J. Cryptology (1998) 11: 235–251

Journal of
CRYPTOLOGY

An Analysis of SAFER*

Sean Murphy
Information Security Group, Royal Holloway, University of London,
Egham, Surrey TW20 0EX, England

Communicated by Don Coppersmith

Received 12 June 1995 and revised 22 July 1997

Abstract. We investigate some of the algebraic properties of the SAFER block cipher when the message space is considered as a $\mathbb{Z}$-module. In particular, we consider the invariant $\mathbb{Z}$-submodules of the PHT layer and show how these invariant $\mathbb{Z}$-submodules give potential cryptographic weaknesses.

Key words. Block cipher, SAFER, Cryptanalysis, Invariant $\mathbb{Z}$-submodules.

1. Introduction

SAFER K-64 is a block cipher that was introduced by Massey at the 1993 Cambridge Security Workshop on Fast Software Encryption [7]. It operates on 64-bit blocks under the control of a 64-bit key. It is a "byte-oriented" cipher in that all the basic encryption operations are on bytes or pairs of bytes. At the 1994 Leuven Workshop on Cryptographic Algorithms, Massey presented a paper [8] which surveyed the first year's research on SAFER K-64 and defined SAFER K-128, which is SAFER with a 128-bit key. In this paper we investigate certain algebraic properties of the SAFER block cipher and show how these properties are a potential source of cryptographic weaknesses.

Following the original submission of this paper and a related key attack by Knudsen [6] in mid-1995, revised block ciphers SAFER SK-64 and SAFER SK-128 were proposed [6]. These differ from the original SAFER K-64/128 block ciphers only by using the new key schedule proposed by Knudsen [6]. There is also an analysis of SAFER based on truncated differentials [1], and an analysis that considers a revised SAFER encryption algorithm with a different nonlinear layer [11].

Each round of SAFER contains only one operation that mixes message bytes, the *PHT layer* (which exists to provide diffusion [7]). This PHT layer is a $\mathbb{Z}$-module homomorphism on the message space. Our analysis concentrates on the $\mathbb{Z}$-submodules of the message space that are preserved by the PHT layer, that is, the invariant $\mathbb{Z}$-submodules.

* The author acknowledges the support of the Nuffield Foundation.

These invariant $\mathbb{Z}$-submodules and their cosets are not diffused by the PHT layer, and so provide a method for the cryptanalyst to cope with the diffusion in SAFER in a variety of attacks, whatever the key schedule. This is the main result of this paper. However, the original key schedule of SAFER K-64/128 did not mix key bytes. This allowed us to find a projection onto a particular 4-byte invariant $\mathbb{Z}$-submodule that does not depend on a quarter of the key (under standard cryptographic assumptions).

We begin this paper by giving a description of SAFER and the original key schedule of SAFER K-64/128. In Section 4 we give a description of the invariant $\mathbb{Z}$-submodules of the PHT layer, and in Section 5 we show how these invariant $\mathbb{Z}$-submodules can be used to construct a Markov chain on cosets. In Section 6 we show how the original key schedule of SAFER K-64/128 gave rise to the property described above, and in Section 7 we give some other ways in which the invariant $\mathbb{Z}$-submodules of the PHT could be used for cryptanalysis. We finish with some conclusions.

2. Description of SAFER

SAFER is a block cipher that operates on 64-bit blocks considered as 8 bytes. It consists of a round transformation iterated r times followed by a final output transformation. Recommended values of r are 6 for SAFER K-64 and 10 for SAFER K-128. The key-

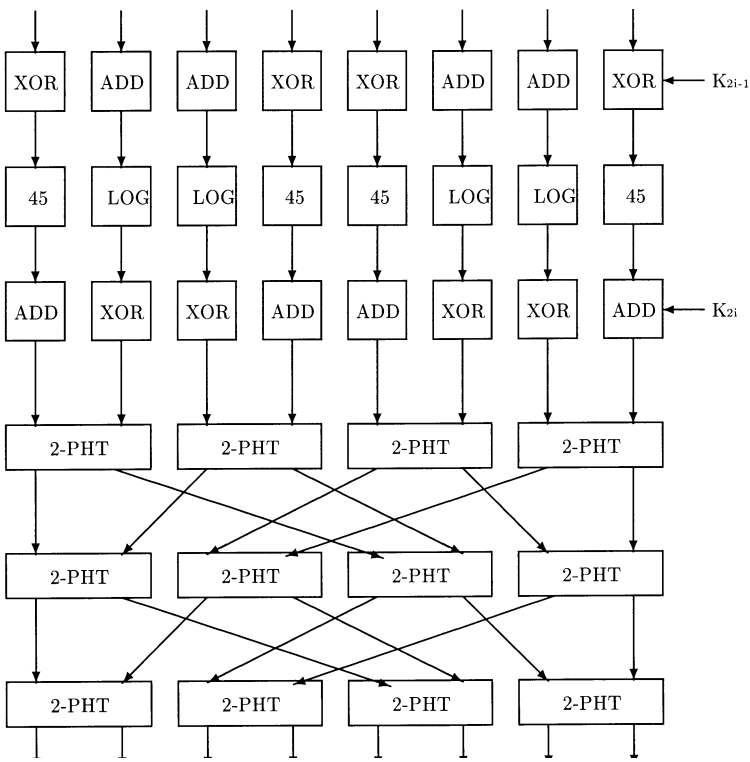

Fig. 1. Encryption round structure of SAFER.

scheduling, described below, gives $(2r + 1)$ 8-byte subkeys $K_1, \ldots, K_{2r+1}$. Subkeys K_{2i-1} and K_{2i} are used in round i, and the subkey K_{2r+1} is used in the output transformation. A diagram of the round function is given in Fig. 1. The ith round function is built from four basic operations.

1. *Mixed XOR/Addition Layer*: Bytes 1, 4, 5, 8 of the round input are XORed with bytes 1, 4, 5, 8 of subkey K_{2i-1}. Bytes 2, 3, 6, 7 of the round input are added bytewise (modulo 256) with bytes 2, 3, 6, 7 of subkey K_{2i-1}.
2. *Nonlinear Layer*: For a byte x, $45^{(x)}$ is defined to be 45^x modulo 257, where x is regarded as a number $0 \le x \le 255$, with the convention that $45^{(128)} = 0$. As 257 is prime and 45 is a primitive element modulo 257, this is an invertible function of a byte, and $\log_{45}(\cdot)$ is defined to be its inverse. The $45^{(\cdot)}$ transformation is applied to bytes 1, 4, 5, 8 of the output of the mixed XOR/addition layer and the $\log_{45}(\cdot)$ transformation to bytes 2, 3, 6, 7.
3. *Mixed Addition/XOR Layer*: Bytes 1, 4, 5, 8 of the output of the nonlinear layer are added bytewise (modulo 256) with bytes 1, 4, 5, 8 of subkey K_{2i}. Bytes 2, 3, 6, 7 of the output of the nonlinear layer are XORed with bytes 2, 3, 6, 7 of subkey K_{2i}.
4. *Pseudo-Hadamard Transform* (*PHT*) *Layer*: The transforms 2-PHT in Fig. 1 map the byte pair (a_1, a_2) to the byte pair $(2a_1 + a_2, a_1 + a_2)$, where addition is modulo 256. The effect of the three layers of 2-PHT transforms on the output v of the mixed addition/XOR layer is to map it to vM, where addition is modulo 256 and

$$M = \begin{pmatrix} 8 & 4 & 4 & 2 & 4 & 2 & 2 & 1 \\ 4 & 2 & 4 & 2 & 2 & 1 & 2 & 1 \\ 4 & 2 & 2 & 1 & 4 & 2 & 2 & 1 \\ 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 \\ 4 & 4 & 2 & 2 & 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

   The output of the PHT layer is the output of the round function.

The final output transformation after r rounds is an application of the mixed XOR/addition layer with the output of the rth round and the subkey K_{2r+1}.

Decryption using SAFER is carried out by reversing these operations and we do not describe it in detail.

3. Key Scheduling

The key scheduling for SAFER is again byte-oriented. For SAFER K-64 with 8-byte key K, let K^j denote the jth byte of K. The jth byte of subkey K_i, K_i^j $(j = 1, \ldots, 8)$, is defined by

$$K_i^j = ROL_{3(i-1)}(K^j) + B_i^j \qquad (i = 1, \ldots, 2r + 1),$$

where ROL_n denotes a left rotation of the byte by n positions, B_i^j are predefined "key biases", and addition is modulo 256. Note that $B_1^j = 0$, so $K_1 = K$.

For SAFER K-128, the 16-byte key K is split into two 8-byte halves K_a, K_b, so $K = (K_a, K_b)$. The jth byte of a subkey is defined by

$$\begin{aligned} K_{2i}^j &= ROL_{6(i-1)+3}(K_a^j) + B_{2i}^j \qquad (i = 1, \ldots, r), \\ K_{2i-1}^j &= ROL_{6(i-1)}(K_b^j) + B_{2i-1}^j \qquad (i = 1, \ldots, r+1), \end{aligned}$$

where the key biases B_i^j are the same as for SAFER K-64. Encryption under key $K = (K', K')$ for SAFER K-128 is identical to encryption under key K' for SAFER K-64.

For SAFER K-64, all of the subkey bytes used in the jth byte position depend solely on the jth byte of the key. For SAFER K-128, all of the even (K_{2i}) subkey bytes used in the jth byte position depend solely on the jth byte of the left half of the key, and all of the odd (K_{2i-1}) subkey bytes used in the jth byte position depend solely on the jth byte of the right half of the key. Thus for SAFER K-128, a byte of the 16-byte key gives subkey bytes that are all used either in an XOR operation or an addition operation, but never both. We term these two types of key bytes XOR and addition key bytes.

The new key schedules for SAFER SK-64 and SAFER SK-128 do mix key bytes [6]. We do not describe them here.

4. Algebraic Structure of the PHT Layer

Our comments on and our analysis of SAFER relate primarily to the algebraic properties of the PHT layer. An introduction to the algebra required is given in [5]. The PHT layer is a collection of transformations based on $\mathbb{Z}_{2^8}$, the ring of integers modulo $2^8 = 256$. Consider the 8-byte message space $V = \mathbb{Z}_{2^8}^8$ of SAFER. We can think of V as a module in three equivalent ways. We can regard V firstly as a free $\mathbb{Z}_{2^8}$-module of rank 8, or secondly as a torsion $\mathbb{Z}$-module that is annihilated by $2^8\mathbb{Z}$, or finally as the quotient $\mathbb{Z}$-module $\mathbb{Z}^8/2^8\mathbb{Z}^8$. In each case the PHT layer is an invertible module homomorphism $\alpha\colon V \to V$, where α has matrix M with respect to the standard basis. Our analysis of SAFER is based on the α-invariant submodules of V. Recall that a $\mathbb{Z}$-submodule U is an α-invariant $\mathbb{Z}$-submodule of V if $U\alpha \leq U$. However, in this case α is invertible, so if U is an α-invariant submodule, then $U\alpha = U$. The α-invariant $\mathbb{Z}$-submodules can be thought of as those $\mathbb{Z}$-submodules that are "preserved" by the PHT layer. For example, the simplest α-invariant submodules are $V_n = 2^{8-n}V$, the submodule obtained by considering the least significant n bits of each byte. We therefore begin our analysis of SAFER with a thorough investigation of α-invariant submodules.

Consider the eight-dimensional rational vector space $V_{\mathbb{Q}} = \mathbb{Q}^8$ and the free $\mathbb{Z}$-module $V_{\mathbb{Z}} = \mathbb{Z}^8$. Now α can be regarded as the linear transformation on $V_{\mathbb{Q}}$ and as the $\mathbb{Z}$-module homomorphism on $V_{\mathbb{Z}}$ given by the matrix M. Considered as a set embedded in $V_{\mathbb{Q}}$, α as a linear transformation on $V_{\mathbb{Q}}$ fixes the subset $V_{\mathbb{Z}}$. The characteristic polynomial of α on $V_{\mathbb{Q}}$, $f(x) = \mathrm{Det}(\alpha - xI)$ is given by

$$f(X) = 1 - 18X + X^2 - 18X^3 + 324X^4 - 18X^5 + X^6 - 18X^7 + X^8,$$

which factorizes over the integers as

$$f(X) = (1 - 18X + X^2)(1 - 3X + X^2)(1 + 3X + 8X^2 + 3X^3 + X^4).$$

$V_{\mathbb{Q}}$ therefore has the following three minimal α-invariant subspaces: $P_{\mathbb{Q}}$ (of dimension 2), $Q_{\mathbb{Q}}$ (of dimension 2) and $R_{\mathbb{Q}}$ (of dimension 4). These α-invariant $V_{\mathbb{Q}}$-subspaces are defined by

$$\begin{aligned} P_{\mathbb{Q}} &= \ker(I - 18\alpha + \alpha^2) \\ &= \langle e_1 - e_8 + d_2\,,\ e_1 + e_8 + d_1 \rangle, \\ Q_{\mathbb{Q}} &= \ker(I - 3\alpha + \alpha^2) \\ &= \langle 3e_1 - d_1 - d_2\,,\ d_1 - d_2 - 3e_8 \rangle, \\ R_{\mathbb{Q}} &= \ker(I + 3\alpha + 8\alpha^2 + 3\alpha^3 + \alpha^4) \\ &= \langle e_2 - e_5, e_3 - e_5, e_6 - e_4, e_7 - e_4 \rangle, \end{aligned}$$

where e_i denotes the ith standard basis vector, and $d_1 = e_2 + e_3 + e_5$ and $d_2 = e_4 + e_6 + e_7$. If we let $B_{\mathbb{Q}}$ denote the basis for $V_{\mathbb{Q}}$ given above in terms of $P_{\mathbb{Q}}$, $Q_{\mathbb{Q}}$ and $R_{\mathbb{Q}}$, then

$$B_{\mathbb{Q}} = \left\{ \begin{matrix} e_1 - e_8 + d_2,\ e_1 + e_8 + d_1,\ 3e_1 - d_1 - d_2, \\ d_1 - d_2 - 3e_8,\ e_2 - e_5,\ e_3 - e_5,\ e_6 - e_4,\ e_7 - e_4 \end{matrix} \right\},$$

and the change of basis transformation from the standard basis to $B_{\mathbb{Q}}$ is given by the matrix A (with determinant -225), where

$$A = \begin{pmatrix} 1 & 1 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 & 0 & -1 & -1 \\ 0 & 1 & -1 & 1 & -1 & -1 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & -1 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & -3 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$V_{\mathbb{Q}}$ has a direct sum decomposition as

$$V_{\mathbb{Q}} = P_{\mathbb{Q}} \oplus Q_{\mathbb{Q}} \oplus R_{\mathbb{Q}} \qquad \text{(where } \oplus \text{ here denotes direct sum).}$$

The $V_{\mathbb{Q}}$ subspaces $P_{\mathbb{Q}}$, $Q_{\mathbb{Q}}$ and $R_{\mathbb{Q}}$ are in fact pairwise orthogonal with respect to the standard inner product on $\mathbb{Q}$. The matrix of α with respect to the basis $B_{\mathbb{Q}}$ of $V_{\mathbb{Q}}$ is a block diagonal M' defined by

$$M' = \begin{pmatrix} M_P & 0 & 0 \\ 0 & M_Q & 0 \\ 0 & 0 & M_R \end{pmatrix},$$

where

$$M_P = \begin{pmatrix} 5 & 8 \\ 8 & 13 \end{pmatrix}, \qquad M_Q = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad M_R = \begin{pmatrix} -2 & 2 & -1 & 1 \\ -2 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix}.$$

We thus have the following three α-invariant $\mathbb{Z}$-submodules of $V_{\mathbb{Z}}$, as $V_{\mathbb{Z}}$ is fixed by α as a subset of $V_{\mathbb{Q}}$:

$$\begin{aligned} P_{\mathbb{Z}} &= P_{\mathbb{Q}} \cap V_{\mathbb{Z}} = \ker(I - 18\alpha + \alpha^2) \\ &= \langle e_1 - e_8 + d_2,\ e_1 + e_8 + d_1 \rangle, \\ Q_{\mathbb{Z}} &= Q_{\mathbb{Q}} \cap V_{\mathbb{Z}} = \ker(I - 3\alpha + \alpha^2) \\ &= \langle 3e_1 - d_1 - d_2,\ d_1 - d_2 - 3e_8 \rangle, \\ R_{\mathbb{Z}} &= R_{\mathbb{Q}} \cap V_{\mathbb{Z}} = \ker(I + 3\alpha + 8\alpha^2 + 3\alpha^3 + \alpha^4) \\ &= \langle e_2 - e_5, e_3 - e_5, e_6 - e_4, e_7 - e_4 \rangle. \end{aligned}$$

We can now define $T_{\mathbb{Z}} \leq V_{\mathbb{Z}}$ as the direct sum of these α-invariant $\mathbb{Z}$-submodules, so

$$T_{\mathbb{Z}} = P_{\mathbb{Z}} \oplus Q_{\mathbb{Z}} \oplus R_{\mathbb{Z}} \qquad \text{(where } \oplus \text{ here denotes direct sum).}$$

We note that $T_{\mathbb{Z}}$ is a proper $\mathbb{Z}$-submodule of $V_{\mathbb{Z}}$, as, for example, $d_1 \notin T_{\mathbb{Z}}$.

We can use this α-invariant decomposition on $T_{\mathbb{Z}}$ to give an α-invariant decomposition on V by using the following lemma:

Lemma [5, Lemma 8.1]. *Let L be a module over a ring (with 1), and suppose L is a direct sum $L = L_1 \oplus \cdots \oplus L_t$ of submodules $L_i \leq L$. For each i, let $N_i \leq L_i$, and let $N = \sum_{i=1}^{t} N_i$. If ν is the natural homomorphism $L \to L/N$, then*

$$\frac{L}{N} = L\nu = L_1\nu \oplus \cdots \oplus L_t\nu \cong \frac{L_1}{N_1} \oplus \cdots \oplus \frac{L_t}{N_t}.$$

Suppose we let ν^* denote the natural $\mathbb{Z}$-module homomorphism

$$\nu^*\colon V_{\mathbb{Z}} \to \frac{V_{\mathbb{Z}}}{2^8\mathbb{Z}^8} = V,$$

then ν^* gives the natural $\mathbb{Z}$-module homomorphism

$$\nu\colon T_{\mathbb{Z}} \to \frac{T_{\mathbb{Z}}}{T_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8} \cong T,$$

where $T = T_{\mathbb{Z}}\nu^*$ is a $\mathbb{Z}$-submodule of V. We apply the above lemma to the natural $\mathbb{Z}$-homomorphism ν and the decomposition $T_{\mathbb{Z}} = P_{\mathbb{Z}} \oplus Q_{\mathbb{Z}} \oplus R_{\mathbb{Z}}$. Noting that

$$T_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8 = (P_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8) + (Q_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8) + (R_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8),$$

we obtain

$$T \cong \frac{T_{\mathbb{Z}}}{T_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8} \cong \frac{P_{\mathbb{Z}}}{P_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8} \oplus \frac{Q_{\mathbb{Z}}}{Q_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8} \oplus \frac{R_{\mathbb{Z}}}{R_{\mathbb{Z}} \cap 2^8\mathbb{Z}^8}.$$

As $2^8\mathbb{Z}^8$ and $T_{\mathbb{Z}}$ are α-invariant $\mathbb{Z}$-submodules, α can be well defined as the induced $\mathbb{Z}$-module homomorphism on the quotient $\mathbb{Z}$-module T. Similarly the three quotient $\mathbb{Z}$-submodules in this decomposition are α-invariant.

We thus have following three α-invariant $\mathbb{Z}$-submodules of T:

$$\begin{aligned} P &= \ker(I - 18\alpha + \alpha^2) \\ &= \langle e_1 - e_8 + d_2,\ e_1 + e_8 + d_1 \rangle, \\ Q &= \ker(I - 3\alpha + \alpha^2) \\ &= \langle 3e_1 - d_1 - d_2,\ d_1 - d_2 - 3e_8 \rangle, \\ R &= \ker(I + 3\alpha + 8\alpha^2 + 3\alpha^3 + \alpha^4) \\ &= \langle e_2 - e_5, e_3 - e_5, e_6 - e_4, e_7 - e_4 \rangle, \end{aligned}$$

and the following decomposition of T as the direct sum of α-invariant $\mathbb{Z}$-submodules:

$$T = P \oplus Q \oplus R \qquad \text{(where } \oplus \text{ here denotes direct sum).}$$

Both T and V can be regarded as free $\mathbb{Z}_{2^8}$-modules of rank 8 with $T \leq V$. T is freely generated by the basis B, where

$$B = \left\{ \begin{matrix} e_1 - e_8 + d_2,\ e_1 + e_8 + d_1,\ 3e_1 - d_1 - d_2, \\ d_1 - d_2 - 3e_8,\ e_2 - e_5,\ e_3 - e_5,\ e_6 - e_4,\ e_7 - e_4 \end{matrix} \right\},$$

and V is freely generated by the standard basis. The change of basis transformation that maps the standard basis to B is given by the matrix A defined above. A is invertible as its determinant (-225) is a unit in $\mathbb{Z}_{2^8}$, so B is a basis for V as a free $\mathbb{Z}_{2^8}$-module [5, Lemmas 7.5 and 7.6]. Therefore as a free $\mathbb{Z}_{2^8}$-module, and hence as a $\mathbb{Z}$-module, $T = V$.

The decomposition of T as the direct sum of α-invariant $\mathbb{Z}$-submodules is thus a decomposition for V, and so we have

$$V = P \oplus Q \oplus R \qquad \text{(where } \oplus \text{ here denotes direct sum).}$$

Clearly, V, P, Q and R have sizes 2^{64}, 2^{16}, 2^{16} and 2^{32}, respectively.

In order to find further α-invariant $\mathbb{Z}$-submodules, we regard V as a (torsion) $\mathbb{Z}[X]$-module. In this module, scalar multiplication of a module element v by an integer polynomial $g(X)$ (an element of $\mathbb{Z}[X]$) is defined by

$$v \cdot g(X) = (v)(g(\alpha)),$$

that is, the image of v under the module homomorphism $g(\alpha)$. The role of this $\mathbb{Z}[X]$-module in finding α-invariant submodules is given by the following theorem.

Theorem. *U is an α-invariant $\mathbb{Z}$-submodule of V if and only if U is a $\mathbb{Z}[X]$-submodule of V.*

Proof. If U is an α-invariant $\mathbb{Z}$-submodule of V, then $U\alpha \leq U$. Thus, for any n and $\lambda_0, \ldots, \lambda_n$,

$$U \sum_{i=0}^{n} \lambda_i \alpha^i = \sum_{i=0}^{n} \lambda_i U \alpha^i \leq U,$$

so $U \cdot g(X) \leq U$ for any polynomial $g(X)$. Hence U is a $\mathbb{Z}[X]$-submodule.

Conversely, if U is a $\mathbb{Z}[X]$-submodule, then $U\alpha \leq U$, so U is an α-invariant submodule. □

We thus need to find the $\mathbb{Z}[X]$-submodules of the $\mathbb{Z}[X]$-module V. For any $\mathbb{Z}[X]$-submodule U, let

$$\text{ann}(U) = \{g(X) \in \mathbb{Z}[X] | U \cdot g(X) = 0\}$$

denote the annihilator of U in $\mathbb{Z}[X]$, an ideal in $\mathbb{Z}[X]$. This function gives an inclusion-reversing mapping from the $\mathbb{Z}[X]$-submodules of V to the ideals of $\mathbb{Z}[X]$. We can also define an inclusion-reversing mapping from the set of ideals of $\mathbb{Z}[X]$ to the $\mathbb{Z}[X]$-submodules of V. Accordingly, for any ideal I of $\mathbb{Z}[X]$, let

$$N_V(I) = \{v \in V | v \cdot I = 0\} \leq V$$

denote the "null" $\mathbb{Z}[X]$-submodule of V of the ideal I. N_P, N_Q and N_R can be similarly defined as null submodules in P, Q and R. However, not every $\mathbb{Z}[X]$-submodule is the null $\mathbb{Z}[X]$-submodule of some ideal, for example, $\mathbb{Z}[X]$-submodule $\{0, 2^7(e_3 + e_7)\}$, which is fixed by α, but is a proper $\mathbb{Z}[X]$-submodule of fixed point $\mathbb{Z}[X]$-submodule $\ker(\alpha - I)$. However, any $\mathbb{Z}[X]$-submodule U is a $\mathbb{Z}[X]$-submodule of the null $\mathbb{Z}[X]$-submodule of its annihilator, so $U \leq N_V(\text{ann}(U))$. We term $N_V(\text{ann}(U))$ the minimal null $\mathbb{Z}[X]$-submodule containing U. The following theorem shows that a null $\mathbb{Z}[X]$-submodule can be decomposed as a direct sum of null $\mathbb{Z}[X]$-submodules of P, Q and R.

Theorem. *For any ideal I of $\mathbb{Z}[X]$,*

$$N_V(I) = N_P(I) \oplus N_Q(I) \oplus N_R(I),$$

with $N_P(I) \leq P$, $N_Q(I) \leq Q$ and $N_R(I) \leq R$.

Proof. Let $n \in N_P(I) + N_Q(I) + N_R(I)$, then n can be written as $n = p + q + r$, where $p \cdot h(X) = q \cdot h(X) = r \cdot h(X) = 0$ for all $h(X) \in I$. Thus $n \cdot h(X) = p \cdot h(X) + q \cdot h(X) + r \cdot h(X) = 0$ for all $h(X) \in I$, so $n \in N_V(I)$. Therefore $N_P(I) + N_Q(I) + N_R(I) \leq N_V(I)$.

Conversely, suppose $n \in N_V(I)$. n can be written as $n = p + q + r$ for some $p \in P$, $q \in Q$ and $r \in R$. For any $h(X) \in I$,

$$0 = n \cdot h(X) = (p + q + r) \cdot h(X) = p \cdot h(X) + q \cdot h(X) + r \cdot h(X) = p' + q' + r'$$

for some $p' \in P$, $q' \in Q$ and $r' \in R$, as P, Q and R are $\mathbb{Z}[X]$-submodules. However, $P \oplus Q \oplus R$ is a direct sum, hence

$$0 = p' = q' = r' = p \cdot h(X) = q \cdot h(X) = r \cdot h(X),$$

and so $p \in N_P(I)$, $q \in N_Q(I)$ and $r \in N_R(I)$. As $n = p + q + r$, we have that $n \in N_P(I) + N_Q(I) + N_R(I)$. Therefore $N_V(I) \leq N_P(I) + N_Q(I) + N_R(I)$.

Therefore we have equality, and we clearly have a direct sum, so

$$N_V(I) = N_P(I) \oplus N_Q(I) \oplus N_R(I). \qquad \square$$

This theorem enables us to find $\mathbb{Z}[X]$-submodules of V because any $\mathbb{Z}[X]$-submodule is contained in its minimal null $\mathbb{Z}[X]$-submodule. This null $\mathbb{Z}[X]$-submodule can be

decomposed into null $\mathbb{Z}[X]$-submodules of P, Q and R. We thus consider the $\mathbb{Z}[X]$-submodules of V, P, Q and R. Any $\mathbb{Z}[X]$-submodule U can be regarded in the natural way as a $(\mathbb{Z}[X]/\text{ann}(U))$-module. We therefore calculate the following annihilators:

$$A_V = \text{ann}(V), \qquad A_P = \text{ann}(P), \qquad A_Q = \text{ann}(Q), \qquad A_R = \text{ann}(R).$$

We know that

$$\langle 2^8, p(X)\rangle = \langle 2^8, 1 - 18X + X^2\rangle \subset A_P,$$

so any element of A_P can be reduced modulo $\langle 2^8, p(X)\rangle$ to a polynomial of the form $a_0 + a_1X$, where $a_0, a_1 \in \{0, \ldots, 2^8 - 1\}$. By considering the effect of module homomorphisms $a_0 + a_1\alpha$ on the generators of P (as a $\mathbb{Z}$-module), we find the only other generator for A_P is $2^5(X + 3)$. Similar calculations give A_Q and A_R, so we have

$$\begin{aligned}
A_P &= \text{ann}(P) = \langle 2^8, 2^5(X + 3), p(X)\rangle \\
&= \langle 2^8, 2^5(X + 3), (1 - 18X + X^2)\rangle \\
A_Q &= \text{ann}(Q) = \langle 2^8, q(X)\rangle = \langle 2^8, (1 - 3X + X^2)\rangle \\
A_R &= \text{ann}(R) = \langle 2^8, 2^5(X^3 + 3), r(X)\rangle \\
&= \langle 2^8, 2^5(X^3 + 3), (1 + 3X + 8X^2 + 3X^3 + X^4)\rangle.
\end{aligned}$$

By considering the effect of elements of A_Q, A_R on P, etc., we find

$$\begin{aligned}
A_V &= \text{ann}(V) = \langle 2^8, 2^5(X + 3), 2q(X)r(X), (X + 1)q(X)r(X)\rangle \\
&= \langle 2^8, 2^5(X + 3), 2(1 - 3X + X^2)(1 + 3X + 8X^2 + 3X^3 + X^4), \\
&\quad (X + 1)(1 - 3X + X^2)(1 + 3X + 8X^2 + 3X^3 + X^4)\rangle
\end{aligned}$$

If we define the following quotient rings,

$$Z_V = \mathbb{Z}[X]/A_V, \qquad Z_P = \mathbb{Z}[X]/A_P, \qquad Z_Q = \mathbb{Z}[X]/A_Q, \qquad Z_R = \mathbb{Z}[X]/A_R,$$

then the elements in these rings can be written as

$$\begin{aligned}
Z_P &= \{a_0 + a_1X | a_0 = 0, \ldots, 2^8 - 1; a_1 = 0, \ldots, 2^5 - 1\}, \\
Z_Q &= \{a_0 + a_1X | a_0, a_1 = 0, \ldots, 2^8 - 1\}, \\
Z_R &= \{a_0 + a_1X + a_2X^2 + a_3X^3 | a_0, a_1, a_2 = 0, \ldots, 2^8 - 1; \\
&\qquad a_3 = 0, \ldots, 2^5 - 1\}, \\
Z_V &= \{a_0 + \cdots + a_6X^6 | a_0, a_1, a_2 = 0, \ldots, 2^8 - 1; \\
&\qquad a_3, a_4, a_5 = 0, \ldots, 2^5 - 1; a_6 = 0, 1\}.
\end{aligned}$$

The rings Z_P, Z_Q, Z_R and Z_V have sizes 2^{13}, 2^{16}, 2^{29} and 2^{40}, respectively. Note that all four of these rings can be regarded as cyclic modules over themselves.

We can thus regard V as a Z_V-module, and P, Q and R as Z_P-, Z_Q- and Z_R-modules, respectively. By considering the effect of α on one of the generators (as $\mathbb{Z}$-modules) p, q and r of P, Q and R, respectively, we find that P, Q and R contain cyclic submodules (generated by one element) that are isomorphic to Z_P, Z_Q and Z_R (as modules over

themselves), respectively. We denote these cyclic submodules $\langle p\rangle_{Z_P}$, $\langle q\rangle_{Z_Q}$ and $\langle r\rangle_{Z_R}$, respectively. These cyclic submodules can also be regarded as $\mathbb{Z}$-submodules, and as such we have the following $\mathbb{Z}$-module isomorphisms:

$$\begin{aligned}
\langle p\rangle_{Z_P} &\cong \langle p\rangle + 2^3 P && (\mathbb{Z}\text{-module isomorphism}),\\
\langle q\rangle_{Z_Q} &\cong Q && (\mathbb{Z}\text{-module isomorphism}),\\
\langle r\rangle_{Z_R} &\cong \langle r\rangle + 2^3 R && (\mathbb{Z}\text{-module isomorphism}).
\end{aligned}$$

Thus Q is a cyclic Z_Q-module of size 2^{16}. The cyclic Z_P- and Z_R-submodules of P and R, $\langle p\rangle_{Z_P}$ and $\langle r\rangle_{Z_R}$, are of size 2^{13} and 2^{29}, respectively. These cyclic submodules intersect non-trivially with any other submodule of P and R, respectively. P and R contain many such cyclic submodules of these sizes. They are each generated by a single generator of P and R (considered as $\mathbb{Z}$-modules). The submodules of the cyclic modules Z_P, Z_Q and Z_R (over themselves) are given by the ideals of Z_P, Z_Q and Z_R, respectively, or equivalently by ideals in $\mathbb{Z}[X]$ containing A_P, A_Q and A_R, respectively [5, Theorem 2.12]. Further analysis of the ideals of the rings can be conducted by using the theory of Gröbner bases [3].

We have given a thorough analysis of the α-invariant $\mathbb{Z}$-submodules of V by considering the equivalent $\mathbb{Z}[X]$-submodules of V. We summarize here the results of the $\mathbb{Z}[X]$-module analysis in terms of α-invariant $\mathbb{Z}$-submodules. We have shown that V can be decomposed as $V = P \oplus Q \oplus R$, where P, Q and R are α-invariant $\mathbb{Z}$-submodules, and that any other α-invariant $\mathbb{Z}$-module is contained in a minimal "null" $\mathbb{Z}$-submodule which decomposes as a sum of α-invariant $\mathbb{Z}$-submodules of P, Q and R. The $\mathbb{Z}$-submodules P and R contain certain α-invariant $\mathbb{Z}$-submodules that intersect non-trivially with any other α-invariant $\mathbb{Z}$-submodule of P and R, respectfully. Further α-invariant $\mathbb{Z}$-submodules of these $\mathbb{Z}$-submodules of P and R, and of the $\mathbb{Z}$-submodule Q, can be calculated by considering the ideals of the various polynomial quotient rings given above.

5. A Markov Chain on Quotient Modules

Let U be an α-invariant $\mathbb{Z}$-submodule of V. We consider the effect of the ith round function on the cosets of U in V, or equivalently on the quotient $\mathbb{Z}$-module V/U. Suppose now that an element $x^i \in V$ is the round input, $y^i \in V$ is the input to the PHT layer and $z^i \in V$ is the round output, and x^i_U, y^i_U and z^i_U are the corresponding cosets of U or elements of V/U. For a given round subkey (K_{2i-1}, K_{2i}), we can calculate the transition probability $P_{(K_{2i-1},K_{2i})}(y^i_U|x^i_U)$ between a coset of U before and a coset of U after the combined mixed XOR/addition, nonlinear and mixed addition/XOR layers. The effect of the PHT layer (or α) is to permute the cosets of U as U is α-invariant. For this round subkey (K_{2i-1}, K_{2i}), we can calculate the transition probability $P_{(K_{2i-1},K_{2i})}(z^i_U|x^i_U)$ between the a coset of U before and after the round function of SAFER. For this round subkey (K_{2i-1}, K_{2i}), the round function of SAFER gives a key-dependent probability transition matrix on the cosets of U. This transition matrix $Q_{(K_{2i-1},K_{2i})}$ is defined by

$$Q_{(K_{2i-1},K_{2i})} = (P_k(z^i_U|x^i_U)).$$

Consider now the SAFER encryption function with r rounds followed by a final output transformation which we regard as the $(r+1)$th round. Let $x^i \in V$ be the input to the

ith round ($i = 1, \ldots, (r+1)$), with x^{r+2} as the output. Let x_U^i ($i = 1, \ldots, (r+2)$) be the corresponding cosets of U or elements of V/U. For a given key K, the sequence $x_U^1, \ldots, x_U^{r+2}$ forms a key-dependent random process with state space the cosets of U. The transition matrix for this random process Q_K is defined by

$$Q_K = (P_K(x_U^{r+2}|x_U^1)).$$

The standard cryptographic assumption, implicitly used in both linear [9] and differential cryptanalysis [2] is that, for a given key, such a random process defined on a state space of cosets forms a first-order Markov chain [4]. In linear cryptanalysis these are usually cosets of a hyperplane of the message space (considered as a binary vector space). In differential cryptanalysis these are usually cosets of

$$\{(m, m)|m \in M\} \leq M \times M \qquad \text{(where } M \text{ is the message space)},$$

and these give the "characteristics". This assumption can be tested empirically. Under this assumption we can write the transition matrix Q_K as a product of transition matrices for each round. Thus if key K gives round subkeys $K_1, \ldots, K_{2r+1}$, the transition matrix Q_K is given by

$$Q_K = Q_{(K_1,K_2)} Q_{(K_3,K_4)} \cdots Q_{(K_{2r-1},K_{2r})} Q_{K_{2r+1}}.$$

6. A Potential Cryptographic Weakness

We have seen that the $\mathbb{Z}$-module V can be written as the direct sum of α-invariant $\mathbb{Z}$-submodules as $V = P \oplus Q \oplus R$. We can define a submodule S by

$$S = P \oplus Q = \langle e_1, d_1, d_2, e_8 \rangle.$$

S is an α-invariant submodule with $V = R \oplus S$. We consider how the Markov chain described in Section 5 applies to cosets of the submodule S. We can define φ to be the natural $\mathbb{Z}$-module homomorphism

$$\varphi\colon V \to \frac{V}{S} \cong R.$$

We can regard the random process on the cosets of S as a random process on elements of R. For an element $v \in V$, we write v_S for this coset of S or equivalently element of R. The value of v_S does not depend on the first and eighth bytes of v as e_1 and e_8 are absent from the basis for R. Consider the ith round of a SAFER encryption under a given ith round subkey (K_{2i-1}, K_{2i}). Suppose now that an element $x^i \in V$ is the round input, $y^i \in V$ is the input to the PHT layer, and $z^i \in V$ is the round output, and x_S^i, y_S^i and z_S^i are the corresponding cosets of S or elements of R. For a given input x^i, the central six bytes of the output y^i of the combined mixed XOR/addition, nonlinear and mixed addition/XOR layers do not depend on the first and eighth bytes of the subkeys K_{2i-1} and K_{2i}. Thus the distribution of y_S^i conditional on x_S^i is constant whenever the central key bytes agree. z_S^i depends only on y_S^i as S is α-invariant. For the given subkey, the distribution of z_S^i conditional on x_S^i therefore does not depend on the first and eighth

bytes of the subkeys K_{2i-1} and K_{2i}. The key-dependent one-round transition matrix on the cosets of S is identical for all subkeys (K_{2i-1}, K_{2i}) that agree on the central six bytes. If we define θ to be the restriction to these central six bytes, then the one-round transition matrix can be written as

$$Q_{(K_{2i-1},K_{2i})\theta} = (P_{(K_{2i-1},K_{2i})\theta}(z_S^i|x_S^i)).$$

Suppose now that we have an r-round plus final output transformation SAFER encryption with message $m = x^1$ and ciphertext $c = x^{r+2}$. For a given key K, under the standard cryptographic assumption that such a process forms a first-order Markov process, the transition matrix between message cosets and ciphertext cosets of S is the product of the round transition matrices. Thus

$$Q_K = (P_K(z_U^i|x_U^i)) = Q_{(K_1,K_2)\theta} \cdots Q_{(K_{2r-1},K_{2r})\theta} Q_{(K_{2r+1})\theta}.$$

The transition matrix Q_K therefore depends only on the central six bytes of the subkeys $K_1, \ldots, K_{2r+1}$. For SAFER K-64/128, the key scheduling restricts bytes of K to the same bytes of any subkey, so we have $Q_K = Q_{K\theta}$. For SAFER K-64, this means that the distribution of c_S conditional on m_S does not depend on the first and eighth bytes of the key, whereas for SAFER K-64 the distribution of c_S conditional on m_S does not depend on the first, eighth, ninth and sixteenth bytes of the key.

For either SAFER K-64 or SAFER K-128, we have found a half-rank algebraic structure (R) of the message/ciphertext space on which the output distributions do not depend on a quarter of the key bytes. In [7] it is stated that "SAFER was designed in accordance with Shannon's principles of confusion and diffusion for obtaining security in secret-key ciphers". Shannon's principle of confusion [10] is "to make the relation between simple statistics of ciphertext and simple statistics of the key a very complex and involved one". As there are simple statistics of the output (that is, its "projection" on R) that do not depend on a quarter of the key bytes, it arguable that SAFER does not satisfy the principle of confusion. Therefore if there are any collections of functions of m_S which have non-negligible correlations with any collections of functions of c_S, then we have a reduced key search.

For practical reasons, we may need to concentrate on the least significant n bits of each byte, that is, the module V_n. The above reasoning also applies to this module, so just by considering the least significant n bits we may still get a reduced key search. We note that when we consider α as a module homomorphism of V_n for small n, it has even more structure. For example, as a module homomorphism of V_3, every submodule of P is α-invariant and R has a submodule of rank 2 in which every submodule is α-invariant.

There are many similar attacks on SAFER using the α-invariant $\mathbb{Z}$-submodule S with different time/space complexity trade-offs. We do not investigate all these attacks to find the best one. The key schedule has already been revised in the light of this paper and [6]. The attack below is given solely as an illustration of the type of attack it may be possible to mount on SAFER using the α-invariant $\mathbb{Z}$-submodule S.

We can attempt to exploit the lack of dependence on certain key bytes by calculating empirical transition probabilities on cosets of S for plaintext/ciphertext data. Given enough data we can see which key-dependent sets of transition probabilities best agree with the empirical probabilities and thus find key information.

In order to calculate these transition matrices, we need to calculate transition matrices for one round. We now explain how the components of the round function affect the transition matrix. We first note that we can write R as the direct sum of two $\mathbb{Z}$-submodules of rank 2, one involving linear combinations of bytes 2, 3, 5 and the other linear combinations of bytes 4, 6, 7. Thus we have

$$R = \langle e_2 - e_5, e_3 - e_5 \rangle \oplus \langle e_6 - e_4, e_7 - e_4 \rangle.$$

In the mixed XOR/addition, nonlinear and mixed addition/XOR phases, SAFER acts independently on these two $\mathbb{Z}$-submodules of R. We call the set of cosets on either of these two submodules half-cosets.

Adding a subkey byte corresponds to permuting the cosets of S in V according to the subkey byte. XORing a subkey byte corresponds to adding one of a small set of other bytes according to a known distribution that depends on the subkey byte. Combining these two operations into the mixed XOR/addition or the mixed addition/XOR phase, we see that either of these phases gives a transition matrix on the cosets that depends on the subkey. The $45^{(\cdot)}$, $\log_{45}(\cdot)$ transforms give transition matrices on each of the half-cosets. We give some examples below. Thus the combined effect of the mixed XOR/addition, nonlinear and mixed addition/XOR layers is to give key-dependent transition matrices on the half-cosets. These transition matrices are effectively key-dependent weighted averages of the permuted transition matrix on the half-cosets for the nonlinear layer. To obtain the key-dependent transition matrix for all of the cosets, we combine the key-dependent transition matrices for both of the component half-cosets using an element-by-element product.

The effect of the PHT layer is to permute the cosets as S is an α-invariant subspace. This permutation is given by M_R, where

$$M_R = \begin{pmatrix} -2 & 2 & -1 & 1 \\ -2 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

was given earlier as the block (corresponding to R) of the matrix M as a block matrix. Thus we can find the key-dependent transition matrix for one round by permuting the columns of the key-dependent matrix for the combined mixed XOR/addition, nonlinear and mixed addition/XOR layers. The key-dependent transition matrix for a number of rounds is just the product of the individual key-dependent one-round transition matrices.

Let $P_{K'}(m_S, c_S)$ denote the probability of transition from message class m_S to ciphertext class c_S under key class K', where $K' = K\theta$ denotes the relevant 6 or 12 key bytes (SAFER K-64 or K-128). $P_{K'}(m_S, c_S)$ is just the relevant entry of the transition matrix corresponding to K'. Suppose we have a number of message/ciphertext pairs and let $N(m_S, c_S)$ be the number of times message class m_S and c_S occur. For a known plaintext attack we can find the value of K' that maximizes the log-likelihood function

$$\sum_{(m_S, c_S)} N(m_S, c_S) \log P_{K'}(m_S, c_S).$$

This is a maximization over at most 2^{48} or 2^{96} elements. For 2-round SAFER, an approximate probability argument using random functions shows that, for a given key class K',

about 37% (e^{-1}) of class pairs (m_S, c_S) do not occur, that is, $P_{K'}(m_S, c_S) = 0$ for 37% of class pairs (m_S, c_S). Thus for 2-round SAFER, we have a trivial attack in which we can identify the true key class with a handful of message/ciphertext pairs.

The above analysis for SAFER with a realistic number of rounds requires the calculation of vast numbers of $2^{32} \times 2^{32}$ transition matrices. In order to make the calculations more tractable, we can consider the least significant n bits of every byte, that is, the module V_n. The general theory given above applies to this module and its relevant submodules. Looking at the least significant bit gives no information as the transition probabilities through the non-linear layer are uniform. Vaudenay [11] considered this situation in the case when the $45^{(\cdot)}$ and $\log_{45}(\cdot)$ functions are replaced by functions that give non-uniform transition probabilities. Thus we consider the least significant two bits of each byte. In this case XORing by 00 or 10 is equivalent to adding 00 or 10, respectively, and XORing by 01 or 11 is equivalent to adding 01 with probability one-half or 11 with probability one-half. The transition matrix for a set of half-cosets of S is $2^{-4}J_2 + 2^{-20}T_2$, where J_2 is the 16×16 matrix with every entry 1 and T_2 is given in Appendix 1. The matrix T_2 is calculated by considering all 2^{24} values of the three bytes involved in the half-coset (as are J_3 and J_4 below).

Whilst these transition probabilities are not uniform, they are not non-uniform enough to launch an attack on SAFER with a realistic number of rounds.

When we consider the least significant three bits of each byte, we obtain the transition matrix on the half-cosets of $2^{-6}J_3 + 2^{-18}T_3$, where J_3 is the 64×64 matrix with every entry 1. The first row of T_3 is given in Appendix 2. It can be seen that a typical entry of T_3 has absolute size about 2^8, so a typical entry of the transition matrix differs from the uniform value of 2^{-6} by about 2^{-10}. The least significant four bits in each byte give the transition matrix on the half-cosets of $2^{-8}J_4 + 2^{-16}T_4$ where J_4 is the 256×256 matrix with every entry 1. The first row of T_4 is given in Appendix 3. A typical entry of T_4 has size about 2^6, so a typical entry of the transition matrix differs from the uniform value of 2^{-8} by about 2^{-10}. We have shown parts of these two matrices in the Appendix to show that even when considering the least few significant bits of every byte, the transition probabilities of the "half-cosets" are highly non-uniform. Even after allowing for the averaging effect of XORing key bytes, the key-dependent transition matrices for the half-cosets across the mixed XOR/addition, nonlinear and mixed addition/XOR layers are still highly non-uniform. By taking the elementwise product of transition matrices we obtain key-dependent transition matrices across the cosets. The PHT layer permutes the columns of this matrix, so we can obtain one-round key-dependent round transition matrices on the cosets which are highly non-uniform. By taking an appropriate product of such matrices, we can calculate key-dependent transition matrices from the message cosets to the ciphertext cosets.

Given sufficient computational power, we can pre-compute all such key-dependent transition matrices. For a set of message/ciphertext pairs, we could use these matrices to calculate the likelihoods as given above. In practice, the keys may naturally occur in classes that give approximately equal transition matrices, which would reduce the key search. In any case, the key search for the attack described above is at worst 48 or 96 bits (SAFER K-64 or K-128).

7. Other Potential Algebraic Weaknesses

The invariant $\mathbb{Z}$-submodules of Section 4 essentially give the cryptanalyst a method for controlling the diffusion of SAFER in the PHT layer. There are several ways in which this may be potentially exploited. We briefly list some of them.

1. In the $\mathbb{Z}$-submodule V_4 or V_8, there are many linear combinations of bytes that are fixed by α. By analysing how the key and nonlinear layers affect these linear combinations, either individually or jointly, it may be possible to find information about the key.
2. α has many small cycles. In particular on V_4, α has order 3, and on V_8, α^3 fixes every $\mathbb{Z}$-submodule of rank 1. We can analyse the $\mathbb{Z}$-submodules generated by such small cycles. Those $\mathbb{Z}$-submodules generated by cycles that involve elements of low (module) weight may provide key information.
3. Constructing affine ($\mathbb{Z}$-module) approximations to the $45^{(\cdot)}$, $\log_{45}(\cdot)$ and XOR functions and relating these to the invariant $\mathbb{Z}$-submodules may give key information.
4. Further investigation of α-invariant $\mathbb{Z}$-submodules of V. In particular it may be possible to relax the requirement of strict invariance and analyse probabilistically invariant $\mathbb{Z}$-submodules, for example, the central six bytes $\langle e_2, e_3, e_4, e_5, e_6, e_7\rangle$. There are many such $\mathbb{Z}$-submodules which could give key information.
5. Differential analysis based on the α-invariant $\mathbb{Z}$-submodules.
6. In SAFER K-128 we saw above that the key bytes divide into two types, the addition key bytes and the XOR key bytes. The effect of adding two key bytes sequentially is the same as adding one key byte equal to their sum. It is therefore possible that, because of the underlying $\mathbb{Z}$-module structure, the transition probabilities would depend only on the value of the overall addition of certain addition key bytes. This would give a greatly reduced key search.

8. Conclusions

In this paper we have given an analysis of the SAFER algorithm based on the algebraic properties of the PHT layer, and, in particular, the invariant $\mathbb{Z}$-submodules. In particular, for a given key, we have found a "projection" (φ) of the 8-byte message/ciphertext space onto a 4-byte $\mathbb{Z}$-submodule so that the probability of any message projection giving any ciphertext projection is independent of one-quarter of the key bytes. This gives the real possibility of reducing a key search to 6 or 12 bytes (SAFER K-64 or K-128). We have given an example of one way in which this may be exploitable given sufficient computational resources.

The key scheduling for SAFER K-64 and SAFER K-128 has been changed to give SAFER SK-64 and SK-128 since the original submission of this paper and [6]. This (amongst other things) ensures that this projection depends on all the key bytes. However, the main contribution of this paper is the use of the invariant $\mathbb{Z}$-submodules of the PHT layer to allow the cryptanalyst to control diffusion, and the algebraic analysis of these invariant $\mathbb{Z}$-submodules. Even with the new key schedule, there remains the possibility of using the invariant $\mathbb{Z}$-submodules of the PHT layer to analyse SAFER.

Acknowledgements

I wish to thank Fred Piper and Simon Blackburn for some interesting discussions and the referees for their helpful comments.

Appendix 1. Matrix T_2

$$
\begin{array}{cccccccc}
2048 & -512 & 512 & -1024 & -512 & 2048 & -1024 & 512 \\
512 & -1024 & 1024 & -1536 & -1024 & 512 & -1536 & 1024 \\
 & & & & & & & \\
-128 & 128 & 512 & 512 & 128 & -128 & 512 & 512 \\
-512 & -512 & 128 & -128 & -512 & -512 & -128 & 128 \\
 & & & & & & & \\
-1536 & 1024 & 0 & 1536 & 1024 & -1536 & 1536 & 0 \\
-1024 & 512 & -1536 & 1024 & 512 & -1024 & 1024 & -1536 \\
 & & & & & & & \\
640 & 384 & 0 & 0 & 384 & 640 & 0 & 0 \\
0 & 0 & -640 & -384 & 0 & 0 & -384 & -640 \\
 & & & & & & & \\
-128 & 128 & -512 & -512 & 128 & -128 & -512 & -512 \\
512 & 512 & 128 & -128 & 512 & 512 & -128 & 128 \\
 & & & & & & & \\
1024 & -1536 & 512 & -1024 & -1536 & 1024 & -1024 & 512 \\
512 & -1024 & 2048 & -512 & -1024 & 512 & -512 & 2048 \\
 & & & & & & & \\
-384 & -640 & 0 & 0 & -640 & -384 & 0 & 0 \\
0 & 0 & 384 & 640 & 0 & 0 & 640 & 384 \\
 & & & & & & & \\
-1536 & 1024 & -1024 & 512 & 1024 & -1536 & 512 & -1024 \\
0 & 1536 & -1536 & 1024 & 1536 & 0 & 1024 & -1536 \\
 & & & & & & & \\
-1536 & 1024 & -1024 & 512 & 1024 & -1536 & 512 & -1024 \\
0 & 1536 & -1536 & 1024 & 1536 & 0 & 1024 & -1536 \\
 & & & & & & & \\
-384 & -640 & 0 & 0 & -640 & -384 & 0 & 0 \\
0 & 0 & 384 & 640 & 0 & 0 & 640 & 384 \\
 & & & & & & & \\
1024 & -1536 & 512 & -1024 & -1536 & 1024 & -1024 & 512 \\
512 & -1024 & 2048 & -512 & -1024 & 512 & -512 & 2048 \\
 & & & & & & & \\
-128 & 128 & -512 & -512 & 128 & -128 & -512 & -512 \\
512 & 512 & 128 & -128 & 512 & 512 & -128 & 128 \\
 & & & & & & & \\
640 & 384 & 0 & 0 & 384 & 640 & 0 & 0 \\
0 & 0 & -640 & -384 & 0 & 0 & -384 & -640 \\
 & & & & & & & \\
-1536 & 1024 & 0 & 1536 & 1024 & -1536 & 1536 & 0 \\
-1024 & 512 & -1536 & 1024 & 512 & -1024 & 1024 & -1536 \\
 & & & & & & & \\
-128 & 128 & 512 & 512 & 128 & -128 & 512 & 512 \\
-512 & -512 & 128 & -128 & -512 & -512 & -128 & 128 \\
 & & & & & & & \\
2048 & -512 & 512 & -1024 & -512 & 2048 & -1024 & 512 \\
512 & -1024 & 1024 & -1536 & -1024 & 512 & -1536 & 1024
\end{array}
$$

Appendix 2. First Row of Matrix T_3

$$
\begin{array}{cccccccc}
1328 & 124 & 140 & -100 & 96 & 164 & 20 & 276 \\
124 & 720 & -116 & -140 & -188 & -48 & -188 & 92 \\
140 & -116 & 608 & -340 & -172 & -100 & -224 & -52 \\
-100 & -140 & -340 & 416 & -404 & -196 & -436 & -80 \\
96 & -188 & -172 & -404 & 464 & -196 & -244 & -124 \\
164 & -48 & -100 & -196 & -196 & 656 & -268 & -12 \\
20 & -188 & -224 & -436 & -244 & -268 & 416 & -100 \\
276 & 92 & -52 & -80 & -124 & -12 & -100 & 1024
\end{array}
$$

Appendix 3. First Row of Matrix T_4

321	−13	18	7	10	−35	11	−21
−8	66	10	−33	46	48	6	31
−13	201	−53	−16	−45	−8	−43	−43
−25	−20	−31	−33	−49	−2	8	28
18	−53	181	−36	−33	−38	0	−78
−16	10	−44	−49	−32	−39	−23	−8
7	−16	−36	194	−72	−27	−19	−14
−49	−17	−36	−56	−7	−25	−49	30
10	−45	−33	−72	144	−56	−22	−65
−8	−43	−36	−52	−52	−2	−46	−38
−35	−8	−38	−27	−56	184	−41	−27
−30	8	−55	−25	−13	−32	−19	−10
11	−43	0	−19	−22	−41	220	−68
29	8	−3	−59	−12	−10	−24	33
−21	−43	−78	−14	−65	−27	−68	124
−35	−5	−24	−38	−56	−8	−38	−20
−8	−25	−16	−49	−8	−30	29	−35
237	7	1	−9	8	−17	−7	66
66	−20	10	−17	−43	8	8	−5
7	325	−17	27	39	50	−39	33
10	−31	−44	−36	−36	−55	−3	−24
1	−17	177	−46	10	2	2	−22
−33	−33	−49	−56	−52	−25	−59	−38
−9	27	−46	186	−6	20	−25	6
46	−49	−32	−7	−52	−13	−12	−56
8	39	10	−6	298	−28	23	23
48	−2	−39	−25	−2	−32	−10	−8
17	50	2	20	−28	234	−31	32
6	8	−23	−49	−46	−19	−24	−38
−7	−39	2	−25	23	−31	214	16
31	28	−8	30	−38	−10	33	−20
66	33	−22	6	23	32	16	344

References

[1] T.A. Berson and L.A. Knudsen. Truncated Differentials of SAFER. In *Fast Software Encryption, Third International Workshop, Cambridge*, 1996, pages 15–26. LNCS 1039, Springer-Verlag, Berlin, 1996.

[2] E. Biham and A. Shamir. *Differential Cryptanalysis of the Date Encryption Standard*. Springer-Verlag, New York, 1993.

[3] D. Cox, J. Little and D. O'Shea. *Ideals, Varieties and Algorithms*. Springer-Verlag, New York, 1992.

[4] G. Grimmett and D. Stirzacker. *Probability and Random Processes*. Clarendon Press, Oxford, 1982.

[5] B. Hartley and T.O. Hawkes. *Rings, Modules and Linear Algebra*. Chapman and Hall, London, 1970.

[6] L.A. Knudsen. A Key-Schedule Weakness in SAFER K-64. In *Advances in Cryptology — CRYPTO* '95, pages 274–286. LNCS 963, Springer-Verlag, Berlin, 1995.

[7] J.L. Massey. SAFER K-64: A Byte-Oriented Block-Ciphering Algorithm. In *Fast Software Encryption, Proceedings of Cambridge Security Workshop* 1993, pages 1–17. LNCS 809, Springer-Verlag, Berlin, 1994.

[8] J.L. Massey. SAFER K-64: One Year Later. In *Fast Software Encryption, Second International Workshop, Leuven* 1994, pages 212–241. LNCS 1008, Springer-Verlag, Berlin, 1995.

[9] M. Matsui. Linear Cryptanalysis Method for DES Cipher. In *Advances in Cryptology—EUROCRYPT* 93, pages 386–397. LNCS 765, Springer-Verlag, Berlin, 1994.

[10] C.E. Shannon. Communication Theory of Secrecy Systems. *Bell System Technical Journal*, 28:656–715, 1949.

[11] S. Vaudenay. On the Need for Multipermutations: Cryptanalysis of MD4 and SAFER. In *Fast Software Encryption, Second International Workshop, Leuven*, 1994, pages 286–297. LNCS 1008, Springer-Verlag, Berlin, 1995.