The Harborth constant of $C_p \oplus C_p$ for every prime $p \geq 11$

P. Graham
https://orcid.org/0009-0007-8493-1505

November 4, 2026

Abstract

The Harborth constant $\mathsf{g}(G)$ of a finite abelian group G is the least k such that every k-element subset of G contains $\exp(G)$ distinct elements whose sum is zero; the exponent $\exp(G)$ is the largest order of an element of G. Gao and Thangadurai conjectured that $\mathsf{g}(C_n \oplus C_n) = 2n-1$ for odd n, and proved it for primes $p \geq 67$; Gao, Geroldinger and Schmid extended this to $p \geq 47$, and Kemnitz had settled $p \in \{3, 5, 7\}$. Bajnok's *Menu of research problems* asks for the primes $11 \leq p \leq 43$ (Problem F.100). We prove that $\mathsf{g}(C_p \oplus C_p) = 2p-1$ for every prime $p \geq 11$, which answers that problem and, with Kemnitz's values, gives the conjecture for every odd prime. As in the work of Gao and Thangadurai and of Gao, Geroldinger and Schmid, we project the set along a line through 0, which sorts its points into p parallel lines (fibers), choose a projection in which many pairs of points share a fiber, and use the lifting criterion of Gao and Thangadurai. What is new is a sufficient condition, proved with the theorem of Dias da Silva and Hamidoune, that depends only on how many points each fiber holds and not on where the fibers lie. The analysis of fiber sizes is by hand for $p \geq 17$ and a finite check for $p = 11$ and 13. The proof for $p \geq 11$ and the lower bound for every odd prime are formalized in Lean.

This work has not yet been reviewed by experts. The proofs, the Lean formalization and the text were produced by AI models (from OpenAI and Anthropic), working under the author's direction.

2020 Mathematics Subject Classification. 11B30, 11P70, 20K01.
Keywords. Zero-sum problems, Harborth constant, restricted sumsets, Dias da Silva–Hamidoune theorem.

1 Introduction

1.1 Main result

For a finite abelian group G, the exponent $\exp(G)$ is the largest order of an element of G, so $\exp(C_p \oplus C_p) = p$. The Harborth constant $\mathsf{g}(G)$ is the least integer k such that every subset of G with k elements has $\exp(G)$ distinct elements whose sum is 0. We identify $C_p \oplus C_p$ with the plane $\mathbb{F}_p^2 = \mathbb{Z}_p^2$, so that we can speak of lines through 0 (directions) and projections.

Theorem 1. *Let $p \geq 11$ be prime. Every subset of $C_p \oplus C_p$ with $2p-1$ elements contains p distinct elements whose sum is 0. Hence $\mathsf{g}(C_p \oplus C_p) = 2p - 1$.*

The lower bound $\mathsf{g}(C_p \oplus C_p) \geq 2p - 1$ holds for every odd prime p and is due to Kemnitz [11] (see [2, Proposition F.97]). In the words of the *Menu* [2, p. 301], "As Kemnitz observed in [127], the set $A = \{0, 1\}^{r-1} \times \{0, 1, \ldots, k-2\}$ [...] is zero-k-sum-free". We use its translate by $(0, 1)$ for $r = 2$: the $2p-2$ points (a, b) with $a \in \{0, 1\}$ and $b \neq 0$. If p of them sum to 0, the first coordinate

of the sum is the number m of chosen points with $a = 1$, so p divides m. Since $m \le p-1$, $m = 0$, but only $p-1$ points have $a = 0$. Theorem 1 and this lower bound are proved in Lean with Mathlib, from the definitions (Section 7).

1.2 Background and prior work

The constant goes back to Harborth's problem on lattice points with integral centroid [10]. Gao and Thangadurai [8] introduced the conjecture in its present form; in the words of Gao, Geroldinger and Schmid [6, Section 5], with our citation labels, "In [8] it is conjectured that $\mathsf{g}(C_n \oplus C_n)$ is equal to $2n-1$ for every odd $n \ge 3$ and equal to $2n+1$ for every even $n \ge 3$, and it is observed that these values are lower bounds. Moreover, it is proved (see [8, Theorem 1]) that $\mathsf{g}(C_p \oplus C_p) = 2p-1$ for all primes $p \ge 67$." Gao, Geroldinger and Schmid [6, Theorem 5.1] described all sets of size $2p-2$ with no zero-sum subset of size p for primes $p \ge 47$, and so proved $\mathsf{g}(C_p \oplus C_p) = 2p-1$ for those primes.

Bajnok's *Menu* [2] states the conjecture as Conjecture F.98, in his notation and as the equivalent statement that for odd k the largest subset of $\mathbb{Z}_k^2$ with no k distinct elements summing to 0 has $2k-2$ elements; his k is the n of the quotation above. It records that "Kemnitz in [127] proved that Conjecture F.98 holds for $k \in \{2, 3, 5, 7\}$" (see [11]), states the result for $p \ge 47$ as Theorem F.99, and poses

> Problem F.100. Prove that Conjecture F.98 holds for prime numbers $11 \le p \le 43$.

Guillot, Marchan, Ordaz, Schmid and Zerdoum [9] recorded the same state of knowledge in 2019: "It is known that $\mathsf{g}(C_p \oplus C_p) = 2p-1$ for prime $p \ge 47$ and for $p \in \{3, 5, 7\}$" [9, p. 614]. Theorem 1 settles Problem F.100, and with Kemnitz's values it gives $\mathsf{g}(C_p \oplus C_p) = 2p-1$ for every odd prime p. The composite case of the conjecture is untouched.

We build directly on three earlier ingredients. The *lifting criterion* (Lemma 2.1) is Lemma 2.5 of Gao and Thangadurai [8]. Gao, Hou, Schmid and Thangadurai extend it to C_p^r as Lemma 5.2 of [7], which "was proved in [19] in the special case $r = 2$", with [19] their label for [8], and Gao, Geroldinger and Schmid quote it as Lemma 5.3(1) of [6]. The *averaging over directions* (Lemma 2.2) is in Gao and Thangadurai [8, pp. 83–84] and is Lemma 5.13 of [6]. Closest to our criterion are Lemmas 5.7 and 5.9 of [6], stated with the projected sequence written as $R'R^3U^2V$ [6, p. 267]. Lemma 5.9 applies the theorem of Dias da Silva and Hamidoune (their Theorem 2.3(1)) to the support RUV and the Cauchy–Davenport theorem to squarefree pairs Q_i from R', and Lemma 5.7 bounds the lifting sum below by $2(v_g(\varphi(S)) - 2)$ for each g in the support of the repeated part of the zero-sum sequence. Lemma 5.9 assumes $p \ge 47$ and Lemma 5.10 $p \ge 41$; Lemmas 5.7, 5.8, 5.11 and 5.12 assume $p \ge 5$, 5, 17 and 23. Gao, Geroldinger and Schmid write [6, p. 272] that "with more involved proofs some of the results could be obtained for (slightly) smaller primes. Our aim is to obtain results (at least) for primes p with $p \ge 47$, which is the bound given by Lemma 5.9." The tools are the theorem of Dias da Silva and Hamidoune [5], in the form of [6, Theorem 2.3(1)], and the Cauchy–Davenport theorem [3, 4].

What this paper adds is the symmetric layer criterion of Section 3. It applies Dias da Silva–Hamidoune to every layer of a depth chosen inside the fibers (both defined in Section 3), and its hypotheses, a single quadratic inequality and a bound on the layer sizes, depend only on the fiber sizes, not on their positions. The analysis of these sizes in Sections 4–6 brings the threshold down to $p = 11$. The hard step was to find a sufficient condition that does not depend on where the fibers sit; once the quadratic form was found, the case analysis became short.

1.3 Proof overview

Let $S \subseteq \mathbb{F}_p^2$ with $|S| = 2p-1$. Fix a direction, that is, a line K through 0, and project S along K onto a complementary line $H \cong \mathbb{Z}_p$. Each $g \in H$ carries a fiber $S \cap (g+K)$ of v_g points, and the vector v of fiber sizes is the *profile*. By the lifting criterion, S has the required p points as soon as there are integers $0 \le t_g \le v_g$ with $\sum t_g = p$, $\sum t_g g = 0$ and enough freedom $\sum t_g(v_g - t_g) \ge p-1$ inside the fibers.

Two points of S share a fiber for exactly one of the $p+1$ directions, so averaging gives a direction in which at least $2p-4$ pairs of points of S share a fiber. Section 2 reduces Theorem 1 in this way to Proposition 2.3 about such profiles.

To prove the proposition we fix a depth r and choose t in the box $\{t : |2t_g - v_g| \le r_g\}$ around $v/2$. The layers $\{g : r_g \ge i\}$ of the box are sets of distinct residues, so Dias da Silva–Hamidoune bounds the sums of half of each layer, and Cauchy–Davenport bounds the sumset of these sets (Section 3). The conditions on r then involve only the multiset $\{v_g\}$, the *shape* of v. For $p \ge 17$ two explicit rules give r for all shapes but four (Section 4). For one of the four, $(p-2, p-2, 3)$, a rule still gives a witness, checked directly; the other three are handled through the positions of their fibers (Section 5). For $p = 11$ and 13 all 526 and 1435 concentrated shapes, respectively, are checked (Section 6).

2 Reduction to a lemma on profiles

Let p be an odd prime and $S \subseteq \mathbb{F}_p^2$ with $|S| = 2p-1$. Fix a line K through 0 and a complementary line H, and let φ be the projection onto H along K. For $g \in H$ the fiber $S \cap (g+K)$ has v_g elements, and $\sum_g v_g = 2p-1$. We call g a *class* and v_g its *size*.

Lemma 2.1 (Lifting criterion [8, 7])**.** *Suppose that $t\colon H \to \mathbb{Z}_{\ge 0}$ satisfies $t_g \le v_g$, $\sum_g t_g = p$, $\sum_g t_g\, g = 0$ in H, and*

$$W(t) := \sum_{g \in H} t_g(v_g - t_g) \;\ge\; p-1.$$

Then S contains p distinct elements whose sum is 0.

Proof. Write the fiber over g as $g + A_g$ with $A_g \subseteq K \cong \mathbb{Z}_p$ and $|A_g| = v_g$. Choosing t_g distinct points in each fiber gives p distinct points with sum $\sum_g t_g g + \sum_g a_g = \sum_g a_g$, where a_g runs over the set $t_g^{\wedge} A_g$ of sums of t_g distinct elements of A_g. By Dias da Silva–Hamidoune, $|t_g^{\wedge} A_g| \ge \min(p, t_g(v_g - t_g) + 1)$; this is trivial when $t_g \in \{0, v_g\}$. By Cauchy–Davenport the sumset $\sum_g t_g^{\wedge} A_g$ has at least $\min(p, W(t)+1) = p$ elements, so it contains 0. □

If some fiber is a whole line $g + K$, its p points sum to $pg + \sum_{k \in K} k = 0$ because p is odd. So we may assume $v_g \le p-1$ for all g.

Lemma 2.2 (Averaging [8, 6])**.** *Some direction K has $\sum_g \binom{v_g}{2} \ge 2p-4$.*

Proof. Two points of S lie in a common fiber for exactly one of the $p+1$ directions, so the sums $\sum_g \binom{v_g}{2}$ over all directions add up to $\binom{2p-1}{2}$. Some direction has at least $(2p-1)(p-1)/(p+1) = 2p - 5 + 6/(p+1) > 2p-5$. □

Since an affine map $g \mapsto \alpha g + \beta$ ($\alpha \ne 0$) of H preserves the conditions of Lemma 2.1 (it multiplies $\sum t_g g$ by α and adds $p\beta = 0$), we may identify H with $\mathbb{Z}_p$. Theorem 1 now follows from Lemmas 2.1 and 2.2 and the following proposition.

Proposition 2.3 (Concentrated profiles)**.** *Let $p \geq 11$ be prime and $v\colon \mathbb{Z}_p \to \{0,\dots,p-1\}$ with $\sum_g v_g = 2p-1$ and $\sum_g \binom{v_g}{2} \geq 2p-4$. Then some t satisfies the hypotheses of Lemma 2.1.*

We call such a v *concentrated*. Write n_j for the number of classes g with $v_g = j$. Then $\sum_j n_j = p$ and $\sum_j (j-1)n_j = p-1$. Since $\binom{j}{2} = (j-1) + \binom{j-1}{2}$ for $j \geq 1$ and $\binom{0}{2} = 0 = (0-1)+1$, we get $\sum_j \binom{j}{2} n_j = (p-1) + n_0 + \sum_{j\geq 3} \binom{j-1}{2} n_j$, so the concentration hypothesis reads

$$n_0 + \sum_{j\geq 3} \tbinom{j-1}{2} n_j \;\geq\; p-3. \tag{1}$$

3 The symmetric layer criterion

The *shape* of a profile v is the multiset $\{v_g\}$ of its sizes. For a profile v and a *depth* $r\colon \mathbb{Z}_p \to \mathbb{Z}_{\geq 0}$, let $x_i = |\{g : r_g \geq i\}|$ for $i \geq 1$, the sizes of the layers of r.

Lemma 3.1. *Let v be as in Proposition 2.3, and suppose that r satisfies*

(D) $r_g \leq v_g$ *and* $r_g \equiv v_g \pmod 2$ *for all g;*

(W) $\sum_g r_g^2 \leq \sum_g (v_g - 2)^2$*;*

(S) $\sum_{i\geq 1} \lfloor x_i^2/4 \rfloor \geq p-1$.

Then some t satisfies the hypotheses of Lemma 2.1.

Proof. Put $a_g = (v_g - r_g)/2$, an integer by (D), and consider t with $a_g \leq t_g \leq a_g + r_g$. Then $u_g = 2t_g - v_g$ satisfies $|u_g| \leq r_g$, and $t_g(v_g - t_g) = (v_g^2 - u_g^2)/4 \geq (v_g^2 - r_g^2)/4$. Since $\sum_g (v_g-2)^2 = \sum_g v_g^2 - 4(2p-1) + 4p$, condition (W) gives $\sum_g (v_g^2 - r_g^2) \geq 4(p-1)$, so $W(t) \geq p-1$ for every such t.

The layers $L_i = \{g : r_g \geq i\}$ are sets of distinct residues. For subsets $X_i \subseteq L_i$ of sizes h_i, the vector $t_g = a_g + |\{i : g \in X_i\}|$ lies in the box, with $\sum_g t_g = \sum_g a_g + \sum_i h_i$ and $\sum_g t_g g = \sum_g a_g g + \sum_i \sum_{x\in X_i} x$. Let $R = \sum_g r_g = \sum_i x_i$. By (D), $R \equiv \sum_g v_g$ is odd, and $\sum_g t_g = p$ means $\sum_i h_i = (R+1)/2$. Since $\sum_i \lfloor x_i/2 \rfloor \leq (R-1)/2$ and $\sum_i \lceil x_i/2 \rceil \geq (R+1)/2$, we can choose $h_i \in \{\lfloor x_i/2 \rfloor, \lceil x_i/2 \rceil\}$ with this sum; then $h_i(x_i - h_i) = \lfloor x_i^2/4 \rfloor$. By Dias da Silva–Hamidoune each $h_i^{\wedge} L_i$ has at least $\min(p, \lfloor x_i^2/4 \rfloor + 1)$ elements, and by Cauchy–Davenport and (S) their sumset is all of $\mathbb{Z}_p$. So some choice of the X_i gives $\sum_g t_g g = 0$. □

Conditions (D), (W) and (S) are unchanged when v and r are permuted together, so they depend only on the shape of v. A depth r meeting them, listed in the order of the shape, is a *witness* for the shape, and then every profile v' with that shape has some t satisfying the hypotheses of Lemma 2.1 for v'. For example, take $p = 11$ and the shape $(6,3,3,3,3,3)$, padded with five zeros. The depth $r = (2,3,1,1,1,1)$, in the order of the shape and 0 on the five empty classes, is a witness: (D) holds, $\sum r_g^2 = 17 \leq 16 + 5\cdot 1 + 5\cdot 4 = 41$, and the layers have sizes $6, 2, 1$, so $9 + 1 + 0 = 10 = p - 1$.

4 The Shape Lemma for $p \geq 17$

Let $N(i) = |\{g : v_g \geq i\}|$, so that $N(1) = p - n_0$ is the number of nonempty classes and $\sum_i N(i) = 2p-1$, and let $v_{(1)} \geq v_{(2)} \geq \cdots$ be the sizes in decreasing order. For $\rho \geq 1$ put $\mathrm{sp}(\rho) = \sum_{i=1}^{\rho} \lfloor N(i)^2/4 \rfloor$. Since $\lfloor N^2/4 \rfloor = \sum_{j=2}^{N} \lfloor j/2 \rfloor$,

$$\mathrm{sp}(\rho) = \sum_{j\geq 2} \lfloor j/2 \rfloor \min(v_{(j)}, \rho). \tag{2}$$

Rule B (basic: every $r_g \le 2$). For $j \ge 3$ let $e(j) = (j-1)(j-3)/4$ if j is odd and $e(j) = j(j-4)/4$ if j is even; this is $((j-2)^2 - r^2)/4$, the slack in (W) of a class of size j with the depth $r \in \{1, 2\}$ given below. Put $E = \sum_{j \ge 3} e(j) n_j$ and $k = \min(n_2, n_0 + E)$. The depth is

- $r_g = 0$ if $v_g = 0$;
- $r_g = 1$ if v_g is odd;
- $r_g = 2$ if $v_g \ge 4$ is even;
- $r_g = 2$ on k of the n_2 classes of size 2, and $r_g = 0$ on the others.

Comparing class by class, $\sum_g (v_g - 2)^2 - \sum_g r_g^2 = 4(n_0 + E - k) \ge 0$, so (W) holds. The layers have sizes $x_1 = N(1) - n_2 + k$, $x_2 = |\{g : v_g \ge 4 \text{ even}\}| + k$ and $x_i = 0$ for $i \ge 3$.

Rule H(ρ) (high: depths up to $\rho + 1$), for $\rho \ge 2$. Put $r_g = v_g$ if $v_g \le \rho + 1$, and otherwise $r_g \in \{\rho, \rho + 1\}$ with $r_g \equiv v_g$. Then $x_i = N(i)$ for $i \le \rho$, so (S) holds when $\mathrm{sp}(\rho) \ge p - 1$. Since $\sum_g (v_g - 2)^2 - \sum_g r_g^2 = \sum_g (v_g^2 - r_g^2) - 4(p-1)$, (W) holds when

$$\Delta(\rho) := \sum_{v_g \ge \rho + 2} \left(v_g^2 - (\rho + 1)^2\right) \;\ge\; 4(p-1).$$

Proposition 4.1 (Shape Lemma)**.** *Let $p \ge 17$ be prime and v concentrated, of none of the shapes $(p-1, p-1, 1)$, $(p-1, p-2, 2)$, $(p-1, p-3, 3)$, $(p-2, p-2, 3)$. Then rule B or rule H(ρ^*) gives a witness, where ρ^* is the least ρ with $\mathrm{sp}(\rho) \ge p - 1$.*

Proof. Case 1: $n_2 > n_0 + E$, so $k = n_0 + E$. Let $A = \sum_{j \ge 3} (j-1) n_j$; the identity $\sum_j (j-1) n_j = p - 1$ gives $n_2 = p - 1 + n_0 - A$, and hence $x_1 = 1 - n_0 + A + E$. Since $n_2 \le p - n_0 - N(3)$, we have $2n_0 \le A + 1 - N(3)$. This gives $x_1 \ge \frac{1}{2} + \sum_{j \ge 3} f(j) n_j$ with $f(j) = j/2 + e(j)$, and, with (1), $\sum_{j \ge 3} q(j) n_j \ge p - \frac{7}{2}$ with $q(j) = j(j-2)/2$. One checks that $f(j) = q(j)/2$ for even j and $f(j) = (j^2 - 2j + 3)/4 \ge q(j)/2$ for odd j. Hence $x_1 \ge p/2 - 5/4$, and as x_1 is an integer and p is odd, $x_1 \ge (p-1)/2$. Then $\lfloor x_1^2/4 \rfloor \ge \lfloor (p-1)^2/16 \rfloor \ge p - 1$ because $p \ge 17$, so (S) holds for rule B.

Case 2: $n_2 \le n_0 + E$ and $\lfloor N(1)^2/4 \rfloor \ge p - 1$. Rule B has $k = n_2$ and $x_1 = N(1)$, so (S) holds.

Case 3: $\lfloor N(1)^2/4 \rfloor \le p - 2$. This case does not use (1). Now $\mathrm{sp}(1) = \lfloor N(1)^2/4 \rfloor < p - 1$, so $\rho = \rho^* \ge 2$, and by (2) $\mathrm{sp}(v_{(2)}) \ge \sum_{j \ge 2} v_{(j)} = 2p - 1 - v_{(1)} \ge p$, so $\rho \le v_{(2)}$. It remains to show $\Delta(\rho) \ge 4(p-1)$, using $\mathrm{sp}(\rho - 1) \le p - 2$.

(3a) $\rho \le (p-1)/2$. Put $y = \rho - 1 \in [1, (p-3)/2]$. From $\lfloor N^2/4 \rfloor \ge (N^2 - 1)/4$ we get $\sum_{i \le y} N(i)^2 \le X := 4p - 8 + y$, so by Cauchy–Schwarz and the monotonicity of N, $L := \sum_{i \le \rho + 1} N(i) \le \sqrt{yX} + 2\sqrt{X/y}$. Each class with $v_g \ge \rho + 2$ contributes $v_g^2 - (\rho+1)^2 \ge (2\rho + 3)(v_g - \rho - 1)$ to $\Delta(\rho)$, and $\sum_g \max(v_g - \rho - 1, 0) = \sum_{i \ge \rho + 2} N(i) = 2p - 1 - L$. Therefore $\Delta(\rho) \ge (2y + 5)(2p - 1 - L)$, and the inequality

$$F(p, y) := (2y + 5)\left(2p - 1 - \sqrt{yX} - 2\sqrt{X/y}\right) - 4(p-1) \;\ge\; 0$$

for $p \ge 17$ and $1 \le y \le (p-3)/2$ is verified in Appendix A.

(3b) $\rho \ge (p+1)/2$. By (2), $\mathrm{sp}(\rho - 1) = (\rho - 1) + \sum_{j \ge 3} \lfloor j/2 \rfloor \min(v_{(j)}, \rho - 1) \le p - 2$, so $v_{(j)} \le \rho - 2$ for all $j \ge 3$ (otherwise $\mathrm{sp}(\rho - 1) \ge 2\rho - 2 \ge p - 1$). With $\sigma = \sum_{j \ge 3} \lfloor j/2 \rfloor v_{(j)}$ this gives $\mathrm{sp}(\rho - 1) = \rho - 1 + \sigma$ and $\mathrm{sp}(\rho) = \rho + \sigma$, so $\rho = p - 1 - \sigma$ and $\sigma \le (p-3)/2$. Only the two largest classes exceed $\rho - 2$, and Appendix A shows $\Delta(\rho) \ge (\sigma - 1)(p + 3)$, which is at least $4(p-1)$ when $\sigma \ge 5$. For $\sigma \le 4$ the shapes are listed there: $\Delta(\rho) \ge 4(p-1)$ except for the four shapes excluded in the statement. □

5 The exceptional shapes

Lemma 5.1. *Proposition 2.3 holds for the shapes* $(p-2,p-2,3)$, $(p-1,p-1,1)$, $(p-1,p-2,2)$ *and* $(p-1,p-3,3)$, *for every prime* $p \geq 7$.

Proof. For $(p-2,p-2,3)$ the depth $r=(p-4,p-4,3)$ of rule H$(p-4)$, with $r_g=0$ on the $p-3$ empty classes, is a witness: $2(p-4)^2+9 \leq 2(p-4)^2+1+4(p-3)$, and the layers have sizes $3,3,3$ and then 2 (for $4 \leq i \leq p-4$), so $3 \cdot 2+(p-7)=p-1$. Only the sufficient test $\Delta(\rho) \geq 4(p-1)$ of Section 4 misses this shape.

The other three shapes are $(p-1,b,c)$ with $b+c=p$ and $c \in \{1,2,3\}$; here the classes' positions matter. By an affine map, put the class A of size $p-1$ at 0, the class B of size b at 1, and the class C of size c at some $z \notin \{0,1\}$. Let $\beta \in \{1,\dots,p-2\}$ be the residue of $-z$, and take $t_C=1$, $t_B=\beta$, $t_A=p-1-\beta$. Then $\sum t=p$, $\sum t_g g=\beta+z=0$, and

$$W(t)=\beta(p-1-\beta)+\beta(b-\beta)+(c-1),$$

with $\beta(p-1-\beta) \geq p-2$. For $b=p-1$, $W \geq 2(p-2) \geq p-1$; for $b=p-2$, $W \geq (p-2)+0+1=p-1$; for $b=p-3$ and $\beta \leq p-3$, $W \geq (p-2)+0+2 > p-1$. If $b=p-3$ and $\beta=p-2$, that is $z=2$, take instead $t_C=2$, $t_B=p-4$, $t_A=2$: then $\sum t_g g=p-4+4=0$ and $W=2(p-3)+(p-4)+2=3p-8 \geq p-1$. □

Proposition 4.1 and Lemmas 3.1 and 5.1 prove Proposition 2.3 for every prime $p \geq 17$.

6 The primes 11 and 13

For $p=11$ and $p=13$ the shapes are finite in number: there are 526 and 1435 concentrated shapes with all sizes at most $p-1$. For each, one checks that it is one of the three shapes $(p-1,b,c)$ of Lemma 5.1, or that rule B, some rule H(ρ) with $2 \leq \rho \leq p-1$, or a depth from Table 1 is a witness. At $p=13$ rules B and H suffice. At $p=11$ three shapes need the depths of Table 1, which can be checked by hand as in the example of Section 3.

shape	witness r	layer sizes
$(6,3,3,3,3,3)$	$(2,3,1,1,1,1)$	$6,2,1$
$(5,5,3,3,3,2)$	$(3,3,3,1,1,0)$	$5,3,3$
$(5,4,3,3,3,3)$	$(3,2,1,1,1,1)$	$6,2,1$

Table 1: The concentrated shapes at $p=11$ that rules B and H miss, with a witness (zeros omitted). In each, (W) reads $17 \leq 41$, $29 \leq 41$ and $17 \leq 37$.

With Lemma 3.1 this proves Proposition 2.3, and hence Theorem 1, for $p=11$ and 13. In Lean the check runs as a kernel computation over all sorted profiles (Section 7). A separate program, written from the statement of the criterion and sharing no code with the formal proof, confirms the counts and the coverage; it searches all depths when the rules fail and finds no uncovered shape. The criterion alone does not reach $p=7$: there the shapes $(5,4,4)$, $(4,3,3,3)$ and $(4,2,2,2,2,1)$ have no witness.

7 Formal verification

Theorem 1 and the lower bound are proved in Lean 4 with Mathlib. In the namespace `Harborth`, the theorems are `upper`, `lower` and `harborth_constant`:

```
def HasZeroSumPSubset (p : ℕ) (S : Finset (ZMod p × ZMod p)) : Prop :=
  ∃ T ⊆ S, T.card = p ∧ ∑ x ∈ T, x = 0

theorem upper (p : ℕ) [Fact p.Prime] (hp : 11 ≤ p)
    (S : Finset (ZMod p × ZMod p)) (hS : S.card = 2 * p - 1) :
    HasZeroSumPSubset p S

theorem lower (p : ℕ) [Fact p.Prime] (hp : 3 ≤ p) :
    ∃ S : Finset (ZMod p × ZMod p), S.card = 2 * p - 2 ∧ ¬ HasZeroSumPSubset p S

theorem harborth_constant (p : ℕ) [Fact p.Prime] (hp : 11 ≤ p) :
    sInf {k : ℕ | ∀ S : Finset (ZMod p × ZMod p), S.card = k → HasZeroSumPSubset p S}
      = 2 * p - 1
```

They match the theorem as follows. `ZMod p × ZMod p` is $C_p \oplus C_p$, and a `Finset` has distinct elements, so `HasZeroSumPSubset p S` says that S has $p = \exp(C_p \oplus C_p)$ distinct elements summing to 0. `upper` is the first sentence of Theorem 1 and `lower` is the lower bound. The set in `harborth_constant` contains $2p-1$ by `upper` and no smaller k by `lower` and passing to subsets, so its least element, `sInf`, is $\mathsf{g}(C_p \oplus C_p)$. Lean's natural-number subtraction truncates at 0; here `2 * p - 1` and `2 * p - 2` have their usual values because $p \geq 3$.

The formalization follows Sections 2–6. It proves the theorem of Dias da Silva and Hamidoune by the polynomial method of Alon, Nathanson and Ruzsa [1], from Mathlib's Combinatorial Nullstellensatz, and takes Cauchy–Davenport from Mathlib. Case 3 of Proposition 4.1 is proved with integer inequalities in place of the square roots of Appendix A. The checks of Section 6 run by `decide +kernel`, evaluation in Lean's kernel, which adds no axiom. The build restates the three theorems with the definition unfolded, checks small examples of the definition (a full line of $C_3 \oplus C_3$, and the lower-bound set at $p = 3$), and checks that each theorem depends only on `propext`, `Classical.choice` and `Quot.sound`. Kemnitz's cases $p \in \{3, 5, 7\}$ are not part of the formal proof. The Lean files accompany this paper; Appendix B says how to check them.

Appendix A The inequalities of Case 3

Case (3a). We show $F(p, y) \geq 0$ for primes $p \geq 17$ and $1 \leq y \leq (p-3)/2$. Note that $X/y = (4p-8)/y + 1$ decreases in y, and that in each range below the factor $2p - 1 - \sqrt{yX} - 2\sqrt{X/y}$ is positive, so $2y + 5$ may be replaced by a lower bound.

- $y = 1$: $X = 4p - 7$ and $F = 10p - 3 - 21\sqrt{4p-7}$, which is nonnegative if and only if $100p^2 - 1824p + 3096 \geq 0$, true for $p \geq 17$.
- $p \geq 23$, $2 \leq y \leq p/8$: $\sqrt{yX} \leq (\sqrt{33}/8)p$ and $2\sqrt{X/y} \leq 2\sqrt{2p-3}$, so with $2y + 5 \geq 9$ it suffices that $7.53p - 5 \geq 18\sqrt{2p-3}$, true for $p \geq 12$.
- $p \geq 23$, $p/8 \leq y \leq p/4$: $2y+5 \geq p/4+5$, $\sqrt{yX} \leq (\sqrt{17}/4)p$ and $2\sqrt{X/y} \leq 2\sqrt{33}$; the resulting quadratic in p has positive leading coefficient and largest root below 21.
- $p \geq 23$, $p/4 \leq y \leq (p-3)/2$: $2y+5 \geq p/2+5$, $\sqrt{yX} \leq 3(p-2)/2$ and $2\sqrt{X/y} \leq 2\sqrt{17 - 32/p}$, so it suffices that $(p/2+5)(p/2+2-2\sqrt{17-32/p}) \geq 4(p-1)$. With $\sqrt{17}$ in place of $\sqrt{17-32/p}$ this holds for $p \geq 24$; at $p = 23$ the left side is $92.3\ldots \geq 88$.
- $p = 17$ and $p = 19$: direct evaluation gives $F(17, y) = 3.0,\ 32.6,\ 47.0,\ 53.0,\ 52.4,\ 45.9,\ 34.0$ for $y = 1, \ldots, 7$ and $F(19, y) = 12.6,\ 48.0,\ 67.4,\ 78.1,\ 81.8,\ 79.4,\ 71.3,\ 57.7$ for $y = 1, \ldots, 8$ (to one decimal).

Case (3b). Let $s = \sum_{j\ge 3} v_{(j)} \le \sigma$. As $v_{(1)}, v_{(2)} \le p-1$, $s \ge 1$. Write $v_{(1)} = p-1-d$ and $v_{(2)} = p-s+d$ with $0 \le 2d \le s-1$. Since $\rho+1 = p-\sigma$ and $s \le \sigma$, both sizes are at least $\rho+1$, the smaller sizes are at most $\rho-2$, and

$$\Delta(\rho) = (\sigma-1-d)(2p-1-d-\sigma) + (\sigma-s+d)(2p-s+d-\sigma).$$

Both first factors are nonnegative, with sum $2\sigma-1-s$, and both second factors are at least $2p-2\sigma \ge p+3$, so $\Delta(\rho) \ge (2\sigma-1-s)(p+3) \ge (\sigma-1)(p+3)$. If $\sigma \le 4$, the tail $(v_{(3)}, v_{(4)}, \dots)$ is one of (1), (2), (3), (4), $(1,1)$, $(2,1)$, and the formula gives:

tail	d	shape	$\Delta(\rho)$
(1)	0	$(p-1, p-1, 1)$	0
(2)	0	$(p-1, p-2, 2)$	$2p-3$
(3)	0	$(p-1, p-3, 3)$	$4p-8$
(3)	1	$(p-2, p-2, 3)$	$4p-10$
$(1,1)$	0	$(p-1, p-2, 1, 1)$	$6p-13$
$(2,1)$	$0,1$	$(p-1, p-3, 2, 1)$, $(p-2, p-2, 2, 1)$	$8p-22$, $8p-24$
(4)	$0,1$	$(p-1, p-4, 4)$, $(p-2, p-3, 4)$	$6p-15$, $6p-19$

Apart from the first four rows, all values are at least $4(p-1)$ for $p \ge 17$.

Appendix B Reproducing the checks

The Lean project pins Lean 4 v4.34.0 and a fixed Mathlib revision. With the Lean toolchain manager `elan` installed, running `lake exe cache get` and then `lake build` in the project's `lean` folder downloads Mathlib's compiled library and checks the proof, the restated theorems, the examples and the axioms. The build takes a few minutes and about 9 GB of memory, most of it in the kernel checks of Section 6. The independent program of Section 6 needs only Python 3 and runs in under a second.

Acknowledgments

The problem was posed in this form by Weidong Gao and Ravindranathan Thangadurai, whose lifting lemma and averaging argument the proof uses, and recorded as Problem F.100 by Béla Bajnok. The layer approach and the result for $p \ge 47$ are due to Weidong Gao, Alfred Geroldinger and Wolfgang Schmid, and the extension of the lifting criterion to C_p^r to Weidong Gao, Qing-Hu Hou, Wolfgang Schmid and Ravindranathan Thangadurai. Arnfried Kemnitz proved the lower bound and settled $p = 3$, 5 and 7.

References

[1] N. Alon, M. B. Nathanson and I. Ruzsa, *The polynomial method and restricted sums of congruence classes*, J. Number Theory 56 (1996), 404–417, doi:10.1006/jnth.1996.0029.

[2] B. Bajnok, *Additive Combinatorics: A Menu of Research Problems*, arXiv:1705.07444v2 (2017); CRC Press (2018). Proposition F.97, Conjecture F.98, Theorem F.99, Problem F.100 and the passages quoted here are from arXiv version 2.

[3] A.-L. Cauchy, *Recherches sur les nombres*, J. École Polytech. 9 (1813), 99–116.

[4] H. Davenport, *On the addition of residue classes*, J. London Math. Soc. 10 (1935), 30–32, doi:10.1112/jlms/s1-10.37.30.

[5] J. A. Dias da Silva and Y. O. Hamidoune, *Cyclic spaces for Grassmann derivatives and additive theory*, Bull. London Math. Soc. 26 (1994), 140–146, doi:10.1112/blms/26.2.140.

[6] W. Gao, A. Geroldinger and W. A. Schmid, *Inverse zero-sum problems*, Acta Arith. 128 (2007), 245–279, doi:10.4064/aa128-3-5.

[7] W. D. Gao, Q. H. Hou, W. A. Schmid and R. Thangadurai, *On short zero-sum subsequences II*, Integers 7 (2007), #A21, math.colgate.edu/~integers/h21/h21.pdf.

[8] W. Gao and R. Thangadurai, *A variant of Kemnitz conjecture*, J. Combin. Theory Ser. A 107 (2004), 69–86, doi:10.1016/j.jcta.2004.03.009.

[9] P. Guillot, L. E. Marchan, O. Ordaz, W. A. Schmid and H. Zerdoum, *On the Harborth constant of* $C_3 \oplus C_{3p}$, J. Théor. Nombres Bordeaux 31 (2019), 613–633, doi:10.5802/jtnb.1097.

[10] H. Harborth, *Ein Extremalproblem für Gitterpunkte*, J. Reine Angew. Math. 262/263 (1973), 356–360.

[11] A. Kemnitz, *On a lattice point problem*, Ars Combin. 16B (1983), 151–160.