Weaknesses in the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ Hashing Scheme

Rainer Steinwandt, Markus Grassl, Willi Geiselmann, and Thomas Beth

Institut für Algorithmen und Kognitive Systeme,
Fakultät für Informatik, Universität Karlsruhe,
Am Fasanengarten 5, 76 128 Karlsruhe, Germany,
{steinwan,grassl,geiselma,EISS_Office}@ira.uka.de.

Abstract. We show that for various choices of the parameters in the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme, suggested by Tillich and Zémor, messages can be modified without changing the hash value. Moreover, examples of hash functions "with a trapdoor" within this family are given. Due to these weaknesses one should impose at least certain restrictions on the allowed parameter values when using the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme for cryptographic purposes.

1 Introduction

At CRYPTO '94 Tillich and Zémor [11] have proposed a class of hash functions based on the group $\mathrm{SL}_2(\mathbb{F}_{2^n})$, the group of 2×2-matrices with determinant 1 over $\mathbb{F}_{2^n}$. The hash functions are parameterized by the degree n and the defining polynomial $f(X)$ of $\mathbb{F}_{2^n}$. The hash value $H(m) \in \mathrm{SL}_2(\mathbb{F}_{2^n})$ of some message m is a 2×2-matrix.

At ASIACRYPT '94 a first "attack" on this hash function was proposed by Charnes and Pieprzyk [3]. They showed that the hash function is weak for some particular choices of the defining polynomial $f(X)$. However, for any chosen hash function it is easy to check if it is resistant against this attack—the order of the generators of $\mathrm{SL}_2(\mathbb{F}_{2^n})$ has to be large. This can easily be calculated. Moreover, Abdukhalikov and Kim [1] have shown that an *arbitrary* choice of $f(X)$ results in a scheme vulnerable to Charnes' and Pieprzyk's attack only with a probability of approximately 10^{-27}.

Some additional structure of the group $\mathrm{SL}_2(\mathbb{F}_{2^n})$ was used by Geiselmann [5] to reduce the problem of finding collisions to the calculation of discrete logarithms in $\mathbb{F}_{2^n}$ or $\mathbb{F}_{2^{2n}}$ (which is feasible for the proposed values of $n \in \{130, \ldots, 170\}$). The drawback of this "attack" is the extremely long message required for such a collision. (E. g., the collision given in [5] for $n = 21$ has a length of about 237 000 bits.).

The main advantage of the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme to other schemes is the algebraic background that yields some proven properties about

distribution, shortest length of collisions [11, 15, 16], and allows a parallelization of the calculation: it holds $H(m_1|m_2) = H(m_1) \cdot H(m_2)$, where $m_1|m_2$ denotes the concatenation of the two messages m_1, m_2. This parallelization property is very helpful in some applications so that Quisquater and Joye have suggested to use this hash function for the authentication of video sequences [10], despite the weaknesses already known (more information on parallelizable hash functions can be found in [2, 4]).

In this paper we describe some weaknesses in the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme that also affect the generalization of this scheme to arbitrary finite fields suggested in [1]: as shown in [11], any collision in the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme involves at least one bitstring of length $\ell \geq n$. Hence it is infeasible to search directly for a collision among the more than 2^n bitstrings of length $\ell \geq n$. But using some structural properties of the group $\mathrm{SL}_2(\mathbb{F}_{p^n})$ we show that for several choices of the parameters it is possible to find short bitstrings that hash to an element of small order in $\mathrm{SL}_2(\mathbb{F}_{p^n})$. Repeating such a bitstring several times, a message can be modified without changing its hash value. In the case where adequate subfields of $\mathbb{F}_{p^n}$ exist, this approach works quite efficiently. Cases where n is prime are left to be resistant to this kind of attack. However, we show that—independent of n being prime or not—for all suggested values $130 \leq n \leq 170$ in the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme one can find a defining polynomial of $\mathbb{F}_{2^n}$ with a prescribed collision.

2 Preliminaries

2.1 The $\mathrm{SL}_2(\mathbb{F}_{2^n})$ Hashing Scheme

By $\mathbb{F}_{p^n}$ we denote the finite field with p^n elements. The hash function $H(m)$ of Tillich and Zémor [11] is based on the group $\mathrm{SL}_2(\mathbb{F}_{2^n})$:

Definition 1. *Let* $A := \left(\begin{smallmatrix} \alpha & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $B := \left(\begin{smallmatrix} \alpha & \alpha+1 \\ 1 & 1 \end{smallmatrix}\right)$ *be elements of* $\mathrm{SL}_2(\mathbb{F}_{2^n})$, *the group of* 2×2*-matrices with determinant* 1 *over* $\mathbb{F}_{2^n}$. *Here* $\alpha \in \mathbb{F}_{2^n}$ *is a root of a generating polynomial* $f(X)$ *of the field* $\mathbb{F}_{2^n} \simeq \mathbb{F}_2[X]/f(X)$.

Then, according to [11], the hash value $H(b_1 \dots b_r) \in \mathbb{F}_{2^n}^{2\times 2}$ *of a binary stream* $b_1 \dots b_r$ *is defined as the product* $M_1 \cdot \ldots \cdot M_r$ *with*

$$M_i := \begin{cases} A \text{ if } b_i = 0 \ ; \\ B \text{ if } b_i = 1 \ . \end{cases}$$

The straightforward generalization of this hashing scheme is to switch to $\mathbb{F}_{p^n}$, i. e., replacing $\mathrm{SL}_2(\mathbb{F}_{2^n})$ by $\mathrm{SL}_2(\mathbb{F}_{p^n})$, the group of 2×2-matrices with determinant 1 over $\mathbb{F}_{p^n}$, generated by $A = \left(\begin{smallmatrix} \alpha & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} \alpha & \alpha-1 \\ 1 & 1 \end{smallmatrix}\right)$ (see Proposition 2).

2.2 Some Properties of $\mathbf{SL_2(\mathbb{F}_{p^n})}$

As stated before, we use some properties of the group $\mathrm{SL}_2(\mathbb{F}_{p^n})$ to find collisions of relatively short length. First we recall the structure of the projective special linear group $\mathrm{PSL}_2(\mathbb{F}_{p^n})$ which will prove useful for analyzing $\mathrm{SL}_2(\mathbb{F}_{p^n})$. Denoting the cyclic group with r elements by C_r we have (see, e. g., [7, Kapitel II, Satz 8.5])

Proposition 1. *Any non-identity element of* $\mathrm{PSL}_2(\mathbb{F}_{p^n})$ *is either contained in an elementary abelian p-Sylow group* $C_p^n \cong \mathfrak{P}$*, in a cyclic subgroup* $C_{(p^n-1)/k} \cong \mathfrak{U}$*, or in a cyclic subgroup* $C_{(p^n+1)/k} \cong \mathfrak{S}$*, where* $k = \gcd(p-1, 2)$*. Thus the group can be written as a disjoint union of sets*

$$\mathfrak{G} := \mathrm{PSL}_2(\mathbb{F}_{p^n}) = \mathfrak{E} \uplus \mathcal{P} \uplus \mathcal{U} \uplus \mathcal{S}$$

where the sets $\mathcal{E}$*,* $\mathcal{P}$*,* $\mathcal{U}$*, and* $\mathcal{S}$ *are defined by the disjoint unions*

$$\mathcal{E} := \mathfrak{E} = \{id\} \qquad \mathcal{U} := \biguplus_{g \in \mathfrak{G}} (\mathfrak{U}^g \setminus \mathfrak{E})$$

$$\mathcal{P} := \biguplus_{g \in \mathfrak{G}} (\mathfrak{P}^g \setminus \mathfrak{E}) \qquad \mathcal{S} := \biguplus_{g \in \mathfrak{G}} (\mathfrak{S}^g \setminus \mathfrak{E}) \ .$$

(Here $\mathfrak{H}^g := \{g^{-1}hg \colon h \in \mathfrak{H}\}$ *is the conjugate of* $\mathfrak{H}$ *with respect to* g*.)*

This yields immediately the structure of the group $\mathrm{SL}_2(\mathbb{F}_{p^n})$:

Corollary 1. *As matrix, any element of* $\mathrm{SL}_2(\mathbb{F}_{p^n})$ *can be written as* $\pm M$ *where* $M \in \mathfrak{E} \cup \mathcal{P} \cup \mathcal{U} \cup \mathcal{S}$*.*

Proof. By definition, $\mathrm{PSL}_2(\mathbb{F}_{p^n}) = \mathrm{SL}_2(\mathbb{F}_{p^n})/(\mathrm{SL}_2(\mathbb{F}_{p^n}) \cap \mathfrak{Z})$ where $\mathfrak{Z} = \{a \cdot I_2 \colon a \in \mathbb{F}_{p^n}^\times\}$ as matrix group (with I_2 denoting the 2×2 identity matrix). For $p = 2$, $\mathrm{SL}_2(\mathbb{F}_{p^n}) \cap \mathfrak{Z} = \mathfrak{E}$ and hence $\mathrm{SL}_2(\mathbb{F}_{p^n}) \cong \mathrm{PSL}_2(\mathbb{F}_{p^n})$. For $p > 2$, $\mathrm{SL}_2(\mathbb{F}_{p^n}) \cap \mathfrak{Z} = \{\pm I_2\} \cong C_2$. So we may conclude that $\mathrm{SL}_2(\mathbb{F}_{p^n}) \cong C_2 \times \mathrm{PSL}_2(\mathbb{F}_{p^n})$. □

Further properties of the group $\mathrm{SL}_2(\mathbb{F}_{p^n})$ are summarized as follows:

Remark 1. For any subfield $\mathbb{F}_{p^m} \leq \mathbb{F}_{p^n}$ of $\mathbb{F}_{p^n}$, $\mathrm{SL}_2(\mathbb{F}_{p^m})$ and all its conjugates are subgroups of $\mathrm{SL}_2(\mathbb{F}_{p^n})$.

Remark 2. The group $\mathrm{SL}_2(\mathbb{F}_{p^n})$ has $p^n(p^n+1)(p^n-1)$ elements. There are $p^{2n}-1$ elements of order p (cf. [7, Kapitel II, Satz 8.2]). Furthermore for all factors f of $(p^n-1)/\gcd(p-1,2)$ and $(p^n+1)/\gcd(p-1,2)$ there are elements of order f. In particular, for $p = 2$ there are elements of order 3.

Lemma 1. *The group* $\mathrm{SL}_2(\mathbb{F}_{2^n})$ *has* $2^{2n} + (-2)^n$ *elements of order* 3.

Proof. Let $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{F}_{2^n})$ be of order 3. Then $M^3 = I_2$ and $M \neq I_2$, hence the characteristic polynomial $f_M(X) = X^2+(a+d)X+(ad+bc)$ of M equals X^2+X+1. The number of common solutions of the equations $a + d = 1$ and $ad + bc = 1$ is computed as follows: for $b = 0$, c is unconstrained, and a and d are specified by $a+d = 1$ and $a^2+a+1 = 0$. The latter equation has two solutions iff n is even. For $b \neq 0$, a can take any value, and c and d are given by $d = a+1$ and $c = (a^2+a+1)/b$. □

Finally, we present relations between bitstrings and their hash value.

Proposition 2. *Let* $\mathbb{F}_{p^n} = \mathbb{F}_p(\alpha)$. *Then, as a matrix group,* $\mathrm{SL}_2(\mathbb{F}_{p^n})$ *is generated by*

$$A = \begin{pmatrix} \alpha & -1 \\ 1 & 0 \end{pmatrix} \quad \textit{and} \quad B = \begin{pmatrix} \alpha & \alpha - 1 \\ 1 & 1 \end{pmatrix} .$$

Furthermore, the hash value of a bitstring $m = b_1 \dots b_\ell$ *of length* ℓ *is of the form* M_{b_ℓ} *(where* b_ℓ *is the last bit of* m*) with*

$$M_0 = \begin{pmatrix} c_\ell(\alpha) & c_{\ell-1}(\alpha) \\ d_{\ell-1}(\alpha) & c_{\ell-2}(\alpha) \end{pmatrix} \quad \textit{and} \quad M_1 = \begin{pmatrix} c_\ell(\alpha) & d_\ell(\alpha) \\ c_{\ell-1}(\alpha) & d_{\ell-1}(\alpha) \end{pmatrix} .$$

Here $c_i, d_i \in \mathbb{F}_p[X]$ *are polynomials of degree* i.

If $M \in \mathrm{SL}_2(\mathbb{F}_{p^n})$ *is the hash value of a bitstring* m *of length* $\ell < n$, *the representation in the form* M_i *is unique and the bitstring* m *can be obtained by successively stripping the factors.*

Proof. (cf. also [11, proof of Lemma 3.5]) For a proof that A and B generate $\mathrm{SL}_2(\mathbb{F}_{p^n})$ see, e. g., [1]. Defining the degree of the zero polynomial to be -1, the statement is true for $H(0) = A = M_0$ and $H(1) = B = M_1$, i. e., bitstrings of length 1.

Assuming that $H(m|0)$ is of the form M_0, the hash value $H(m|00) = H(m|0) \cdot A$ is computed as

$$\begin{pmatrix} c_\ell(\alpha) & c_{\ell-1}(\alpha) \\ d_{\ell-1}(\alpha) & c_{\ell-2}(\alpha) \end{pmatrix} \cdot A = \begin{pmatrix} \alpha \cdot c_\ell(\alpha) + c_{\ell-1}(\alpha) & -c_\ell(\alpha) \\ \alpha \cdot d_{\ell-1}(\alpha) + c_{\ell-2}(\alpha) & -d_{\ell-1}(\alpha) \end{pmatrix} ,$$

i. e., $H(m|00)$ is of the form M_0 where the degrees of all polynomials are increased by one. Analogously, we can show that $H(m|10)$ is of the form M_0 and that $H(m|01)$ and $H(m|11)$ are of the form M_1.

Note that the minimal polynomial $f_\alpha(X) \in \mathbb{F}_p[X]$ of α over $\mathbb{F}_p$ has degree n. Hence, in polynomial representation of $\mathbb{F}_{p^n}$, no reduction occurs

when $\ell < n$ and the representation is unique. Furthermore, by inspection of the degrees of the polynomials, it is easy to decide if a given matrix M is of the form M_0 or M_1. This yields the final bit b_ℓ of a bitstring $m = m'|b_\ell$ hashing to M. Using the identities $H(m') = H(m'|0) \cdot A^{-1}$ and $H(m') = H(m'|1) \cdot B^{-1}$, we can strip off one factor and proceed similarly to determine the bitstring m'. □

3 Finding Elements of Small Order

If we know a bitstring that hashes to the identity matrix then this bitstring can be inserted into a given message at arbitrary positions without changing the hash value of that message. For practical purposes we are of course particularly interested in *short* bitstrings that hash to the identity matrix. In order to find such bitstrings we want to exploit Proposition 1 and Remark 2 which imply that in case of $(p^n - 1)/\gcd(p-1, 2)$ or $(p^n + 1)/\gcd(p-1, 2)$ having several small factors, the group $\mathrm{SL}_2(\mathbb{F}_{p^n})$ contains various elements of small order: instead of looking for arbitrary bitstrings hashing to the identity matrix I_2 we try to find very short bitstrings (say less than 50 bits) which hash to a matrix of small order (say less than 300).

One family of matrices which are promising candidates for being of small order is formed by the elements $M \in \mathrm{SL}_2(\mathbb{F}_{p^n})$ whose coefficients are contained in a proper subfield $\mathbb{F}_{p^m} \subsetneq \mathbb{F}_{p^n}$ already, because according to Proposition 1 the order of such a matrix is bounded by $p^m + 1$. Moreover, as the orders of similar matrices coincide, we are also interested in matrices $M \in \mathrm{SL}_2(\mathbb{F}_{p^n})$ that are similar to some $M' \in \mathrm{SL}_2(\mathbb{F}_{p^m})$ with $\mathbb{F}_{p^m} \subsetneq \mathbb{F}_{p^n}$ (i. e. $M = N^{-1} \cdot M' \cdot N$ for some non-singular matrix N with coefficients in an extension field of $\mathbb{F}_{p^m}$). By means of the trace operation (which computes the sum of the diagonal entries of a matrix) we can give the following characterization:

Proposition 3. *Let $M \in \mathrm{SL}_2(\mathbb{F}_{p^n})$ and $\mathbb{F}_{p^m} \leq \mathbb{F}_{p^n}$. Then*

$$M \text{ is similar to a matrix } M' \in \mathrm{SL}_2(\mathbb{F}_{p^m}) \iff \mathrm{Trace}(M) \in \mathbb{F}_{p^m} \ .$$

Proof. $\Longrightarrow$: Trivial.
$\Longleftarrow$: If the minimal polynomial of M is linear then $M = \pm I_2 \in \mathrm{SL}_2(\mathbb{F}_p) \leq \mathrm{SL}_2(\mathbb{F}_{p^m})$. So we assume w. l. o. g. that the minimal polynomial $m(X)$ of M is quadratic, i. e., $m(X) = X^2 - \mathrm{Trace}(M) \cdot X + 1 \in \mathbb{F}_{p^n}[X]$. Let λ_1, λ_2 be the (not necessarily distinct) eigenvalues of M. Then the Jordan normal form of M (as a matrix in the general linear group $\mathrm{GL}_2(\mathbb{F}_{p^n}(\lambda_1, \lambda_2))$) is either $\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$ or $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

In the former case we have $\lambda_1 = \lambda_2$ and $\det(M) = 1$ implies $\lambda_1 = \pm 1$, i. e., the Jordan normal form of M is contained in $\mathrm{SL}_2(\mathbb{F}_p) \leq \mathrm{SL}_2(\mathbb{F}_{p^m})$. In the latter case M has two different eigenvalues and is similar to the matrix $\left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$, which is also similar to $M' := \left(\begin{smallmatrix} \mathrm{Trace}(M) & -1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{F}_{p^m})$, as the characteristic polynomials of M and M' coincide. □

Corollary 2. *For $M \in \mathrm{SL}_2(\mathbb{F}_{p^n})$ with $\mathrm{Trace}(M) \in \mathbb{F}_{p^m} \leq \mathbb{F}_{p^n}$ we have $\mathrm{ord}(M) \leq p^m + 1$.*

Proof. Immediate from Proposition 1 and Proposition 3. □

So if the trace θ of a matrix $M \in \mathrm{SL}_2(\mathbb{F}_{p^n})$ generates only a small subfield $\mathbb{F}_p(\theta) \lneq \mathbb{F}_{p^n}$ then M is of small order. Of course, it is not sufficient to know a matrix $M \in \mathrm{SL}_2(\mathbb{F}_{p^n})$ of small order—we also need a short bitstring which hashes to M. Subsequently we want to verify that for certain choices of $\mathbb{F}_{p^n}$ such matrices and corresponding bitstrings can indeed be found.

3.1 Elements of Small Order, Functional Decomposition, and Intermediate Fields

Let $\mathbb{F}_{p^n} = \mathbb{F}_p[X]/f(X)$ where $f(X) \in \mathbb{F}_p[X]$ is an irreducible polynomial with a root α. Moreover, assume that $f(X)$ can be expressed as a functional composition $f(X) = (g \circ h)(X) = g(h(X))$ with (non-linear) "composition factors" $g(X), h(X) \in \mathbb{F}_p[X]$—such decompositions can be found efficiently (for more information about the problem of computing functional decompositions of polynomials cf., e. g., [6, 13, 14]):

Proposition 4. *(see [8, Theorem 9]) An irreducible polynomial of degree n over $\mathbb{F}_p$ can be tested for the existence of a nontrivial decomposition $g \circ h$ in NC (parallel time $\log^{O(1)} np$ on $(np)^{O(1)}$ processors). If such a decomposition exists, the coefficients of g and h can be computed in NC.*

If the trace of a matrix $M \in \mathrm{SL}_2(\mathbb{F}_{p^n})$ equals $h(\alpha)$, then we know that the extension $\mathbb{F}_p(\mathrm{Trace}(M))/\mathbb{F}_p$ is of degree $\deg(g(Y))$—note that irreducibility of $f(X)$ implies irreducibility of $g(Y)$. So according to Corollary 2 we have $\mathrm{ord}(M) \leq p^{\deg(g(Y))} + 1$. Consequently, if M can be expressed as a product in the generators $A, B \in \mathrm{SL}_2(\mathbb{F}_p(\alpha))$ with ℓ factors, then we obtain a bitstring of length $\leq \ell \cdot (p^{\deg(g(Y))} + 1)$ hashing to the identity $I_2 \in \mathrm{SL}_2(\mathbb{F}_2(\alpha))$.

So the idea for exploiting a decomposition $f(X) = (g \circ h)(X) \in \mathbb{F}_p[X]$ is to construct a bitstring that hashes to a matrix with trace $h(\alpha)$. As for

practical purposes we are only interested in short bitstrings, we restrict ourselves to bitstrings of length $\leq n$ (for the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme suggested in [11] we have $130 \leq n \leq 170$). In order to obtain a matrix with trace $h(\alpha)$, Proposition 2 suggests to choose the length of our bitstring equal to $\deg(h(X))$, and if $\deg(h(X))$ is not too large we can simply use an exhaustive search over all $2^{\deg(h(X))}$ bitstrings of length $\deg(h(X))$ to check whether a product of this length with the required trace exists.

3.2 Application to the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ Hashing Scheme of Tillich and Zémor

To justify the relevance of the above discussion we apply these ideas to the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme suggested in [11] (i. e., we choose $130 \leq n \leq 170$). As irreducible trinomials are of particular interest when implementing an $\mathbb{F}_{2^n}$ arithmetic in hardware, we first give an example of a decomposable irreducible trinomial:

Example 1. For the representation $\mathbb{F}_{2^{147}} \cong \mathbb{F}_2(\alpha)$ with α being a root of $X^{147} + X^{98} + 1 = (Y^3 + Y^2 + 1) \circ (X^{49}) \in \mathbb{F}_2[X]$ the orders of A and B compute to $\mathrm{ord}(A) = 2^{147} - 1, \mathrm{ord}(B) = 2^{147} + 1$, and the bitstring 11111011111110001000111110010100101010000111110110 (of length 49) hashes to a matrix with trace α^{49} and order 7. So we obtain a bitstring of length $7 \cdot 49 = 343$ that hashes to the identity.

Some more examples based on decomposable polynomials are listed in Table 2 in the appendix. Here we continue with an example in characteristic 3 which demonstrates that using characteristic 2 is not vital:

Example 2. For the representation $\mathbb{F}_{3^{90}} \cong \mathbb{F}_3(\alpha)$ with α being a root of
$X^{90} + X^{78} + X^{75} - X^{69} + X^{66} - X^{63} + X^{57} + X^{56} + X^{55} + X^{54} - X^{53} - X^{48} + X^{45} - X^{44} - X^{43} - X^{40} - X^{39} + X^{38} + X^{35} + X^{34} - X^{32} + X^{30} + X^{26} + X^{25} - X^{23} + X^{22} - X^{21} + X^{20} + X^{19} + X^{18} + X^{16} + 1 = (Y^6 - Y^4 + Y^2 + 1) \circ (X^{15} - X^{11} - X^{10} + X^8) \in \mathbb{F}_3[X]$
the orders of A and B are $\mathrm{ord}(A) = (3^{90} - 1)/4, \mathrm{ord}(B) = (3^{90} - 1)/52$, and the bitstring 000101111111100 (of length 15) hashes to a matrix with trace $\alpha^{15} - \alpha^{11} - \alpha^{10} + \alpha^8$ and order 56. So we obtain a bitstring of length $15 \cdot 56 = 840$ that hashes to the identity.

All of the examples mentioned make use of the existence of a nontrivial intermediate field $\mathbb{F}_p[Y]/g(Y)$ of the extension $\mathbb{F}_p \leq \mathbb{F}_p[X]/f(X)$ where $f(X) = (g \circ h)(X)$. But even in the case when $f(X)$ is indecomposable, there may exist short bitstrings where the trace of the hash value lies in a small intermediate field:

Example 3. The polynomial

$f(X) = X^{140} + X^{139} + X^{137} + X^{135} + X^{133} + X^{132} + X^{127} + X^{122} + X^{120} + X^{119} + X^{116} + X^{114} + X^{113} + X^{112} + X^{111} + X^{106} + X^{104} + X^{101} + X^{100} + X^{94} + X^{93} + X^{91} + X^{90} + X^{88} + X^{87} + X^{84} + X^{83} + X^{82} + X^{80} + X^{79} + X^{73} + X^{71} + X^{69} + X^{67} + X^{65} + X^{63} + X^{62} + X^{60} + X^{59} + X^{57} + X^{56} + X^{55} + X^{53} + X^{52} + X^{51} + X^{49} + X^{47} + X^{46} + X^{45} + X^{43} + X^{40} + X^{39} + X^{38} + X^{37} + X^{35} + X^{34} + X^{33} + X^{32} + X^{31} + X^{30} + X^{28} + X^{27} + X^{25} + X^{23} + X^{17} + X^{16} + X^{15} + X^8 + X^7 + X^6 + X^4 + X + 1$

is indecomposable. For $\mathbb{F}_{2^{140}} \cong \mathbb{F}_2(\alpha)$ with α a root of $f(X)$ we have $\mathrm{ord}(A) = \mathrm{ord}(B) = 2^{140} + 1$, and the bitstring $m := 1111111110101110$ hashes to a matrix $H(m)$ with $\mathrm{Trace}(H(m)) \in \mathbb{F}_{2^{10}}$ and $\mathrm{ord}(H(m)) = 25$.

To prevent the above attack we may choose $\mathbb{F}_{p^n} \cong \mathbb{F}_p[X]/f(X)$ in such a way that n is prime. Then f does not permit a nontrivial functional decomposition, and there are no nontrivial intermediate fields of $\mathbb{F}_{p^n}/\mathbb{F}_p$. Moreover, in order to make the search for elements of small order not unnecessarily easy, we may also try to fix n in such a way that the orders $(p^n \mp 1)/\gcd(p-1, 2)$ of the cyclic groups $\mathfrak{U}$, $\mathfrak{S}$ (cf. Proposition 1) do not have many small factors. Ideally, for $p = 2$ we have the following conditions fulfilled: n is prime and $(2^n - 1) \cdot (2^n + 1) = 3 \cdot p_1 \cdot p_2$ for some prime numbers p_1, p_2.

Using a computer algebra system like MAGMA one easily checks that for $\mathbb{F}_{p^n} = \mathbb{F}_{2^n}$ with $120 \leq n \leq 180$ the only possible choice for satisfying these conditions is $n = 127$; in particular none of the parameter values $130 \leq n \leq 170$ suggested in [11] meets these requirements. Furthermore, the next section shows that independent of the degree of the extension $\mathbb{F}_{p^n}/\mathbb{F}_p$ one should be careful about who is allowed to fix the actual representation of $\mathbb{F}_{p^n}$ as $\mathbb{F}_p[X]/f(X)$ used for hashing.

4 Deriving "Hash Functions with a Trapdoor"

In [16, Section 5.3] it is pointed out that for the $\mathrm{SL}_2(\mathbb{F}_p)$ hashing scheme discussed in [16] " ... some care should be taken in the choice of the prime number p, because finding simultaneously two texts and a prime number p such that those two texts collide for the hash function associated to p, is substantially easier than finding a collision for a given p ... "

In the sequel we shall discuss a related problem with the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme of [11]—being allowed to choose a representation of $\mathbb{F}_{2^n}$ we can select a hash function "with a trapdoor":

Example 4. For the representation $\mathbb{F}_{2^{167}} \cong \mathbb{F}_2(\alpha)$ with α being a root of

$$X^{167}+X^{165}+X^{161}+X^{160}+X^{158}+X^{157}+X^{156}+X^{155}+X^{154}+X^{152}+X^{150}+X^{148}+X^{145}+X^{143}+X^{142}+X^{140}+X^{138}+X^{137}+X^{134}+X^{131}+X^{130}+X^{126}+X^{125}+X^{123}+X^{119}+X^{118}+X^{117}+X^{116}+X^{115}+X^{113}+X^{112}+X^{111}+X^{107}+X^{105}+X^{104}+X^{99}+X^{96}+X^{93}+X^{91}+X^{89}+X^{88}+X^{86}+X^{85}+X^{82}+X^{81}+X^{80}+X^{77}+X^{76}+X^{74}+X^{73}+X^{71}+X^{65}+X^{64}+X^{62}+X^{61}+X^{58}+X^{57}+X^{54}+X^{53}+X^{51}+X^{49}+X^{48}+X^{47}+X^{45}+X^{40}+X^{38}+X^{37}+X^{35}+X^{34}+X^{33}+X^{30}+X^{29}+X^{27}+X^{26}+X^{25}+X^{21}+X^{19}+X^{17}+X^{14}+X^{13}+X^{12}+X^{10}+X^{6}+X^{5}+X^{2}+X+1 \in \mathbb{F}_2[X]$$

the orders of A and B are $\mathrm{ord}(A) = 2^{167}+1, \mathrm{ord}(B) = (2^{167}+1)/3$. Moreover, 167 is prime, and the prime factorizations of $(2^{167} \mp 1)$ compute to

$$2^{167}-1 = 2349023 \cdot 79638304766856507377778616296087448490695649\ ,$$
$$2^{167}+1 = 3 \cdot 623574031927851911766905528625614088386531218336 43\ .$$

At first glance these parameters look reasonable. However, the bitstring

	01010100	01101000	01101001	01110011	00100000	01101001
(ASCII)	T	h	i	s		i
	01110011	00100000	01110100	01101000	01100101	00100000
	s		t	h	e	
	01110111	01100001	01111001	00100000	01100001	00100000
	w	a	y		a	
	01110100	01110010	01100001	01110000	01100100	01101111
	t	r	a	p	d	o
	01101111	01110010	00100000	01100011	01100001	01101110
	o	r		c	a	n
	00100000	01101100	01101111	01101111	01101011	00100000
		l	o	o	k	
	01101100	01101001	01101011	01100101	00101110	00100000
	l	i	k	e	.	

(of length $42 \cdot 8 = 336$) hashes to a matrix with trace 0 and order 2. So we obtain a bitstring of length $336 \cdot 2 = 672$ that hashes to the identity.

The phenomenon of hash functions with a trapdoor is well-known (see, e. g., [9, 12]). For the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme deriving parameters with a trapdoor as in Example 4, is comparatively easy—the basic idea is to exploit the fact that, independent of the value of n, the group $\mathrm{SL}_2(\mathbb{F}_{2^n})$ always contains elements of order 2 and 3 (see Proposition 1 and Lemma 1): we start by fixing a bitstring which consists of two or three (depending on whether we want to have an element of order two or three) repetitions of an arbitrary bit sequence m. Then we compute the

"generic hash value" $H = H(X)$ of this bitstring, i. e., instead of using the matrices A and B we use the matrices

$$A_X = \begin{pmatrix} X & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B_X = \begin{pmatrix} X & X+1 \\ 1 & 1 \end{pmatrix} \ ,$$

where the generator α is replaced by an indeterminate X. Next, we compute the irreducible factors $f_1, \ldots, f_r$ of the greatest common divisor of the entries of the matrix $H - I_2$. Then choosing the field $\mathbb{F}_{2^n}$ as $\mathbb{F}_{2^n} \cong \mathbb{F}_2(\alpha)$ with α a root of some f_i guarantees that $H - I_2$ is in the kernel of the specialization $X \mapsto \alpha$. In other words, the bit sequence m hashes to a matrix of order at most two resp. three. Experiments show that it is quite easy to derive weak parameters for the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme in this way for all $127 \leq n \leq 170$ (see Table 1 in the appendix for some examples).

5 Constructing "Real" Collisions

We conclude by a simple example that illustrates the use of elements of small order for deriving "real" collisions, i. e. (short) non-empty bitstrings $m_1 \neq m_2$ that hash to the same value.

Remark 3. Let m be a bitstring hashing to a matrix of order $\mathrm{ord}(m)$. Then also each bitstring $\mathrm{rot}(m)$ derived from m through bit-wise left- or right-rotation hashes to a matrix of order $\mathrm{ord}(m)$.

Proof. Rotating the bitstring simply translates into conjugating the hash value with a non-singular matrix. Hence, as the order of a matrix is invariant under conjugation, the claim follows. □

In the following example Remark 3 is used for deriving a collision:

Example 5. We use the representation of $\mathbb{F}_{2^{140}}$ of Example 3. Applying a brute-force approach for constructing short products of A and B of small order one can derive the identities

$$(B^9ABAB^3A)^{25} = I_2 = (B^3ABAB^9A)^{25} \ .$$

As B^3ABAB^9A is similar to B^9AB^3ABA we get

$$(B^9ABAB^3A)^{25} = I_2 = (B^9AB^3ABA)^{25}$$

resp. after multiplication with $(B^9AB)^{-1}$ from the left and $(BA)^{-1}$ from the right

$$AB^3A(B^9ABAB^3A)^{23}B^9ABAB^2 = B^2ABA(B^9AB^3ABA)^{23}B^9AB^3A \ .$$

So we obtain two different bitstrings of length $5 + 16 \cdot 23 + 14 = 387$ that hash to the same value.

6 Summary and Conclusion

We have shown that for various choices of the parameters in the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme, suggested in [11], messages can be modified without changing the hash value. Moreover, we have given several examples of hash functions "with a trapdoor" within this family.

In order to avoid the attacks based on functional decomposition and intermediate fields presented in Section 3, one should choose n being prime. We dissuade from using the $\mathrm{SL}_2(\mathbb{F}_{2^n})$ hashing scheme or its generalization to $\mathrm{SL}_2(\mathbb{F}_{p^n})$ in case of n being composite. Moreover, Section 4 demonstrates that even in the case of n being prime it is fairly easy to find defining polynomials yielding hash functions with a trapdoor. Consequently, appropriate care should be taken in fixing the representation of $\mathbb{F}_{2^n}$ which is used for hashing (concerning the problem of avoiding trapdoors in hash functions cf., e. g., [12]).

Acknowledgments The first author is supported by grant DFG - GRK 209/3-98 "Beherrschbarkeit komplexer Systeme". Moreover, we would like to thank the referees for several useful remarks and references.

References

1. K. S. Abdukhalikov and C. Kim, *On the Security of the Hashing Scheme Based on* SL_2, in Fast Software Encryption – FSE '98, S. Vaudenay, ed., vol. 1372 of Lecture Notes in Computer Science, Springer, 1998, pp. 93–102.
2. M. Bellare and D. Micciancio, *A New Paradigm for Collision-Free Hashing: Incrementality at Reduced Cost*, in Advances in Cryptology – EUROCRYPT '97, W. Fumy, ed., vol. 1233 of Lecture Notes in Computer Science, Springer, 1997, pp. 163–192.
3. C. Charnes and J. Pieprzyk, *Attacking the* SL_2 *hashing scheme*, in Advances in Cryptology – ASIACRYPT '94, J. Pieprzyk and R. Safavi-Naini, eds., vol. 917 of Lecture Notes in Computer Science, Springer, 1995, pp. 322–330.
4. I. B. Damgård, *A Design Principle for Hash Functions*, in Advances in Cryptology – CRYPTO '89, G. Brassard, ed., vol. 435 of Lecture Notes in Computer Science, Springer, 1989, pp. 416–427.
5. W. Geiselmann, *A Note on the Hash Function of Tillich and Zemor*, in Cryptography and Coding, C. Boyd, ed., vol. 1025 of Lecture Notes in Computer Science, Springer, 1995, pp. 257–263.
6. J. Gutiérrez, T. Recio, and C. Ruiz de Velasco, *Polynomial decomposition algorithm of almost quadratic complexity*, in Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (AAECC-6), Rome, Italy, 1988, T. Mora, ed., vol. 357 of Lecture Notes in Computer Science, Springer, 1989, pp. 471–475.
7. B. Huppert, *Endliche Gruppen I*, vol. 134 of Grundlehren der mathematischen Wissenschaften, Springer, 1967. Zweiter Nachdruck der ersten Auflage.

8. D. Kozen and S. Landau, *Polynomial Decomposition Algorithms*, Journal of Symbolic Computation, 7 (1989), pp. 445–456.
9. B. Preneel, *Design principles for dedicated hash functions*, in Fast Software Encryption, R. Anderson, ed., vol. 809 of Lecture Notes in Computer Science, Springer, 1994, pp. 71–82.
10. J.-J. Quisquater and M. Joye, *Authentication of sequences with the* SL_2 *hash function: Application to video sequences.*, Journal of Computer Security, 5 (1997), pp. 213–223.
11. J.-P. Tillich and G. Zémor, *Hashing with* SL_2, in Advances in Cryptology – CRYPTO '94, Y. Desmedt, ed., vol. 839 of Lecture Notes in Computer Science, Springer, 1994, pp. 40–49.
12. S. Vaudenay, *Hidden Collisions on DSS*, in Advances in Cryptology – CRYPTO '96, N. Koblitz, ed., vol. 1109 of Lecture Notes in Computer Science, Springer, 1996, pp. 83–88.
13. J. von zur Gathen, *Functional Decomposition of Polynomials: The Tame Case*, Journal of Symbolic Computation, 9 (1990), pp. 281–300.
14. ———, *Functional Decomposition of Polynomials: The Wild Case*, Journal of Symbolic Computation, 10 (1990), pp. 437–452.
15. G. Zémor, *Hash Functions and Graphs With Large Girths*, in Advances in Cryptology – EUROCRYPT '91, D. W. Davies, ed., vol. 547 of Lecture Notes in Computer Science, Springer, 1991, pp. 508–511.
16. ———, *Hash Functions and Cayley Graphs*, Designs, Codes and Cryptography, 4 (1994), pp. 381–394.

Appendix: Examples

For each $127 \leq n \leq 170$ we easily found representations of $\mathbb{F}_{2^n} \cong \mathbb{F}_2[X]/f(X)$ together with a bitstring m of length n that hashes to a matrix of order 3. In Table 1, we list for each prime number in this range such a representation together with the corresponding bitstring. To illustrate that neither A nor B is of small order, the orders of A and B are also included in the table.

Note that all the examples have been derived by means of the computer algebra system MAGMA on usual SUN workstations at a university institute; neither specialized hard- or software nor extraordinary computational power have been used.

In Table 2 some representations of $\mathbb{F}_{2^n} \cong \mathbb{F}_2[X]/f(X)$ are listed, where the defining polynomial $f(X) = (g \circ h)(X)$ allows a nontrivial decomposition. In addition to a bitstring m that hashes to a matrix $H(m)$ of small order $\mathrm{ord}(H(m))$, the orders of A and B are also included. In the last column, we list the total length of the resulting bitstring hashing to the identity. As low weight polynomials are of particular interest for hardware implementations of $\mathbb{F}_{2^n}$ arithmetic, a main focus is on decomposable trinomials.

Table 1. Weak parameters (bitstrings m of length n and $\mathrm{ord}(H(m)) = 3$)

n	$f(X)$	$\mathrm{ord}(A)$	$\mathrm{ord}(B)$	m
127	$X^{127}+X^{125}+X^{117}+X^{113}+X^{109}+$ $X^{103}+X^{87}+X^{85}+X^{81}+X^{77}+$ $X^{71}+X^{23}+X^{21}+X^{17}+X^{13}+X^{7}+1$	$2^{127}-1$	$2^{127}-1$	$1^2 01^3 01^2 0^3 10^{114}$
131	$X^{131}+X^{123}+X^{121}+X^{115}+$ $X^{113}+X^{109}+X^{99}+X^{97}+X^{77}+$ $X^{67}+X^{65}+X^{13}+X^{3}+X+1$	$2^{131}+1$	$2^{131}-1$	$1^5 0^4 10^{121}$
137	$X^{137}+X^{135}+X^{125}+X^{119}+$ $X^{113}+X^{105}+X^{103}+X^{97}+X^{73}+$ $X^{71}+X^{65}+X^{9}+X^{7}+X+1$	$2^{137}+1$	$2^{137}-1$	10101010^{130}
139	$X^{139}+X^{133}+X^{127}+X^{123}+X^{121}+$ $X^{117}+X^{113}+X^{107}+X^{101}+X^{97}+$ $X^{75}+X^{69}+X^{65}+X^{11}+X^{5}+X+1$	$2^{139}+1$	$2^{139}-1$	$1^3 0^2 1^3 0^{131}$
149	$X^{149}+X^{143}+X^{137}+X^{133}+X^{129}+$ $X^{113}+X^{111}+X^{105}+X^{101}+X^{97}+$ $X^{85}+X^{79}+X^{73}+X^{69}+X^{65}+$ $X^{21}+X^{15}+X^{9}+X^{5}+X+1$	$\frac{2^{149}+1}{3}$	$2^{149}-1$	$10^3 10^{144}$
151	$X^{151}+X^{147}+X^{139}+X^{135}+$ $X^{131}+X^{107}+X^{103}+X^{99}+X^{87}+$ $X^{83}+X^{75}+X^{71}+X^{67}+X^{23}+$ $X^{19}+X^{11}+X^{7}+X^{3}+1$	$2^{151}-1$	$2^{151}-1$	1010^{148}
157	$X^{157}+X^{155}+X^{153}+X^{147}+X^{141}+$ $X^{137}+X^{131}+X^{109}+X^{105}+$ $X^{99}+X^{93}+X^{91}+X^{89}+X^{83}+$ $X^{77}+X^{73}+X^{67}+X^{29}+X^{27}+$ $X^{25}+X^{19}+X^{13}+X^{9}+X^{3}+1$	$2^{157}-1$	$2^{157}-1$	$101^2 0^2 10^{150}$
163	$X^{163}+X^{162}+X^{157}+X^{156}+X^{141}+$ $X^{140}+X^{109}+X^{108}+X^{99}+X^{98}+$ $X^{93}+X^{92}+X^{77}+X^{76}+X^{35}+$ $X^{34}+X^{29}+X^{28}+X^{13}+X^{12}+1$	$2^{163}-1$	$2^{163}-1$	$1^5 010^2 10^{153}$
167	$X^{167}+X^{165}+X^{163}+X^{153}+X^{149}+$ $X^{133}+X^{129}+X^{113}+X^{103}+X^{99}+$ $X^{97}+X^{89}+X^{85}+X^{69}+X^{65}+X^{39}+$ $X^{37}+X^{35}+X^{25}+X^{21}+X^{5}+X+1$	$2^{167}+1$	$2^{167}-1$	$101^3 0^2 1^2 010^2 10^{153}$

Table 2. Decomposable trinomials and non-monomial decompositions

n	$f(X)$	$\mathrm{ord}(A)$	$\mathrm{ord}(B)$	m	$\mathrm{ord}(H(m))$	total
147	$X^{147}+X^{49}+1=(Y^3+Y+1)\circ(X^{49})$	$2^{147}-1$	$2^{147}-1$	$0^3101^40^3101015^30^31^70^31^30^21010^51^2$	9	441
147	$X^{147}+X^{98}+1=(Y^3+Y^2+1)\circ(X^{49})$	$2^{147}-1$	$2^{147}+1$	$0^3101^40^3101015^30^31^70^31^30^21010^51^2$	7	343
155	$X^{155}+X^{62}+1=(Y^5+Y^2+1)\circ(X^{31})$	$2^{155}-1$	$\frac{2^{155}+1}{11}$	$1^70^411010001^7010^4$	31	961
155	$X^{155}+X^{93}+1=(Y^5+Y^3+1)\circ(X^{31})$	$2^{155}-1$	$2^{155}-1$	$1^70^411010001^7010^4$	31	961
156	$X^{156}+X^{91}+1=(Y^{12}+Y^7+1)\circ(X^{13})$	$\frac{2^{156}-1}{2^{12}-1}$	$2^{156}+1$	0110000011011	273	3549
162	$X^{162}+X^{81}+1=(Y^6+Y^3+1)\circ(X^{27})$	$2^{162}-1$	$2^{162}+1$	$0^210^610^2101^2010^21^20^31^2$	63	1701
162	$X^{162}+X^{135}+1=(Y^6+Y^5+1)\circ(X^{27})$	$2^{162}-1$	$2^{162}+1$	$0^210^610^2101^2010^21^20^31^2$	21	567
130	$(Y^{10}+Y^8+Y^4+Y^3+1)\circ(X^{13}+X^9)$	$2^{130}-1$	$2^{130}-1$	0010011001101	93	1209
133	$(Y^7+Y+1)\circ(X^{19}+X^{17}+X^7)$	$2^{133}-1$	$2^{133}-1$	1110101011100000000	43	817
135	$(Y^9+Y^8+Y^7+Y^6+Y^5+Y+1)\circ(X^{15}+X^{11})$	$2^{135}-1$	$2^{135}-1$	010101000001011	171	2565
140	$(Y^7+Y+1)\circ(X^{20}+X^{19}+X^{12}+X^{11}+X^4+X^3)$	$2^{140}-1$	$2^{140}+1$	10000000000000000000	43	860
147	$(Y^7+Y+1)\circ(X^{21}+X^{15}+X^5)$	$2^{147}-1$	$2^{147}-1$	111100001010000000000	43	903
152	$(Y^8+Y^4+Y^3+Y^2+1)\circ(X^{19}+X^{13}+X^{11}+X^5)$	$2^{152}-1$	$2^{152}-1$	1101101000001000000	51	969
153	$(Y^9+Y^8+Y^6+Y^5+Y^4+Y+1)\circ$ $(X^{17}+X^{11}+X^7+X^5)$	$2^{153}-1$	$2^{153}-1$	00000010000101111	19	323
160	$(Y^8+Y^4+Y^3+Y+1)\circ(X^{20}+X^{19}+X^{10}+X^9)$	$2^{160}-1$	$2^{160}-1$	11000010100100000000	257	5140
162	$(Y^9+Y^8+Y^6+Y^5+Y^4+Y+1)\circ$ $(X^{18}+X^{17}+X^{12}+X^{11}+X^8+X^7+X^6+X^5)$	$2^{162}-1$	$2^{162}-1$	101010010100000000	19	342
168	$(Y^8+Y^7+Y^5+Y^3+1)\circ(X^{21}+X^{17}+X^{11})$	$2^{168}-1$	$2^{168}+1$	110010001101000000000	51	1071
170	$(Y^{10}+Y^6+Y^5+Y^3+Y^2+Y+1)\circ(X^{17}+X^7+X)$	$2^{170}+1$	$2^{170}+1$	00000100011110001	25	425