ACTA ARITHMETICA
XCVIII.2 (2001)

The number of $(2,3)$-sum-free subsets of $\{1,\ldots,n\}$

by

Tomasz Schoen (Kiel and Poznań)

1. Introduction. A subset A of a group G is *sum-free* if the equation $x+y=z$ has no solutions in A. Denote by $\mathcal{SF}(G)$ and $\mathcal{SF}(n)$ the family of all sum-free subsets of G and of $\{1,\ldots,n\}\subseteq\mathbb{Z}$, respectively. A well known conjecture of Cameron and Erdős [6] states that

$$|\mathcal{SF}(n)| = O(2^{n/2}). \tag{1}$$

Notice that in view of $\{\lfloor n/2\rfloor+1,\ldots,n\}\in\mathcal{SF}(n)$, if (1) is true it is best possible.

This problem was extensively studied, but in spite of many partial results the conjecture is still open. Alon [1] and Calkin [4] proved that

$$|\mathcal{SF}(n)| = 2^{n/2+o(n)} \tag{2}$$

as $n\to\infty$. Alon also showed that

$$|\mathcal{SF}(G)| = 2^{n/2+o(n)}$$

for any group G of order n. Cameron and Erdős [6] proved that the number of sum-free subsets of $\{\lfloor n/3\rfloor,\ldots,n\}$ is $O(2^{n/2})$. They also observed [7] that it is sufficient to count sum-free sets with at least $n/10$ elements, because $\binom{n}{\lfloor n/10\rfloor}=o(2^{n/2})$.

Deshouillers, Freiman, Sós, and Temkin [8] gave a characterization of dense sum-free sets. Their result implies that there are at most $O(2^{n/2})$ sum-free subsets $A\subseteq\{1,\ldots,n\}$ satisfying

$$|A|\ge 2n/5.$$

On the other hand, Bilu proved in a recent paper [3] that for any fixed $\varepsilon>0$ there are at most $O_\varepsilon(2^{n/2-\varepsilon^2 n/16})$ sum-free sets with

$$|A|\le(1/4-\varepsilon)n. \tag{3}$$

2000 *Mathematics Subject Classification*: 11B75, 11A99.

In the next section we will give a further improvement of the above results, proving that the number of sum-free sets satisfying either

$$|A| < (1/4-\varepsilon)n \quad \text{or} \quad |A| > (1/4+\varepsilon)n$$

is $O_\varepsilon(2^{n/2-\varepsilon^2 n})$.

The importance of Bilu's paper [3] lies not only in (3). He proposed to consider the following modified version of the problem of Cameron and Erdős. For given positive integers $k > l$ call a set $A \subseteq \{1,\ldots,n\}$ (k,l)-*sum-free* if there are no solutions to the equation

$$x_1 + \ldots + x_k = y_1 + \ldots + y_l$$

in A (see [5]). Furthermore, denote by $\mathcal{SF}_k(n)$ the family of all $(k+1,k)$-sum-free subsets of $\{1,\ldots,n\}$. Then clearly, for any $k \geq 3$, we have

$$\mathcal{SF}_k(n) \subseteq \mathcal{SF}_{k-1}(n) \subseteq \ldots \subseteq \mathcal{SF}_2(n) \subseteq \mathcal{SF}(n);$$

but on the other hand, the set of all odd natural numbers less than or equal to n belongs to $\mathcal{SF}_k(n)$, so that we still have

$$|\mathcal{SF}_k(n)| \geq 2^{\lceil n/2\rceil}.$$

One may ask a natural question: Is it true that $|\mathcal{SF}_k(n)| = (1+o(1))2^{\lceil n/2\rceil}$, for every $k \geq 2$? Obviously the problem seems to be "easier" for large k. A theorem of Lev [9] imposes very strong restrictions on the structure of dense $(k+1,k)$-sum-free sets, and (3) can be applied to bound the number of sparse $(k+1,k)$-sum-free sets. Implementing this idea, Bilu was able to prove that

$$|\mathcal{SF}_3(n)| = (1+o(1))2^{\lceil n/2\rceil}.$$

He also conjectured that

$$|\mathcal{SF}_2(n)| = (1+o(1))2^{\lceil n/2\rceil}. \tag{4}$$

The main result of this paper establishes this conjecture by proving the following theorem, which can be viewed as the next step towards the conjecture of Cameron and Erdős.

THEOREM. *There is an absolute positive constant c such that*

$$|\mathcal{SF}_2(n)| = 2^{\lceil n/2\rceil} + O(2^{n/2-cn}).$$

Let us also quote a theorem obtained very recently by Lev, Łuczak, and the present author [11] (for related results and methods see also [10] and [12]).

THEOREM A. *There is an absolute positive constant c' such that for any abelian group G of cardinality $n = |G|$,*

$$|\mathcal{SF}(G)| = (2^{e(G)}-1)2^{n/2} + O(2^{(1/2-c')n}),$$

where $e(G)$ *is the number of even order components in the canonical decomposition of* G *into a direct sum of its cyclic subgroups.*

Notice that the above theorem solves the problem of Cameron–Erdős for any finite abelian group. It will also play a crucial role in the proof of our main result. (The first proof of (4) obtained by the author did not use Theorem A, but was much more complicated.)

Notation. We will use the following notation. For subsets $A,\ B$ of a group put

$$A+B=\{a+b : a\in A,\ b\in B\},$$
$$A-B=\{a-b : a\in A,\ b\in B\}.$$

If $A=\{a\}$ we write $a\pm B$ instead of $A\pm B$. The $O(\ldots)$-symbol without a subscript means that the implied constant is absolute.

2. Sum-free sets. In this section we prove some results related to the conjecture of Cameron–Erdős. We start with a lemma which can be viewed as a generalization of a result of Calkin's (see also Bilu's Lemma 2.3 in [3]).

Lemma 1. *Let* $i,\ l,\ d,\ k,$ *and* $t\in\mathbb{N}$ *be natural numbers, and let*

$$P=\{2i+l-1-(k-1)d,\ldots,2i+l-1,\ldots,2i+l-1+(k-1)d\}$$

be an arithmetic progression with difference d. *Then the number of sets*

$$A\subseteq\{i,i+1,\ldots,i+l-1\}$$

such that

$$(A+A)\cap P=\emptyset \quad \textit{and} \quad |A|=t, \tag{5}$$

where $0\le t\le l/2$, *is less than or equal to* $2^{kd+l/(2k)}\binom{l/2}{t}$.

Proof. Let r and m be non-negative integers such that

$$l=2kdm+r \quad \text{and} \quad 0\le r<2kd.$$

Put

$$I'=\{i,i+1,\ldots,i+kdm-1\},$$
$$J=\{i+kdm,i+kdm+1,\ldots,i+l-kdm-1\},$$
$$I''=\{i+l-kdm,i+l-kdm+1,\ldots,i+l-1\}.$$

We will count the number of sets $A\subseteq\{i,i+1,\ldots,i+l-1\}$ such that $(A+A)\cap P=\emptyset$, $|A|=t$ and $|A\cap J|=t'$, for a fixed t' $(0\le t'\le r)$.

First, we estimate the number of possible sets $A\cap J$. Clearly, $t'\le\lfloor r/2\rfloor$, otherwise one could have $i+kdm+j, i+l-kdm-1-j\in A$ for some $0\le j\le\lfloor r/2\rfloor+1$. To build a set $A\cap J$ we choose a subset S of t' elements from the interval $\{i+kdm,i+kdm+1,\ldots,i+kdm-\lfloor r/2\rfloor-1\}$, which can be done in $\binom{\lfloor r/2\rfloor}{t'}$ ways. Then from every pair $(i+kdm+j, i+l-kdm-1-j)$, $i+kd+j \in S$ we take exactly one element. For this we have $2^{t'}$ choices. Thus, there are at most

$$2^{t'}\binom{\lfloor r/2\rfloor}{t'} \le 2^{kd}\binom{\lfloor r/2\rfloor}{t'} \tag{6}$$

possible sets $A \cap J$ of t' elements.

To count the number of possible sets $A \cap (I' \cup I'')$ we decompose the intervals I' and I'' in the following way:

$$I' = \bigcup_{\substack{0\le u\le d-1\\ 0\le v\le m-1}} P'_{uv},$$

where $P'_{uv} = \{i+u+vkd, i+u+vkd+d, \ldots, i+u+(v+1)kd-d\}$, and

$$I'' = \bigcup_{\substack{0\le u\le d-1\\ 0\le v\le m-1}} P''_{uv},$$

where $P''_{uv} = l+2i-1-P'_{uv}$. Notice that for every $0 \le u \le d-1$ and $0 \le v \le m-1$,

$$P'_{uv} + P''_{uv} = P. \tag{7}$$

Write

$$I' \cup I'' = \bigcup_{\substack{0\le u\le d-1\\ 0\le v\le m-1}} (P'_{uv} \cup P''_{uv}).$$

From (7) it follows that for any set A fulfilling (5), we have either $A \cap P'_{uv} = \emptyset$ or $A \cap P''_{uv} = \emptyset$. Thus every set $A_1 \subseteq I' \cup I''$ with $(A_1 + A_1) \cap P = \emptyset$ is contained in at least one set of the form

$$\bigcup_{\substack{0\le u\le d-1\\ 0\le v\le m-1}} Q_{uv},$$

where Q_{uv} is equal to either P'_{uv} or P''_{uv} for all $0 \le u \le d-1$ and $0 \le v \le m-1$. Notice that there are $2^{md} = 2^{(l-r)/(2k)}$ sets of the above form and each of them has exactly $(l-r)/2$ elements. Therefore there are no more than $2^{l/(2k)}\binom{(l-r)/2}{t-t'}$ choices for the set $A \cap (I' \cup I'')$, so that by (6), the number of sets satisfying (5) and $|A \cap J| = t'$ does not exceed

$$2^{kd+(l-r)/(2k)}\binom{(l-r)/2}{t-t'}\binom{\lfloor r/2\rfloor}{t'}.$$

Furthermore, since

$$\sum_{t'=0}^{\lfloor r/2\rfloor}\binom{(l-r)/2}{t-t'}\binom{\lfloor r/2\rfloor}{t'} \le \binom{l/2}{t},$$

we have at most

$$2^{kd+l/(2k)}\binom{l/2}{t}$$

sets with cardinality t satisfying

$$A \subseteq \{i, i+1, \ldots, i+l-1\} \quad \text{and} \quad (A+A) \cap P = \emptyset.$$

This completes the proof. ∎

To prove the main result of this section we need the following well known theorem of Szemerédi [13]. For a given natural number k denote by $s_k(n)$ the maximum cardinality of a set $A \subseteq \{1, \ldots, n\}$ not containing any arithmetic progression of length k. Then Szemerédi's result states that

$$s_k(n) = o(n) \quad \text{as } n \to \infty. \tag{8}$$

Lemma 2 (see [4]). *Let f be a function such that $f(n) = o(n)$ as $n \to \infty$. Then for any $\varepsilon > 0$ the number of sets $A \subseteq \{1, \ldots, n\}$ satisfying $|A| \le f(n)$ does not exceed $O_{\varepsilon,f}(2^{\varepsilon n})$.*

Theorem 1. *Let ε be a positive constant. Then the number of sum-free subsets of $\{1, \ldots, n\}$ with t elements*, $0 \le t \le n/2$, *is*

$$O_\varepsilon\left(2^{\varepsilon n}\binom{n/2}{t}\right).$$

Proof. Let $k = \lceil 1/\varepsilon \rceil$ and $L = \lceil \sqrt{n} \rceil$. For every $A \subseteq \{1, \ldots, n\}$ denote by $\mu = \mu(A)$ the maximum integer such that $A \cap [\mu, \mu + L - 1]$ contains an arithmetic progression P of length $2k-1$. If such an integer does not exist, put $\mu = 0$.

First of all, we estimate the number of sum-free sets A such that $\mu > 0$ and $|A \cap [l, n]| = t'$ for some fixed t', $0 \le t' \le t$, where l stands for the middle term of the progression P. Then, by Lemma 1, there are at most

$$2^{kd+l/(2k)}\binom{(l-1)/2}{t-t'} \le 2^{L+l/(2k)}\binom{l/2}{t-t'} \tag{9}$$

subsets $A' \subseteq \{1, \ldots, l-1\}$ such that $(A' + A') \cap P = \emptyset$. Moreover, since $A \cap [l, n]$ does not contain any arithmetic progression of length $2k-1$, one can deduce from Szemerédi's theorem that

$$|A \cap [l, n]| \le s_{2k-1}(n-l) \le s_{2k-1}(n).$$

Thus, by (8) and Lemma 2, there are no more than

$$O_\varepsilon(2^{\varepsilon n/4}) \tag{10}$$

possible sets $A \cap [l, n]$. Finally, as we have no more than n^2 ways to choose P, combining (9) and (10) we see that the number of sum-free subsets of

$\{1,\dots,n\}$ with t elements does not exceed

$$\sum_{l=1}^{n}\sum_{t'=0}^{s_{2k-1}(n)} n^2 2^{L+l/(2k)}\binom{l/2}{t-t'}O_\varepsilon(2^{\varepsilon n/4})$$
$$\le n^4 2^{L+n/(2k)}\binom{n/2}{t-t'}O_\varepsilon(2^{\varepsilon n/4}) \le O_\varepsilon(2^{3\varepsilon n/4})\binom{t'+n/2}{t}$$
$$\le O_\varepsilon(2^{3\varepsilon n/4})2^{t'}\binom{n/2}{t} = O_\varepsilon(2^{\varepsilon n})\binom{n/2}{t}.$$

To complete the proof, we have to estimate the number of sets A which do not contain any arithmetic progression of length $2k-1$ in an interval of length L. Again, by (8), we have $t=|A|\le s_{2k-1}(n)$. Thus, by Lemma 2 there are at most $O_{\varepsilon,k}(2^{\varepsilon n}) = O_\varepsilon(2^{\varepsilon n})\binom{n/2}{t}$ sum-free sets with t elements. ∎

We will also use the following corollary from Chernoff's inequality (see Theorem A.1 of [2]):

$$\sum_{t=0}^{\lfloor(1/2-\varepsilon)n\rfloor}\binom{n}{t} \le 2^{n/2-2\varepsilon^2 n/\ln 2}. \tag{11}$$

THEOREM 2. *Let ε be a positive constant. Then the number of sum-free subsets $A\subseteq\{1,\dots,n\}$ such that either*

$$|A|<(1/4-\varepsilon)n \quad or \quad |A|>(1/4+\varepsilon)n$$

is $O_\varepsilon(2^{n/2-\varepsilon^2 n})$.

Proof. By Theorem 1, the number of sum-free subsets of $\{1,\dots,n\}$ with either at most $(1/4-\varepsilon)n$ or at least $(1/4+\varepsilon)n$ elements does not exceed

$$O_\varepsilon(2^{\varepsilon^2 n/3})\sum_{t=0}^{\lfloor(1/4-\varepsilon)n\rfloor}\binom{n/2}{t}.$$

Using (11) one can estimate the last expression by

$$O_\varepsilon(2^{n/2+\varepsilon^2 n/3-\varepsilon^2 n/\ln 2}) = O_\varepsilon(2^{n/2-\varepsilon^2 n}),$$

as claimed. ∎

3. (2, 3)-Sum-free sets. In this section we present the proof of (4), which can be outlined as follows. First, using an elementary argument we will show that for every $A\in\mathcal{SF}_2(n)$ and a suitable choice of $m=m_A$, the set A considered as a subset of $\mathbb{Z}_m$, is sum-free. Then we will see that for almost all (2, 3)-sum-free sets (the exceptional set is of size $o(|\mathcal{SF}_2(n)|)$) we can take $m\sim n$. This will allow us to apply Theorem A, and consequently prove the main result.

LEMMA 3. *Let $A \subseteq \{1, \dots, n\}$ be a $(2,3)$-sum-free set. Suppose that $m \in A + A$ and $m \geq n$. Then A is a sum-free subset of $\mathbb{Z}_m$.*

Proof. Assume that there are $a, a', a'' \in A$ such that $a + a' \equiv a'' \pmod{m}$. Thus, either

$$a + a' = a'' \quad \text{or} \quad a + a' = a'' + m.$$

Clearly, the first equality is not possible, because the set A belongs to $\mathcal{SF}_2(n)$, and consequently is sum-free. Further, since $m = b + b'$ for some $b, b' \in A$, the second equality would yield

$$a + a' = a'' + b + b',$$

contradicting the assumption. ∎

LEMMA 4. *Let δ be a positive constant. There are at most $O_\delta(2^{n/2-\delta n/7})$ sum-free subsets $A \subseteq \{1, \dots, n\}$ such that*

$$(A + A) \cap [n, (1+\delta)n] = \emptyset. \tag{12}$$

Proof. From Alon–Calkin's theorem (2) it follows that the number of sum-free subsets A such that $M := \max A \leq (1 - \delta/3)n$ is $O_\delta(2^{n/2-\delta n/7})$, so that we can assume the opposite, $M > (1 - \delta/3)n$.

Put $n' = n - \lceil \delta n/2 \rceil + 1$ and suppose that $q := \lfloor 2/\delta \rfloor$ is an even natural number (otherwise put $q := \lfloor 2/\delta \rfloor + 1$). Furthermore, let $d \in \mathbb{N}$ be such that $n' = qd + r$, where $0 \leq r < q$. We decompose the set $\{\lceil \delta n/2 \rceil, \lceil \delta n/2 \rceil + 1, \dots, n\}$ into disjoint successive intervals $I_1, \dots, I_q$ such that $|I_i| = d$ for every $i \neq q/2 + 1$ and $|I_{q/2+1}| = r$. Obviously,

$$I_i + I_{q-i} \subseteq [n, (1+\delta)n],$$

so that for every A satisfying (12) we have either $A \cap I_i = \emptyset$ or $A \cap I_{q-i} = \emptyset$, for each i, $0 \leq i \leq q/2$. In particular, the middle interval $I_{q/2+1}$ cannot share any element with A, whence every subset of $\{\lceil \delta n/2 \rceil, \lceil \delta n/2 \rceil + 1, \dots, n\}$ satisfying (12) is contained in a set of the form

$$\bigcup_{i=1}^{q/2} I'_i,$$

where $I'_i = I_i$ or I_{q-i}, for each $1 \leq i \leq q/2$. There are $2^{q/2}$ choices for the above sets, and each contains exactly $(n' - r)/2$ elements. Thus, the number of subsets $A \cap [\lceil \delta n/2 \rceil, n]$ fulfilling $(A + A) \cap [n, (1+\delta)n] = \emptyset$ does not exceed

$$2^{q/2} 2^{(n'-r)/2} \leq 2^{1/\delta} 2^{n'/2}.$$

Notice that $M > (1 - \delta/3)n$ implies $A \cap [\delta n/3, \lceil \delta n/2 \rceil - 1] = \emptyset$, so that there are at most $2^{\delta n/3}$ possible sets $A \cap [1, \lceil \delta n/2 \rceil]$. Since there are no more than n choices for M, we have at most

$$n 2^{1/\delta} 2^{n'/2 + \delta n/3} = O_\delta(2^{n/2-\delta n/7})$$

sets with the required property. ∎

Proof of Theorem. The main term $2^{\lceil n/2\rceil}$ is given by all subsets of $\{1,3,\dots,2\lceil n/2\rceil-1\}$, so it is sufficient to estimate $|\mathcal{SF}'_2(n)|$, the number of $A\in\mathcal{SF}_2(n)$ not contained in the set of odd numbers.

In view of Lemma 4, it is enough to estimate the number of sets $A\in\mathcal{SF}_2(n)$ such that

$$(A+A)\cap[n,(1+\delta)n]\neq\emptyset,$$

where $\delta:=14c'/9$, and c' is given by Theorem A. Let

$$m=m_A:=\min((A+A)\cap[n,(1+\delta)n]),$$

and observe that Lemma 3 gives $A\in\mathcal{SF}'(\mathbb{Z}_m)$, where the family $\mathcal{SF}'(\mathbb{Z}_m)$ consists of all sum-free subsets of $\mathbb{Z}_m$ which contain at least one even element.

A consequence of Theorem A, applied to $G=\mathbb{Z}_m$, is

$$|\mathcal{SF}'(\mathbb{Z}_m)|=O(2^{m/2-c'm}). \tag{13}$$

Indeed, if m is odd then $e(\mathbb{Z}_m)=0$, and by Theorem A one has

$$|\mathcal{SF}'(\mathbb{Z}_m)|\le|\mathcal{SF}(\mathbb{Z}_m)|=O(2^{m/2-c'm}).$$

If m is even then $e(\mathbb{Z}_m)=1$, and Theorem A gives

$$|\mathcal{SF}(\mathbb{Z}_m)|=2^{m/2}+O(2^{m/2-c'm}).$$

On the other hand, every subset of $\{0,\dots,m-1\}$ consisting of odd integers is sum-free in $\mathbb{Z}_m$. Hence

$$|\mathcal{SF}(\mathbb{Z}_m)|=2^{m/2}+|\mathcal{SF}'(\mathbb{Z}_m)|,$$

and (13) follows.

Thus, by Lemma 4 and (13), we have

$$\begin{aligned}|\mathcal{SF}'_2(n)|&=O\Big(2^{n/2-2c'n/9}+\sum_{m=n}^{\lfloor(1+\delta)n\rfloor}|\mathcal{SF}'(\mathbb{Z}_m)|\Big)\\&=O(2^{n/2-2c'n/9}+n2^{(n/2-c'n)(1+14c'/9)})=O(2^{n/2-2c'n/9}),\end{aligned}$$

and

$$|\mathcal{SF}_2(n)|=2^{\lceil n/2\rceil}+O(2^{n/2-2c'n/9}),$$

which completes the proof. ■

References

[1] N. Alon, *Independent sets in regular graphs and sum-free subsets of finite groups*, Israel J. Math. 73 (1991), 247–256.

[2] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, New York, 1992.

[3] Y. Bilu, *Sum-free sets and related sets*, Combinatorica 18 (1998), 449–459.

[4] N. J. Calkin, *On the number of sum-free sets*, Bull. London Math. Soc. 22 (1990), 141–144.
[5] N. J. Calkin and J. M. Thomson, *Counting generalized sum-free sets*, J. Number Theory 68 (1996), 151–159.
[6] P. J. Cameron and P. Erdős, *On the number of sets of integers with various properties*, in: Number Theory, R. A. Mollin (ed.), de Gruyter, Berlin, 1990, 61–79.
[7] —, —, *Notes on sum-free and related sets*, Combin. Probab. Comput. 8 (1999), 95–107.
[8] J.-M. Deshouillers, G. Freiman, V. Sós and M. Temkin, *On the structure of sum-free sets*, *2*, Astérisque 258 (1999), 149–161.
[9] V. F. Lev, *Structure theorem for multiple addition and the Frobenius problem*, J. Number Theory 58 (1996), 79–88.
[10] V. F. Lev and T. Schoen, *Cameron–Erdős modulo a prime*, submitted.
[11] V. F. Lev, T. Łuczak and T. Schoen, *Sum-free sets in abelian groups*, submitted.
[12] T. Łuczak and T. Schoen, *On the number of maximal sum-free sets*, Proc. Amer. Math. Soc., to appear.
[13] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.

Mathematisches Seminar
Universität zu Kiel
Ludewig-Meyn-Str. 4
24098 Kiel, Germany
E-mail: tos@numerik.uni-kiel.de

Department of Discrete Mathematics
Adam Mickiewicz University
Matejki 48/49
60-769 Poznań, Poland

Received on 3.4.2000
and in revised form on 27.5.2000 (3793)