Erdős #819

REFLECTED SIDON LOWER BOUND

YU LEON LIU

ABSTRACT. We sharpen the lower bound for Erdős Problem #819: writing $f(N) := \max\{|(A+A)\cap[1,N]| : A \subseteq \{1,\dots,N\},\ |A| = \lfloor\sqrt{N}\rfloor\}$, we prove $\liminf_{N\to\infty} f(N)/N \geq (16\sqrt{2}-17)/12 \approx 0.4690$, improving the classical Erdős–Freud bound of $3/8 = 0.375$ and closing most of the gap to the trivial upper bound $f(N)/N \leq 1/2 + o(1)$. The construction is a reflected two-copy of a Bose–Chowla Sidon set with a small uniform random shift; the proof uses Pikhurko's uniformity lemma.

Erdős Problem #819 [1] concerns the maximum size of the sumset of a square-root-density subset of $[1, N]$:

Definition 1. For $N \geq 1$, set

$$f(N) \;:=\; \max\Big\{\, |(A+A)\cap[1,N]| \;:\; A\subseteq\{1,\dots,N\},\ |A| = \lfloor N^{1/2}\rfloor \Big\}.$$

Since $|A+A| \leq \binom{|A|+1}{2}$ for any set A, we have the trivial upper bound

$$f(N) \;\leq\; \binom{\lfloor\sqrt{N}\rfloor+1}{2} \;\leq\; N/2 + O(\sqrt{N}).$$

The classical lower bound, due to Erdős and Freud [2],

$$f(N) \;\geq\; \left(\tfrac{3}{8} - o(1)\right) N, \tag{2}$$

is given by an explicit construction (a union of an arithmetic progression and a structured complement).

In this note we sharpen the lower bound:

Theorem 3. *We have*

$$\liminf_{N\to\infty} \frac{f(N)}{N} \;\geq\; \frac{16\sqrt{2}-17}{12} \;\approx\; 0.4690.$$

The proof outline is as follows: the construction is a reflected two-copy of a maximum Sidon set with some displacement, for the subsequence $N = 4q^2$ with q a sufficiently large positive integer. We control the overlap of the Sidon set and its reflected copy by Pikhurko's uniformity lemma [3, Lemma 10] for asymptotically maximum Sidon sets, which is the core technical statement of this note. Using random-shift optimization, we get a closed form formula (21) for the expected score, which we optimize (Corollary 25) to obtain the bound. Lastly, an interpolation argument (proof of Theorem 3) lifts from the subsequence $N = 4q^2$ to all sufficiently large N. This construction and proof strategy are inspired by those of Pikhurko [3, Lemma 12].

Date: May 15, 2026.

We start the proof by recalling the notion of a Sidon set:

Definition 4. A finite set S is a *Sidon* set if all ordered sums $a+b$ with $a,b\in S$ are distinct up to the unordered pairs $\{a,b\}$.

The core technical input we need is the following uniformity statement due to Pikhurko [3, Lemma 10]. It is a common generalization of an earlier result of Erdős–Freud [2, Lemma 1] (uniformity in subintervals) and of Lindström [6] (uniformity in residue classes).

Lemma 5 (Pikhurko's uniformity lemma)**.** *Let q be large and let $S\subset[0,M]$ be an asymptotically maximum Sidon set of size q, so $M=(1+o(1))\,q^2$. Then for any subinterval $I\subseteq[0,M]$, any fixed integer $m\geq 1$, and any residue class $\ell \pmod m$,*

$$\bigl|S\cap I\cap(\ell+m\mathbb{Z})\bigr| \;=\; \frac{|I|}{m\,q} \;+\; o(q).$$

Remark 6. Conceptually, S has density $\approx 1/q$ in $[0,M]$, and the lemma says this density is preserved on every macroscopic subinterval and every residue class simultaneously, with additive error $o(q)$. (This is the form of Pikhurko's [3, Lemma 10] statement obtained by the substitution $n=M$, $\sqrt{n}=q$.)

We will need such asymptotically maximum Sidon sets to exist for every large q (not merely along a sparse subsequence such as prime powers). This is classical:

Lemma 7 (existence of asymptotically maximum Sidon sets)**.** *For every sufficiently large integer q, there exists a Sidon set $S\subset[0,M]$ with $|S|=q$ and $M=(1+o(1))q^2$.*

Proof. Let p be the smallest prime with $p\geq q$; by the prime number theorem, $p=(1+o(1))q$. The Bose–Chowla construction [5] produces a Sidon set $S^*\subset[0,p^2-2]$ with $|S^*|=p$. Take any subset $S\subseteq S^*$ with $|S|=q$ and set $M:=\max S$; any subset of a Sidon set is Sidon, so S is Sidon. Then $M\leq p^2-2=(1+o(1))q^2$. Conversely, the classical upper bound $|T|\leq M^{1/2}+O(M^{1/4})$ for any Sidon set $T\subset[0,M]$ (Erdős–Turán; see also Lindström [6]) applied to S gives $q\leq M^{1/2}+O(M^{1/4})$, so $M\geq(1-o(1))q^2$. □

We now consider the sumset and difference set of S:

Definition 8. We write

$$\Sigma\coloneqq S+S,\quad \Delta\coloneqq S-S.$$

By the Sidon property,

$$|\Sigma|=\binom{q+1}{2}=\frac{q(q+1)}{2},\quad |\Delta|=q(q-1)+1.$$

Lemma 5 allows us to give a density description of Σ and Δ on macroscopic windows. Recall that $N=4q^2$, so $2M/N\to 1/2$. Let us define the triangular densities:

$$\phi(y)\;:=\;\begin{cases} y, & 0\leq y\leq 1/2,\\ 1-y, & 1/2\leq y\leq 1,\\ 0, & \text{otherwise,}\end{cases}\qquad \psi(y)\;:=\;\max\bigl(0,\;1-2|y|\bigr). \tag{9}$$

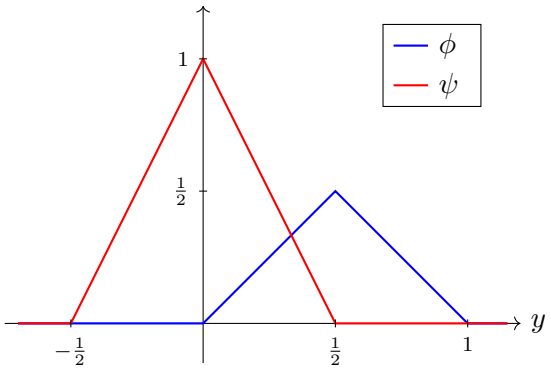

FIGURE 1. The triangular densities ϕ (blue) and ψ (red).

We illustrate ϕ and ψ below:

We now derive density descriptions for Σ and Δ on macroscopic windows.

Lemma 10. *Let q be large, and $S \subset [0, M]$ be an asymptotically maximum Sidon set of size q, so $M = (1 + o(1))q^2$, and $N = 4q^2$. Fix $\eta > 0$ and $x \in \mathbb{R}$, and let $J \subset \mathbb{Z}$ be the integer interval $[\lfloor x \cdot N/2 \rfloor,\ \lfloor x \cdot N/2 \rfloor + \lfloor \eta N/2 \rfloor)$. It has size $|J| = \lfloor \eta N/2 \rfloor$. Then for any fixed integer $m \geq 1$ and any residue class $\ell \pmod m$,*

$$\frac{|(S+S) \cap J \cap (\ell + m\mathbb{Z})|}{|J|/m} = \phi(x) + O(\eta) + o_q(1),$$
$$\frac{|(S-S) \cap J \cap (\ell + m\mathbb{Z})|}{|J|/m} = \psi(x) + O(\eta) + o_q(1).$$

Here the implicit constant in $O(\eta)$ is absolute (i.e., independent of the other parameters), and $o_q(1)$ is a quantity that, for each fixed $\eta > 0$ and m, tends to 0 as $q \to \infty$ uniformly in $x \in \mathbb{R}$ and $\ell \pmod m$.

Proof. We prove the statement for $S + S$; the argument for $S - S$ is analogous.

Fix a small parameter $0 < \rho \ll \eta$. Partition $[0, M]$ into intervals I_k of length $\rho N/2$. By Lemma 5 (with fixed modulus m), for each residue class $r \pmod m$ and each interval I_k,

$$|S \cap I_k \cap (r + m\mathbb{Z})| = \frac{|I_k|}{m\, q} + o(q),$$

uniformly in k and r.

Now fix the window $J = [\lfloor xN/2 \rfloor,\ \lfloor xN/2 \rfloor + \lfloor \eta N/2 \rfloor)$. We count ordered pairs $(a, b) \in S \times S$ with $a + b \in J$ and $a + b \equiv \ell \pmod m$. Writing $a \in I_i$ and $b \in I_j$, this count equals

$$\sum_{i,j} \sum_{\substack{r_i, r_j \pmod m \\ r_i + r_j \equiv \ell}} |S \cap I_i \cap (r_i + m\mathbb{Z})| \cdot |S \cap I_j \cap (r_j + m\mathbb{Z})|,$$

where the outer sum is restricted to pairs (i, j) for which $I_i + I_j$ intersects J. By the uniformity estimate above and a Riemann-sum approximation, this count equals

$$(\phi(x) + O(\eta)) \cdot \frac{|J|\, q}{m} + o(q\,|J|).$$

Passing from ordered pairs to unordered pairs introduces an $O(q) = o(|J|)$ correction from the diagonal $a = b$. By the Sidon property, distinct unordered pairs map injectively to distinct elements of $\Sigma = S + S$. Therefore

$$|\Sigma \cap J \cap (\ell + m\mathbb{Z})| = \big(\phi(x) + O(\eta)\big) \cdot \frac{|J|}{m} + o(|J|),$$

and division by $|J|/m$ gives the first statement.

The proof for $\Delta = S - S$ is analogous, with ψ in place of ϕ (the density of differences of two independent uniform points on $[0, 1/2]$). The zero difference $a = b$ contributes only $O(q) = o(|J|)$, and all nonzero differences are unique by the Sidon property. □

We now turn to the construction, which is probabilistic; we derandomize at the end. Fix any $u \in (0, 1/4)$ and any $\eta \in (0, u/2)$. Let $S \subset [0, M]$ be an asymptotically maximum Sidon set of size q with $M = (1 + o(1))q^2$, and $N = 4q^2$. Let a, b be independent integers chosen uniformly at random from $\{1, \dots, \lfloor \eta N/2 \rfloor\}$. Define

$$B := a + S, \quad C := N/2 + \lfloor uN/2 \rfloor - b - S, \quad A := B \cup C.$$

Observation 11. Note that

$$B \subset [1,\ (\tfrac{1}{2} + \eta + o(1))\tfrac{N}{2}], \quad C \subset [(\tfrac{1}{2} + u - \eta - o(1))\tfrac{N}{2},\ (1 + u)\tfrac{N}{2}].$$

Both are subsets of $[1, N]$, so $A \subset [1, N]$. Moreover, B and C are disjoint for q sufficiently large:

$$\max B \le (\tfrac{1}{2} + \eta + o(1))\tfrac{N}{2} \ < \ (\tfrac{1}{2} + u - \eta - o(1))\tfrac{N}{2} \le \min C \tag{12}$$

since $\eta < u/2$. Hence $|A| = |B| + |C| = q + q = 2q = \sqrt{N}$.

Observe that $A + A = (B + B) \cup (B + C) \cup (C + C)$. We compute the pointwise densities:

Lemma 13. *Uniformly for $v \in [1, N]$, with $x := 2v/N$,*

$$\Pr(v \in B + B) \ = \ \phi(x) + O(\eta) + o_q(1), \tag{14}$$

$$\Pr(v \in C + C) \ = \ \phi(2 + 2u - x) + O(\eta) + o_q(1), \tag{15}$$

$$\Pr(v \in B + C) \ = \ \psi(x - 1 - u) + O(\eta) + o_q(1), \tag{16}$$

$$\Pr(v \in (B + B) \cap (B + C)) \ = \ \phi(x)\,\psi(x - 1 - u) + O(\eta) + o_q(1), \tag{17}$$

$$\Pr(v \in (B + C) \cap (C + C)) \ = \ \psi(x - 1 - u)\,\phi(2 + 2u - x) + O(\eta) + o_q(1). \tag{18}$$

Moreover, for q sufficiently large (depending on η and u), $(B+B)\cap(C+C) = \emptyset$; in particular, $\Pr(v \in (B + B) \cap (C + C)) = 0$ and $\Pr(v \in (B + B) \cap (B + C) \cap (C + C)) = 0$.

Proof. Lines (14)–(16) follow directly from Lemma 10 (in each case as a single-family density query for Σ or Δ over a window of a or b, uniformly in v).

For the disjointness $(B+B)\cap(C+C) = \emptyset$ when q is large enough, observe that $\max(B+B) = 2\max B \le (1 + 2\eta + o(1))\frac{N}{2}$ while $\min(C + C) = 2\min C \ge (1 + 2u - 2\eta - o(1))\frac{N}{2}$. Since $\eta < u/2$ is strict, the gap $2u - 4\eta > 0$ is fixed; for q large enough that the $o(1)$ correction is smaller than this gap, $\max(B + B) < \min(C + C)$ strictly.

It remains to prove (17). First, note that the event $\{v \in B + B\}$ is determined by a alone, while the event $\{v \in B + C\}$ depends on both a and b. Conditioning on $a = a_0$,

$$\Pr(v \in B + C \mid a = a_0) = \Pr_b(v - N/2 - \lfloor uN/2 \rfloor - a_0 + b \in \Delta),$$

which is a single-family density query for $\Delta = S - S$ over a window of b of length $\lfloor \eta N/2 \rfloor$ centered at $(x - 1 - u)N/2 - a_0 \in \mathbb{Z}$. By Lemma 10 this equals $\psi(x - 1 - u) + O(\eta) + o_q(1)$ (the a_0 shift of the window center is $a_0/(N/2) = O(\eta)$, absorbed by Lipschitz continuity of ψ). Since this conditional probability is approximately the same for all a_0 (up to $O(\eta) + o_q(1)$), it follows that

$$\Pr(v \in (B + B) \cap (B + C)) = \Pr(v \in B + B) \cdot (\psi(x - 1 - u) + O(\eta) + o_q(1)),$$

so (17) follows from (14). Lastly, (18) follows from the same argument with a, b swapped. □

Remark 19. Lemma 13 says that the pointwise densities of $B + B$, $B + C$, $C + C$ and their pairwise intersections are approximately constant (up to error $O(\eta) + o_q(1)$) on macroscopic windows.

We are now ready for our main theorem:

Theorem 20. *With the random construction above and $N = 4q^2$, for any $u \in (0, 1/4)$ and $\eta \in (0, u/2)$ we have*

$$\mathbb{E}(|(A + A) \cap [1, N]|) \geq \tfrac{N}{2} \cdot F(u) - O(\eta N) - o(N),$$

where

$$F(u) := 1 - 2u^2 - \frac{(1 - 2u)^3}{12}. \tag{21}$$

Proof. By Observation 11 we have $A \subset [1, N]$ and $|A| = 2q$. For q sufficiently large, $(B + B) \cap (C + C) = \emptyset$ by Lemma 13, and a fortiori the triple intersection also vanishes. Applying inclusion-exclusion to $A + A = (B + B) \cup (B + C) \cup (C + C)$,

$$\begin{aligned}\Pr(v \in A + A) = {} & \Pr(v \in B + B) + \Pr(v \in C + C) + \Pr(v \in B + C) \\ & - \Pr(v \in (B + B) \cap (B + C)) - \Pr(v \in (B + C) \cap (C + C)).\end{aligned}$$

By Lemma 13, uniformly in v,

$$\Pr(v \in A + A) = (p + d + r - pd - dr)(x) + O(\eta) + o_q(1),$$

where we use the shorthand:

$$p(x) = \phi(x), \quad d(x) = \psi(x - 1 - u), \quad r(x) = \phi(2 + 2u - x).$$

Summing over $v \in [1, N]$ yields a Riemann sum in $x = 2v/N \in [0, 2]$,

$$\mathbb{E}(|(A + A) \cap [1, N]|) \;\geq\; \frac{N}{2} \int_0^2 (p + d + r - pd - dr)\, dx - O(\eta N) - o(N). \tag{22}$$

It remains to compute those integrals for $u \in (0, 1/4)$. Recall that the area under ϕ on $[0, 1]$ is $\int_0^1 \phi(y)\,dy = 1/4$, and the area under ψ on $[-1/2, 1/2]$ is $\int_{-1/2}^{1/2} \psi(y)\,dy = 1/2$. Hence

$$\int_0^2 p\,dx \;=\; \int_0^1 \phi(y)\,dy \;=\; \tfrac{1}{4}, \qquad \int_0^2 d\,dx \;=\; \int_{-1/2}^{1/2} \psi(y)\,dy \;=\; \tfrac{1}{2}.$$

For r, use the substitution $y = 2 + 2u - x$:

$$\int_0^2 r\,dx \;=\; \int_{2u}^{2+2u} \phi(y)\,dy \;=\; \int_0^1 \phi(y)\,dy - \int_0^{2u} y\,dy \;=\; \tfrac{1}{4} - 2u^2,$$

valid since ϕ vanishes outside $[0, 1]$ and $2u \le 1/2$. Thus

$$\int_0^2 (p + d + r)\,dx \;=\; \tfrac{1}{4} + \tfrac{1}{2} + \tfrac{1}{4} - 2u^2 \;=\; 1 - 2u^2. \tag{23}$$

For the overlaps, pd is supported on $\operatorname{supp}(\phi(x)) \cap \operatorname{supp}(\psi(x - 1 - u)) = [1/2 + u, 1]$, where $\phi(x) = 1 - x$ and $\psi(x - 1 - u) = 2x - (1 + 2u)$. With the substitution $y = x - (1/2 + u)$,

$$\int_0^2 pd\,dx \;=\; \int_0^{1/2-u} \big((1 - 2u)y - 2y^2\big)\,dy \;=\; \frac{(1/2 - u)^3}{3} \;=\; \frac{(1 - 2u)^3}{24}.$$

By the reflection symmetry $x \mapsto 2(1 + u) - x$ (which swaps p with r while fixing d, hence swaps pd with dr while preserving the integral),

$$\int_0^2 pd\,dx \;=\; \int_0^2 dr\,dx \;=\; \frac{(1 - 2u)^3}{24}. \tag{24}$$

Combining (23) and (24),

$$\int_0^2 (p + d + r - pd - dr)\,dx \;=\; 1 - 2u^2 - \frac{(1 - 2u)^3}{12} \;=\; F(u).$$

Substituting into (22) gives the displayed expectation bound. □

We now optimize u to maximize $F(u)$:

Corollary 25. *The maximum of $F(u)$ on $(0, 1/4)$ is achieved at $u^* = \frac{3}{2} - \sqrt{2} \approx 0.0858$, with value*

$$F(u^*) \;=\; \tfrac{16\sqrt{2} - 17}{6}.$$

Setting $c := \frac{F(u^*)}{2} = \frac{16\sqrt{2} - 17}{12} \approx 0.4690$, by Theorem 20 we have

$$\mathbb{E}(|(A + A) \cap [1, N]|) \;\ge\; N \cdot c - O(\eta N) - o(N). \tag{26}$$

Proof. $F'(u) = -4u + \frac{(1-2u)^2}{2} = \frac{4u^2 - 12u + 1}{2}$. The roots are $u = \frac{3 \pm 2\sqrt{2}}{2} = \frac{3}{2} \pm \sqrt{2}$; only $u^* = \frac{3}{2} - \sqrt{2} \in (0, 1/4)$, and $F''(u) = -4 - 2(1 - 2u) < 0$ on this interval, so u^* is the unique maximum. Direct substitution gives the displayed value of $F(u^*)$. □

Figure 2 illustrates the three density profiles p, d, r at the three values $u = 0$, $u = u^* = \frac{3}{2} - \sqrt{2}$, and $u = 1/4$. At $u = 0$ the supports of d and r overlap heavily with p and d respectively (large overlap defects, low score). At $u = 1/4$ the support of r is pushed past $x = 2$ and gets truncated (also low score). The optimum $u = u^*$ balances these two failure modes.

We are now ready to prove Theorem 3.

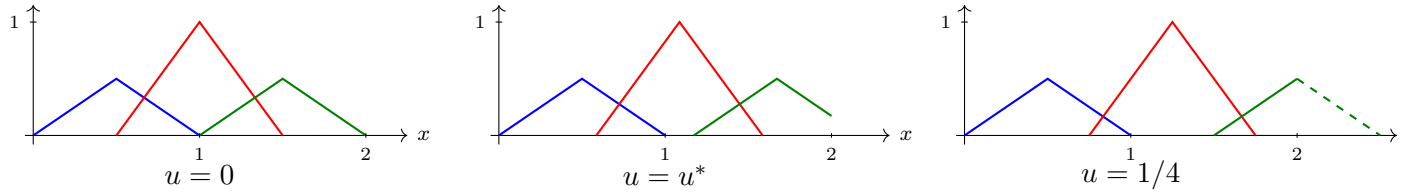

FIGURE 2. The three local density profiles p, d, r on $[0, 2]$ at $u = 0$, $u = u^* = \frac{3}{2} - \sqrt{2} \approx 0.0858$, and $u = 1/4$. At $u = u^*$ the support of r just begins to spill past $x = 2$ (small truncation); at $u = 1/4$ the truncation is substantial (dashed portion).

Proof of Theorem 3. We first establish the bound

$$\liminf_{N\to\infty} \frac{f(N)}{N} \;\geq\; c \;\approx\; 0.4690 \tag{27}$$

along the subsequence $N = 4q^2$. Fix any $\eta \in (0, u^*/2)$. By Corollary 25, for q large,

$$\mathbb{E}(|(A + A) \cap [1, N_q]|) \geq N_q \cdot c - O(\eta N_q) - o(N_q),$$

where the implicit constant in $O(\eta)$ is absolute and $o(N_q)$ vanishes as $q \to \infty$ (for fixed η). Pick, for each q, a realization (a, b) achieving at least the expectation; this produces deterministic sets $A_q^{(\eta)} \subset [1, N_q]$ with $|A_q^{(\eta)}| = 2q$ and

$$|(A_q^{(\eta)} + A_q^{(\eta)}) \cap [1, N_q]| \geq c \cdot N_q - O(\eta N_q) - o(N_q).$$

Dividing by N_q and taking $\liminf_{q\to\infty}$ gives, for every $\eta \in (0, u^*/2)$,

$$\liminf_{q\to\infty} \frac{f(N_q)}{N_q} \geq c - O(\eta).$$

Since the $O(\eta)$ constant is absolute, letting $\eta \downarrow 0$ yields $\liminf_{q\to\infty} f(N_q)/N_q \geq c$, which is (27) along the subsequence $N = 4q^2$.

It remains to lift this bound to all N.

Given any large N, set $q := \lfloor \sqrt{N}/2 \rfloor$, so that

$$N_q = 4q^2 \;\leq\; N \;<\; 4(q+1)^2 = N_{q+1}, \qquad N_{q+1} - N_q = 8q + 4 = O(\sqrt{N}). \tag{28}$$

In particular $N - N_q \leq 8q + 4 = o(N)$ and $q \to \infty$ as $N \to \infty$.

Let $A_q \subset [1, N_q]$ be any admissible competitor for $f(N_q)$ (i.e., $|A_q| = 2q$). Since $A_q \subset [1, N_q] \subseteq [1, N]$ and $\lfloor \sqrt{N} \rfloor - 2q \in \{0, 1\}$, we may extend A_q to a set $A := A_q \cup E$ with $|A| = \lfloor \sqrt{N} \rfloor$ and $A \subset [1, N]$ by adjoining at most one extra point. Adding points only enlarges the sumset, so $|(A + A) \cap [1, N]| \geq |(A_q + A_q) \cap [1, N_q]|$. Taking the maximum over A_q gives $f(N) \geq f(N_q)$. Hence

$$\frac{f(N)}{N} \;\geq\; \frac{f(N_q)}{N} \;=\; \frac{f(N_q)}{N_q} \cdot \frac{N_q}{N} \;\geq\; (c - o_q(1)) \cdot (1 + O(N^{-1/2})) \;=\; c - o(1),$$

where we used the subsequence bound $\liminf_{q\to\infty} f(N_q)/N_q \geq c$ in the second inequality. Taking $\liminf_{N\to\infty}$ gives the displayed inequality. □

Acknowledgements

The proof in this note was discovered with the assistance of OpenAI's GPT-5.5 together with the Rethlas open-source agentic mathematics-research pipeline. All mathematical arguments and claims in the final manuscript were independently verified by the author, who takes full responsibility for the paper.

References

[1] T. F. Bloom, *Erdős Problem #819*, https://www.erdosproblems.com/819, accessed May 15, 2026. 1

[2] P. Erdős and R. Freud, *On Sidon sequences and related problems*, Mat. Lapok (N.S.) **1** (1991), 1–44. 1, 2

[3] O. Pikhurko, *Dense edge-magic graphs and thin additive bases*, Discrete Mathematics **306** (2006), 2097–2107. 1, 2

[4] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Trans. Amer. Math. Soc. **43** (1938), 377–385.

[5] R. C. Bose and S. Chowla, *Theorems in the additive theory of numbers*, Comment. Math. Helv. **37** (1962/63), 141–147. 2

[6] B. Lindström, *Well distribution of Sidon sets in residue classes*, J. Number Theory **69** (1998), 197–200. 2

1 Oxford St, Cambridge, MA 02139

Email address: yuleonliu@math.harvard.edu