CREATIVE COMPONENT
THE GAME SET AND FINITE FIELDS: AN ALGEBRAIC APPROACH

MELISSA MILLS

Abstract. In 1973, the card game SET was invented. This game has many applications to mathematics. In this paper we will answer the question, "What is the maximum number of SET cards that does not contain a SET?" We will do this by considering connections between the game and vector spaces over finite fields.

Contents

1. Introduction 2
2. Combinatorics and SET 4
3. Finite Fields 5
4. Vector Spaces and SET 8
5. Finite Fields and SET 12
6. Finding the SET-less Set 16
7. Proofs About Maximal Caps For the Two Dimensional and Three Dimensional Cases 20
8. The Four Dimensional Case 23
References 28

Date: April 2006.

1. Introduction

The game SET has become popular in many circles, from university math departments, to elementary school math classrooms, to MENSA. SET is a game of visual perception that is very fast paced and addictive. It was invented in 1974 by Marsha Jean Falco. Falco was a geneticist, studying epilepsy in German Shephards. She was representing different genetic qualities on cards using shape, color, shading, and number. She was then searching for patterns in the data, and soon realized that finding these patterns could be a game. Her family played it for years before it was marketed in 1991. This game is so simple that a child can play it, but it raises a lot of mathematical questions and has caught the attention of the mathematical societies worldwide. It won the "Best New Mind Game of 1991" award and was named one of the six best games of 1991 [1].

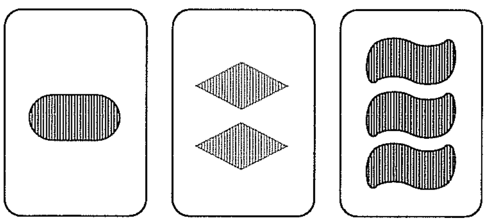
A deck of 81 cards makes up the game SET. Each card is different. Each card has shapes on it and differs with color, number of shapes, kind of shapes, and shading. The possibilities are:

Number:	Two	Three	One
Color:	Red	Green	Purple
Shading:	Open	Striped	Solid
Shape:	Oval	Squiggle	Diamond

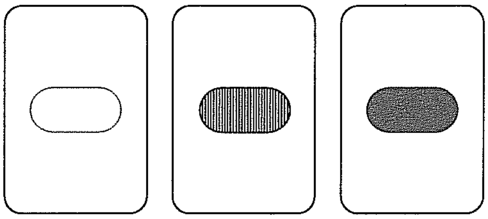
So, there are four characteristics on each card and each aspect has three possible values each. A SET is a collection of three cards given by the following rule:

> **Definition:** Three cards are a SET if, with respect to each characteristic, they are either all alike or all different.

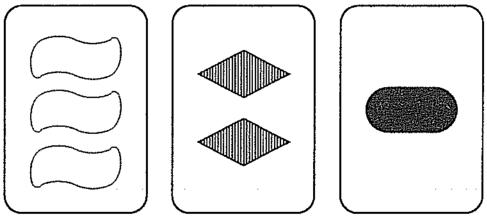
Some examples of sets are given below:



The first collection of three cards is a SET because the numbers are all different, the colors are all different, the shadings are all the same, and the shapes are all different.



This second collection is a SET because the numbers are all the same, the colors are all different, the shapes are all the same, and the shadings are all different.



This collection is a SET because in every aspect the cards are all different.

To play the game, the dealer deals twelve cards, and the players look for SETs. When a set is found, the player says, "SET," and picks it up. Then three more cards are dealt. If the player says "SET" and there actually is no SET, the player loses a point. When you have gone through the entire deck, the game is over and each player counts the number of SETs that they have collected and subtracts any points that were lost. Sometimes there are no sets in the twelve cards that were originally dealt. In this case, three more cards are dealt. If there are no SETs in that 15 cards, three more cards are dealt. It is natural to ask the question, "How many cards must you lay down to guarantee that there is a SET?" This project will deal with that question in multiple ways. The problem has been approached by using applications to vector spaces, lines, and planes. We will attempt to make some progress by imposing an algebraic structure on the cards and using finite field theory.

The main goal of this paper is to answer the question of how many cards are necessary to gurantee a SET. This will be done in Section 8, with a reproduction of a proof by Maclagan and Davis [1]. The second main goal is to show an example of a maximal no-SET by using finite fields. This will be done in Section 6.

Many of the proofs about SET use combinatorics, therefore it is necessary to make some basic combinatorial claims about the deck of SET cards. This will be done in the section Combinatorics and SET. We then give the reader a short refresher on finite field theory, and clarify some of the notation that will be used throughout the paper. The next section describes vector spaces over finite fields in more detail and how these vector spaces can be used to prove things about SET. In the section Finite Fields and SET, we use fields of order 3^n to make claims about the number of elements of a collection of cards that does not contain a SET. Since we have proved that there are a certain number of cards not containing a SET, the reader may be interested in actually seeing an example of a "SET-less set." This section is inspired mainly by Anthony Kable [4]. His idea was that finding this collection of cards with no SET could be done by imposing a field structure on the cards. In this section we show how a collection of cards that doesn't contain a SET can be drawn out of the abstraction of finite field theory. In sections 7 and 8, we give proofs to show that the collections we previously found are actually maximal collections not containing a SET.

2. Combinatorics and SET

Here we will briefly give some of the very basic combinatorial aspects of the deck of SET cards. Since each card has four aspects and three options for each aspect, the number of possible cards is $3 \times 3 \times 3 \times 3 = 3^4 = 81$. So, to find the number of green cards, the green cards have different shapes, shading, and number. So there are $3 \times 3 \times 3 = 3^3 = 27$ cards that are green, and similarly 27 cards with any one fixed quality. Let's try fixing two qualities by looking for the number of cards with green diamonds. The green diamonds could still vary in shading and number, so there are $3 \times 3 = 3^2 = 9$ green diamonds. Similarly, there are 9 cards with any two aspects fixed. Hence if we fix three aspects, then there are only 3 cards, and there is only one card with all four aspects specified. So each card in the deck of 81 cards is unique.

How many SETs are there containing a given card? The strategy we will use is to count the number of SETs with different numbers of aspects fixed. So we have a given card, with a particular color, number, shape and shading. Say we want to make a SET of three cards that had exactly the same color, shape, shading and number as the given card. Since we know that each card is unique, fixing all four aspects would only give us a collection of one card because each SET card is unique. Now, we can start by choosing 3 aspects to be fixed while one aspect is different. Say we want to make a SET of three cards that has the same color, shape, and shading as our original card, but the number has to be different. If three aspects are fixed, there is only one possible SET so there are $1 \times \binom{4}{3}$ SETs of that type. Next we choose 2 aspects to be fixed. Say, for example, we fix color and shape letting number and shading vary. Then there are two possible SETs containing the given card with color and shape fixed, hence there are $2 \times \binom{4}{2}$ SETs of that type. Then we fix only one aspect, allowing three to vary. Say we fix only color. With number, shading, and shape varying, there are four possible SETs There are then 2^2 SETs containing the card with color fixed. There are $2^2 \times \binom{4}{1}$ SETs of that type. Lastly we allow all four aspects to be different. There are 2^3 such SETs containing the given card. Total there are

$$\sum_{k=1}^{3} 2^{k-1} \times \binom{4}{4-k} = 4 + 12 + 16 + 8 = 40$$

SETs containing a given card.

Turning to a much simpler combinatorial argument, if we choose one card, there are 80 cards left. We know that the choice of the second SET card determines a SET, however, the same SET is determined by precisely two choices of a second card. Hence, there are 80/2=40 SETs containing the original card.

Now let's count how many possible SETs there are in the deck. Given any two cards, we can find the third card to make a SET. First we can look at the shapes on the two cards. If those two cards have the same shape on them, then the third card must also have the same shape. If they have different shapes on them, then the third card must have the missing shape. We can continue this process for each aspect, and then we have all four aspects specified for the third card, so there can only be one such card. So, given any two cards in the SET deck, there is exactly one card that completes the SET. To count the total number of possibilities of SETs, we must first choose 2 cards from the 81, and then to make sure that we haven't double counted any SETs, we must divide by 3 choose 2, because those two cards must have come out of a SET of three cards. Therefore there are

$$\binom{81}{2} \bigg/ \binom{3}{2} = 2160$$

possible SETs in the deck.

3. Finite Fields

Since this paper will consider many relationships between the game SET and finite fields, it is necessary first to define some things about fields [2].

Definition 1. *A **field** is a nonempty set $\mathbb{F}$ equipped with two operations (usually written as addition and multiplication) that satisfy the following axioms.*

(1) *$\mathbb{F}$ is a commutative group under addition with additive identity $0_{\mathbb{F}}$.*
(2) *$\mathbb{F}^{\times} = \mathbb{F} \setminus \{0\}$ is a commutative group under multiplication with multiplicative identity $1_{\mathbb{F}}$.*
(3) *For all $a, b, c \in \mathbb{F}$, $a(b + c) = ab + ac$.*

So a field has two operations that have the properties of closure, associativity, commutativity, identity, and inverses. It also has a distributive law across the two operations.

Fields can be infinite or finite. Since we are relating fields to our SET cards, we will be dealing with finite fields. In Abstract Algebra, the concept of finite and infinite fields is crucial. The simplest finite fields are called the prime fields, which can be written $\mathbb{F}_p$, where p is a prime number. The structure of these fields comes from arithmetic modulo p which has the property that any multiple of p is equal to zero. So $\mathbb{F}_p = \{0, 1, 2, \ldots, p - 1\}$, with the usual rules of integer addition and multiplication modulo p.

For this project, the SET cards relate particularly to the finite field of order 3. $\mathbb{F}_3$ has elements $0, 1$, and 2, but we sometimes write 2 as -1 in modulo 3, because $2 = -1 + 3$. We say that any multiple of three is equal to zero in this field. This field is the basis of the structure of the SET cards because each aspect of the SET cards has three options.

Definition 2. *A field F is said to have **characteristic** n if n is the smallest positive integer such that $n1_F = 0_F$. If no such $n \neq 0$ exists, then we say that the field has characteristic zero, and the field is infinite.*

The field $\mathbb{F}_3$ has characteristic 3, because multiples of 3 are equal to zero in arithmetic modulo 3.

Proposition 1. *Every finite field F has characteristic p for some prime p.*

Proof. Since F is finite, F does not have characteristic 0. Assume that F has characteristic $n > 0$. If n were not prime, then there exist positive integers k, t such that $n = kt$ with $k < n$ and $t < n$. Then

$$(k1_F)(t1_F) = (kt)1_F = n1_F = 0_F.$$

But, since a finite field is also an integral domain, either $k1_F = 0_F$ or $t1_F = 0_F$, contradicting that n is the smallest positive integer such that $n1_F = 0_F$. Therefore, n is prime. □

If p is the characteristic of a finite field F, then the field F contains a subfield isomorphic to $\mathbb{F}_p$. Since F contains $\mathbb{F}_p$, it also has the structure of a *vector space* over the field $\mathbb{F}_p$. That is, the usual addition law on F and the multiplication by elements of $\mathbb{F}_p$ satisfy all of the axioms of a vector space. Since F has a finite number of elements, it also has a finite dimension n as a vector space [2]. That means that every element of F can be described by n coordinates chosen from $\mathbb{F}_p$. Therefore, the number of elements of F (also called the *order* of F) is p^n. We shall also refer to the dimension n as the *degree* of F over $\mathbb{F}_p$, written $[F : \mathbb{F}_p]$. This proves the following theorem.

Proposition 2. *A finite field F has order p^n where p is the characteristic of F and $n = [F : \mathbb{F}_p]$.*

This theorem limits the possiblilities for the order of a finite field, because each finite field must be a finite extension of $\mathbb{F}_p$.

Since we know that the only possible orders of finite fields are p^n for primes p, it is necessary to state another theorem about possible finite fields.

Proposition 3. *Two finite fields of the same order are isomorphic* [2].

This theorem tells us that all finite fields of the same order are the same up to an isomorphism. Let's take a moment to talk about what we mean by a finite extension. Take, for example, the field $\mathbb{F}_p$. A finite extension of order 2 would be of the form $\mathbb{F}_p[\alpha]$ where α is a root of a second degree polynomial with coefficients in $\mathbb{F}_p$, with the added condition that $\alpha \notin \mathbb{F}_p$. This extension has order p^2. The next theorem will give us more insight into the structure of this field.

Proposition 4. *All nonzero elements of $\mathbb{F}_{p^n}$ satisfy $x^{p^n-1} = 1$. In fact, the nonzero elements of $\mathbb{F}_{p^n}$ form a cyclic group of order $p^n - 1$.*

Define $F^\times$ to be the field F minus its zero element. Then $\mathbb{F}_{p^2}^\times$ is a cyclic multiplicative group of order $p^2 - 1$. Subgroups of this group are also cyclic groups of the form $\{x^d = x\}$ for $d|p^2$.

The collection $\{x \in \mathbb{F}_{81} | x^9 = x\}$ contains all elements of the field $\mathbb{F}_9$. We can factor the equation $x^9 - x = 0$ in mod 3, and all of the roots of this equation are the elements of $\mathbb{F}_9$. The equation $x^9 - x = 0$ factors into $x(x+1)(x-1)(x^2+1)(x^2+x-1)(x^2-x-1) = 0$, and we can see that the first three roots are the elements of $\mathbb{F}_3 = \{0, 1, -1\}$ and the roots of the three quadratics give us the other elements of the field $\mathbb{F}_9$. We can let α be a root of the irreducible quadratic $x^2 + 1$, and then, according to finite field theory, the other elements of $\mathbb{F}_9$ are of the form $a + b\alpha$ where $a, b \in \mathbb{F}_3$.

Another concept is that of a vector space over a finite field. A n-dimensional vector space over a finite field F contains n-dimensional vectors that have entries from the finite field F. For example, the vector space $\mathbb{F}_3^2$ contains all two dimensional vectors with entries from $\mathbb{F}_3$. In other words, $\mathbb{F}_3^2 = \{(u, v)|u, v \in \mathbb{F}_3\}$.

The finite field theory relates to SET in two different ways. Since there are $81 = 3^4$ cards in a deck, we can relate the deck of cards to $\mathbb{F}_{81}$, or we can view the deck of cards as the vector space $\mathbb{F}_3^4$ and each card would represent a vector. This paper will use both methods to prove things about the game SET.

4. Vector Spaces and SET

Let's begin by approaching the deck of SET cards as a vector space over $\mathbb{F}_3$. We can let each of the aspects be a dimension in this vector space. Each aspect has three options, so we can assign values to each of the options. Say we assign values as such:

Aspect	-1	0	1
Number:	Two	Three	One
Color:	Red	Green	Purple
Shading:	Open	Striped	Solid
Shape:	Oval	Squiggle	Diamond

Then a card that has three green open diamonds can be written as the vector $< 0, 0, -1, 1 >$. Each element belongs to this four dimensional vector space, and can be represented as a vector $< v_1, v_2, v_3, v_4 >$ with $v_i \in \mathbb{F}_3$. Since we have such a representation for the individual cards, it is now natural to use it to determine an alternative definition of a SET.

Proposition 5. *A* SET *can be represented as a collection of three vectors in $\mathbb{F}_3^4$ that add to give zero.*

Proof. Let $u, v, w \in \mathbb{F}_3^4$ represent three cards that make a SET. Then since the cards make a SET, they are either all alike or all different in each aspect. Now we have in $\mathbb{F}_3$ that $1+1+1 = 0$. Also, $(-1)+(-1)+(-1) = 0$, $0+0+0 = 0$, and $(-1)+0+1 = 0$. So we get zero whenever we are adding values that are either all alike or all different. We can re-write our three vectors as $u =< u_1, u_2, u_3, u_4 >, v =< v_1, v_2, v_3, v_4 >, w =< w_1, w_2, w_3, w_4 >$. Now when we add $u, v,$and w, we have $u+v+w =< u_1 + v_1 + w_1, u_2 + v_2 + w_2, u_3 + v_3 + w_3, u_4 + v_4 + w_4 >$, and since u_1, v_1, w_1 are either all alike or all different, $u_1 + v_1 + w_1 = 0$. Similarly, $u_i + v_i + w_i = 0$ for $i \in \{1, 2, 3, 4\}$, therefore $u + v + w = 0$. □

Now we can see that a SET corresponds to a collection of three vectors that add to give you the zero vector. So, each card in the deck represents a point in the vector space $\mathbb{F}_3^4$, and a SET corresponds to an *affine line* in $\mathbb{F}_3^4$. An *affine line* is a fixed vector v_0 added to all scalar multiples of a direction vector v_1. Since we are in $\mathbb{F}_3$, the only scalars are 0,1, and 2. Then an affine line would be the three points:

$$v_0, v_0 + v_1, v_0 + 2v_1$$

Notice that they add to zero. conversely, any three points that add to zero can be written in the above form.

We're using the vector space $\mathbb{F}_3^4$ because we have four aspects with three options each. What if we had five aspects with three options each? Say, for example, we added another aspect like position on the card. The position on the card could be aligned to the left, right, or center. Now each card would be represented as a five dimensional vector, and the deck of cards would represent $\mathbb{F}_3^5$. If we played with only the red cards, then there would only be three aspects left. There would only be shape, number, and shading. Each of the three aspects has three options each, so then the deck would correspond to $\mathbb{F}_3^3$. If we played with only the red ovals, we would have $\mathbb{F}_3^2$.

Notice that we are just changing the dimension of our vectors. We still had three options for each aspect, because we defined a set as a collection of three cards, so we are still working with vectors with elements in $\mathbb{F}_3$. If we changed the cards to have only two options for each aspect, then we would be working with vectors that have elements in $\mathbb{F}_2$. We would then have to change the definition of a SET to be a collection of two cards that are either alike or different in regard to each aspect.

Since each of the vector spaces has entries from $\mathbb{F}_3$, the property that $u + v + w = 0$ for $u, v, w \in \mathbb{F}_3^4$ implies $\{u, v, w\}$ is a SET is still maintained. We define a d-cap to be a subset of $\mathbb{F}_3^d$ that does not contain a SET. In other words, a d-cap is a collection, S, of elements of $\mathbb{F}_3^d$ such that no three elements of S add to give zero. So now our question, "How many cards must you lay down to guarantee that there is a SET?" is related to the question, "What is the maximal size for a 4-cap?" If we can know the maximal size of a 4-cap, then adding one more card would guarantee that there is a SET. This question of a 4-cap was answered in 1971 by Giuseppe Pellegrino [1], three years before the game SET was invented!

Tackling the question of a 4-cap is pretty substantial, so we'll start with dimension 2. This would correspond to playing with only the red ovals. Consider Figure 1.

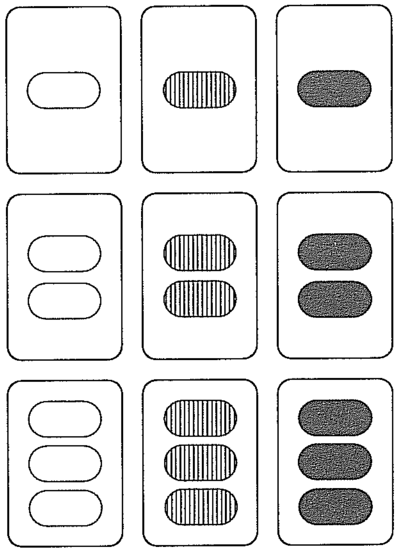


FIGURE 1. A picture of $\mathbb{F}_3^2$ as represented by SET cards.

If we set up the cards like a tic-tac toe board as in the figure, we can see that the SETs corrrespond to the "winning" horizontal, vertical, and diagonal lines. They also correspond to diagonal lines that wrap around to the opposite corner. So, in this two-dimensional space, what is the maximum number of cards that does not contain a SET? Let's use the term no-SET to mean a collection of cards that does not contain a SET.

Proposition 6. *The collection of the four corner cards in Figure 1 is a no-*SET.

Proof. Choose the four corner cards in Figure 1. We claim that any collection of three of these four cards will not be a SET. Notice that two of those cards have one shape and two of the four have three shapes. Any collection of three of these cards will have either two cards with one shape and one card with three shapes, or two cards with three shapes and one card with one shape. But, in order for this collection to be a SET, the number of shapes would either be all alike or all different. so any collection of three of these cards is not a SET. □

There are other collections of four cards that do not contain SETs, for example, if we choose the bottom right, bottom middle, middle left and top left, that collection does not contain a SET. It is a more difficult task to prove that any collection of five cards must contain a SET. This will be tackled in later sections.

We can ask ourselves if these two no-SETs are related to each other, and how they are related. In a vector space, there is a relationship between points called a linear transformation. A linear transformation is a mapping from a vector space V onto another vector space W that satisfies these two conditions:

1. $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$.
2. $T(\alpha v_1) = \alpha T(v_1)$ for all $v \in V$ and $\alpha \in \mathbb{F}_3$.

When the linear transformation is between two vector spaces of the same dimension, it is possible for the linear transformation to be invertible. In our case, we are concerned with linear transformations between the vector space $\mathbb{F}_3^d$ and itself. In any case, it is true that $T(0) = 0$ if T is a linear transformation. An *affine* transformation has the form $T(v) = Av + b$ where A is a $d \times d$ matrix and b is a vector. These do not satisfy the prevoius conditions, and also $T(0) = b$ for an affine transformation T.

So, how do affine transformations affect SETs and no-SETs?

Proposition 7. *An invertible affine transformation from $\mathbb{F}_3^d$ onto itself takes* SETs *to* SETs *and no-*SETs *to no-*SETs.

Proof. Let $T(v) = Av+u$ where $u, v \in \mathbb{F}_3^d$ and A is an $n \times d$ matrix with entries from $\mathbb{F}_3$. Then if for $v_1 \neq v_2 \neq v_3$ we have that $v_1 + v_2 + v_3 = 0$, since the transformation T is a linear map, $T(v_1) + T(v_2) + T(v_3) = Av_1 + u + Av_2 + u + Av_3 + u = A(v_1 + v_2 + v_3) + 3u = A0 + 0 = 0$. Assume by way of contradiction that $T(v_i) = T(v_j)$ for $i \neq j, i, j \in \{1, 2, 3\}$. Then, since T is invertible, T is also injective, and this implies $v_i = v_j$, but this is a contradiction. Hence, $T(v_1) \neq T(v_2) \neq T(v_3)$ and therefore $\{T(v_1), T(v_2), T(v_3)\}$ is a SET.

T^{-1} is also an affine transformation, because $T^{-1}(v) = A^{-1}(v) - A^{-1}(u)$. Thus, the same argument shows that T^{-1} carries SETs to SETs.

If S is a no-SET, suppose $T(S)$ contains a SET $\{v_1, v_2, v_3\}$. Since T^{-1} is an affine transformation $\{T(v_1), T(v_2), T(v_3)\}$ is a SET contained in $T^{-1}(T(S)) = S$. This contradicts the assumption that S is a no-SET. Thus, T carries no-SETs to no-SETs.

Since T^{-1} is also an affine transformation, it carries no-SETs to no-SETs. □

We have seen how affine transformations affect SETs in two dimensions, but what about in n dimensions?

Proposition 8. *In n dimensions, all* SETs *are the same up to an affine transformation* [3].

Proof. A SET is specified by a fixed vector v_0 and a nonzero direction vector v_1. Suppose we have another SET specified by a fixed vector w_0 and direction vector w_1. We will find a transformation T that takes one SET onto the other.

There is an invertible $n \times n$ matrix A such that $Av_1 = w_1$. To find such a matrix, first we say that since v_1 is nonzero, it contains a nonzero coordinate. Say it's the j-th coordinate $v_{1,j}$. If $(e_1 e_2 ... e_n)$ is the identity matrix in n dimensions, let

$$B = (e_1 ... v_1 ... e_n)$$

where v_1 replaces the j-th column. Then $Be_j = v_1$ and B is invertible because $det(B) = v_{1,j} \neq 0$. Similarly, define C so that $Ce_k = w_1$ for some index k. Finally, the permutation matrix P defined by switching the j-th and k-th columns of the identity matrix has the property that $Pe_j = e_k$. Then $A = CPB^{-1}$ satisfies $Av_1 = CPB^{-1}v_1 = CPe_j = Ce_k = w_1$. Finally, define $T(v) = Av + (w_0 - Av_0)$ so that $T(v_0) = w_0$. Then

$$\begin{aligned} T(v_0 + \alpha v_1) &= A(v_0 + \alpha v_1) + w_0 - Av_0 \\ &= w_0 + \alpha w_1 \end{aligned}$$

for any scalar α. □

This proves that all SETs are the same up to an affine transformation, but what about no-SETs? In this paper, the no-SETs that are of interest are those that are maximal. Define an d-cap to be a no-SET in d dimensions. The maximum possible size for a d-cap is given in the table below, with the maximal size of a d-cap denoted a_d.

d	1	2	3	4	5	6	7
a_d	2	4	9	20	45	$112 \leq a_6 \leq 114$	?

Computer programmers used exhaustive search methods to find the maximal size of an d-cap for dimension $d < 5$, but it becomes extremely complicated. Programmer Donald Knuth was using the concept of affine transformations to reduce the number of cases needed in an exhaustive search [1].

5. Finite Fields and SET

Because there are three options for each aspect of the SET cards, and 3 is prime, we can relate the cards to fields of order 3^n for $n = 1, 2, 3, 4$. Since the field of order 3 contains exactly one SET, the maximal number of cards not containing a SET must be 2. The more interesting cases are $n = 2, 3$, and 4. Since one of our main goals is to prove that a collection of, say, m cards must contain a SET, we must first prove that there is a collection of $m-1$ cards that does not contain a SET. Here we will do this using finite fields of size 3^n for the cases $n = 2$ and $n = 4$. The case $n = 3$ was much too complicated, and therefore will not be presented using this method.

If we only deal with part of the deck, say we let only two aspects vary, then we have a deck of nine cards. An example is pictured in Figure 1. This deck of nine cards corresponds to the field $\mathbb{F}_9$. We will first show that there are four elements in $\mathbb{F}_9$ that do not contain a SET, and in later sections, we will prove that this is actually the maximal size of a collection containing no SETs. We will do this by assuming that there is a collection of five elements that does not contain a SET and then deriving a contradiction.

Proposition 9. *There exists a collection of four elements in $\mathbb{F}_9$ that does not contain a* SET.

Proof. Let $S = \{x \in \mathbb{F}_9 : x^4 = 1\}$. We claim that S contains no SETs. We will proceed by contradiction. Suppose S does contain a SET. Then there are three distinct elements, $u, v, w \in S$ such that $u + v + w = 0$. This means that $\frac{u}{u} + \frac{v}{u} + \frac{w}{u} = 0$, alternatively $1 + \frac{v}{u} + \frac{w}{u} = 0$. Now we can say that there exists some $x, y \in S$ so that $1 + y + x = 0$, so $y = -1 - x$. Hence, x is a solution to $(1+x)^4 - 1 = 0$ and $x^4 - 1 = 0$. By factoring these two polynomials modulo 3, we have that

$$x^4 - 1 = (x^2-1)(x^2+1) = (x-1)(x+1)(x^2+1)$$

and by substituting $x + 1$ for x, we get

$$(x+1)^4 - 1 = ((x+1)-1)((x+1)+1)((x+1)^2+1) = (x)(x-1)(x^2-x-1).$$

The only factor that these two have in common is $(x-1)$, so the only solution to both of them is $x = 1$. This means that

$$x = \frac{v}{u} = 1,$$

hence $v = u$. This contradicts the assumption that u, v, and w are distinct elements of S.

□

When we let $n = 4$, we are dealing with the field $\mathbb{F}_{81}$. Since we will be proving that the number of cards guaranteeing a SET is 21, here we will prove that there does exist a collection of 20 cards not containing a SET. We will later show how to find such a collection of cards and then prove that it is actually the maximal size of a no-SET in Section 8.

Theorem 1. *There exists a collection of 20 elements in $\mathbb{F}_{81}$ that does not contain a* SET.

Proof. Let $S = \{x \in \mathbb{F}_{81} : x^{20} = 1\}$. We claim that S contains no SETs. We will proceed by contradiction. Suppose S does contain a SET. Then there are three distinct elements, u, v, and w, such that $u^{20} = v^{20} = w^{20} = 1$ and $u + v + w = 0$. From the above proof we can see this implies that there exist distinct elements $1, x, y \in S$ such that $1 + x + y = 0$. Then $y = -1 - x$, so x is a solution to $(1+x)^{20} - 1 = 0$ and $x^{20} - 1 = 0$.

Let's begin by factoring $x^{20} - 1 = 0$. We will need some standard formulas about geometric sums.

$$x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x^2 + x + 1) \qquad \text{for any positive integer } n$$
$$x^n + 1 = (x+1)(x^{n-1} - x^{n-2} + - \cdots + x^2 - x + 1) \qquad \text{for odd } n$$

If we apply these formulas for various n, we get

$$\begin{aligned} x^{20} - 1 &= (x^{10} - 1)(x^{10} + 1) \\ &= (x^5 - 1)(x^5 + 1)(x^2 + 1)(x^8 - x^6 + x^4 - x^2 + 1) \\ &= (x-1)(x^4 + x^3 + x^2 + x + 1)(x+1)(x^4 - x^3 + x^2 - x + 1) \cdot \\ &\quad \cdot (x^2 + 1)(x^8 - x^6 + x^4 - x^2 + 1) \\ &= (x-1)(x+1)(x^2+1) \cdot \\ &\quad \cdot (x^4 + x^3 + x^2 + x + 1)(x^4 - x^3 + x^2 - x + 1) \cdot \\ &\quad \cdot (x^8 - x^6 + x^4 - x^2 + 1) \end{aligned}$$

Now we must establish which of these are irreducible and reduce the polynomial completely. Since $\mathbb{F}_{81}$ is a fourth degree extension of $\mathbb{F}_3$, all elements of $\mathbb{F}_{81}$ are solutions to polynomials in $\mathbb{F}_3$ of degree 1, 2, or 4 (the divisors of 4). The irreducible polynomials of degree 1 are simply x, $x - 1$, and $x + 1$. The irreducible polynomials of degree 2 have roots in $\mathbb{F}_9$ but not $\mathbb{F}_3$. They are the degree 2 factors of

$$\begin{aligned} x^8 - 1 &= (x^4 - 1)(x^4 + 1) \\ &= (x^2 - 1)(x^2 + 1)(x^4 + 1) \\ &= (x-1)(x+1)(x^2+1)(x^4+1) \\ &= (x-1)(x+1)(x^2+1)(x^2 - x - 1)(x^2 + x - 1) \end{aligned}$$

The last two factors can be found by trial and error, since the only possible choices are $x^2 \pm x \pm 1$. Since the roots of $x^4 + 1$ satisfy $x^4 = -1$, we know that they are *not* roots of $x^{20}-1$ since $x^{20}=(x^4)^5=(-1)^5=-1$. Hence, the last two quadratics are not factors of $x^{20}-1$.

This proves that the degree 1, 2, and 4 factors that we already had are irreducible, and that the degree 8 factor must factor into two irreducible degree 4 factors [3].

To begin to find the factors, we will prove that the constant term in both quartics must be equal to 1.

If a is a root of a polynomial $f(x)$ over the field $\mathbb{F}_3$, then since $f(x)^3=f(x^3)$ by the binomial expansion and since 3=0, it is also true that a^3 is a root of $f(x)$. If $f(x)$ is irreducible of degree 4, then the four roots are a, a^3, a^9, and a^{27}. So, the polynomial would be

$$f(x)=(x-a)(x-a^3)(x-a^9)(x-a^{27}).$$

Notice that when this is multiplied, the constant term is a^{40}. Since a satisfies $a^{20}=1$, then $a^{40}=(a^{20})^2=1$. Hence, the constant term of both quartics is 1.

Now we know that

$$x^8-x^6+x^4-x^2+1=$$
$$(x^4+Ax^3+Bx^2+Cx+1)(x^4+Dx^3+Ex^2+Fx+1)$$

for some constants $A,B,C,D,E,F\in\{0,\pm 1\}$. Since there is no x^7 or x term, $A+D=0$ and $F+C=0$. That gives

$$x^8-x^6+x^4-x^2+1=$$
$$(x^4+Ax^3+Bx^2+Cx+1)(x^4-Ax^3+Ex^2-Cx+1)$$

Now, if we replace x with $-x$, the degree 8 polynomial on the left doesn't change. That means that the two factors on the right are transposed. Hence, $B=E$, and so

$$x^8-x^6+x^4-x^2+1=$$
$$(x^4+Ax^3+Bx^2+Cx+1)(x^4-Ax^3+Bx^2-Cx+1)$$

The coefficient of x^6 is now $2B-A^2=-1$ and the coefficient of x^2 is $2B-C^2=-1$. If $A=0$, then $B=1$ and $C=0$. If $A\neq 0$, then $A^2=1$, $B=0$ and $C\neq 0$ as well. By trying these possibilities, we find

$$x^8-x^6+x^4-x^2+1=$$
$$(x^4+x^3-x+1)(x^4-x^3+x+1)$$

This completes our factorization of $x^{20}-1$.

$$\begin{aligned}x^{20}-1=(x-1)(x+1)(x^2+1)\cdot\\ \cdot(x^4+x^3+x^2+x+1)(x^4-x^3+x^2-x+1)\cdot\\ \cdot(x^4+x^3-x+1)(x^4-x^3+x+1)\end{aligned}$$

The factorization of $(x+1)^{20}-1$ can be done simply by replacing x by $x+1$ in all of the factors, remembering that 3=0 and 2=-1. It helps to observe that

$$\begin{aligned}(x+1)^4&=x^4+x^3+x+1\\ (x+1)^3&=x^3+1\\ (x+1)^2&=x^2-x+1\end{aligned}$$

Using these, we obtain:

$$\begin{aligned}(x+1)^{20}-1=x(x-1)(x^2-x-1)\cdot\\ \cdot(x^4-x^3+x^2+x+1)(x^4+x^2+x+1)\cdot\\ \cdot(x^4-x^3-1)(x^4-x-1)\end{aligned}$$

We can see that the only common factor of $x^{20}-1$ and $(x+1)^{20}-1$ is $(x-1)$. So, if x satisfies both of them, then $x-1=0$, hence $x=1$. This contradicts the choice of 3 distinct elements $1,x,y\in S$. □

6. Finding the SET-less Set

In the previous section we proved that there exists a collection of 20 SET cards that does not contain a SET, therefore it is of interest to use the method in the proof to generate the 20 card no-SET.

Since the collection S is defined as $S=\{x\in\mathbb{F}_{81}:x^{20}=1\}$, we need to decide how to place a multiplication operation onto the SET cards. To begin, we will look at the structure of the field $\mathbb{F}_{81}$. This field is a fourth degree extension of $\mathbb{F}_3$, or we can view it as a second degree extension of $\mathbb{F}_9$. As previously stated in this paper, $\mathbb{F}_9$ is the collection $\{a+b\alpha|a,b\in\mathbb{F}_3,\alpha^2=-1\}$. Now, $\mathbb{F}_{81}$ viewed as a second degree extension of $\mathbb{F}_9$ has elements of the form $a+b\beta$ with $a,b\in\mathbb{F}_9$ and β is the root of a polynomial with coefficients in $\mathbb{F}_9$. We want to choose $\beta\notin\mathbb{F}_9$, because if $\beta\in\mathbb{F}_9$, then the collection of elements $a+b\beta$ with $a,b\in\mathbb{F}_9$ would just be $\mathbb{F}_9$, since it is closed. Let's choose β so that $\beta^2=1+\alpha$. This means that elements of $\mathbb{F}_{81}$ are of the form $a+b\alpha+(c+d\alpha)\beta=a+b\alpha+c\beta+d\alpha\beta$ with $a,b,c,d\in\mathbb{F}_3$. Now that we know what elements of $\mathbb{F}_{81}$ look like, we can define multiplication and addition of these polynomials in the usual way with the coefficients written in arithmetic modulo 3.

Since $\mathbb{F}_{81}^{\times}$ is a cyclic group under multiplication, we can try to find a generator. A generator would be a polynomial $a + b\alpha + c\beta + d\alpha\beta$ as defined above such that $(a+b\alpha+c\beta+d\alpha\beta)^n = 1$ only when $n = 80$. We can choose the polynomial $1 + \alpha + \beta$ and since the only divisors of 80 are 2, 4, 5, 8, 16, 20, and 40, we need only to check that $g = 1+\alpha+\beta$ has the property that $g^{40} = -1$ and that $g^{16} \neq 1$ since each of the other powers are divisors of these two.

To generate powers of g, we can start by finding g^2. We must remember the rules that in $\mathbb{F}_3$, $3 = 0$ and $2 = -1$. Also, because of the way we chose α and β, $\alpha^2 = -1$ and $\beta^2 = 1 + \alpha$.

Hence,

$$\begin{aligned} g^2 = (1+\alpha+\beta)^2 &= 1+\alpha+\beta+\alpha-1+\alpha\beta+\beta+\alpha\beta+\alpha+1 \\ &= 1+3\alpha+2\beta+2\alpha\beta \\ &= 1-\beta-\alpha\beta \end{aligned}$$

Continuing in this manner, we can keep multiplying g by itself and simplifying to get the other powers. Once we have found g^4, we can just multiply that by itself to find the fourth powers of the generator. We found that

$$\begin{aligned} (1+\alpha+\beta)^4 &= -1-\alpha+\beta+\alpha\beta \\ (1+\alpha+\beta)^{16} &= -1+\alpha-\beta+\alpha\beta \\ (1+\alpha+\beta)^{40} &= -1 \end{aligned}$$

hence, we know that g will generate the entire multiplicative group $\mathbb{F}_{81}^{\times}$.

Since calculating powers of g in this manner is fairly tedious, the same calculations can be carried out using matrix multiplication [3]. We can represent multiplication by g as multiplication by a matrix A with entries in $\mathbb{F}_3$. Recall that $\alpha^2 = -1$ and $\beta^2 = 1 + \alpha$. These imply

$$\begin{array}{lcrcccccc} g\cdot 1 &=& 1 &+& \alpha &+& \beta & & \\ g\cdot\alpha &=& -1 &+& \alpha & & &+& \alpha\beta \\ g\cdot\beta &=& 1 &+& \alpha &+& \beta &+& \alpha\beta \\ g\cdot\alpha\beta &=& -1 &+& \alpha &-& \beta &+& \alpha\beta \end{array}$$

We can read the columns of the corresponding matrix A from the coefficients of these results.

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

Then the coefficients of g^n in

$$g^n = a + b\alpha + c\beta + d\alpha\beta$$

can be calculated by repeated matrix multiplication

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

From the proof in the previous section, the subgroup defined by $S = \{x \in \mathbb{F}_{81}^{\times} | x^{20} = 1\}$ has twenty elements and contains no SETs. By SET, we mean a collection of three elements $x, y, z \in \mathbb{F}_{81}$ such that $x + y + z = 0$. We claim that every fourth power of our generator will have the property that $x^{20} = 1$ beacuse $((1 + \alpha + \beta)^{4n})^{20} = (1 + \alpha + \beta)^{80n} = 1^n = 1$. So, we can find all of the elements of S by taking every fourth power of our generator [4]. The coefficients of the fourth powers of g that we calculated in the above manner are represented in the following table.

	1	α	β	$\alpha\beta$
g^4	-1	-1	1	1
g^8	1	1	0	-1
g^{12}	-1	1	-1	0
g^{16}	-1	1	-1	1
g^{20}	0	1	0	0
g^{24}	1	-1	-1	1
g^{28}	-1	1	1	0
g^{32}	-1	-1	0	-1
g^{36}	-1	-1	-1	-1
g^{40}	-1	0	0	0
g^{44}	1	1	-1	-1
g^{48}	-1	-1	0	1
g^{52}	1	-1	1	0
g^{56}	1	-1	1	-1
g^{60}	0	-1	0	0
g^{64}	-1	1	1	-1
g^{68}	1	-1	-1	0
g^{72}	1	1	0	1
g^{76}	1	1	1	1
g^{80}	1	0	0	0

Above we have shown that $(1+\alpha+\beta)^4 = -1-\alpha+\beta+\alpha\beta$. What does this mean in relation to the SET cards? Let the coefficients of the element $a + b\alpha + c\beta + d\alpha\beta$ of $\mathbb{F}_{81}$ be assigned an aspect of the SET cards. Say that a tells us the number of shapes, b tells us the color, c tells us the shading, and d tells us the shape on the card. By using the table below we can decipher what card $(1+\alpha+\beta)^4$ represents.

Aspect	-1	0	1
Number:	Two	Three	One
Color:	Red	Green	Purple
Shading:	Open	Striped	Solid
Shape:	Oval	Squiggle	Diamond

We can see that $(1+\alpha+\beta)^4 = -1-\alpha+\beta+\alpha\beta$ represents the card with two red solid diamonds. Continuing in this manner, we can generate the whole set S of twenty cards that does not contain a SET, as shown in Figure 2.

By looking at the no-SET of twenty cards, we can see some patterns emerging. If you focus only on the red cards, it would be like playing with three choices and three aspects. This represents the field $\mathbb{F}_3^3$ or the field $\mathbb{F}_{27}$. Notice that there are nine red cards in the collection. In the next section, we will prove that there are a maximum of nine cards that do not contain a SET. This collection of nine red cards is an example of nine cards that does not contain a SET.

Notice the pattern that the colors have. There are nine red cards, nine purple cards, and two green cards. Later we will call this a $\{9,9,2\}$ triple. Notice that there are 8 squiggles, 6 ovals, and 6 diamonds to make an $\{8,6,6\}$ triple. There are eight possible triples that could show up, and they correspond to the hyperplane triples that we will discuss in Section 8. It turns out that the only triples that will actually show up are $\{9,9,2\}$ and $\{8,6,6\}$, but the proof of that fact is beyond the scope of this paper.

If you concentrate only on the red cards that have two shapes on them, you will notice that there are four of those. If we played with only the red cards with two shapes, it could be represented with $\mathbb{F}_3^2$ which is isomorphic to $\mathbb{F}_9$. In the next section we will prove that the maximal size of a no-SET is four, and we can clearly see that this collection of four red cards with two shapes on them is a collection of four cards in $\mathbb{F}_3^2$ that does not contain a SET.

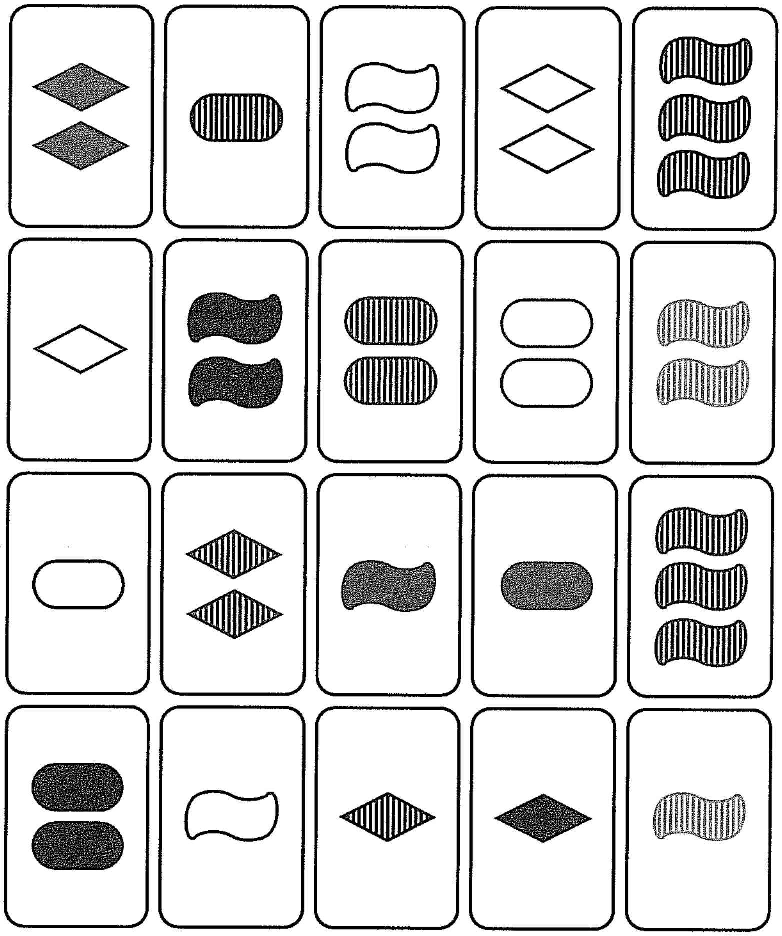


FIGURE 2. The 20 card no-SET that we generated.

7. PROOFS ABOUT MAXIMAL CAPS FOR THE TWO DIMENSIONAL AND THREE DIMENSIONAL CASES

A d-cap is defined as a no-SET in a d-dimensional vector space, alternatively, a collection of elements or vectors in a d-dimensional vector space over $\mathbb{F}_3$ that does not contain a SET. Therefore a maximal d-cap is the biggest such collection. The maximal d-cap has been found and proven for $d \leq 5$, but for higher dimensions, it is still unknown [1].

Here we will prove that a maximal 2-cap contains 4 elements. This is the same as saying that the maximal number of elements of $\mathbb{F}_3^2$ that does not contain a SET is 4. We will also prove that the maximal 3-cap contains 9 elements. We have shown an example of a collection of 9 cards that does not contain a SET when letting three aspects vary, so here we will prove that 9 is actually the maximal size of such a collection.

In this section, we will use "lines" and "planes" in our proofs. Here we are assuming that these are affine lines and planes, i.e. ones that do not necessarily pass through the origin.

Proposition 10. *The maximal 2-cap contains 4 elements.*

Proof. We have seen an example of a 2-cap that has four elements. Let's proceed by contradiction. Assume that there exists a 2-cap of size 5. Let C be such a 2-cap. We can divide $\mathbb{F}_3^2$ into spreads consisting of three parallel lines, and each parallel line is a copy of $\mathbb{F}_3$. We can define each spread by the line through the origin that is orthogonal to that spread. Let's proceed to count the number of spreads. Consider the number of lines through the origin. Since $|\mathbb{F}_3^2| = 3^2 = 9$, after we remove the origin itself, there are 8 nonzero points. Since each line contains 2 points other than the origin, there are $\frac{8}{2} = 4$ lines in $\mathbb{F}_3^2$ that contain the origin. Each of these lines containing the origin determines a spread of three parallel lines in $\mathbb{F}_3^2$, hence there are four possible spreads in $\mathbb{F}_3^2$. Let L_1, L_2, L_3 be such a spread. Define a triple to be $|L_1 \cap C|, |L_2 \cap C|, |L_3 \cap C|$. Since any three points in $\mathbb{F}_3$ contains a line, the only possible triple is $\{2, 2, 1\}$. Define a 2-marked line as a line in $\mathbb{F}_3^2$ that contains two specified points from C. Let's proceed to count the number of two marked lines.

There are 5 points in C, so we need to choose 2 points to determine a 2-marked line. Since a line is determined by two points, there is only one line for each pair of points that we choose from C. So, the number of 2-marked lines is

$$\binom{5}{2} \times 1 = 10.$$

On the other hand, for each spread, we have a certain number of 2-marked lines. Since we have assumed that C is a 2-cap, C contains at most two points from any given line, and hence these lines are distinct. So for each spread of parallel planes, we want to calculate the number of 2-marked lines. To do this, we determine a line by picking one of the three planes and choosing two points in C out of that plane. Since the planes are disjoint, we can add the numbers of 2-marked lines that are calculated this way. That gives

$$\binom{2}{2} + \binom{2}{2} + \binom{1}{2} = 1 + 1 + 0 = 2$$

2-marked lines per spread. Since there are four spreads, we count $2 \times 4 = 8$ 2-marked lines in $\mathbb{F}_3^2$. This contradicts that we already found ten 2-marked lines. Therefore, a 2-cap of size 5 does not exist. $\square$

Considering the relationship between the vector space and finite fields of order p^n, we can construct an alternative proof of the above theorem about a maximal 2-cap. Here we will use the field $\mathbb{F}_9$ instead of the 2-dimensional vector space $\mathbb{F}_3^2$, because the structures of these two are the same.

Proof. Suppose $S \subset \mathbb{F}_9$ and $|S| = 5$. Suppose by way of contradiction that S contains no SETs. This means that $x_i + x_j + x_k \neq 0$ for any $x_i, x_j, x_k \in S$. So $S = x_1, x_2, x_3, x_4, x_5$. To simplify the alleged no-SET, we can perform linear transformations of the form $x \mapsto ax + b$ for $a, b \in \mathbb{F}_9$. Since these are linear transformations, by Proposition 5 the image is also a no-SET. We can use the linear transformation defined by $\sigma(x) = x - x_1$ on $\mathbb{F}_9$. Now $\sigma(S)$ contains no SETs. We can then perform another transformation on $\mathbb{F}_9$, defined by $\tau(x) = \frac{x}{x_2 - x_1}$. We can see that τ is also an invertible linear transformation of $\mathbb{F}_9$ onto itself. Then $\tau(\sigma(x_1)) = 0$, and $\tau(\sigma(x_2)) = 1$. Say that $\tau(\sigma(x_i)) = y_i$, and $\tau(\sigma(S)) = T$. Then $\tau(\sigma(S)) = T = \{0, 1, y_3, y_4, y_5\}$. So, we know that $-1 \notin T$, because $1+0+(-1) = 0$, and similarly, $-y_3, -y_4 \notin T$ because $0+y_i-y_i = 0$. Also, $-1-y_3, -1-y_4 \notin T$ because $1+y_i+(-1-y_i) = 0$. So that makes five elements $\{-1, -y_3, -y_4, -1-y_3, -1-y_4\}$ that are not in T, but since $|\mathbb{F}_9| = 9$ and $|S| = 5$, there are only four elements in $\mathbb{F}_9 \setminus S$. So, y_5 must equal one of these four elements. Hence T does contain a SET, and since $S = \sigma^{-1}(\tau^{-1}(T))$, S must also contain a SET. $\square$

Proposition 11. *A maximal 3-cap contains 9 elements.*

Proof. We have displayed a 3-cap with 9 elements. Let's proceed by contradiction. Assume C is a 3-cap containing 10 elements. We can divide $\mathbb{F}_3^3$ into spreads of three parallel planes. Each of these spreads intersects C at all 10 points of C. Consider a spread H_1, H_2, H_3. We can define a triple to be $|H_1 \cap C|, |H_2 \cap C|, |H_3 \cap C|$. Since the maximal 2-cap is 4, $|H_i \cap C| \leq 4$. Also, the minimum order for $H_i \cap C$ must be 2 or 3, because if it were 1, we could have a maximum of $1+4+4 = 9$ points in C. So, there are two possibilities for triples in $\mathbb{F}_3^3$, $\{4, 4, 2\}$ and $\{4, 3, 3\}$.

Let a be the number of $\{4,4,2\}$ triples and b be the number of $\{4,3,3\}$ triples. To find how many triples we have, we can count each spread of three planes by the line orthogonal to the spread. We need to look for the number of lines through the origin, and that will tell us the number of normal lines to possible for a spread. Since $|\mathbb{F}_3^3| = 27$, there are 26 nonzero points, and since each line contains two nonzero points, there are $\frac{26}{2} = 13$ unique normal lines, and hence 13 spreads of parallel planes. So, if a and b represent the number of triples of a certain type, then $a + b = 13$.

Define a 2-marked plane to be a plane containing two specified points from C. Each spread has exactly three planes. We can mark a plane in the spread by choosing a pair of the points in C that lie on the plane. When we choose 2 points, there are 25 points left over, but we want to exclude the point that makes a line with our two chosen points, so there are 24 possible points left over. Each plane containing the line is determined by one point not on the line. Also, each plane has a distint set of six points not on the line. Hence, there are $\frac{24}{6} = 4$ planes containing our two points. This means there must be $\binom{10}{2} \times 4 = 180$ 2-marked planes.

On the other hand, each spread has a certain number of 2-marked planes contained in it. For each $\{4,4,2\}$ triple we count

$$\binom{4}{2} + \binom{4}{2} + \binom{2}{2} = 13$$

2-marked planes, and for each $\{4,3,3\}$ triple we count

$$\binom{4}{2} + \binom{3}{2} + \binom{3}{2} = 12$$

2-marked planes.

This gives us a second equation $13a + 12b = 180$. We already had that $a + b = 13$, hence $a = 13 - b$. By substituting, we can see that $13(13 - b) + 12b = 180$ implies $b = -11$. Since we defined b to be the number of hyperplane triples of type $\{4,3,3\}$, b must be a nonnegative integer. This contradiction proves that our initial assumption that there exists a 3-cap with 10 elements is false. □

8. The Four Dimensional Case

We have established that the SET cards correspond to the field $\mathbb{F}_3^4$. In this section we will prove that the maximal collection of SET cards that does not contain a SET is 20. In previous sections we have shown that a no-SET of size 20 exists. Here we will prove that 20 is the maximal size of such a collection by considering the four dimensional vector space over $\mathbb{F}_3$, assuming that a 4 cap of size 21 exists, and deriving a contradiction.

In this four-dimensional case, the proof is a bit more complicated than in the 2 and 3 dimensional case in the previous section. In order to start this proof, we need to prove some things about how many hyperplanes can be constructed containing a fixed number of points. We can define a k-flat K to be a translation of a k-dimensional subspace W of a vector space by a fixed vector v_0. Thus, $K = \{v + v_0 | v \in W\}$. If a k-flat contains the origin, then it is a k-dimensional subspace. We will use the term "hyperplane" to refer to a $(d-1)$-flat, where d is the dimension of the vector space.

Proposition 12. *The number of hyperplanes that contain a fixed k-flat in $\mathbb{F}_3^d$ is given by*

$$\frac{3^{d-k}-1}{2}.$$

Proof. Let K be a k-flat. Then there is a vector v_0 and a k-dimensional vector space V such that $K = v_0 + V$. There is a one-to-one correspondence between the hyperplanes H that contain V and affine hyperplanes $v_0 + H$ that contain K. Thus, we are reduced to counting the number of hyperplanes that contain a k-dimensional subspace.

For that purpose, we use the notation of *quotient vector space.* We can define two vectors v, w in $\mathbb{F}_3^d$ to be equivalent if $v - w$ belongs to the subspace V. The set of equivalence classes of points in $\mathbb{F}_3^d$ forms a vector space of dimension $d-k$, and so is isomorphic to $\mathbb{F}_3^{d-k}$. This is called the *quotient space* of $\mathbb{F}_3^d$ modulo the subspace V, and denoted $\mathbb{F}_3^d/V$.

There is a natural map $v \mapsto v + V$ from the hyperplanes in $\mathbb{F}_3^d$ containing V to the hyperplanes in $\mathbb{F}_3^d/V$ (which is isomorphic to $\mathbb{F}_3^{d-k}$) simply containing the origin, since V is mapped to 0. Thus, we are now reduced to just counting the number of hyperplane subspaces in $\mathbb{F}_3^{d-k}$.

Each hyperplane subspace can be determined by a normal line. Since there are two nonzero points determining a line, there are half as many hyperplanes as there are nonzero points. There are $3^{d-k} - 1$ nonzero points in $\mathbb{F}_3^{d-k}$ so there must be $(3^{d-k}-1)/2$ hyperplanes containing the origin in $\mathbb{F}_3^{d-k}$ and hence there are $(3^{d-k}-1)/2$ hyperplanes containing K in $\mathbb{F}_3^d$. □

Now, we can apply this proposition to prove the maximal cap for dimension 4 has 20 elements. We will use the proposition in exactly three ways. First, a 0-flat is just a point, so the proposition says that the number of hyperplanes through a point is $(3^d - 1)/2$. Second, any two distinct points determine an affine line, which is just a 1-flat. Hence, the number of hyperplanes containing two distinct points is $(3^{d-1} - 1)/2$. The last case that you apply Prop. 12 to is the number of hyperplanes through three non-collinear points. They determine a unique affine plane or 2-flat. Hence, the number of hyperplanes is $(3^{d-2} - 1)/2$.

Theorem 2. *The maximum number of cards that does not contain a* SET *is 20.*

Proof. We have shown a collection of twenty cards that did not contain a SET. Now we will proceed by contradiction. Suppose that C is a collection of 21 elements of $\mathbb{F}_3^4$ that contains no lines. We can divide $\mathbb{F}_3^4$ into spreads of three hyperplanes. Let x_{ijk} be the number of $\{i, j, k\}$ hyperplane triples as defined in the previous proofs. The intersection of C with any hyperplane is a subset of points in an affine 3-dimensional space that contains no lines. Thus, by Proposition 9, that intersection has at most 9 points. Hence there are only seven possibilities for hyperplane triples: $\{9, 9, 3\}, \{9, 8, 4\}, \{9, 7, 5\}\{9, 6, 6, \}, \{8, 8, 5\}, \{8, 7, 6\}$, and $\{7, 7, 7\}$. Since each spread of hyperplanes contains precisely one hyperplane subspace passing through the origin the total number of hyperplane triples is the same as the number of hyperplane subspaces containing the 0-dimensional flat of the origin. Thus, by Proposition 12 with $d = 4$ and $k = 0$ there are $\frac{3^4-1}{2} = 40$ hyperplane triples.

Hence we have the equation

$$(1) \qquad x_{993} + x_{984} + x_{975} + x_{966} + x_{885} + x_{876} + x_{777} = 40.$$

To find a second equation using these variables, we can count the number of *2-marked hyperplanes.* Since there are $\frac{3^3-1}{2} = 13$ distinct hyperplanes containing a pair of points by Proposition 12, there are a total of $13\binom{21}{2} = 2730$ 2-marked hyperplanes. On the other hand, we can count the number of 2-marked hyperplanes for each kind of hyperplane triple. For each hyperplane triple of type $\{i, j, k\}$ there are exactly three hyperplanes, and they can each be marked by a choice of two points from C that lie on the hyperplane. Since there are i points in C lying on the first of the three hyperplanes in the spread, there are $\binom{i}{2}$ 2-marked hyperplanes coming from that hyperplane. Similar consideraions for the other hyperplanes in the spread leads to

$$\binom{i}{2}+\binom{j}{2}+\binom{k}{2}$$

total number of 2-marked hyperplanes for that particular spread of hyperplanes. Counting in this way leads to the equation

$$\left[\binom{9}{2}+\binom{9}{2}+\binom{3}{2}\right]x_{993}+\cdots+\left[\binom{7}{2}+\binom{7}{2}+\binom{7}{2}\right]x_{777}$$

By calculating the binomials and setting that equation equal to the total number of 2-marked hyperplanes, we derive that

$$75x_{993}+70x_{984}+67x_{975}+66x_{966}+66x_{855}+64x_{876}+63x_{777}=2730. \tag{2}$$

We need to obtain another equation in these variables, so we can count the number of *3-marked hyperplanes* in a similar way. 3-marked hyperplanes are hyperplanes containing three points from our set C. Since the three points are from the no-SET C, they determine a unique affine plane or 2-flat. From Proposition 12 with $d = 4$ and $k = 2$, we have that there are $(3^{4-2}-1)/2 = 4$ hyperplanes containing three fixed points from C. Thus, there are $4\binom{21}{3} = 5320$ 3-marked hyperplanes. We can count the number of hyperplanes in each triple similarly to the way we counted *2-marked hyperplanes* in each triple.

$$\left[\binom{9}{3}+\binom{9}{3}+\binom{3}{3}\right]x_{993}+\cdots+\left[\binom{7}{3}+\binom{7}{3}+\binom{7}{3}\right]x_{777}$$

3-marked hyperplanes. Hence we have the equation

(3)

$$169x_{993}+144x_{984}+129x_{975}+124x_{966}+122x_{885}+111x_{876}+105x_{777}=5320.$$

Now we have three equations and seven variables, which we would normally not be able to solve. Because we are looking for only non-negative integer solutions, however, we can narrow it down. Adding 693 times equation (1) to 3 times equation (3) and subtracting 16 times equation (2) gives

$$5x_{984}+8x_{975}+9x_{966}+3x_{885}+2x_{876}=0.$$

The only way this can happen with non-negative coefficients is for all of those variables to equal zero.

Thus we have

(1) $$x_{993} + x_{777} = 40$$

(2) $$75x_{993} + 63x_{777} = 2730$$

Then equation (2) minus 63 times equation (1) is $12x_{993} = 210$, or $x_{993} = \frac{210}{12}$, but each of the x_{ijk} are non-negative integers. This contradiction proves that our initial assumption that there exists a collection of 21 elements of $\mathbb{F}_3^4$ that does not contain a SET is false. $\square$

References

[1] Benjamin L. Davis and Diane MacLagan (2003). The Card Game SET. *The Mathematical Intelligencer*, Vol 25(3), Pages 33-40.

[2] Thomas W. Hungerford. *Abstract Algebra: An Introduction*, Saunders College Publishing, 1997.

[3] David Wright *Project Instructional Meetings*, Oklahoma State University, Stillwater, OK, 2005-2006.

[4] Anthony Kable *Private Communication*, Oklahoma State University, Stillwater, OK, 2005