Dense Edge-Magic Graphs and Thin Additive Bases

Oleg Pikhurko

Department of Mathematical Sciences

Carnegie Mellon University

Pittsburgh, PA 15213-3890

Web: `http://www.math.cmu.edu/~pikhurko`

November 8, 2003

Abstract

A graph G of order n and size m is *edge-magic* if there is a bijection $l : V(G)\cup E(G) \to [n+m]$ such that all sums $l(a)+l(b)+l(ab)$, $ab \in E(G)$, are the same. We present new lower and upper bounds on $\mathcal{M}(n)$, the maximum size of an edge-magic graph of order n, being the first to show an upper bound of the form $\mathcal{M}(n) \le (1-\varepsilon)\binom{n}{2}$. Concrete estimates for ε can be obtained by knowing $s(k,n)$, the maximum number of distinct pairwise sums that a k-subset of $[n]$ can have.

So, we also study $s(k,n)$, motivated by the above connections to edge-magic graphs and by the fact that a few known functions from additive number theory can be expressed via $s(k,n)$. For example, our estimate

$$s(k,n) \le n + k^2\left(\frac{1}{4} - \frac{1}{(\pi+2)^2} + o(1)\right)$$

implies new bounds on the maximum size of quasi-Sidon sets, a problem posed by Erdős and Freud [J. Number Th. **38** (1991) 196–205]. The related problem for differences is considered as well.

Keywords: additive basis, edge-magic graph, Sidon set, quasi-Sidon set, sum-set.

AMS Subject Classification: 05C78, 11B75.

1 Introduction

Let $[k]$ stand for $\{1, \ldots, k\}$. Let G be a graph with n vertices and m edges. An *edge-magic labelling* with the *magic sum* s is a bijection $l : V(G) \cup E(G) \to [m+n]$ such that $l(a) + l(b) + l(ab) = s$ for any edge ab of G. (We always assume that $V(G) \cap E(G) = \emptyset$.) This definition appeared first in Kotzig and Rosa [13] (but under the name *magic valuation*). The graph G is *edge-magic* if it admits an edge-magic labelling (for some s). We refer the reader to Gallian [8] and Wood [21] for plentiful references on edge-magic graphs.

Not all graphs are edge-magic nor is this property in any way monotone with respect to the subgraph relation. In 1996 Erdős asked (see [3]) for $\mathcal{M}(n)$, the maximum number of edges that an edge-magic graph of order n can have.

This function has been computed exactly for $n \leq 6$ but for large n the best known bounds were $\lfloor n^2/4 \rfloor \leq \mathcal{M}(n) \leq \binom{n}{2} - 1$, see Craft and Tesar [3].

Here we improve both these bounds if n is large.

Theorem 1

$$\frac{2}{7}\, n^2 + O(n) \leq \mathcal{M}(n) \leq (0.489... + o(1))\; n^2. \tag{1}$$

It turns out that edge-magic labellings have strong relations to some problems from additive number theory, especially to additive bases.

Section 2 can serve as a warm-up where we improve the bounds of Wood [21] on so-called *edge-magic injections*. Our proof uses some classical results about *Sidon sets*, that is, sets $A \subset \mathbb{Z}$ such that all sums $a + b$, with $a, b \in A$ and $a \leq b$, are distinct.

For a set A of integers define its *sum-set* by $A + A := \{a + b : a, b \in A\}$; A is called an *additive basis* for X if $A + A \supset X$. In Section 3, we prove the lower bound in (1) by using known (explicit) constructions of a thin additive basis for some suitable interval of integers.

But the most interesting connections were found during our quest for an upper bound on $\mathcal{M}(n)$. This research led us to the following problem. What is

$$s(k,n) := \max\left\{|A+A| : A \in \binom{[n]}{k}\right\},$$

that is, the maximum size of the sum-set of a k-subset of $\{1,\ldots,n\}$?

The trivial upper bound is

$$s(k,n) \le \min\left\{\binom{k}{2} + k, 2n-1\right\}. \qquad (2)$$

We have $s(k,n) = \binom{k}{2} + k$ if and only if there exists a Sidon k-set $A \subset [n]$; the classical results of Singer [20] and Erdős and Turán [6] (see [10, Chapter II]) state that for a given n the largest such k is $(1+o(1))\,n^{1/2}$. The open question whether the maximum size of a Sidon subset of $[n]$ is $n^{1/2}+O(1)$ has the \$500-dollar reward of Erdős [4] attached.

We have $s(k,n) = 2n-1$ if and only if there is an additive k-basis $A \subset [n]$ for $[2,2n]$. How small can k be then? A simple construction of Rohrbach [19, Satz 2] gives $(2\sqrt{2}+o(1))\,n^{1/2}$ for k (see Section 7). The trivial lower bound is $k \ge (2+o(1))\,n^{1/2}$; the current best known bound $k \ge (2.17...+o(1))\,n^{1/2}$ of Moser, Pounder and Riddell [17] is only slightly bigger.

As we see, already the question when we have equality in (2) leads to very difficult open problems. The computation of $s(k,n)$ for other values is likely to be even harder. We present the following upper bound which improves on (2) for a range of k around $2\,n^{1/2}$.

Theorem 2

$$s(k,n) \le n + k^2\left(\frac{1}{4} - \frac{1}{(\pi+2)^2} + o(1)\right). \qquad (3)$$

Here is an application of Theorem 2. Erdős and Freud [5] call a set $A \in \binom{[n]}{k}$ with $|A+A| = (1+o(1))\binom{k}{2}$ *quasi-Sidon* and ask how large k can be. (It is obvious what is meant here so we do not bother writing out any formal definitions.) They constructed quasi-Sidon subsets of $[n]$ with

$$k = (2/\sqrt{3}+o(1))\,n^{1/2} = (1.154...+o(1))\,n^{1/2}. \qquad (4)$$

As $A + A \subset [2n]$, a trivial upper bound is $\binom{k}{2} \le (2 + o(1))\, n$, that is, $k \le (2 + o(1))\, n^{1/2}$. Erdős and Freud [5, p. 204] promised to publish the proof of $k \le (1.98 + o(1))\, n^{1/2}$ in a follow-up paper. Unfortunately, it has not been published, but their bound is superseded by the following easy corollary of Theorem 2 anyway.

Theorem 3 *Let $A \subset [n]$ be quasi-Sidon. Then*

$$|A| \le \left(\left(\frac{1}{4} + \frac{1}{(\pi + 2)^2} \right)^{-1/2} + o(1) \right) n^{1/2} = (1.863... + o(1))\, n^{1/2}. \blacksquare$$

As another application of Theorem 2 let us show that $\mathcal{M}(n) \le (1 - \varepsilon)\binom{n}{2}$. Indeed, if G is an edge-magic graph of order n and size $(\frac{1}{2} + o(1))\, n^2$, then its vertex labels form a quasi-Sidon set, which contradicts Theorem 3. This way we do not obtain any explicit value for ε but one can get one by using Theorem 2 with a little bit of work. A slightly better bound, the one in (1), is deduced in Section 5 from a generalisation of Theorem 2.

Given these applications of $s(k, n)$, we present some lower bounds on $s(k, n)$ in Section 7. It is interesting to compare them with the upper bounds, see Figure 1.

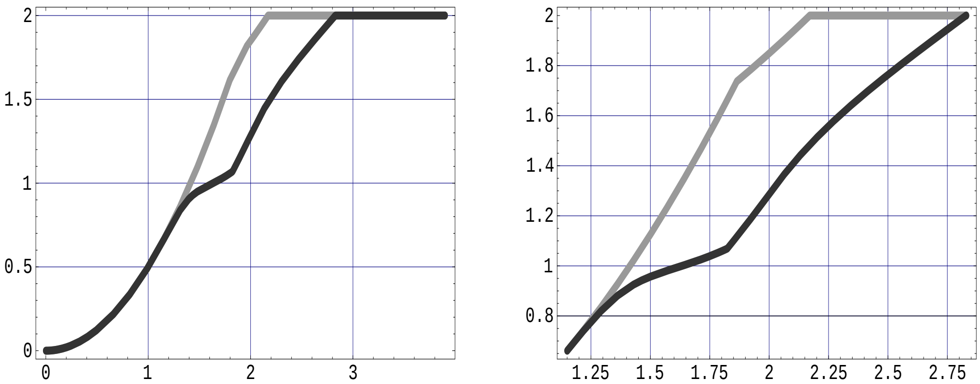


Figure 1: Our bounds on $s(k, n)$: $x = kn^{-1/2}$, $y = s(k, n)/n$.

Our auxiliary Lemma 10 states that any asymptotically maximum Sidon subset of $[n]$ is uniformly distributed in subintervals and in residue classes simultaneously. This places the corresponding results of Erdős and Freud [5] and Lindström [14] under a common roof.

Besides being a natural and interesting question on its own, the $s(k,n)$-problem demonstrates new connections between Sidon sets and additive bases. This helped the author to realise that the technique of Moser [16] which was used in the context of additive bases can be applied to $s(k,n)$ (and to quasi-Sidon sets). In fact, our proof of Theorem 2 goes by modifying Moser's [16] method. Although the determination of $s(k,n)$ is apparently very hard, it seems a worthwhile direction of research.

In Section 8 we study the analogous problem for differences.

2 Edge-Magic Injections

Wood [21] defines an *edge-magic injection* of a graph G as an injection $l : V(G) \cup E(G) \to \mathbb{Z}_{>0}$ (into positive integers) such that for any edge $ab \in E(G)$ the sum $l(a)+l(b)+l(ab) = s$ is constant. Note that the labels need not sweep a contiguous interval of integers (but must be pairwise distinct). It is easy to show that any graph G admits an edge-magic injection.

The general question is how economical such a labelling can be. One possible way to state it formally is to ask about $\mathcal{I}(G)$, the smallest value of the magic sum s over all edge-magic injections of G. If $v(G) = n$, then clearly $\mathcal{I}(G) \le \mathcal{I}(K_n)$, so here we investigate $\mathcal{I}(K_n)$. Wood [21, Theorem 1] showed that $\mathcal{I}(K_n) \le (3 + o(1))\, n^2$. Here we improve on it.

Theorem 4

$$\mathcal{I}(K_n) \le \left(\frac{288}{121} + o(1)\right) n^2 = (2.380... + o(1))\, n^2. \tag{5}$$

Proof. Choose $m = \lceil (\frac{12}{11} + \delta) n \rceil$ for some small constant $\delta > 0$. Take a Sidon set

$$A = \{a_1, \dots, a_m\} \text{ with } 1 \le a_1 < a_2 < \cdots < a_m \le (1 + o(1))\, m^2, \qquad (6)$$

that is, asymptotically maximum. Explicit such sets were constructed by Singer [20] and by Bose and Chowla [1] (Theorems 1 and 3 of Chapter II in [10]).

The case $m = 1$ of our Lemma 10 (or Lemma 1 in Erdős and Freud [5]) shows that A is almost uniformly distributed in $[a_m]$. This implies that if define T to consist of all triple sums $a_f + a_g + a_h$, $1 \le f \le g \le h \le m$, counted with their multiplicities, then we know the asymptotic distribution of T. We are interested in the interval $[2m^2, 3m^2]$, where the 'density' of T at xm^2, $2 \le x \le 3$, is

$$\int_{x-2}^{1} \mathrm{d}y \int_{x-y-1}^{1} \mathrm{d}z + o(1) = \frac{(3-x)^2}{2} + o(1).$$

For example, the number of elements of T lying between $2m^2$ and $3m^2$ is

$$(1 + o(1)) \binom{m}{3} \int_2^3 \frac{(3-x)^2}{2}\, \mathrm{d}x = \left(\frac{1}{36} + o(1)\right) m^3.$$

The interval $I := [2a_m, (2+\delta)m^2]$ has about $\frac{\delta}{2}\binom{m}{3}$ elements of T, so some $s \in I$ has multiplicity $k \le (\frac{1}{12} + o(1))\, m$. For each of the k representations $s = a_f + a_g + a_h$ remove one of the summands from A. Let $B \subset A$ be the remaining set. By removing further elements we can assume that $|B| = n$.

Label the vertices of K_n by the elements of B. We want s to be the magic sum. This determines uniquely the edge labels which are positive (because $s \ge 2a_m$) and pairwise distinct (because $B \subset A$ is a Sidon set). Also, as $s \notin B + B + B$, no edge label equals a vertex label. As δ can be chosen arbitrarily small, we obtain $s = (2 + o(1))\, m^2 = (\frac{288}{121} + o(1))\, n^2$, proving the theorem. ∎

3 Lower Bound on $\mathcal{M}(n)$

For $A \subset \mathbb{Z}$ let $A \oplus A := \{a+b : a, b \in A,\ a \neq b\}$. We have $A \oplus A \subset A + A$.

Lemma 5 *Suppose that there is a set $A := \{a_1 = 1 < a_2 < \cdots < a_n\}$ of integers such that $A \oplus A$ contains an interval of length m (that is, $A \oplus A \supset [k, k+m-1]$ for some k). If $a_n \leq m$, then $\mathcal{M}(n) \geq m - n$.*

Proof. We construct an edge-magic graph G on $[n]$ with $m-n$ edges. Label $i \in [n]$ by $l(i) := a_i$. The magic sum will be $s := k+m$. For every $a \in A \oplus A$ with $s - a \in [m] \setminus A$ choose a representation $l(i) + l(j) = a$, $1 \leq i < j \leq n$, and add the pair $\{i, j\}$ (with label $s-a$) to $E(G)$.

Clearly, no two labels are the same. We have

$$\{s - a : a \in A \oplus A\} \supset [m] \supset A$$

So the label set is $[m]$ and we do have an edge-magic graph. The number of edges is $|[m] \setminus A| = m - n$, as required. ∎

Mrose [18] constructed a set $A \subset [0, 10t^2 + 8t]$ of size $7t+3$ such that $A + A \supset L := [0, 14t^2 + 10t - 1]$. In fact, $A = \cup_{i=1}^{5} A_i$ is the union of five disjoint arithmetic progressions. Namely, let

$$[a, (d), b] := \{a + id : i = 0, 1, \ldots, \lfloor (b-a)/d \rfloor\};$$

then

$$\begin{aligned}
A_1 &:= [0, (1), t], \\
A_2 &:= [2t, (t), 3t^2 + t], \\
A_3 &:= [3t^2 + 2t, (t+1), 4t^2 + 2t - 1], \\
A_4 &:= [6t^2 + 4t, (1), 6t^2 + 5t], \\
A_5 &:= [10t^2 + 7t, (1), 10t^2 + 8t],
\end{aligned}$$

Fried [7] independently discovered a similar construction, giving almost the same bounds.

For any arithmetic progression B we have $|(B+B)\setminus(B\oplus B)| \le 2$ (because $2b_i = b_{i-1} + b_{i+1}$). Hence, $A \oplus A$ contains all but at most 10 elements from $I := [0, 14t^2 + 10t - 1]$. Inspecting each of the ten suspicious elements, we see that $I \setminus (A \oplus A) = \{0, 8t^2 + 4t - 2\}$. Applying Lemma 5 to, for example, the set $\{a+1 : a \in A\} \cup \{8t^2 + 4t - 1\}$ with $n = 7t + 4$, $k = 3$, $m = 14t^2 + 10t - 1$, we obtain that $\mathcal{M}(7t+4) \ge 14t^2 + 3t - 5$ for any $t \ge 1$. Now, the lower bound in (1) follows from the following lemma.

Lemma 6 *For any n we have $\mathcal{M}(n) \le \mathcal{M}(n+1)$.*

Proof. Let G be a maximum edge-magic graph of order n with a labelling l. The graph G' obtained by adding an extra isolated vertex x to G is edge-magic: extend l to G' by defining $l(x) = v(G) + e(G) + 1$. ∎

Problem 7 *Does the ratio $\mathcal{M}(n)/n^2$ tend to a limit as $n \to \infty$?*

4 The Number of Pairwise Sums

The following result is proved via the modification of the argument in Moser, Pounder and Riddell [17, Lemma 1] which in turn is built upon the generating function method of Moser [16]. We also refer the reader to a few related papers: Klotz [11], Green [9], Cilleruelo, Ruzsa and Trujillo [2], Martin and O'Bryant [15].

Theorem 8 *Let $\lambda = \frac{1}{4}(2\sqrt{2} - 4 + \pi(4 - \sqrt{2})) = 0.323\ldots$. Let n be large, $A \subset \mathbb{Z}$, $m := |A \setminus [n]|$, and $k := |A \cap [n]|$. If $k \ge \lambda m$, then*

$$|(A+A) \cap [2n]| \le n + \frac{|A|^2}{4} - \frac{(|A| - \pi m)^2}{(\pi+2)^2} + o(n), \tag{7}$$

where the $o(n)$ term depends on n only.

Proof. Assume that $|A| = O(n^{1/2})$ for otherwise we are done. Let $A = \{a_1, \cdots, a_{k+m}\}$ with $a_1, \dots, a_k \in [n]$. Correspond to A its generating function

$$f(x) := \sum_{j=1}^{k+m} x^{a_j}.$$

Let $g(x) = (f^2(x) + f(x^2))/2$. Clearly, the coefficient at x^j in $g(x)$ is the number of representations of j of the form $a_s + a_t$ with $1 \le s \le t \le k+m$.

Let $h(x) := \sum_{j=1}^{2n} x^j$. Define δ_j, $j \in \mathbb{Z}$, by the formal identity

$$\sum_{j\in\mathbb{Z}} \delta_j x^j := g(x) - h(x).$$

We have $\sum_{j=0}^{2n} \delta_j = \binom{k+m+1}{2} - 2n$.

Let $t \in [2n-1]$. Then $h(\mathrm{e}^{\pi i t/n}) = 0$, where i is a square root of -1. Hence,

$$\sum_{j\in\mathbb{Z}} \delta_j \mathrm{e}^{\pi i t j/n} = g(\mathrm{e}^{\pi i t/n}).$$

Also observe that each δ_j is non-negative with the exception of j lying in $L := [2n] \setminus (A + A)$ when $\delta_j = -1$. Let $l := |L|$.

Putting all together we obtain, for $t \in [2n-1]$,

$$\begin{aligned} \frac{1}{2}\left(|f^2(\mathrm{e}^{\pi i t/n})| - |f(\mathrm{e}^{2\pi i t/n})|\right) &\le |g(\mathrm{e}^{\pi i t/n})| \le \sum_{j\in\mathbb{Z}\setminus L} \delta_j + \left|\sum_{j\in L} \mathrm{e}^{\pi i t j/n}\right| \\ \le \sum_{j\in\mathbb{Z}} \delta_j + 2l + o(n) &= \binom{k+m+1}{2} - 2n + 2l + o(n). \quad (8) \end{aligned}$$

Let z denote the right-hand side of (8), including the $o(p)$-term.

Let $b_t := \frac{2}{t^2-1}$ for even $t > 0$ and $b_t := 0$ otherwise. Clearly, $|f(\mathrm{e}^{2\pi i t/n})| \le k + m$ while

$$|f^2(\mathrm{e}^{\pi i t/n})| = |f(\mathrm{e}^{\pi i t/n})|^2 = \Big(\sum_{j\in A} \sin(\pi t a_j/n)\Big)^2 + \Big(\sum_{j\in A} \cos(\pi t a_j/n)\Big)^2. \quad (9)$$

Hence, from (8) and (9) we deduce that

$$\frac{\pi}{2}\,(2z)^{1/2} \ \ge\ \frac{\pi}{2} \sum_{j\in A} \sin(\pi a_j/n), \quad (10)$$

$$b_t\,(2z)^{1/2} \ \ge\ b_t \sum_{j\in A} \cos(\pi t a_j/n), \quad t \in [2, 2n-1]. \quad (11)$$

Note that $\sum_{t=2}^{2n-1} b_t = 1 - \frac{1}{2n-1} < 1$. By adding (10) and (11) we obtain

$$\left(\frac{\pi}{2}+1\right)(2z)^{1/2} \geq \sum_{j\in A}\left(\frac{\pi}{2}\sin(\pi a_j/n) + \sum_{t=2}^{2n-1} b_t\cos(\pi t a_j/n)\right) \tag{12}$$

It is routine to see that $S(x) := \frac{\pi}{2}\sin(x) + \sum_{t=2}^{\infty} b_t\cos(tx)$ is the Fourier series of the function

$$r(x) = \begin{cases} 1, & 0 \leq x \leq \pi, \\ 1+\pi\sin(x), & \pi \leq x \leq 2\pi. \end{cases}$$

(This series appears in [17, p. 400].) As the sum $\sum_{t=2}^{\infty} |b_t|$ converges and $r(x) : \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{R}$ is a continuous function, it follows from Körner [12, Theorem 9.1] that $S(x)$ converges uniformly to $r(x)$. Noting that $0 \leq \pi a_j/n \leq \pi$ for any $j \in [k]$, we conclude that

$$\left(\frac{\pi}{2}+1\right)(2z)^{1/2} \geq k + (1-\pi)m + o(m+k). \tag{13}$$

Assume that $(\pi-1)m > k$ for otherwise we obtain the required by squaring (13).

Now, (13) is vacuous but we can use the obvious upper bounds on $|(A + A) \cap [2n]|$ such as $2n$ and $\binom{k+m+1}{2} - \frac{m^2}{4}$. (The latter follows from the fact that the pairwise sums in $\{x \in A : x > n\}$ and in $\{x \in A : x < 1\}$ lie outside $[2n]$.) If neither of these bounds implies (7), then

$$\frac{(k+m)^2}{4} - \frac{(k+m(1-\pi))^2}{(\pi+2)^2} < n < \frac{(k+m)^2}{4} - \frac{m^2}{4} + \frac{(k+m(1-\pi))^2}{(\pi+2)^2}.$$

Solving the obtained quadratic inequality in k (and using $k < m(\pi-1)$), we obtain $k < \lambda m$, as required. ∎

Note that Theorem 2 easily follows from (7).

5 Upper Bound on $\mathcal{M}(n)$

To prove an upper bound on $\mathcal{M}(n)$ we study the following function first. Let $b(k)$ be the largest n such that for some k-set $A \subset \mathbb{Z}$ we have

$$|(A+A) \cap [n]| = (1-o(1))\, n. \tag{14}$$

It is not hard to see that $b(k)$ has order $\Theta(k^2)$. To state the problem formally, we consider the following constant:

$$b_{\sup} := \limsup_{\varepsilon \to 0} \limsup_{k \to \infty} \frac{\max\{n : \exists A \in \binom{\mathbb{Z}}{k},\ |(A+A) \cap [n]| \ge (1-\varepsilon)n\}}{k^2}. \tag{15}$$

This definition is related to the question of Rohrbach [19] which (when correspondingly reformulated) asks about $b'(k)$, the largest n such that $[0,n] \subset A+A$ for some k-set $A \subset \mathbb{Z}_{\ge 0}$. (Note that here A must consist of *non-negative* integers.) The currently best known upper bound

$$b'(k) \le (0.480... + o(1))\, k^2,$$

is due to Klotz [11]. In fact, Klotz's argument gives the same bound if we weaken the assumption $[0,n] \subset A+A$ to (14). The two-side restricted function $b''(k)$ (when we require that $A \subset [0, (\frac{1}{2} + o(1))n]$) has also been studied with the present record

$$b''(k) \le (0.424... + o(1))\, k^2,$$

belonging to Moser, Pounder and Riddell [17] (valid with the weaker assumption (14) as well).

However, it seems that nobody has considered $b(k)$. Here we fill this gap as this is the function needed for our application.

Theorem 9

$$b_{\sup} \le \frac{1}{2} - \frac{2}{(2 + (1+2\sqrt{2})\,\pi)^2} = 0.489\ldots.$$

Proof. Let $A \subset \mathbb{Z}$ have size k and satisfy (14). We can assume that n is even. Let $m := |A \setminus [n/2]|$. As at least $2\binom{m/2}{2} = (\frac{1}{4} + o(1))\, m^2$ sums in $A + A$ fall outside $[n]$, we have

$$n \le \binom{k}{2} - \frac{m^2}{4} + o(k^2) \tag{16}$$

If $m \ge k/\pi$, then we have

$$n \le \left(\frac{1}{2} - \frac{1}{4\pi^2} + o(1)\right) k^2 = (0.474... + o(1))\, k^2, \tag{17}$$

and we are done. Otherwise, by (7) we obtain

$$n \le \frac{n}{2} + \frac{k^2}{4} - \frac{(k - \pi m)^2}{(\pi + 2)^2} + o(k^2). \tag{18}$$

We conclude that

$$b_{\sup} \le \min_{m \in [0,k]} \left(\frac{1}{2} - \frac{(m/k)^2}{4},\ \frac{1}{2} - \frac{2(1 - \pi m/k)^2}{(\pi + 2)^2}\right),$$

and the claim routinely follows. ∎

Let us return to the original problem. Let l be an edge-magic labelling with the magic sum s of a graph G of order n and size m. Let $A := l(V(G))$. We have

$$(A + A) \cap [s - m - n, s - 1] \supset \{s - l(\{x, y\}) : xy \in E(G)\}, \tag{19}$$

that is, $A + A$ contains almost whole interval of length $m + n$ (assuming, obviously, $n = o(m)$). We conclude that $m \le (b_{\sup} + o(1))\, n^2$, which establishes the upper bound in (1).

6 Asymptotically Maximum Sidon Sequences

As we have already mentioned, the maximum size of a Sidon subset of $[n]$ is $(1 + o(1))\, n^{1/2}$. Erdős and Freud [5, Lemma 1] showed that a set achieving this bound is almost uniformly distributed among subintervals of $[n]$. Lindström [14, Theorem 1] proved the analogue of this result with respect to residue classes.

Here we prove a common generalisation of these results which we will need in Section 7. Our proof is based on the method of Erdős and Freud [5, Lemma 1].

Lemma 10 *Let n be large. Let A be an asymptotically maximum Sidon subset of $[n]$ (that is, having size $(1+o(1))\,n^{1/2}$). Then for any subinterval $I \subset [n]$ and for any integers m and l, we have*

$$|A \cap I \cap M_l| = \frac{|I|}{m\,n^{1/2}} + o(n^{1/2}). \tag{20}$$

where $M_l := \{x \in \mathbb{Z} : x \equiv l \pmod{m}\}$.

Proof. It is enough to prove the lemma for $I = [k]$, an initial interval, as any other interval is the set-theoretic difference of two such intervals. Assume that $k = \Omega(n)$ and $m = O(1)$ for otherwise (20) trivially holds.

Choose an integer $t = \Theta(n^{3/4})$. Let $J = \{jm : j \in [t]\}$. For $i \in [-mt+1, n-1]$ let $A_i := A \cap (J+i)$ and $a_i := |A_i|$. By the Sidon property of A, the difference set $(A_i - A_i) \cap \mathbb{Z}_{>0} \subset J$ has $\binom{a_i}{2}$ elements. On the other hand, a difference $jm \in J$ is counted $t-j$ times. Hence, we conclude that

$$\sum_{j=1}^{t}(t-j) = \binom{t}{2} \geq \sum_{i=-mt+1}^{n-1} \binom{a_i}{2} = \frac{1}{2}\sum_{i=-mt+1}^{n-1} a_i^2 - \frac{1}{2}\sum_{i=-mt+1}^{n-1} a_i \tag{21}$$

The left-hand size of (21) has magnitude $t^2 = \Theta(n^{3/2})$. All $o(n^{3/2})$-expressions will be dumped into the error term. In particular, $\sum_i a_i = t|A|$ goes there.

To estimate $\sum_i a_i^2$ we split the summation interval into smaller parts

$$R_j := [-mt+1, k] \cap M_j \text{ and } S_j := [k+1, n-1] \cap M_j,\ \ j \in [m].$$

Now we apply the arithmetic-geometric mean inequality.

$$\begin{aligned}\sum_{i=-mt+1}^{n-1} a_i^2 &\geq \sum_{j\in[m]} \left(\frac{\left(\sum_{i\in R_j} a_i\right)^2}{|R_j|} + \frac{\left(\sum_{i\in S_j} a_i\right)^2}{|S_j|} \right) \\ &= mt^2 \left(\sum_{j\in[m]} \frac{|A\cap I\cap M_j|^2}{k} + \sum_{j\in[m]} \frac{|(A\setminus I)\cap M_j|^2}{n-k} \right) + o(n^{3/2}).\end{aligned}$$

(Note that $|R_j| = \frac{k}{m} + O(t)$, $|S_j| = \frac{n-k}{m} + O(1)$, and $a_i = O(t^{1/2})$.)

We can estimate the first summand as follows, by using the arithmetic-geometric mean inequality.

$$\frac{mt^2}{k} \sum_{j\in[m]} |A\cap I\cap M_j|^2 \geq \frac{t^2}{k} \left(\sum_{j\in[m]} |A\cap I\cap M_j| \right)^2 = \frac{t^2}{k}|A\cap I|^2.$$

We obtain the analogous bounds for $A\setminus I$. Let $|A\cap I| = \alpha n^{1/2}$. Then $|A\setminus I| = (1-\alpha+o(1))\, n^{1/2}$. In summary, starting with (21), we obtain

$$\binom{t}{2} \geq \frac{t^2}{2} \left(\frac{|A\cap I|^2}{k} + \frac{|A\setminus I|^2}{n-k} \right) + o(n^{3/2}) = t^2 \left(\frac{1}{2} + \frac{(\alpha n - k)^2}{k(n-k)} \right) + o(n^{3/2}).$$

Thus, up to an error term of $o(n^{3/2})$, we must have equality throughout. We conclude that $\alpha = k/n + o(1)$ and for any $j \in [m]$ we have $\sum_{i\in R_j} a_i = (1+o(1))\, k/mn^{1/2}$. (Recall that $m = O(1)$.) The lemma is proved. ∎

7 Lower Bounds on $s(k,n)$

We know that the range of interest is $k = \Theta(n^{1/2})$. We will be proving lower bounds on the following 'scaled' one-parameter version of $s(k,n)$:

$$s(c) := \liminf_{n\to\infty} \frac{s(\lfloor cn^{1/2}\rfloor, n)}{n}. \tag{22}$$

Note that in (22) we could have replaced $\lfloor cn^{1/2}\rfloor$ by anything of the form $(c+o(1))\, n^{1/2}$ without affecting the value of $s(c)$. However, we have to write lim inf as the following question is open.

Problem 11 *Let c be a fixed positive real. Suppose that n tends to the infinity and $k = (c + o(1))\, n^{1/2}$. Does the ratio $s(k, n)/n$ tend to a limit?*

Our lower bound on $s(c)$, provided by the following lemmas, will be given by different formulae for different ranges of c.

The bound (4) of Erdős and Freud [5] implies that

$$s(c) = \frac{c^2}{2}, \quad c \le 2/\sqrt{3}. \tag{23}$$

Their construction can be generalised to give lower bounds on $s(c)$ for larger c.

Lemma 12

$$s(c) \ge \begin{cases} -\frac{5c^2}{8} + \frac{9}{2} - \frac{6}{c^2} + \frac{8}{3c^4}, & 2/\sqrt{3} \le c \le \sqrt{2}, \\ \frac{3c^2}{8} - \frac{3}{2} + \frac{6}{c^2} - \frac{16}{3c^4}, & \sqrt{2} \le c \le 2. \end{cases} \tag{24}$$

Proof. Let $\alpha = c^2/4$. Choose an integer $m = (\alpha + o(1))\, n$. Let $A \subset [m]$ be a Sidon set with $(1 + o(1))\, m^{1/2}$ elements. The main idea (which we borrow from Erdős and Freud [5]) is to consider the set $X := A \cup (n - A)$, where $n - A := \{n - a : a \in A\}$. It is easy to see that, as A is a Sidon set, all pairwise sums in $A + (n - A)$ are distinct.

However, the set $A + (n - A)$ might intersect $A + A$. In order to control the intersection size we introduce some randomness into the definition of X. In what follows, $\varepsilon > 0$ is a sufficiently small constant. Let s, t be two integers chosen uniformly and independently from between 1 and $\varepsilon^2 n$. We define

$$X := B \cup C, \quad \text{where } B := s + A \text{ and } C := n - t - A.$$

Let us compute the densities in $X + X$ which are well defined because of Lemma 10. For example, if we denote

$$\delta_{B+B}(x) := \frac{|(B + B) \cap I|}{|I|},$$

where I is an interval of integers of length $(\varepsilon + o(1))n$ around xn, then

$$\delta_{B+B}(x) = (\text{error term}) + \begin{cases} \frac{x}{2\alpha}, & 0 \le x \le \alpha, \\ -\frac{x}{2\alpha} + 1, & \alpha \le x \le 2\alpha, \\ 0, & \text{otherwise}, \end{cases}$$

where the error term tends to zero if $\varepsilon > 0$ is sufficiently small and $n \ge n_0(\varepsilon)$. Similarly,

$$\delta_{B+C}(x) = (\text{error term}) + \begin{cases} 0, & 0 \le x \le 1-\alpha, \\ \frac{x}{\alpha} - \frac{1}{\alpha} + 1, & 1-\alpha \le x \le 1. \end{cases}$$

As the picture is symmetric with respect $x = 1$ (given our scaling), we do not bother about $x \ge 1$ (or about $C + C$).

Thus when one takes some $v \in [n]$ then the probability that $v \in B + B$ is approximately $\delta_{B+B}(v/n)$. Indeed, this is equivalent to $v - 2s \in A + A$. The case $m = 2$ of Lemma 10 implies that the number of odd and even elements of $A + A$ in the vicinity of v is about the same, so their relative density is $\delta_{A+A}(v) + o(1)$. The analogous claim about the probability of $v \in B + C$ is also true. Moreover,

$$\Pr\{v \in (B+B) \cap (B+C)\} = \delta_{B+B}(v/n) \times \delta_{B+C}(v/n) + o(1),$$

because the event is equivalent to $v - 2s \in A + A$ and then, conditioned on this, to $(v - s - n) + t \in A - A$, which has probability $\delta_{A-A}(\frac{v-s-n}{n}) + o(1) = \delta_{B+C}(\frac{v}{n}) + o(1)$.

Hence, by simple inclusion-exclusion, the expectation of $|X+X|$ is at least

$$(2 + o(1))\, n \int_0^1 (\delta_{B+B}(x) + \delta_{B+C}(x) - \delta_{B+B}(x)\delta_{B+C}(x))\, \mathrm{d}x. \qquad (25)$$

(Recall that we use the symmetry around $x = 1$.) The points α, $1 - \alpha$, and 2α partition the x-range into intervals on each of which the function in the integral (25) is given by an explicit polynomial in x. We have to be careful with the relative positions of the dividing points: for $\alpha = 1/2$ (that is, for $c = \sqrt{2}$), the points α and $1-\alpha$ swap places while 2α disappears from the interval. This is why we have two cases in the bound (24) which is obtained by straightforward although somewhat lengthy calculations (omitted).

Finally observe that there exist s and t such that $|X+X|$ is at least its expectation, proving the lemma. ▮

A construction of Rohrbach [19, Satz 2] shows that

$$s(x) = 2, \qquad \text{if } x \geq 2\sqrt{2}. \tag{26}$$

We can extend it for smaller x in the following way.

Lemma 13 *Let $c_0 := 7/(2\sqrt{3}) = 2.02...$ and $c_1 := 2\sqrt{2} = 2.82...$ Then*

$$s(c) \geq \begin{cases} \frac{9c^2}{28}, & c \leq c_0 \\ -c^2 + 7\alpha c + \frac{c}{\alpha} - 11\alpha^2 - 2 - \frac{1}{4\alpha^2}, & c_0 \leq c \leq c_1, \end{cases} \tag{27}$$

where $\alpha = \alpha(c)$ is the linear function with $\alpha(c_0) = \sqrt{3}/4$ and $\alpha(c_1) = 1/\sqrt{2}$.

Proof. Let $k = (c + o(1))\, n^{1/2}$ and let $l := (3c/14 + o(1))\, n^{1/2}$ for $c \leq c_0$ and $l := (\alpha + o(1))\, n^{1/2}$ otherwise.

Let $A := [l]$, $B := [n-l+1, n]$. Let C and D be two arithmetic progressions each of length $\frac{k}{2} - l$ starting at $(1/2 + o(1))\, n$ but with differences $-l$ and $l+1$ respectively. Let $X := A \cup B \cup C \cup D$.

All pairwise sums in $A + (C \cup D)$ are distinct, lying within an interval $[a_0, a_1]$, where $a_0 = \frac{n}{2} - m + o(n)$ and $a_1 = \frac{n}{2} + m + o(n)$, where $m := (\frac{k}{2} - l)l$.

Now let us consider $C + D$. Suppose that $c' + d' = c'' + d''$ for some $c' < c''$ in C and $d' > d''$ in D. Now, the difference $c'' - c' = d' - d''$ is divisible by both l and $l+1$, hence, it is at least $l(l+1)$. It is routine to check that $2l^2 > m + o(1) \geq l^2$ for $0 < c \leq c_1$. This implies that $o(n)$ elements of $C + D$

have multiplicity at least 3 and $(\frac{k}{2}-2l)^2+o(1)$ elements have multiplicity 2 (and all others have multiplicity 1).

Observe also that $C+D\subset[b_0,b_1]$, where $b_0=n-m+o(n)$ and $b_1=n+m+o(n)$.

Let $c\le c_0$. Then $b_0\ge a_1+o(n)$, that is, $A+(C\cup D)$ and $C+D$ have $o(n)$ elements in common. Therefore, by a sort of symmetry around n, we obtain

$$|X+X|=4(k/2-l)l+(k/2-l)^2-(k/2-2l)^2, \tag{28}$$

giving the claimed bound.

However, for $c_0\le c\le c_1$, we have $b_0\le a_1+o(n)$. Hence, we have to subtract from the bound (28) twice (by the symmetry) the number of elements of $C+D$ lying in $[b_0,a_1]$. This correction term is

$$2\times n\int_{b_0/n}^{a_1/n}\left(\frac{x}{\alpha^2}+\frac{c}{2\alpha}-1-\frac{1}{\alpha^2}\right)\mathrm{d}x+o(n)$$

Computing the value of the integral and plugging it into (28), the reader should be able to derive the stated bound. ∎

Remark. The choice of l for $c_0\le c\le c_1$ in Lemma 13 is not best possible. It seems that there is no closed expression for the optimal choice. So we took a linear interpolation, given the optimal values for $c=c_0$ and $c=c_1$.

Figure 1 (drawn in *Mathematica*) contains the graphical summary of our findings.

8 Differences

Similar questions can be asked about differences. For example, let us define

$$d(k,n):=\max\left\{|A-A| : A\in\binom{[n]}{k}\right\}.$$

The obvious upper bounds are $2n-1$ and $k(k-1)+1$ (where the last summand 1 counts $0\in A+A$). These bounds can be improved when $\sqrt{n}\le(1+o(1))k\le\frac{3}{2}\sqrt{n}$ as the following theorem demonstrates.

Theorem 14 *Let n be large and $k \geq \sqrt{n}$. Then*

$$d(k,n) \leq 2k\sqrt{n} - n + o(n). \tag{29}$$

Proof. Let $c := k/\sqrt{n} > 1$. Assume that $c - 1 = \Theta(1)$ for otherwise we are trivially done. Define $t := \lfloor (c-1)n \rfloor$,

$$A_i := A \cap [i, i+t-1], \text{ and } a_i := |A_i|, \quad i \in [2-t, n].$$

Let $\mathcal{X}$ consist of all quadruples (a, b, i, x) such that $x = a - b > 0$ and $a, b \in A_i$. Using the identity $\sum_{i=2-t}^{n} a_i = kt$ and the quadratic-arithmetic mean inequality, we obtain

$$|\mathcal{X}| = \sum_{i=2-t}^{n} \binom{a_i}{2} = \frac{1}{2} \sum_{i=2-t}^{n} a_i^2 - \frac{kt}{2} \geq (1+o(1)) \frac{(kt)^2}{2(n+t)}. \tag{30}$$

For $x \in \mathbb{Z}$, let g_x be the number of representations $x = a - b$ with $a, b \in A$. Then, each $x \in [t-1]$ is included in $g_x(t-x)$ quadruples. Hence,

$$|\mathcal{X}| \leq \sum_{x=0}^{t-1} (t-x) g_x. \tag{31}$$

The above sum can be bounded by

$$\sum_{i=0}^{t-1} (t-i) + \frac{1}{2} \sum_{x=-n+1}^{n-1} t(g_x - 1)_+ = \frac{t^2}{2} + \frac{t(k^2 - |A-A|)}{2} + o(n^2),$$

where $f_+ := f$ if $f > 0$ and $f_+ := 0$ if $f \leq 0$. Putting all together we obtain:

$$\frac{(kt)^2}{2(n+t)} \leq \frac{t^2}{2} + \frac{t(k^2 - |A-A|)}{2} + o(n^2).$$

Routine simplifications yield the claim. ∎

Let us briefly discuss the lower bounds on

$$d(c) := \liminf_{n \to \infty} \frac{d(\lfloor cn^{1/2} \rfloor, n)}{n}.$$

Problem 15 *Does the ratio $d(k,n)/n$ tend to a limit as $n \to \infty$ and $k = (c+o(1))\,n^{1/2}$ where c is fixed?*

Sidon sets show that $d(c) = c^2$ for $0 \le c \le 1$.

Lemma 16 *For $1 \le c \le \sqrt{2}$,*

$$d(c) \ge -\frac{c^4}{3} + 2c^2 - 2 + \frac{4}{3c^2}.$$

Proof. Let $\beta = 1/c^2$ and $b = \lfloor \beta n \rfloor$. Let $B \subset [b]$ be a maximal Sidon set. Let $C = [n] \cap (B+b)$ and $A = B \cup (C+t)$, where t is a small random integer. As B is uniformly distributed in $[b]$, it is easy to see that $|A| = (c+o(1))\sqrt{n}$ is as required.

All differences in $C - B$ are pairwise distinct, except the difference b of multiplicity $O(\sqrt{n})$. So, the densities of $B - B$ and $C - B$ at xn, $0 \le x \le 1$, are respectively $f(x) = 1 - x/\beta$ if $0 \le x \le \beta$ (while $f(x) = 0$ for $x \ge \beta$) and

$$g(x) = \begin{cases} \frac{x}{\beta}, & 0 \le x \le 1-\beta, \\ \frac{1-\beta}{\beta}, & 1-\beta \le x \le \beta, \\ \frac{1-x}{\beta}, & \beta \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(Note that $C - C \subset B - B$, so there is no point to consider $C - C$.)

Now, similarly to our analysis in Lemma 12, the expected size of $A - A$ is

$$(2+o(1))n \int_0^1 (f(x) + g(x) - f(x)g(x))\mathrm{d}x = n\left(\frac{4\beta}{3} - 2 + \frac{2}{\beta} - \frac{1}{3\beta^2} + o(1)\right).$$

By taking t so that $|A-A|$ is at least its expectation, we complete the proof. ∎

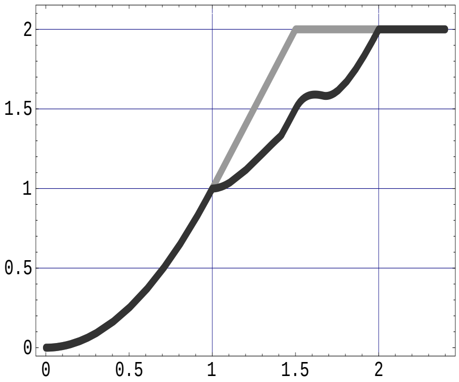
The following construction provides best known lower bounds for the remaining values of c.

Choose some $\beta \le c$ (to be specified later). Let $b = \lfloor \beta\sqrt{n} \rfloor$. Define $B = [b]$. Let C and D be arithmetic progressions of length $\frac{c-\beta}{2}\sqrt{n}$ starting at

$(1-\frac{\beta(c-\beta)}{2}+o(1))n$ but the differences $-b$ and $b-1$ respectively. (Thus, for example, D ends around n.) Let $A=B\cup C\cup D$. Clearly, $(C\cup D)-B$ covers an interval $[(1-\beta(c-\beta)+o(1))n, n-1]$. Also, the distribution of $D-C$ can be explicitly written, which allows us to compute $|A-A|$ asymptotically.

For $c\geq 2$, we can ensure that $A-A=[-n+1,n-1]$; thus $d(c)=2$ then. For $\sqrt{2}\leq c\leq 3/2$, the optimal choice is $\beta=c/3$, giving

$$d(c)\geq 2c^2/3, \quad \sqrt{2}\leq c\leq 3/2.$$



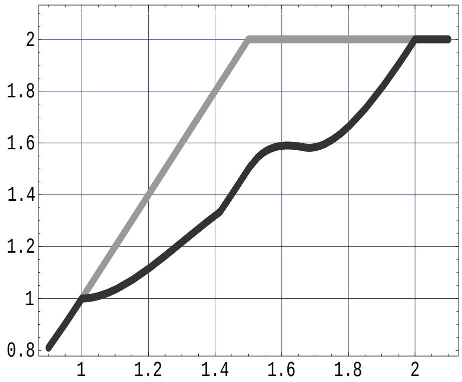


Figure 2: Our bounds on $d(k,n)$: $x=kn^{-1/2}$, $y=d(k,n)/n$.

Unfortunately, it seems that there is no closed formula the optimal $\beta=\beta(c)$ for other values of c. (And, in fact, $\beta(c)$ is not a continuous function.) But, as an illustration, we choose $\beta(c)=c-1$, the linear interpolation given the optimal choices $\beta(3/2)=1/2$ and $\beta(2)=1$. Routine calculations show that the density of $C-B$ at xn, for $0\leq x\leq 1-\beta(c-\beta)=2-c$, is $f(x)=x/(c-1)^2$ if $\frac{5}{3}\leq c\leq 2$ and

$$f(x)=\begin{cases}\frac{x}{(c-1)^2}, & 0\leq x\leq\frac{c-1}{2},\\ \frac{1}{c-1}-\frac{x}{(c-1)^2}, & \frac{c-1}{2}\leq x\leq 2-c,\end{cases}$$

if $\frac{3}{2} \leq x \leq \frac{5}{3}$. This gives us the following lower bounds.

$$d(c) \geq \begin{cases} \frac{4c^3-19c^2+34c-21}{2(c-1)^2}, & \frac{3}{2} \leq c \leq \frac{5}{3}, \\ \frac{2c^3-5c^2+2c+2}{(c-1)^2}, & \frac{5}{3} \leq c \leq 2. \end{cases}$$

Figure 2 contains the corresponding graph. (The obtained bounds are not always monotone with respect to c because our choice of β was not optimal.)

References

[1] R. C. Bose and S. Chowla, *Theorems in the additive theory of numbers*, Comment. Math. Helvet. **37** (1962–63), 141–147.

[2] J. Cilleruelo, I. Z. Ruzsa, and C. Trujillo, *Upper and lower bounds for finite $B_h[g]$ sequences*, J. Number Theory **97** (2002), 26–34.

[3] D. Craft and E. Tesar, *On a question of Erdős about edge-magic graphs*, Discrete Math. **207** (1999), 271–279.

[4] P. Erdős, *A survey of problems in combinatorial number theory*, Ann. Discrete Math. **6** (1980), 89–115.

[5] P. Erdős and R. Freud, *On sums of a Sidon-sequence*, J. Number Theory **38** (1991), 196–205.

[6] P. Erdős and P. Turán, *On a problem of Sidon in additive number theory, and on some related problems*, J. Lond. Math. Soc. **16** (1941), 212–215.

[7] K. Fried, *Rare bases for finite intervals of integers*, Acta Sci. Math. **52** (1988), 303–305.

[8] J. A. Gallian, *A dynamic survey of graph labeling*, Electronic J. Combin. **DS6** (2002), 106pp.

[9] B. Green, *The number of squares and $B_h[g]$ sets*, Acta Arithm. **100** (2001), 365–390.

[10] H. Halberstam and K. F. Roth, *Sequences*, Springer Verlag, 1983.

[11] W. Klotz, *Eine obere Schranke für die Reichweite einer Extremalbasis zweiter ordnung*, J. Reine Angew. Math **238** (1969), 161–168.

[12] T. W. Körner, *Fourier analysis*, Cambridge Univ. Press, 1988.

[13] A. Kotzig and A. Rosa, *Magic valuations of finite graphs*, Can. Math. Bull. **13** (1970), 451–561.

[14] B. Lindström, *Well distributin of Sidon sets in residue classes*, J. Number Theory **69** (1998), 197–200.

[15] G. Martin and K. O'Bryant, *Continuous Ramsey theory and Sidon sets*, Preprint, 2002.

[16] L. Moser, *On the representation of* $1, 2, \ldots, n$ *by sums*, Acta Arithm. **6** (1960), 11–13.

[17] L. Moser, J. R. Pounder, and J. Riddell, *On the cardinality of h-bases for n*, J. Lond. Math. Soc. **44** (1969), 397–407.

[18] A. Mrose, *Untere Schranken für die Reichweiten von Extremalbasen fester Ordnung*, Abh. Math. Sem. Univ. Hamburg **48** (1979), 118–124.

[19] H. Rohrbach, *Ein Beitrag zur additiven Zahlentheorie*, Math. Zeitschrift **42** (1937), 1–30.

[20] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Trans. Amer. Math. Soc. **43** (1938), 377–85.

[21] D. R. Wood, *On vertex-magic and edge-magic total injections of graphs*, Australas. J. Comb. **26** (2002), 49–63.