Conjectures.io

Formal proof verified through Conjectures

GREEN'S PROBLEM 51 AT DENSITY ONE HALF: SUBSPACES OF BOUNDED CODIMENSION IN BINARY SUMSETS

JENSEN KOHLMEYER AND LIAM KRUER

Abstract. Let $A \subseteq \mathbb{F}_2^n$ have density at least $1/2 - \varepsilon$, where $0 \le \varepsilon \le 1/4$. We prove that $A + A$ contains a linear subspace of codimension at most $1 + 256en\varepsilon^2$. Consequently, for every fixed $k > 0$, density at least $1/2 - k/\sqrt{n}$ forces a subspace of codimension at most $\lceil 1 + 256ek^2 \rceil$ for all sufficiently large n. This answers the near-half-density question in Sanders' Question 5.1 and Green's Problem 51. The proof combines a variance bound for ternary separators of the Boolean cube, a Fourier argument producing linear cuts of missing sums, and a polynomial interpolation bound on minimal obstructions to covering by such cuts. A Lean formalization and a public verification record accompany the argument; the final section gives the proof correspondence and reproduction details.

1. Introduction

Write $G = \mathbb{F}_2^n$, give G normalized counting measure, and let

$$\alpha = \frac{|A|}{2^n}, \qquad A + A = \{a + a' : a, a' \in A\}.$$

Green's Problem 51 [1] asks for the largest affine subspace guaranteed in a sumset at a given density. Its near-half-density question, also posed for linear subspaces in Sanders' Question 5.1 [2], asks whether every fixed $k > 0$ in the hypothesis $\alpha > 1/2 - k/\sqrt{n}$ allows a bound on codimension independent of n. We prove the following quantitative result.

Theorem 1.1. *Let $n \ge 0$, $0 \le \varepsilon \le 1/4$, and $A \subseteq \mathbb{F}_2^n$ with $|A|/2^n \ge 1/2 - \varepsilon$. There is a linear subspace $H \le \mathbb{F}_2^n$ such that*

$$H \subseteq A + A, \qquad \operatorname{codim} H \le 1 + 256en\varepsilon^2. \tag{1}$$

Here $e = \exp(1)$ and $\operatorname{codim} H = n - \dim H$.

Corollary 1.2. *Fix $k > 0$, and set*

$$c(k) = \lceil 1 + 256ek^2 \rceil, \qquad N(k) = \lceil 16k^2 \rceil + 1.$$

For every $n \ge N(k)$, any set $A \subseteq \mathbb{F}_2^n$ of density at least $1/2 - k/\sqrt{n}$ has a sumset containing a linear subspace of dimension at least $n - c(k)$.

If $\alpha > 1/2$, then $A + A = G$: for every $x \in G$, the two sets A and $x + A$ intersect. Sanders [2, Theorem 1.2] proved that a sufficiently small absolute constant $c > 0$ in the hypothesis $\alpha > 1/2 - c/\sqrt{n}$ still guarantees a subspace of codimension one. Corollary 1.2 allows every fixed k, with

Date: Working draft — 9 September 2026.

Key words and phrases. Sumsets, vector spaces over finite fields, Boolean cube, Walsh transform, polynomial interpolation, Lean formalization.

a quadratic bound on codimension. Sanders also gives an upper bound $O(n/\log(\varepsilon^{-1}))$ at density $1/2-\varepsilon$ [2, Theorem 1.4].

Following Question 5.1, Sanders suggested that an $O(k^2)$ bound might be accessible. Corollary 1.2 establishes that quadratic scale, with an explicit constant. The Green–Ruzsa construction recorded in [2, Theorem 1.3] requires codimension $\Omega(k)$ for sufficiently large k at this density scale. Thus the optimal dependence on k remains between linear and quadratic growth.

The theorem concerns the near-half-density subproblem, rather than the exact extremal function at all densities. Its conclusion supplies a linear subspace, which is also an affine coset. No optimality of the constant $256e$, or of the quadratic dependence on k, is asserted.

1.1. **Outline of the argument.** Let $B = G \setminus (A+A)$ be the set of missing sums. We first show that every nonempty $D \subseteq B$ admits a linear functional $\ell : G \to \mathbb{F}_2$ such that

$$|D \cap \ker \ell| \le 256\varepsilon^2 |D|. \tag{2}$$

The proof lifts a difference of indicator functions to a Boolean cube with one coordinate for each element of D. A parity factor turns the lift into a ternary separator. Small zero mass then forces a nonzero Fourier coefficient of small degree, which yields (2).

Iterating cuts gives exponential decay in the number of remaining points. To obtain the required dependence on $n\varepsilon^2$, we combine this decay with a bound on a minimal subset of B that cannot be covered by r complements of kernels. Interpolation by polynomials of degree at most r bounds such a subset by $\sum_{j\le r}\binom{n}{j}$. The resulting inequality bounds the number of cuts by $1+256en\varepsilon^2$. Their common kernel is contained in $A+A$.

All vector spaces and linear functionals below are over $\mathbb{F}_2$. Expectations use uniform probability measure unless stated otherwise; Fourier coefficients and variances are real-valued. Sections 2–6 give the mathematical proof, and Section 7 records its formal correspondence.

2. A variance bound for cube separators

Put $Q_m = \{0,1\}^m$, with addition modulo two and uniform probability measure. For a real function f on a finite probability space, write

$$z(f) = \mathbb{P}(f=0), \qquad \operatorname{Var}(f) = \mathbb{E}f^2 - (\mathbb{E}f)^2.$$

Definition 2.1. A *ternary separator* on Q_m is a function $f : Q_m \to \{-1,0,1\}$ such that adjacent vertices never have opposite nonzero values: if x, y differ in one coordinate, then $f(x)f(y) \ne -1$.

Lemma 2.2 (Separator variance)**.** *Every ternary separator f on Q_m satisfies*

$$\operatorname{Var}(f) \le 2\sqrt{m}\, z(f). \tag{3}$$

Proof. We induct on m. For $m=0$ the cube has one point and the variance is zero. For the induction step, write f_0, f_1 for the two slices of a separator on Q_{m+1}, and put $z_i = z(f_i)$ and $\mu_i = \mathbb{E}f_i$. Both slices are separators. Across the two slices, the separation condition gives pointwise

$$|f_0(x) - f_1(x)| \le \mathbf{1}_{\{f_0(x)=0\}} + \mathbf{1}_{\{f_1(x)=0\}}.$$

Consequently $|\mu_0 - \mu_1| \le z_0 + z_1$. With $z = z(f) = (z_0+z_1)/2$, the variance decomposition and induction give

$$\begin{aligned}\operatorname{Var}(f) &= \frac{\operatorname{Var}(f_0)+\operatorname{Var}(f_1)}{2} + \frac{(\mu_0-\mu_1)^2}{4}\\ &\le 2\sqrt{m}\, z + z^2.\end{aligned} \tag{4}$$

Also $\operatorname{Var}(f) \le 1$, since $|f| \le 1$.

If $\sqrt{m+1}\, z \ge 1$, this last bound implies the claimed estimate. Otherwise,

$$z < \frac{1}{\sqrt{m+1}} \le \frac{2}{\sqrt{m+1}+\sqrt{m}} = 2(\sqrt{m+1}-\sqrt{m}).$$

Substituting this bound for one factor of z^2 in (4) proves (3) in dimension $m+1$. □

The estimate applies to each coordinate face. Averaging over random faces will relate the zero mass to the Fourier degrees of the separator.

3. Random faces and a low-degree Walsh coefficient

For $S \subseteq [m] = \{1, \dots, m\}$, let

$$\chi_S(x) = (-1)^{\sum_{i\in S} x_i}, \qquad \widehat{f}(S) = \mathbb{E}_x f(x)\chi_S(x).$$

The characters are orthonormal: $\mathbb{E}\chi_S\chi_T$ equals one for $S = T$, and otherwise vanishes by averaging in a coordinate of $S \triangle T$. There are 2^m characters, so they form a basis. This gives Fourier inversion and Parseval:

$$f = \sum_{S\subseteq[m]} \widehat{f}(S)\chi_S, \qquad \sum_S \widehat{f}(S)^2 = \mathbb{E}f^2. \tag{5}$$

For ternary f, the last quantity is $1 - z(f)$.

Lemma 3.1 (Random-face energy)**.** *Let f be a ternary separator on Q_m. For $0 \le t \le 1$,*

$$\sum_{S\subseteq[m]} \widehat{f}(S)^2\big(1-(1-t)^{|S|}\big) \le 2\sqrt{mt}\, z(f). \tag{6}$$

Proof. Fix a subset $J \subseteq [m]$ of free coordinates. For each choice of the other coordinates, restrict f to the corresponding J-face. Let V_J be the average variance of these restrictions. Their average zero mass is $z(f)$, so Lemma 2.2 gives

$$V_J \le 2\sqrt{|J|}\, z(f). \tag{7}$$

The face average of a character χ_S is zero unless $S \cap J = \varnothing$, in which case it is the same character on the fixed coordinates. By orthogonality, the average square of the face mean is therefore $\sum_{S: S\cap J=\varnothing} \widehat{f}(S)^2$. Subtracting this from $\mathbb{E}f^2$ shows

$$V_J = \sum_{S: S\cap J\neq\varnothing} \widehat{f}(S)^2.$$

Now include each coordinate in J independently with probability t. Then $\mathbb{P}(S\cap J = \varnothing) = (1-t)^{|S|}$, while $\mathbb{E}|J| = mt$ and $\mathbb{E}\sqrt{|J|} \le \sqrt{mt}$ by Cauchy–Schwarz. Averaging (7) proves (6). □

Lemma 3.2 (A low-degree coefficient)**.** *If f is a ternary separator on Q_m and $z = z(f) \le 1/2$, there exists $S \subseteq [m]$ such that*

$$\widehat{f}(S) \neq 0, \qquad |S| \le 64mz^2. \tag{8}$$

Proof. Parseval gives $\sum_S \widehat{f}(S)^2 = 1 - z \ge 1/2$, so at least one coefficient is nonzero. Let s be the least degree of a nonzero coefficient. If $s = 0$, there is nothing to prove. Otherwise apply Lemma 3.1 with $t = 1/s$.

For every integer $l \ge s$,

$$1 - (1-1/s)^l \ge \frac{1}{2}.$$

For $s = 1$ this is immediate. For $s > 1$, Bernoulli's inequality gives $(1-1/s)^{-s} = (1+1/(s-1))^s \ge 1 + s/(s-1) \ge 2$, which proves the assertion for $l = s$ and hence for $l \ge s$. All nonzero coefficients have degree at least s, so

$$\frac{1-z}{2} \le \sum_S \widehat{f}(S)^2\big(1-(1-1/s)^{|S|}\big) \le 2\sqrt{m/s}\, z.$$

Squaring nonnegative quantities yields $s(1-z)^2 \le 16mz^2$. Since $(1-z)^2 \ge 1/4$, we obtain $s \le 64mz^2$. □

The coefficient in Lemma 3.2 may be the constant coefficient. The next argument uses nonvanishing and the degree bound, and does not require the corresponding subset S to be nonempty.

4. Linear cuts of missing sums

Return to $G = \mathbb{F}_2^n$. A vector $b \notin A + A$ has the property

$$x \in A \quad \Longrightarrow \quad x + b \notin A. \tag{9}$$

Indeed, otherwise $b = x + (x + b) \in A + A$.

Proposition 4.1 (The cut estimate)**.** *Suppose $0 \le \varepsilon \le 1/4$ and $\alpha \ge 1/2 - \varepsilon$. For every nonempty $D \subseteq G \setminus (A + A)$ there is a linear functional $\ell : G \to \mathbb{F}_2$ such that*

$$|D \cap \ker \ell| \le 256\varepsilon^2 |D|. \tag{10}$$

Proof. Choose $b_0 \in D$ and define

$$f(x) = \mathbf{1}_A(x) - \mathbf{1}_A(x + b_0).$$

The two indicator supports are disjoint by (9). Thus f is ternary and

$$z(f) = 1 - 2\alpha \le 2\varepsilon. \tag{11}$$

In particular, the existence of the missing sum b_0 implies $\alpha \le 1/2$.

For every $b \in D$ and $x \in G$,

$$f(x)f(x + b) \ne 1. \tag{12}$$

To see this, a product equal to one would mean that both values are 1 or both are -1. In the first case $x, x + b \in A$. In the second case $x + b_0, x + b + b_0 \in A$. Both contradict (9) for b.

Enumerate $D = \{b_1, \dots, b_m\}$, and define a linear map on the cube by

$$L(z) = \sum_{i=1}^{m} z_i b_i, \qquad z \in Q_m.$$

The b_i need not be linearly independent. For each $y \in G$, put

$$g_y(z) = (-1)^{z_1 + \cdots + z_m} f(y + L(z)). \tag{13}$$

These functions are ternary. Flipping coordinate i changes the parity factor's sign and translates the argument of f by b_i. Therefore

$$g_y(z)g_y(z + e_i) = -f(y + L(z))f(y + L(z) + b_i) \ne -1$$

by (12). Each g_y is a ternary separator.

For every fixed z, translation by $L(z)$ preserves the uniform measure on G. Averaging first over y and then over z yields

$$\mathbb{E}_y z(g_y) = z(f) \le 2\varepsilon.$$

Choose y with $z(g_y) \le 2\varepsilon \le 1/2$. Lemma 3.2 supplies a subset $S \subseteq [m]$ such that

$$\widehat{g_y}(S) \ne 0, \qquad |S| \le 64 m z(g_y)^2 \le 256 m\varepsilon^2. \tag{14}$$

It remains to interpret this coefficient. Use the standard dot product on G, and write its characters as $\chi_a(x) = (-1)^{a \cdot x}$ for $a \in G$. Fourier inversion for f in (13) gives

$$g_y(z) = \sum_{a \in G} \widehat{f}(a)\chi_a(y)\chi_{S(a)}(z), \qquad S(a) = \{i : a \cdot b_i = 0\}. \tag{15}$$

The formula for $S(a)$ follows because the parity factor adds one to every coordinate of the lifted frequency. In particular,

$$\widehat{g_y}(S) = \sum_{a : S(a) = S} \widehat{f}(a)\chi_a(y).$$

Since this is nonzero, some a has $S(a) = S$; no injectivity of the map $a \mapsto S(a)$ is needed. With $\ell(x) = a \cdot x$, equation (14) becomes (10). $\square$

5. Minimal obstructions and polynomial interpolation

The rest of the argument is a finite-dimensional covering statement.

Definition 5.1. A set $B \subseteq \mathbb{F}_2^n$ is *covered by r functionals* if there are linear functionals $\ell_1, \dots, \ell_r$ such that for every $x \in B$, some $\ell_i(x)$ equals one. Equivalently,

$$B \cap \bigcap_{i=1}^{r} \ker \ell_i = \varnothing.$$

Covering by zero functionals means $B = \varnothing$.

Let

$$M(n,r) = \sum_{j=0}^{\min(r,n)} \binom{n}{j}.$$

The $\mathbb{F}_2$-vector space of multilinear polynomial functions of degree at most r is spanned by the $M(n,r)$ squarefree monomials $\prod_{i \in S} x_i$, $|S| \le r$. Its dimension is therefore at most $M(n,r)$. Products of r affine-linear functions belong to this space: on $\mathbb{F}_2^n$ the relations $x_i^2 = x_i$ reduce every monomial to a squarefree monomial without increasing degree.

Lemma 5.2 (Interpolation bound). *If $B \subseteq \mathbb{F}_2^n$ is not covered by r functionals, there exists $D \subseteq B$ which is not covered by r functionals and satisfies*

$$|D| \le M(n,r). \tag{16}$$

Proof. Choose an inclusion-minimal subset $D \subseteq B$ which is not covered by r functionals. For every $b \in D$, the set $D \setminus \{b\}$ has a cover by functionals $\ell_{b,1}, \dots, \ell_{b,r}$. All these functionals vanish at b: if one did not, the same list would cover D.

Define the polynomial function

$$P_b(x) = \prod_{i=1}^{r} (1 + \ell_{b,i}(x)),$$

with arithmetic in $\mathbb{F}_2$. Then $P_b(b) = 1$, while $P_b(x) = 0$ for every $x \in D \setminus \{b\}$. Consequently the functions $(P_b)_{b \in D}$ are linearly independent: evaluating a linear relation at each b isolates its coefficient. Each P_b belongs to the space of degree-at-most-r multilinear polynomial functions, giving (16). The argument includes $r = 0$, with the empty product interpreted as one. $\square$

Lemma 5.3 (Counting monomials). *If $1 \le r \le n$, then*

$$M(n,r) \le \left(\frac{en}{r}\right)^r. \tag{17}$$

Proof. Set $t = r/n \in (0,1]$. Since $t^r \le t^j$ for $j \le r$,

$$M(n,r)t^r \le \sum_{j=0}^{n} \binom{n}{j} t^j = (1+t)^n \le e^{nt} = e^r.$$

Divide by t^r. $\square$

Proposition 5.4 (A cover from uniform cuts). *Let $B \subseteq \mathbb{F}_2^n$ with $0 \notin B$, and let $\delta \ge 0$. Suppose that every nonempty $D \subseteq B$ admits a linear functional ℓ with*

$$|D \cap \ker \ell| \le \delta |D|. \tag{18}$$

Then B is covered by c functionals for some nonnegative integer c with

$$c \le 1 + en\delta. \tag{19}$$

Proof. Every nonzero vector has a nonzero coordinate. The n coordinate functionals therefore cover B, so there is a least possible cover size $c \le n$. If $c = 0$ or $c = 1$, (19) is immediate. We may thus write $c = r + 1$ with $1 \le r < n$.

Since B is not covered by r functionals, Lemma 5.2 supplies a set $D \subseteq B$, also not covered by r functionals, with $|D| \le M(n,r)$. Starting with this D, apply (18) successively to the points remaining in the kernels of the previously chosen functionals. After r steps there are at most $\delta^r|D|$ remaining points. The remaining set cannot be empty, because that would give a cover of D by r functionals. Thus

$$1 \le \delta^r|D| \le \delta^r M(n,r) \le \left(\frac{en\delta}{r}\right)^r, \tag{20}$$

where the last inequality is Lemma 5.3. Taking the positive rth root gives $r \le en\delta$, proving the result. In the case $\delta = 0$, (20) is impossible, so the already handled cases $c \le 1$ are the only possibilities. □

The crucial hypothesis is that the cut estimate holds for every subset of B, including the minimal obstruction and all of its subsequent remainders. It is this hereditary property that permits the interpolation bound to replace the size of the full set of missing sums.

6. Completion of the sumset argument

Proof of Theorem 1.1. The density hypothesis implies $\alpha \ge 1/4$, so A is nonempty. Hence $0 \in A+A$, since $a + a = 0$ for any $a \in A$. Set $B = G \setminus (A + A)$; then $0 \notin B$.

Proposition 4.1 supplies (18) for every nonempty $D \subseteq B$ with $\delta = 256\varepsilon^2$. By Proposition 5.4, there are $c \le 1 + 256en\varepsilon^2$ functionals $\ell_1, \dots, \ell_c$ covering B. Their common kernel

$$H = \bigcap_{i=1}^{c} \ker \ell_i$$

is disjoint from B and therefore lies in $A+A$. It is the kernel of the linear map $x \mapsto (\ell_1(x), \dots, \ell_c(x))$ into $\mathbb{F}_2^c$. Rank–nullity gives $\operatorname{codim} H \le c$, completing the proof. If B is empty, the empty list of functionals gives $H = G$. This also handles $n = 0$. □

Proof of Corollary 1.2. Set $\varepsilon = k/\sqrt{n}$. The choice of $N(k)$ ensures $n > 16k^2 > 0$ and $0 < \varepsilon < 1/4$, while $n\varepsilon^2 = k^2$. Theorem 1.1 gives $H \subseteq A + A$ with $\operatorname{codim} H \le 1 + 256ek^2 \le c(k)$. □

6.1. The formulation using a guaranteed coset dimension.

To compare the conclusion with the formal target, define

$$d(S) = \max\{\dim V : x + V \subseteq S,\ V \le \mathbb{F}_2^n\}$$

for nonempty $S \subseteq \mathbb{F}_2^n$, and set $d(\varnothing) = 0$. For $\beta \le 1$, put

$$F(n,\beta) = \min_{\substack{A\subseteq \mathbb{F}_2^n \\ |A|/2^n \ge \beta}} d(A + A).$$

The family in this minimum is nonempty because $A = G$ is admissible; all dimensions lie in $\{0, \dots, n\}$. This is the finite extremal quantity represented by `Green51.guaranteedMaxCosetDim`.

Let $k > 0$, $n \ge N(k)$, and $1/2 - k/\sqrt{n} < \beta \le 1$. Corollary 1.2 applies to every admissible A, giving $n \le d(A + A) + c(k)$. Taking the minimum proves

$$n \le F(n,\beta) + c(k).$$

Thus the uniform statement proved here implies the exact quantified one-half target. The upper restriction $\beta \le 1$ matters when expressing the theorem through this minimum, because it ensures that the defining family is nonempty.

7. Lean proof correspondence and verification evidence

The formalization develops real Walsh transforms, face averages, the lifting argument, and multilinear polynomial functions in the namespace `Bounty.Green51Proof`. Conjectures' public verification report records successful compilation, acceptance by Lean's default kernel, comparison with `Green51.green_51.one_half`, and a permitted-axiom check. The final declaration is `Bounty.target`. The correspondence between the mathematical argument and the formal declarations is as follows.

Paper result	Formal declaration
Lemma 2.2	`cube_separator_variance`
Lemma 3.1	`noise_energy_bound`
Lemma 3.2	`exists_low_degree_separator_coefficient`
Proposition 4.1	`exists_good_cut`
Lemma 5.2	`critical_interpolation`
Lemma 5.3	`monomial_count_bound`
Proposition 5.4	`small_cover`
Theorem 1.1, construction	`commonKernel`
Theorem 1.1, dimension bound	`sumset_codimension_bound`
Corollary 1.2 and the target	`green_one_half`, `target`

The named theorem `sumset_codimension_bound` concludes a bound on the maximum affine-coset dimension; its proof explicitly constructs the common linear kernel used in Theorem 1.1. Accordingly, the linear subspace assertion here is supported by that construction, and is not being inferred merely from a statement about affine cosets.

7.1. Public artifacts and reproducibility.

The artifacts are linked from the public result record. The complete Lean source and machine-readable verification report are publicly accessible. The source is also available as the `source` field of the solution API response. The UTF-8 source consists of 74,582 bytes with SHA-256

`95efbc149a5b09ff8f34a83fa673b5bd63b9591c071d9cf7a495182706ce5db3`.

The public report identifies the full original report by SHA-256

`f2f20d91eabb88ad101e556135173d7981dd28122fbafbfd14bd6804a86337c8`.

This second digest commits to the full report, not to its public JSON projection.

The verification used Lean 4.33.1 and the revisions below. Linked identifiers resolve to the pinned files or repositories; the complete revisions and dependency fingerprints are in the dependency lock file.

Component	Revision
Mathlib	`0df444a360ea`
Formal Conjectures, patched	`8432eac99811`
Conjectures verifier	`428166381eba`
Task collection	`275ef4824c41`
Statement comparator	`68a064109f01`

To reproduce the check, clone the pinned verifier and task repositories into sibling directories named `conjectures-validator` and `conjectures-tasks`. The verifier's Docker build instructions reconstruct the patched Formal Conjectures revision from its upstream base and the hash-checked patch in the task repository. Save the solution API's `source` field as UTF-8 bytes in `submissions/Main.lean`, preserving line endings, and check the source digest above. Build with `docker compose build verifier`, then run `docker compose run --rm verifier doctor` and the following check from the verifier directory:

```
bundle=sha256:
bundle+=9e80ef6216da982ff582d20419ed0b0393952ef9fa1c197726f891246db4d327
FC_SUBMISSION_FILE=./submissions/Main.lean \
  docker compose run --rm verifier verify \
  --task /inputs/tasks/pool/tier-1/green-51-one-half-formalized \
  --submission /inputs/submissions/Main.lean \
  --expected-task-sha256 "$bundle"
```

These commands use Bash and the Linux Docker Compose profile. The expected outcome is exit status zero with `accepted: true` and reason `VERIFIED`; durations and diagnostic output may differ between runs. The verification permits `propext`, `Quot.sound`, and `Classical.choice`. The optional Nanoda checker was disabled. The recorded checks concern the Lean target and its dependencies; they do not mechanically verify the prose exposition.

7.2. **Quantitative and structural questions.** The constant $256e$ comes from two estimates: 64 in the least Fourier degree bound, multiplied by four when $z \le 2\varepsilon$, and e in the monomial count. These estimates have not been optimized. The $\Omega(k)$ obstruction from [2, Theorem 1.3] and the $O(k^2)$ bound in Corollary 1.2 leave open whether the guaranteed codimension can always be bounded linearly in k.

At $\varepsilon = 0$ the theorem yields codimension at most one. For general fixed $\varepsilon > 0$, its bound is linear in n and may be numerically uninformative; the near-half scale $\varepsilon = k/\sqrt{n}$ is the application emphasized here. The proof does not determine Green's full extremal function, nor does it address the arithmetic-progression version of the problem over the integers.

Acknowledgments

OpenAI Codex assisted with drafting, revising, and typesetting the exposition from the Lean proof and its verification records.

References

[1] B. Green, *100 Open Problems*, Problem 51, p. 25, online collection, version consulted 9 September 2026. https://people.maths.ox.ac.uk/greenbj/papers/open-problems.pdf.

[2] T. Sanders, *Green's sumset problem at density one half*, Acta Arithmetica **146** (2011), no. 1, 91–101. See Theorems 1.2 and 1.4 and Question 5.1. https://arxiv.org/abs/1003.5649.