathbf{f}'$. Then

$$d(\mathbf{e}+\mathbf{f}, \mathbf{e}+\mathbf{f}') = d(\mathbf{f}, \mathbf{f}') = w(\mathbf{f}-\mathbf{f}') \leq w(\mathbf{f}) + w(\mathbf{f}') \leq 2t < d.$$

As $\mathbf{e}+\mathbf{f}$ and $\mathbf{e}+\mathbf{f}'$ are two distinct members of $\mathcal{C}$ at distance less than d, we arrive at a contradiction. A similar argument shows that the weight of $\mathbf{f}$ must be equal to t and that $d = 2t+1$.

ii) Let

$$B = \{(\mathbf{e}, \mathbf{f}) \in \mathbb{F}_q^n \times \mathbb{F}_q^n \mid \mathbf{e}+\mathbf{f} \in C,\ w(\mathbf{e}) = t+1,\ w(\mathbf{f}) = t\}.$$

Let $\mathbf{c}$ be a codeword in $\mathcal{C}$ of weight $2t+1$. Then each $(t+1)$–subset of the support of $\mathbf{c}$ gives rise to a unique pair $(\mathbf{e}, \mathbf{f}) \in B$ (change the remaining t non–zero coordinates of $\mathbf{c}$ into a zero, resp. the $t+1$ coordinates themselves). Conversely, each element $(\mathbf{e}, \mathbf{f}) \in B$ can be obtained this way. Hence it follows that $|B| = A_{2t+1}\binom{2t+1}{t+1}$.

The total number of 'error–vectors' of weight $t+1$ in $\mathbb{F}_q^n$ is $\binom{n}{t+1}(q-1)^{t+1}$. The probability P is the quotient of $|B|$ and this number.

iii) By assumption the relation

$$A_{2t+1} \approx \frac{\binom{n}{2t+1}(q-1)^d}{q^{n-k}}$$

holds. It follows from ii) that

$$P \approx \frac{\binom{n}{2t+1}(q-1)^{2t+1}\binom{2t+1}{t+1}}{q^{n-k}\binom{n}{t+1}(q-1)^{t+1}} = \frac{\binom{n-(t+1)}{t}(q-1)^t}{q^{n-k}}.$$

To arrive at the inequality in iii), note that the binomial theorem implies the inequality

$$n^n = (m+(n-m))^n = \sum_{i=0}^{n} \binom{n}{i} m^i (n-m)^{n-i} \geq \binom{n}{m} m^m (n-m)^{n-m}$$

for each $0 \leq m \leq n$. This can be rewritten as

$$\binom{n}{m} \leq \frac{n^n}{m^m(n-m)^{n-m}} = \frac{2^{n\log_2 n}}{2^{m\log_2 m}2^{(n-m)\log_2(n-m)}} = 2^{-m\log_2(m/n)-(n-m)\log_2((n-m)/n)} = 2^{nh(m/n)}.$$

Note that since

$$n^n = (m+(n-m))^n = \sum_{i=0}^{n}\binom{n}{i}m^i(n-m)^{n-i} \leq (n+1)\binom{n}{m}m^m(n-m)^{n-m},$$

it also follows that

$$\binom{n}{m} \leq 2^{nh(m/n)} \leq (n+1)\binom{n}{m}.$$

Note that this inequality is often used in the literature, in the form

$$2^{nh(m/n)} \geq \binom{n}{m} \geq \frac{1}{n+1}2^{nh(m/n)} = 2^{n(h(m/n)-\log_2(n+1)/n)}$$

to prove that for $0 \leq \alpha \leq 1$ and as n tends to infinity, $\binom{n}{\alpha n} = 2^{nh(\alpha)}$.

iv) With the assumption that the weight distributions of the Goppa codes are indeed approximately binomial, the probability P mentioned in ii) can be approximated using iii). If the parameters proposed by McEliece are used, i.e. $n = 1024$, $k = 524$, $t = 50$, the probability P can be approximated by

$$P \leq \frac{2^{975*h(50/975)}}{2^{500}} \approx \frac{2^{285}}{2^{500}} = 2^{-215},$$

which is negligible. Similarly, $P \leq 2^{-54}$ if the parameters are as in Assumption 2.2.1. So the same holds for general McEliece cryptosystems, provided of course the weight distribution of the used code is approximately binomial.

□

The following observation may be of interest to the reader. It is well known that for a perfect t–error correcting code $\binom{2t+1}{t}A_{2t+1} = \binom{n}{t+1}(q-1)^{t+1}$ (see for instance [Til93a, Problem 3.4.9]). Substitution of this relation into (2.3) gives $P = 1$, as it should be for a perfect code: each word at distance $t+1$ from one codeword lies at distance t from exactly one other codeword. Thus a Sloppy Alice attack on McEliece–like cryptosystem which uses a perfect code will not work, since each vector can be decoded and thus no error messages will be generated.

Lemma 2.3.5 *The probability that a random $k \times (k+\log_q k)$ q–ary matrix has rank k is at least $1 - \frac{1}{k}$.*

Proof: Let $P(k,m),\ m \geq k$, denote the probability that a random $k \times m$ binary matrix A has rank k. Looking at the rows of A we observe that the first row of A should be non–zero, the second row should be independent of the first, etc. This argument leads to

$$\begin{aligned} P(k,m) &= \frac{\prod_{i=0}^{k-1}(q^m - q^i)}{\prod_{i=0}^{k-1} q^m} = \prod_{i=m-k+1}^{m}(1-\frac{1}{q^i}) \\ &\overset{(*)}{\geq} 1 - \sum_{i=m-k+1}^{m} \frac{1}{q^i} = 1 - \frac{1-\frac{1}{q^k}}{(q-1)q^{m-k}} \geq 1 - \frac{1}{q^{m-k}}, \end{aligned}$$

where $(*)$ follows quite easily with an induction argument. Now substituting $m = k + \log_q k$ in the above relation gives

$$P(k, k+\log_q k) \geq 1 - \frac{1}{k}.$$

□

Now the main theorem can be proven:

Proof of Theorem 2.3.3: Consider any $\mathbf{r} = \mathbf{c} + \mathbf{e}$ where $\mathbf{c} \in C$ and $w(\mathbf{e}) = s \leq t$. If the i^{th} coordinate of $\mathbf{r}$ is changed, which can be described by adding a vector $\mathbf{u}$ of weight 1 and with support $\{i\}$ to $\mathbf{r}$, then there are three possibilities for the resulting $\mathbf{r}' = \mathbf{r} + \mathbf{u} = \mathbf{c} + \mathbf{e}'$ (only two if $q = 2$):

(i). $w(\mathbf{e}') = s - 1$ if and only if $e_i \neq 0$ and $u_i = -e_i$,

(ii). $w(\mathbf{e}') = s$ if and only if $e_i \neq 0$ and $u_i \neq -e_i$
(This is impossible if $q = 2$, because both are also non–zero in this case),

(iii). $w(\mathbf{e}') = s + 1$ if and only if $e_i = 0$.

Consider Step 1 of Algorithm 2.3.1. For the range $0 \leq i \leq 2t+1$, let $\mathbf{e}^{(i)} = \mathbf{r}^{(i)} - \mathbf{c}$, that is, $\mathbf{r}^{(i)} = \mathbf{c} + \mathbf{e}^{(i)}$. Of course each $\mathbf{e}^{(i)}$ is unknown to Eve and $\mathbf{e}^{(0)} = \mathbf{e}_A$. As $w(\mathbf{e}^{(0)}) = w(\mathbf{e}_A) \leq t$ it follows that there exists a first $0 < i \leq 2t+1$ in Step 1, such that $w(\mathbf{e}^{(i)}) = t$ and $w(\mathbf{e}^{(i+1)}) = t+1$. So, for $0 \leq j \leq i$ the execution of the decoding algorithm $\mathcal{A}_t$ applied by Bob to $\mathbf{r}^{(j)}$ does not result in an error–message.

Note that i can only reach the value $2t+1$ in the (extremely unlikely) case that the $2t+1$ errors introduced by Eve in this step of the attack algorithm include all the errors that Alice originally has added to $\mathbf{c}$. In this case, we can immediately proceed to Step 3 of the algorithm since all other coordinates will be error–free, and contain enough independent columns of the generator matrix G'.

Next, it is claimed that the $\mathcal{A}_t$ applied to $\mathbf{r}^{(i+1)} = \mathbf{c} + \mathbf{e}^{(i+1)}$ with $\mathbf{e}^{(i+1)}$ of weight $t+1$, will result in an error–message (and we go to Step 2). Indeed, if $\mathcal{A}_t$ applied to $\mathbf{r}^{(i+1)}$ does not issue an error–message then $\mathbf{r}^{(i+1)}$ lies at distance at most t from $\mathcal{C}$ by the maximal error property. This means that there exist a codeword $\mathbf{c}'$ in $\mathcal{C}$ and

a vector $\mathbf{e}'$ in $\mathbb{F}_q^n$ of weight at most t such that $\mathbf{r}^{i+1} = \mathbf{c}' + \mathbf{e}'$. Hence, the situation of Lemma 2.3.4 applies, from which it follows that this situation has a negligible probability of occurring.

In Step 2, write $\mathbf{r}' = \mathbf{c} + \mathbf{e}'$, with $w(\mathbf{e}') = t$. Following the same reasoning as above, when the f–th coordinate in $\mathbf{r}$ is changed, the obtained vector $\hat{\mathbf{r}}$ will not be accepted by $\mathcal{A}_t$ (i.e. without error–messages) if and only if (with a negligible probability of failure) the f–th coordinate of $\mathbf{e}'$ is zero. □

Theorem 2.3.6 *Let the number of times the loop in Step 1 is executed be S_1. With large probability the loop in Step 2 will be executed at most $k+\log_q(k)+X+1$ times, where $X = \min(t, 2t+1-S_1)$.*

Corollary 2.3.7 *With large probability Algorithm 2.3.1 is an adaptive chosen ciphertext attack which uses at most $k + \log_q(k) + 2t + 2$ oracle queries.*

Note that if a binary code is used, there is only one possibility for an error–coordinate. In that case Algorithm 2.3.1 can be improved to an attack of at most $k + 2t + 1$ oracle queries. Also note that if in Step 2 of Algorithm 2.3.1 t rejections are encountered, all errors introduced by Alice have been found and all remaining coordinates are error–free. In this case deciphering can be done much faster. Of course, the probability that this occurs is negligible.

Proof of Theorem 2.3.6: The number of loops in Step 2 of the algorithm follows directly from Lemma 2.3.5. Since in this step only coordinates different from those selected in Step 1 are taken, an extra term $+X$ must be added that reflects the (marginal) possibility that X times a change made in this step canceled out an error introduced by Alice. Thus the number of oracle queries needed in Step 2 of the algorithm is equal to $k + \log_q(k) + S_1 + X$.

However, if a change in Step 2 canceled out an error introduced by Alice, this error could not have been canceled in Step 1. Also, the number of oracle queries in Step 1 is at most $2t + 1$. Thus the sharper bound $S_1 + X \leq 2t + 1$ holds. Because $S_1 + X \leq 2t+1$, with large probability the algorithm uses at most $k+\log_q(k)+2t+1$ oracle queries. □

2.4 Countermeasures

Countermeasures to the adaptive chosen ciphertext attack on the McEliece cryptosystem should at least aim to achieve that there is no correlation between deciphering problems and the number of errors applied to the plaintext.

First, in the case that Bob is Eve's oracle, Bob could come up with the idea of checking for repeated messages. This would detect an attack as described above, but nothing prevents the attacker Eve from adding a random codeword from $\mathcal{C}$ to her probe each time she queries the oracle. This preserves the error–vector e, and will allow Eve to conduct her attack as usual with little additional effort.

As a second idea one might consider to fix the weight of the error vector to exactly t, (or any $t' \leq t$). (cf. Equation (2.1)) and to return an error–message when in the deciphering process an error vector is encountered of weight unequal to t (or the chosen t'), that is, irrespective of whether successful decoding is still possible.

However, in this setting effective attacks are still possible. First of all, suppose that the used code is non–binary. If one 'probes' a coordinate in Step 2, say the i–th coordinate, twice but with different values, then Bob will always return an error–message if that coordinate is error–free (since there are $t+1$ errors in both probes). However, if there's an error on that coordinate, at most one probe will return no error message (since Eve only alters the value of $\mathbf{e}_i$ and not the weight of the error–vector).

So we have the following situation:

- if both probes give an error–message, then $e_i = 0$,
- if only one probe gives an error–message, then $e_i \neq 0$ and e_i is in fact determined,
- if none of the probes gives an error–message, then $e_i \neq 0$.

It easily follows that the plaintext can be found with approximately $k+t$ probes, i.e. in approximately $2(k+t)$ oracle queries.

If the used code is binary, then one starts by changing any coordinate, say the i–th, of $\mathbf{r}'$. In the setting here, this will always lead to an error–message from the oracle. Now we distinguish two possibilities.

With probability $(n-t)/n$ one has that $e_i = 0$. If one changes an additional coordinate in all possible ways, then $n-t-1$ times another error will have been introduced, resulting in an error–message from the oracle. Further, t times an error (introduced originally by the original sender) will be eliminated and so in this case one is back at t errors, leading to a correct decryption. In this way, all errors introduced by Alice can be found.

If $e_i = 1$ (with probability t/n), then additionally introduced single errors will $n-t$ times lead to correct decryptions (and coordinates with $e_i = 0$) and $t-1$ times to an error message.

In the case that Bob is in fact serving as Eve's oracle, a better countermeasure to the attack technique may be to introduce further redundancy in the system to enable Bob to check if an active eavesdropper is altering a proper ciphertext.

For instance, let Alice choose her plaintext $\mathbf{m}$ from $\mathbb{F}_q^{k-128}$ instead of $\mathbb{F}_q^k$. As before, she chooses a random error vector of weight $\leq t$. Bob has published as part of his public key a cryptographically secure hash–function h that maps bit strings of arbitrary length to elements in $\mathbb{F}_q^{128}$ (this hash function can also be a system parameter). To encrypt $\mathbf{m}$ Eve computes (cf. Equality (2.1)):

$$\mathbf{r} = (\mathbf{m}||h(\mathbf{m}||\mathbf{e}))G' + \mathbf{e} \quad (= (\mathbf{m}||h(\mathbf{m}||\mathbf{e}))SGP + \mathbf{e}), \tag{2.4}$$

where $||$ stands for concatenation. When Bob receives a vector $\mathbf{r}'$, he will attempt

to decode it. If this works, he will find an error vector $\mathbf{e}'$ and an element $\mathbf{m}' \in \mathbb{F}_q^k$ satisfying

$$\mathbf{r}' = \mathbf{m}'G' + \mathbf{e}'.$$

Finally, he checks the hash value of the message and the error vector. If this verification fails, an error–message is issued. This error–message will not give any information to Eve about the original choice of $\mathbf{e}$ by Alice, since any error results in rejection. It follows that the attack as described in Section 2.3 fails.

Note that the choice of 128 bits in the above example is arbitrary: this is a security parameter of the system which indicates how many message bits are used to increase security. We refer to [FO99] for a general description of this construction. Also observe that as a side effect of this variation the McEliece cryptosystem loses its inherent error–correcting capabilities. This seems to be inevitable.

2.5 Conclusion

In this chapter an adaptive chosen ciphertext attack was introduced, which is based on the assumption that (ordinary) users may see no problem in revealing whether or not an encrypted message deciphers correctly. Such a Sloppy Alice attack on the McEliece public–key cryptosystem was analyzed in detail.

The general conclusion is that such error–messages can be used to efficiently decrypt any message encrypted with the McEliece cryptosystem. Therefore, at the very least, error–messages should be as non–descriptive as possible and users should be alerted when many encrypted messages do not decrypt properly. Also, a variant of the McEliece cryptosystem which is immune to attacks was proposed.

Chapter 3

Digital signature schemes based on error–correcting codes

Summary. In this chapter the security of digital signature schemes based on error–correcting codes is discussed. Several attacks against the Xinmei scheme are surveyed and some reasons for the failure of the Xinmei scheme are given. Another weakness is found in the Alabbadi–Wicker scheme. This weakness leads to a universal forgery attack against it. Further analysis shows that this new weakness also applies to the Xinmei scheme. This chapter is based on joint work with S.B. Xu and H.C.A. van Tilborg [XD99; XDT03].

3.1 Introduction

The concept of digital signatures was proposed by Diffie and Hellman when they introduced public–key cryptography in their pioneering paper [DH82]. When Alice wants to sign a message $\mathbf{m}$ and send the signature to Bob, she sends the pair $(\mathbf{m}, \mathbf{s})$ where $\mathbf{s}$ is the signature $\mathbf{s} = \text{Sign}(\mathbf{m})$. Bob can then verify the signature by applying Alice's public verification algorithm Ver to $\mathbf{s}$ (thus the relation $\text{Ver}(\mathbf{s}) = \text{Ver}(\text{Sign}(\mathbf{m})) = \mathbf{m}$ must hold). Sometimes, the message $\mathbf{m}$ is even omitted. In that case, the verification algorithm will return the message $\mathbf{m}$. Such a scheme is called a *message recovery scheme*. See for instance [MOV97, Chapter 11] for more information on this subject.

Digital signatures play an important role in electronic commerce because they can replace a written signature. Several digital signature schemes are based on the integer factorization problem (e.g. RSA [RSA78]) and the discrete logarithm problem (e.g. ElGamal [ElG85]). People are now trying to design new digital signature schemes based on other mathematically hard problems. The problem of decoding general linear codes is such a problem, which has been proven to be NP–complete by Berlekamp, McEliece and Van Tilborg [BMT78]. McEliece [McE78] first proposed a public–key cryptosystem based on linear error–correcting codes, which derives its security from the above general decoding problem. No efficient attack on

McEliece's cryptosystem has been found until now, though several computationally intensive attacks have been discussed in the literature [Cha95; Ber97; CS98] and attacks on implementations of the McEliece cryptosystem have been published (see e.g. [HGS99; VDT02] and also Chapter 2 of this thesis). Since 1978, several other cryptosystems based on error–correcting codes have been proposed, such as the Rao–Nam private–key cryptosystem [RN89], the Xinmei signature scheme [Wan90] and the Stern identification scheme [Ste93]. These schemes are either used to protect the secrecy or to provide the authenticity of the message according to different needs. In this chapter the security of digital signature schemes based on error–correcting codes is discussed.

Some public–key cryptosystems can directly be used as digital signature schemes, for instance the RSA cryptosystem [RSA78]. However, McEliece's public–key cryptosystem cannot be used directly as a digital signature scheme because its encryption function maps binary k–tuples to binary n–tuples. Since this mapping is not surjective [MOV97, Section 8.5], it can not be inverted. Thus almost no messages can be signed, since no k-tuple exists that maps onto this particular message.

In 1990, Xinmei Wang proposed the first digital signature scheme based on error–correcting codes [Wan90], referred to as the Xinmei scheme. In the Xinmei scheme, the signature is generated in a manner similar to the way plaintext is encrypted in the Rao–Nam private–key cryptosystem [RN89]. The Xinmei scheme was claimed to have its security rely on the large number of generator matrices of a particular error–correcting code and the difficulty of retrieving a particular one from its scrambled public equivalents. In 1992, several methods were proposed to attack the Xinmei scheme. Harn and Wang [HW92] first proposed a homomorphism attack on the Xinmei scheme without factoring large matrices, and presented an improved scheme (here called the Harn–Wang scheme) in which a nonlinear function is introduced to defeat the homomorphism attack. Then, Alabbadi and Wicker [AW92b] showed that the Xinmei scheme is vulnerable to a chosen plaintext attack with complexity $O(n^3)$. They [AW92a] also showed that the Harn–Wang scheme can be broken completely by a known plaintext attack with complexity $O(k^3)$. Later, Van Tilburg [Til92] showed that one can directly obtain the signature key from the public keys in both the Xinmei scheme and the Harn–Wang scheme. In 1993, Alabbadi and Wicker [AW93] proposed a new digital signature scheme based on error–correcting codes. In the same year, Van Tilburg [Til93b] showed that this new scheme is not secure if one is able to verify n signatures (with linearly independent error vectors). In 1994, Alabbadi and Wicker [AW94] proposed a universal forgery attack on the Xinmei scheme and their own scheme. Later that year, Alabbadi and Wicker [AW95] presented another digital signature scheme based on error–correcting codes, which will here be referred to as the Alabbadi–Wicker scheme. They claimed that the proposed scheme is resistant to the attacks that proved successful when used against the aforementioned digital signature schemes.

Courtois, Finiasz and Sendrier recently proposed [CFS01] a digital signature scheme based directly on the McEliece cryptosystem. They argued that the problem of the non–surjective encryption mapping is not insurmountable: the probability

that a random McEliece ciphertext is valid, i.e. the probability that it can be decrypted is about $\frac{1}{t!}$. Thus, by successively trying about $t!$ different ciphertext candidates, formatted as the concatenation of a hash of the message m and a counter $0 \leq i \leq t!$, a valid signature can be obtained.

The subsequent sections are organized as follows: in the second section, the Xinmei scheme will be introduced and studied. In the third section the Alabbadi–Wicker scheme will be described. Then a universal forgery attack against the Alabbadi–Wicker scheme will be presented. Finally, some comments about the security of digital signature schemes based on error–correcting codes are made.

3.2 Security analysis of the Xinmei scheme

3.2.1 Description of the Xinmei scheme

The Xinmei scheme works as follows:

Setup phase: Alice takes a $(k \times n)$ generator matrix G of a binary Goppa code with an error–correcting capability of t errors of which t' errors can be corrected efficiently. She also chooses a right inverse matrix G^{-R} of G, so G^{-R} satisfies $GG^{-R} = I_k$, where I_k is the $k \times k$ identity matrix. Furthermore, she chooses a $n \times n$ full rank random matrix R and a $k \times k$ full rank matrix S, called the scrambling matrix.

Alice then publishes her public keys $t, t'(< t), H, J, W$ and T, which are given by:

$$\begin{aligned} J &= R^{-1}G^{-R}S^{-1}, \\ W &= G^{-R}S^{-1}, \\ T &= R^{-1}H^T, \end{aligned}$$

where H is the parity check matrix of the Goppa code in the usual form (1.1), in particular $GH^T = 0$. Alice's private keys are R and the matrix product SG.

Signature phase: The signature $\mathbf{s}$ of a k–bit message $\mathbf{m}$ is obtained by computing

$$\mathbf{s} = (\mathbf{e} + \mathbf{m}SG)R,$$

where $\mathbf{e}$ is a random n–bit error vector of Hamming weight $w(\mathbf{e}) \leq t'$, chosen by Alice. After the signature $\mathbf{s}$ is calculated, Alice sends the pair $(\mathbf{m}, \mathbf{s})$ to Bob.

Verification phase: The authenticity of a message–signature pair $(\mathbf{m}, \mathbf{s})$ can be checked in the following way:

(i). Calculate the syndrome

$$\mathbf{s}T = [(\mathbf{e} + \mathbf{m}SG)R]R^{-1}H^T = \mathbf{e}H^T + \mathbf{m}SGH^T = \mathbf{e}H^T.$$

(ii). Let $\mathbf{e}'$ be the result of the Berlekamp–Massey algorithm applied to the above syndrome $\mathbf{s}T$. It is possible to calculate this since the parity check matrix H

is in the usual form (1.1). If $t' < w(\mathbf{e}') < t$, [1] Bob rejects the signature. Note that $\mathbf{e} = \mathbf{e}'$, since $w(\mathbf{e}) \leq t'$.

(iii). Verify whether the relation $\mathbf{m} = \mathbf{s}J - \mathbf{e}'W$ holds as it should do, since

$$\mathbf{s}J = (\mathbf{e} + \mathbf{m}SG)\, R \left(R^{-1}G^{-R}S^{-1}\right) = \mathbf{e}G^{-R}S^{-1} + \mathbf{m} = \mathbf{e}W + \mathbf{m},$$

and $\mathbf{e}' = \mathbf{e}$. If this is the case, the signature is valid. Otherwise reject the signature.

Note that if in the setup phase a permutation matrix was taken for R, the Xinmei scheme reduces to the Rao–Nam private–key cryptosystem. It has been shown in [RN89] that the matrix SGR can be determined through majority voting if the Hamming weight of the n–bit error vector e is not in the neighborhood of $n/2$.

3.2.2 Some weaknesses in the Xinmei scheme

As mentioned in the introduction, the Xinmei scheme is vulnerable to several types of attacks. In the following these attacks will be surveyed and an analysis of why the Xinmei scheme is susceptible to them will be presented.

- *Homomorphism attack [HW92].* Since the error vectors $\mathbf{e}$ are revealed during the verification, Eve can choose two message–signature pairs satisfying $w(\mathbf{e}_1 + \mathbf{e}_2) \leq t'$. Then $\mathbf{s}_1 + \mathbf{s}_2$ will be a valid signature for the message $\mathbf{m}_1 + \mathbf{m}_2$. Obviously, the linearity of the signature in the Xinmei scheme results in this homomorphism attack. To thwart this attack, Harn and Wang suggested modifying the Xinmei scheme with a hash function h by setting $\mathbf{s} = (\mathbf{e} + h(\mathbf{m})SG)R$.

- *Chosen–plaintext attacks [AW92b].* If Eve is able to get $n+1$ different pairs of signatures and error patterns for the same message m in which n signatures are linearly independent, Eve can obtain the secret matrix R using the relation $D = ER$ where D and E are the $n \times n$ matrices with as i^{th} row $\mathbf{s}_i$ respectively $\mathbf{e}_i$ $(1 \leq i \leq n)$.

  Once R is known, Eve can obtain the other secret key SG through the following chosen–plaintext attack: suppose Eve has obtained k message–signature pairs for a set of linearly independent messages. Using the error patterns from the verification procedure, Eve can calculate SG from the equation $E' = MSG$ where E' and M are the $k \times n$ matrices with as i^{th} row $\mathbf{s}_i R^{-1} - \mathbf{e}_i$ respectively $\mathbf{m}_i$.

  The linearity of the signature enables Eve to successfully recover the secret keys R and SG in the above chosen–plaintext attacks. The knowledge of the error patterns also plays an important role in this attack. In the Xinmei scheme the random error vector is used to improve the security of scheme. Unfortunately,

[1] The original equations are $2t' - t > w(\mathbf{e}') > t$. Obviously, they are wrong because $t > t'$.

the leakage of the error vector results in the failure of the Xinmei scheme. To defeat the above attack, Alabbadi and Wicker suggested to introduce a hash function $h(x, y)$ into the signature scheme [AW93]. In their scheme $h(x, y)$ is used to hash the k–bit message $\mathbf{m}$ and the n–bit error vector $\mathbf{e}$ to replace the k–bit message in the signature generation of the Xinmei scheme.

In addition, Alabbadi and Wicker proved that the Harn–Wang scheme is susceptible to a known–plaintext attack [AW92a].

- *Directly recovering the secret keys from the public keys.* In the above attacks, Eve can calculate the signers secret keys from some triplets of messages, signatures and error patterns. Van Tilburg [Til92] also showed that the secret keys in the Xinmei scheme can be recovered directly from the public keys. Since G and H^T are orthogonal matrices, one can find a so–called *analogous* generator matrix $\tilde{G} = QG$ where Q is an unknown nonsingular $k \times k$ matrix. Following this, an *analogous* scrambling matrix $\tilde{S}$ can be obtained by inverting $\tilde{G}W = QGG^{-R}S^{-1} = QS^{-1} = (\tilde{S})^{-1}$. The original secret key SG then follows from $\tilde{S}\tilde{G} = SQ^{-1}QG = SG$. Finally R can be recovered from the equation $[J\tilde{S}\tilde{G}|T] = R^{-1}[W\tilde{S}\tilde{G}|H^T]$. Van Tilburg proved that $[W\tilde{S}\tilde{G}|H^T]$ has rank n. Thus the Xinmei scheme can be totally broken. The same attack also applies to the Harn–Wang scheme.

  Alabbadi and Wicker also tried to recover G from H and they estimated that the search is infeasible because it has complexity $O(k!)$ [AW92b].

  Without question, it is the redundancy in the public keys that results in the above attack. However, in order for Bob to check the validity of a signature, Alice has to publish some necessary public keys. Firstly Bob needs to have the ability to decode the signature. Thus the public key has to include some information about the parity check matrix. In the Xinmei scheme and also in other schemes, Bob is supposed to have the ability to recover the error pattern from the signature by means of the Berlekamp–Massey algorithm. However, the Berlekamp–Massey algorithm requires the parity check matrix to be in the usual form (1.1) [MS77]. Thus the parity check matrix has to be public, if either Euclid's or the Berlekamp–Massey algorithm is used for decoding.

  Furthermore, Bob needs to recover the message, whether in hashed form or not, from the signature using some public keys. These public keys and the verification procedure undoubtedly leak information about the secret keys.

- *Potential threats from analogous matrices.* Is it possible for Bob to completely defeat forgery attacks by recovering the message and checking if it is equal to the received message in the Xinmei scheme and other schemes [AW93; AW95]? Some potential threats from analogous matrices of secret keys are explored in this chapter.

  Firstly, the generator matrix G is the most important secret key in the Xinmei scheme and other schemes [AW93; AW95]. Even though Eve knows the

parameters (length n, dimension k and minimum distance d) of the code used in these schemes, it is still difficult for her to find G. For each $[n, k]$ binary linear code, there are $(2^k-1)(2^k-2)\cdots(2^k-2^{k-1})$ different generator matrices. As a secret key the generator matrix G is protected by two nonsingular random matrices S and R against direct calculation by the attacker (by using the public keys and the verification procedure).

Different generator matrices define different mappings from messages to signatures. But it is difficult to design a verification procedure which can check whether the signature satisfies the real mapping. This is because the real mapping is not known by either Bob or Eve (otherwise the scheme would be broken). However, Eve can obtain an analogous generator matrix $\tilde{G}$ from $\tilde{G}H^T = 0_{k\times(n-k)}$ (where 0 is the all-zero matrix) because she knows the parity check matrix H. This analogous matrix $\tilde{G}$ can be found in polynomial time. Then Eve can use $\tilde{G}$ to forge a signature. It is possible for the forged signature to pass the verification procedure because all items related to $\tilde{G}$ in the signature usually can be canceled in the procedure of calculating the syndrome. In addition, this can also happen to other secret keys. Thus it is possible for Eve to forge a signature which can pass the other checks in the verification procedure.

In Section 3.5 this method will be used to break the Alabbadi–Wicker scheme.

3.3 The Alabbadi–Wicker scheme

The three phases of the Alabbadi–Wicker scheme can be described as follows:

Setup phase: in the setup phase, each user chooses a t–error correcting binary irreducible Goppa code C with length $n = 2^m$ and dimension k. The code is described by an irreducible polynomial $G(z)$ of degree t and coefficients in $\mathbb{F}_{2^m}$. Alice then selects a $k \times n$ binary generator matrix G and a $(n-k)\times n$ binary parity check matrix H for the chosen code. She then chooses two $k \times n$ binary matrices W and V such that

$$G = W + V, \tag{3.1}$$

where the rank of W is k. This means that there exists an $n \times k$ binary right–inverse matrix W^{-R} such that

$$WW^{-R} = I_k, \tag{3.2}$$

where I_k is the identity matrix. The matrix W^{-R} is chosen such that GW^{-R} has nonzero rank $k' < k$. Then she generates a nonsingular $n \times n$ binary matrix R. The final step of initializing the signature scheme is the computation of the following

matrices:

$$H' = R^{-1}H^T, \tag{3.3}$$
$$W' = R^{-1}W^{-R}, \tag{3.4}$$
$$W'' = W^{-R}GW^{-R}. \tag{3.5}$$

Alice then publishes $G(z)$, W^{-R}, H', W', W'', t and $t' < t$ as public keys. The private key consists of the matrices V, W, G, $W^{-R}G$, and R.

In addition a hash function $h : \mathbb{F}_{2^k} \times \mathbb{F}_{2^n} \to \mathbb{F}_{2^k}$ is made available to all users of the system.

Signature phase: suppose user Alice wants to sign a k–bit message $\mathbf{m}$. She then selects two binary vectors at random: a n–bit vector $\mathbf{e}$ of weight t', and a k–bit vector $\mathbf{r}$ of arbitrary but nonzero weight. The signature $(\mathbf{s}, \mathbf{x})$ of the message $\mathbf{m}$ is then computed as follows:

$$\begin{aligned} \mathbf{x} &= (\mathbf{r}G + h(\mathbf{m},\mathbf{e})V)\,R, \\ \mathbf{s} &= \left(\mathbf{e} + h(\mathbf{m},\mathbf{e})W + \mathbf{x}W^{-R}G\right)R. \end{aligned}$$

Finally, Alice transmits the triplet $\mathbf{x}$, $\mathbf{s}$, and $\mathbf{m}$ to Bob.

Verification phase: Bob gets a signature $(\mathbf{x}, \mathbf{s})$ along with the message $\mathbf{m}$. The signature validation is then performed as follows:

(i). The following expression is computed:

$$\begin{aligned} \mathbf{x} + \mathbf{s} &= \left(\mathbf{r}G + h(\mathbf{m},\mathbf{e})V + \mathbf{e} + h(\mathbf{m},\mathbf{e})W + \mathbf{x}W^{-R}G\right)R \\ &= \left(\mathbf{r}G + h(\mathbf{m},\mathbf{e})G + \mathbf{e} + \mathbf{x}W^{-R}G\right)R. \end{aligned}$$

(ii). The syndrome is calculated:

$$\begin{aligned} (\mathbf{x} + \mathbf{s})H' &= \left(\mathbf{r}G + h(\mathbf{m},\mathbf{e})G + \mathbf{e} + \mathbf{x}W^{-R}G\right)RR^{-1}H^T \\ &= \mathbf{e}H^T. \end{aligned}$$

(iii). The Berlekamp–Massey algorithm is applied to the above syndrome to obtain the error vector $\mathbf{e}$. If $w(\mathbf{e}) \neq t'$, Bob rejects the signature.

(iv). The hash $h(\mathbf{m}, \mathbf{e})$ of the message and the error vector is recovered from $\mathbf{x}$, $\mathbf{e}$, and $\mathbf{s}$ by computing

$$\begin{aligned} \mathbf{s}W' + \mathbf{x}W'' + \mathbf{e}W^{-R} &= \mathbf{s}R^{-1}W^{-R} + \mathbf{x}W^{-R}GW^{-R} + \mathbf{e}W^{-R} \\ &= \left(\left(\mathbf{e} + h(\mathbf{m},\mathbf{e})W + \mathbf{x}W^{-R}G\right)R\right)R^{-1}W^{-R} + \\ &\quad + \mathbf{x}W^{-R}GW^{-R} + \mathbf{e}W^{-R} \\ &= \mathbf{e}W^{-R} + h(\mathbf{m},\mathbf{e}) + \mathbf{x}W^{-R}GW^{-R} + \\ &\quad + \mathbf{x}W^{-R}GW^{-R} + \mathbf{e}W^{-R} \\ &= h(\mathbf{m},\mathbf{e}). \end{aligned}$$

(v). Finally, Bob compares the result of the above computation to $h(\mathbf{m}, \mathbf{e})$, which he can calculate himself. If they are equal, he accepts the signature as valid.

Apparently, the proposers of this scheme overlooked the fact that step (iii) of this verification procedure will not work, since applying the Berlekamp–Massey algorithm requires the parity check matrix to be in the usual form (1.1) [MS77, Section 12.9]. Obviously, if this step is to work, Alice has to select H to be in the usual form, and should thus make H public.

3.4 Modifying the Alabbadi–Wicker scheme

Since the verification phase of the Alabbadi–Wicker scheme will not work unless the parity check matrix H is in the usual form (1.1), and thus public, a revision of the scheme is needed. This revision is made here by modifying the three phases as follows:

Setup phase: Alice first calculates the public keys as in the original Alabaddi-Wicker scheme, as described in the previous section. Suppose that the order of coordinates in $\mathbb{F}_{2^m}$ is chosen to be canonical (it could also be chosen by each user individually, but it would then have to be part of the public key as well). Furthermore, let H_Γ be the parity check matrix of the chosen Goppa code in the usual form (1.1). Since the chosen matrix H is also a parity check matrix, Alice can find a non-singular matrix M such that

$$H_\Gamma = MH.$$

Alice then publishes $G(z)$, W^{-R}, H', W', W'', t and $t' < t$ as public keys. The private key consists of the matrices V, W, G, $W^{-R}G$, R and M.

Signature phase: suppose Alice wants to sign a k–bit message $\mathbf{m}$. She then selects two binary vectors at random: a n–bit vector $\mathbf{e}$ of weight t' and a k–bit vector $\mathbf{r}$ of arbitrary but nonzero weight. The signature $(\mathbf{s}, \mathbf{x})$ of the message $\mathbf{m}$ is then computed in the following steps:

(i). Alice finds a solution $\mathbf{f}$ to the equation $\mathbf{f}H_\Gamma^T = \mathbf{e}H_\Gamma^T M^T$ directly. This is possible, since a solution surely exists (every syndrome occurs among all vectors of $\mathbb{F}_2^n$ since we are dealing with a linear code). However, the solution will obviously not be unique since no requirement on the weight is given. Thus $\mathbf{f}$ satisfies

$$\mathbf{f}H^T = \mathbf{e}H_\Gamma^T.$$

(ii). Alice computes

$$\begin{aligned} \mathbf{x} &= (\mathbf{e} + \mathbf{f} + \mathbf{r}G + h(\mathbf{m}, \mathbf{e})V)\, R, \\ \mathbf{s} &= \left(\mathbf{e} + h(\mathbf{m}, \mathbf{e})W + \mathbf{x}W^{-R}G\right) R. \end{aligned}$$

Finally, Alice transmits the signature ($\mathbf{x}$, $\mathbf{s}$) and the message $\mathbf{m}$ to Bob.

Verification phase: Bob gets a signature $(\mathbf{x}, \mathbf{s})$ along with the message $\mathbf{m}$. He validates the signature in the following steps:

(i). The following expression is computed:

$$\begin{aligned} \mathbf{x}+\mathbf{s} &= \left(\mathbf{r}G + h(\mathbf{m},\mathbf{e})V + \mathbf{f} + h(\mathbf{m},\mathbf{e})W + \mathbf{x}W^{-R}G\right) R \\ &= \left(\mathbf{r}G + h(\mathbf{m},\mathbf{e})G + \mathbf{f} + \mathbf{x}W^{-R}G\right) R. \end{aligned}$$

(ii). The syndrome is calculated:

$$\begin{aligned} (\mathbf{x}+\mathbf{s})H' &= \left(\mathbf{r}G + h(\mathbf{m},\mathbf{e})G + \mathbf{f} + \mathbf{x}W^{-R}G\right) RR^{-1}H^T \\ &= \mathbf{f}H^T = \mathbf{e}H_\Gamma^T. \end{aligned}$$

(iii). The Berlekamp–Massey algorithm is applied to the above syndrome. The result of the algorithm will be the error vector $\mathbf{e}$. If $w(\mathbf{e}) \neq t'$, Bob rejects the signature.

(iv). The hash $h(\mathbf{m}, \mathbf{e})$ is recovered from the signature ($\mathbf{x}$, $\mathbf{s}$) and the error vector $\mathbf{e}$ by computing

$$\begin{aligned} \mathbf{s}W' + \mathbf{x}W'' + \mathbf{e}W^{-R} &= \mathbf{s}R^{-1}W^{-R} + \mathbf{x}W^{-R}GW^{-R} + \mathbf{e}W^{-R} \\ &= \mathbf{e}W^{-R} + h(\mathbf{m},\mathbf{e}) + \mathbf{x}W^{-R}GW^{-R} + \\ &\quad + \mathbf{x}W^{-R}GW^{-R} + \mathbf{e}W^{-R} \\ &= h(\mathbf{m},\mathbf{e}). \end{aligned}$$

(v). Finally, Bob calculates the value $h(\mathbf{m}, \mathbf{e})$ and compares it to the last result. If they are equal, he accepts the signature as valid.

3.5 Cryptanalysis of the Alabbadi–Wicker scheme

Alabbadi and Wicker claimed that their scheme is resistant to the attacks that proved successful against the Xinmei scheme and also to other attacks. First the resistance of the Alabbadi–Wicker against the attacks described in Section 3.2 will be investigated.

3.5.1 Resistance of the Alabbadi–Wicker scheme against attacks

The Alabbadi–Wicker scheme looks similar to the Xinmei scheme (with a hash function) if the signatures $\mathbf{x}$ and $\mathbf{s}$ are added:

$$\begin{aligned} \mathbf{x}+\mathbf{s} &= \left(\mathbf{r}G + h(\mathbf{m},\mathbf{e})G + \mathbf{e} + \mathbf{x}W^{-R}G\right) R \\ &= \left(\mathbf{e} + [\mathbf{r} + h(\mathbf{m},\mathbf{e}) + \mathbf{x}W^{-R}]G\right) R \\ &= \left(\mathbf{e} + h'(\mathbf{m},\mathbf{e})S'G\right) R. \end{aligned}$$

Note that the modified scheme retains this property. Alabbadi and Wicker adopted different methods to defeat the attacks which are successful against the Xinmei scheme. First a hash function h is applied to both the message $\mathbf{m}$ and the error vector $\mathbf{e}$ to prevent the homomorphism attack in Section 3.2.2. Furthermore a k–bit vector $\mathbf{r}$ of arbitrary (but nonzero) weight has been introduced to the signature $\mathbf{x}$. Bob cannot solve $\mathbf{r}$ from the signature $\mathbf{x}$, and only Alice knows it. Thus the Alabbadi–Wicker scheme defeats both the chosen–plaintext and the known–plaintext attack in Section 3.2.2. Lastly the generator matrix G has been split into two matrices W and V and the public keys (namely W^{-R}, W' and W'') include only partial information about G (of course, the null space of G is determined by both H and the polynomial $G(z)$). So at least it is difficult to directly derive the secret key G from the public keys. A total break appears to be infeasible, primarily because the public keys do not completely describe G (this is true because the matrix $W'' = W^{-R}GW^{-R}$ is not of full rank). Eve thus seems to be forced to perform an exhaustive search through all possible generator matrices for the code C.

However, the Alabbadi–Wicker scheme is not as secure as they claimed. They state that their digital signature scheme derives its security from the complexity of three problems: the decoding of general linear error–correcting block codes, the factoring of large matrices, and the derivation of a matrix from its right inverse. In the following sections a universal forgery attack against the Alabbadi–Wicker scheme will be presented.

3.5.2 A universal forgery of the Alabbadi–Wicker scheme

In [AW95], Alabbadi and Wicker analyzed the possibility of a universal forgery, i.e. being able to sign an arbitrary message given only the public keys. Even though their attack did not succeed, it did motivate the following attack using analogous matrices.

Recovering the parity check matrix H

Even if H is not in the usual form, it is still possible for Eve to recover H from the public keys and the verification procedure. From the second and the third step in the verification of a signature the following equation can be obtained:

$$(\mathbf{x} + \mathbf{s})H' = \mathbf{e}H^T, \tag{3.6}$$

where $\mathbf{x}$, $\mathbf{s}$ and $\mathbf{e}$ and H' are known to Eve. Note that the matrix dimensions of H^T and H' are the same.

Suppose Eve is able to obtain signatures with n independent error vectors $\mathbf{e}_i$ and thus the corresponding $(\mathbf{x}_i + \mathbf{s}_i)H'$ $(1 \leq i \leq n)$. Then she can solve the parity check matrix H^T from the n Equations (3.6) by setting $H^T = E^{-1}(X + S)H'$ where E, X and S are the matrices with as i^{th} row respectively $\mathbf{e}_i$, $\mathbf{x}_i$ or $\mathbf{s}_i$ $(1 \leq i \leq n)$. The complexity of solving H^T in this way is only $O(n^3)$.

Calculating an analogous matrix $\tilde{R}$

After Eve has successfully recovered the parity check matrix H, she can try to find the nonsingular matrix R according to the following method. From (3.3) and (3.4) the following expression follows:

$$[H'|W'] = R^{-1}[H^T|W^{-R}], \tag{3.7}$$

where H' and H^T are $n \times (n-k)$ matrices and W' and W^{-R} are $n \times k$ matrices. So $[H'|W']$, $[H^T|W^{-R}]$ and R^{-1} are $n \times n$ matrices. Alabaddi and Wicker proved that $[H^T|W^{-R}]$ is a singular matrix, so Eve cannot find R^{-1} from Equation (3.7). Even so, she can obtain an analogous matrix $\tilde{R}^{-1}$ which can also be used to forge a signature.

Even though $[H^T|W^{-R}]$ is not a full rank matrix, Eve can obtain a nonsingular row transformation matrix $\tilde{R}^{-1}$ from (3.7), which satisfies the following equations:

$$H' = \tilde{R}^{-1}H^T, \tag{3.8}$$
$$W' = \tilde{R}^{-1}W^{-R}. \tag{3.9}$$

Of course, it would be best if the matrix $\tilde{R}^{-1}$ is equal to R^{-1}. However, Eve has no way of knowing this. The attack still goes through, even if the two matrices are not equal. Eve may calculate the inverse matrix $\tilde{R}$ from $\tilde{R}^{-1}$ in polynomial time. The matrix $\tilde{R}$ will play an important role in the following universal forgery.

Universal Forgery for the Alabaddi-Wicker scheme

Eve will now calculate an analogous generator matrix $\tilde{G}$ which should satisfy

$$\tilde{G}H^T = 0_{k\times(n-k)}. \tag{3.10}$$

Note that $\tilde{G}$ is in general not equal to G, the generator matrix used by Alice (just like with the above $\tilde{R}$), but again this does not matter.

Since W^{-R} is a public key, Eve can calculate a left inverse $\tilde{W}$ of W^{-R}, so

$$\tilde{W}W^{-R} = I_k. \tag{3.11}$$

Then Eve calculates $\tilde{V} = \tilde{G} + \tilde{W}$. Again, in general $V \neq \tilde{V}$ and $W \neq \tilde{W}$.

Since W'' and W^{-R} are public keys, Eve can calculate a matrix $\tilde{Y}$ by simple algebraic means which satisfies the following equation.

$$W'' = W^{-R}GW^{-R} = \tilde{Y}W^{-R}. \tag{3.12}$$

Now Eve is able to forge the signature of any message $\mathbf{m}$. According to Alabbadi–Wicker scheme, an n–bit error vector $\mathbf{e}$ of weight t' is chosen at random. Since $\mathbf{r}$ is only used to protect G from the attacks in Section 2, it is discarded (after all, she is trying to forge a signature, not to obscure G).

To obtain a signature for the message $\mathbf{m}$, Eve first calculates the vector $\mathbf{x}$ of the signature pair $(\mathbf{x}, \mathbf{s})$ from the implicit equation

$$\mathbf{x} = \left(h(\mathbf{m}, \mathbf{e})\tilde{V} + \mathbf{x}\tilde{Y} \right) \tilde{R}. \tag{3.13}$$

Then she can calculate $\mathbf{s}$ from

$$\mathbf{s} = \left(\mathbf{e} + h(\mathbf{m}, \mathbf{e})\tilde{W} + \mathbf{x}\tilde{Y} \right) \tilde{R}. \tag{3.14}$$

Eve can now send the triple $(\mathbf{m}, \mathbf{x}, \mathbf{s})$ as a message with a forged signature to Bob.

This triple will be shown to pass the signature validation of the Alabbadi–Wicker scheme. Bob follows the verification procedure to get

$$\begin{aligned} (\mathbf{x}+\mathbf{s})H' &= \left(h(\mathbf{m}, \mathbf{e})\tilde{V} + \mathbf{e} + h(\mathbf{m}, \mathbf{e})\tilde{W} \right) \tilde{R}H' \\ &= \left(h(\mathbf{m}, \mathbf{e})\tilde{G} + \mathbf{e} \right) \tilde{R}\tilde{R}^{-1}H^T \\ &= \mathbf{e}H^T. \end{aligned}$$

It is obvious that the signature $(\mathbf{x}, \mathbf{s})$ will pass the first three steps of the verification phase. Now Bob looks at the fourth step:.

$$\begin{aligned} \mathbf{s}W' + \mathbf{x}W'' + \mathbf{e}W^{-R} &= \mathbf{s}R^{-1}W^{-R} + \mathbf{x}W'' + \mathbf{e}W^{-R} \\ &= \mathbf{e}W^{-R} + h(\mathbf{m}, \mathbf{e}) + \mathbf{x}\tilde{Y}W^{-R} + \mathbf{x}W'' + \mathbf{e}W^{-R} \\ &= \mathbf{e}W^{-R} + h(\mathbf{m}, \mathbf{e}) + \mathbf{x}W'' + \mathbf{x}W'' + \mathbf{e}W^{-R} \\ &= h(\mathbf{m}, \mathbf{e}). \end{aligned}$$

So the forged signature has passed all steps of the verification and Bob will accept the signature as a valid one (from Alice).

Example

As an example, Eve will now forge a signature using the Alabaddi–Wicker scheme. The same $[6, 3, 3]$–code that Alabbadi and Wicker chose for their example in [AW95] will be used here. The public and private keys for their example of the scheme are:

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}, H = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}, W = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \end{pmatrix},$$

$$V = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, R = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, R^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$W^{-R} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}, H' = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, W' = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, W'' = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Suppose that Eve has recovered the parity check matrix H as described in Section 3.5.2 (or otherwise). Now she will calculate the analogous matrices $\tilde{R}, \tilde{G}, \tilde{W}$ and $\tilde{Y}$ from the public keys of the Alabbadi–Wicker scheme.

First Eve calculates $\tilde{R}^{-1}$ from Equation (3.7):

$$\tilde{R}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that many choices are possible here since $\left(H^T | W^{-R}\right)$ is not a full–rank matrix. Now she inverts this matrix to get $\tilde{R}$ (note that $\tilde{R}^{-1}$ is a full–rank matrix, so this is possible).

$$\tilde{R} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The matrices $\tilde{R}$ and $\tilde{R}^{-1}$ are indeed not equal to R and R^{-1}. However, this does not effect Eve's ability to forge a signature.

Similarly, Eve calculates the analogous matrices $\tilde{G}, \tilde{W}$ and $\tilde{Y}$ from equations (3.10), (3.11) and (3.12):

$$\tilde{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}, \tilde{W} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \tilde{Y} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

The matrix $\tilde{V}$ follows from the equation $\tilde{G} = \tilde{W} + \tilde{V}$:

$$\tilde{V} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Suppose Eve wants to sign the message $\mathbf{m} = (001)$ and selects the error vector $\mathbf{e} = (000001)$. As in [AW95], the hash value is taken to be $h(\mathbf{m}, \mathbf{e}) = (011)$. According to Equations (3.13) and (3.14) of the forgery steps, Eve calculates the signature pair $(\mathbf{x}, \mathbf{s})$ of the message $\mathbf{m}$ as $\mathbf{x} = (101111)$ and $\mathbf{s} = (111100)$.

The verification of this signature pair goes as follows: in the first step, $\mathbf{x} + \mathbf{s} = (010011)$. Then the syndrome is calculated to be $(010011)H' = (001)$ in the second step. Now suppose that Bob is able to recover the error vector from the above syndrome. Clearly, the recovered error vector will be $\mathbf{e} = (000001)$, since the last column of H is exactly $(001)^T$. Finally, the expression

$$\begin{aligned} \mathbf{s}W' + \mathbf{x}W'' + \mathbf{e}W^{-R} &= (111100)W' + (101111)W'' + (000001)W^{-R} \\ &= (111) + (001) + (101) = (011). \end{aligned}$$

Clearly, this is equal to $h\big((001), (000001)\big)$, and Bob will accept the signature as valid.

3.5.3 Cryptanalyzing the modified Alabbadi–Wicker scheme

Since the error vector $\mathbf{f}$ is not revealed to Bob, and thus also remains hidden from Eve, the recovery of H is no longer feasible. However, a universal forgery is still possible. Eve should first construct an analogous (non-degenerate) matrix $\tilde{R}$ by finding a solution to the equation

$$\tilde{R}W' = W^{-R}.$$

Then she should find analogous matrices $\tilde{W}$ and $\tilde{Y}$ as described in Section 3.5.2.

Suppose Eve wants to forge a signature of the message $\mathbf{m}$. To that end, she picks a random $\mathbf{e}$ of weight $w(\mathbf{e}) = t'$, and then calculates a solution $\mathbf{f}$ to the equation $\mathbf{e}H_\Gamma^T = \mathbf{f}\tilde{R}H'$ directly. Note that a solution always exists, since all syndromes are taken by the vectors in $\mathbb{F}_2^n$. Since this means solving a system of k linear equations for n variables, it is possible to do this effectively. Note that H_Γ is effectively public, since both $G(z)$ and the order of coordinates $\{\gamma_i\}$ of $\mathbb{F}_{2^m}$ are public. She then sets

$$\mathbf{x} = (\mathbf{e} + \mathbf{f} + h(\mathbf{m}, \mathbf{e})\tilde{W} + \mathbf{x}Y)\tilde{R}.$$

$$\mathbf{s} = (\mathbf{e} + h(\mathbf{m}, \mathbf{e})\tilde{W} + \mathbf{x}Y)\tilde{R}.$$

Bob will accept this pair as a valid signature of $\mathbf{m}$, since it passes the second step of the verification procedure:

$$\begin{aligned} (\mathbf{x} + \mathbf{s})H' &= (\mathbf{e} + \mathbf{f} + h(\mathbf{m}, \mathbf{e})\tilde{W} + \mathbf{x}Y + \mathbf{e} + h(\mathbf{m}, \mathbf{e})\tilde{W} + \mathbf{x}Y)\tilde{R}H' \\ &= \mathbf{f}\tilde{R}H' = \mathbf{e}H_\Gamma^T, \end{aligned}$$

as well as the last step of the verification:

$$\begin{aligned} \mathbf{s}W' + \mathbf{x}W'' + \mathbf{e}W^{-R} &= (\mathbf{e} + h(\mathbf{m}, \mathbf{e})\tilde{W} + \mathbf{x}Y)\tilde{R}W' + \mathbf{x}W'' + \mathbf{e}W^{-R} \\ &= \mathbf{e}W^{-R} + h(\mathbf{m}, \mathbf{e})\tilde{W}W^{-R} + \mathbf{x}YW^{-R} + \mathbf{x}W'' + \mathbf{e}W^{-R} \\ &= h(\mathbf{m}, \mathbf{e}) + \mathbf{x}W'' + \mathbf{x}W'' = h(\mathbf{m}, \mathbf{e}). \end{aligned}$$

Thus Eve can construct a signature for any message **m**. Note that this (second) universal forgery can be slightly adapted to apply to the unmodified Alabbadi–Wicker scheme as well, given that the decoding step in the verification is made possible.

3.6 Discussion

The above universal forgery makes use of analogous matrices such as $\tilde{G}, \tilde{W}$ and $\tilde{Y}$. There are other possible drawbacks of the Alabbadi–Wicker scheme which shall not be discussed here. The aim of this discussion is to find the reason behind the above universal forgery. This may help to improve the Alabbadi–Wicker scheme or design new signature schemes using error–correcting codes.

From the description of the universal forgery attack, it seems very difficult to prevent this kind of forgery. The designers of the scheme were hoping to hide the real secret key G between its analogous matrices using the simple fact that there are many analogous matrices. Thus it will be infeasible for Eve to find the original map between the message and signature, which is defined by the secret key G. However, they did not realize that Bob does not have the ability to check this original map between the message and its signature. Even though the analogous matrices $\tilde{G}, \tilde{W}$ and $\tilde{Y}$ define a different map, Bob cannot detect it because he does not know the secret key.

In addition, the universal forgery in Section 3.5.2 also shows that neither the Alabbadi–Wicker scheme nor the Xinmei scheme have the important property that it should be infeasible for Eve to find a signature algorithm that passes the verification step given only this verification algorithm and the public keys. In fact, there are many such signature algorithms she can come up with. It is very difficult for the designer to defeat them when he designs a digital signature scheme using the same method as the Xinmei scheme.

Up to now all attempts to design a secure digital signature scheme based in this way on error–correcting codes have failed. Why is it so hard to design such a scheme? This is because these signature schemes do not really depend on the hardness of the decoding problem of general error–correcting codes.

Van Tilburg [Til94] showed that signature schemes, where the security is based only on the bounded, hard–decision decoding problem for linear codes, do not exist. However Kabatianskii, Krouk and Smeets proposed a digital signature scheme based on random error–correcting codes in 1997 [KKS97]. They exploited a hardly known fact that for every linear code the set of its correctable syndromes contains a linear subspace of relatively large dimension L. Unfortunately their scheme can only be used once. Even so, it does give some new ideas to further explore the use of this intractability feature of the decoding problem.

3.7 Conclusion

In this chapter the security of digital signature schemes based on error–correcting codes was discussed and some existing weaknesses in the Xinmei scheme were surveyed. Potential threats from matrices that have the same properties as some of the secret matrices, so–called analogous matrices, were explored as well. As an example a universal forgery of the Alabbadi–Wicker scheme was presented.

Chapter 4

Two families of Mersenne–like primes

Summary. In this chapter two families of numbers are introduced which can efficiently be tested for primality. These families naturally extend the Mersenne numbers to the area of elliptic curves. The first few primes in these families are also presented and compared to the generalized Wagstaff conjecture. This chapter is based on joint work with Peter Beelen [BD03].

4.1 Introduction

Two families of prime numbers, not unlike the Mersenne primes, for which an efficient primality test exists, will be studied in this chapter. For one of these families a primality test which is as fast as the test for Mersenne numbers is presented. For Mersenne numbers this test has led to the discovery of very big primes. In fact, the largest explicitly known prime number at this moment (April 2003) is a Mersenne prime. For a list of large known prime numbers see [YC03].

4.2 Testing for primality

In this section some known results concerning primality testing will be reviewed. For an exposition of these and other tests see e.g. [Rib96].

Proposition 4.2.1 (Proth) *Let n be an odd natural number and suppose that $n-1 = 2^e R$, where R is odd and less than $\sqrt{n}$. If a number a exists such that $a^{(n-1)/2} \equiv -1 \pmod{n}$, then n is a prime.*

Proof: Let p be the smallest prime divisor of n and let b be the order of $a \pmod{p}$. Thus b is a divisor of $p-1$. Since $a^{(n-1)/2} \equiv -1 \pmod{n}$ the greatest common divisor of n and $a^{(n-1)/2} - 1$ is equal to one. Thus the greatest common divisor of p and $a^{(n-1)/2} - 1$ must also be equal to one. But $a^b \equiv 1 \pmod{p}$, and since

$a^{(n-1)/2} \not\equiv 1 \pmod{n}$, b cannot divide $(n-1)/2 = 2^{e-1}R$. Since $a^{n-1} \equiv 1 \pmod{n}$, b divides $n-1 = 2^eR$. The conclusion is that 2^e divides b and hence that 2^e divides $p-1$. Thus $p \geq 2^e + 1 > \sqrt{n}$ and a contradiction arises. So one can conclude that n has no prime factors. □

Proth's test can be extended:

Theorem 4.2.2 ([HW79, Theorem 102]) *Let $n = 2^eR + 1$ with $e \geq 2$ and $R < 2^e$. Furthermore, let n satisfy $(\frac{n}{p}) = -1$ for some odd number p. Then a necessary and sufficient condition for n to be prime is that $p^{(n-1)/2} \equiv -1 \pmod{n}$.*

Proof: If n is prime, then $p^{(n-1)/2} \pmod{n} \equiv (\frac{n}{p}) = -1$ per definition of the Jacobi symbol. Since

$$\sqrt{n} = \sqrt{2^eR + 1} > \sqrt{R^2 + 1} > R,$$

the other half of the theorem follows directly from Proposition (4.2.1). □

Note that Hardy and Wright state a slightly different version of this theorem. They only look at odd prime numbers p. Here the Jacobi symbol is used, which coincides with the Legendre symbol if p is a prime.

More general than Proth's test is the following well-known theorem:

Theorem 4.2.3 (Pocklington) *Let n be a natural number and suppose that $n - 1 = FR$, where F has known prime factorization $F = p_1^{e_1} \cdots p_k^{e_k}$ and where R is relatively prime to F and less than $\sqrt{n}$. If for each $1 \leq i \leq k$ there exists a number a_i such that $a_i^{n-1} \equiv 1 \pmod{n}$ and $\gcd(a_i^{(n-1)/p_i} - 1, n) = 1$, then n is a prime.*

Proof: The proof follows directly from the proof of Proposition 4.2.1 and the Chinese remainder theorem. □

Mersenne numbers are numbers of the form $M_n = 2^n - 1$. It is easy to see that $2^n - 1$ can only be a prime if the exponent n is a prime. One of the nice things about Mersenne prime numbers is that there exists a very efficient primality test for them, namely the Lucas-Lehmer primality test. See for example [Rib96, page 92] or [Bre89, page 27]. The test is described in the following proposition for later reference.

Proposition 4.2.4 *Consider the Lucas sequence $\{V_n(2, -2)\}$ (the indices (2,-2) will be omitted from now on) defined recursively by $V_0 = 2, V_1 = 2$ and $V_{j+1} = 2V_j + 2V_{j-1}$. Now M_n is prime if and only if M_n divides $V_{(M_n+1)/2}$.*

A proof of this proposition can be found in either of the aforementioned references.

4.3 Prime-generating elliptic curves

Let $e \geq 1$ be a natural number and $\mathcal{E}$ an elliptic curve defined over a finite field $\mathbb{F}_q$. The curve $\mathcal{E}$ over the extension field $\mathbb{F}_{q^e}$ will also be considered. More precisely, denote by $\mathcal{E}(\mathbb{F}_{q^e})$ the group of those points of $\mathcal{E}$ whose coordinates are contained in the field $\mathbb{F}_{q^e}$, and denote by E_e the cardinality of this group. The following theorem by Hasse and Weil gives some information about these numbers.

Theorem 4.3.1 (Hasse-Weil) *There exists an algebraic integer α, depending on the elliptic curve $\mathcal{E}$, with absolute value $\sqrt{q}$ such that the cardinality E_e of $\mathcal{E}(\mathbb{F}_{q^e})$ satisfies*

$$E_e = q^e + 1 - (\alpha^e + \overline{\alpha}^e)$$

for all natural numbers $e \geq 1$. This α is called a Frobenius eigenvalue *of the elliptic curve $\mathcal{E}$.*

For a proof of this theorem, see for example page 134 of [Sil86]. Note that $\mathcal{E}(\mathbb{F}_q)$ is a subgroup of $\mathcal{E}(\mathbb{F}_{q^e})$. Hence E_1 divides E_e for all e.

Definition 4.3.2 *Let $\mathcal{E}$ be an elliptic curve defined over a finite field $\mathbb{F}_q$. A* prime-generating elliptic curve *is a curve $\mathcal{E}$ such that there exist infinitely many prime numbers of the form E_e/E_1.*

It is unfortunately not known whether or not there exist prime-generating elliptic curves, but some good candidates will be given later. First, the possible values of e for which E_e/E_1 can be prime will be narrowed down.

Proposition 4.3.3 *Let $\mathcal{E}$ be an elliptic curve defined over a finite field $\mathbb{F}_q$. Suppose that E_e/E_1 is a prime. Then one of the following statements holds:*

(i) *e is prime,*

(ii) *$q = 2$ and $e \in \{4, 6, 9\}$*[1]*,*

(iii) *$q = 3$ and $e = 4$.*

Proof: Suppose that e is not a prime and let f be a non-trivial divisor of e. Then $\mathbb{F}_{q^f}$ is a non-trivial subfield of $\mathbb{F}_{q^e}$. Hence $\mathcal{E}(\mathbb{F}_{q^f})$ is a subgroup of $\mathcal{E}(\mathbb{F}_{q^e})$, which implies that E_f/E_1 divides E_e/E_1. The remaining problem is to show that, except for the last two cases mentioned in the proposition, this is a non-trivial divisor of E_e/E_1.

Since f is a non-trivial divisor of e, it can be assumed without loss of generality that f satisfies $\lceil \sqrt{e} \, \rceil \leq f \leq e/2$. Here $\lceil f \rceil$ denotes the ceiling function applied to f. Hence, if either $q \geq 3$ or $e \geq 5$,

$$E_f \leq (\sqrt{q^f} + 1)^2 \leq (\sqrt{q^{e/2}} + 1)^2 < (\sqrt{q^e} - 1)^2 \leq E_e.$$

[1] In [Kob01] two of these cases are missing.

To obtain the first and the last inequality, Hasse-Weil's theorem was used. If $q = 2$ and $e = 4$, the strict inequality becomes an equality. So for $q = 2$ either $E_2 < E_4$ or $E_2 = E_4 = 9$.

Now the case $E_f/E_1 = 1$ will be investigated. The relation

$$E_1 \leq (\sqrt{q} + 1)^2 < (\sqrt{q^f} - 1)^2 \leq E_f,$$

holds whenever $q = 2$ and $e > 9$, or $q = 3$ and $e > 4$, or $q = 4$ and $e > 4$, or $q \geq 5$. Here the relation $f \geq \lceil \sqrt{e} \rceil$ is used. Note that for $q = 2$ and $e = 8$ strict inequality holds as well, since in this case $f = 4$.

So far it has been proven that E_f/E_1 is a non-trivial divisor of E_e/E_1 whenever $q = 2$ and $e \notin \{4, 6, 9\}$, or $q = 3$ and $e \neq 4$, or $q = 4$ and $e \neq 4$, or $q \geq 5$.

The case $q = 4$ and $e = 4$ still has to be excluded. For E_2/E_1 to be a trivial divisor of E_4/E_1, a curve with $E_1 = E_2$ is necessary. It is not hard to see that in this case $\alpha = -2$ (from Hasse-Weil's theorem), and hence $E_4/E_1 = 25$, which is not a prime. □

Note that all of the above cases actually occur, as will be shown in Table 4.2 for $q = 2$. Further note that it is not true that whenever e is a prime the number E_e/E_1 is a prime as well. The situation is similar to the case of the Mersenne numbers. Not all elliptic curves are prime-generating, as is shown in the following example.

Example 4.3.4 Consider the curve defined over $\mathbb{F}_4$ having equation $Y^2 + Y = X^3 + \vartheta$, where ϑ is a primitive element of $\mathbb{F}_4$. In this case $E_1 = 1$ and hence $\alpha = 2$. This implies $\frac{E_e}{E_1} = (2^e - 1)^2$, which obviously never is a prime number.

Up to isomorphism there exist exactly 5 elliptic curves defined over $\mathbb{F}_2$. As a matter of fact their isomorphism classes can be characterized by E_1, the number of points defined over $\mathbb{F}_2$. From Hasse-Weil's theorem the relation $1 \leq E_1 \leq 5$ immediately follows. In Table 4.1 representatives of these isomorphism classes $\mathcal{E}$ and their Frobenius eigenvalues α will be listed.

E_1	Weierstrass equation of $\mathcal{E}$	α
1	$Y^2 + Y = X^3 + X + 1$	$1 \pm i$
2	$Y^2 + XY = X^3 + X + 1$	$1/2 \pm i\sqrt{7}/2$
3	$Y^2 + Y = X^3$	$\pm i\sqrt{2}$
4	$Y^2 + XY = X^3 + 1$	$-1/2 \pm i\sqrt{7}/2$
5	$Y^2 + Y = X^3 + X$	$-1 \pm i$

TABLE 4.1: The non-isomorphic elliptic curves over $\mathbb{F}_2$.

For which values of e the number E_e/E_1 is a (probable) prime has been investigated for these five curves. Probable prime means that the number passed two probabilistic primality tests: both the Mathematica function PrimeQ and the quadratic Frobenius test as introduced in [Gra01]. The results will be listed in Table 4.2. This table extends the table given in [Kob01] for the five curves defined over $\mathbb{F}_2$.

Whether E_e/E_1 is a (probable) prime has been checked for all values of e less than or equal to 17389. Since the characteristic is 2, the only candidates for e are prime numbers and the numbers in the set $\{4, 6, 9\}$ (see Proposition 4.3.3).

The case $E_1 = 1$ will be examined more closely in the next section.

E_1	e
1	$2, 3, 5, 7, 11, 19, 29, 47, 73, 79, 113, 151, 157, 163, 167, 239, 241, 283, 353, 367,$ $379, 457, 997, 1367, 3041, 10141, 14699$
2	$2, 3, 5, 7, 11, 17, 19, 23, 101, 107, 109, 113, 163, 283, 311, 331, 347, 359, 701,$ $1153, 1597, 1621, 2063, 2437, 2909, 3319, 6011, 12829$
3	$2, 3, 4, 5, 7, 11, 13, 17, 19, 23, 31, 43, 61, 79, 101, 127, 167, 191, 199, 313, 347,$ $701, 1709, 2617, 3539, 5807, 10501, 10691, 11279, 12391, 14479$
4	$2, 5, 7, 9, 13, 19, 23, 41, 83, 97, 103, 107, 131, 233, 239, 277, 283, 349, 409, 571, 1249,$ $1913, 2221, 2647, 3169, 3527, 4349, 5333, 5903, 5923, 6701, 9127, 9829, 16187$
5	$4, 5, 6, 7, 9, 11, 13, 17, 29, 43, 53, 89, 283, 557, 563, 613, 691, 1223, 2731, 5147,$ $5323, 5479, 9533, 10771, 11257, 11519, 12583$

TABLE 4.2: All $e \leq 17389$ for which E_e/E_1 is a (probable) prime.

4.4 A primality test for certain elliptic curves

If either $N + 1$ or $N - 1$ can be factored, there exists an efficient way of testing N for primality. The numbers which will be tested have the form E_e/E_1, where E_1 is some small number and $E_e = q^e + 1 - (\alpha^e + \overline{\alpha}^e)$ for some complex number α of absolute value $\sqrt{q}$. In general there seems no hope to factor $E_e/E_1 \pm 1$, since the nice structure of the number E_e is lost by the division by E_1. However, for some elliptic curves E_1 is equal to 1. In the following proposition all elliptic curves with $E_1 = 1$ will be given.

Proposition 4.4.1 *Let $\mathcal{E}$ be an elliptic curve defined over $\mathbb{F}_q$ with the property that $E_1 = 1$. Then there are three possibilities:*

(i) *$q = 2$ and the curve $\mathcal{E}$ is isomorphic to the curve with Weierstrass equation $Y^2 + Y = X^3 + X + 1$,*

(ii) *$q = 3$ and the curve $\mathcal{E}$ is isomorphic to the curve with Weierstrass equation $Y^2 = X^3 - X - 1$,*

(iii) *$q = 4$ and the curve $\mathcal{E}$ is isomorphic to the curve with Weierstrass equation $Y^2 + Y = X^3 + \vartheta$, where ϑ is a primitive element of $\mathbb{F}_4$.*

This is a well-known fact from the theory of elliptic curves. As a matter of fact all curves defined over a finite field $\mathbb{F}_q$ whose Jacobians have exactly one $\mathbb{F}_q$-rational point have been classified (see for example [LMQ75]). Nevertheless, the proof of the above proposition will be given for the reader's convenience.

Proof: First, the possible Frobenius eigenvalues α are determined. Write $\alpha = a + bi$, with a and b real numbers. Then the relation $a^2 + b^2 = q$ holds by Hasse-Weil's theorem. Further, the assumption $E_1 = 1$ implies the equation $a = q/2$. Hence $b^2 = q - q^2/4$. Since $b^2 \geq 0$, this implies $q \in \{2, 3, 4\}$. This gives rise to the three cases mentioned in the proposition. □

Note that for all of the three cases mentioned above, $\alpha + \overline{\alpha} = q$, so the three curves are supersingular. We shall now determine the numbers E_e for these three cases.

Proposition 4.4.2 *Let $\mathcal{E}$ be the elliptic curve defined over $\mathbb{F}_2$ with Weierstrass equation $Y^2 + Y = X^3 + X + 1$. Then the Frobenius eigenvalues are given by $1 \pm i$. Furthermore*

$$E_e = \begin{cases} 2^e - 2^{e/2+1} + 1 & \text{if } e \equiv 0 \pmod 8, \\ 2^e - 2^{(e+1)/2} + 1 & \text{if } e \equiv 1, 7 \pmod 8, \\ 2^e + 1 & \text{if } e \equiv 2, 6 \pmod 8, \\ 2^e + 2^{(e+1)/2} + 1 & \text{if } e \equiv 3, 5 \pmod 8, \\ 2^e + 2^{e/2+1} + 1 & \text{if } e \equiv 4 \pmod 8. \end{cases}$$

Let $\mathcal{E}$ be the elliptic curve defined over $\mathbb{F}_3$ with Weierstrass equation $Y^2 = X^3 - X - 1$. Then the Frobenius eigenvalues are given by $(3 \pm i\sqrt{3})/2$ and

$$E_e = \begin{cases} 3^e - 2 \cdot 3^{e/2} + 1 & \text{if } e \equiv 0 \pmod{12}, \\ 3^e - 3^{(e+1)/2} + 1 & \text{if } e \equiv 1, 11 \pmod{12}, \\ 3^e - 3^{e/2} + 1 & \text{if } e \equiv 2, 10 \pmod{12}, \\ 3^e + 1 & \text{if } e \equiv 3, 9 \pmod{12}, \\ 3^e + 3^{e/2} + 1 & \text{if } e \equiv 4, 8 \pmod{12}, \\ 3^e + 3^{(e+1)/2} + 1 & \text{if } e \equiv 5, 7 \pmod{12}, \\ 3^e + 2 \cdot 3^{e/2} + 1 & \text{if } e \equiv 6 \pmod{12}. \end{cases}$$

Let $\mathcal{E}$ be the elliptic curve defined over $\mathbb{F}_4$ with Weierstrass equation $Y^2 + Y = X^3 + \vartheta$, where ϑ is a primitive element of $\mathbb{F}_4$. Then the Frobenius eigenvalue is given by 2 and the number of $\mathbb{F}_4$-rational points satisfies

$$E_e = 4^e - 4^{(e+1)/2} + 1 = (2^e - 1)^2.$$

Proof: The results follow directly from Theorem 4.3.1. □

The third curve can never give (new) primes, since E_e is the square of a Mersenne number for any e (see Example 4.3.4). For the first two curves the following corollary is stated:

Corollary 4.4.3 *Let e be a prime not equal to 2 or 3. Then $E_e = 2^e - \left(\frac{2}{e}\right) 2^{(e+1)/2} + 1$ for the first curve in Proposition 4.4.2. For the second curve in Proposition 4.4.2, $E_e = 3^e - \left(\frac{3}{e}\right) 3^{(e+1)/2} + 1$. Here $\left(\frac{2}{e}\right)$ and $\left(\frac{3}{e}\right)$ denote Legendre symbols.*

These numbers can be viewed as an equivalent of Mersenne numbers in the ring of Gaussian (respectively Eisenstein) integers, hence the following names:

Definition 4.4.4 *Let e be an odd prime. Define the* Gauss-Mersenne number GM_e *as*

$$GM_e = 2^e - \left(\frac{2}{e}\right) 2^{(e+1)/2} + 1,$$

and the Eisenstein-Mersenne number EM_e *as*

$$EM_e = 3^e - \left(\frac{3}{e}\right) 3^{(e+1)/2} + 1.$$

For these two numbers, $GM_e - 1$ (respectively $EM_e - 1$) can be factored easily to such an extent that a primality test for the numbers GM_e and EM_e immediately follows:

Theorem 4.4.5 *Let e be an odd prime and $n \geq 2$. Define*

$$a = \begin{cases} 3 & \text{if } e \equiv 5, 7 \pmod 8, \\ 5 & \text{if } e \equiv 3 \pmod 8, \\ 2^{2^n} + 1 & \text{if } e \equiv 2^{n+1} + 1 \pmod{2^{n+2}}. \end{cases}$$

The number GM_e is a prime if and only if

$$a^{(GM_e - 1)/2} \equiv -1 \pmod{GM_e}.$$

Proof: According to Theorem 4.2.2, the only thing that has to be shown is that $\left(\frac{GM_e}{a}\right) = -1$. However, this follows immediately by repeatedly using the quadratic reciprocity law for Jacobi symbols. In the case $e \equiv 7 \pmod 8$ we have

$$\left(\frac{GM_e}{a}\right) = \left(\frac{2^e - 2^{\frac{e+1}{2}} + 1}{3}\right) = \left(\frac{3}{2^e - 2^{\frac{e+1}{2}} + 1}\right) = \left(\frac{3}{2 - 1 + 1}\right),$$

because if we set $e = 7 + 8k$ we get $2^{7+8k} \equiv 2^7 2^{8k} \equiv 128256^k \equiv 2 \pmod 3$ and $2^{4+4k} \equiv 2^4 2^{4k} \equiv 1616^k \equiv 1 \pmod 3$. So

$$\left(\frac{GM_e}{a}\right) = \left(\frac{3}{2}\right) = -1.$$

When $e \equiv 5 \pmod 8$ we have

$$\left(\frac{GM_e}{a}\right) = \left(\frac{2^e + 2^{\frac{e+1}{2}} + 1}{3}\right) = -\left(\frac{3}{2^e + 2^{\frac{e+1}{2}} + 1}\right) = -\left(\frac{3}{2 - 2 + 1}\right),$$

since writing $e = 5 + 8k$ yields $2^{5+8k} \equiv 2^5 2^{8k} \equiv 32256^k \equiv 5 \pmod 3$ and $2^{3+4k} \equiv 2^3 2^{4k} \equiv 816^k \equiv 2 \pmod 3$. So

$$\left(\frac{GM_e}{a}\right) = -\left(\frac{3}{1}\right) = -1.$$

If $e \equiv 3 \pmod 8$, the relation

$$\left(\frac{GM_e}{a}\right) = \left(\frac{2^e + 2^{\frac{e+1}{2}} + 1}{5}\right) = \left(\frac{5}{2^e + 2^{\frac{e+1}{2}} + 1}\right) = \left(\frac{5}{3+4+1}\right)$$

holds, since if we write $e = 3 + 8k$ we have $2^{3+8k} \equiv 2^3 2^{8k} \equiv 3256^k \equiv 3 \pmod 5$ and $2^{2+4k} \equiv 2^2 2^{4k} \equiv 416^k \equiv 4 \pmod 5$. Thus

$$\left(\frac{GM_e}{a}\right) = \left(\frac{5}{3}\right) = \left(\frac{3}{5}\right) = \left(\frac{3}{2}\right) = -1.$$

Lastly, if $e \equiv 1 \pmod 8$ there exists an integer $n \geq 1$, such that $e \equiv 2^{n+1} + 1 \pmod{2^{n+2}}$. Then

$$\left(\frac{GM_e}{a}\right) = \left(\frac{2^e - 2^{\frac{e+1}{2}} + 1}{2^{2^n} + 1}\right) = \left(\frac{2^{2^n} + 1}{2^e - 2^{\frac{e+1}{2}} + 1}\right) = \left(\frac{2^{2^n} + 1}{2^{2^n} - 2^{2^n} + 1}\right) = -1,$$

since $e > (e+1)/2 > 2^n$. □

To perform this primality test e squarings and 1 multiplication modulo GM_e has to be computed since $a^{(GM_e-1)/2} = (a^{2^{(e-1)/2}} a^{\pm 1})^{2^{(e+1)/2}}$. Since a squaring takes far more work than an addition, additions will be ignored in this estimate. The modular reduction can be implemented in a few additions because of the special form of the numbers. Therefore the computational complexity of this primality test is $O(\log_2(GM_e))$ squarings.

Now the computational complexity of the Lucas-Lehmer test will be considered. The primality test of Proposition 4.2.4 can be rewritten as follows:

Proposition 4.4.6 *Define the sequence $\{s_n\}$ inductively by $s_1 = 4$ and $s_{n+1} = s_n^2 - 2$. For an odd prime p the Mersenne number $M_p = 2^p - 1$ is a prime if and only it divides s_{p-1}.*

Note that the two sequences s_j and V_j are connected by the relation $s_j = V_{2^j}/2^{2^{j-1}}$. This form of the Lucas-Lehmer test is easier to compute, since the calculation of $s_{p-1} \pmod{M_p}$ only involves $p-2$ squarings modulo the Mersenne number M_p. Note that in this case the modular reduction can be done with a single addition because of the special form of the Mersenne number. Thus the computational complexity of the Lucas-Lehmer test is $O(\log_2(M_p))$ squarings.

Therefore it is fair to say that the test of Theorem 4.4.5 is about as efficient as the Lucas-Lehmer test for the Mersenne primes. An actual implementation of this

test has revealed that Gauss-Mersenne primes exist. The number GM_e is a prime for each e in the set
$\{2, 3, 5, 7, 11, 19, 29, 47, 73, 79, 113, 151, 157, 163, 167, 239, 241, 283, 353, 367, 379,$
$457, 997, 1367, 3041, 10141, 14699, 27529, 49207, 77291, 85237\}$.

While we were looking for larger e, M. Oakes et al. [Oak00] found that GM_e is prime for $e \in \{106693, 160423, 203789, 364289\}$ as well. It was shown by us that this list is complete for $e \leq 188369$. Also, no prime number GM_e was found with e in the closed interval $[566000, 695566]$.

Now consider the Eisenstein-Mersenne numbers:

Theorem 4.4.7 *The number EM_e is a prime if and only if either $2^{(EM_e-1)/3} \equiv -3^e \pmod{EM_e}$ or $2^{(EM_e-1)/3} \equiv 3^e - 1 \pmod{EM_e}$.*

Proof: The "if"-part follows directly from Pocklington's theorem. To show the "only if"-part, suppose that EM_e is a prime. Note that in this case the solutions of the equation $x^3 \equiv 1 \pmod{EM_e}$ are given by 1, -3^e, and 3^e-1. Hence all that needs to be shown is that 2 is not a cubic residue modulo EM_e. Define $\omega = (-1+i\sqrt{3})/2$ and $\alpha = 2 + \omega$. Using the cubic reciprocity law in the ring $\mathbb{Z}[\omega]$ (see for example [Cox89, pages 78-80]), it can be shown that 2 is not a cubic residue modulo EM_e (still assuming that EM_e is a prime). The main fact that is used is that EM_e factors over $\mathbb{Z}[\omega]$ as $EM_e = (\alpha^p - 1)(\overline{\alpha}^p - 1)$. □

Using a C program it was shown that EM_e is a prime for each e in the set
$\{2, 5, 7, 11, 17, 19, 79, 163, 193, 239, 317, 353, 659, 709, 1049, 1103, 1759, 2029, 5153,$
$7541, 9049, 10453, 23743, 255361\}$.
In order to find the last value in the list, a modified version of Chris Nash's OpenPFGW C++ program was used as well as our own program. This list is complete for values of e up to 268154.

4.5 The Wagstaff conjecture

The Wagstaff conjecture for the Mersenne numbers can be generalized to the families of numbers arising from elliptic curves discussed in the preceding paragraph. This was done independently in [Kob01]. The Wagstaff conjecture for Mersenne numbers is the following (see [Wag83]):

Conjecture 4.5.1 (Wagstaff) *The distribution of Mersenne primes is characterized by*

(i) *The number of Mersenne primes less than or equal to x is approximately $(e^\gamma / \log 2) \log\log x$. (Here γ is Euler's constant).*

(ii) *The expected number of Mersenne primes $2^p - 1$ with p between n and $2n$, is approximately e^γ.*

(iii) *The probability that $2^p - 1$ is prime, is approximately $(e^\gamma \log ap)/p \log 2$ where $a = 2$ if $p \equiv 3 \pmod 4$ and $a = 6$ if $p \equiv 1 \pmod 4$.*

Although no proof of this statement is known, it is supported by the following heuristic argument:

Heuristic argument. First recall that $M_p = 2^p - 1$ can only be prime if p itself is a prime. According to the well-known prime number theorem (see for instance [HW79, Theorem 6]), the probability that a random number r is prime is asymptotically equal to $1/\log r$. However, M_p is not a random number, and also does not behave as such. Namely, if q divides M_p, then obviously $2^p = 1 \pmod q$ and the order of 2 modulo q divides the prime p, and thus must be equal to p. By Fermat's little theorem we have that the order of 2 modulo q divides $q-1$, thus any divisor q of M_p must be of the form $q = kp + 1$ where k is an integer. Also observe that $2^{(q-1)/2} = 2^{pq} = 1 \pmod q$, so 2 is a quadratic residue modulo q and so q must be equal to ± 1 modulo 8. Thus the smallest divisor of M_p is at least $ap + 1$ where $a = 2$ if p is 1 modulo 4, and $a = 6$ if p is 3 modulo 4.

Since prime numbers q that are smaller than $ap + 1$ cannot divide M_p, the probability that M_p is prime is enlarged by a factor of $\frac{1}{(1-1/q)}$ for each prime $q < ap + 1$. Thus an estimate of the above probability is

$$\frac{1}{\log(2^p - 1)} \prod_{q<2ap+1} \frac{1}{(1 - 1/q)},$$

where the product is taken only over prime numbers q. Merten's theorem (see for example [BS96]) states that

$$\lim_{n\to\infty} \frac{1}{\log n} \prod_{i=1}^{n} \frac{p_i}{p_i - 1} = e^{\gamma},$$

where p_i is the i^{th} prime number. Thus this probability is asymptotically equal to $(e^{\gamma} \log ap)/p \log 2$. The first two claims easily follow from this observation.

For prime-generating elliptic curves Wagstaff's conjecture can be generalized:

Conjecture 4.5.2 *Let $\mathcal{E}$ be a prime-generating elliptic curve defined over the field $\mathbb{F}_q$. Then the following statements can be made about the distribution of the primes generated by $\mathcal{E}$:*

(i) The number of primes of the form E_k/E_1 less than or equal to x, is approximately $(e^{\gamma}/\log q)\log\log x$.

(ii) The expected number of primes of the form E_k/E_1 with k between n and qn, is approximately e^{γ}.

(iii) The probability that E_k/E_1 is prime, is approximately $e^{\gamma}\log ak/k\log q$, where a depends on the specific choice of $\mathcal{E}$ (in general $a = 2$).

The a in the above conjecture is due to the fact that for all prime divisors d of E_e/E_1 we have $d \equiv 1 \pmod{ae}$, if e is odd and q is not too large [Bee01, Theorem

5.3.2]. In general, $a = 2$ gives a sharp bound. However, for some specific curves this bound can be improved, as is shown for the two families of Mersenne-like numbers in the next proposition.

Proposition 4.5.3 *Let e be an odd prime larger than 3. Then the following statements hold:*

(i) Suppose that l is a prime divisor of the number $GM_e = 2^e - \left(\frac{2}{e}\right) 2^{(e+1)/2} + 1$. Then $l \equiv 1 \pmod{4e}$.

(ii) If l is a prime divisor of the number $EM_e = 3^e - \left(\frac{3}{e}\right) 3^{(e+1)/2} + 1$, we have that $l \equiv 1 \pmod{6e}$.

Proof: Note that GM_e is a divisor of $2^{2e}+1$. This implies that $2^{2e} \equiv -1 \pmod{l}$. Hence the multiplicative order of 2 in the group $\mathbb{F}_l^*$ is a divisor of $4e$. This implies after some manipulation of the formulas that the multiplicative order of 2 equals $4e$. Thus $l \equiv 1 \pmod{4e}$. For the other half of the proposition, note that EM_e is a divisor of $3^{3e} + 1$ and proceed analogously. □

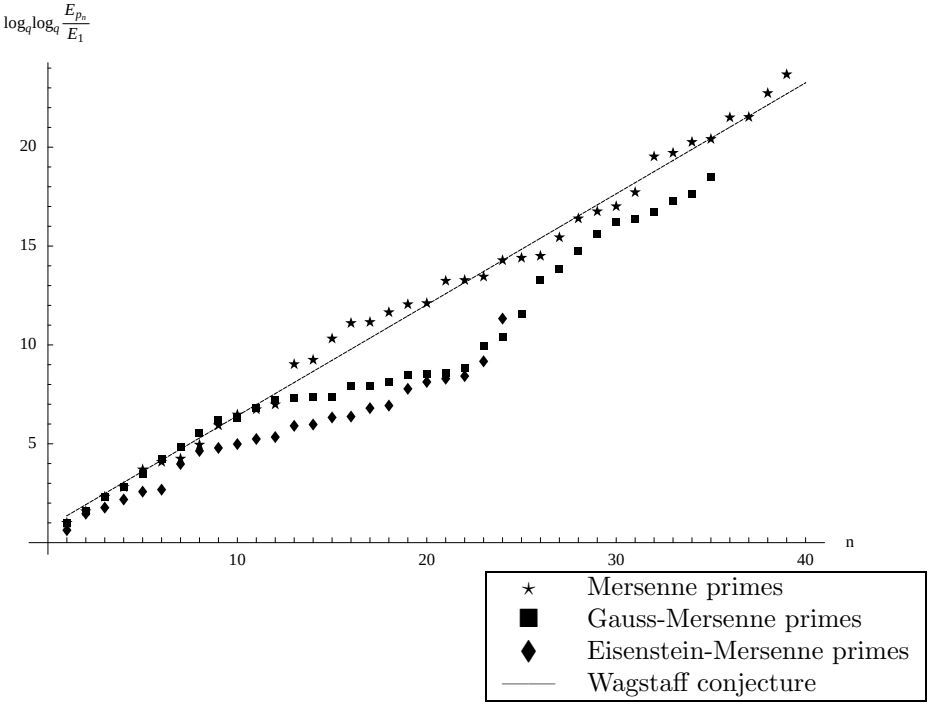

Figure 4.1: Generalized Wagstaff's conjecture

Table 4.2 and the results concerning the Mersenne-like primes found in the previous section give an idea of how accurate both Wagstaff's conjecture and its generalization are. Define p_n to be the n-th number in the respective family such that

E_{p_n}/E_1 is a prime and suppose that the elliptic curve is defined over $\mathbb{F}_q$. According to Wagstaff's conjecture the plot of n against $\log_q(\log_q(E_{p_n}/E_1))$ should lie on a straight line with slope $e^{-\gamma}$. In this way the accuracy of Conjecture 4.5.2 can be observed graphically. In Figure 4.1 a line with slope $e^{-\gamma}$ and the points $(n, \log_2(\log_2(GM_{p_n})))$ have been drawn. The respective points for the (currently) known Mersenne primes M_p and the (currently) known Eisenstein-Mersenne primes EM_p are included in the figure as well.

Chapter 5

Pseudorandom sequences from elliptic curves

Summary. In this chapter some known constructions to produce pseudorandom sequences with the aid of elliptic curves will be generalized. Both additive and multiplicative characters on elliptic curves will be used for this purpose. This chapter is based on joint work with Peter Beelen [BD02].

5.1 Introduction

Nowadays, many applications call for random numbers. One of the most preferable ways to generate those would be to take a monkey, give him a coin to flip, and write down the result of each coin flip. Unfortunately this process is quite slow, and a faster way to generate random numbers is needed. On second thought, a sequence of numbers that *appears* random would be just as good - who could tell the difference? Such a sequence will be called pseudorandom.

Many people have constructed pseudorandom number generators using many, diverse methods (see for example [MOV97, Chapter 5] for an overview). The first study of using linear congruences on elliptic curves to generate pseudorandom sequences was done in [Hal94]. Further results on these generators were obtained in [MS02; GBS00; KS00; VW00]. Some of these constructions will be generalized here and another construction using linear recurrence relations on elliptic curves will be introduced. An instance of this last construction was investigated in [GL02].

5.2 Some properties of elliptic curves

As elliptic curves will be used heavily throughout this chapter, first some notation will be fixed and some elementary properties of elliptic curves will be stated. The algebraic closure of a field F will be denoted by $\overline{F}$.

Definition 5.2.1 *An elliptic curve $\mathcal{E}$ is the set of projective points in $\mathbb{P}^2$ satisfying*

the Weierstrass equation $F(X,Y,Z) = 0$, where F is defined as

$$F(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3.$$

A curve $\mathcal{E}$ is said to be defined over $\mathbb{F}_q$ if $a_1, \ldots, a_6 \in \mathbb{F}_q$.

However, usually one considers the (affine) non-homogeneous equation obtained by putting $x = X/Z$ and $y = Y/Z$ in the equation $F(X,Y,Z) = 0$:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

always remembering to include the *point $\mathcal{O}$ at infinity,* which has projective coordinates $[0:1:0]$. We will write the affine equation as $f(x,y) = F(X/Z, Y/Z, 1)$. A point P satisfying $\frac{\partial f}{\partial x}(P) = \frac{\partial f}{\partial y}(P) = 0$, will be called singular. In this chapter, we will only consider non-singular elliptic curves, that is elliptic curves which have no singular point(s).

Elliptic curves have an Abelian group structure [Sil86, Section 3.2]. Adding two points P and Q is done by first taking the line through P and Q (if $P = Q$, take the tangent at P). This line will intersect the curve $\mathcal{E}$ in a third point R, since f is a cubic equation. Note that points are counted with multiplicity. The sum of P and Q is now defined to be the intersection point of the line through R and $\mathcal{O}$. Scalar multiplication of a point P on an elliptic curve by n, i.e. adding P to itself n times, will be denoted by $[n]P$.

A point is called $\mathbb{F}$-rational if its coordinates are in $\mathbb{F}$. The group of $\mathbb{F}$-rational points (see [Sil86, page 57] for proof that these points form a group) on the curve $\mathcal{E}$ will be denoted by $\mathcal{E}(\mathbb{F})$. A point $P \in \mathcal{E}$ is called a n-torsion point if it satisfies $[n]P = \mathcal{O}$. The n-torsion subgroup of $\mathcal{E}$ will be denoted by $\mathcal{E}[n]$.

The function field of an elliptic curve $\mathcal{E}$ defined over $\mathbb{F}_q$ is defined to be the quotient field of $\overline{\mathbb{F}_q}[x,y]/(f(x,y))$. It will be denoted by $\mathcal{F}(\mathcal{C})$, or by $\mathcal{F}_q(\mathcal{C})$ if only the functions with coefficients in $\mathbb{F}_q$ are considered. The valuation map $v_P(g)$ is a map from $\mathcal{F}(\mathcal{C})$ to $\mathbb{Z} \cup \infty$, which takes as value minus the pole order of g at P. So if g has a zero of degree k at point P and a pole of degree l at Q, one has $v_P(g) = k$ and $v_Q(g) = -l$. For every (rational) function $f \in \mathcal{F}(\mathcal{C})$, there are finitely many points where $v_P(f) \neq 0$ [Sil86, Proposition 2.1.2].

An algebraic curve $\mathcal{C}$ is a projective variety of dimension 1, with an arbitrary genus g. Elliptic curves are algebraic curves with genus $g = 1$. For more information on algebraic curves, see for instance [Sil86, Chapter 2] or [Sti93].

The following famous result by Hasse and Weil concerns the number of points of an algebraic curve. For a proof of this theorem, see for instance [Sti93, Section 5.2].

Theorem 5.2.2 (Hasse-Weil bound) *Let $\mathcal{C}$ be an algebraic curve defined over $\mathbb{F}_q$ of genus g. Then there exist $\alpha_1, \ldots, \alpha_g \in \mathbb{C}$ of length $\sqrt{q}$, such that for all $e \in \mathbb{N}$*

$$\#\mathcal{C}(\mathbb{F}_{q^e}) = q^e + 1 - \sum_{i=1}^{g} (\alpha_i^e + \overline{\alpha}_i^e)\,.$$

Corollary 5.2.3 *The number of points of an algebraic curve $\mathcal{C}$ defined over $\mathbb{F}_q$ of genus g satisfies*

$$|\#\mathcal{C}(\mathbb{F}_{q^e}) - q^e - 1| \leq 2g\sqrt{q^e}.$$

Proof: This follows easily from Theorem 5.2.2. □

In the case of an elliptic curve, we have

Corollary 5.2.4 *The number of points of an elliptic curve $\mathcal{E}$ defined over $\mathbb{F}_q$ satisfies*

$$|\#\mathcal{C}(\mathbb{F}_{q^e}) - q^e - 1| \leq 2\sqrt{q^e}.$$

Proof: This follows easily from Theorem 5.2.2 and the fact that the genus of an elliptic curve is 1. □

For the following proposition, see [Sil86, page 145].

Proposition 5.2.5 *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_q$. Then there exist numbers k and l such that the elliptic curve as an Abelian group satisfies*

$$\mathcal{E}(\mathbb{F}_q) \cong \mathbb{Z}/k\mathbb{Z} \times \mathbb{Z}/kl\mathbb{Z}.$$

Furthermore, k divides $(q-1)$.

The k and l from the above proposition will be denoted by $k(\mathcal{E})$ and $l(\mathcal{E})$ respectively or by $k(\mathcal{E}, \mathbb{F}_q)$ and $l(\mathcal{E}, \mathbb{F}_q)$ if the field of definition is not immediately clear. Note that if $k = 1$, the elliptic curve will have a cyclic structure.

Since $k = k(\mathcal{E}, \mathbb{F}_q)$ divides $q-1$, the map from $\mathcal{E}$ to $\mathcal{E}$ obtained by multiplying with k, is of degree k^2 and unramified, i.e. the equation $[k]P = Q$ has k solutions for any point Q. Further note that $\mathcal{E}[k] \subset \mathcal{E}(\mathbb{F}_q)$.

5.3 Pseudorandom sequences

In this section some basic definitions concerning pseudorandom sequences are given. First a very handy tool is introduced:

Definition 5.3.1 *The trace map $\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}$ is a function from $\mathbb{F}_{q^e}$ to $\mathbb{F}_q$ defined by*

$$\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(x) = x + x^q + x^{q^2} + \cdots + x^{q^{e-1}}.$$

Note that the trace map is $\mathbb{F}_q$–linear, since $(a+b)^q = a^q + b^q$ in $\mathbb{F}_{q^e}$, and $(\alpha a)^q = \alpha a^q$ for $\alpha \in \mathbb{F}_q$ and $a \in \mathbb{F}_{q^e}$.

Now, some basic properties concerning pseudorandom sequences are discussed.

Definition 5.3.2 *Let $S = \{s(0), s(1), \ldots, s(N-1)\}$ be a sequence of elements of $\mathbb{F}_q$ and let $\alpha \in \mathbb{F}_q^*$. Denote the characteristic of $\mathbb{F}_q$ by p. The* balance of S with respect to α *is defined in the following way:*

$$B_S(\alpha) = \frac{1}{N}\sum_{i=0}^{N-1} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(\alpha s(i))},$$

where ζ_p is the p^{th} root of unity $\exp(2\pi i/p)$*. Furthermore, the* balance of S *is defined as*

$$B_S = \max_{\alpha\in\mathbb{F}_q^*}\{|B_S(\alpha)|\}.$$

Note that for binary sequences, this definition is equivalent to

$$B_S = \frac{1}{N}\sum_{i=0}^{N-1}(-1)^{s(i)} = \frac{1}{N}\big(\#\{i : s(i) = 0\} - \#\{i : s(i) = 1\}\big).$$

Thus the balance is a measure of any bias in the sequence S. Also, if $s(i)$ takes every value in $\mathbb{F}_q$ equally often, then $B_S = B_S(\alpha) = 0$ for any α.

Now a similar concept for sequences defined over $\mathbb{Z}/m\mathbb{Z}$ is introduced. It will be assumed that m divides $q-1$. Then it is possible to identify $\mathbb{Z}/m\mathbb{Z}$ with $(\mathbb{F}_q^*)^{(q-1)/m}$. Thus there exists a surjective homomorphism of groups $\chi_m : \mathbb{F}_q^* \to \mathbb{Z}/m\mathbb{Z}$.

Definition 5.3.3 *If $S = \{s(0), s(1), \ldots, s(N-1)\}$ is a sequence of elements of $(\mathbb{F}_q^*)^{(q-1)/m}$, then the balance with respect to α is defined as*

$$B_S(\alpha) = \frac{1}{N}\sum_{i=0}^{N-1} \zeta_m^{\chi_m(\alpha s(i))},$$

with $\alpha \in \mathbb{F}_q^*$ and $\zeta_m = \exp(2\pi i/m)$.

Now the autocorrelation of a sequence can be defined in s similar way.

Definition 5.3.4 *Let $\{s(0), s(1), \ldots, s(N-1)\}$ be a sequence S of period N defined over the finite field $\mathbb{F}_q$. Write p for the characteristic of this field. Furthermore, let $\alpha, \beta \in \mathbb{F}_q^*$.*
The autocorrelation of S with respect to α and β *is defined as:*

$$C_S(d, \alpha, \beta) = \frac{1}{N}\sum_{i=0}^{N-1} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(\alpha s(i+d) - \beta s(i))},$$

with $0 \le d < N$ and $\zeta_p = \exp(2\pi i/p)$. For a sequence S defined over $(\mathbb{F}_q^*)^{(q-1)/m}$ the definition is

$$C_S(d, \alpha, \beta) = \frac{1}{N}\sum_{i=0}^{N-1} \zeta_m^{\chi_m(\alpha s(i+d) - \beta s(i))},$$

with $\zeta_m = \exp(2\pi i/m)$.

Note that in the above definition $i + d$ should be read modulo N. Also note that for binary sequences this definition amounts to

$$C_S(d) = \frac{1}{N} \sum_{i=0}^{N-1} (-1)^{s(i+d)+s(i)},$$

which is the usual definition of the autocorrelation (see for example [MOV97, Section 5.4]).

Another useful object is the crosscorrelation of two sequences:

Definition 5.3.5 *Let $S = \{s(i)\}$ and $T = \{t(i)\}$ be two sequences defined over $\mathbb{F}_q$ having the same period N. Denote the characteristic of $\mathbb{F}_q$ by p and let $\alpha, \beta \in \mathbb{F}_q^*$. The* crosscorrelation of S and T with respect to α and β *is defined as*

$$C_{S,T}(d, \alpha, \beta) = \frac{1}{N} \sum_{i=0}^{N-1} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(\alpha s(i+d) - \beta t(i))},$$

with $\zeta_p = \exp(2\pi i/p)$ and $0 \leq d < N$. For sequences S and T defined over $(\mathbb{F}_q^*)^{(q-1)/m}$ the crosscorrelation is defined as

$$C_{S,T}(d, \alpha, \beta) = \frac{1}{N} \sum_{i=0}^{N-1} \zeta_m^{\chi_m(\alpha s(i+d) - \beta t(i))},$$

with $\zeta_m = \exp(2\pi i/m)$.

Note that the crosscorrelation $C_{S,S}(d, \alpha, \beta)$ of a sequence with itself is equal to the autocorrelation $C_S(d, \alpha, \beta)$, as it should be.

The problem is to find a family of sequences $\Sigma = \{S_i | i \in I\}$ such that for all $i, j \in I$ the crosscorrelations $C_{S_i,S_j}(d, \alpha, \beta)$ are small. Such a family of sequences could be used for CDMA purposes, for instance in mobile telephony. Also, the autocorrelation of a sequence obtained by concatenating two or more sequences from such a family is essentially bounded by these crosscorrelations. When the generator has to switch to another member of the family, one would still want certain properties to hold for the autocorrelation of its combined output.

5.4 Using additive characters

Some generalizations of known constructions of pseudorandom sequences from elliptic curves will be given in this section.

Let $\mathcal{E}$ be an elliptic curve defined over a finite field $\mathbb{F}_{q^e}$ of characteristic p. Suppose for now that this group is cyclic of order $N = l(\mathcal{E})$ (i.e. $k(\mathcal{E}) = 1$) and has generator P. Let $f \in \mathcal{F}_{q^e}(\mathcal{E})$ be a function on $\mathcal{E}$ defined over $\mathbb{F}_{q^e}$. The pseudorandom sequence $S = \{s(i)\}$ will be studied in this section, where the $s(i)$ are given by:

$$s(i) = \mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f([i]P)),$$

with $0 \leq i < N$.

The condition that $\mathcal{E}(\mathbb{F}_{q^e})$ is a cyclic group is a natural one, since an ordering of the points in $\mathcal{E}(\mathbb{F}_{q^e})$ is needed. In the literature this assumption is often made. Moreover, the field $\mathbb{F}_q$ is usually assumed to be $\mathbb{F}_p$. Both these restrictions will be removed in this section. First some concepts are introduced.

Definition 5.4.1 *Let $\mathcal{C}$ be an algebraic curve of genus g defined over $\mathbb{F}_q$. Let $f \in \mathcal{F}_q(\mathcal{C})$ be a rational function on $\mathcal{C}$ defined over $\mathbb{F}_q$ as well. Define $\mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_q)$ to be the set of all $\mathbb{F}_q$-rational points Q on $\mathcal{C}$ such that there exists a function $g \in \mathcal{F}_q(\mathcal{C})$ (depending on Q) with the property that $f - g^p + g$ is defined at Q.*

Here g^p denotes taking the p^{th} power of each function value, so that $\mathrm{Tr}(g^p - g) = 0$. Note that for every function $f \in \mathcal{F}_q(\mathcal{C})$ and point $Q \in \mathcal{C}(\overline{\mathbb{F}_q})$ there exists a function $g \in \mathcal{F}_q(\mathcal{C})$ such that either $v_Q(f - g^p + g) \geq 0$ or $v_Q(f - g^p + g) < 0$ and $p \not| v_Q(f - g^p + g)$. Define $m_Q = -1$ in the former case and $m_Q = -v_Q(f - g^p + g)$ in the latter. Of course m_Q depends on f as well. When this is to be made explicit $m_Q(f)$ will be written instead of m_Q. For more details see [Sti93, page 114].

Also observe that the quantity $\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}((f - g^p + g)(Q))$ does not depend on g as long as the function $f - g^p + g$ is defined at Q. Thus, even if f itself is not defined in Q, the notation $\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(f(Q))$ will be used for $Q \in \mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_q)$.

Now define the sequence to be studied is given by:

Definition 5.4.2 *Let $\mathcal{E}$ be an elliptic curve defined over $\mathbb{F}_{q^e}$. Suppose that P is a generator of the group $[k(\mathcal{E}, \mathbb{F}_{q^e})]\mathcal{E}(\mathbb{F}_{q^e})$ and denote its order by N. Let $f \in \mathcal{F}_{q^e}(\mathcal{E})$. The sequence $S^{\mathrm{AS}}(f, P) = \{s(i)\}_{0 \leq i < N}$ is defined by*

$$s(i) = \mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f([i]P)).$$

Here the convention that $\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(Q)) = 0$ if $Q \notin \mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e})$ is used. Of course this sequence depends on the elliptic curve as well, but this is not made explicit in the notation.

Note that in the above notation $N = l(\mathcal{E}, \mathbb{F}_{q^e})$ and that if $\mathcal{E}(\mathbb{F}_{q^e})$ is cyclic, the situation has already been studied in the literature [GBS00; VW00].

From the point of view of coding theory an ordering of the points in $\mathcal{E}(\mathbb{F}_{q^e})$ is unnecessary. In this case any change of ordering gives rise to an equivalent code. Indeed there is in that case no need of restricting oneself to elliptic curves. The resulting codes are called trace-codes. They have been studied in for example [Sti93, Chapter 8] and [Vos93].

Before some estimates for the parameters of the pseudorandom sequences defined above are given, some more theory is needed. The key to the results is the proposition about the following exponential sum:

Definition 5.4.3 *Let $\mathcal{C}$ be an algebraic curve defined over $\mathbb{F}_q$. Let $f \in \mathcal{F}_q(\mathcal{C})$. Now look at the following exponential sum:*

$$ES^{\mathrm{AS}}(\mathcal{C}, f) = \sum_{P \in \mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_q)} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(f(P))},$$

with $\zeta_p = \exp(2\pi i/p)$.

A known upper bound for this exponential sum is given in the next proposition.

Proposition 5.4.4 *Let $\mathcal{C}$ be an algebraic curve of genus g defined over $\mathbb{F}_q$. Let $f \in \mathcal{F}_q(\mathcal{C})$ and suppose that $f \neq z^p - z$ for all $z \in \overline{\mathcal{F}(\mathcal{C})}$. Then the following holds:*

$$\left|ES^{\mathrm{AS}}(\mathcal{C}, f)\right| \leq \left(2g - 2 + \sum_{P \in \mathcal{C}(\overline{\mathbb{F}}_q)} (m_P + 1)\right) \sqrt{q}.$$

This proposition was proven in [Bom66; KS00; VW00]. The gist of the proof is given here for the convenience of the reader.

Proof: The proposition can be rephrased by considering the curve $\mathcal{D}$, defined over $\mathbb{F}_q$ whose function field is given by $\mathcal{F}_q(\mathcal{C})(z)$ with $z^p - z = f$. Denote its genus by h. The L-function of $\mathcal{D}$ is the product of the L-function of $\mathcal{C}$ with the following $p-1$ expressions ($1 \leq i \leq p-1$):

$$\exp\left(\sum_{e \geq 1} \frac{T^e}{e} \sum_{P \in \mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_{q^e})} \zeta_p^{i\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_p}(P)}\right).$$

As a matter of fact the above expressions turn out to be polynomials. By Hasse-Weil's theorem these polynomials have roots of length $1/\sqrt{q}$ and hence for all $e \geq 1$ the following relation holds.

$$\left|\sum_{P \in \mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_{q^e})} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_p}(P)}\right| \leq \frac{2(h-g)}{p-1} \sqrt{q^e}.$$

Using the theory for Artin-Schreier extensions an explicit expression for the genus h can be derived (see for example [Sti93, Section 3.7]). This leads to the upper bound given in the proposition. □

By applying the above result an estimate for the parameters of the sequences $S^{\mathrm{AS}}(f, P)$ can now be given.

Theorem 5.4.5 *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_{q^e}$ of characteristic p. Set $k = k(\mathcal{E}, \mathbb{F}_{q^e})$. Furthermore, let $f \in \mathcal{F}_{q^e}(\mathcal{E})$ be a function and P be a generator of the group $[k]\mathcal{E}(\mathbb{F}_{q^e})$.*

Suppose that for all $z \in \overline{\mathcal{F}}(\mathcal{E})$ the relation $z^p - z \neq f \circ [k]$ holds and that

$$\mathcal{E}^{\mathrm{AS}}(f \circ [k], \mathbb{F}_{q^e}) = [k]^{-1} \left(\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle \right).$$

Then $B_{S^{\mathrm{AS}}(f,P)}$ *is bounded from above by*

$$\frac{1}{N} \left(N - \#(\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle) + \frac{1}{k(\mathcal{E})^2} \sum_{Q} (m_Q(f \circ [k]) + 1) \sqrt{q^e} \right),$$

with $N = \#\langle P \rangle$.

Proof: Denote by S the sequence $\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(Q))\}$ with $Q \in \mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle$. Further denote by T the sequence $\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f \circ [k](Q))\}$ with $Q \in \mathcal{E}^{\mathrm{AS}}(f \circ [k], \mathbb{F}_{q^e})$. The relation $\mathcal{E}^{\mathrm{AS}}(f \circ [k], \mathbb{F}_{q^e}) = [k]^{-1} \left(\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle \right)$ holds and hence for each point R in $\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle$, there exist exactly k^2 points Q in the set $\mathcal{E}^{\mathrm{AS}}(f \circ [k], \mathbb{F}_{q^e})$ such that $[k]Q = R$. Thus for any $\alpha \in \mathbb{F}_q^*$

$$\#\mathcal{E}^{\mathrm{AS}}(f \circ [k], \mathbb{F}_{q^e}) B_T(\alpha) = k^2 \#(\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle) B_S(\alpha).$$

So

$$B_T(\alpha) = B_S(\alpha).$$

Since

$$N \cdot |B_{S^{\mathrm{AS}}(f,P)}(\alpha)| \leq N - \#(\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle) + \#(\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle) |B_S(\alpha)|,$$

an upper bound for $B_T(\alpha)$ results in an upper bound for $B_{S^{\mathrm{AS}}(f,P)}$. However, since $\#\mathcal{E}^{\mathrm{AS}}(f \circ [k], \mathbb{F}_{q^e}) B_T(\alpha) = ES^{\mathrm{AS}}(\mathcal{E}, f \circ [k])$, such an upper bound is available from Proposition 5.4.4. This concludes the proof. □

Note that the technical condition

$$\mathcal{E}^{\mathrm{AS}}(f \circ [k], \mathbb{F}_{q^e}) = [k]^{-1} \left(\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) \cap \langle P \rangle \right)$$

is fulfilled if $k = 1$. Also note that the righthandside set is always contained in the lefthandside set.

Now a special case of the above lemma is considered. Denote by $\mathrm{wdeg}(f(x, y))$ the weighted degree of a polynomial in two variables defined by $\mathrm{wdeg}(x) = 2$ and $\mathrm{wdeg}(y) = 3$. Thus the weighted degree of a polynomial is equal to the highest weighted degree of its monomials. Note that the choice of degrees for the monomials x and y is natural for elliptic curves, since one works modulo the polynomial $f(x, y)$, and thus the weighted degree of y^2 should be equal to the weighted degree of x^3 for a consistent definition.

Corollary 5.4.6 *Let* $\mathcal{E}$ *be an elliptic curve defined over the finite field* $\mathbb{F}_{q^e}$ *of characteristic* p *given by a Weierstrass equation. Let* f *be a polynomial in the coordinate*

functions x and y such that $\deg_y(f) \leq 1$. *Further let P be a generator of the group* $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_{q^e})$ *and define* $N = \#\langle P \rangle$. *Suppose that p does not divide* $\mathrm{wdeg}(f)$. *Then*

$$B_{S^{\mathrm{AS}}(f,P)} \leq \frac{1}{N}\left(1 + (1 + \mathrm{wdeg}(f))\sqrt{q^e}\right).$$

Note that the above corollary also follows from [KS00, Theorem 1] or the work of Bombieri [Bom66]. The condition $\deg_y(f) \leq 1$ is not a real restriction, because the Weierstrass equation can be used to reduce this degree if $\deg_y(f) \geq 2$. Further note that if this condition is met, the relation $v_{\mathcal{O}}(f) = -\mathrm{wdeg}(f)$ holds where $\mathcal{O}$ is the point at infinity $[0:1:0]$. For the proof of the above theorem one uses that $\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) = \mathcal{E}(\mathbb{F}_{q^e}) \setminus \{\mathcal{O}\}$ and that $\mathcal{E}^{\mathrm{AS}}(f \circ [k(\mathcal{E})], \mathbb{F}_{q^e}) = \mathcal{E}(\mathbb{F}_{q^e}) \setminus \mathcal{E}[k(\mathcal{E})]$.

In the same way the autocorrelation of the sequences $S^{\mathrm{AS}}(f, P)$ can be investigated. Before moving to the main theorem on this, a lemma is stated.

Lemma 5.4.7 *Let $\mathcal{E}$ be an elliptic curve defined over the field $\mathbb{F}_{q^e}$ of characteristic p. Let $f \in \mathcal{F}_{q^e}(\mathcal{E})$ and choose $\alpha, \beta \in \mathbb{F}_q^*$. Write $k = k(\mathcal{E}, \mathbb{F}_{q^e})$ and choose a generator P of the group $[k]\mathcal{E}(\mathbb{F}_{q^e})$ and a number d satisfying $1 \leq d < N$ with $N = \#\langle P \rangle$. Define $h \in \mathcal{F}_{q^e}(\mathcal{E})$ by*

$$h(X) = \alpha f(X \oplus [d]P) - \beta f(X).$$

Denote by S the sequence $\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_p}(h \circ [k](Q))\}$ with $Q \in \mathcal{E}^{\mathrm{AS}}(h \circ [k], \mathbb{F}_{q^e})$. Finally suppose that $\mathcal{E}^{\mathrm{AS}}(h \circ [k], \mathbb{F}_{q^e}) = [k]^{-1}\left(\mathcal{E}^{\mathrm{AS}}(h, \mathbb{F}_{q^e}) \cap \langle P \rangle\right)$. Then

$$N \cdot C_{S^{\mathrm{AS}}(f,P)}(d, \alpha, \beta) = c + \#(\mathcal{E}^{\mathrm{AS}}(h, \mathbb{F}_{q^e}) \cap \langle P \rangle) B_S.$$

Here

$$c = \sum_{Q} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_p}(h(Q))},$$

where the sum is over points Q such that $Q \in \langle P \rangle$ and $Q \notin \mathcal{E}^{\mathrm{AS}}(h, \mathbb{F}_{q^e})$.

Proof: Note that $C_{S^{\mathrm{AS}}(f,P)}(d, \alpha, \beta) = B_{S^{\mathrm{AS}}(h,P)}(1)$. Using similar techniques as in the proof of Lemma 5.4.5, the result is obtained. □

Using an upper bound for exponential sums, an upper bound for the autocorrelation can be derived, if some technical conditions are met. More explicitly, the following theorem is proven in the case that f is a polynomial in the coordinate functions.

Theorem 5.4.8 *Let $\mathcal{E}$ be an elliptic curve defined over the field $\mathbb{F}_{q^e}$ given by a Weierstrass equation. Let f be a polynomial in the two coordinate functions x and y, such that* $\deg_y(f) \leq 1$. *Choose $\alpha, \beta \in \mathbb{F}_q^*$. Further choose a generator P of the group* $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_{q^e})$ *and a number d satisfying $1 \leq d < N$ with $N = \#\langle P \rangle$. Suppose that the characteristic of $\mathbb{F}_q$ does not divide* $\mathrm{wdeg}(f)$. *Then the autocorrelation can be bounded as follows:*

$$|C_{S^{\mathrm{AS}}{}_{(f,P)}}(d,\alpha,\beta)| \leq \frac{1}{N}\left(2+2(1+\mathrm{wdeg}(f))\sqrt{q^e}\right).$$

Analogously to the autocorrelation, properties about the crosscorrelations of sequences can be derived. Some results are stated in the following theorem.

Theorem 5.4.9 *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_{q^e}$ of characteristic p, given by a Weierstrass equation. Let P be a generator of the group $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_{q^e})$ and write $N = \#\langle P\rangle$. Let f_1 and f_2 be two polynomials in the coordinate functions x and y such that $\deg_y(f_i) \leq 1$ for $i = 1, 2$, and such that for all $(\alpha,\beta) \in \mathbb{F}_{q^e}^2 \setminus \{(0,0)\}$ the relation $p \not| \mathrm{wdeg}(\alpha f_1 - \beta f_2)$ holds. Write $S_1 = S^{\mathrm{AS}}(f_1, P)$ and $S_2 = S^{\mathrm{AS}}(f_2, P)$. For all $\alpha, \beta \in \mathbb{F}_q^*$ and $0 \leq d < N$ the crosscorrelation can be bounded as*

$$|C_{S_1,S_2}(d,\alpha,\beta)| \leq \frac{1}{N}\left(2+(2+\mathrm{wdeg}(f_1)+\mathrm{wdeg}(f_2))\sqrt{q^e}\right),$$

unless $d = 0$ and $\alpha f_1 = \beta f_2$.

Proof: This is a straightforward generalization of the proof of Theorem 5.4.8. □

Now an example of a family of sequences having good crosscorrelations will be given. In this example it is assumed that the characteristic is 2, since this is the most interesting case for applications.

Example 5.4.10 Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_{2^e}$. Denote by P a generator of the group $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_{q^e})$ and write $N = \#\langle P\rangle$. Let $\mathbf{a}$ be defined by $\mathbf{a} = (a_0, \cdots, a_m) \in \mathbb{F}_{2^e}^{m+1}$ and let $S_{\mathbf{a}}$ be the binary sequence $S^{\mathrm{AS}}(f_{\mathbf{a}}, P)$ with defining function $f_{\mathbf{a}} = a_0 y + \cdots + a_m x^m y$.
For any number $0 \leq d < N$ and $\mathbf{a}, \mathbf{b} \in \mathbb{F}_{q^e}^{n+1} \setminus \{\mathbf{0}\}$ the crosscorrelation of the two sequences can be bounded as

$$\begin{aligned} |C_{S_{\mathbf{a}},S_{\mathbf{b}}}(d)| &\leq \tfrac{1}{N}\left(2+(2+\mathrm{wdeg}(f_{\mathbf{a}})+\mathrm{wdeg}(f_{\mathbf{b}}))\sqrt{2^e}\right) \\ &\leq \tfrac{1}{N}\left(2+(8+4m)\sqrt{2^e}\right), \end{aligned}$$

unless $d = 0$ and $\mathbf{a} = \mathbf{b}$.

5.5 Using multiplicative characters

Now multiplicative characters and Kummer extensions will be used to obtain similar results as in the previous section, instead of additive characters and Artin-Schreier extensions. Codes have been obtained using this approach in [Per91]. Sequences have been constructed in this way using the projective line in [Bar97]. Here sequences will be constructed using elliptic curves.

Definition 5.5.1 *Let $\mathcal{C}$ be an algebraic curve defined over $\mathbb{F}_q$. Choose $1 < m < q-1$ a divisor of $q-1$ and let $f \in \mathcal{F}_q(\mathcal{C})$. Define $\mathcal{C}^K(f, \mathbb{F}_q)$ to be the set of $\mathbb{F}_q$-rational points on $\mathcal{C}$ such that there exists a $g \in \mathcal{F}_q(\mathcal{C})$ such that $v_P(f \cdot g^m) = 0$.*

Note that if $\phi : \mathbb{F}_q^* \to \mathbb{Z}/m\mathbb{Z}$ is a homomorphism of groups, the quantity $\phi((f \cdot g^m)(Q))$ does not depend on g, as long as $v_Q(f \cdot g^m) = 0$. Hence, $\phi(f(Q))$ will be written instead of $Q \in \mathcal{C}^K(f, \mathbb{F}_q)$, even if $v_Q(f) \neq 0$. In particular the quantity $f(Q)^{(q-1)/m}$ is well-defined for $Q \in \mathcal{C}^K(f, \mathbb{F}_q)$. If $Q \notin \mathcal{C}^K(f, \mathbb{F}_q)$, a $g \in \mathcal{F}_q(\mathcal{C})$ can always be found such that $(f \cdot g^m)(Q) = 0$. Hence $f(Q)^{(q-1)/m}$ is defined to be zero for $Q \notin \mathcal{C}^K(f, \mathbb{F}_q)$, even if f has a pole in Q. In the same way $\phi(f(Q)) = 0$ is defined for $Q \notin \mathcal{C}^K(f, \mathbb{F}_q)$.

Definition 5.5.2 *Let $\mathcal{E}$ be an elliptic curve defined over $\mathbb{F}_q$. Fix a natural number $1 < m < q-1$ dividing $q-1$. Denote by $\chi_m : \mathbb{F}_q^* \to \mathbb{Z}/m\mathbb{Z}$ some fixed, surjective homomorphism of groups. Let $P \in \mathcal{E}(\mathbb{F}_q)$ and $f \in \mathcal{F}_q(\mathcal{E})$. Define $S^K(f, P) = \{s(i)\}$ by*

$$s(i) = \chi_m(f([i]P)).$$

The homomorphism χ_m is needed to define the balance and correlations of the sequence $S^K(f, P)$ (see Section 5.3). Note that $\chi_m(f(Q)) = 0$ if $Q \notin \mathcal{E}^K(f, \mathbb{F}_q)$.

Example 5.5.3 Let $\mathbb{F}_p$ be a prime field with p odd. Let α be a generator of the multiplicative group $\mathbb{F}_p^*$. Let $f[X]$ be a polynomial in $\mathbb{F}_p[X]$ of degree m. By evaluating this polynomial in all elements $\alpha, \alpha^2, \ldots$ of $\mathbb{F}_p^*$ a codeword from a Reed-Solomon code (RS-code) is obtained. A binary sequence is obtained from this codeword by applying the map $\chi_2 : \mathbb{F}_p \to \mathbb{Z}/2\mathbb{Z}$ coordinatewise, defined by

$$\chi_2(a) = \begin{cases} 0 & \text{if } a = 0 \text{ or } \left(\frac{a}{p}\right) = 1, \\ 1 & \text{if } \left(\frac{a}{p}\right) = -1. \end{cases}$$

Here, $\left(\frac{a}{p}\right)$ denotes the Legendre symbol, which indicates whether a is a quadratic residue modulo p. If for example $p = 13$, $f[X] = X^2 + X$ and $\alpha = 2$ are chosen the codeword

$$(2, 6, 7, 7, 12, 3, 0, 2, 12, 4, 6, 4)$$

and corresponding binary sequence

$$(1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0)$$

are found.

As said before, it is possible to obtain codes using this construction. This was done in [Per91]. There, non-linear codes were found and investigated. It is possible to find interesting linear codes as is shown in the following example. After the example, the study of sequences is resumed.

Example 5.5.4 Choose $\mathcal{C}$ to be the projective line defined over some finite field $\mathbb{F}_q$ of odd characteristic. For $\mathcal{M} \subset \mathbb{F}_q(x)/(\mathbb{F}_q(x))^2$ the group generated by the residue classes of $x - \beta$ is chosen, with β in some non-empty subset S of $\mathbb{F}_q$. Then every element of $\mathcal{M}$ has a representative of the form $\prod_{\beta \in S}(x-\beta)^{\epsilon_\beta}$ with $\epsilon_\beta \in \{0,1\}$. As a vector space over $\mathbb{F}_2$, the group $\mathcal{M}$ has dimension $\#S$. Define χ_2 as in Example 5.5.3. Evaluating $\chi_2 \circ f$ in the set $\mathbb{F}_q \backslash S$ for all functions $f \in \mathcal{M}$, yields a binary linear code C of length $\#(\mathbb{F}_q \backslash S)$. Its dimension is less than or equal to $\min(\#S, \#(\mathbb{F}_q \backslash S))$.

Equality need not hold in this equation. Suppose for example that q is a square. The evaluation of the polynomial $f(X) = X^{\sqrt{q}} + X$ in an element a of $\mathbb{F}_q$ is either zero or a square. To see this note that for $a \in \mathbb{F}_q$ the relation $f(a)^{\sqrt{q}} = f(a)$ holds, and hence either $f(a) = 0$ or $f(a)^{(q-1)/2} = 1$. Thus, if S is chosen to be $S = \{a \in \mathbb{F}_q \,|\, a^{\sqrt{q}} + a = 0\}$, the polynomial $X^{\sqrt{q}} + x$ will correspond to the all-zero codeword. This means that in this case the dimension of the code cannot equal the cardinality of S.

Note that the curve given by the equation $Y^2 = X^{\sqrt{q}} + X$ has the maximum number of $\mathbb{F}_q$-rational points for its genus. As a matter of fact the number of $\mathbb{F}_q$-rational points can be seen to be $2q - \sqrt{q} + 1$, while its genus equals $(\sqrt{q}-1)/2$. The fact that it is maximal also follows from the fact that it can be covered by the Hermitian curve which has equation $Y^{\sqrt{q}+1} = X^{\sqrt{q}} + X$. Using the Hasse-Weil bound and investigating the curve $Y^2 = f(X)$, one can show that for sets S with cardinality strictly smaller than $\sqrt{q}$, only the zero-polynomial can give the all-zero codeword. This means that in this case the dimension of the resulting codes equals the cardinality of the set S.

Refining this argument, it follows that the following statement holds for the minimum distance d of these codes:

$$d \geq \frac{q - (\#S - 1)(\sqrt{q}-1)}{2} - \#S,$$

which is a non-trivial lower bound if $\#S < \sqrt{q}$.

In a similar way as in the previous section statements about the balance, autocorrelation and crosscorrelation of the sequences $S^K(f, P)$ will be given.

Definition 5.5.5 *Let $\mathcal{C}$ be an algebraic curve defined over $\mathbb{F}_q$. Let $1 < m < q-1$ be a divisor of $q-1$ and denote by $\chi_m : \mathbb{F}_q^* \to \mathbb{Z}/m\mathbb{Z}$ some surjective homomorphism of groups. Let $f \in \mathcal{F}_q(\mathcal{C})$. Define the following exponential sum:*

$$ES^K(\mathcal{C}, f) = \sum_{P \in \mathcal{C}^K(f, \mathbb{F}_q)} \zeta_m^{\chi_m(f(P))},$$

with $\zeta_m = exp(2\pi i/m)$.

For this exponential sum a bound exists similar to that of the exponential sum defined in Definition 5.4.3. See for example [Per91], where similar upper bounds are derived. The proof of the following proposition is analogous to that of Proposition

5.4.4. Instead of Artin-Schreier extensions, Kummer extensions are used (see for example [Sti93, Section[3.7]).

Proposition 5.5.6 *Let $\mathcal{C}$ be an algebraic curve of genus g defined over $\mathbb{F}_q$. Let $f \in \mathcal{F}_q(\mathcal{C})$ and suppose that $f \neq z^l$ for all $z \in \overline{\mathcal{F}(\mathcal{C})}$ and all divisors $l > 1$ of m. Write $r_P = \gcd(m, v_P(f)) > 0$. Then the following holds:*

$$\left|ES^K(\mathcal{C}, f)\right| \leq \left(2g - 2 + \sum_{P \in \mathcal{E}(\overline{\mathbb{F}}_q)} (1 - \frac{r_P - 1}{m-1})\right) \sqrt{q}.$$

The r_P occurring in the above proposition are standard in the theory of Kummer extensions. When the role of f is to be stressed, $r_P(f)$ will be written instead.

Theorem 5.5.7 *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_q$ of characteristic p. Let $f \in \mathcal{F}_q(\mathcal{E})$ be a function and write $k = k(\mathcal{E}, \mathbb{F}_q)$. Let P be a generator of the group $[k]\mathcal{E}(\mathbb{F}_q)$. Suppose that the polynomial $T^m - f \circ [k]$ is absolutely irreducible. Then*

$$B_{S^K(f,P)} \leq \frac{1}{N}\left(N - \#(\mathcal{E}^K(f, \mathbb{F}_q) \cap \langle P \rangle) + \sum_Q (1 - \frac{r_Q(f) - 1}{m-1})\sqrt{q}\right),$$

with $N = \#\langle P \rangle$.

Proof: Note that $v_Q(f \circ [k]) = v_{[k]Q}(f)$ and hence $r_Q(f \circ [k]) = r_{[k]Q}(f)$. Moreover, note that $r_Q = m$ for $Q \in \mathcal{E}(\mathbb{F}_q)$, if and only if $Q \in \mathcal{E}^K(f, \mathbb{F}_q)$. Hence it follows that $\mathcal{E}^K(f \circ [k], \overline{\mathbb{F}}_q) = [k]^{-1}\mathcal{E}^K(f, \mathbb{F}_q) \cap \langle P \rangle$. The rest of the proof is similar to that of Lemma 5.4.5. □

In the following corollary the weighted degree of a polynomial in two variables defined by $\mathrm{wdeg}(x) = 2$ and $\mathrm{wdeg}(y) = 3$ is used again.

Corollary 5.5.8 *Let the notation be as in the above theorem and suppose that $\mathcal{E}$ is given by a Weierstrass equation. Suppose that f is a non-trivial polynomial of total degree Δ in the coordinate functions x and y satisfying $\deg_x(f) \leq 2$. Suppose that $\gcd(\mathrm{wdeg}(f), m) = 1$. Then the following bound holds:*

$$B_{S^K(f,P)} \leq \frac{1}{N}\left(N - \#(\mathcal{E}^K(f, \mathbb{F}_q) \cap \langle P \rangle) + (3\Delta + 1)\sqrt{q}\right).$$

If $(\langle P \rangle \setminus \{\mathcal{O}\}) \subset \mathcal{E}^K(f, \mathbb{F}_q)$ is demanded as well, the bound

$$B_{S^K(f,P)} \leq \frac{1}{N}\left(1 + (3\Delta + 1)\sqrt{q}\right)$$

holds.

Proof: This follows from the above theorem by remarking that by Bézout's theorem (see for example [Wae40, Section 83]) f has at most 3Δ zeros on $\mathcal{E}$. Further note that the point $\mathcal{O}$ is the only pole f has on $\mathcal{E}$. These zeros and poles are the only points Q for which it can happen that $r_Q < m$. Using that $r_Q \geq 1$ for these points Q, the result follows. Note that $r_{\mathcal{O}}(f \circ [k(\mathcal{E})]) = r_{\mathcal{O}}(f) = \gcd(\mathrm{wdeg}(f), m) = 1$. Hence the polynomial $T^m - f \circ [k(\mathcal{E})]$ is absolutely irreducible by the theory of Kummer extensions. □

Note that it can be useful to rewrite f, using the equation of $\mathcal{E}$, in such a form that the total degree is minimal. This explains why the assumption $\deg_x(f) \leq 2$ is made instead of $\deg_y(f) \leq 1$, as was done previously.

Now some statements about the autocorrelation and crosscorrelations of these sequences will be given. Most of the proofs will be omitted since they are analogous to the proofs in the Artin-Schreier case.

Theorem 5.5.9 *Let $\mathcal{E}$ be an elliptic curve defined over the field $\mathbb{F}_q$ of characteristic p. Let $f \in \mathcal{F}_q(\mathcal{E})$ and choose $\alpha, \beta \in \mathbb{F}_q^*$. Write $k = k(\mathcal{E}, \mathbb{F}_q)$ and choose a generator P of the group $[k]\mathcal{E}(\mathbb{F}_q)$ and a number d satisfying $1 \leq d < N$ with $N = \#\langle P \rangle$. Define $h \in \mathcal{F}_q(\mathcal{E})$ by*

$$h(X) = \alpha f(X \oplus [d]P) - \beta f(X).$$

Suppose that the polynomial $T^m - h \circ [k]$ is absolutely irreducible. Then

$$|C_{S^K(f,P)}(d, \alpha, \beta)| \leq \frac{1}{N}\left(N - \#(\mathcal{E}^K(h, \mathbb{F}_q) \cap \langle P \rangle) + \sum_Q (1 - \frac{r_Q(h) - 1}{m-1})\sqrt{q}\right).$$

Corollary 5.5.10 *Let the notation be the same as in the above theorem. Suppose that $\mathcal{E}$ is given by a Weierstrass equation. Further assume that f is a non-trivial polynomial in the coordinate functions of total degree Δ satisfying $\deg_x(f) \leq 2$ and $\gcd(\mathrm{wdeg}(f), m) = 1$. Then*

$$C_{S^K(f,P)}(d, \alpha, \beta) \leq \frac{1}{N}\left(N - \#(\mathcal{E}^K(h, \mathbb{F}_q) \cap \langle P \rangle) + (3\Delta + 3\mathrm{wdeg}(f) + 2)\sqrt{q}\right).$$

If $(\langle P \rangle \setminus \{\mathcal{O}, [-d]P\}) \subset \mathcal{E}^K(h, \mathbb{F}_q)$ is demanded as well, the bound

$$C_{S^K(f,P)}(d, \alpha, \beta) \leq \frac{1}{N}\left(2 + (3\Delta + 3\mathrm{wdeg}(f) + 2)\sqrt{q}\right)$$

is found.

Proof: Again Bézout's theorem is to be used to estimate the number of zeros of the function h. Using the addition formula (see for example [Sil86, Section 3.2]), $(x, y) \oplus (a, b)$ can be written as $(g_1(x, y)/(x - a)^2, g_2(x, y)/(x - a)^3)$ with g_1 (respectively g_2) a polynomial in x and y of total degree less than or equal

to 2 (respectively 3). This means that $\alpha f((x,y) \oplus (a,b))$ can be written as $k(x,y)/(x-a)^{\mathrm{wdeg}(f)}$ with k a polynomial of total degree less than or equal to $\mathrm{wdeg}(f)$. Hence, after multiplying the rational function h with $(x-a)^{\mathrm{wdeg}(f)}$, a polynomial of total degree less than or equal to $\Delta + \mathrm{wdeg}(f)$ is obtained. This gives an upper bound for the total number of zeros of the function h while its poles are $\mathcal{O}$ and $-[d]P$. The rest of the proof is analogous to the proof of Corollary 5.5.8. □

Theorem 5.5.11 *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_q$ of characteristic p. Let P be a generator of the group $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_q)$ and write $N = \#\langle P\rangle$. Let f_1 and f_2 be two functions and choose $\alpha, \beta \in \mathbb{F}_q^*$ as well as a natural number $0 \le d < N$. Write $S_1 = S^K(f_1, P)$ and $S_2 = S^K(f_2, P)$. Define $h \in \mathcal{F}_q(\mathcal{E})$ by*

$$h(X) = \alpha f_1(X \oplus [d]P) - \beta f_2(X)$$

and suppose that the polynomial $T^m - h \circ [k(\mathcal{E})]$ is absolutely irreducible. Then the crosscorrelation is bounded by

$$|C_{S_1,S_2}(d,\alpha,\beta)| \le \frac{1}{N}\left(N - \#(\mathcal{E}^K(h,\mathbb{F}_q) \cap \langle P\rangle) + \sum_Q (1 - \frac{r_Q(h)-1}{m-1})\sqrt{q}\right).$$

Corollary 5.5.12 *Let the notation be the same as in the above theorem. Suppose that $\mathcal{E}$ is given by a Weierstrass equation. Further assume that f_1 and f_2 are non-trivial polynomials in the coordinate functions of total degree Δ_1 and Δ_2 with $\Delta_1 \ge \Delta_2$ and satisfying $\deg_x(f_i) \le 2$ with $i = 1, 2$. Further suppose that $\gcd(\mathrm{wdeg}(f_1), m) = 1$ if $d \ne 0$ and $\gcd(\mathrm{wdeg}(\alpha f_1 - \beta f_2), m) = 1$ if $d = 0$. Then*

$$C_{S_1,S_2}(d,\alpha,\beta) \le \frac{1}{N}\left(N - \#(\mathcal{E}^K(h,\mathbb{F}_q) \cap \langle P\rangle) + (3\mathrm{wdeg}(f_1) + 3\Delta_2 + 2)\sqrt{q}\right).$$

If $(\langle P\rangle \setminus \{\mathcal{O}, [-d]P\}) \subset \mathcal{E}^K(h,\mathbb{F}_q)$ holds additionally, the following bound holds:

$$C_{S_1,S_2}(d,\alpha,\beta) \le \frac{1}{N}\left(2 + (3\Delta_2 + 3\mathrm{wdeg}(f_1) + 2)\sqrt{q}\right).$$

5.6 Using linear recurrence relations on elliptic curves

In this section the balance and the period of a family of sequences obtained by using linear recurrence relations on the points of $\mathcal{E}$ will be investigated.

Suppose that G is a cyclic subgroup of $\mathcal{E}$ of order N generated by a point $P \in \mathcal{E}$. In this section, N will be assumed to be a prime number.

Let $r(X) = X^n + r_{n-1}X^{n-1} \cdots + r_0$ be a monic polynomial of degree $n > 1$ over $\mathbb{Z}/N\mathbb{Z}$ with $\gcd(r_0, N) = 1$ and let $\Omega(r, G)$ be the vector space over $\mathbb{Z}/N\mathbb{Z}$ of bi-infinite sequences of points in G that satisfy the linear recurrence relation with characteristic polynomial $r(X)$. This vector space has dimension n.

Suppose from now on that the characteristic polynomial $r(X)$ of the recursion is irreducible over $\mathbb{Z}/N\mathbb{Z}$. It is known from the theory of linear recurrencies (see for example [Shp99, Chapter 7]) that if $r(X)$ is irreducible, every sequence, apart from the zero sequence, has the same period $k(N,r)$. As a matter of fact $k(N,r)$ is the smallest positive integer k such that for every root α of $r(X)$ the relation $\alpha^k = 1$ holds.

Define $\Psi(r,G)$ to be the set of sequences $\Omega(r,G)$ modulo cyclic shifts.

Lemma 5.6.1 *Every point $Q \in G$ occurs the same number of times in sequences in $\Psi(r,G)$, i.e. the number of pairs*

$$\#\left\{(i,\psi) \mid 0 \leq i < k(N,r); \psi(i) = Q; \psi \in \Psi(r,G)\right\}$$

is independent of the choice of Q.

Proof: Since by definition of the recursion polynomial r the greatest common divisor of r_0 and N is one and this polynomial has degree n, each sequence in $\Psi(r,G)$ is uniquely determined by the choice of n consecutive points. Conversely, each n-tuple of points occurs exactly once in $\Psi(r,G)$ (note that this is modulo cyclic shifts). Since each point Q occurs equally often in the set of all n-tuples of points, this is the case in $\Psi(r,G)$ as well. □

Let $f \in \mathcal{F}_{q^e}$ be a function on $\mathcal{E}$. Now look at the sequence $S^{\mathrm{AS}}(f,P)$ which was defined earlier by

$$S^{\mathrm{AS}}(f,P) = \left\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}\left(f\left([i]P\right)\right)\right\}_{0 \leq i < N}.$$

Furthermore, define the set of sequences $\Psi_f(r,G)$ by applying the function f to each point in each sequence in $\Psi(r,G)$, and then taking the trace to the ground field $\mathbb{F}_q$ of the result:

$$\Psi_f(r,G) = \left\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}\left(f\left(\psi\right)\right) | \psi \in \Psi(r,G)\right\}.$$

Note that each sequence of points in $\Psi(r,G)$ corresponds with a sequence in $\Psi_f(r,G)$.

Here, the same convention as before is used, namely that $\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(Q)) = 0$ if $Q \notin \mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})$.

Theorem 5.6.2 *Choose a point $P \in \mathcal{E}$. Let G be the subgroup of $\mathcal{E}$ generated by P and suppose that its order is a prime N. Furthermore, let r be a monic recursion polynomial of degree n whose zeroth coefficient is coprime to N. Then the average balance of a sequence in $\Psi_f(r,G)$ is the same as the balance of the sequence $S^{\mathrm{AS}}(f,P)$.*

Proof: Start by defining the sequence T by merging all sequences of $\Psi_f(r,G)$. Since the order of points is not important in the definition of balance and because according to Lemma 5.6.1 each point occurs an equal number of times, we can reorder the points of T such that we get a number of copies of the sequence $S^{\mathrm{AS}}(f,P)$. Of course, this is the same sequence as $S^{\mathrm{AS}}(f,P)$ itself. Thus the average balance of sequences in $\Psi_f(r,G)$ is equal to the balance of the sequence $S^{\mathrm{AS}}(f,P)$. □

It is well known from the theory of linear recurrencies that the period of a sequence can be larger than the group order. So sequences defined in the above way can have a larger period than the sequences described in the previous sections. But this only applies to the sequences of points on $\mathcal{E}$. The next theorem links this period to the period of the generated pseudorandom sequence. The polynomial $r(X)$ is still supposed to be irreducible.

Theorem 5.6.3 *Let r and G be defined as in the above theorem. Suppose that the order of G is a prime N and that the degree of the recursion polynomial is n. Denote by $T_f(a)$ the number of points Q in G for which $\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(Q)) = a$. Suppose that all sequences in $\Psi_f(r, G)$ have period dividing $k(N, r)/d$. Then d is a divisor of $\gcd(N - k(N, r), T_f(a), N^n - 1)$ for all $a \in \mathbb{F}_q \backslash \{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(\mathcal{O}))\}$.*

Proof: All non-zero sequences in $\Psi_f(r, G)$ have as period a divisor of $k(N, r)/d$. Hence the number of times a occurs in the corresponding sequences is divisible by d. Write $b = \mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(\mathcal{O}))$. Then d divides $N^n T_f(a)$ with $a \in \mathbb{F}_q \setminus \{b\}$, and d divides $N^n T_f(b) - k(N, r)$. Since $k(N, r)$ divides $N^n - 1$, d divides $\gcd(N^n T_f(b) - k(N, r), T_f(a), N^n - 1)$ for all $a \in \mathbb{F}_q \setminus \{b\}$. Using $\sum_{a \in \mathbb{F}_q} T_f(a) = N$, for all $a \in \mathbb{F}_q \backslash \{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(\mathcal{O}))\}$ it is found (after eliminating $T_f(b)$) that d divides

$$\gcd(N^{n+1} - k(N, r), T_f(a), N^n - 1).$$

The result follows directly from this. □

Example 5.6.4 Let $\mathcal{E}$ be the elliptic curve defined over $\mathbb{F}_2$ given by the equation

$$Y^2 + Y = X^3 + X + 1.$$

Let G be a prime-order subgroup of $\mathcal{E}(\mathbb{F}_q)$ with q an odd power of 2. Using the addition formulas on $\mathcal{E}$, one can derive that for this curve $T_x(0) - 1 = T_x(1) = (N - 1)/2$. Hence, according to the above theorem, d divides $\gcd((N-1)/2, k(N, r) - 1)$. Take for example $r(X) = X^2 - X - 1$, the Fibonacci-recursion. The polynomial $r(X)$ is irreducible if and only if $N \equiv 2, 3 \pmod 5$. Assuming this, $k(N, r)$ divides $2N + 2$, since for any root ρ of $r(X)$ the relation $\rho^{2N+2} = (\rho \cdot \rho^N)^2 = (-1)^2 = 1$ holds. Here the fact was used that $r(X)$ is absolutely irreducible and hence that its roots are given by ρ and ρ^N. If furthermore the assumption $k(N, r) = 2N + 2$ is made, d divides 3. So the period of the resulting binary sequence is either the full period of the sequence of points on the elliptic curve, or a third of that.

5.7 Conclusion

Sequences obtained by using elliptic curves, as described in this section, certainly appear to be pseudorandom. When one generates such sequences in practice, all important statistical properties are well within the ranges one would expect a random sequence to have (they pass the FIPS 140-1 statistical tests for randomness).

However, only a few statistical properties are proven to hold. The balance and autocorrelation of these sequences are within the range that would be expected for a random sequence. The question of whether or not these sequences satisfy the expectation of other statistical tests (such as the next-bit test), remains open. Especially in the area of using linear recurrencies on elliptic curves, more research is needed to determine the cryptographic strength of sequences generated in this way.

References

[AW92a] M. Alabbadi and S. B. Wicker, *Cryptoanalysis of the Harn and Wang modification of the Xinmei digital signature scheme*, Electronic Letters **28** (1992), no. 18, 1756–1758.

[AW92b] M. Alabbadi and S. B. Wicker, *Security of Xinmei digital signature scheme*, Electronic Letters **28** (1992), no. 9, 890–891.

[AW93] M. Alabbadi and S. B. Wicker, *Digital signature scheme based on error–correcting codes*, Proceedings of 1993 IEEE International Symposium on Information Theory, 1993, p. 199.

[AW94] M. Alabbadi and S. B. Wicker, *Susceptibility of digital signature schemes based on error-correcting codes to universal forgery*, Error control, cryptology, and speech compression (Moscow, 1993), Springer, Berlin, 1994, pp. 6–12.

[AW95] M. Alabbadi and S. B. Wicker, *A digital signature scheme based on linear error-correcting block codes*, Advances in cryptology—ASIACRYPT '94 (Wollongong, 1994), Springer, Berlin, 1995, pp. 238–248.

[Bar97] A. Barg, *A large family of sequences with low periodic correlation*, Discrete Math. **176** (1997), no. 1-3, 21–27.

[BD02] P. H. T. Beelen and J. M. Doumen, *Pseudorandom sequences from elliptic curves*, Finite Fields with Applications to Coding Theory, Cryptography and Related Areas, Springer Verlag, 2002, pp. 37–52.

[BD03] P. H. T. Beelen and J. M. Doumen, *Two Mersenne-like families of prime numbers*, Manuscript, 2003.

[BDL97] D. Boneh, R. A. DeMillo, and R. J. Lipton, *On the importance of checking cryptographic protocols for faults (extended abstract)*, Advances in cryptology—EUROCRYPT '97 (Konstanz), Springer, Berlin, 1997, pp. 37–51.

[Bee01] P. H. T. Beelen, *Algebraic geometry and coding theory*, Ph.D. thesis, Eindhoven University of Technology, 2001.

[Ber97] T. A. Berson, *Failure of the McEliece public-key cryptosystem under message–resend and related–message attack*, Advances in Cryptology –

CRYPTO '97, Lecture Notes in Computer Science 1294 (B. S. Kaliski Jr., ed.), Springer-Verlag, 1997, pp. 213–220.

[BKT99] A. Barg, E. Krouk, and H. C. A. v. Tilborg, *On the complexity of minimum distance decoding of long linear codes*, IEEE Trans. Inform. Theory **45** (1999), no. 5, 1392–1405.

[Ble98] D. Bleichenbacher, *Chosen ciphertext attacks against protocols based on the RSA encryption standard PKCS #1*, Advances in Cryptology - CRYPTO 1998, Springer-Verlag, 1998, pp. 1–12.

[BMT78] E. R. Berlekamp, R. J. McEliece, and H. C. A. v. Tilborg, *On the inherent intractability of certain coding problems*, IEEE Trans. Information Theory **IT-24** (1978), no. 3, 384–386.

[Bom66] E. Bombieri, *On exponential sums in finite fields*, Amer. J. Math. **88** (1966), 71–105.

[Bre89] D. Bressoud, *Factorization and primality testing*, Springer-Verlag, New York, 1989.

[BS93] E. Biham and A. Shamir, *Differential cryptanalysis of the data encryption standard*, Springer-Verlag, New York, 1993.

[BS96] E. Bach and J. Shallit, *Algorithmic number theory. Vol. 1*, MIT Press, Cambridge, MA, 1996, Efficient algorithms.

[BS97] E. Biham and A. Shamir, *Differential fault analysis of secret key cryptosystems*, Lecture Notes in Computer Science **1294** (1997), 513–522.

[CFS01] N. Courtois, M. Finiasz, and N. Sendrier, *How to achieve a McEliece-based digital signature scheme*, Advances in Cryptology - ASIACRYPT 2001, Springer-Verlag, 2001, pp. 157–174.

[Cha95] F. Chabaud, *On the security of some cryptosystems based on error-correcting codes*, Advances in cryptology—EUROCRYPT '94 (Perugia), Springer, Berlin, 1995, pp. 131–139.

[Cox89] D. Cox, *Primes of the form $x^2 + ny^2$*, John Wiley & Sons Inc., New York, 1989, Fermat, class field theory and complex multiplication.

[CS98] A. Canteaut and N. Sendrier, *Cryptanalysis of the original McEliece cryptosystem*, Advances in cryptology—ASIACRYPT'98 (Beijing), Springer, Berlin, 1998, pp. 187–199.

[DH82] W. Diffie and M. E. Hellman, *New directions in cryptography*, Secure communications and asymmetric cryptosystems, Westview, Boulder, CO, 1982, pp. 143–180.

[DR98] J. Daemen and V. Rijmen, *Aes proposal: Rijndael*, 1998, www.esat.kuleuven.ac.be/~rijmen/rijndael/.

[Dum96] I. Dumer, *Suboptimal decoding of linear codes: partition technique*, IEEE Trans. Inform. Theory **42** (1996), no. 6, part 1, 1971–1986, Codes and complexity.

[ElG85] T. ElGamal, *A public key cryptosystem and a signature scheme based on discrete logarithms*, Advances in cryptology (Santa Barbara, Calif., 1984), Springer, Berlin, 1985, pp. 10–18.

[FO99] E. Fujisaki and T. Okamoto, *How to enhance the security of public-key encryption at minimum cost*, Public Key Cryptography, 1999, pp. 53–68.

[GBS00] G. Gong, T. Berson, and D. Stinson, *Elliptic curve pseudorandom sequence generators*, Selected areas in cryptography (Kingston, ON, 1999) (Berlin), Springer, 2000, pp. 34–48.

[GL02] G. Gong and C. C. Y. Lam, *Recursive sequences over elliptic curves*, Sequences and their Applications - SETA '01, Springer, London, 2002, pp. 182–196.

[GMT82] S. Goldwasser, S. Micali, and P. Tong, *Why and how to establish a private code on a public network*, 23rd annual symposium on foundations of computer science (Chicago, Ill., 1982), IEEE, New York, 1982, pp. 134–144.

[Gra01] J. Grantham, *Frobenius pseudoprimes*, Math. Comp. **70** (2001), no. 234, 873–891.

[Hal94] S. Hallgren, *Linear congruential generators over elliptic curves*, Tech. Report CS-94-143, Dept. of Comp. Sc., Carnegie Mellon Univ., 1994.

[Ham50] R. W. Hamming, *Error detecting and error correcting codes*, Bell System Technical Journal **29** (1950), 147–160.

[HGS99] C. Hall, I. Goldberg, and B. Schneier, *Reaction attacks against several public-key cryptosystems*, Proceedings of Information and Communication Security, ICICS'99, Springer-Verlag, 1999, pp. 2–12.

[HW79] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, fifth ed., The Clarendon Press Oxford University Press, New York, 1979.

[HW92] L. Harn and D. C. Wang, *Cryptoanalysis and modification of digital signature scheme based on error–correcting codes*, Electronic Letters **28** (1992), no. 2, 157–159.

[Kah67] D. Kahn, *The codebreakers: the story of secret writing*, MacMillan Publishing Company, New York, NY, USA, 1967.

[KFL85] T. Kasami, T. Fujiwara, and S. Lin, *An approximation to the weight distribution of binary linear codes*, IEEE Trans. Inform. Theory **31** (1985), no. 6, 769–780.

[KJJ99] P. Kocher, J. Jaffe, and B. Jun, *Differential power analysis*, Advances in Cryptology - CRYPTO 1999, Springer-Verlag, 1999, pp. 388–397.

[KKS97] G. Kabatianskii, E. Krouk, and B. Smeets, *A digital signature scheme based on random error-correcting codes*, Cryptography and coding (Cirencester, 1997), Springer, Berlin, 1997, pp. 161–167.

[KL95] I. Krasikov and S. Litsyn, *On the accuracy of the binomial approximation to the distance distribution of codes*, IEEE Trans. Inform. Theory **41** (1995), no. 5, 1472–1474.

[Kob01] N. Koblitz, *Almost primality of group orders of elliptic curves defined over small finite fields*, Experiment. Math. **10** (2001), no. 4, 553–558.

[Koc96] P. C. Kocher, *Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems*, Lecture Notes in Computer Science **1109** (1996), 104–113.

[KS00] D. R. Kohel and I. E. Shparlinski, *On exponential sums and group generators for elliptic curves over finite fields*, Algorithmic number theory (Leiden, 2000), Springer, Berlin, 2000, pp. 395–404.

[LMQ75] J. Leitzel, M. Madan, and C. Queen, *Algebraic function fields with small class number*, J. Number Theory **7** (1975), 11–27.

[Mat93] M. Matsui, *Linear cryptanalysis method for the DES cipher*, Advances in Cryptology: EUROCRYPT '93, Proceedings, Lofthus, Norway, May, 1993, Lecture Notes in Computer Science 765 (Berlin, Heidelberg, New York) (T. Helleseth, ed.), Springer Verlag, 1993, pp. 386–397.

[McE78] R. J. McEliece, *A public–key cryptosystem based on algebraic coding theory*, DSN Progress Report 42–44, Jet Propulsion Laboratory, Pasadena, 1978, pp. 114–116.

[MOV97] A. J. Menezes, P. C. v. Oorschot, and S. A. Vanstone, *Handbook of applied cryptography*, CRC Press, Boca Raton, FL, 1997, With a foreword by Ronald L. Rivest.

[MS77] F. MacWilliams and N. Sloane, *The theory of error-correcting codes. I*, North-Holland Publishing Co., Amsterdam, 1977, North-Holland Mathematical Library, Vol. 16.

[MS02] E. E. Mahassni and I. Shparlinski, *The uniformity of distribution of congruential generators over elliptic curves*, Sequences and their Applications - SETA '01, Springer, London, 2002, pp. 257–264.

[Oak00] M. Oakes, Private communication, 2000.

[Per91] M. Perret, *Multiplicative character sums and nonlinear geometric codes*, Eurocode '90 (Udine, 1990), Springer, Berlin, 1991, pp. 158–165.

[Rib96] P. Ribenboim, *The new book of prime number records*, Springer-Verlag, New York, 1996.

[RN89] T. R. N. Rao and K.-H. Nam, *Private-key algebraic-code encryptions*, IEEE Trans. Inform. Theory **35** (1989), no. 4, 829–833.

[RSA78] R. L. Rivest, A. Shamir, and L. Adleman, *A method for obtaining digital signatures and public-key cryptosystems*, Comm. ACM **21** (1978), no. 2, 120–126.

[Sha48] C. E. Shannon, *A mathematical theory of communication*, Bell System Technical Journal **27** (1948), 379–423 and 623–656.

[Shp99] I. E. Shparlinski, *Finite fields: theory and computation*, Kluwer Academic Publishers, Dordrecht, 1999, The meeting point of number theory, computer science, coding theory and cryptography.

[Sil86] J. H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, Berlin, 1986.

[Ste93] J. Stern, *A new identification scheme based on syndrome decoding*, Advances in Cryptology—CRYPTO '93 (D. R. Stinson, ed.), Lecture Notes in Computer Science, vol. 773, Springer-Verlag, 1993, pp. 13–21.

[Sti93] H. Stichtenoth, *Algebraic function fields and codes*, Springer-Verlag, Berlin, 1993.

[Til92] J. v. Tilburg, *Cryptanalysis of Xinmei digital signature scheme*, Electronic Letters **28** (1992), no. 20, 1935–1936.

[Til93a] H. C. A. v. Tilborg, *Error–correcting codes – a first course*, Chartwell Bratt Ltd, 1993.

[Til93b] J. v. Tilburg, *Cryptanalysis of the Alabbadi–Wicker digital signature scheme*, Proceedings of Fourteenth Symposium on Information Theory in the Benelux, 1993, pp. 114–119.

[Til94] J. v. Tilburg, *Security-analysis of a class of cryptosystems based on linear error-correcting codes*, Technische Universiteit Eindhoven, Eindhoven, 1994, Dissertation, Technische Universiteit Eindhoven, Eindhoven, 1994.

[VDT02] E. Verheul, J. M. Doumen, and H. C. A. v. Tilborg, *Sloppy Alice attacks! Adaptive chosen ciphertext attacks on the McEliece cryptosystem*, Information, Coding and Mathematics, Kluwer Academic Publishers, Boston etc., 2002, pp. 99–119.

[Vos93] C. Voss, *Abschätzungen der Parameter von Spurcodes mit Hilfe algebraischer Funktionenkörper*, Ph.D. thesis, Universität Essen, 1993.

[VW00] J. F. Voloch and J. L. Walker, *Euclidean weights of codes from elliptic curves over rings*, Trans. Amer. Math. Soc. **352** (2000), no. 11, 5063–5076 (electronic).

[Wae40] B. L. v. d. Waerden, *Moderne Algebra*, J. Springer, Berlin, 1940.

[Wag83] S. S. Wagstaff, Jr., *Divisors of Mersenne numbers*, Math. Comp. **40** (1983), no. 161, 385–397.

[Wan90] X. M. Wang, *Digital signature scheme based on error–correcting codes*, Electronics Letters **26** (1990), no. 13, 898–899.

[XD99] S. Xu and J. M. Doumen, *An attack against the Alabbadi–Wicker scheme*, The 20th symposium on information theory in the Benelux, 1999.

[XDT03] S. Xu, J. M. Doumen, and H. C. A. v. Tilborg, *On the security of digital signature schemes based on error–correcting codes*, Designs, Codes and Cryptography **28** (2003), no. 2, 187–199.

[YC03] S. Yates and C. Caldwell, *The largest known primes*, 2003, `http://www.utm.edu/research/primes/ftp/all.txt`.

Index

Acknowledgements

First of all I would like to thank my utilization committee, under the adept leadership of Henk van Tilborg, for the guidance they have given to my research. Considering the comparison of the original planning for my four years to the subjects discussed here, I am grateful for the freedom they granted me in choosing my favourite subjects.

My thanks go out as well to Henk van Tilborg, Arjen Lenstra and Berry Schoenmakers for suffering through the iterations of the thesis in front of you. Their remarks, both mathematical and linguistic, have improved my thesis considerably. Also, many thanks to the innumerable Ph.D. students who preceded me and who continually refined the used LaTeX style, to which I now made my own contribution.

Furthermore I want to thank many other people with whom I had many interesting mathematical discussions. To name but a few, I am thinking of Peter Beelen, Aart Blokhuis, Andries Brouwer, Jan Draisma, Ralf Gramlich, O.P. Lossers, Sander van Rijnswou, Berry Schoenmakers, Martijn Stam and all the other people I will mention here only implicitly.

Finally I would like to thank the technology foundation STW for funding my Ph.D. position in the project *Strong Authentication Methods*, number EWI.4536.

Samenvatting

In dit proefschrift worden een aantal toepassingen van coderingstheorie binnen de cryptografie behandeld. Het klassieke voorbeeld hiervan is het McEliece cryptosysteem, waarin een willekeurige lineaire code wordt gebruikt om berichten te kunnen versleutelen. De veiligheid van dit systeem is gebaseerd op één van de fundamentele problemen uit de coderingstheorie, namelijk het feit dat het decoderen in een willekeurige code een NP-compleet probleem is.

Alhoewel de veiligheid van dit cryptosysteem nog fier overeind staat, moet (zoals altijd) de implementatie van het cryptosysteem zorgvuldig gebeuren. Mocht een aanvaller de mogelijkheid hebben om (herhaaldelijk) informatie te krijgen over het al dan niet juist gevormd zijn van een cijfertekst, dan is het mogelijk om de versleuteling van een willekeurige cijfertekst ongedaan te maken, zoals in dit proefschrift is aangetoond. In de praktijk kan het eenvoudig zijn om zulke informatie te vergaren: de aanvaller zou bijvoorbeeld (op wat voor manier dan ook) toegang kunnen hebben tot de foutmeldingen die het ontcijfer-apparaat van de ontvanger produceert, of het zou kunnen zijn dat dit apparaat geautomatiseerd is en standaard een foutmelding terug stuurt als de ontvangen cijfertekst niet kan worden ontcijferd. Voor het geval dat het niet kan worden uitgesloten dat een aanvaller deze informatie bemachtigt, worden een aantal suggesties gedaan om de besproken aanval tegen te gaan.

Een andere belangrijke toepassing van cryptografie zijn zogenaamde digitale handtekeningen. Het doel hiervan is het digitale equivalent van een schriftelijke handtekening onder een papieren document. In 1990 werd door Wang een systeem voorgesteld dat gebaseerd is op coderingstheorie. Daarna werd dit systeem in relatief korte tijd verscheidene malen gebroken en weer gerepareerd. In dit proefschrift wordt een korte historie hiervan gegeven. Tevens wordt er aangetoond dat in de laatste incarnatie van deze familie handtekeningenschema's een vervalsing van een handtekening mogelijk is, als men alleen de publieke sleutel kent. Tevens wordt nogmaals beargumenteerd dat een dergelijke opzet voor een handtekeningenschema tot falen gedoemd is.

Een ander aspect van cryptografie is het genereren van pseudorandom rijen. Deze worden niet alleen in vele cryptografische protocollen direct gebruikt, maar ook heeft elk cryptosysteem een geheime sleutel nodig. De praktijk leert dat het niet verstandig is om die sleutel door mensen te laten kiezen (als het cryptosysteem dat al toelaat), maar dat het beter is om deze willekeurig te nemen. Helaas is het genereren van willekeurige getallen relatief duur. Edoch is een iets zwakkere

eis voldoende, namelijk dat deze getallen random lijken te zijn. In dit proefschrift wordt een methode beschreven om pseudorandom rijen te genereren met behulp van recurrente betrekkingen op elliptische krommen. Tevens wordt een aanzet gedaan tot het bewijzen van de cryptografische sterkte van deze generator.

In cryptografie worden vaak priemgetallen gebruikt. Onder andere hebben ze een directe toepassing in het RSA cryptosysteem. Deze praktische toepassing van priemgetallen heeft het fundamentele onderzoek ernaar zeker gestimuleerd. In dit proefschrift wordt daaraan een steentje bijgedragen met de introductie van twee nieuwe families getallen, waarvan de primaliteit efficiënt getest kan worden. Beide families vertonen grote verwantschap met de Mersenne getallen. Onder andere is in dit proefschrift het Wagstaff-vermoeden gegeneraliseerd naar de distributie van priemgetallen in de twee nieuwe families.

Curriculum Vitae

Jeroen Doumen werd geboren op 1 juli 1975 te Warstein, Duitsland. Na te zijn teruggekeerd naar Nederland, behaalde hij aldaar bij het Katholiek Gymnasium Rolduc te Kerkrade in 1993 het gymnasium diploma. Daarna verhuisde hij naar Leiden om aan de Leidse universiteit wis– en natuurkunde te gaan studeren. Beide studies werden met succes afgerond: hij studeerde eind 1998 af bij prof.dr. J.M.J. van Leeuwen in de theoretische natuurkunde met de scriptie *„Criticality in the Transverse Ising Model“*, en in de wiskunde studeerde hij begin 1999 bij drs. M.A.J.G. van der Vlugt af met de scriptie *„Gewichten van deelcodes en krommen met veel punten“*.

Eind 1998 begon hij aan de Technische Universiteit Eindhoven als AIO bij prof.dr.ir. H.C.A. van Tilborg aan het STW-projekt EWI.4536, getiteld *„Strong Authentication Methods“*. Het resultaat van deze arbeid ligt voor U.