Centrum voor Wiskunde en Informatica

REPORT *RAPPORT*

PNA

Probability, Networks and Algorithms

Probability, Networks and Algorithms

Balanced subset sums of dense sets of integers

G. Károlyi

REPORT PNA-R0801 JANUARY 2008

Centrum voor Wiskunde en Informatica (CWI) is the national research institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organisation for Scientific Research (NWO).
CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

ISSN 1386-3711

Balanced subset sums of dense sets of integers

ABSTRACT

Given n different positive integers not greater than 2n-2, we prove that more than n^2/12 consecutive integers can be represented as the sum of half of the given numbers. This confirms a conjecture of Lev.

2000 Mathematics Subject Classification: 11B75
Keywords and Phrases: subset sum problem
Note: This work was carried out under project PNA1-Spinoza Award

BALANCED SUBSET SUMS IN DENSE SETS OF INTEGERS

GYULA KÁROLYI[1] Department of Algebra and Number Theory, Eötvös University, Pázmány P. sétány 1/C, Budapest, H–1117 Hungary

ABSTRACT. Let $1 \le a_1 < a_2 < \ldots < a_n \le 2n-2$ denote integers. We prove that there exist $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ such that $|\varepsilon_1 + \ldots + \varepsilon_n| \le 1$ and $|\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n| \le 1$, at least when n is large enough. This result is sharp and, in turn, confirms a conjecture of V.F. Lev. We also prove that more than $n^2/12$ consecutive integers can be reperesented as the sum of roughly $n/2$ elements of the sequence.

1. INTRODUCTION

At the Workshop on Combinatorial Number Theory held at DIMACS, 1996, V.F. Lev proposed the following problem. Suppose that $1 \le a_1 < a_2 < \ldots < a_n \le 2n-1$ are integers such that their sum $\sigma = \sum_{i=1}^{n} a_i$ is even. Does there always exist $I \subset \{1, 2, \ldots, n\}$ such that $\sum_{i \in I} a_i = \sigma/2$? The answer is in the affirmative if n is large enough. Note that such a restriction has to be imposed on n, since the sequences $(1, 4, 5, 6)$ and $(1, 2, 3, 9, 10, 11)$ provide counterexamples otherwise. In fact, it follows from a result of Lev [3], that if n is large enough, then every integer in the interval $[560n, \sigma - 560n]$ can be expressed as the sum of

[1]Visiting the CWI in Amsterdam. Research partially supported by Hungarian Scientific Research Grants OTKA T043631 and K67676.

different a_i's, see [1]. In this paper we prove the following much stronger version of Lev's conjecture.

Theorem 1. *Let $1 \leq a_1 < a_2 < \ldots < a_n \leq 2n-1$ denote integers such that $a_{\nu+1} - a_\nu = 1$ holds for at least one index $1 \leq \nu \leq n-1$. If $n \geq n_0 = 89$, then there exist $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ such that $|\varepsilon_1 + \ldots + \varepsilon_n| \leq 1$ and $|\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n| \leq 1$.*

Corollary 2. *Let $1 \leq a_1 < a_2 < \ldots < a_n \leq 2n-2$ denote integers. If $n \geq 89$, then there exist $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ such that $|\varepsilon_1 + \ldots + \varepsilon_n| \leq 1$ and $|\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n| \leq 1$.*

Note that although most likely the condition $n \geq 89$ can essentially be relaxed, it is not merely technical. The sequence $(1, 2, 3, 8, 9, 10, 14, 15)$ demonstrates that Theorem 1 is not valid with $n_0 = 8$. An other formulation of the condition in the above theorem is the requirement that there exists an index $1 \leq \nu \leq n$ such that a_ν is even.

Now the conjecture of Lev, assumed that $n \geq 89$, follows immediately from the above Theorem, unless $a_i = 2i - 1$ for $1 \leq i \leq n$. Even in that case, it is easy to check that Theorem 1 remains valid if $n \equiv 0,\ 1$ or $3 \pmod 4$. This is not the case, however, if $n \equiv 2 \pmod 4$.

Indeed, let $n = 4k+2$ and suppose that $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ such that $|\varepsilon_1 + \ldots + \varepsilon_n| \leq 1$. Consider $I = \{1 \leq i \leq n \mid \varepsilon_i = +1\}$, then $|I| = 2k+1$. Therefore $A = \sum_{i \in I} a_i$ and $B = \sum_{i \notin I} a_i$ are odd numbers. However, $A + B = \sum_{i=1}^n a_i = (4k+2)^2$ is divisible by 4, hence $A - B \equiv 2 \pmod 4$, and $|\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n| = |A - B| \geq 2$. Nevertheless, choosing

$$I = \{1, 2, 3, 5\} \cup \bigcup_{i=2}^{k} \{4i, 4i+1\} \subseteq \{1, 2, \ldots, n\}$$

we find that

$$\sum_{i \in I} a_i = \frac{1}{2} \sum_{i=1}^{n} a_i = \frac{\sigma}{2}\ ,$$

confirming the conjecture of Lev in this remaining case, too.

The method of the proof of Theorem 1 allows us to obtain the following generalization.

Theorem 3. *For every $\varepsilon > 0$ there is an integer $n_0 = n_0(\varepsilon)$ with the following property. If $n \geq n_0$, $1 \leq a_1 < a_2 < \ldots < a_n \leq 2n-2$ are integers, and N is an* *integer such that* $|N| \le (\frac{9}{100} - \varepsilon)n^2$*, then there exist* $\varepsilon_1, \dots, \varepsilon_n \in \{-1, +1\}$ *such that* $|\varepsilon_1 + \ldots + \varepsilon_n| \le 1$ *and* $|\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n - N| \le 1$.

Consequently, every integer in a long interval can be expressed as a 'balanced' subset sum:

Corollary 4. *If* n *is large enough and* $1 \le a_1 < a_2 < \ldots < a_n \le 2n-2$ *are integers, then for every integer*

$$k \in [\sigma/2 - n^2/24, \sigma/2 + n^2/24]$$

there exists a set of indices $I \subset \{1, 2, \dots, n\}$ *such that* $|I| \in \{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}$ *and* $\sum_{i \in I} a_i = k$.

Proof. We apply Theorem 3 with $\varepsilon = 9/100 - 1/12$. If $k = \sigma/2 + x$ is an integer in the prescribed interval, then for the integer $N = 2x$ there exist $\varepsilon_1, \dots, \varepsilon_n \in \{-1, +1\}$ such that $|\varepsilon_1 + \ldots + \varepsilon_n| \le 1$ and $|\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n - N| \le 1$. Since $N = 2x \equiv \sigma \equiv \varepsilon_1 a_1 + \ldots + \varepsilon_n a_n \pmod 2$, it follows that $\varepsilon_1 a_1 + \ldots + \varepsilon_n a_n = N$, and with $I = \{i \mid \varepsilon_i = +1\}$ we have $|I| \in \{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}$ and

$$\sum_{i \in I} a_i = \frac{1}{2}\Big(\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} \varepsilon_i a_i\Big) = \frac{\sigma}{2} + x = k.$$

□

Note that all these results can be extended to less dense sequences under the assumption that the sequence contains sufficiently many small gaps. We do not elaborate on this here.

Finally we note that if balancedness is not required, then the following ultimate result is now available, see [1].

Theorem 5. *Let* $1 \le a_1 < a_2 < \ldots < a_n \le \ell \le 2n-6$ *denote integers. If* n *is large enough, then every integer in the interval*

$$[2\ell - 2n + 1, \sigma - (2\ell - 2n + 1)]$$

can be expressed as the sum of different a_i*'s. Neither the length of this interval can be extended, nor the condition imposed on* ℓ *can be relaxed.*

2. The Proof of Theorem 1

First we note that it is enough to prove Theorem 1 when n is an even number. Indeed, let n be odd, and assume that the statement has been proved for $n+1$. Consider the sequence

$$b_1 = 1 < b_2 = a_1 + 1 < \ldots < b_{n+1} = a_n + 1 < 2(n+1) - 1.$$

There exist $\eta_1, \ldots, \eta_{n+1} \in \{-1, +1\}$ such that,

$$|\eta_1 + \ldots + \eta_{n+1}| \leq 1 \quad \text{and} \quad |\eta_1 b_1 + \ldots + \eta_{n+1} b_{n+1}| \leq 1.$$

Since $n+1$ is even, it follows that $\eta_1 + \ldots + \eta_{n+1} = 0$. Let $\varepsilon_i = \eta_{i+1}$, then $|\varepsilon_1 + \ldots + \varepsilon_n| = |-\eta_1| = 1$, and

$$\Big|\sum_{i=1}^{n} \varepsilon_i a_i\Big| = \Big|\sum_{i=1}^{n} \eta_{i+1} a_i + \sum_{i=1}^{n+1} \eta_i\Big| = \Big|\sum_{i=1}^{n+1} \eta_i b_i\Big| \leq 1.$$

Accordingly, we assume that $n = 2m$ with an integer $m \geq 45$. To illustrate the initial idea of the proof, consider the differences $e_i = a_{2i} - a_{2i-1}$ for $i = 1, 2, \ldots, m$. If we found $\delta_1, \ldots, \delta_m \in \{-1, +1\}$ such that $|\sum_{i=1}^{m} \delta_i e_i| < 2$, then the choice $\varepsilon_{2i} = \delta_i$, $\varepsilon_{2i-1} = -\delta_i$ would clearly give the desired result. This is the case, in fact, when $\sum_{i=1}^{m} e_i \leq 2m - 2$, as it can be easily derived from the following two simple lemmas. They are formulated so that their application is not limited to integer sequences.

Lemma 6. *Let $e_1, \ldots, e_k \geq 1$ and suppose that*

$$E = \sum_{i=1}^{k} e_i \leq \beta k - (\beta^2 - \beta)$$

for some positive real number β. Then

$$\sum_{e_i < s+1} e_i \geq s$$

holds for every positive integer $\beta - 1 \leq s \leq k - \beta$.

Proof. If s is a positive integer then, obviously,

$$\sum_{e_i < s+1} e_i \geq \sum_{e_i < s+1} 1 = k - \sum_{e_i \geq s+1} 1 \geq k - \frac{E}{s+1} \ .$$

As long as

$$(k-1)^2 - 4(E-k) \geq (k-\alpha)^2 \ , \tag{1}$$

we have

$$k - \frac{E}{s+1} \geq s$$

for every $(\alpha-1)/2 \leq s \leq k-(\alpha+1)/2$. To complete the proof we only have to notice that (1) is satisfied if $\alpha = 2\beta - 1$. □

Lemma 7. *Let $e_1, \ldots, e_k \geq 1$ and suppose that*

$$\sum_{e_i < s+1} e_i \geq s \tag{2}$$

holds for every integer $1 \leq s \leq \max\{e_i \mid 1 \leq i \leq k\}$. Let F be any number such that

$$|F| < \sum_{i=1}^{k} e_i + 2 \ . \tag{3}$$

Then there exist $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, +1\}$ such that

$$\Big|\sum_{i=1}^{k} \varepsilon_i e_i - F\Big| < 2 \ ,$$

in particular $F = \sum_{i=1}^{k} \varepsilon_i e_i$ if the e_i's are integers and $F \equiv \sum_{i=1}^{k} e_i \pmod 2$.

Proof. Without loss of generality, we may suppose that that $e_1 \geq e_2 \geq \ldots \geq e_k$, then $e_k < 2$. The point is, that the condition allows us to construct $\varepsilon_1, \ldots, \varepsilon_k$ sequentially so that the sequence of partial sums $\sum_{j=1}^{i} \varepsilon_j e_j$ oscillates about F with smaller and smaller amplitude, until it eventually approximates F with the desired accuracy.

More precisely, let $\Delta_0 = F$, and define ε_n and Δ_n recursively as follows. Let, for $n = 1, 2, \ldots, k$,

$$\varepsilon_n = \begin{cases} 1 & \text{if } \Delta_{n-1} \geq 0 \\ -1 & \text{if } \Delta_{n-1} < 0 \end{cases}$$

and let $\Delta_n = \Delta_{n-1} - \varepsilon_n e_n$, then

$$\Delta_n = F - \varepsilon_1 e_1 - \varepsilon_2 e_2 - \ldots - \varepsilon_n e_n$$

for every $0 \leq n \leq k$. We prove, by induction, that

$$|\Delta_n| < e_{n+1} + \ldots + e_{k-1} + e_k + 2 \tag{4}$$

for $n = 0, 1, \ldots, k$.

This is true for $n = 0$. Thus, let $1 \leq n \leq k$, and suppose that (4) is satisfied with $n-1$ in place of n. Assume, w.l.o.g, that $\Delta_{n-1} \geq 0$. Then, by definition,

$$-e_n \leq \Delta_n = \Delta_{n-1} + (-1)e_n < e_{n+1} + \ldots + e_k + 2 \ .$$

Thus, to verify (4), it suffices to show that $e_n < e_{n+1} + \ldots + e_k + 2$. This is definitely true, if $e_{n+1} = e_n$ or $n = k$. Otherwise we can write

$$\sum_{i=n+1}^{k} e_i = \sum_{e_i < e_n} e_i \geq \sum_{e_i < \lfloor e_n \rfloor} e_i \geq \lfloor e_n \rfloor - 1 > e_n - 2 \ ,$$

proving the assertion. Letting $n = k$ in (4), the statement of the lemma follows. □

The main idea of the proof of Theorem 1 is to find a partition

$$\{a_1, a_2, \ldots, a_n\} = \bigcup_{i=1}^{k} \{x_i, y_i\} \cup \{z_1, \ldots, z_{n-2k}\} \tag{5}$$

such that $e_i = x_i - y_i$ $(1 \leq i \leq k)$ and $F = \sum_{i=1}^{n-2k} (-1)^i z_i$ satisfy the conditions of Lemma 7. Then Theorem 1 follows immediately.

To achieve this we will construct the above partition so that

$$\sum_{i=1}^{k} e_i \leq 4k - 12 \quad (\text{or} \quad \sum_{i=1}^{k} e_i \leq 3k - 6), \tag{6}$$

$$e_i \leq k - 4 \quad (\text{or} \quad e_i \leq k - 3) \quad \text{for} \quad i = 1, 2, \ldots, k \ , \tag{7}$$

$$|F| \leq k + 1 \ , \quad \text{and} \tag{8}$$

$$\sum_{e_i \leq s} e_i \geq s \quad \text{if} \quad s = 1 \quad \text{or} \quad s = 2 \ . \tag{9}$$

Then an application of Lemma 6 with $\beta = 4$ (or with $\beta = 3$) will show that e_i $(1 \leq i \leq k)$ and F satisfy the conditions of Lemma 7. More precisely, it follows from (6) and (9) that condition (2) holds for $s \leq k - \beta$, hence for every integer $1 \leq s \leq \max\{e_i \mid 1 \leq i \leq k\}$ in view of (7). Finally, (3) follows from (8), given that $\sum_{i=1}^{k} e_i \geq k$. Therefore, once we found a partition (5) with properties (6)–(9), the proof of Theorem 1 will be complete.

First we take care of the condition (9). If we take $x_k = a_{\nu+1}$ and $y_k = a_\nu$, then $e_k = 1$. Moreover, since

$$\sum_{i=1}^{n-1} (a_{i+1} - a_i) \leq 2n - 2,$$

there must be an index $\mu \notin \{\nu-1, \nu, \nu+1, n\}$, such that $a_{\mu+1} - a_\mu \leq 2$. Taking $x_{k-1} = a_{\mu+1}$ and $y_{k-1} = a_\mu$, condition (9) will be satisfied. Enumerating the remaining $n-4$ elements of the sequence (a_i) as

$$1 \leq b_1 < b_2 < \ldots < b_{2m-4} \leq 4m-1,$$

with $f_i = b_{2i} - b_{2i-1}$ we find that

$$\sum_{i=1}^{m-2} f_i = \sum_{i=1}^{m-2} (b_{2i} - b_{2i-1}) \leq (4m-2) - (m-3) = 3m+1. \tag{10}$$

Since $m > 21$, there cannot be 3 different indices i with $f_i \geq m-5$. We distinguish between three cases.

Case 1) If $f_i \leq m-6$ for $1 \leq i \leq m-2$, then we can choose $k = m$, $F = 0$. Taking $x_i = b_{2i}$ and $y_i = b_{2i-1}$ for $1 \leq i \leq k-2$, conditions (7) and (8) are obviously satisfied, whereas (6) follows easily form (10):

$$\sum_{i=1}^{k} e_i \leq \sum_{i=1}^{m-2} f_i + 3 \leq 3m+4 \leq 4m-12,$$

given that $m \geq 16$.

Case 2) There exist indices u, v such that $m-5 \leq f_u \leq f_v$. In view of (10) we have $f_u + f_v \leq (3m+1) - (m-4) = 2m+5$, and consequently $m-5 \leq f_u \leq f_v \leq m+10$ and $0 \leq f_v - f_u \leq 15$. Therefore we may choose $k = m-2$, $z_1 = b_{2v-1}$, $z_2 = b_{2v}$, $z_3 = b_{2u}$, $z_4 = b_{2u-1}$. Constructing x_i, y_i $(1 \leq i \leq m-4)$ from the remaining elements of the sequence (b_i) in the obvious way we find that $|F| \leq 15 < m-2 = k$, each e_i satisfies $e_i \leq m-6 = k-4$, and once again (10) gives

$$\sum_{i=1}^{k} e_i \leq \sum_{i=1}^{m-2} f_i - 2(m-5) + 3 \leq m+14 < 4m-20 = 4k-12.$$

Case 3) There exists exactly one index u with $m-5 \leq f_u$. From (10) it follows that $f_u \leq (3m+1) - (m-3) = 2m+4$. We claim that there exist indices v, w different from u such that

$$|b_{2w} + b_{2w-1} - b_{2v} - b_{2v-1} - f_u| \leq m-2. \tag{11}$$

In that case we can choose $k = m-3$ and $z_1 = b_{2u}$, $z_2 = b_{2u-1}$, $z_3 = b_{2v}$, $z_4 = b_{2w}$, $z_5 = b_{2w-1}$, $z_6 = b_{2u-1}$ to have $|F| \leq m-2 = k+1$. Constructing

x_i, y_i $(1 \le i \le m-4)$ from the remaining elements of the sequence (b_i) in the obvious way this time we find that each e_i satisfies $e_i \le m-6 = k-3$, and

$$\sum_{i=1}^{k} e_i \le \sum_{i=1}^{m-2} f_i - (m-5) - 2 + 3 \le 2m+7 \le 3m-15 = 3k-6.$$

It only remains to prove the above claim. The idea is to find v, w such a way that f_v, f_w are small and at the same time $b_{2w} - b_{2v}$ lies in a prescribed interval that depends on the size of f_u. It turns out that the optimum strategy for such an approach is the following. First, for any positive integer $\kappa \ge 2$, introduce

$$I_\kappa = \{i \mid 1 \le i \le m-2,\ i \ne u,\ f_i \le \kappa\}.$$

Denote by x the number of indices $i \ne u$ for which $f_i > \kappa$. Then

$$(m-3-x) + (\kappa+1)x \le \sum_{i=1}^{m-2} f_i - f_u \le (3m+1) - (m-5) = 2m+6.$$

Thus, $\kappa x \le m+9$, and $m-3-x \ge (1-1/\kappa)m - 3 - 9/\kappa$. We have proved

Claim 8. $|I_\kappa| \ge \dfrac{\kappa-1}{\kappa} m - \dfrac{9}{\kappa} - 3$. *In particular* $t = |I_7| \ge \dfrac{6m-30}{7}$.

Write $c_0 = 0$ and let

$$\bigcup_{i \in I_7} \{b_{2i-1}, b_{2i}\} = \{c_1 < c_2 < \ldots < c_{2t-1} < c_{2t}\}.$$

Now we separate two subcases as follows.

Case 3A) $m-5 \le f_u \le 2m-14$. We will prove that there exist $1 \le i < j \le t$ such that

$$\frac{m}{2} - 3 \le \Delta_{i,j} = c_{2j} - c_{2i} \le m-7. \tag{12}$$

Since we have

$$1 \le c_{2i} - c_{2i-1}, c_{2j} - c_{2j-1} \le 7, \tag{13}$$

we can argue that

$$m-12 \le 2\Delta_{i,j} - 6 \le c_{2j} + c_{2j-1} - c_{2i} - c_{2i-1} \le 2\Delta_{i,j} + 6 < 2m-7,$$

and that implies (11). If there exists $1 \le i \le t-1$ such that

$$\frac{m}{2} - 3 \le c_{2i+2} - c_{2i} \le m-7,$$

then (12) is immediate. Otherwise we have

$$c_{2i+2} - c_{2i} \le \frac{m}{2} - \frac{7}{2} \quad \text{or} \quad c_{2i+2} - c_{2i} \ge m-6$$

for every integer $1 \leq i \leq t-1$. This way we distinguish between 'small gaps' and 'large gaps' in the sequence $c_2, c_4, \ldots, c_{2t}$. The large gaps partition this sequence into 'blocks', where the gap between two consecutive elements within a block is always small. For such a block $B = (c_{2i}, c_{2i+2}, \ldots, c_{2i'})$, the quantity $\ell(B) = 2(i' - i)$ we call the length of the block. Since

$$2 \cdot \left(\frac{m}{2} - \frac{7}{2}\right) < m - 6,$$

in order to have a pair i, j with (12), it is enough to prove that at least one block has a length $\geq m/2 - 3$. Then the smallest integer j satisfying $c_{2j} - c_{2i} \geq m/2 - 3$ will do the job.

We claim that there cannot be more than 3 blocks. Indeed, since every gap is at least 2, were there 3 or more large gaps, we would find that

$$\begin{aligned} 4m - 1 &\geq \sum_{i=0}^{t-1} (c_{2i+2} - c_{2i}) \geq 3(m-6) + (t-3)2 \\ &\geq 3m - 18 + 2\left(\frac{6m-30}{7} - 3\right), \end{aligned}$$

implying $m \leq 221/5 < 45$, a contradiction.

Since there are at most 3 blocks, one must contain at least $t/3$ different c_{2i}'s, and thus its length

$$\ell(B) \geq 2\left(\frac{t}{3} - 1\right) \geq \frac{4m-20}{7} - 2.$$

Given that $m \geq 26$ we conclude that indeed $\ell(B) \geq m/2 - 3$.

Case 3B) $2m - 13 \leq f_u \leq 2m + 4$. This time we prove that

$$\frac{m}{2} + 6 \leq \Delta_{i,j} \leq \frac{3}{2}m - \frac{21}{2} \tag{14}$$

holds with suitable $1 \leq i < j \leq t$. In view of (13) this implies

$$m + 6 \leq 2\Delta_{i,j} - 6 \leq c_{2j} + c_{2j-1} - c_{2i} - c_{2i-1} \leq 2\Delta_{i,j} + 6 \leq 3m - 15,$$

and from that (11) follows. Similarly to the previous case, we may assume that there are only small and large gaps, which in this case means that

$$c_{2i+2} - c_{2i} \leq \frac{m}{2} + \frac{11}{2} \quad \text{or} \quad c_{2i+2} - c_{2i} \geq \frac{3}{2}m - 10$$

holds for every integer $1 \leq i \leq t-1$. Given that (here we use $m \geq 44$)

$$2 \cdot \left(\frac{m}{2} + \frac{11}{2}\right) < \frac{3}{2}m - 10,$$

it suffices to prove that there is a block B with $\ell(B) \geq m/2 + 6$.

Were there 2 or more large gaps, we would find that

$$\begin{aligned} 4m-1 &\geq \sum_{i=0}^{t-1}(c_{2i+2}-c_{2i}) \geq 2\Big(\frac{3}{2}m-10\Big)+(t-2)2 \\ &\geq 3m-20+2\Big(\frac{6m-30}{7}-2\Big), \end{aligned}$$

implying $m \leq 221/5 < 45$, a contradiction. Therefore there are at most 2 blocks, one of which containing at least $t/2$ different c_{2i}'s. The length of that block thus satisfies

$$\ell(B) \geq 2\Big(\frac{t}{2}-1\Big) \geq \frac{6m-30}{7}-2.$$

Since $m \geq 172/5$, we find that $\ell(B) \geq m/2+6$, and the proof is complete.

3. The Proof of Theorem 3

Obviously we may assume that $\varepsilon > 0$ is small enough so that all the below arguments work. We fix such an ε and assume that n is large enough. As in the proof of Theorem 1, we may assume that $n = 2m$ is an even number. Put $c = 1/5 - 2\varepsilon$. We will prove that there exists an integer $k \geq (1-c)m - 7$ and a partition in the form (5) such that for $e_i = x_i - y_i$ $(1 \leq i \leq k)$ and $F = N + \sum_{i=1}^{n-2k}(-1)^i z_i$ the following conditions hold:

$$\sum_{i=1}^{k} e_i \leq 4k-12, \tag{15}$$

$$e_i \leq (1-c)m - 11 \leq k-4 \quad \text{for} \quad i = 1,2,\ldots,k\ , \tag{16}$$

$$|F| \leq (1-c)m - 6 \leq k+1\ , \quad \text{and} \tag{17}$$

$$\sum_{e_i \leq s} e_i \geq s \quad \text{if} \quad s=1 \quad \text{or} \quad s=2\ . \tag{18}$$

As in the proof of Theorem 1, we can apply Lemma 6 with $\beta = 4$, and then Lemma 7 gives the result.

Clearly there exist $1 \leq \mu, \nu \leq n-1$, $\mu \notin \{\nu-1, \nu, \nu+1\}$ such that $a_{\nu+1} - a_\nu = 1$ and $a_{\mu+1} - a_\mu \leq 2$. Putting $x_1 = a_{\nu+1}$, $y_1 = a_\nu$, $x_2 = a_{\mu+1}$, $y_2 = a_\mu$ then takes care of (18). Enumerate the remaining $n-4$ elements of the sequence (a_i) as

$$1 \leq b_1 < b_2 < \ldots < b_{2m-4} \leq 4m-2.$$

Take $q = \lceil cm \rceil$. Since

$$\begin{aligned}\sum_{i=1}^{q}(b_{2m-3-i}-b_i) &\geq \sum_{i=1}^{q}(2m-2i-3) = 2qm - q(q+4)\\ &> 2cm^2-(cm+1)(cm+5) = (2c-c^2)m^2-(6cm+5)\\ &> \Big(\frac{9}{25}-\frac{16}{5}\varepsilon-4\varepsilon^2\Big)m^2-2m > \Big(\frac{9}{25}-4\varepsilon\Big)m^2 \geq |N|\end{aligned}$$

and $b_{2m-3-i}-b_i \leq 4m-3$ for every i, there exists an integer $0 \leq r < cm+1$ such that

$$\Big|N - \operatorname{sgn}(N)\sum_{i=1}^{r}(b_{2m-3-i}-b_i)\Big| \leq 2m-2,$$

where $\operatorname{sgn}(N) = +1$, if $N \geq 0$ and $\operatorname{sgn}(N) = -1$, if $N < 0$. Consider

$$r+1 \leq b_{r+1} < b_{r+2} < \ldots < b_{2m-4-r} \leq 4m-2-r,$$

and let $f_i = b_{r+2i} - b_{r+2i-1}$ for $1 \leq i \leq m-2-r$, then

$$\sum_{i=1}^{m-r-2} f_i \leq \big((4m-2-r)-(r+1)\big)-(m-r-3) \leq 3m. \tag{19}$$

Were there 3 or more indices i with $f_i > (1-c)m - 11$, it would imply

$$\sum_{i=1}^{m-r-2} f_i > 3\big((1-c)m-11\big)+(m-r-5) > (4-4c)m-39 > 3m,$$

a contradiction, if m is large enough. Thus there exist an integer $s \in \{0,1,2\}$ and indices $i_1,\ldots,i_s$ such that $f_i > (1-c)m-11$ if and only if $i \in \{i_1,\ldots,i_s\}$. Moreover, if $s \geq 1$, then for each $j \in \{1,\ldots,s\}$ we have

$$f_{i_j} \leq 3m-(m-r-3) < (2+c)m+4.$$

Consequently, there exist $\delta_1,\ldots,\delta_s \in \{-1,+1\}$ such that

$$\Big|N-\operatorname{sgn}(N)\sum_{i=1}^{r}(b_{2m-3-i}-b_i)-\sum_{j=1}^{s}\delta_j f_{i_j}\Big| < (2+c)m+4. \tag{20}$$

Put $\kappa = \lceil 3/\varepsilon \rceil \leq (1-c)m-11$ and introduce

$$I_\kappa = \{i \mid 1 \leq i \leq m-r-2,\ f_i \leq \kappa\}.$$

Denoting by x the number of indices i with $f_i > \kappa$ we have

$$(m-r-2-x)+(\kappa+1)x \leq \sum_{i=1}^{m-r-2} f_i \leq 3m,$$

implying $\kappa x < (2+c)m+3$, and thus

$$t = |I_\kappa| = m - r - 2 - x > \Big(1 - c - \frac{2+c}{\kappa}\Big)m - 3 - \frac{3}{\kappa} > \Big(\frac{4}{5} + \varepsilon\Big)m.$$

Write $c_0 = 0$ and let

$$\bigcup_{i \in I_\kappa} \{b_{r+2i-1}, b_{r+2i}\} = \{c_1 < c_2 < \ldots < c_{2t-1} < c_{2t}\}.$$

We prove that there exist $1 \le i_1 < j_1 \le t$ such that

$$\frac{2}{5}m \le \Delta_1 = c_{2j_1} - c_{2i_1} \le \frac{4}{5}m. \tag{21}$$

This is immediate if there exists $1 \le i \le t-1$ such that

$$\frac{2}{5}m \le c_{2i+2} - c_{2i} \le \frac{4}{5}m,$$

otherwise we have

$$c_{2i+2} - c_{2i} < \frac{2}{5}m \quad \text{or} \quad c_{2i+2} - c_{2i} > \frac{4}{5}m$$

for every integer $1 \le i \le t-1$. Gaps in the sequence $c_2, c_4, \ldots, c_{2t}$, which are larger than $4m/5$, partition this sequence into blocks, where the gap between two consecutive elements within a block is always smaller than $2m/5$. We claim that there cannot be more than 3 such blocks. Were there on the contrary at least 3 large gaps, we would find that

$$4m - 2 \ge \sum_{i=0}^{t-1}(c_{2i+2} - c_{2i}) > 3 \cdot \frac{4}{5}m + (t-3) \cdot 2 > (4 + 2\varepsilon)m - 6,$$

a contradiction. Now one of the blocks must contain at least $t/3$ different c_{2i}'s, and thus its length satisfies

$$\ell(B) \ge 2\Big(\frac{t}{3} - 1\Big) > \frac{2}{5}m.$$

Consequently, (21) holds with suitable elements c_{2i_1}, c_{2j_1} of B. Removing i_1, j_1 from I_κ and repeating the argument we find $1 \le i_2 < j_2 \le t$ such that $\{i_2, j_2\} \cap \{i_1, j_1\} = \emptyset$ and $2m/5 \le \Delta_2 = c_{2j_2} - c_{2i_2} \le 4m/5$. Since for $\alpha = 1, 2$ we have

$$1 \le c_{2i_\alpha} - c_{2i_\alpha - 1}, c_{2j_\alpha} - c_{2j_\alpha - 1} \le \kappa, \tag{22}$$

we can argue that

$$2\Delta_\alpha - \kappa + 1 \le \Gamma_\alpha = c_{2j_\alpha} + c_{2j_\alpha - 1} - c_{2i_\alpha} - c_{2i_\alpha - 1} \le 2\Delta_\alpha + \kappa - 1,$$

that is,

$$\frac{4}{5}m - \frac{3}{\varepsilon} < \Gamma_\alpha < \frac{8}{5}m + \frac{3}{\varepsilon}. \tag{23}$$

In view of (20) and (23), there exist an integer $p \in \{0,1,2\}$ and $\eta_1,\dots,\eta_p \in \{-1,+1\}$ such that

$$\Big| N - \operatorname{sgn}(N)\sum_{i=1}^{r}(b_{2m-3-i}-b_i) - \sum_{j=1}^{s}\delta_j f_{i_j} - \sum_{\alpha=1}^{p}\eta_\alpha\Gamma_\alpha \Big| < \frac{4}{5}m + \frac{3}{2\varepsilon} \le (1-c)m-6.$$

Consequently, we can choose $k = m-r-s-2p > (1-c)m-7$, and the elements of the set

$$\bigcup_{i=1}^{r}\{b_i, b_{2m-3-i}\} \cup \bigcup_{j=1}^{s}\{b_{r+2i_j}, b_{r+2i_j-1}\} \cup \bigcup_{\alpha=1}^{p}\{c_{2i_\alpha}, c_{2i_\alpha-1}, c_{2j_\alpha}, c_{2j_\alpha-1}\}$$

can be enumerated as $z_1,\dots,z_{n-2k}$ so that $F = N + \sum_{i=1}^{n-2k}(-1)^i z_i$ satisfies (17). Since $f_i \le (1-c)m-11$ holds for every $1 \le i \le m-r-2$, $i \notin \{i_1,\dots,i_s\}$, removing $z_1,\dots,z_{n-2k}$ form the sequence $b_1,\dots,b_{2m-4}$, the rest can be rearranged as $x_3, y_3, \dots, x_k, y_k$ such that $1 \le e_i = x_i - y_i$ satisfies (16). Finally, it follows from (19) that

$$\sum_{i=1}^{k} e_i \le \sum_{i=1}^{m-r-2} f_i + 3 \le 3m+3 \le (4-4c)m - 40 \le 4k-12,$$

therefore condition (15) is also fulfilled. This completes the proof of Theorem 3.

Acknowledgment I am thankful to Seva Lev who introduced me to the subject and whose suggestions simplified the presentation of this paper. I also gratefully acknowledge the hospitality and support of the CWI in Amsterdam, and of the Institute for Advanced Study in Princeton where the main part of this research was done with the support of the Alfred Sloan Foundation, NSF grant DMS-9304580 and by DIMACS under NSF grant STC-91-19999.

References

[1] Gy. Károlyi, Long arithmetic progressions in subset sums and a conjecture of Alon, submitted

[2] V.F. Lev, On consecutive subset sums, *Discrete Math.* **187** (1998) 151–160

[3] V.F. Lev, Blocks and progressions in subset sum sets, *Acta Arithm.* **106** (2003) 123–142

E-mail address: karolyi@cs.elte.hu