Classification and computational search for planar functions in characteristic 3

Alise Haukenes
Supervisors: Lilya Budaghyan and Nikolay S. Kaleyski

Department of Informatics
University of Bergen

2022

Abstract

Planar functions are mappings over a finite field $\mathbb{F}_{p^n}$ having the best possible differential uniformity. The differential uniformity is a measurement of the resistance of a function to differential cryptanalysis, which is one of the most powerful attacks known today that can be used against block ciphers. While this makes the study of planar functions highly relevant from the point of view of cryptography, these functions also correspond to important algebraic and combinatorial structures such as commutative semifields which makes their study important in a much broader context in mathematics and computer science. Unfortunately, planar functions are difficult to construct and analyze. One of the reasons that new constructions are difficult is that planar functions are classified up to a certain notion of equivalence (typically, CCZ-equivalence), and showing that a given function is new involves demonstrating that it is CCZ-inequivalent to all known planar functions. The known constructions, or infinite families, of planar functions can produce very large numbers of planar functions. While most of these will end up being CCZ-equivalent to one another, it is necessary to test all of them for CCZ-equivalence with any newly found instance, and doing so can be a very laborious and time-consuming process.

We classify all known planar functions over $\mathbb{F}_{3^n}$ for n from 3 to 8 up to CCZ-equivalence and give a table of representatives covering all of their CCZ-equivalence classes. At the time of writing, such a classification had only been done for $n \leq 6$. Even for $n \leq 6$, it did not include some recently discovered planar functions, and therefore needed to be updated. We also compute invariants for these representatives, including lists of representatives from their right orbits (as defined in a paper due to Ivkovic and Kaleyski) that can be used to speed up equivalence tests in the future. We organize all the representatives and invariants in tables for ease of reference.

We run expansion searches (that is, we try to find new functions by adding terms to existing functions) over $\mathbb{F}_{3^n}$ with $n \leq 8$ and document their results in order to evaluate the efficiency of this method for finding planar functions. We describe some tricks and optimizations that can be used to speed up such a search and reduce the number of functions that have to be classified. We find seven new planar functions (up to CCZ-equivalence) over $\mathbb{F}_{3^6}$ and we confirm that no new planar functions can be obtained in a number of cases.

A sporadic instace over $\mathbb{F}_{3^8}$ was introduced in 2007 by Coulter et al., and the problem of classifying it into an infinite family had remained open since then. We show that this sporadic instance is, in fact, equivalent to an instance from the Zhou-Pott family of planar functions. We also find a quadrinomial that is CCZ-equivalent to this sporadic instance, and has a significantly simpler univariate polynomial representation than both the sporadic instance itself and the Zhou-Pott functions. Based on this, we generalize the quadrinomial into an infinite family of planar functions. The instances from this family are equivalent to ones from the Zhou-Pott family for small dimensions n; we leave the question open of whether this family is new in general.

Acknowledgements

I would first like to thank my supervisors Lilya Budaghyan and Nikolay Stoyanov Kaleyski at the Selmer Center for their help and guidance during the writing of this thesis. A special thanks to Nikolay for always being available whenever I had questions or needed motivation. The accomplishments of this thesis would not have been possible without him and his extensive knowledge in the field.

Secondly, I would like to thank my fellow students and the Department of Informatics for providing me with a great study environment.

Last but not least, a big thanks to my family and friends for supporting me throughout my studies.

Contents

Chapter 1

Introduction

Vectorial functions, that is, functions from the vector space $\mathbb{F}_p^n$ over the finite field $\mathbb{F}_p$ to the vector space $\mathbb{F}_p^m$, are natural objects in many branches of mathematics. For instance, when $p = 2$, they take a sequence of n bits as input and produce an output sequence of m bits. In the general case (for values of p other than two), the input consists of n digits from among $\{0, 1, \ldots, p-1\}$, and the output consists of m such digits.

More precisely, for a prime number p and natural numbers n, m, we call a function from $\mathbb{F}_p^n$ to $\mathbb{F}_p^m$ an (n, m, p)-function. This is also called a vectorial function if the values of n, m, p are understood from the context. In this thesis, we will mostly focus on the case when $n = m$ and $p = 3$, and we will investigate functions F for which the equation $F(x) - F(x + a) = b$ has a small number δ of solutions x from $\mathbb{F}_p^n$ for every nonzero element a from $\mathbb{F}_p^n$ and every b in $\mathbb{F}_p^n$. When δ is equal to 1, these functions are called planar functions, and provide optimal resistance to differential cryptanalysis, which is one of the most powerful attacks that can be used against block ciphers. These functions also have many other special properties that make them interesting to study.

Planar functions can be defined in several different ways, depending on the field of study. They where first formally introduced by P. Dembowski and T. G. Ostrom in 1968, during their work on projective planes with special properties in finite geometry [23]. However, already in 1965, J. E. H. Elliot and A. T. Butson published a paper on relative difference sets [28] and later it was proved that relative difference sets are equivalent to planar functions. This was shown using the notion that planar mappings may also be seen as projections of relative difference sets [38]. Therefore, it is obvious that planar functions are a central part of many different research areas, where they have implicitly played a role, even before their formal introduction.

Since 1968, there has been an increasing interest for the study of planar functions, and

not just in relation to projective planes where they originated. They have been studied in several different contexts, among which are design theory, the study of semifields, cryptography and coding theory. The reason for this increase in interest were the results of Dembowski and Ostrom, where they proved that the existence of a planar function is equivalent to the existence of an affine plane whose projective closure satisfies certain criteria [17]. Planar functions are also important for cryptography because of their optimal differential uniformity, and also in a broader mathematical context, for their revolutionary role in the study of commutative semifields. More precisely, it was shown that there is a correspondence between quadratic planar functions and commutative semifields, and this connection was exploited to construct new families of semifields that had eluded researchers before.

Even though planar functions have been studied for a long time and are related to many other fields of study in mathematics, new planar functions have proved to be difficult to find. We can clearly see this based on the number of infinite families of planar functions and sporadic instances that we know of so far, which are far from many. There are many reasons for why finding new planar functions is a difficult task, among which is the vast search space which results in a huge number of possible candidates for new planar functions. In addition to this, we have identified the following two problems that add to the complexity of the task . We will attempt to ease these problems by the work done in this thesis.

The first problem stems from the fact that planar functions are classified up to certain equivalence relations. When researchers find planar functions they believe might be new, they need to verify that these functions are inequivalent to all of the previously known ones. In other words, they have to classify the newly discovered functions up to equivalence. Classifying functions is difficult because they have to be compared to all representatives from the known families and sporadic instances. Despite us knowing very few families, in some cases these families can generate hundreds of thousands of functions (which will however only fall into very few equivalence classes), which makes this comparison a time-consuming and intense task.

The second problem is that constructing new families appears to be very difficult per se. Just finding new planar functions seems to be extremely hard. Finding instances of planar functions with other desirable properties, such as having a short polynomial representation, is of course even harder. Besides providing a more compact representation, a short polynomial form may allow us to more clearly see patterns and structure that the function possesses. For example for one of the known sporadic instances in dimension 8 with characteristic 3, the shortest previously known representation consisted of 16 terms. One of our results is that we found a quadrinomial (that is, a function of only 4 terms) equivalent to it, which allowed us to generalize this sporadic instance into an infinite family of quadrinomials. We show that the sporadic instance is also equivalent to an instance from the Zhou-Pott family of planar

functions (although the Zhou-Pott representation is still significantly more complicated than the quadrinomial that we find). While both the sporadic instance and the Zhou-Pott family had been known for some time, this equivalence had never been observed before. In fact, to the best of our knowledge the problem of classifying the sporadic instance into an infinite family had been open since its introduction in 2007.

In this thesis, we first generate all possible functions from the known families with characteristic 3 in dimension n for n from 3 to 8 and we classify them up to CCZ-equivalence, together with the known sporadic instances. We then select a representative for each equivalence class, and give a table of CCZ-inequivalent representatives from all the known infinite families and sporadic instances. Using this table, one only has to compare newly found planar functions for equivalence against a very small number of functions (instead of generating all instances from the known families and classifying them up to equivalence from scratch). Furthermore, we compute the values of some of the most useful known invariants for the selected representatives, including the so-called right automorphism orbits as defined in [31]. Knowledge of these orbits allows the computation time for performing equivalence tests to be reduced even further when using the algorithm from [31], which at the moment appears to be the most efficient way for testing equivalence.

We then use these tables to run computational searches (based on the "polynomial expansion" technique) to try and find new planar functions, as well as more compact representatives of known functions. In this way, we evaluate the efficiency of the "polynomial expansion" approach for some of the dimensions up to 8. More precisely, we focus on $n = 6$ and $n = 8$ since it is possible to restrict the coefficients in searches to subfields (while for $n = 5$ and $n = 7$, being prime, it is only possible to restrict the coefficients to the prime field $\mathbb{F}_3$, and all quadratic planar functions of this form have already been classified in [22]). Among these, we dedicate most of our computational efforts to $n = 8$, since computational tools for efficiently checking equivalence in this dimension such as [31] have only recently become available, and we expect that it is more likely to find new instances of planar functions there.

Using polynomial expansion searches, we find seven new classes of planar functions (up to CCZ-equivalence) over $\mathbb{F}_{3^6}$. Over $\mathbb{F}_{3^8}$ (which we explore in even more depth using expansion searches), we do not find any new planar functions. However, we show that a known sporadic instance from [19] over $\mathbb{F}_{3^8}$ is in fact CCZ-equivalent to an instance from the Zhou-Pott family. In this way, as mentioned above, we solve the problem of classifying this sporadic instance into an infinite family, which to the best of our knowledge had remained open since its introduction in 2007.

Furthermore, we find a significantly more compact polynomial representation (as compared both to its original representation, and the equivalent instances arising from the Zhou-Pott

family) of this sporadic instance. Based on this compact representation, we construct an infinite family of planar quadrinomials that contains the sporadic instance. While our family intersects the Zhou-Pott functions for $n = 4$ and $n = 8$, it provides functions with a significantly simpler univariate representation, and appears promising from the point of view of further generalization. Furthermore, the question of whether the two families continue to coincide for dimensions n greater than 8 remains open.

Chapter 2

Preliminaries

2.1 Functions over finite fields and their representations

Let p be a prime number and n be a natural number. Then we denote the finite field with p^n elements by $\mathbb{F}_{p^n}$, and we denote the n-dimensional vector space over $\mathbb{F}_p$ by $\mathbb{F}_p^n$.

Mappings from $\mathbb{F}_p^n$ to $\mathbb{F}_p^m$ (for some prime number p and some natural numbers n and m) are natural mathematical objects that have numerous applications in many branches of mathematics and computer science. These are called (n, m, p)**-functions**, or simply **vectorial functions**.

Although (n, m, p)-functions are formally defined as mappings between vector spaces, we can also see them as functions between finite fields. This is due to the fact that the vector space $\mathbb{F}_p^n$ can be identified with the finite field $\mathbb{F}_{p^n}$. In this thesis, we will mostly be considering vectorial functions as functions over finite fields.

Vectorial functions can be represented in a number of different ways, all of which have advantages and disadvantages. In the following, we introduce some of the most frequently used representations.

The simplest way to define an (n, m, p)-function F is to represent it as a **truth table** (TT). This means to define the function F by explicitly listing the output values $F(x) \in \mathbb{F}_p^m$ for all possible inputs $x \in \mathbb{F}_p^n$. In the case of an (n, m, p)-function with $p = 3$, each input consists of n variables which can take the values 0, 1 or 2 and corresponds to an output $F(x)$ of length m. An example of a truth table of a $(2, 2, 3)$-function is given in Table 2.1.

x	$F(x)$
00	00
01	11
02	12
10	10
11	20
12	21
20	22
21	01
22	02

Table 2.1: Truth table of an $(2, 2, 3)$-function

The TT representation is conceptually very simple, and it allows values of the function to be determined very quickly. On the other hand, the TT can occupy a lot of memory (especially when n and m are large) so it is only possible to use it for relatively small dimensions. Another shortcoming is that the TT reveals very little about the structural properties of the function: for instance, from the algebraic normal form, which will be presented shortly, we can immediately derive the algebraic degree of the function, but this is very hard to do using the TT.

In the example above, in Table 2.1, we see the TT used to represent a vectorial function, with two input and two output variables. We can alternatively express this by two functions, f_1 and f_2, that represent the outputs of $F : \mathbb{F}_3^2 \to \mathbb{F}_3^2$ as a vector of two variables:

$$F(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2)).$$

Generalizing this to arbitrary m, it follows that any (n, m, p)-function F can be seen as a vector of $(n, 1, p)$-functions:

$$F = (f_1, ..., f_m),$$

where the $(n, 1, p)$-functions $f_1, ..., f_m$ are called the **coordinates**, or the **coordinate functions** of F. Any non-zero linear combination of the coordinates of a function F is called a **component function** of F. The component functions are used, for example, in the definition of nonlinearity, which is an important cryptographic parameter. Since we do not directly deal with nonlinearity in our work, we do not go into further details.

Another way to represent a vectorial function, is in ANF, short for **algebraic normal form**. This is a uniquely defined multivariate polynomial representation with coefficients from $\mathbb{F}_p^m$ and variables from $\mathbb{F}_p$, meaning that for a function $F : \mathbb{F}_p^n \to \mathbb{F}_p^m$, the ANF of F is of the

form

$$F(x_1, ..., x_n) = \sum_{u \in \mathbb{F}_p^n} a_u \prod_{i=1}^{n} x_i^{u_i}, a_u \in \mathbb{F}_p^m.$$

One of the advantages of the ANF over the TT is that some important properties of the function can be deduced immediately from the ANF. One such property is the so-called algebraic degree, which has many important connotations to the implementation and cryptographic strength of the function. For example, the higher the degree, the more resistant the function is to higher-order differential attacks [26] [33]. For an (n, m, p)-function F, the **algebraic degree** of F, denoted by $\deg(F)$, is the degree (as a multivariate polynomial) of its ANF.

In this thesis, we are mostly interested in the case of (n, m, p)-functions where $n = m$. In this case, representing these functions is often most convenient using the **univariate representation**. The vector space $\mathbb{F}_p^n$ can be identified with $\mathbb{F}_{p^n}$, and we can consider functions from $\mathbb{F}_p^n$ to itself as mappings from $\mathbb{F}_{p^n}$ to itself. Any such function F has a unique representation as a univariate polynomial over $\mathbb{F}_{p^n}$ of degree smaller than p^n:

$$F(x) = \sum_{i=0}^{p^n-1} c_i x^i, c_i \in \mathbb{F}_{p^n}.$$

The algebraic degree can also easily be derived from the univariate representation. More precisely, $\deg(F)$ is the largest p-ary weight of any exponent i for which the coefficient c_i is non-zero. For example, the 3-ary expansion of the integer 17 is $(1, 2, 2)$ since $17 = 3^2 + 2 \cdot 3^1 + 2 \cdot 3^0$, and therefore its 3-ary weight is $1 + 2 + 2 = 5$. Consequently, any function having a non-zero coefficient in front of x^{17} has algebraic degree at least 5.

Besides the cryptographic significance related to higher-order differential attacks, the algebraic degree can be used to define affine, linear and quadratic functions. An (n, m, p)-function F is **affine** if the algebraic degree is at most 1, **quadratic** if the algebraic degree is exactly 2 and **cubic** if the algebraic degree is exactly 3. If an affine function A satisfies $A(0) = 0$, we say that A is **linear**.

Alternatively, a function F from $\mathbb{F}_{p^n}$ to itself is linear if

$$F(x) = \sum_{0 \le i < n} c_i x^{p^i}, c_i \in \mathbb{F}_{p^n}.$$

If F is a sum of a linear function and a constant, it is affine. Furthermore, we say that the

function F is a **Dembowski-Ostrom polynomial** (DO polynomial) if

$$F(x) = \sum_{0 \leq k \leq j < n} c_{kj} x^{p^k + p^j}, c_{kj} \in \mathbb{F}_{p^n}.$$

Thus, if a function is a sum of a DO polynomial and an affine function, it is quadratic [9].

These types of functions are important when working on planar functions for different reasons. One reason is for example that there is a relation between commutative semifields and quadratic planar functions [18]. Affine and linear functions are used in the definition of various equivalence relations such as linear equivalence and EA-equivalence that we will see later.

2.2 Cryptographic parameters

Block ciphers are widely known and much used in modern cryptography. The idea is to combine simple operations in order to construct a complex encryption transformation. Blocks of data are encrypted by transforming the input (plaintext) block into an output (ciphertext) block, usually with the same length as the original block. The vast majority of modern block cipher designs use (n, m, p)-functions as fundamental components of these transformations. In order for the encryption of these blocks of data to be strong, certain properties are required of the functions used. If these properties are lacking, it leaves the encryption vulnerable to different attacks, which might give an attacker the opportunity to obtain information about the original data block, or even to decrypt the message altogether.

One of the most powerful attacks against block ciphers known today is the so-called differential cryptanalysis [7]. Very briefly, the basic idea of differential cryptanalysis is to study how the difference $d_y = y_2 - y_1$ of two outputs y_1 and y_2 of a function depends on the difference $d_x = x_2 - x_1$ of their corresponding inputs x_1 and x_2. If some output difference d_y is more likely than uniform for a given d_x, this might give an attacker information about the encryption that can be used to break the cipher. Therefore, the output differences of a cryptographically strong function should be as uniformly distributed as possible in order to provide good security against this type of attack.

Given an (n, m, p)-function, its derivative is a function that expresses the relation between the input and output differences. The formal definition is as follows.

Let F be an (n, m, p)-function. The map $D_a F(x) = F(x + a) - F(x)$ is the **derivative of F** in the direction of $a \in \mathbb{F}_p^n$.

Alternatively, the derivative of F can be defined as $\Delta_a F(x) = F(x+a)-F(x)-F(a)+F(0)$. This has the advantage of having 0 map to 0; in particular, if F is quadratic, then $\Delta_a F$ is linear. We will later use this when constructing a family of planar functions (more precisely, we will use this in the proof that the functions from this family are planar).

The derivative $D_a F$ expresses exactly the output difference of all pairs of inputs whose input difference is equal to a; and we want $D_a F$ to be as uniformly distributed as possible in order to resist differential cryptanalysis. The concept of being uniformly distributed is captured by the following notion.

An (n, m, p)-function F from $\mathbb{F}_p^n$ to $\mathbb{F}_p^m$, is called **balanced** if it takes every value of $\mathbb{F}_p^n$ the same number of times, p^{n-m}.

In this thesis, we mostly concentrate on the case of $n = m$. For $n = m$, the balanced functions over $\mathbb{F}_p^n$ are precisely the permutations. While balanced functions represent the optimal case, the notion of differential uniformity is used to measure how good the resistance of a function to differential cryptanalysis is even if it is not balanced.

If the equation $D_a F(x) = F(x) - F(x + a) = b$ has at most δ solutions for every nonzero element a of $\mathbb{F}_p^n$ and every b in $\mathbb{F}_p^m$, it is called **differentially δ-uniform**.

The maximum number of solutions to the above equation through all $a \neq 0$ and b is called the **differential uniformity** of F and is denoted by Δ_F.

In the following, suppose $n \geq m$. By the above discussion, (n, m, p)-functions F with the value $\Delta_F = p^{n-m}$, which is the smallest possible value, contribute an optimal resistance to the differential attack. These functions are called PN, and are the main topic of study in this thesis. This optimal case occurs precisely when all the derivatives of F are balanced, which motivates the following definition.

If for any $a \in \mathbb{F}_p^n$, the derivative of F in the direction of a, $D_a F(x) = F(x + a) - F(x)$, is balanced, then the function is called **perfect nonlinear** (PN), or **planar** if $n = m$.

In the case of $n = m$, a function F is clearly PN if and only if all of its derivatives $D_a F$ for $a \neq 0$ are permutations. These functions exist only for p odd, because if p is even and x_0 is a solution of $D_a F(x) = F(x + a) - F(x)$, then $x_0 + a$ is also a solution.

2.3 Presemifields and semifields

As mentioned before, planar functions are useful not only for their cryptographic properties but because they correspond to optimal objects in other areas of mathematics. Perhaps the best example are algebraic objects called presemifields and semifields. After the classification of finite fields was complete, researchers started investigating such more general structures defined by relaxing conditions and axioms. Semifields are a relaxation of finite fields (they satisfy all the axioms except that they do not have to be associative). Although the term semifield was not used in the earlier literature, the study of these algebraic objects was initiated in the beginning of the 20th century by Dickson [24]. For approximately 60 years they were referred to as "nonassociative division rings" or "distributive quasifields", until Knuth [34] introduced the term *semifields* in 1965. For a long time after the study of semifields was initiated, there was little progress in the area. One of the major factors that contributed to the recent development of semifields was the introduction of planar functions in the 1970s by Dembowski and Ostrom [23] and the important connection between commutative semifields and quadratic planar functions, which was precisely formulated in [18]. Another major factor was the recent progress in the construction of APN polynomials [9]. APN polynomials are optimal with respects to differential uniformity in the case of even characteristic. Many of the constructions of APN polynomials could be adopted to the planar case, and therefore to the construction of semifields.

2.3.1 Definitions

Before defining the notion of a commutative semifield, we first define that of a presemifield. A finite presemifield is a finite set S with two binary operations $+$ and $*$ satisfying the following axioms:
$\mathbf{S_1}$: $(S,+)$ is an Abelian group with identity 0.
$\mathbf{S_2}$: The left and right distributive law holds, i.e. $a*(b+c) = a*b+a*c$ and $(a+b)*c = a*c+b*c$ for any $a, b, c \in S$.
$\mathbf{S_3}$: If $a * b = 0$ then a or b is 0.
If, in addition to the axioms for a presemifield, we also have
$\mathbf{S_4}$: There exists an element $1 \in S$ such that $1 \neq 0$ and $1 * a = a = a * 1$ for all $a \in S$ then the presemifield is called a semifield.

In short, a **presemifield** is a ring with left and right distributivity and no zero divisors, and a presemifield with a multiplicative identity is called a **semifield**. A **commutative** semifield (or presemifield) is one whose multiplicative operation $*$ is commutative. A presemifield can be represented by $\mathbb{S} = (\mathbb{F}_{p^n}, +, *)$, where $\mathbb{F}_{p^n}$ is the finite field with p^n elements. The prime

number p is then called the characteristic of $\mathbb{S}$, and the integer n is called the dimension of $\mathbb{S}$ [34].

Given any presemifield, we can construct a semifield from it as follows. Assume the presemifield $\mathbb{S}$ is commutative and does not contain an identity. In order to create a semifield from $\mathbb{S}$, we choose any $a \in \mathbb{F}_{p^n}^*$ and define a new multiplication $\star$ by $(x * a) \star (a * y) = x * y$, for all $x, y, \in \mathbb{F}_{p^n}$. Now $\mathbb{S}' = (\mathbb{F}_{p^n}, +, \star)$ is a commutative semifield related to $\mathbb{S}$ with identity $a * a$. In fact, $\mathbb{S}$ and $\mathbb{S}'$ are isotopic to each other; we will discuss isotopism in Section 2.4.2.

2.3.2 Commutative semifields and planar functions

Every commutative presemifield defines a planar DO polynomial and vice versa [9]. If we have a quadratic PN function F over $\mathbb{F}_{p^n}$, then we can define the commutative presemifield $\mathbb{S} = (\mathbb{F}_{p^n}, +, *)$, with $x * y = F(x+y) - F(x) - F(y)$ for any $x, y \in \mathbb{F}_{p^n}$. We can then denote the commutative semifield corresponding to the commutative presemifield $\mathbb{S}$ by $\mathbb{S}_F = (\mathbb{F}_{p^n}, +, \star)$. Then $\mathbb{S}_F = (\mathbb{F}_{p^n}, +, \star)$ is a commutative semifield defined by the quadratic PN function F. It is possible to go the other way as well, from a commutative presemifield to a planar DO polynomial as follows. Given a commutative presemifield $\mathbb{S} = (\mathbb{F}_{p^n}, +, *)$ of odd order, we can define a planar DO polynomial given by the function $F(x) = \frac{1}{2}(x * x)$.

Even though commutative semifields have been a popular field of study since their introduction in 1906, there are few cases of commutative semifields of odd order known today [9]. Since almost all known planar functions are DO polynomials, researchers have been able to find new commutative semifields by exploiting the connection between quadratic PN functions and commutative semifields. Despite this, constructing new semifields and presemifields remains a very difficult problem, and is one of the major reasons that planar functions and their constructions are important objects of study.

2.4 Equivalence relations

Finding new planar functions among all (n, m, p)-functions by doing an exhaustive search is an impossible task to undertake, even for relatively small numbers n, m and p. This is because the number of (n, m, p)-functions is $(p^m)^{(p^n)}$, which increases drastically even when incrementing only one of m or n by 1. For instance, if we consider the functions that take 2 input variables and return 2 output variables which can take the values 0, 1 or 2, we have $9^9 = 387420489$ distinct $(2, 2, 3)$-functions. Furthermore, we're usually more interested in values of n and m larger than 2, meaning that with today's technology it would be impossible to go through all

of the (n, m, p)-functions when searching for new planar functions.

Even if we are able to find all planar functions in some field $\mathbb{F}_{p^n}$, it is safe to say that they are going to be too many to handle. If we just find one planar function for every 1000 (n, m, p)-functions we search through, it would leave us with 387420 planar functions, even if using the uninteresting example of the $(2, 2, 3)$-functions above. Normally this number would be significantly larger, which means that going through this set of functions looking for useful properties would be extremely time consuming, maybe even impossible. Because of this, we use equivalence relations both to reduce the number of functions found and to also make searches for new functions easier. The equivalence used must be such that it preserves the properties of interest, such as being planar in our case. In this section, we introduce the most important equivalence relations that preserve planarity.

2.4.1 CCZ-equivalence

The most universal equivalence relation used in practice that preserves differential uniformity for any function is the CCZ-equivalence, short for Carlet-Charpin-Zinoviev-equivalence, introduced in [15]. Of course, not all properties are invariant under CCZ-equivalence; for example, one such property is the algebraic degree. This can be used constructively, since functions of high algebraic degree are known to resist higher-order differential attacks. In the case of planar functions, CCZ-equivalence coincides with EA-equivalence, which does preserve the algebraic degree; however, CCZ-equivalence can be used constructively to find higher-degree representatives of other classes of functions such as APN functions.

The CCZ-equivalence of two functions is defined in terms of their graphs, where the graph of a function $F : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ is the set $\Gamma_F = \{(x, F(x)) : x \in \mathbb{F}_{p^n}\}$. Note that the graph is a set of pairs from $\mathbb{F}_{p^n} \times \mathbb{F}_{p^m}$. But $\mathbb{F}_{p^n} \times \mathbb{F}_{p^m}$ can be identified with the finite field $\mathbb{F}_{p^{n+m}}$, which allows us to define affine permutations of $\mathbb{F}_{p^n} \times \mathbb{F}_{p^m}$. Now, two functions F and G are **CCZ-equivalent** if there is an affine permutation A mapping Γ_F to Γ_G, i.e. $\{A(x, F(x)) : x \in \mathbb{F}_{p^n}\} = \Gamma_G$.

The only known way to efficiently test CCZ-equivalence in general is through linear codes. A linear code $C(F)$ is associated with each function F, and then F and G are CCZ-equivalent if and only if $C(F)$ and $C(G)$ are isomorphic. More details can be found in e.g. [27]. Algorithms for testing isomorphism of codes have existed for quite some time, but they are slow and memory-intensive, which means there is currently no known efficient way of testing CCZ-equivalence in higher dimensions. This is also one of the reasons why little progress has been made when it comes to finding new planar functions, as the code isomorphism test has been the only known way for testing the CCZ-equivalence of planar functions. The recently introduced algorithms for testing linear equivalence of planar functions [31] can be used to

test CCZ-equivalence much more efficiently, since in the case of quadratic planar functions CCZ-equivalence coincides with linear equivalence; more details are given below.

2.4.2 Isotopic equivalence

Isotopic equivalence is even more general than CCZ-equivalence, but it only applies in the case of quadratic planar polynomials. The notion of isotopic equivalence is based on that of isotopism, which was originally defined in the study of presemifields and semifields [1]. Two presemifields, $\mathbb{S}_1 = (\mathbb{F}_{p^n}, +, *)$ and $\mathbb{S}_2 = (\mathbb{F}_{p^n}, +, \star)$, of the same order are said to be **isotopic** if there exist three linear permutations L, M, N over $\mathbb{F}_{p^n}$ such that for any $x, y \in \mathbb{F}_{p^n}$

$$L(x * y) = M(x) \star N(y).$$

The triplet (M, N, L) is called the **isotopism** between $\mathbb{S}_1$ and $\mathbb{S}_2$ and if $M = N$ then $\mathbb{S}_1$ and $\mathbb{S}_2$ are **strongly isotopic**.

Because of the correspondence between quadratic planar functions and semifields shown in section 2.3.2, we can use isotopism to define a notion of equivalence between two quadratic planar functions. Given two quadratic planar (n, n, p)-functions F and F, they are called **isotopic equivalent** if their corresponding presemifields $\mathbb{S}_F$ and $\mathbb{S}_{F'}$ are isotopic.

Isotopic equivalence is related to CCZ-equivalence as follows:

- two quadratic planar polynomials are CCZ-equivalent if and only if their corresponding commutative semifields are strongly isotopic; [13]
- two commutative presemifields of order p^n with n odd are strongly isotopic if and only if they are isotopic; [18]
- a commutative presemifield of order p^n with n even can generate at most two CCZ-equivalence classes of quadratic planar polynomials. [18]

2.4.3 EA-equivalence

Extended affine equivalence, or EA-equivalence for short, is useful because it can be simpler to use or test than CCZ-equivalence. Actually, in some cases, like in the case of planar functions, testing CCZ-equivalence can be reduced to testing EA-equivalence; more precisely, two planar functions are CCZ-equivalent if and only if they are EA-equivalent [12] [11]. Two (n, m, p)-functions F and G are **EA-equivalent** if there are affine permutations A_1, A_2 of $\mathbb{F}_{p^n}$ and $\mathbb{F}_{p^m}$,

respectively and an affine (n, m, p)-function A such that

$$A_1 \circ F \circ A_2 + A = G,$$

where $\circ$ represents composition.

One reason that EA-equivalence is relevant in practice is because testing it may sometimes be done in a simpler way than for CCZ-equivalence. For example, in the case of even characteristic, simple algorithms for testing EA-equivalence without going through linear codes have recently been developed, e.g. [14], [32]. While these algorithms cannot directly be applied to the case of planar functions, they show that EA-equivalence can be more tractable than CCZ-equivalence.

2.4.4 Linear equivalence

Linear equivalence is one of the simplest notions of equivalence that can be defined for (n, m, p)-functions. Two (n, m, p)-functions F and G are **linear equivalent** if there exist linear permutations A_1 and A_2 of $\mathbb{F}_{p^n}$ and $\mathbb{F}_{p^m}$, respectively, such that

$$A_1 \circ F \circ A_2 = G.$$

As we can see from the equation above, linear equivalence is the special case of EA-equivalence when $A = 0$ and A_1, A_2 are linear. In addition to this, we also have that linear equivalence coincides with CCZ-equivalence for DO planar polynomials [11] [12]. This makes our searches for new planar functions significantly easier; we mostly focus on DO planar polynomials and it is therefore enough to test linear equivalence. We use the algorithm for testing linear equivalence between 2-to-1 planar functions from [31]. This method is significantly faster that the code isomorphism test, and allows us to test many more functions than would have otherwise been possible. In addition, code isomorphism does not work reliably on the hardware that we have available in $\mathbb{F}_{3^n}$ for $n > 7$ due to memory constraints, so the algorithm from [31] is the only possible way at the moment to classify functions over $\mathbb{F}_{3^8}$.

In addition, since in the case of quadratic planar functions CCZ-equivalence reduces to linear equivalence, we make use of the following simple trick to significantly reduce the search space.

Suppose that we are expanding a monomial x^d by adding a number $K+1$ of terms to it, the first of which is cx^e. We thus consider all functions of the form $x^d + cx^e + c_1x^{e_1} + \cdots + c_Kx^{e_K}$ for all possible choices of $c, c_1, \ldots, c_K$ and $e, e_1, \ldots, e_K$. Let $f(x) = x^d + cx^e + c_1x^{e_1} + \cdots + c_Kxc^{e_K}$.

The function

$$f(x^{p^{n-1}})^p = x^d + c^p x^e + c_1^p x^{e_1} + \cdots + c_K^p x^{e_K}$$

is clearly linear-equivalent to $f(x)$. We can see that it can be obtained from the same exact polynomial expansion, i.e. by starting with x^d and adding terms to it. The difference is that the coefficient c of the first term has been replaced with c^p. Therefore, when searching for functions up to equivalence, it suffices to only consider one representative from each cyclotomic coset when guessing the value c of the first term added to x^d. Unfortunately, this cannot be applied to the coefficient c_1 of the second term, since raising c_1 to the power p also changes c. Nonetheless, this reduces the number of guesses for c by a factor of approximately n, and has the effect of significantly reducing the search space.

A similar trick can be performed by taking $f(x)$ as above and taking the linear-equivalent

$$\frac{1}{a^d} f(ax) = x^d + ca^{e-d}x^e + c_1 a^{e_1-d}x^{e_1} + \cdots + c_K a^{e_K-d}x^{e_K},$$

where a is some non-zero element from $\mathbb{F}_{p^n}$. This means that if we first guess the exponent e of the first term, then only one coefficient c from each set of the form $\{cm : m \in M\}$ where $M = \{a^{d-e} : 0 \neq a \in \mathbb{F}_{p^n}\}$ has to be considered.

2.5 Invariants

Invariants are properties that are preserved under a given equivalence relation. Examples of such properties include differential uniformity for CCZ-equivalence and both differential uniformity and the algebraic degree for EA-equivalence. The fact that the differential uniformity is invariant under CCZ-equivalence is the reason that CCZ-equivalence can be meaningfully used to classify planar functions. However, we also have some invariants that can take a multitude of distinct values even when the functions have the same differential uniformity. These invariants are an important facilitator for classifying functions up to equivalence, because they can be used to quickly distinguish between inequivalent functions. If for example two given functions have different values of an invariant under CCZ-equivalence, we can immediately conclude that they are CCZ-inequivalent. Invariants have been very successfully used in this way in the case of e.g. APN functions over finite fields of even characteristic. Unfortunately, for planar functions we do not know that many invariants. In the following, we summarize the invariants that we do know.

The orders of the nuclei is one of the invariants that has been found to be useful for quadratic planar functions; the nuclei are restricted to the quadratic case because they are defined in terms of the corresponding semifield. The **nuclei** are invariant under isotopism

and measure how far the semifield $\mathbb{S}$ is from being associative. Given a finite semifield $\mathbb{S} = (\mathbb{F}_{p^n}, +, *)$, the subsets

$$N_l(\mathbb{S}) = \{\alpha \in \mathbb{S} : (\alpha * x) * y = \alpha * (x * y) \text{ for all } x, y \in \mathbb{S}\},$$
$$N_m(\mathbb{S}) = \{\alpha \in \mathbb{S} : (x * \alpha) * y = x * (\alpha * y) \text{ for all } x, y \in \mathbb{S}\},$$
$$N_r(\mathbb{S}) = \{\alpha \in \mathbb{S} : (x * y) * \alpha = x * (y * \alpha) \text{ for all } x, y \in \mathbb{S}\}$$

are called the **left, middle** and **right nucleus** of $\mathbb{S}$, respectively. The set $N(\mathbb{S}) = N_l(\mathbb{S}) \cap N_m(\mathbb{S}) \cap N_r(\mathbb{S})$ is called the **nucleus**. The subsets $N_l(\mathbb{S}), N_m(\mathbb{S}), N_r(\mathbb{S})$ are multiplicative cosets of finite fields, and if $\mathbb{S}$ is commutative then $N_l(\mathbb{S}) = N_r(\mathbb{S}) \subseteq N_m(\mathbb{S})$ [34]. We note in [34], it is claimed that the nuclei are subfields of $\mathbb{F}_{p^n}$, but in fact they are multiplicative cosets.

Another invariant that can be used to distinguish between CCZ-inequivalent functions is the order of the monomial automorphism group of a particular linear code associated with the function [39]. This order is invariant under CCZ-equivalence, and it can take distinct values for CCZ-inequivalent planar functions. As part of our work, we compute the order of this automorphism group for representatives from all known CCZ-classes of planar functions.

Finally, the right orbits as defined in [31] can be used to significantly reduce the computation time for deciding equivalence using the algorithm from the same paper. The number and sizes of these orbits is an invariant under linear equivalence. In our work, we also compute the exact orbits for all the represenatives from the known families.

2.6 Known cases of planar functions and commutative semifields

Constructing planar functions and commutative semifields is known to be very hard. Despite many years of research, we only know very few constructions. Here, we give a summary of the known infinite families and sporadic instances of planar functions.

Some of the functions are given by a simple univariate representation. Others are defined as corresponding to commutative semifields. We list these functions below. The ones given in univariate form are labeled U_i, and the ones defined through semifields are labeled S_i.

(U_1) x^2 in $\mathbb{F}_{p^n}$ (folklore).

(U_2) x^{p^k+1} in $\mathbb{F}_{p^n}$, $k \leq n/2$ and $n/\gcd(k, n)$ is odd, which corresponds to Albert's commutative twisted fields [21, 23].

(U_3) $x^{10} + x^6 - x^2$ in $\mathbb{F}_{3^n}, n \geq 5$ is odd or $n = 2$ [21]. Corresponds to the Coulter-Matthews-Ding-Yang semifields (CMDY[1]).

(U_4) $x^{10} - x^6 - x^2$ in $\mathbb{F}_{3^n}, n \geq 5$ is odd or $n = 2$ [25, 18]. Corresponds to the Coulter-Matthews-Ding-Yang semifields (CMDY[2]).

(U_5) $x^{\frac{3^k+1}{2}}$ in $\mathbb{F}_{3^n}, k \geq 3$ is odd and $gcd(k, n) = 1$. They are referred to as the Coulter-Matthews (CM) planar functions [21, 30] and they do not correspond to any semifield.

(U_6) $x^{1+q'} - vx^{q^2+q'q}$ over $\mathbb{F}_{q^3}$ where p is an odd prime, $q = p^s, q' = p^t, s' = s/gcd(s,t), t' = t/gcd(s,t), s'$ is odd, $ord(v) = q^2 + q + 1$, and at least one of the following conditions holds:

$$s' + t' \equiv 0 \bmod 3,$$
$$q \equiv q' \equiv 1 \bmod 3.$$

This planar function family corresponds to the Zha-Kyureghyan-Wang (ZKW) semifields [41, 4]. This construction was motivated by the APN binomials from [10].

(U_7) $x^{1+q'} - vx^{q^3+q'q}$ over $\mathbb{F}_{q^4}$, where p is an odd prime, $q = p^s, q' = p^t$ such that $2s/\gcd(2s,t)$ is odd, $q \equiv q' \equiv 1 \bmod 4$, and $\mathrm{ord}(v) = q^3 + q^2 + q + 1$. It corresponds to the Bierbrauer semifields [5]. This was also motivated by the APN binomial construction from [10].

(U_8) The following two families due to Budaghyan and Helleseth (BH):

(a) $(bx)^{p^s+1} - \left((bx)^{p^s+1}\right)^{p^k} + \sum_{i=0}^{k-1} c_i x^{p^{i(p^k+1)}}$, where p is an odd prime, s and k are positive integers such that $\gcd(p^s+1, p^k+1) \neq \gcd(p^s+1, (p^k+1)/2)$, $\gcd(k+s, 2k) = \gcd(k+s, k)$, and $n = 2k$, $b \in \mathbb{F}_{p^n}^*$ and $\sum_{i=0}^{k-1} c_i x^{p^i}$ is a permutation over $\mathbb{F}_{p^k}$ with coefficients in $\mathbb{F}_{p^k}$ [11].

(b) $\mathrm{Tr}_k^{2k}(bx^{p^s+1}) + cx^{p^k+1} + \sum_{i=1}^{k-1} r_i x^{p^{k+i}+p^i}$, where p is an odd prime, s and k are positive integers, $n = 2k$, $\gcd(k+s, n) = \gcd(k+s, k)$, $b \in \mathbb{F}_{p^n}^*$ is not a square, $c \in \mathbb{F}_{p^n} \setminus \mathbb{F}_{p^k}$, and $r_i \in \mathbb{F}_{p^k}$ for $0 \leq i < k$ [11].

In [6], Bierbrauer presented an alternative expression of the two families of the form $\mathrm{Tr}(x^{q+1}) + \mathrm{Tr}(\beta x^{p^s+1})\omega$ over $\mathbb{F}_{q^2}$, where p is an odd prime, $q = p^m, \mathrm{Tr}(\cdot)$ is the trace function from $\mathbb{F}_{q^2}$ to $\mathbb{F}_q, \omega, \beta \in \mathbb{F}_{q^2}, \mathrm{Tr}(\omega) = 0$ and s is a positive integer such that the following holds:

a) β^{q-1} is not contained in the subgroup of order $(q+1)/\gcd(q+1, p^s+1)$ in $(\mathbb{F}_{q^2}, *)$;

b) there is no $0 \neq a \in \mathbb{F}_{q^2}$, such that $\mathrm{Tr}(a) = 0$ and $a^{p^s} = -a$.

We use this representation in our classification, since it has less parameters than the original Budaghyan-Helleseth functions, and therefore generates a smaller amount of functions that we need to classify.

(U_9) $\mathrm{Tr}(x^2) + G(x^{q^2+1})$ over $\mathbb{F}_{q^{2m}}$, where q is a power of an odd prime $p, m = 2k+1, \mathrm{Tr}(\cdot)$ is the trace function from $\mathbb{F}_{q^{2m}}$ to $\mathbb{F}_{q^m}$, and $G(x) = h(x - x^{q^m})$, where $h \in \mathbb{F}_{q^{2m}}[x]$ is defined as

$$h(x) = \sum_{i=0}^{k} (-1)^i x^{q^{2i}} + \sum_{j=0}^{k-1} (-1)^{k+j} x^{q^{2j+1}}.$$

This family corresponds to Bierbrauer's generalization of the semifields discovered by Lunardon, Marion, Polverino and Trombetti over q^6, see [6]. They are referred to as Lunardon-Marino-Polverino-Trombetti-Bierbrauer (LMPTB) semifields [35].

(U_{10}) $x^{162} + x^{108} - x^{84} + x^2$ over $\mathbb{F}_{3^5}$ discovered by Coulter R.S. and Kosick P. (CK[1]) [20] [40].

(U_{11}) The function over $\mathbb{F}_{5^5}$ given by the form

$$f(x) = L(t^2(x)) + D(t(x)) + \frac{1}{2}x^2,$$

where $L(x) = x^{125} + x^{25} + 2x^5 + 3x$ and $D(x) = 0$, or $L(x) = 2x^{25} + x^5$ and $D(x) = 2x^{130} + 2x^{26}$, and $t(x) = x^5 - x$, see [20]. Discovered by Coulter R.S. and Kosick P. (CK[2]).

Families of planar functions given by their corresponding semifields:

(S_{12}) The functions over $\mathbb{F}_{p^{2m}}$, corresponding to the Zhou-Pott (ZP) semifields. Defined as follows: Let m, k be positive integers, such that $\frac{m}{gcd(m,k)}$ is odd. Define $x \circ_k y = x^{p^k} y + y^{p^k} x$. For elements $(a, b), (c, d) \in \mathbb{F}_{p^{2m}}$, define a binary operation $\star$ as follows,

$$(a, b) \star (c, d) = (a \circ_k c + \alpha(b \circ_k d)^\sigma, ad + bc),$$

where α is a nonsquare in $\mathbb{F}_{p^m}$ and σ is an automorphism of $\mathbb{F}_{p^m}$. Then, $(F_{p^{2m}}, +, \star)$ is a presemifield [42].

(S_{13}) The functions over $\mathbb{F}_{p^{2k}}$, corresponding to the Dickson semifields [24].

For $p = 3$, there are some more families as follows:

(S_{14}) The functions over $\mathbb{F}_{3^{2k}}$, corresponding to the Cohen-Ganley (CG) semifields [16].

(S_{15}) The functions over $\mathbb{F}_{3^{2k}}$, with k odd, corresponding to the Ganley semifield [29].

(S_{16}) The function $x^2 + x^{90}$ over $\mathbb{F}_{3^5}$, corresponding to the At-Cohen-Weng (ACW) semifield [3].

(S_{17}) The function over $\mathbb{F}_{3^8}$, corresponding to the Coulter-Henderson-Kosick (CHK) semifields [19].

(S_{18}) The function over $\mathbb{F}_{3^{10}}$, corresponding to the Penttila-Wiliams semifield [36].

One of the contributions of this thesis is the construction of an infinite family of planar quadrinomials, as described in Chapter 5. While we can observe through equivalence tests that the family intersects the Zhou-Pott family for $n = 4$ and $n = 8$ (as well as the sporadic instance (S_{17}) in the case of $n = 8$), it does provide functions of a significantly simpler univariate representation than both (S_{17}) and Zhou-Pott (both of which consist of functions having many terms in their univariate representation). We also expect that our family can be generalized further and may then provide other CCZ-classes; we leave this as a problem for future work. Whether the family is new in general (for dimensions greater than 8) is left as a problem for future work. In any case, we have resolved the problem of classifying (S_{17}) into an infinite family which had been open since 2007. We note that the classification from [39] claims that (S_{17}) is not part of any of the currently known infinite families of planar functions. While this may be true, we have shown that (S_{17}) is CCZ-equivalent to a representative from the Zhou-Pott family.

We have also found seven new instances (up to CCZ-equivalence) of quadratic planar functions over $\mathbb{F}_{3^6}$. More details, including polynomial representations of these instances, are given in Section 4.2.

Chapter 3

Classification of the known planar functions

One of the goals of this thesis is to classify the planar functions obtained from the known infinite families up to CCZ-equivalence. We are able to present CCZ-inequivalent representatives that cover all of the known infinite families for all finite fields $\mathbb{F}_{3^n}$ for n up to 8. The tables of representatives that we give can then be used by researchers to easily check whether or not a planar function that they have newly discovered is equivalent to one of the known planar functions, while saving a lot of time. Without tables like these one would have to generate all planar functions from the families and then classify them up to equivalence, or alternatively compare the newly found planar functions for equivalence against every single instance from the known families, which can be quite many. For instance, even just generating the list of functions for dimension 8 takes more than 10 days of computation.

Planar functions in characteristic 3 from the known families have been classified up to dimension 6 in [39] in 2010. The same year a new sporadic planar function over $\mathbb{F}_{3^5}$ was discovered in [20]. Therefore, the table for $\mathbb{F}_{3^5}$ needed to be updated, and we have done so. In addition, we have classified the planar functions from the known families in dimension 7 and 8. The results are presented in table 3.2. This has previously not been possible to do because there was no efficient enough method for checking CCZ-equivalence between planar functions for dimensions 7 and above. Until recently, the only algorithm for testing equivalence was the code isomorphism test. This is quite slow in dimension 7, and is frequently impossible to use for dimension 8 due to its heavy memory consumption. For classifying the functions for $n = 7, 8$, we use the recently proposed algorithm from [31], which is significantly faster and much less memory intensive than the code isomorphism test. For comparison, testing a single pair of functions for equivalence in dimension 7 using the code isomorphism test takes around 3 hours; using the algorithm from [31], it takes about 1 minute. Even using this faster method,

classifying planar functions is a very time-consuming process: in dimension 8, this took just over 25 days to complete.

3.1 Representatives from the known infinite families and sporadic instances

In this section, we present our computational results over $\mathbb{F}_{3^n}$ for $3 \leq n \leq 8$. The following tables list CCZ-inequivalent representatives covering the CCZ-classes of all the known infinite families and sporadic instances of planar functions. Table 3.1 covers dimensions 3 through 6, while Table 3.2 contains the results for dimensions 7 and 8. We have also calculated invariants for each, and we give them in separate tables. Table 3.3 contains invariants for the representatives over dimension 3 through 6, while Table 3.4 covers the invariants for dimension 7 and 8. Table 3.5 gives the orbit representatives for the functions for dimension 7 and 8.

The number of functions that needed to be classified is quite large; for instance, for $n = 8$, we needed to classify more than 20 000 planar functions. Comparing functions for CCZ-equivalence over $\mathbb{F}_{3^8}$ is a very time-consuming process; in order to appreciate this, we remark that the finite field has $3^8 = 6561$ elements. As pointed out in [31], a test for equivalence over $\mathbb{F}_{3^8}$ can take around 10 minutes in the positive case; furthermore, such tests were impossible to conduct using the code isomorphism test due to lack of memory. Creating these tables has taken almost 60 days, just in computation time. In addition to the time spent on computation, a significant amount of time was spent implementing the code needed to run the computations. By creating these tables, we wish to help researchers with further work in this field by saving them the effort of having to classify planar functions from the known families up to CCZ-equivalence.

We also compute invariants such as the order of the auotmorphism group described in [39], and the sizes of the nuclei of the associated semifield in the case of quadratic functions. In order to compute the nuclei efficiently, we use the following approach. It has been stated in e.g. [34] that the nuclei of a semifield are always finite fields. This is, in fact, only true if the identity $1 \in \mathbb{F}_{p^n}$ of the finite field is the identity of the semifield as well. If the semifield has a different identity element, then the nuclei are mutliplicative cosets of fields obtained by multiplying the elements of a subfield by this identity element. In the following, we assume that we have accounted for this multiplication by the semifield's identity element, and treat the nuclei as subfields.

In our approach, we test whether a given subfield $\mathbb{F}_{p^m}$ of $\mathbb{F}_{p^n}$ belongs to e.g. the left nucleus by taking some random element from $\mathbb{F}_{p^m}$ that does not belong to any proper subfield of $\mathbb{F}_{p^m}$ and checking whether it belongs to this nucleus from the definition. If not, then the nucleus cannot be $\mathbb{F}_{p^m}$; conversely, if the tested element does belong to the nucleus, then the nucleus must be $\mathbb{F}_{p^m}$. In this way, we are able to compute the nuclei much more quickly then by testing each element of $\mathbb{F}_{p^n}$ for membership in a given nucleus individually.

The automorphism group of the associated linear code from [39] is computed in a straightforward way using the Magma algebra system. Unfortunately, this is only possible to do for $\mathbb{F}_{3^n}$ with $n \leq 6$; for higher dimensions, the memory needed to perform the computation becomes prohibitive.

Another useful computational result that we provide in Table 3.5 are lists of representatives from the right orbits (as defined in [31]) for all the representatives from Table 3.2. In the case of the newly discovered planar functions over $\mathbb{F}_{3^6}$ (described in more detail in Section 4.2), the number of orbits is always 62, and the representatives can be taken to be α^i, where α is a primitive element in $\mathbb{F}_{3^6}$ and i is from

0, 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 23, 24, 26, 28, 29, 30, 31, 32, 33, 35, 37, 38, 39, 40, 46, 47, 48, 49, 51, 53, 55, 56, 57, 58, 60, 69, 71, 73, 74, 76, 78, 80, 91, 92, 94, 96, 98, 101, 114, 119, 121, 137, 139.

We note that orbit representatives for all previously known planar functions for $n \leq 6$ are given in [31], and so together with the results in Table 3.5 and the above paragraph, we now know such representatives for all dimensions up to 8. As explained in [31], knowing these representatives allows us to significantly reduce the computation time for checking equivalence with these functions. In some cases (for functions 7.5, 7.6 and 8.8) we have not listed the representatives from the orbits, but have denoted the partition by the star symbol (*). In the case of $n = 7$, the orbit of each element $x \in \mathbb{F}_{3^n}$ is of the form $\{(\pm x)^{3^k} : 0 \leq k \leq n-1\}$, i.e. it consists of the cyclotomic coset of x and its additive inverse. In the case of $n = 8$, the orbits have a similar form, but the situation is slightly more complicated, as some of the sets as for $n = 7$ belong to a single orbit. Due to lack of space, we do not give the orbit representatives in the table; they are the elements α^i, where α is primitive in $\mathbb{F}_{3^8}$, and i is as follows:

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 62, 63, 64, 65, 66, 67, 68, 69, 70, 72, 73, 75, 76, 78, 79, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 99, 100, 101, 102, 103, 104, 105, 107, 109, 110, 111, 112, 113, 114, 115, 117, 118, 119, 120, 121, 122, 123, 125, 126, 127, 129, 130, 131, 132, 133, 134, 135, 136, 137, 139, 142, 143, 146, 147, 148, 149, 150, 151, 153, 154, 155, 157, 158, 159, 161, 162, 163, 164, 165, 166, 167, 171, 172, 173, 174, 175, 177, 178, 179, 181, 182, 183, 185, 186, 188, 189, 190, 191, 192, 195, 196, 197, 199, 200, 201, 202, 203, 204, 205, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 219, 220, 222, 224, 225, 226, 227, 228, 231, 232, 233, 234, 235, 236, 237, 238, 239, 241, 242, 243, 244, 246, 247, 248, 249, 253, 254, 256, 257, 258, 260, 261, 263, 264, 266, 267, 269, 270, 271, 274, 275, 278, 281, 282, 284, 287, 289, 291, 292, 293, 294, 299, 304, 305, 306, 307, 308, 309, 311, 312, 314, 315, 317, 320, 321, 323, 324, 326, 329, 333, 334, 336, 340, 342, 346, 347, 349, 351, 359, 360, 361, 363, 365, 369, 372, 373,

374, 377, 378, 380, 381, 382, 383, 386, 390, 394, 396, 397, 400, 401, 402, 405, 406, 412, 414, 415, 419, 420, 427, 429, 431, 434, 435, 439, 442, 445, 447, 449, 455, 456, 457, 461, 462, 469, 470, 472, 473, 475, 477, 493, 495, 496, 497, 498, 500, 504, 506, 507, 513, 515, 519, 520, 521, 523, 525, 529, 533, 539, 540, 544, 547, 552, 557, 558, 560, 564, 565, 567, 568, 576, 583, 584, 591, 593, 598, 601, 611, 613, 620, 621, 623, 637, 641, 643, 646, 649, 652, 653, 656, 658, 659, 667, 668, 670, 679, 681, 685, 694, 697, 699, 703, 704, 716, 721, 724, 736, 775, 784, 787, 791, 796, 807, 810, 816, 822, 830, 837, 845, 855, 865, 867, 880, 933, 935, 939, 954, 975, 986, 994, 1080, 1107, 1134, 1194.

Due to time constraints, we were not able to compute the orbits for function 8.10. We will report on them in an upcoming work.

Table 3.1: Representatives from the known infinite families over $\mathbb{F}_{3^n}$ for n from 3 to 6

Dim	**N$^{\mathbf{O}}$**	**Functions**	**Family**
3	3.1	x^2	Finite Field (U_1)
	3.2	x^4	Albert (U_2)
4	4.1	x^2	Finite Field (U_1)
	4.2	x^{14}	CM (U_5)
	4.3	$x^{36} + 2x^{10} + 2x^4$	BH (U_8)
5	5.1	x^2	Finite Field (U_1)
	5.2	x^4	Albert (U_2)
	5.3	x^{10}	Albert (U_2)
	5.4	$x^{10} + x^6 + 2x^2$	CMDY[1] (U_3)
	5.5	$x^{10} + 2x^6 + 2x^2$	CMDY[2] (U_4)
	5.6	x^{14}	CM (U_5)
	5.7	$x^{90} + x^2$	ACW (S_{16})
	5.8	$x^{162} + x^{108} - x^{84} + x^2$	CK[1] (U_{10})
6	6.1	x^2	Finite Field (U_1)
	6.2	x^{10}	Albert (U_2)
	6.3	$x^{162} + x^{84} + \alpha^{58}x^{54} + \alpha^{58}x^{28} + x^6 + \alpha^{531}x^2$	Dickson (S_{13})
	6.4	$\alpha^{75}x^{2214} + x^{756} + \alpha^{205}x^{82} + x^{28}$	BH (U_8)
	6.5	$2x^{270} + x^{246} + 2x^{90} + x^{82} + x^{54} + 2x^{30} + x^{10} + x^2$	LMPTB (U_9)
	6.6	$x^{270} + 2x^{244} + \alpha^{449}x^{162} + \alpha^{449}x^{84} + \alpha^{534}x^{54} + 2x^{36}$ $+\alpha^{534}x^{28} + x^{10} + \alpha^{449}x^6 + \alpha^{279}x^2$	Ganley (S_{15})
	6.7	$x^{486} + x^{252} + \alpha^{561}x^{162} + \alpha^{561}x^{84} + \alpha^{183}x^{54} + \alpha^{183}x^{28}$ $+x^{18} + \alpha^{561}x^6 + \alpha^{209}x^2$	CG (S_{14})
	6.8	x^{122}	CM (U_5)
		$\alpha^{438}x^{486} + \alpha^{180}x^{324} + \alpha^{458}x^{270} + \alpha^{672}x^{252} + \alpha^{622}x^{246}$	
	6.9	$+\alpha^{94}x^{244} + \alpha^{650}x^{162} + \alpha^{441}x^{108} + \alpha^{50}x^{90} + x^{84}$ $+\alpha^{77}x^{82} + \alpha^{328}x^{36} + \alpha^{583}x^{30} + \alpha^{407}x^{28} + \alpha^{178}x^{18}$ $+\alpha^{492}x^{12} + \alpha^{692}x^{10} + \alpha^{78}x^6 + \alpha^{219}x^4 + \alpha^{69}x^2$	ZP (S_{12})
	6.10	$\alpha^{91}x^{30} + x^{10} + x^2$	New
	6.11	$\alpha^{91}x^{486} + x^{10} + x^2$	New

Table 3.2: Representatives from the known infinite families over $\mathbb{F}_{3^n}$ for n from 6 to 8

Dim	**N$^{\mathbf{O}}$**	**Functions**	**Family**
6	6.12	$\alpha^{182}x^{82} + 2x^{10} + \alpha^{91}x^6 + x^2$	New
	6.13	$\alpha^{182}x^{82} + 2x^{10} + \alpha^{273}x^6 + x^2$	New
	6.14	$\alpha^{91}x^{486} + \alpha^{182}x^{90} + 2x^{10} + x^2$	New
	6.15	$\alpha^{273}x^{486} + \alpha^{182}x^{90} + 2x^{10} + x^2$	New
	6.16	$\alpha^{273}x^{246} + \alpha^{182}x^{82} + \alpha^{91}x^6 + x^2$	New
7	7.1	x^2	Finite Field (U_1)
	7.2	x^4	Albert (U_2)
	7.3	x^{10}	Albert (U_2)
	7.4	x^{28}	Albert (U_2)
	7.5	$x^{10} + x^6 + 2x^2$	CMDY[1] (U_3)
	7.6	$x^{10} + 2x^6 + 2x^2$	CMDY[2] (U_4)
	7.7	x^{14}	CM (U_5)
	7.8	x^{122}	CM (U_5)
8	8.1	x^2	Finite Field (U_1)
	8.2	x^{14}	CM (U_5)
	8.3	x^{122}	CM (U_5)
	8.4	x^{1094}	CM (U_5)
	8.5	$\alpha^{3994}x^{244} + \alpha^{5354}x^{84} + 2x^{82}$	BH (U_8)
	8.6	$\alpha^{264}x^{1458} + x^{82}$	Bierbrauer (U_7)
	8.7	$\alpha^{3698}x^{2188} + \alpha^{1058}x^{108} + 2x^{82}$	BH (U_8)
	8.8	$x^{4374} + x^{2430} + x^{2214} + 2x^{2190} + 2x^{1458} + 2x^{810} + x^{486} + 2x^{270}$ $+x^{246} + x^{82} + x^{54} + x^{30} + x^{18} + x^{10} + x^6 + x^2$	CHK (S_{17})
	8.9	$\alpha^{3608}x^{1458} + \alpha^{3608}x^{738} + \alpha^{3810}x^{486} + \alpha^{3810}x^{246} + \alpha^{3413}x^{162}$ $+\alpha^{3413}x^{82} + \alpha^{3608}x^{18} + \alpha^{3810}x^6 + \alpha^{2565}x^2$	CG (S_{14})
	8.10	$\alpha^{164}x^{1458} + \alpha^{164}x^{738} + \alpha^{950}x^{486} + \alpha^{950}x^{246} + \alpha^{616}x^{162}$ $+\alpha^{616}x^{82} + \alpha^{164}x^{18} + \alpha^{950}x^6 + \alpha^{6297}x^2$	CG (S_{16})

Table 3.3: Invariants for the representatives over $\mathbb{F}_{3^n}$ for n from 3 to 6

Dim	**N$^{\mathbf{O}}$**	Left nucleus	Middle nucleus	Right nucleus	Code invariant
3	3.1	3^3	3^3	3^3	4212
	3.2	3	3	3	4121
4	4.1	3^4	3^4	3^4	51840
	4.2	–	–	–	640
	4.3	3	3^2	3	10368
5	5.1	3^5	3^5	3^5	588060
	5.2	3	3	3	588060
	5.3	3	3	3	588060
	5.4	3	3	3	4860
	5.5	3	3	3	4860
	5.6	–	–	–	2420
	5.7	3	3	3	53460
	5.8	3	3	3	4860
6	6.1	3^6	3^6	3^6	6368544
	6.2	3^2	3^2	3^2	6368544
	6.3	3	3^3	3	5832
	6.4	3	3	3	454896
	6.5	3	3	3	17496
	6.6	3	3	3	113724
	6.7	3	3	3	2916
	6.8	–	–	–	8736
	6.9	3	3	3	227448
	6.10	3	3^2	3	17496
	6.11	3	3^2	3	17496
	6.12	3	3^2	3	17496
	6.13	3	3^2	3	17496
	6.14	3	3^2	3	17496
	6.15	3	3^2	3	17496
	6.16	3	3^2	3	17496

Table 3.4: Invariants for the representatives over $\mathbb{F}_{3^n}$ for n from 7 to 8

Dim	$\mathbf{N^O}$	Left nucleus	Middle nucleus	Right nucleus	Code invariant
7	7.1	3^7	3^7	3^7	–
	7.2	3	3	3	–
	7.3	3	3	3	–
	7.4	3	3	3	–
	7.5	3	3	3	–
	7.6	3	3	3	–
	7.7	–	–	–	–
	7.8	–	–	–	–
8	8.1	3^8	3^8	3^8	–
	8.2	–	–	–	–
	8.3	–	–	–	–
	8.4	–	–	–	–
	8.5	3	3	3	–
	8.6	3^2	3^2	3^2	–
	8.7	3	3	3	–
	8.8	3	3^4	3	–
	8.9	3	3^4	3	–
	8.10	3	3^4	3	–

Table 3.5: Right orbit representatives for representatives from all known CCZ-classes of planar function over $\mathbb{F}_{3^n}$ for $n = 7, 8$

Dim	$\mathbf{N^O}$	Number of orbits	Orbit representatives
7	7.1	1	1
	7.2	1	1
	7.3	1	1
	7.4	1	1
	7.5	157	*
	7.6	157	*
	7.7	1	1
	7.8	1	1
8	8.1	1	1
	8.2	2	$1, \alpha$
	8.3	2	$1, \alpha$
	8.4	2	$1, \alpha$
	8.5	6	$1, \alpha, \alpha^2, \alpha^4, \alpha^5, \alpha^{29}$
	8.6	6	$1, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^7$
	8.7	6	$1, \alpha, \alpha^2, \alpha^3, \alpha^6, \alpha^{35}$
	8.8	410	*
	8.9	12	$1, \alpha, \alpha^2, \alpha^4, \alpha^7, \alpha^{10}, \alpha^{13}, \alpha^{14}, \alpha^{17}, \alpha^{20}, \alpha^{41}, \alpha^{43}$
	8.10	?	?

Chapter 4

Expansion search for planar functions

Finding new inequivalent planar functions has proven to be a difficult task despite being investigated (in one form or another) since the beginning of the last century. We can see this from the fact that we know very few instances of inequivalent families. One reason for this is of course the vast search space of (n, m, p)-functions, rendering the chances of finding a new planar function quite small. In theory, we can find sporadic PN instances by exhaustive searches over small subclasses of functions. This can be done in a multitude of ways, but no widely successful method has been found so far. However, in the case of APN functions, a method called expansion search, or polynomial expansion, has been used successfully to find short and elegant instances of APN functions [2]. Because of the success of expansion search, one of the main goals of this thesis was to evaluate how successful this approach would be in the case of planar functions.

The concept of expansion search is as follows; we take an (n, m, p)-function $F(x)$, represented as a univariate polynomial, and then we add a number of different terms to it, going over all possible combinations of coefficients and exponents. If no new planar functions of the form $F(x) + c_1 x^{e_1}$ are obtained after adding one term, then we try adding two terms, i.e. we consider all functions of the form $F(x) + c_1 x^{e_1} + c_2 x^{e_2}$; then we try three terms, four terms, and so forth, until the search becomes too slow. Whenever we reach the point where it is too slow, we can choose to restrict the coefficients to a subfield. The exponents can also be restricted to a smaller set in order to make the search faster.

The method of expansion search has various advantages when it comes to constructing planar functions. One of them is that we always find the shortest and simplest polynomial from its equivalence class first, as we are always adding the smallest number of terms first. This approach is also a conceptually simple one and can be easily parallelized. Furthermore, it can be optimized for example by using equivalence relations (as described in Section 2.4.4).

Possible goals of this approach could be to find completely new planar polynomials (up to equivalence), or to try to find simpler representations (up to equivalence) of already known planar polynomials.

In this thesis, we evaluated the efficiency of this approach, and were able to find seven new instances of planar functions (up to CCZ-equivalence) over $\mathbb{F}_{3^6}$. We were unable to find any new planar functions over $\mathbb{F}_{3^8}$, but we did show that the sporadic instance (S_{17}) over $\mathbb{F}_{3^8}$ is equivalent to a function from the Zhou-Pott family. Furthermore, we obtained a significantly simpler representation of this instance (compared to both (S_{17}) and to Zhou-Pott) which we generalized into an infinite family of planar quadrinomials (described in Chapter 5).

4.1 Implementation

The expansion search was implemented in Magma V2.24-3 [8] since the Magma programming language makes it easy to work with functions in univariate polynomial form.

Our implementation was continuously improved and modified based on the results that we observed. For instance, in the beginning we searched through all possible choices of coefficients in $\mathbb{F}_{3^n}$ and all possible exponents up to $3^n - 1$. Obviously, this made the searches quite slow. However, we observed that most of the functions that we found are quadratic, and so when conducting searches with more terms, we restricted the exponents to quadratic ones, i.e. of the form $3^i + 3^j$, with $0 \leq i, j \leq n - 1$. We note that the Coulter-Matthews planar functions, i.e. (U_5), are the only known instance of planar functions that are not quadratic, and finding other such cases has been an open problem for quite some time. This suggests that the chances of finding non-quadratic planar functions are very low. One can also use the case of APN functions over finite fields of even characteristic as a comparison, where we currently know thousands of CCZ-inequivalent APN instances, but only one instance that is not equivalent to a monomial or a quadratic function; computational searches for APN functions are also typically restricted to the quadratic case.

A further advantage that this restriction gives us, is that we only construct functions with even exponents (since $3^i + 3^j$ is even for any choice of i and j). We know that a quadratic function $F(x)$ with zero constant term and even exponents is planar if and only if it is 2-to-1, i.e. if $F(0) = 0$, and every non-zero element in the image set of F has precisely 2 pre-images [40]. This allows us to check whether a function of this form is planar much more quickly than in the general case. More precisely, we only have to count the number of distinct images of F; this operation is linear in the size of the field. On the other hand, checking the definition by verifying that all derivatives of F are permutations is quadratic in the size of the field.

This also allows us to use the much more efficient algorithm from [31] for classifying the functions up to CCZ-equivalence. The reason for this is that the algorithm only applies to functions having even exponents. When running the algorithm, we make use of the orbit representatives that we have computed in Table 3.5 to reduce the running times (see [31] for details).

Furthermore, we use the equivalence trick described in Section 2.4.4 to reduce the number of coefficients that we have to consider for the first term that we add. This not only makes the search faster but also reduces the number of functions that have to be classified.

An additional trick that we can use to quickly eliminate a large number of functions CCZ-equivalent to monomials is the following. We know that any planar function with exponents in a single cyclotomic coset must be equivalent to a monomial. For instance, if all the exponents of F are in the coset of 2, then F is equivalent to x^2, and can thus be removed from consideration. As an example of the efficiency of this approach, in one of our searches in dimension 8, we ended up with more than 300 000 planar function to classify. By removing the ones whose exponents lie in a single coset, we were able to reduce the number of functions to around 30 000.

Even with all of these optimizations, the search eventually gets too slow. When it does, we restrict the set of possible coefficients to a subfield, which allows us to consider more terms.

The results of our searches are summarized in the next section.

4.2 Computational results

The following tables summarize the expansion searches that we conducted, and the times necessary to find and classify all the planar functions in each case. We have mostly concentrated on dimension 8 since prior to the introduction of [31], comparing functions for equivalence over $\mathbb{F}_{3^8}$ was impossible, and so we naturally could expect higher chances of finding new functions there. We have also run some searches for lower dimensions, such as 5, in order to experimentally gauge the efficiency of polynomial expansion. In particular, we note that planar functions over $\mathbb{F}_{3^n}$ with coefficients in the prime field have been completely classified up to $n = 7$ [22], and since 5 and 7 are prime numbers, searches over $\mathbb{F}_{3^5}$ and $\mathbb{F}_{3^7}$ cannot be restricted to any other subfield. This is why we do not carry out any searches for $n = 7$, and concentrate on $n = 8$ instead.

We also run some searches for $n = 6$ and, somewhat surprisingly, find several instances of

planar functions that appear to be inequivalent to all the known representatives from Table 3.1. This is especially curious since we have run many more and more extensive searches for $n = 8$, and found no new planar functions there. Our results are summarized below.

The entries in each of the following tables give the total time (in hours) necessary for conducting the search and classifying the functions. In the case when a cell is labeled with "N", this means that no planar functions were found. Entries labeled with a star symbol ($*$) give the time of the expansion searches that we were not able to classify (due to lack of time). Entries marked with "-" means that no expansion searches where run for the given function with the given number of terms.

The cells in the tables below are given different colors depending on the restrictions used for the coefficients in the expansion. For some expansion searches we had to restrict the coefficients to a certain subfield, in order to obtain the results within a reasonable amount of time. For example, when we tried to run expansion search on x^2 in dimension 8 with two terms and coefficients from $\mathbb{F}_{3^8}$, just the expansion search took 269.3 hours and the classification of the expansion results ran for more than 50 days without being close to finishing.

We also recall that the exponents used in the expansion search are always quadratic.

In the tables showing the results for dimension 6 and 8, cells are colored as follows:

- Results in cells marked with green represent expansion searches where the coefficients were from the field $\mathbb{F}_{3^6}$ for dimension 6 and $\mathbb{F}_{3^8}$ for dimension 8.
- Results in cells marked with orange represent expansion searches where the coefficients were from the field $\mathbb{F}_{3^3}$ for dimension 6 and $\mathbb{F}_{3^4}$ for dimension 8.
- Results in cells marked with yellow represent expansion searches where the coefficients were from the field $\mathbb{F}_{3^2}$.
- Results in cells marked with blue represent expansion searches where the coefficients were from the field $\mathbb{F}_3$.

Table 4.1: Expansion search and classification time for dimension 6 (time in hours)

Dim 6	One term	Two terms	Three terms	Four terms	Five terms
x^2	0	49	0.9	52.7	-
x^{10}	0	55	11.2	27.6	0.4

Table 4.2: Expansion search and classification time for dimension 8 (time in hours)

Dim 8	One term	Two terms	Three terms	Four terms	Five terms
x^2	8.8	24.8	79.9	–	–
x^4	N	4.9	17.7	642.2	73.6
x^{28}	N	682.4	7.2*	640.5*	–
x^{82}	17.4	266.1	901.4*	757.2	87.4

The following subsections describe in more detail the functions that we have observed in dimension 6 and 8.

4.2.1 Dimension 6

In dimension 6, we find seven new CCZ-classes of planar functions. All of these can be obtained by expanding x^2 and x^{10} by two or three terms with coefficients in $\mathbb{F}_{3^2}$. Two of the classes can be represented by trinomials, namely

$$f_1(x) = \alpha^{91}x^{30} + x^{10} + x^2$$

and

$$f_2(x) = \alpha^{91}x^{486} + x^{10} + x^2,$$

where α is a primitive element of $\mathbb{F}_{3^6}$. Note also that α^{91} is primitive in $\mathbb{F}_{3^2}$, and so all the coefficients of these representations lie in the subfield $\mathbb{F}_{3^2}$.

The remaining five classes do not appear to have a trinomial representation, but can be expressed using quadrinomials with coefficients in $\mathbb{F}_{3^3}$. These are:

$$\begin{aligned}
f_3(x) &= \alpha^{182}x^{82} + 2x^{10} + \alpha^{91}x^6 + x^2,\\
f_4(x) &= \alpha^{182}x^{82} + 2x^{10} + \alpha^{273}x^6 + x^2,\\
f_5(x) &= \alpha^{91}x^{486} + \alpha^{182}x^{90} + 2x^{10} + x^2,\\
f_6(x) &= \alpha^{273}x^{486} + \alpha^{182}x^{90} + 2x^{10} + x^2,\\
f_7(x) &= \alpha^{273}x^{246} + \alpha^{182}x^{82} + \alpha^{91}x^6 + x^2.
\end{aligned}$$

We have verified that these functions are inequivalent to each other both using the duplicate equivalence algorithm from [31], as well as the code isomorphism test.

All of the seven classes represented by these functions have a middle nucleus of size 9, and a left and right nucleus of size 3, and all of them have exactly 62 right orbits. The orbit representatives are given above around the start of Section 3.1.

Furthermore, the order of the automorphism of the associated code is equal to 17496 for all seven classes. We note that the only function from the previously known planar instances to have this value is 6.5; however, that function has a middle nucleus of size 3, whereas the aforementioned functions have a middle nucleus of size 9, and so their inequivalence to the known planar functions can also be established based on invariants alone.

4.2.2 Dimension 8

Throughout all of our searches in $\mathbb{F}_{3^8}$, we were not able to find any new planar functions (up to equivalence). However, we did find the planar quadrinomial

$$x^{246} + x^{82} + 2x^6 + x^2 \tag{4.1}$$

which is equivalent to the sporadic instance (S_{17}) with univariate form

$$x^{4374} + x^{2430} + x^{2214} + 2x^{2190} + 2x^{1458} + 2x^{810} + x^{486} + \\ 2x^{270} + x^{246} + x^{82} + x^{54} + x^{30} + x^{18} + x^{10} + x^6 + x^2.$$

We can also observe that the quadrinomial is CCZ-equivalent to an instance from the Zhou-Pott family; nonetheless, the latter has a significantly more complicated univariate representation.

In the next chapter, we describe how we generalize this function into an infinite family of planar quadrinomials. While this family intersects the Zhou-Pott functions (up to equivalence) in dimensions 4 and 8, it is not clear at present whether this will still be the case for higher dimensions. We also expect that our construction may be further generalized, and so might provide instances of functions whose CCZ-classes lie outside the Zhou-Pott family.

Chapter 5

New infinite family of planar functions

In the following, we define a family of quadrinomials which contains the one that we found in $\mathbb{F}_{3^8}$, i.e. (4.1), as a special case, and we prove their planarity.

Recall that the derivative of $F : \mathbb{F}_{3^n} \to \mathbb{F}_{3^n}$ in direction $a \in \mathbb{F}_{3^n}$ can be defined as

$$\Delta_a F(x) = F(a + x) - F(a) - F(x).$$

For convenience, we define the conjugate of an element $x \in \mathbb{F}_{3^n}$ for $n = 2m$ as $\overline{x} = x^{3^m}$.

The proof of planarity is contained in the following theorem.

Theorem 1. *Let $n = 2m = 4k$ for some $k \in \mathbb{N}$, and let us denote $q = 3^m$. Then the function*

$$F(x) = x^2 - (x^2)^3 + x^{q+1} + (x^{q+1})^3 \tag{5.1}$$

is planar over $\mathbb{F}_{3^n}$.

Proof. Let $a \in \mathbb{F}_{3^n}$. Then the derivative of F has the form

$$\Delta_a F(x) = A - A^3 + B + B^3, \tag{5.2}$$

where $A = (x + a)^2 - x^2 - a^2 = 2ax = -ax$ and $B = (x + a)^q - x^q - a^q = xa^q + x^q a = x\overline{a} + \overline{x}a$. Since F is quadratic, it suffices to show that $\Delta_a F(x) = 0$ implies $a = 0$ or $x = 0$.

Since $\overline{B} = \overline{x\overline{a} + \overline{x}a} = \overline{x}a + x\overline{a} = B$, we see that $B \in \mathbb{F}_q$. Since $A - A^3 + B + B^3 = 0$, this implies that $A - A^3 \in \mathbb{F}_q$ as well, i.e. $A - A^3 = \delta$ for some $\delta \in \mathbb{F}_q$. This, however, implies that $A \in \mathbb{F}_q$, which can be seen as follows: the absolute trace of the left-hand side is $\mathrm{Tr}(A - A^3) = 0$, so the same must be true for the right-hand side as well, i.e. we must have $\mathrm{Tr}(\delta) = 0$. This is

satisfied by precisely one third of the elements $\delta \in \mathbb{F}_{3^m}$. Furthermore, $x \mapsto x - x^3$ is a 3-to-1 function over $\mathbb{F}_q$, and any element $A \in \mathbb{F}_q$ satisfies $A - A^3 \in \mathbb{F}_q$. This means that the elements $A \in \mathbb{F}_q$ account for all possible solutions of $A - A^3 \in \mathbb{F}_q$, and therefore we must necessarily have $A \in \mathbb{F}_q$. This then means that $A = -ax \in \mathbb{F}_q$, so that also $ax \in \mathbb{F}_q$.

By substituting $A = -ax$ and $B = a\overline{x} + \overline{a}x$ in (5.2), we get

$$-ax + a^3x^3 + a\overline{x} + \overline{a}x + a^3\overline{x}^3 + \overline{a}^3x^3 = 0;$$

regrouping the terms, we obtain

$$(\overline{a} - a)x + (a^3 + \overline{a}^3)x^3 + a\overline{x} + a^3\overline{x}^3 = 0. \tag{5.3}$$

The conjugate of (5.3) is

$$(a - \overline{a})\overline{x} + (a^3 + \overline{a}^3)\overline{x}^3 + \overline{a}x + \overline{a}^3x^3 = 0. \tag{5.4}$$

Adding (5.3) to (5.4), we obtain

$$(-a - \overline{a})x + (a^3 - \overline{a}^3)x^3 + (-a - \overline{a})\overline{x} + (-a^3 + \overline{a}^3)\overline{x}^3 = 0. \tag{5.5}$$

This becomes

$$(a + \overline{a})(x + \overline{x}) = ((a - \overline{a})(x - \overline{x}))^3. \tag{5.6}$$

Since $ax \in \mathbb{F}_q$, we can write $ax = \varepsilon$ for some $\varepsilon \in \mathbb{F}_q$. Then we have $a = \varepsilon/x$. Substituting this into (5.6), we get

$$\left(\frac{\varepsilon}{x} + \frac{\varepsilon}{\overline{x}}\right)(x + \overline{x}) = \left(\left(\frac{\varepsilon}{x} - \frac{\varepsilon}{\overline{x}}\right)(x - \overline{x})^3\right),$$

which becomes

$$\frac{\varepsilon(x + \overline{x})^2}{x\overline{x}} = -\left(\frac{\varepsilon(x - \overline{x})^2}{x\overline{x}}\right)^3.$$

This simplifies to

$$(x + \overline{x})^2 = -\frac{\varepsilon^2(\overline{x} - x)^6}{(x\overline{x})^2}.$$

Substituting ax for ε in this, we arrive at

$$-a^2 = \left(\frac{(x + \overline{x})x\overline{x}}{x(\overline{x} - x)^3}\right)^2,$$

that is,

$$-a^2 = \left(\frac{(x + \overline{x})\overline{x}}{(\overline{x} - x)^3}\right)^2.$$

We have an equation of the form $-a^2 = X^2$ for some expression X, i.e. $a^2 = -X^2$. The solutions to this equation are $\pm\omega X$, where ω is a square root of -1, i.e. $\omega^2 = -1$. We can take $\omega = \alpha^{(3^n-1)/4}$ since then we see that $\omega^2 = \alpha^{(3^n-1)/2}$, and this must be -1 since its square is $(\alpha^{(3^n-1)/2})^2 = \alpha^{3^n-1} = 1$, and since clearly $\alpha^{(3^n-1)/2} \neq 1$.

In other words, we have

$$a = \pm\omega\frac{(x+\overline{x})\overline{x}}{(\overline{x}-x)^3}.$$

Multiplying both sides by x, we get

$$ax = \pm\omega\frac{(x+\overline{x})x\overline{x}}{(\overline{x}-x)^3}.$$

Since $ax \in \mathbb{F}_q$, we must have

$$\omega\frac{(x+\overline{x})x\overline{x}}{(\overline{x}-x)^3} \in \mathbb{F}_q;$$

and since $x+\overline{x}$ and $x\overline{x}$ are in $\mathbb{F}_q$ for any x, this is equivalent to

$$\frac{\omega}{(\overline{x}-x)^3} \in \mathbb{F}_q. \tag{5.7}$$

Since the dimension n is doubly even, i.e. $n = 2m = 4k$, we have $\overline{\omega} = \omega$. This is because $\overline{\omega} = \omega^q = \omega \cdot \omega^{q-1} = \omega \cdot (-1)^{(q-1)/2}$; thus, $\overline{\omega}$ is either ω or $-\omega$. When $q = 3^m$ for even m, we can see by induction on m that $(q-1)/2$ is even, and so $\overline{\omega} = \omega$. Thus, (5.7) is equivalent to

$$\frac{\omega}{(\overline{x}-x)^3} = \frac{\omega}{(x-\overline{x})^3},$$

that is, $\overline{x} - x = x - \overline{x}$ (since $x \mapsto x^3$ is a permutation of $\mathbb{F}_{3^n}$), i.e. $x = \overline{x}$, i.e. $x \in \mathbb{F}_q$. Since $ax \in \mathbb{F}_q$, this also means that $a \in \mathbb{F}_q$ as well. Substituting $\overline{x} = x$ and $\overline{a} = a$ in (5.3), we have

$$-a^3x^3 + ax + a^3x^3 = 0,$$

that is

$$ax = 0,$$

implying that either $a = 0$ or $x = 0$. Thus, for $a \neq 0$, $\Delta_a F(x) = 0$ only has $x = 0$ as a solution, and hence F is planar. □

It only remains to show that for $n = 8$, the function $F(x)$ above is CCZ-equivalent to the sporadic instance $S_{18}(x)$ from [19]. Using the algorithm from [31], we can see that taking

$$\begin{aligned} L_1(x) =& \alpha^{3936}x^{2187} + \alpha^{5084}x^{729} + \alpha^{2132}x^{243} + \alpha^{1804}x^{81} + \alpha^{820}x^{27} + \\ & \alpha^{1804}x^9 + \alpha^{5412}x^3 + \alpha^{5084}x \end{aligned}$$

and

$$L_2(x) = \alpha^{3485}x^{729} + \alpha^{3567}x^{81} + \alpha^{6273}x^9 + \alpha^{3567}x$$

where α is a primitive element of $\mathbb{F}_{3^8}$, we have $L_1 \circ F \circ L_2 = S_{18}$, so that F and (S_{17}) are indeed equivalent.

Unfortunately, this family is not new (at least for $n \leq 8$) up to equivalence, since we can observe that for both $n = 4$ and $n = 8$, its instances are equivalent to ones from the Zhou-Pott family. Nonetheless, (S_{17}) was listed as a sporadic instance in [37], and so we have resolved the problem of classifying it into an infinite family which had been open since its introduction in 2007.

More precisely, we can show that the instance from our quadrinomial family is linear-equivalent to

$$G(x) = \alpha^{602}x^{2268} + \alpha^{3882}x^{2188} + \alpha^{411}x^{162} + \alpha^{6042}x^{108} + \alpha^{2542}x^{82} + \alpha^{2762}x^{28} + \alpha^{491}x^2, \quad (5.8)$$

which is a representative from the Zhou-Pott family that we obtain for $\sigma(x) = x^3$ and $k = 4$. Taking

$$L_1'(x) = \alpha^{3478}y^{2187} + \alpha^{1518}y^{729} + \alpha^{1518}y^{243} + \alpha^{1518}y^{81} + \alpha^{1716}y^{27} + \alpha^{4798}y^9 + \alpha^{4798}y^3 + \alpha^{4798}y$$

and

$$L_2'(x) = \alpha^{779}y^{243} + \alpha^{6233}y^{81} + \alpha^{4059}y^3 + \alpha^{3033}y,$$

we can see that $L_1'(x) \circ F \circ L_2'(x) = G(x)$.

We note that the Zhou-Pott instance (5.8) is significantly simpler than the sporadic instance (S_{17}) (the former having only 7 terms, while the latter has 16 terms). The representation coming from our quadrinomial family is simpler still, since not only does it have four terms, but all of the coefficients lie in the prime field $\mathbb{F}_3$.

We also remark that the family presented in the above theorem may very well be susceptible to further generalizations; for instance, it could potentially be adapted to charactristics other than 3, or to a wider family of functions that may include other CCZ-classes. Whether our family continues to intersect Zhou-Pott in dimensions higher than 8 is anothr question that we leave for future work.

Chapter 6

Conclusion and future work

We provided a classification up to CCZ-equivalence of all known planar functions (from both infinite families and sporadic instances) over $\mathbb{F}_{3^n}$ for $n \leq 8$. We computed values of the known invariants for each equivalence class, and compiled all the data into tables that we hope will facilitate searches for new planar functions undertaken by researchers in the future.

We ran numerous expansion searches for planar functions in order to test their efficiency, and found seven new instances (up to CCZ-equivalence) of planar functions over $\mathbb{F}_{3^6}$. We also concluded that no new planar functions (up to equivalence) can be found in many cases. We described tricks and methods that can be used to speed up the search and reduce the number of resulting functions. We found a shorter representation for the equivalence class of a known sporadic instance over $\mathbb{F}_{3^8}$, and generalized it into an infinite family of planar quadrinomials, as well as shown that it is equivalent to an instance from the Zhou-Pott functions in small dimensions. The question of whether this quadrinomial family is new in general is left open. We note that the problem of classifying (S_{17}) into an infinite family had been open since its introduction by Coulter et al. in 2007. In [39], it was pointed out that (S_{17}) was not contained in any of the known infinite families of planar functions; nonetheless, we have shown that it is equivalent to a representative from the Zhou-Pott family (S_{12}).

There are many directions left for future work. Using newly developed approaches such as [31], it might be possible to extend the classification of the known functions to higher dimensions, such as 9 or 10, and to characteristics other than 3. More expansion searches can be run, both for dimensions that have not been covered by our searches (including the prime dimensions 5 and 7, as well as dimensions greater than 8), as well as for more terms and a broader range of coefficients and exponents in the dimensions that we investigated.

It might be worth considering the structure of the newly introduced family, and considering

whether it can be generalized to obtain further constructions. The properties of this family and its exact relationship to e.g. the Zhou-Pott family, still need to be investigated.

Bibliography

[1] Abraham Adrian Albert. Finite division algebras and finite planes. In *Proc. Sympos. Appl. Math*, volume 10, page 53, 1960.

[2] Maren Hestad Aleksandersen, Lilya Budaghyan, and Nikolay Stoyanov Kaleyski. Searching for APN functions by polynomial expansion. In *Norsk IKT-konferanse for forskning og utdanning*, number 3, 2021.

[3] Nuray At and Stephen D Cohen. A new tool for assurance of perfect nonlinearity. In *International Conference on Sequences and Their Applications*, pages 415–419. Springer, 2008.

[4] Jürgen Bierbrauer. New commutative semifields and their nuclei. In *International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes*, pages 179–185. Springer, 2009.

[5] Jürgen Bierbrauer. New semifields, PN and APN functions. *Designs, Codes and Cryptography*, 54(3):189–200, 2010.

[6] Jürgen Bierbrauer. Commutative semifields from projection mappings. *Designs, Codes and Cryptography*, 61(2):187–196, 2011.

[7] Eli Biham and Adi Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of CRYPTOLOGY*, 4(1):3–72, 1991.

[8] Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system I: The user language. *Journal of Symbolic Computation*, 24(3-4):235–265, 1997.

[9] Lilya Budaghyan. *Construction and analysis of cryptographic functions*. Springer, 2015.

[10] Lilya Budaghyan, Claude Carlet, and Gregor Leander. Two classes of quadratic APN binomials inequivalent to power functions. *IEEE Transactions on Information Theory*, 54(9):4218–4229, 2008.

[11] Lilya Budaghyan and Tor Helleseth. New perfect nonlinear multinomials over $\mathbb{F}_{p^{2k}}$ for any odd prime p. In *International Conference on Sequences and Their Applications*, pages 403–414. Springer, 2008.

[12] Lilya Budaghyan and Tor Helleseth. New commutative semifields defined by new PN multinomials. *Cryptography and communications*, 3(1):1–16, 2011.

[13] Lilya Budaghyan and Tor Helleseth. On isotopisms of commutative presemifields and CCZ-equivalence of functions. *International Journal of Foundations of Computer Science*, 22(06):1243–1258, 2011.

[14] Anne Canteaut, Alain Couvreur, and Léo Perrin. Recovering or testing extended-affine equivalence. *IEEE Transactions on Information Theory*, 2022.

[15] Claude Carlet, Pascale Charpin, and Victor Zinoviev. Codes, bent functions and permutations suitable for DES-like cryptosystems. *Designs, Codes and Cryptography*, 15(2):125–156, 1998.

[16] Stephen D Cohen and Michael J Ganley. Commutative semifields, two dimensional over their middle nuclei. *Journal of Algebra*, 75(2):373–385, 1982.

[17] Robert S Coulter. Planar functions and related topics in finite fields. *Bulletin of the Australian Mathematical Society*, 59(1):173–174, 1999.

[18] Robert S Coulter and Marie Henderson. Commutative presemifields and semifields. *Advances in Mathematics*, 217(1):282–304, 2008.

[19] Robert S Coulter, Marie Henderson, and Pamela Kosick. Planar polynomials for commutative semifields with specified nuclei. *Designs, Codes and Cryptography*, 44(1):275–286, 2007.

[20] Robert S Coulter and Pamela Kosick. Commutative semifields of order 243 and 3125. *Finite Fields: Theory and Applications, in: Contemp. Math*, 518:129–136, 2010.

[21] Robert S Coulter and Rex W Matthews. Planar functions and planes of Lenz-Barlotti class II. *Designs, Codes and Cryptography*, 10(2):167–184, 1997.

[22] Diana Davidova and Nikolay S Kaleyski. Classification of all DO planar polynomials with prime field coefficients over $\mathbb{F}_{3^n}$ for $n \leq 7$.

[23] Peter Dembowski and Thedore G Ostrom. Planes of order n with collineation groups of order n^2. *Mathematische Zeitschrift*, 103(3):239–258, 1968.

[24] Leonard Eugene Dickson. On commutative linear algebras in which division is always uniquely possible. *Transactions of the American Mathematical Society*, 7(4):514–522, 1906.

[25] Cunsheng Ding and Jin Yuan. A family of skew Hadamard difference sets. *Journal of Combinatorial Theory, Series A*, 113(7):1526–1535, 2006.

[26] Itai Dinur and Adi Shamir. Breaking Grain-128 with dynamic cube attacks. In *International Workshop on Fast Software Encryption*, pages 167–187. Springer, 2011.

[27] Yves Edel and Alexander Pott. On the equivalence of nonlinear functions. In *Enhancing cryptographic primitives with techniques from error correcting codes*, pages 87–103. IOS Press, 2009.

[28] Jaqueline Edit Hind Elliott. *Relative difference sets.* PhD thesis, University of Miami, 1965.

[29] Michael J Ganley. Central weak nucleus semifields. *European Journal of Combinatorics*, 2(4):339–347, 1981.

[30] Tor Helleseth and Daniel Sandberg. Some power mappings with low differential uniformity. *Applicable Algebra in Engineering, Communication and Computing*, 8(5):363–370, 1997.

[31] Ivana Ivkovic and Nikolay Kaleyski. Deciding and reconstructing linear equivalence of uniformly distributed functions. Cryptology ePrint Archive, Paper 2022/666, 2022. https://eprint.iacr.org/2022/666.

[32] Nikolay Kaleyski. Deciding EA-equivalence via invariants. *Cryptography and Communications*, 14(2):271–290, 2022.

[33] Lars R Knudsen. Truncated and higher order differentials. In *International Workshop on Fast Software Encryption*, pages 196–211. Springer, 1994.

[34] Donald Ervin Knuth. *Finite semifields and projective planes.* PhD thesis, California Institute of Technology, 1963.

[35] G Lunardon, G Marino, O Polverino, and R Trombetti. Symplectic spreads and quadric Veroneseans. *Preprint*, 2009.

[36] Tim Penttila and Blair Williams. Ovoids of parabolic spaces. *Geometriae Dedicata*, 82(1):1–19, 2000.

[37] Alexander Pott. Almost perfect and planar functions. *Designs, Codes and Cryptography*, 78(1):141–195, 2016.

[38] Alexander Pott, Kai-Uwe Schmidt, and Yue Zhou. Semifields, relative difference sets, and bent functions. In *Algebraic curves and finite fields*, volume 16, pages 161–178. De Gruyter, 2014.

[39] Alexander Pott and Yue Zhou. Switching construction of planar functions on finite fields. In *International Workshop on the Arithmetic of Finite Fields*, pages 135–150. Springer, 2010.

[40] Guobiao Weng and Xiangyong Zeng. Further results on planar DO functions and commutative semifields. *Designs, Codes and Cryptography*, 63(3):413–423, 2012.

[41] Zhengbang Zha, Gohar M Kyureghyan, and Xueli Wang. Perfect nonlinear binomials and their semifields. *Finite Fields and Their Applications*, 15(2):125–133, 2009.

[42] Yue Zhou and Alexander Pott. A new family of semifields with 2 parameters. *Advances in Mathematics*, 234:43–60, 2013.