CARLETON UNIVERSITY

SCHOOL OF
MATHEMATICS AND STATISTICS

HONOURS PROJECT

TITLE: Semi-Involutions over Finite Fields

AUTHOR: Trevor Thompson

SUPERVISOR: Daniel Panario and Qiang Wang

DATE: May 3rd, 2019

Semi-Involutions over Finite Fields

Trevor Thompson

3 May, 2019

Contents

Abstract

We study semi-involutions over finite fields. After motivating and studying basic properties of semi-involutions, we describe several classes of polynomials that are semi-involutions.

Chapter 1

Introduction

A bijection f over a finite field $\mathbb{F}_q$ is a semi-involution if there exists scalars $a, b \in \mathbb{F}_q$ for which $f(f(x+a)+b)$ is the identity function on $\mathbb{F}_q$. Semi-involutions were originally defined in fields of even characteristic in [12] in the context of constructing DES-like ciphers. However, there has not been research concerning semi-involutions in fields of odd characteristic.

This project starts with some background information concerning finite fields and functions that map a finite field to itself, followed by a demonstration that some properties of semi-involutions described in [12] generalize to fields of odd characteristic, and a description of various classes of semi-involutions in fields of any characteristic.

In Chapter 2 we review some basic theory of finite fields. Then, in Chapter 3 we motivate the need of studying semi-involutions and review some basic properties of them. In Chapter 4 we describe several classes of polynomials including linearized functions, monomials and Möbius functions that are semi-involutions. Finally, in Chapter 5 we conclude the project and propose some questions to study in the future.

Chapter 2

Background Results

In this chapter, we state some basic results in finite fields that are required to understand the remainder of the project. These results can be found in chapters 3, 4, and 6 of W. Keith Nicholson's *Introduction to Abstract Algebra*, [9] and so are briefly glossed over here.

2.1 Some Basic Definitions

Definition 1. *A field is a commutative division ring with identity. A finite field is a field with a finite number of elements. The number of elements of a finite field is called the order of the field.*

The ring of integers, modulo n, is not a field when n is composite, for then n would have a non-trivial divisor d that would not have an inverse in $\mathbb{Z}_n$. If n is prime, however, then $\mathbb{Z}_n$ is a field, as every non-zero element of $\mathbb{Z}_n$ has a unique inverse.

The elements of any ring, and therefore finite field, form a group under addition. Also, the non-zero elements of any finite field form a group under multiplication.

Definition 2. *The order of 1 in the additive group $(F, +)$ is called the characteristic of the field F. If the order of 1 is infinite, then, by convention, we say that F has characteristic zero.*

For example, the characteristic of $\mathbb{Z}_p$ for p prime is p.

Proposition 1. *The characteristic of a finite field is a prime number.*

Proof. Recall from group theory that the order of an element in a finite group divides the number of elements in the group. Therefore, a finite field has a non-zero characteristic. If the characteristic was not prime, then it must be equal to mn for some integers $m, n > 1$. It follows that $(m \cdot n) \cdot 1 = 0$, which implies that $(m \cdot 1)(n \cdot 1) = 0$. Thus we have two non-zero quantities multiplying to yield zero, which is a contradiction. □

Definition 3. *Given a field F, if there exists $K \subseteq F$ such that K is a field under the operations of F, then we say that K is a subfield of F, and that F is an extension field of K.*

We often treat an extension of a field as a vector space over that field, as field extensions satisfy the definition of a vector space. It follows that, if a finite field F is an extension of a finite field K, then the order of F is a power of the order of K, and the exponent is the dimension of F. Since a field of characteristic p contains $\mathbb{Z}_p$ as a subfield (this follows from the definition of the characteristic), a finite field of characteristic p contains p^n elements for some positive integer n.

Proposition 2. *The order of a finite field is always a prime power.*

It is possible to have two finite fields that are essentially the same, but with different names assigned to their elements. This is formally defined below.

Definition 4. *If E and F are two fields, a function $\sigma : E \to F$ is called an isomorphism of fields if σ is a bijection that preserves addition and multiplication.*

2.2 Polynomial Extensions of Finite Fields

Given a finite field F, we consider the set of polynomials in x with co-efficients in F. We denote this set $F[x]$.

Definition 5. *The highest power of x with a non-zero term in a polynomial is called the degree of that polynomial (the degree of the zero polynomial, for which no such term exists, is defined to be $-\infty$). If p and q are elements of $F[x]$, the degree of p is m, the degree of q is n, and*

$$p(x) = \sum_{i=0}^{m} a_i x^i,\ q(x) = \sum_{i=0}^{n} b_i x^i$$

then the sum and product of p and q in $F[x]$ are defined as follows:

$$(p+q)(x) = \sum_{i=0}^{\max\{m,n\}} (a_i + b_i)x^i$$

$$(p \cdot q)(x) = \sum_{i=0}^{m+n} \sum_{j=0}^{m} a_j b_{i-j} x^i$$

where $a_i = 0$ whenever $i > m$ and $b_i = 0$ whenever $i > n$.

It can be shown that, under this definition, $F[x]$ is a commutative integral domain with identity (see [9, Th. 4.1.2]). However, $F[x]$ is not a division ring, and therefore not a field. The only units of $F[x]$ are the non-zero constants.

Definition 6. *We say that a polynomial $f \in F[x]$ is irreducible in F if f is not a constant, and $f = g \cdot h$ for $g, h \in F[x]$ implies that either g or h is a* *constant.*

Similar to how any positive integer can be written uniquely as a product of positive primes, a monic polynomial can be written uniquely as a product of irreducible monic polynomials, which is proved in [9, Th.4.2.12]. Since the ring of integers modulo p for p prime was a field, this suggests that we can construct a field from the residue classes of $F[x]$ modulo f for f irreducible. Since f is irreducible, there are no zero divisors in $F[x]/(f)$, and so every non-zero polynomial in this ring has an inverse. Therefore, $F[x]/(f)$ is a field. The elements of $F[x]/(f)$ are all of the polynomials with coefficients in F and degree less than the degree of f. Therefore, the order of $F[x]/(f)$ is the order of F raised to the power of the degree of f.

2.3 Splitting Fields

Recall from Proposition 2 that the order of a finite field is a prime power. From this, one may ask if every prime power is the order of a finite field. In order to answer this, we must introduce the concept of a splitting field.

Definition 7. *If K is a subfield of F, and $u_1, u_2, \ldots, u_m$ are elements of F, then the field $K(u_1, u_2, \ldots, u_m)$ is the intersection of all subfields of F containing K and the elements $u_1, u_2, \ldots, u_m$.*

Definition 8. *Let f be a monic non-constant polynomial in $K[x]$. We call an extension field F of K a splitting field of f over K if $f(x) = (x-u_1)(x-u_2)\ldots(x-u_m)$ for some $u_1, u_2, \ldots, u_m \in F$, and $F = K(u_1, u_2, \ldots, u_m)$.*

For example, $\mathbb{Z}_2[\alpha]/(\alpha^2+1)$ is a splitting field of x^4+x over $\mathbb{Z}_2$, since x^4+x can be factored as $x(x-1)(x-\alpha)(x-(\alpha+1))$ and there is no subfield of $\mathbb{Z}_2[\alpha]/(\alpha^2+1)$ over which x^4+x can be written as a product of linear factors.

Proposition 3. *Let f be a monic non-constant polynomial in $K[x]$. Then there exists a splitting field of f over K, and this field is unique up to isomorphism.*

Existance of splitting fields is proved in [9, Th. 6.3.2]; uniqueness of splitting fields is proved in [9, Th. 6.3.4]. Both results are proved by induction on the degree of f.

From this one can prove the existance and uniqueness of fields of arbitrary prime power order.

Proposition 4. *Let $q = p^n$, where p is prime and n is a positive integer. Then the splitting field of $x^q - x$ over $\mathbb{Z}_p$ has q elements, and every field with q elements is isomorphic to this splitting field.*

The proof requires derivatives of polynomials in polynomial rings, as well as the Frobenius automorphism, and is not described here. It is given in [9, Th. 6.4.4].

If q is a prime power, we denote the unique finite field of q elements by $\mathbb{F}_q$. The multiplicative group of non-zero elements of $\mathbb{F}_q$ is denoted by $\mathbb{F}_q^*$. Galois showed that this multiplicative group is always cyclic [9, Th. 6.4.7].

2.4 Functions on Finite Fields

Over the course of the project we use several classes of functions from $\mathbb{F}_q$ to itself. In this section we define these classes and state a few properties.

Proposition 5. *Any function from a finite field to itself can be expressed as a polynomial.*

Proof. A function defined on a finite field is defined on a finite number of values. Therefore, one can use Lagrangian interpolation to determine a polynomial whose value equals that of the function for every element of its domain. It follows that this polynomial is equal to the function. □

Definition 9. *A bijection on $\mathbb{F}_q$ is called a permutation, or a permutation polynomial. If f is a permutation on $\mathbb{F}_q$ and f^2 (that is, f composed with itself) is equal to the identity function, we say that f is an involution.*

Definition 10. *We say that a function f over a finite field $\mathbb{F}_{q^n}$ is linearized if, for every $x, y \in \mathbb{F}_{q^n}$ and $c \in \mathbb{F}_q$, we have*

$$f(x+y) = x + y \text{ and } f(cx) = cf(x).$$

It is a well-known result that a linearized function over $\mathbb{F}_{q^n}$ contains only terms whose exponents are powers of q. In addition, a linearized function never contains a constant term.

Definition 11. *For $u, v, w, z \in \mathbb{F}_q$, $w \neq 0$, a function of the form*

$$f(y) = \frac{uy+v}{wy+z}$$

over the set $\mathbb{F}_q \cup \infty$, where $f(-zw^{-1}) = \infty$ and $f(\infty) = uw^{-1}$, is a Möbius function.

It is well known that f is a permutation of $\mathbb{F}_q \cup \infty$ if and only if $uz - vw \neq 0$.

Definition 12. *Given a constant $a \in \mathbb{F}_q$, the nth Dickson polynomial of the first kind is given by the following:*

$$D_n(x,a) = \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{n}{n-j} \binom{n-j}{j} x^{n-2j} a^j.$$

These functions are called Dickson polynomials because many of their properties were first discovered by Dickson in 1896 [6]:

Proposition 6. *i. $D_n(x,a)$ is equal to $D_{n+q^2-1}(x,a)$, that is, the Dickson polynomials repeat with periodicity q^2-1.*

ii. A Dickson polynomial is a permutation if and only if n and $q^2 - 1$ are relatively prime.

iii. For any integers m and n, $D_m(x, a)$ composed with $D_n(x, a)$ equals $D_{mn}(x, a)$. In particular, this means that if $D_n(x, a)$ is a bijection, its inverse is also a Dickson polynomial, specifically $D_m(x, a)$ where $mn \equiv 1(\mod q^2 - 1)$.

The following definition of Rédei polynomials over fields of odd characteristic is due to Qureshi and Panario, [10] based on a formula of Carlitz: [2]

Definition 13. *Given a constant $a \in \mathbb{F}_q$, where q is odd, the nth Rédei polynomial on $\mathbb{F}_q \cup \infty$ is given by the following:*

$$R_n(x, a) = \sqrt{a}\frac{\gamma(x)^n + 1}{\gamma(x)^n - 1}$$

where

$$\gamma(x) = \frac{x + \sqrt{a}}{x - \sqrt{a}}$$

and $\gamma(\sqrt{a}) = \infty, \gamma(\infty) = 1, R_n(x, a) = \infty$ whenever $\gamma(x)^n = 1$, and $R_n(\infty, a) = \infty$.

Qureshi and Panario also give several useful properties of Rédei polynomials:

Proposition 7. *i. $R_n(x, a)$ is equal to $R_{n+q-1}(x, a)$, that is, the Rédei polynomials repeat with periodicity $q + 1$.*

ii. A Rédei polynomial is a permutation if and only if n and $q - \chi(a)$ are relatively prime, where $\chi(a) = 1$ if a has a square root in $\mathbb{F}_q$ and $\chi(a) = -1$ otherwise.

iii. For any integers m and n, $R_m(x, a)$ composed with $R_n(x, a)$ equals $R_{mn}(x, a)$. In particular, this means that if $R_n(x, a)$ is a bijection, its inverse is also a Rédei polynomial, specifically $R_m(x, a)$ where $mn \equiv 1(\mod q^2 - 1)$.

Chapter 3

Semi-Involutions

3.1 Motivation

The Data Encryption Standard (DES) was introduced in 1977 [8]. At the time, it was believed to be so secure that it could only be broken by trying every possible key, which would take too much time to be practical. However, in 1991 Biham and Shamir [1] used a technique called differential cryptanaylsis to efficiently break DES. While differential cryptanalysis was still too long to be practical on contemporary computers, as it required 2^{61} operations, this took far less time than simply running the decryption algorithm using every possible set of sixteen 48-bit keys, of which there are 2^{768}, until one obtained a message that made sense. (In practice, the sixteen round keys are derived from a 56-bit "master key" [8], so only a very tiny subset of the possible keys are used. It has been suspected that the NSA did not want the cipher to be so secure that they would not be able to read intercepted foreign intelligence that was encrypted in DES [5, p. 13].) It only seemed to be a matter of time until commercially-available computers could break DES in a reasonable amount of time!

This led to interest into creating ciphers that were similar to, yet more secure than, DES. The *substitution-permutation network* (SPN) was proposed as a generalization of the concepts of DES. The central mechanism of an SPN is the S-box. An S-box is a permutation of the values of a vector space over $\mathbb{F}_q$. As there are very few computers today that operate in anything other than base two, typically $q = 2$. An SPN cipher encrypts a word in $\mathbb{F}_q^n$ over several rounds, each of which consists of the following [12]:

1. The key for the round is added (by componentwise vector addition in $\mathbb{F}_q^n$) to the message. Similar to DES, an SPN is designed so that even if one knew the precise workings of the cryptosystem, it is nigh- impossible to retrieve the original message without knowledge of the keys used to encrypt it.

2. The message is split into sub-words of equal length, which are then passed through a set of S-boxes. Every sub-word of every round can have a different S-box. Despite the S-boxes being permutation functions, this is typically referred to as the "substitution phase" of the substitution-permutation network.

3. The scalar components of the message are permuted. This is the "permutation phase" of the substitution-permutation network.

After the desired number of rounds, one final key is added to the message and the result is the encrypted message. If the S-boxes and permutations are chosen carefully, such a cipher is resistant to differential cryptanalysis [7].

The problem with substitution-permutation networks are that, in general, they are difficult to reverse. Thus, two SPNs, one for encryption and one for decryption, must be stored and implemented separately [12]. By contrast, since a round of DES only affects half of the message at a time, and DES worked in characteristic two (that is, addition is self-inverse in $\mathbb{F}_2^{64}$), one could decrypt a DES-encrypted message by encrypting the encrypted message using the sixteen keys used to encrypt it in reverse order [8]. Thus, only one algorithm needs to be implemented and only one set of S-boxes needs to be

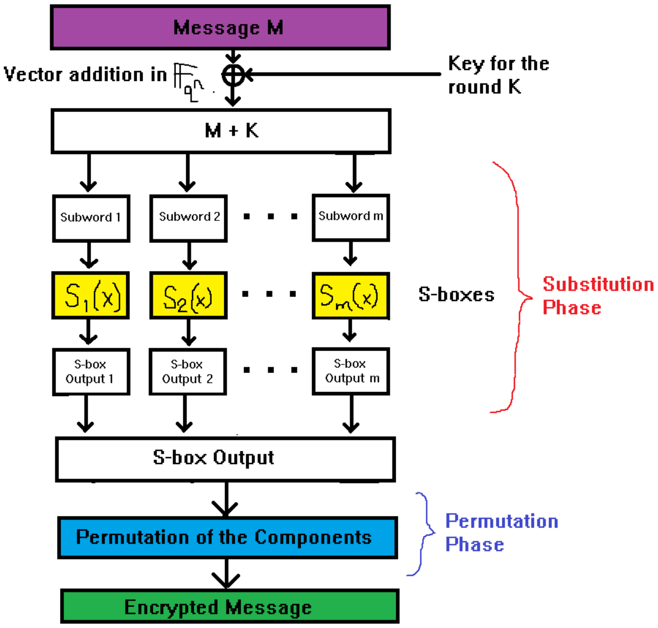

Figure 3.1: A round of an SPN cipher. A typical SPN cipher would consist of many rounds.

stored.

The encryption algorithm can be easily reversed for decryption if we use a class of functions called *semi-involutions* as the functions for the S-boxes.

Definition 14. *A function f over a finite field F is a semi-involution if there exist constants $a, b \in F$ such that $f(f(x+a)+b)$ is the identity function.*

This was originally defined by Youssef, Tavares, and Heys in 1996 [12] for fields of characteristic 2, but the concept of a semi-involution can be generalized to finite fields of any characteristic. Applications of semi-involutions are usually in characteristic 2 as this is the base that virtually all computers use for computation.

3.2 Properties of Semi-Involutions

In this section we cover general results on semi-involutions in finite fields.

We first observe that adding a constant to an involution produces semi-involutions. Youssef et al. [12] gives a construction of semi-involutions from involutions as follows.

Proposition 8. *If f is an involution in $\mathbb{F}_q$, then for any constant $a \in \mathbb{F}_q$, $f - a$ is a semi-involution.*

Proof. Let $g = f - a$. Then for every $x \in \mathbb{F}_q$, we have

$$\begin{aligned} g(g(x+a)+a) &= g(f(x+a)+a-a) \\ &= g(f(x+a)) \\ &= f(f(x+a)) - a \\ &= x + a - a \\ &= x. \end{aligned}$$

□

Remark. The proof given by Youssef et al. does not distinguish addition from subtraction, hence it only works in fields of characteristic two. The above proof works in finite fields of any characteristic.

We now note that $f(f(x+a)+b) = x$ is equivalent to

$$f^{-1}(x) = f(x+a) + b \tag{3.1}$$

Thus, if our S-box was an involution, then we do not need to store an inverse S-box. Instead we can add a to the subword before sending it through the S-box, and add b to the output of the S-box, and we will have the same result as if we sent the subword through the inverse S-box.

We can subtract b from both sides of Equation (3.1) to obtain the following:

$$f^{-1}(x) - b = f(x+a) \tag{3.2}$$

which generally makes it easier to test if a given function f is a semi-involution. In particular, we have the following result.

Lemma 1. *If a polynomial f is a semi-involution over $\mathbb{F}_q$, then, when both f and f^{-1} are reduced modulo $x^q - x$, they will have the same degree (where the zero polynomial is considered to have the same degree as any other constant polynomial).*

Proof. From inspection of Equation (3.2) it is clear that the left- and right-hand sides of the equation must have the same degree modulo $x^q - x$. Since the left-hand side differs from f^{-1} by only a constant, it has the same degree as f^{-1}. Similarly, $f(x+a)$ has the same degree as $f(x)$. Therefore, f^{-1} and f have equal degree modulo $x^q - x$. □

Theorem 1. *[12] If f is a semi-involution with a, b given as in Definition 14, then for every x we have:*

$$f(x+b-a) = f(x) + b - a.$$

Proof. Since f is a semi-involution, we have that Equation (3.2) holds for all x. Let $z = f^{-1}(x)$. Then $z - b = f(x+a)$. Isolating x yields that $x = f^{-1}(z-b) - a$, implying that

$$f(z) = f^{-1}(z-b) - a. \tag{3.3}$$

Since $z = f^{-1}(x)$, and f^{-1} is a bijection, it follows that Equation (3.3) holds for all z. So let $z = u + b$. Then:

$$\begin{aligned} f(u+b) &= f^{-1}(u) - a \\ \Rightarrow f(u+b) &= f(u+a) + b - a, \end{aligned} \tag{3.4}$$

and by similar logic as before, Equation (3.4) holds for all u. So let $w = u+a$. Then we have:

$$f(w+b-a) = f(w) + b - a$$

which is the desired result. □

Remark. The proof given in [12] conflates addition with subtraction and only works in fields of characteristic 2. The proof shown above works in fields of any characteristic.

To understand the ramifications of this theorem, consider sending two subwords through an S-box that used the function f. One of them we call x, the other $x + b - a$. Then since $f(x + b - a) = f(x) + b - a$, the outputs of the two subwords are $f(x)$ and $f(x) + b - a$, respectively. Thus, if the input subwords differ by $b - a$, this guarantees that the output subwords differ by $b - a$. Furthermore, since every S-box in our SPN is a semi-involution, every S-box will have some value of $b - a$ where if the input subwords differ by $b - a$, the output subwords will be guaranteed to differ by $b - a$. This renders the SPN especially vulnerable to differential cryptanalysis [1].

Thus, Theorem 1 renders semi-involutions undesirable for cryptography... unless it so happens that for every pair (a, b) satisfying Definition 14 a equals b, in which case we only have that the same inputs will guarantee the same outputs, which is trivial. While every semi-involution function that we have covered in section 2 has at least one pair (a, b) where a equals b, this is not good enough; this must be true for every such pair.

Youssef et al. [12] seem to be confident that all semi-involutions that satisfy this criterion differ from involutions by the addition of a constant term. For instance, consider the polynomials x^5+x^3+x and $x^5+x^4+x^3+x^2+x$ in the field $\mathbb{F}_8$. For both polynomials we have that the only pair (a, b) that satisfies the definition is (1,1), so if we use these polynomials in our S-boxes we can feel relatively safe from differential cryptanalysis. If we add 1 to both of these polynomials we obtain the polynomials $x^5 + x^3 + x + 1$ and $x^5 + x^4 + x^3 + x^2 + x + 1$, which are both involutions in $\mathbb{F}_8$, as verified in SAGE [11]. However, Youssef et al. does not prove that every "useful" semi-involution differs from an involution by a constant, and as there is very little literature concerning semi-involutions, this is still an open conjecture today. It still remains to be seen if every "useful" semi-involution is a constant added to an involution, or if some "useful" semi-involution cannot be generated in this manner.

Chapter 4

Classes of Semi-Involutions

In this chapter we outline several families of semi- involutions. Some are already known, but we believe that the result concerning Möbius functions has not been covered before. We also state the results of an attempt at classifying Dickson polynomials that are semi-involutions and note topics for possible further research.

4.1 Linearized Functions

First we show that any linearized semi-involution polynomial is an involution.

Lemma 2. *Every linearized semi-involution over a finite field is also an involution.*

Proof. Let L be a linearized semi-involution over a finite field $\mathbb{F}_{q^n}$. Then there exists $a, b \in \mathbb{F}_{q^n}$ such that for all $x \in \mathbb{F}_{q^n}$

$$\begin{aligned} L(L(x+a)+b) &= x \\ \Rightarrow L(L(x)+L(a)+b) &= x \\ \Rightarrow L^2(x)+L^2(a)+L(b) &= x \end{aligned} \tag{4.1}$$

Since L is linearized, L^2 is also linearized and therefore does not have a constant term. The left-hand side of Equation (4.1) therefore has the constant term $L^2(a)+L(b)$, whereas the right-hand side has no constant term. It follows that $L^2(a)+L(b)=0$. Substituting this back into Equation (4.1) gives us $L^2(x)=x$, as desired. □

Since every linearized semi-involution polynomial is an involution we only need to study linearized involutions. In [4], the authors constructed some classes of linearized involutions over finite fields of even characteristic.

4.2 Monomials

Let $f(x)=x^n$ in the field $\mathbb{F}_q$, $0<n<q-1$. Let $y=x+a$. Substituting this into $f(f(x+a)+b)=x$ yields $f(f(y)+b)=y-a$. Since $f(y)=y^n$, we have

$$(y^n+b)^n = y-a. \tag{4.2}$$

We wish to determine when there exist a, b such that Equation (4.2) holds.

Theorem 2. *If $f(y)$ is a semi-involution, then, $n^2 \equiv 1$ mod $q-1$ and $f(y)$ is an involution.*

Proof. Since a function must be a bijection to be a semi-involution, we can safely assume that n is coprime with $q-1$. Thus, n has a unique inverse in the group of units modulo $q-1$. So we define m such that $0 < m < q-1$ and $mn \equiv 1(\text{mod } q-1)$. Raising both sides of Equation (4.2) to the mth power yields:

$$y^n + b = (y-a)^m. \tag{4.3}$$

Since the right-hand side contains a y^m term, the left-hand side must also contain such a term. This cannot be b, as a non-constant term cannot reduce to a constant term modulo $y^q - y$. Therefore, $y^n = y^m$, and $n \equiv m \pmod{q-1}$. This implies that $n^2 \equiv mn \equiv 1 \pmod{q-1}$, which is what we set out to prove. □

4.3 Möbius Functions

Theorem 3. *Any Möbius permutation is a semi-involution.*

Proof. We must find $a, b \in \mathbb{F}_q$ such that

$$f(f(y)+b) = y - a.$$

We first consider the general case where y is neither infinity nor $-zw^{-1}$. Expanding $f(f(y)+b)$ yields

$$\frac{(u^2 + buw + vw)y + (uv + buz + vz)}{(uw + bw^2 + wz)y + (vw + bwz + z^2)} = y - a. \tag{4.4}$$

We then multiply both sides by the denominator of the left-hand side to obtain

$$\begin{aligned}
&(u^2+buw+vw)y+(uv+buz+vz)\\
=&(uw+bw^2+wz)y^2+(vw+bwz+z^2-auw-abw^2-awz)y\\
&\quad-(avw+abwz+az^2). \qquad (4.5)
\end{aligned}$$

Equating the y^2 terms of Equation (4.5) gives us:

$$w(u+z+bw)=0. \qquad (4.6)$$

Since w is nonzero by definition, it follows that $u+z+bw=0$. Equating the y terms gives us:

$$\begin{aligned}
vw+bwz+z^2-auw-abw^2-awz &= u^2+buw+vw\\
uz+bwz+z^2-auw-abw^2-awz &= u^2+buw+uz\\
z(u+z+bw)-aw(u+z+bw) &= u(u+z+bw)\\
(u+z+aw)(u+z+bw) &= 0,
\end{aligned}$$

which tells us nothing new as both the left and right sides equal zero.

Equating the constant terms gives us:

$$\begin{aligned}
uv + buz + vz + avw + abwz + az^2 &= 0 \\
v(u+z) + buz + a(vw + bwz + z^2) &= 0 \\
-bvw + buz + a(vw - uz) &= 0 \\
(b-a)(uz - vw) &= 0. \qquad (4.7)
\end{aligned}$$

Since $uz - vw \neq 0$, it follows that $b - a = 0$ and that $a = b = -(u+z)w^{-1}$.

We now verify the special cases $y = -zw^{-1}$ and $y = \infty$. When $y = -zw^{-1}$, $f(y) = \infty$, and $f(f(y)+b) = f(\infty) = uw^{-1}$. Therefore, $-zw^{-1} - a = uw^{-1}$, or $a = -(u+z)w^{-1}$, a condition we already have.

When $y = \infty$, then we have $f(uw^{-1} + b) = \infty$. It follows that (since f is a bijection) $uw^{-1} + b = -zw^{-1}$. So $b = -(u+z)w^{-1}$, which we have already shown above.

We now show sufficiency. Substituting $a = b = -(u+z)w^{-1}$ into Equation (4.4) results in

$$\begin{aligned}
\frac{-(uz - vw)y + (-uzw^{-1} + v)(u+z)}{-(uz - vw)} &= y + (u+z)w^{-1} \\
y + (u+z)w^{-1} &= y + (u+z)w^{-1},
\end{aligned}$$

therefore, $a = b = -(u+z)w^{-1}$ works for every Möbius permutation. It follows that every Möbius permutation is a semi-involution. □

Remark. It is worth noting that a Möbius permutation is an involution if and only if $a = b = 0$, which is equivalent to $u + z = 0$. Thus any Möbius permutation where $u + z \neq 0$ is a semi-involution that is not an involution.

This is a large family of semi-involutions that are not involutions. If one thinks of these functions as permutations of the set of elements of $\mathbb{F}_q$, then these permutations can have cycles comprising three or more elements.

4.4 Other Classes of Semi-Involutions

We performed an exhaustive search in SAGE [11] to determine all of the Dickson semi-involutions in every field of 47 elements or fewer, as well as every prime field of 71 elements or fewer. We ignored semi-involutions that were involutions as these were already covered in [3]. Surprisingly, every Dickson semi-involution that SAGE found was also an involution, except for these two polynomials in $\mathbb{F}_8$ which are inverses of each other:

$$D_5(x,1) = D_{23}(x,1) = D_{40}(x,1) = D_{58}(x,1) = x^5 + x^3 + x \tag{4.8}$$

$$D_{11}(x,1) = D_{25}(x,1) = D_{38}(x,1) = D_{52}(x,1) = x^5 + x^4 + x^3 + x^2 + x \tag{4.9}$$

It remains to be seen whether these are the only two Dickson semi-involutions that are not involutions, or whether there are more examples of this in fields of higher order.

We also performed an exhaustive search in SAGE [11] to determine all of the Rédei semi-involutions in every field of 17 elements or fewer with odd characteristic. All of them are also involutions, whose properties are covered in [10].

Source code for our SAGE implementations are given in Appendix 5.

Chapter 5

Conclusion

Thus far, little research has been done on the subject of semi-involutions. This is partially because of the introduction of the Advanced Encryption Standard, so more attention has been given to the cryptanalysis of AES as opposed to developing more secure ciphers, and partially because [12] convincingly claimed, without proof, that the semi-involutions useful to cryptography are all offset from involutions by a constant term. However, there is still much to be done in this area. There is no proof that if every pair (a, b) satisfying the definition of a semi-involution itself satisfies $a = b$, then this function is an involution plus a constant term. Nothing is currently known about the cycle structure of semi-involutions. For many classes of functions, there is no characterization of when they are semi-involutions. Even if the study of semi-involutions does not lead to useful cryptographic applications today, they might lead to useful applications decades or even centuries from now that no one today could have imagined.

Appendix

This is the SAGE program we used to output all Dickson semi-involutions that are not involutions over the field of q elements.

```
from sage.rings.finite_rings.integer_mod import lucas_q1

myFile = open(DATA+ ’dickson8.txt’, ’w’)
written = false # This flag is set if something is output.

q = 8 # This is the number of elements in the finite field.

K = GF(q, ’z’)

# Allows us to use ’z’ to refer to the generator of the (multiplicative group
# of the) finite field.
K.inject_variables()

R = PolynomialRing(K, ’x’)

x = R.gen()

# Initialise f and g
```

```
f = x
f = f - x

g = x
g = g - x

# f and g were set to x initially, so that SAGE treats f and g as polynomials
# in the ring R, not as polynomials with real co-efficients (which is what SAGE
# does by default).

for j in range(1, q^2-2):
 if gcd(j, q^2-1) == 1: # If D_j(x,c) is a permutation
  k = inverse_mod(j,(q^2-1)) # The inverse of D_j is D_k
  for c in K:
   f = 0
   g = 0

    # In SAGE, range(m,n) means all of the integers from m to n,
    # including m but not n.
    for i in range(0,floor(j/2)+1):
     f = f + K(j*binomial(j-i,i)/(j-i))*K(c^j)*x^(j-2*i)

    for i in range(0,floor(k/2)+1):
     g = g + K(k*binomial(k-i,i)/(k-i))*K(c^k)*x^(k-2*i)

    ff = f.mod(x^(q)-x)
    gg = g.mod(x^(q)-x)
    #ff is now the Dickson polynomial D_j, and gg is D_k.
```

```
        # Test to see if there exists a, b in K such that
        # ff(x+a) == gg(x)+ b. Only the first pair (a,b) found is recorded.
        # Since the first ordered pair (a,b) that SAGE checks is (0,0), if
        # ff is an involution, SAGE will find (a,b) = (0,0) first and know
        # to ignore ff.
        flag = true
        for a in K:
          for b in K:
            if (ff(x+a) == gg(x)+ b) and flag:
              flag = false
              if c != 0 and (a != 0 or b != 0):
                myFile.write("c is "+str(c)+", j is "+str(j)+", k is "+str(k)+
                myFile.write("the function is "+str(f.mod(x^(q)-x))+"\r\n")
                myFile.write("the inverse is "+str(g.mod(x^(q)-x))+"\r\n")
                myFile.write("semi-involution"+" "+str(a)+" "+str(b)+"\r\n")
                written = true


if written == false: # If nothing has been output
  myFile.write("In this field, all Dickson semi-involutions are involutions.")
myFile.close()
```

This is the SAGE program we used to output all Rédei semi-involutions that are not involutions over the field of q elements, q odd.

```
from sage.rings.finite_rings.integer_mod import lucas_q1


myFile = open(DATA+ 'redei19.txt', 'w')
written = false
```

```
q = 19

K = GF(q, ’z’)
K.inject_variables()

# Here the Redei function is implemented as a subroutine.


def gamma(c):

  G(x) = (x + c^((q-1)/2))/(x - c^((q-1)/2))
  return G




def redeipoly(j,c,x):

  if x == c^((q-1)/2):
    r = c^((q-1)/2)

  else:
    g1 = gamma(c)
    r = (c^((q-1)/2))*((g1(x))^j+1)/((g1(x)^j)-1)
  return r


for j in range(1, q):
  if gcd(j, q+1) == 1:
```

```
  k = inverse_mod(j,q+1)
  for c in K:
   if c == 0:
    continue
   else:


    flag = true
    for a in K:
     for b in K:
      if (redeipoly(j,c,x+a) == redeipoly(k,c,x)+ b) and flag:
       flag = false
       if (a != 0 or b != 0):
        myFile.write("c is "+str(c)+", j is "+str(j)+", k is "+str(k)+
        myFile.write("semi-involution"+" "+str(a)+" "+str(b)+"\r\n")
        written = true


if written == false:
 myFile.write("In this field, all Redei semi-involutions are involutions.")
myFile.close()
```

Both programs have a line of code that extends beyond the right margin of the page: in both programs the line should read ‘myFile.write(”c is ”+str(c)+”, j is ”+str(j)+”, k is ”+str(k)+”\r\n”)’.

Bibliography

[1] E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *J. Cryptology.* **4**:3-72, 1991.

[2] C. Carlitz. A note on permutation functions over a finite field. *Duke Math. J.*, **29**:325-332, 1962.

[3] P. Charpin, S. Mesnager and S. Sarkar. Dickson polynomials that are involutions. *Contemporary Developments in Finite Fields and Applications*, World Scientific, 22-47, 2016.

[4] P. Charpin, S. Mesnager and S. Sarkar. Involutions over the Galois field $\mathbb{F}_{2^n}$. *IEEE Trans. Inform. Theory*, **62**:2266-2276, 2016.

[5] M. Curtin. *Brute Force: Cracking the Data Encryption Standard*, Copernicus Books, 2005.

[6] L.E. Dickson. The Analytic Representation of Substitutions on a Power of a Prime Number of Letters with a Discussion of the Linear Group. *Ann. of Math.*, **11**:65-120, 1896.

[7] H.M. Heys and S.E. Tavares. Substitution-permutation networks resistant to differential and linear cryptanalysis. *J. Cryptology.* **9**:1-19, 1996.

[8] National Institute of Standards and Technology. Data encryption standard. *Federal Information Processing Standards Publication 46-3.* 1999.

[9] W. Keith Nicholson. *Introduction to Abstract Algebra*, 4th ed. John Wiley & Sons, Inc., 2012.

[10] C. Qureshi and D. Panario. Rédei actions on finite fields and multiplication map in cyclic group. *SIAM J. Discrete Math.*. **29**:1486-1503, 2015.

[11] W.A. Stein et al. Sage mathematics software (Version 6.8). The Sage Development Team, http://www.sagemath.org, 2009.

[12] A.M. Youssef, S.E. Tavares and H.M. Heys. A new class of substitution-permutation networks, *Workshop on Selected Areas in Cryptography, SAC '96*, Workshop Record, 132-147, 1996.