PERSPECTIVES ON AN OPEN QUESTION

IOANA BERCEA

Abstract. Starting with the card game SET, we ask the question of how many cards should be laid on the table in order to guarantee the existence of a set. This question is inherently related to the maximum size of an independent set in F_3^k. Although some specific values are known, the open question resides in its asymptotic evaluation. We set up the question and then explore its meaning in the context of hypergraphs, where its echoes are at least an interesting exercise in graph theory.

Contents

1. Introductory concepts

1.1. **Short overview of the card game SET.** The card game SET is played with a deck of 81 cards, each of which displays four attributes: number, shading, color and shape. Each of the attributes assumes three values. The goal of the game is to find as many collections of cards as possible. Three cards form a collection if, with respect to each of their attributes, they are either all the same or they are all different. We ask the question of how many cards one must lay on the table in order to guarantee the existence of a collection. In order to answer this, we will study the maximum size of a set of cards that does not contain a collection, also know as independence number.

1.2. **The mathematical perspective.** Let $\mathbb{F}_3$ be the field of three elements, and consider the vector space $\mathbb{F}_3^4$. A point of $\mathbb{F}_3^4$ is a 4-tuple where each coordinate can assume three possible values. From this perspective, we can think of SET cards as points in $\mathbb{F}_3^4$.

Date: August 19, 2008.

In order to enrich the question we are asking, we will generalize the game SET to a vector space of dimension k, for $k \in \mathbb{N}, k \geq 1$, in which the following rules apply.

Definition 1.1. A subset consisting of three elements $\{x, y, z\} \subset \mathbb{F}_3^k$ is an affine line if $x + y + z = 0$ and x, y, z are all distinct.

A collection of SET cards corresponds, therefore, to an affine line in $\mathbb{F}_3^4$.

Definition 1.2. An independent set $S \subseteq \mathbb{F}_3^k$ is a set S such that if $x, y, x \in S$ and $x + y + z = 0$, then $x = y = z = 0$.

A set of cards which does not contain a collection is, therefore, equivalent to an independent set in $\mathbb{F}_3^4$.

Let α_k be the maximum size of an independent set in $\mathbb{F}_3^k$.

2. Properties of α_k

Proposition 2.1 (Supermultiplicity). *We have that* $\alpha_{k+l} \geq \alpha_k \alpha_l$.

Proof. If $S_1 \subseteq \mathbb{F}_3^k$ and $S_2 \subseteq \mathbb{F}_3^l$ are maximal independent sets, then $S_1 \times S_2$ is also an independent set in $\mathbb{F}_3^{k+l}$. The size of $S_1 \times S_2$ is therefore at most equal to the size of a maximal independent set in $\mathbb{F}_3^{k+l} \Rightarrow \alpha_k \alpha_l \leq \alpha_{k+l}$. □

Now we will discuss the asymptotics of this independence number.

Lemma 2.2 (Fekete's Lemma). [1]*Let* $f : \mathbb{N} \to \mathbb{N}$ *be supermultiplicative(*$f(k+l) \geq f(k)f(l), \forall m, n \in \mathbb{N}$*). Then, there exists*

$$\lim_{k \to \infty} f(k)^{1/k}.$$

Proof. Fix m and l such that $l \leq m$. By induction and using the supermultiplicity of f, we have that $f(l + km) \geq f(l)[f(m)]^k$. From here, we get that

$$\liminf_{k \to \infty} f(l + km)^{1/(l+km)} \geq \liminf_{k \to \infty} f(l)^{1/(l+km)} \cdot \liminf_{k \to \infty} f(m)^{k/(l+km)} = f(m)^{1/m}.$$

Because $l \leq m$, we have that, for $\forall n, \exists k$ such that $n = l + km$, and therefore:

$$\liminf f(n)^{1/n} \geq f(m)^{1/m}.$$

Now, if we let $m \to \infty$, we get that:

$$\liminf f(n)^{1/n} \geq \limsup f(m)^{1/m},$$

but these are equal. Therefore, the limit exists. □

Proposition 2.3. *There exists*

$$\lim_{k \to \infty} {\alpha_k}^{1/k} = \sup_k {\alpha_k}^{1/k}.$$

Proof. First, let's prove that

$$\limsup_{k \to \infty} {\alpha_k}^{1/k} = \sup_k {\alpha_k}^{1/k}.$$

First, let's define the "solidity" of an independent set S in $\mathbb{F}_3^k$ to be:

$$\sigma(S) = \sqrt[k]{|S|}.$$

Let's observe that, for any natural nonnegative value of n, we have that

$$\sigma(S^n) = \sigma(S).$$

Therefore, we can say that

$$\limsup_{k\to\infty} \sigma(k) = \sup_k \sigma k.$$

But

$$\sup_k \sigma(k) = \sup_k {\alpha_k}^{1/k}.$$

Therefore,

$$\limsup_{k\to\infty} \sigma(k) = \sup_k {\alpha_k}^{1/k}. \tag{2.4}$$

Now, let $f : \mathbb{N} \to \mathbb{N}$ be defined such that $f(k) = \alpha_k$. Then, by using Fekete's Lemma, we get that the following limit exists:

$$\lim_{k\to\infty} {\alpha_k}^{1/k}. \tag{2.5}$$

We can conclude, therefore, from the two previous results that:

$$\lim_{k\to\infty} {\alpha_k}^{1/k} = \sup_k {\alpha_k}^{1/k}.$$

□

3. Known bounds

Proposition 3.1. $2^k \leq \alpha_k \leq 3^k$

Proof. Since $|\mathbb{F}_3^k| = 3^k$, we have that $\alpha_k \leq 3^k$. On the other hand, consider only the vectors with all components 0 or 1, and you get a an independent set because the only way to get an affine line in that set is to have all of the three vectors equal, which is a contradiction with the definition of an affine line. Since the size of this independent set is 2^k, we get that the independence number is greater than or equal than that. □

Going back to the properties of solidity, we observe that any known size of an independent set will give us a lower bound on $\lim_k \alpha_k$. The best lower bound on the solidity of maximal independent sets was given by Yves Edel who constructed an independent set with solidity 2.21739 in $\mathbb{F}_3^{480}$.

We know, therefore, that

$$2.21739 \leq \sup_{k\to\infty} \alpha_k^{1/k} \leq 3.$$

The major open question regards whether

$$\sup_{k\to\infty} \alpha_k^{1/k} = 3.$$

An improvement in this direction was given by Roy Meshulam who used Fourier Analysis to give an upper bound on α_k.[2] The proof is a surprising application of a seemingly unrelated field. We will mention the result but leave the proof to the interested reader.

Theorem 3.2 (Roy Meshulam)**.** *We have that*

$$\alpha_k \leq \frac{2 \cdot 3^k}{k}.$$

This upper bound, however, does not give an answer to our question, since

$$\lim_{k\to\infty} (\frac{2\cdot 3^k}{k})^{1/k} = 3.$$

4. The Combinatorial Approach

In search for the answer, one interesting approach is to translate the problem into a graph theory question and explore the possibilites in that realm.

4.1. **Construction of the Hypergraph.** Let us first give some basic defitions that will guide us through the construction.

Definition 4.1. A hypergraph H is a pair (V, E), where V is the set of vertices and E is a collection of non-empty subsets of V, called edges.

A hypergraph is just a generalization of the concept of graph(which is a hypergraph in which every edge has cardinality 2).

Definition 4.2. A hypergraph is called uniform if every edge has the same cardinality.

Definition 4.3. The degree of a vertex is the number of edges(subsets of V) that contain it.

Definition 4.4. A hypergraph is called regular if every vertex has the same degree.

Definition 4.5. A hypergraph automorphism is a function $\varphi : V \to V$ that preserves the edges, i.e. $\varphi(E) = E$.

Definition 4.6. A hypergraph is vertex- transitive if for any two vertices x and y, there exists an automorphism φ such that $\varphi(x) = y$.

Definition 4.7. A set $S \subsetneqq V$ is an independent set if it does not contain any edges, i.e. $S \cap E = \emptyset$.

Definition 4.8. The independence number or a hypergraph H, $\alpha(H)$ is the maximum size of an independent set of H, i.e. $\alpha(H) = \sup\{|S| : S$ is an independent set of $H\}$.

Let us think of the points in $\mathbb{F}_3^k$ as vertices of hypergraph. Three vertices in the hypergraph form an edge if their vector correspondents form an affine line in $\mathbb{F}_3^k$.

If $H = (V, E)$ would be the hypergraph such obtained, we would have the following property.

Proposition 4.9. *H is a uniform regular vertex-transitive hypergraph.*

Proof. The size of each edge is equal to the number of summands in an affine line, which is $3 \Rightarrow H$ is uniform.

H is also regular. We can fix x to be a vertex in V and for any $y \in V$, we have that there exists another $z \in V$ such that $\{x, y, z\} \in E$.(this happens in accordance to the definition of an affine line in $\mathbb{F}_3^k$) Since there are $3^k - 1$ vertices not equal to x and because $\{x, y, z\} = \{x, z, y\}$(we doublecount the edges), we have that the degree of each vertex is $\frac{3^k-1}{2}$.

Next, let's observe that no vertex can be distinguished from any other, based only on the vertices and edges surrounding it. Every vertex has the same local environment, therefore the hypergraph is vertex- transitive. ☐

The bridge between our previous approach to the question and our current one is the fact that the independence number in $\mathbb{F}_3^k$ is actually the independence number of our constructed hypergraph.

$$\alpha_k = \alpha(H)$$

4.2. **Bounds on the Hypergraph.** Let us, from now on, consider the hypergraph $H = (V, E)$ with $|V| = n = 3^k$. In order to explore the various bounds on $\alpha(H)$, we need to define one more concept.

Definition 4.10. The chromatic number of a hypergraph, $\chi(H)$ is the minimum number of maximal independent sets that cover V.

The concept of chromatic number of a hypergraph is usually known to be the minimum number of colors needed to color the vertices of H so that no edge is monochromatic. The two definitions are equivalent, because the partitions that the coloring makes out of the set of vertices correspond exactly to independent sets.

From the first definition of a chromatic number, we have that:

$$\alpha(H) \cdot \chi(H) \geq n.$$

Now we will explore the properties of the created hypergraph, through this theorem.

Theorem 4.11. *If H is vertex-transitive, then*

$$\alpha(H) \cdot \chi(H) \leq n \cdot (\ln(n) + 1).$$

Proof. Let us consider $S \subseteq V$ be a maximal independent set of the hypergraph, $|S| = \alpha(H)$ and let $G = Aut(H)$ be the automorphism group of H.

First, let's observe that for any $\varphi \in G$, we have that $\varphi(S)$ is also a maximal independent set.

Starting with the definition of the chromatic number, we will try to apply automorphisms to S and probabilistically determine when we will get a cover of V.

In order to do that, we need to first establish that:

Lemma 4.12. *For $x \in V$ a fixed vertex and let $\varphi \in G$ be a random automorphism. Then, for any $y \in V$,*

$$Pr[\varphi(x) = y] = \frac{1}{n}.$$

Proof. Let us consider $G_x = \{\varphi \in G : \varphi(x) = x\}$ to be the stabilizer of x in G. Let $y \in V$ be any vertex. Because of vertex-transitivity, we know that $\exists \varphi \in G$ such that $\varphi(x) = y$. Composing every automorphism in G_x with φ will give us a coset G' isomorphic to G_x. Not only that, but $G' = \{\varphi \in G : \varphi(x) = y\}$. Because y was arbitrarily chosen, we get that $|G_x| = |G'| = \frac{|G|}{n}$. Therefore, if we randomly consider $\varphi \in G$, then for any $y \in V$, we will have that:

$$Pr[\varphi(x) = y] = \frac{|G'|}{|G|} = \frac{\frac{|G|}{n}}{|G|} = \frac{1}{n}.$$

□

Now, going back to our initial proof, let $\varphi \in G$ be a random automorphism and fix $x \in V$. Then we get that:

$$Pr[x \in \varphi(S)] = \frac{|S|}{n}.$$

Therefore,

$$Pr[x \notin \varphi(S)] = 1 - \frac{|S|}{n}.$$

Now, let us randomly pick t automorphism from G, $g_1, g_2, ..., g_t \in G$.

$$Pr[x \notin \bigcup_{i=1}^{t} \varphi_i(S)] = (1 - \frac{|S|}{n})^t.$$

We have that

$$Pr[x \notin \bigcup_{i=1}^{t} \varphi_i(S)] = Pr[\bigwedge_{i=1}^{t} (x \notin \varphi_i(S))] = \prod_{i=1}^{t} Pr[v \notin \varphi_i(S)] = (1 - \frac{|S|}{n})^t.$$

We were able to pass to the product because the events considered are independent.(as functions of independently chosen automorphisms)

Now let us consider the probability that $\bigcup_{i=1}^{t} \varphi_i(S)$ will not cover V.

$$Pr[\exists x \in V, x \notin \bigcup_{i=1}^{t} \varphi(S)] = Pr[\bigvee_{x \in V} (x \notin \bigcup_{i=1}^{t} \varphi(S))].$$

But, because of union bound, we have that:

$$Pr[\bigvee_{x \in V} (x \notin \bigcup_{i=1}^{t} \varphi(S))] \leq \sum_{x \in V} Pr[x \notin \bigcup_{i=1}^{t} \varphi(S)] = n(1 - \frac{|S|}{n})^t.$$

Let us find a more suggestive upper bound on our probability. We know that

$$(1 - x)^t \leq e^{-xt}, \forall x \geq 0.$$

We therefore get that

$$(1 - \frac{|S|}{n})^t \leq e^{-\frac{|S| \cdot t}{n}}.$$

We observe that for $t = [\frac{n \cdot \ln(n)}{|S|}] + 1$, we get that

$$(1 - \frac{|S|}{n})^t \leq e^{-\frac{|S|}{n} \cdot \frac{n \cdot \ln(n)}{|S|}} = 1.$$

This means that for $t = [\frac{n \cdot \ln(n)}{|S|}] + 1$, we get that the probability of not covering V is smaller than 1, which automatically means that the probability of covering V is positive. In that case, we obtain that t automorphisms provide a covering of V, so, by the condition of minimality on the chromatic number, we get that:

$$\chi(H) \leq [\frac{n \cdot \ln(n)}{|S|}] + 1 \leq \frac{n \cdot \ln(n)}{|S|} + 1 \leq \frac{n \cdot (\ln(n) + 1)}{|S|}.$$

But $|S| = \alpha(H)$, so we get that:

$$\alpha(H) \cdot \chi(H) \leq n \cdot (\ln(n) + 1).$$

□

Therefore, we get that

$$n \leq \alpha(H) \cdot \chi(H) \leq n \cdot (\ln(n) + 1).$$

Let us consider not what bounds we can obtain on $\chi(H)$ from here.

$$\frac{n}{\alpha(H)} \leq \chi(H) \leq \frac{n \cdot (\ln(n) + 1)}{\alpha(H)}.$$

And we also know, since $n = 3^k$, that

$$2^k \leq \alpha(H) \leq \frac{2 \cdot 3^k}{k}.$$

Therefore, we get that:

$$\frac{k}{2} \leq \chi(H) \leq (k \ln 3 + 1) \cdot (\frac{3}{2})^k.$$

With these results, we enter a whole different world of determining the chromatic and independence numbers of hypergraphs, a problem that touches upon NP-complete problems and optimization algorithms.

5. Conclusions

As we have seen, a simple and natural question can involve much more complicated attemtps to answer it. From Fourier Analysis to hypergraphs, combinatorial questions span a lot of fields and can translate into other seemingly unrelated questions. Pursuing them can provide a challenging trip that leaves one, if not with an answer, at least with insight into different interesting fields that converge in surprising ways.

References

[1] J. H. van Lint and R. M. Wilson, A course in combinatorics, Cambridge University Press, Cambridge, second edition, 2001.

[2] Roy Meshulam, On Subsets of Finite Abelian Groups with No 3-Term Arithmetic Progressions, Journal of Combinatorial Theory, 1995