A Uniform Distribution Theorem Concerning Intersections of Finite Sets

Borbély József, Csala-Takács Éva

Óbudai Egyetem, Székesfehérvár, Hungary

borbely.jozsef@amk.uni-obuda.hu
csala.takacs@amk.uni-obuda.hu

***Abstract*— Families of finite sets with restricted intersections stood always in the centre of mathematical researches. Yet in the sixties there were papers which treated this topic. In this paper we will solve a Kürschák problem of 2019 of this topic, and we will write about its possible generalizations. The aspect is something new in respect of the former research work.**

I. Introduction

Families of finite sets that can intersect each other only in some prescribed ways have been investigated for a very long time. Throughout a problem of the Kürschák competition held in 2019 we will raise some research problems that are worth being investigated.

II. Some Problems in Extremal Set Theory

First of all, we want to give a survey about the former research work concerning the topic of restricted intersections. The first result about extremal sets was published in 1928, this was Sperner's theorem (see [4]):

Sperner's Theorem: *Let n be a positive integer, and let H be the set H={1,2,...,n}. If F is a family of sets consisiting of some subsets of H such that none of the sets in F is a subset of another set in F, then F contains at most* $\binom{\boldsymbol{n}}{[\boldsymbol{n}/\boldsymbol{2}]}$ *sets, and this bound is the best possible.*

One of the most important results about intersecting sets was the Erdős-Ko-Rado theorem, published in 1961 (see [2]).

Erdős-Ko-Rado Theorem: *Let n and k be positive integers, k ≤ n, and let H be H={1,2,...,n}. If F is a family of sets consisiting of some k-element subsets of H, and every two sets in F intersect each other (i.e. they have at least one common element), then there are at most* $\binom{\boldsymbol{n-1}}{\boldsymbol{k-1}}$ *sets in F, and this bound is the best possible.*

Later a lot of problems were raised, in which one gives some restrictions concerning the intersections of the sets modulo a prime number. Perhaps the most classic examples are the Oddtown and Eventown type theorems.

The famous Oddtown theorem was published by Berlekamp in 1969 (see [1]).

Oddtown Theorem: *Let n be positive integer, and let H be H = {1,2,...,n}. Assume that* $\boldsymbol{H_1}$, $\boldsymbol{H_2}$, ..., $\boldsymbol{H_m}$ *are distinct subsets of H such that for every index 1≤j≤n the number* $|\boldsymbol{H_j}|$ *is odd, and for every pair of indices*

1 ≤ k <l ≤ n the number $|\boldsymbol{H_k} \cap \boldsymbol{H_l}|$ *is even. Then m ≤ n, and this estimate is the best possible.*

The Eventown theorem was published independently by Berlekamp and Graver (see [1] and [3]).

Eventown Theorem: *Let n be positive integer, and let H be H = {1,2,...,n}. Assume that* $\boldsymbol{H_1}$, $\boldsymbol{H_2}$, ..., $\boldsymbol{H_m}$ *are distinct subsets of H such that for every index 1≤j≤n the number* $|\boldsymbol{H_j}|$ *is even, and for every pair of indices*

1 ≤ k <l ≤ n the number $|\boldsymbol{H_k} \cap \boldsymbol{H_l}|$ *is even. Then m ≤* $\boldsymbol{2^{[n/2]}}$, *and this estimate is the best possible.*

Since there were a lot of other results concerning the size of intersections modulo a prime number. Here we mention a last one, which without doubt merits attention. Xin, Wei and Ge recently published the following result, which verifies an older conjecture of Babai, Alon and Suzuki in a stronger form (see [5]).

Theorem (Xin, Wei, Ge): *Let p be a prime number and L={*$\boldsymbol{l_1}$, $\boldsymbol{l_2}$, ..., $\boldsymbol{l_s}$*} and K={*$\boldsymbol{k_1}$, $\boldsymbol{k_2}$, ..., $\boldsymbol{k_r}$*} two disjoint subsets of {0, 1, 2, ..., p-1}. Suppose that F is a family of subsets of {1, 2,..., n} such that the number of elements in every set of F is congruent to some* $\boldsymbol{l_i}$, *and the size of the intersection of every two sets in F is congruent to some* $\boldsymbol{k_j}$. *If n ≥ s + max{*$\boldsymbol{k_1}$, $\boldsymbol{k_2}$, ..., $\boldsymbol{k_r}$*}, then F contains at most*

$\sum_{t=0}^{2r-1} \binom{\boldsymbol{n-1}}{\boldsymbol{s-t}}$ *sets.*

III. Our Main Theorem

We formulate the second problem of the Kürschák competition held in 2019 as our main theorem:

Main Theorem : *Let n be positive integer, and let H be the set H={1,2,...,n}. If F is a family of subsets of H such that for every fixed nonempty subset X of H, exactly half of the values |X ∩ A| is odd, whenever A runs through all of the sets of F, then F contains of all of the subsets of H.*

Proof of the Main Theorem:

First of all we will prove that choosing all subsets of H as the family F, we get an appropriate construction. Let X be a subset of H that contains the integer number k (1 ≤ k ≤ n) as an element. Let A be a subset of H that does not contain the integer number k as an element.

In this case one of the two numbers |X∩A| and |X∩AU{k}| is even, and the other is odd. Thus ordering all subsets of

H in pairs in the manner such that such an A is a pair of the subset AU{k}, then one can see that the conditions of the theorem are fulfilled.

Now we have to prove that there is not any other possibility to choose such a family F. We will prove this statement by induction on n. For n=1 it is easy to check, that the statement is true. Let us assume that we

proved the statement of the theorem for all positive integers less than m. Now we have to prove it for n=m.

In this case H={1,2,...,m}. Let X be a nonempty subset of H, that does not contain the element m. Let us consider all of the values |X∩A|, if A runs through all of the sets of F that does not contain m as an element.

Let us assume that a_1 of these values are odd, and a_2 of them are even numbers.

Now let us consider the numbers |X∩B|, if B runs through all of the sets of F that contain m as an element. Let us assume that b_1 of these values are odd, and b_2 of them are even numbers.

Using the conditions of the theorem, a_1+b_1 = a_2 + b_2.

Now let us consider all of the values |(XU{m})∩A|, if A runs through all of the sets of F that does not contain m as an element.

Exactly a_1 of these values are odd, and a_2 of them are even numbers.

Similary, taking the numbers |(XU{m})∩B|, if B runs through all of the sets of F that contain m as an element, one can easily see, that b_2 of these numbers are odd, and b_1 of them are even numbers.

Using the condition of the theorem we get

a_1+b_2 = a_2 + b_1.

Thus we have to solve the sytem of equations

$a_1+b_1 = a_2 + b_2$

$a_1+b_2 = a_2 + b_1$

From the first equation we get $a_1 - a_2 = b_2 - b_1$, from the second one we get $a_1 - a_2 = b_1 - b_2$. This implies that $b_1=b_2$, and $a_1=a_2$.

Thus half of the values |X∩A| is odd, if X is a fixed nonempty subset of H, that does not contain the element m, and A runs through all of the sets of F that does not contain m as an element. Using the induction hypothesis, we get that all of the subsets of the set H-{m} must be present in F.

Now let us choose X as X={m}. Using the conditions we get, that half of the sets in F contain m as an element, thus $|F| \geq 2 \cdot 2^{m-1} = 2^m$, which means that F must contain all subsets of H.

IV. A Linear Algebraic Consequence of the Main Theorem

Let us identify every subset of {1, 2, ..., n} with an n-dimensional vector in the standard way, which means that for every index $1 \leq i \leq n$, the ith coordinate of the vector equals one, if i is an element of the subset, otherwise this coordinate will be zero. Now let us prepare the matrix I with (2^n-1) rows and n columns, where each row corresponds to exactly one nonempty subset of H. Let M be a matrix with (2^n-1) rows and k columns such that in every row of M exactly half of the elements is one, the other elements are zeros, and the rows of M are distinct vectors (the condition implies that k must be even). If we can find a matrix F* such that IF*=M over the field F_2, then F* has n rows and k columns. The columns of F* represent the sets of a good family F, as prescribed in our main theorem. But we proved that this is only possible if k = 2^n, which means that for $0 \leq k < 2^n$ we cannot solve the equation IF*=M over the field F_2 for the unknown matrix F*. If k = 2^n, then the rows of M are uniquely determined, they must be the distinct n-dimensional 0-1 vectors in some order.

V. Some Research Problems

First we will reformulate our main theorem in order to be able to some generalization ideas.

Main Theorem Revisited: *Let n be positive integer, and let H be the set H={1,2,...,n}. If F is a family of subsets of H such that for every fixed nonempty subset X of H, the number $|X \cap A|$ is even exactly with probability 1 / 2 (where A denotes a random set of F), then F contains of all of the subsets of H.*

A natural way to extend our investigation is to replace H by an infinite set (which can be countable or not), and one could try to choose an appropriate family F from the finite subsets of H. The problem could be the following (or one can search for any other similar variants):

Problem 1: *Let H be an infinite set, and let F be a family of sets containing some finite subsets of H. Assume that for every X fixed finite subset of H, the number |X∩A| is even with probabillity 1 / 2, where A is a random set of F. Are there any families F that correspond to these requirements? If there are any, can we describe all of them?*

Here appeared the term probability, which can lead to new research problems concerning this topic. We could replace the modulus two by any other number. In a general form, we can raise the following questions:

Problem 2: *Let n be positive integer, and let H be the set H={1,2,...,n}, and let F be a family of distinct subsets of H. Let m be a positive integer greater than 1, and let 0<p<q<1. Assume that for every X fixed subset of H, the number |X∩A| is divisible by m with a probabillity greater than p, but less than q, where A is a random set of F (this probability can vary, if we choose different sets A of F). Are there any families F that correspond to these requirements? If there are any, can we describe all of them?*

Problem 3: *Let H be an infinite set, and let F be a family of sets containing some finite subsets of H. Let m be a positive integer greater than 1, and let 0<p<q<1. Assume that for every X fixed finite subset of H, the number |X∩A| is divisible by m with a probabillity greater than p but less than q, where A is a random set of F (this probability can vary, if we choose different sets A of F). Are there any families F that correspond to these requirements? If there are any, can we describe all of them?*

Unfortunately, our ideas applied in the proof of the main theorem do not seem to work in the general cases.

If we shift an element modulo 2 by one, then we can be sure that we will get the other element of the field F_2.

But if we want to make calculations modulo m, where m is greater than 1, then this will not work. It seems probable that these problems can we solved (at least partially) using some linear algebraic methods, as for example in the case of Oddtown and Eventown type theorems.

Thus it seems impossible that these problems have a purely elementary solution like our main theorem. But even that is why these slight modifications can be good research problems in the future.

REFERENCES

[1] Berlekamp, On Subsets with Intersections of Even Cardinality, Canad. Math. Bull., 1969.

[2] Erdős; Ko; Rado, Intersection theorems for systems of finite sets, Quarterly Journal of Mathematics, Oxford. Second Series, ***12, 1961.***

[3] Graver, Boolean designs and self-dual matroids, Linear Algebra Appl., 1975.

[4] Sperner, Ein Satz über Untermengen einer endlichen Menge, Mathematische Zeitschrift ***27, 1928.***

[5] Xin, Wei, Ge: A strengthened inequality of Alon-Babai-Suzuki's conjecture on set systems with restricted intersections modulo p, Discrete Mathetics, 2017.