Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Finding pairwise disjoint vector pairs in $\mathbb{F}_2^n$ with a prescribed sequence of differences

(Extended abstract)

Benedek Kovács*

Abstract

We consider the following question by Balister, Győri and Schelp: given 2^{n-1} nonzero vectors in $\mathbb{F}_2^n$ with zero sum, is it always possible to partition $\mathbb{F}_2^n$ into pairs such that the difference between the two elements of the i-th pair is equal to the i-th given vector? An analogous question in $\mathbb{F}_p$ was resolved by Preissmann and Mischler in 2009. In this paper, we prove the conjecture in $\mathbb{F}_2^n$ in the case when there are at most $n - 2\log n - 1$ distinct values among the given differences, and also in the case when at least a fraction $\frac{28}{29}$ of the differences are equal.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-092`

1 Introduction

We consider the following conjecture of Balister, Győri and Schelp [2] from 2008:

Main conjecture 1.1. *Let $n \geq 2$ be an integer and $m = 2^{n-1}$. If the nonzero difference vectors $\mathbf{d_1}$, $\mathbf{d_2}$, $\ldots$, $\mathbf{d_m}$ are given in $\mathbb{F}_2^n$ such that $\sum_{i=1}^{m} \mathbf{d_i} = \mathbf{0}$ (and the $\mathbf{d_i}$'s are not necessarily distinct), then $\mathbb{F}_2^n$ can be partitioned into disjoint pairs $\{\mathbf{a_i}, \mathbf{b_i}\}$ $(1 \leq i \leq m)$ such that $\mathbf{a_i} - \mathbf{b_i} = \mathbf{d_i}$ holds for every i.*

*ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. E-mail: `benoke98@student.elte.hu`. Supported by the ÚNKP, New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

In 2008, Bacher [1] has independently posed another analogous version of this conjecture, where instead of $\mathbb{F}_2^n$, we partition the elements of $\mathbb{F}_p\setminus\{0\}$ into pairs, where p is an odd prime, and there is no restriction on the sum of the given (nonzero) differences. For this case, Preissmann and Mischler gave a positive answer [9]; their method relies on summing the values of an appropriate multivariate polynomial over $\mathbb{F}_p$.

Theorem 1.2 (Preissmann, Mischler)**.** *Let p be an odd prime and $M = \frac{p-1}{2}$. If in $\mathbb{F}_p$, the nonzero differences $d_1, d_2, \ldots, d_M$ are given, then $\mathbb{F}_p \setminus \{0\}$ can be partitioned into disjoint pairs $\{a_i, b_i\}$ $(1 \leq i \leq M)$ such that for each i, $a_i - b_i = d_i$ holds.*

Later, Kohen and Sadofschi [6] gave a new proof of this claim using the Combinatorial Nullstellensatz.

The statement can also be investigated for other cyclic groups as well. The following conjecture of Adamaszek pertaining to cyclic groups of even order has been proven by Kohen and Sadofschi [7]:

Theorem 1.3 (Kohen, Sadofschi)**.** *Let $n = 2M$ be even. If the elements $d_1, d_2, \ldots, d_M \in (\mathbb{Z}/n\mathbb{Z})^\times$ are arbitrarily given, then $\mathbb{Z}/n\mathbb{Z}$ can be partitioned into disjoint pairs $\{a_i, b_i\}$ such that for each i, we have $a_i - b_i = d_i$.*

Another way to generalize Theorem 1.2 is if we consider the problem for $\mathbb{F}_p^n\setminus\{\mathbf{0}\}$ instead of $\mathbb{F}_p\setminus\{0\}$. Karasev and Petrov showed that in this case, the same statement does not hold (by considering the case when every $\mathbf{d_i}$ is equal to the same nonzero vector $\mathbf{d}$). However they have shown the following claim [5, Theorem 3]:

Theorem 1.4 (Karasev, Petrov)**.** *Let p be an odd prime and $M = \frac{p^n-1}{2}$. If the sets $\{\mathbf{d_{1,1}}, \ldots, \mathbf{d_{1,n}}\}, \{\mathbf{d_{2,1}}, \ldots, \mathbf{d_{2,n}}\}, \ldots, \{\mathbf{d_{M,1}}, \ldots, \mathbf{d_{M,n}}\}$ are given in $\mathbb{F}_p^n$ such that each set is a basis of $\mathbb{F}_p^n$, then there exists a function $g : [M] \to [n]$ such that $\mathbb{F}_p^n \setminus \{\mathbf{0}\}$ can be subdivided into disjoint pairs $\{\mathbf{a_i}, \mathbf{b_i}\}$, $1 \leq i \leq M$ with $\mathbf{a_i} - \mathbf{b_i} = \mathbf{d_{i,g(i)}}$ for every i.*

If we investigate the statement in $\mathbb{F}_2^n$ instead of $\mathbb{F}_p^n$, then to obtain a perfect matching, we also need to include the zero vector in the set of elements to be matched. Even in this case, the claim does not hold for arbitrary nonzero differences, as the sum of differences has to be equal to the sum of all elements of the vector space, which is zero. By the main conjecture, this would be a sufficient condition for an adequate perfect matching to exist.

The authors of [2] have also verified this conjecture for the case $n \leq 5$, and they have proved the main conjecture in the following special case [2, Theorem 4]:

Theorem 1.5 (Balister, Győri, Schelp)**.** *The main conjecture is true in the case when the vectors $\mathbf{d_1}, \mathbf{d_2}, \ldots, \mathbf{d_{\frac{m}{2}}}$ are all equal, and for every integer $1 \leq i \leq \frac{m}{2}$ we have $\mathbf{d_{2i-1}} = \mathbf{d_{2i}}$.*

In 2021, Correia, Pokrovskiy and Sudakov [3] published the following result:

Theorem 1.6 (Correia, Pokrovskiy, Sudakov)**.** *Let G be a multigraph whose edges are t-coloured, so that each colour class is a matching of size at least $t + 20t^{15/16}$. Then there exists a rainbow matching (that is, a matching with t edges of all distinct colours).*

Applying this result to the graph on the vertex set $\mathbb{F}_2^n$ with colour class i consisting of the edges between pairs of difference $\mathbf{d_i}$, we get that for any $M \leq \frac{1}{2}N - C \cdot N^{15/16}$ nonzero differences $\mathbf{d_i}$, where $N = 2^n$, we can find disjoint pairs $\{\mathbf{a_i}, \mathbf{b_i}\}$ such that $\mathbf{a_i} - \mathbf{b_i} = \mathbf{d_i}$. However this method does not result in perfect matchings.

Gao, Ramadurai, Wanless and Wormald [4] conjectured that Theorem 1.6 holds for $t+2$ in place of $t+20t^{15/16}$, which would resolve this problem for any $M \leq \frac{1}{2}N-2$ nonzero difference vectors.

In this paper, we prove the main conjecture in the following two special cases:

- when the number of distinct values among the $\frac{1}{2}N$ difference vectors is at most $n - 2\log n - 1$;
- and when n is sufficiently large and at least a fraction $\frac{28}{29}$ of the difference vectors are all equal.

2 Perfect matching in the case of few difference classes

Let the nonzero differences $\mathbf{d_1}, \mathbf{d_2}, \ldots, \mathbf{d_m}$ be given such that $\sum_{i=1}^m \mathbf{d_i} = \mathbf{0}$, where $m = 2^{n-1}$. The collections containing all differences equal to a fixed vector $\mathbf{d}$ will be called *difference classes*. For a given configuration $\{\mathbf{d_1}, \mathbf{d_2}, \ldots, \mathbf{d_m}\}$, let t denote the number of nonempty difference classes. We would like to give a value $T(n)$ as large as possible, for which we can guarantee the existence of a suitable perfect matching of $\mathbb{F}_2^n$ in the case $t \leq T(n)$.

In the case $t = 1$ the task is trivial: take the $\langle \mathbf{d} \rangle$-cosets of $\mathbb{F}_2^n$ for the difference $\mathbf{d}$.

In the case $t = 2$, the task can be solved using Theorem 1.5, as $\sum \mathbf{d_i} = \mathbf{0}$ means that both difference classes have even size. So we have the structure that half of the differences are the same and the rest of the differences can be partitioned into equal-valued pairs.

Theorem 2.1. *The main conjecture is true in the case when the number of difference classes is at most $n - 2\log n - 1$.*

Lemma 2.2. *Let $n \geq 4$, and let P_n denote the power set of $[n]$ as a poset ordered by containment. Let H be a subset of P_n of size at most $n+1$, for which $\emptyset \notin H$ and $[n] \notin H$. Moreover assume that H does not contain all of the one-element sets and does not contain all of the $n-1$ element sets either. Then $P_n \setminus H$ contains a chain of size $n+1$.*

Proof sketch. This can be proved using the fact that P_n admits a decomposition into disjoint symmetric chains (i.e. chains containing one set of each integer cardinality between k and $n-k$ for some k); the proof of this fact can be found in [8, Proposition 2]. □

Proof sketch of theorem 2.1. Let the distinct values of the given differences be $\mathbf{u_1}$, $\mathbf{u_2}$, $\ldots$, $\mathbf{u_t}$ where for each $1 \leq i \leq t$, $\mathbf{u_i}$ appears n_i times with $n_1 \geq n_2 \geq \cdots \geq n_t$.

Let $U = \langle \mathbf{u_1}, \ldots, \mathbf{u_t} \rangle$ and $k = \dim U$. (Then $k \leq t$.) We can assume that $k \geq 2$, as otherwise $t = 1$, a case already seen. Call the U-cosets of $\mathbb{F}_2^n$ *layers*. We create perfect matchings of each layer separately, and will not modify any finished layers later. Our algorithm consists of 3 phases.

Phase 1: We create perfect matchings in some (less than t) layers in such a way that an even number of vectors will remain in each difference class.

Phase 2: We create perfect matchings in some (less than t) layers in such a way that in each difference class, the number of remaining vectors will be divisible by 2^{k-1}.

Phase 3: All of the remaining differences are used to create homogeneous layers (i.e. layers consisting of differences from only one class).

Phase 1. Let $H = \{\mathbf{u_i} : 2 \leq i \leq t, n_i \equiv 1 \pmod 2\}$, and we use the notation $\mathbf{u} = \mathbf{u_1}$.

We call a subset S of H a *circuit* if its elements are linearly dependent mod $\mathbf{u}$, and this property does not hold for any proper subset of S. A circuit S has *good parity* if the sum of its elements is $|S|\mathbf{u}$, and *bad parity* if the sum of its elements is $(|S|+1)\mathbf{u}$.

As we have $\sum n_i \mathbf{u_i} = \mathbf{0}$, the sum of the elements of H is equal to $n_1\mathbf{u}$, so it is $\mathbf{0}$ mod $\mathbf{u}$. We will apply the following step repeatedly: as long as H is nonempty, we select some of its elements (at least three of them), and we will create a perfect matching of a full layer using one copy of each of the selected vectors, and a suitable number of copies of $\mathbf{u}$. The vectors of H used in this process will be removed from H. When all elements of H are depleted, we move on to Phase 2.

A sequence $(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_i})$ consisting of nonzero vectors in H is *diverse* if $\mathbf{0}$, $\mathbf{v_1}$, $\mathbf{v_1}+\mathbf{v_2}$, $\ldots$, $\mathbf{v_1}+\mathbf{v_2}+\cdots+\mathbf{v_{i-1}}$ are all distinct mod $\mathbf{u}$. If $(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_i})$ is a diverse sequence for which $\mathbf{v_1}+\mathbf{v_2}+\cdots+\mathbf{v_i} = i\mathbf{u}$ holds, then we can make a layer with one copy of each of $\mathbf{v_1}$, $\ldots$, $\mathbf{v_i}$ as differences, and all remaining differences being $\mathbf{u}$.

In each step, take A to be a subset of H of minimal size with $\mathbf{0}$ sum mod $\mathbf{u}$. If A has good parity, then any ordering of it is diverse, and a layer can be created. If it has bad parity, then from $H \setminus A$ we can select another minimal-size subset B with $\mathbf{0}$ sum mod $\mathbf{u}$, which can also be assumed to have bad parity. Then $A \cup B$ can be put into a diverse order (by applications of Lemma 2.2), which will be used to create the layer.

Phase 2. For each $2 \leq i \leq n$, if the number of remaining copies of $\mathbf{u_i}$ has a remainder m_i mod 2^{k-1}, then a single layer containing m_i copies of $\mathbf{u_i}$ and $2^{k-1} - m_i$ copies of $\mathbf{u}$ is made, which is possible by the main conjecture for two difference classes.

Phase 3. As the number of remaining vectors in each class is divisible by 2^{k-1}, this phase can be trivially performed, completing the required perfect matching of $\mathbb{F}_2^n$.

By calculation, it can be seen that altogether we used $\frac{4}{3}(t-1)\cdot 2^{k-1}$ copies of $\mathbf{u}$ during the first two phases, which is less than $\frac{2^{n-1}}{t}$, and so less than n_1, as required. $\square$

3 Perfect matching in the case of many equal vectors

In this chapter, we resolve the main conjecture (for sufficiently large n) in the special case when at least a fraction $\frac{28}{29}$ of the difference vectors are all equal, and the others are arbitrary. So in contrast to the theorem of Balister, Győri and Schelp (see Theorem 1.5), here we do not require that all differences appear an even number of times.

Lemma 3.1. *Let G be a finite abelian group, and let $X \subseteq G$. Then in G, we can select at least $\frac{|G|}{|X|(|X|-1)+1}$ pairwise disjoint translates of X.*

Proof idea. Keep choosing translates of X greedily which are disjoint from the previously chosen ones. $\square$

Remark 3.2. If the group G has exponent 2, the lemma can be improved to say that at least $\frac{|G|}{\binom{|X|}{2}+1}$ pairwise disjoint translates of X can be selected.

Lemma 3.3. *Let $n \geq 2$ and $a \geq t \geq 2$ be integers, for which $\sum_{i=0}^{\lfloor t/2 \rfloor} \binom{a}{i} > 2^n$. Then in $\mathbb{F}_2^n$, among any a vectors one can find at most t which are linearly dependent.*

Proof idea. From the assumption, there exist two distinct subsets of the given vectors of size $\leq \left\lfloor \frac{t}{2} \right\rfloor$ with the same sum. Take the symmetric difference of these two subsets. $\square$

Theorem 3.4. *The main conjecture is true in the case when at least a fraction $\frac{28}{29}$ of the differences are all equal, and n is sufficiently large.*

Proof sketch. Let $\mathbf{u} \in \mathbb{F}_2^n$ appear more than $\frac{28}{29}m$ times among the given differences $\mathbf{d_1}$, $\mathbf{d_2}$, $\ldots$, $\mathbf{d_m}$. (Here $m = \frac{1}{2} \cdot 2^n$.) Let H denote the multiset of vectors $\mathbf{d_i}$ not equal to $\mathbf{u}$. Then $|H| < \frac{1}{29}m$.

We will partition $\mathbb{F}_2^n$ into pairs in the following way. In each step we select some elements $(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_i})$ of the multiset H, and among the elements of $\mathbb{F}_2^n$ not yet used, for each $1 \leq j \leq i$ we select a pair of elements with difference $\mathbf{v_j}$ (so that these pairs are disjoint from each other). The set of elements used in each step will be a union of some $\langle \mathbf{u} \rangle$-cosets; so at the end of the process, after having used all elements of H, all the remaining differences will be equal to $\mathbf{u}$, and these can be assigned to one coset each.

Similarly to the notions used in the proof of Theorem 2.1, define diverse sequences, and circuits and their parity.

If the nonzero vectors $(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_i})$ form a diverse sequence, and $\mathbf{v_1}+\mathbf{v_2}+\cdots+\mathbf{v_i} = i\mathbf{u}$, then we can take i pairwise disjoint vector pairs which use each vector in the given sequence as a difference precisely once, and whose union is equal to the union of some $\langle \mathbf{u} \rangle$-cosets. In each step of the partitioning of $\mathbb{F}_2^n$, we will use such a pattern.

We partition H into circuits by always removing the smallest circuit from it, and then in this partition, we pair up bad circuits according to increasing order of their size. Then similarly to the proof of Theorem 2.1, each good circuit, or pair of bad circuits can be arranged in a diverse order. We will use these diversely-ordered classes in decreasing order of size, always trying to find a translate of the corresponding vector set in $\mathbb{F}_2^n$ that does not contain any previously-used vectors. Classes of size greater than 8 will be called *large*, and otherwise a class is called *small*.

When we selected the circuits C_i in H (always selecting the smallest possible circuit within the remaining vectors), then because of Lemma 3.3, as long as the number of remaining vectors (a) fulfilled the inequality $\binom{a}{0} + \binom{a}{1} + \binom{a}{2} > 2^n$, we always found a circuit of size at most 4, leading to small classes. Therefore the total size of large classes is at most $4 \cdot 2^{n/2}$.

For sets of vectors X corresponding to large classes, by a calculation via Remark 3.2, we can find more disjoint translates of X in $\mathbb{F}_2^n$ then there are previously-used points, hence there will always be a translate of X which is completely unused.

For X corresponding to small classes, the total size of previous classes is less than $\frac{1}{58} \cdot 2^n$, and calculating by Remark 3.2, using the fact that $\frac{|X|}{2} \leq 8$, there will again be a sufficient number of pairwise disjoint translates of X. $\square$

References

[1] Question posed by R. Bacher on MathLinks in 2008, currently available here: https://artofproblemsolving.com/community/c6h183554

[2] Balister, P. N., Győri, E., & Schelp, R. H. (2011). Coloring vertices and edges of a graph by nonempty subsets of a set. European Journal of Combinatorics, 32(4), 533-537.

[3] Correia, D. M., Pokrovskiy, A., & Sudakov, B. (2021). Short proofs of rainbow matching results. arXiv preprint arXiv:2108.07734.

[4] Gao, P., Ramadurai, R., Wanless, I. M., & Wormald, N. (2021). Full rainbow matchings in graphs and hypergraphs. Combinatorics, Probability and Computing, 1-19.

[5] Karasev, R. N., & Petrov, F. V. (2012). Partitions of nonzero elements of a finite field into pairs. Israel Journal of Mathematics, 192(1), 143-156.

[6] Kohen, D., & Sadofschi, I. (2010). A New Approach on the Seating Couples Problem. arXiv:1006.2571

[7] Kohen, D., & Sadofschi, I. (2016). On a generalization of the seating couples problem. Discrete Mathematics, 339(12), 3017-3019.

[8] Mička, O. (2017). Hypercube problems. `http://ktiml.mff.cuni.cz/ gregor/hypercube/lecture17.pdf`

[9] Preissmann, E., & Mischler, M. (2009). Seating Couples Around the King's Table and a New Characterization of Prime Numbers. The American Mathematical Monthly, 116(3), 268-272.