

The lowest-degree polynomial with nonnegative coefficients divisible by the n -th cyclotomic polynomial

John P. Steinberger*

Institute for Theoretical Computer Science
Tsinghua University
jpsteinb@gmail.com

Submitted: Oct 7, 2011; Accepted: Sep 23, 2012; Published: XX

Mathematics Subject Classifications: 11R09, 90C57, 90C05, 05E99

Abstract

We pose the question of determining the lowest-degree polynomial with nonnegative coefficients divisible by the n -th cyclotomic polynomial $\Phi_n(x)$. We show this polynomial is $1 + x^{n/p} + \dots + x^{(p-1)n/p}$ where p is the smallest prime dividing n whenever $2/p > 1/q_1 + \dots + 1/q_k$, where $q_1, \dots, q_k$ are the other (distinct) primes besides p dividing n . Determining the lowest-degree polynomial with nonnegative coefficients divisible by $\Phi_n(x)$ remains open in the general case, though we conjecture the existence of values of n for which this degree is, in fact, less than $(p-1)n/p$.

1 Introduction

The n -th cyclotomic polynomial $\Phi_n(x)$ is $\prod(x - \zeta)$ where the product is taken over the distinct primitive n -th roots of unity ζ . It is well-known that $\Phi_n(x)$ has integer coefficients and that $\Phi_n(x)$ is irreducible over $\mathbb{Q}$.

In this paper we consider the problem of determining the polynomial of lowest degree with nonnegative coefficients divisible by $\Phi_n(x)$. The problem doesn't make sense for $n = 1$ since $\Phi_1(x) = x - 1$ and $x = 1$ cannot be the root of any polynomial with nonnegative coefficients (by "polynomial" we always mean "nonzero polynomial"). But when $n > 1$ then $\zeta_n = e^{2\pi i/n}$ is a root of

$$1 + x + x^2 + \dots + x^{n-1}$$

*Supported by the National Basic Research Program of China Grant 2011CBA00300, 2011CBA00301, the National Natural Science Foundation of China Grant 61033001, 61061130540, 61073174, and by NSF grant 0994380.

so $\Phi_n(x)$ divides a polynomial with nonnegative coefficients. The question is to determine the lowest degree such polynomial. It does not matter whether we consider only polynomials in $\mathbb{R}[x]$, $\mathbb{Q}[x]$ or even $\mathbb{Z}[x]$ since the problem can be reduced to the feasibility of a rational system of linear inequalities, and since rational systems have solutions over $\mathbb{Q}$ if and only if they have solutions over $\mathbb{R}$ (and there is obviously no difference between $\mathbb{Q}$ and $\mathbb{Z}$). In fact every solution in $\mathbb{R}[x]$ is a convex combination of solutions in $\mathbb{Q}[x]$, by the same considerations. A priori, the solution may not be unique up to scalar multiplication.

Working over $\mathbb{Z}[x]$ suggests a physical analogy. Assume n equally spaced holes are drilled around the circumference of a circular plate and that stackable pegs are made to fit the holes. Then putting a stack of a_i pegs in the i -th hole, $i = 0 \dots n-1$, yields a plate that is perfectly balanced around the center if and only if

$$a_0 + a_1\zeta_n + \dots + a_{n-1}\zeta_n^{n-1} = 0 \quad (1)$$

which happens if and only if $\Phi_n(x) \mid a_0 + a_1x + \dots + a_{n-1}x^{n-1}$. Therefore balanced configurations of pegs are in one-to-one correspondence with polynomials with nonnegative integer coefficients divisible by $\Phi_n(x)$. The degree of a polynomial associated with a balanced configuration of pegs is equal to the largest index of a nonempty hole. Since the balance of the plate is not affected by rotation it is clearly advantageous to rotate the plate until the longest sequence of consecutive empty holes occurs between hole 0 (that contains a peg) and the peg with largest index. Thus finding the lowest degree polynomial with nonnegative coefficients divisible by $\Phi_n(x)$ is equivalent to finding a (nonempty) balanced configuration of pegs with the longest possible consecutive sequence of empty holes, for a plate with n equally spaced holes.

(We note that for the above analogy to apply integral numbers of pegs must be used; otherwise if we are allowed to saw off our pegs such as to place some arbitrary height of pegs in a given hole we can always balance the plate with just 3 nonempty holes, say hole 0, $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$. The polynomial corresponding to this configuration will indeed have ζ_n as a root but will not be divisible by $\Phi_n(x)$, due to the fact that $\Phi_n(x)$ is irreducible over $\mathbb{Q}[x]$ but not over $\mathbb{R}[x]$.)

Balanced configurations can be quite intricate [10] but some are simple. In particular for every prime p dividing n one can make a regular p -gon, which is balanced. To maximize the gaps between consecutive pegs one should take p to be the smallest prime dividing n . In this case gaps have $n/p - 1$ holes, and the associated polynomial (after rotation to place a peg in hole 0) is $1 + x^{n/p} + \dots + x^{(p-1)n/p}$. This polynomial seems like a “natural enough” candidate for the lowest degree polynomial with nonnegative coefficients divisible by $\Phi_n(x)$, which leads to the following conjecture.

Conjecture 1. *The lowest degree monic polynomial with nonnegative coefficients divisible by $\Phi_n(x)$, $n > 1$, is $1 + x^{n/p} + \dots + x^{(p-1)n/p}$ where p is the smallest prime dividing n .*

In particular, Conjecture 1 would imply that the solution is unique up to scalar multiplication.

In fact, we do *not* believe Conjecture 1 is true, even if we have not disproved it. Our skepticism stems from the fact that Conjecture 1 implies the feasibility of a certain system

of rational linear inequalities that, when n is a product of large number of distinct large primes, seems very unlikely to be feasible. Nonetheless, due to the size of the systems involved, finding a counterexample is quite challenging from a computational standpoint, and our only “hard” results are positive results in favor the conjecture.

Since the polynomial $1 + x^{n/p} + \dots + x^{(p-1)n/p}$ has smaller degree when p is small, it seems, intuitively, that it should be easier to prove Conjecture 1 when p is small. This is reflected in our main result.

Theorem 1. *Conjecture 1 holds when n is even or when n is a prime power or when $2/p > 1/q_1 + \dots + 1/q_k$ where $q_1, \dots, q_k$ are the other primes besides p dividing n .*

The smallest value of n not covered by Theorem 1 is $n = 11 \cdot 13 \cdot 17 \cdot 19 = 46189$ and the next few values are $n = 96577, 215441, 392863, 508079, \dots$. Checking Conjecture 1 for these (specific) values of n is, in theory, a finite problem each time (expressible as a linear program), but we found the computational scale of these problems too large to handle conveniently (our measure of convenience being a few days computation on a laptop).

The case of Theorem 1 when $n = p^\alpha$ is a prime power follows from the fact that

$$\begin{aligned}\Phi_{p^\alpha}(x) &= 1 + x^{p^{\alpha-1}} + x^{2p^{\alpha-1}} + \dots + x^{(p-1)p^{\alpha-1}} \\ &= 1 + x^{n/p} + \dots + x^{(p-1)n/p}\end{aligned}$$

and $\Phi_n(x)$ cannot divide a polynomial of lower degree than itself or divide a non-scalar-multiple of same degree (when n is divisible by two or more primes $\Phi_n(x)$ no longer contains only nonnegative terms). The case when n is even follows from elementary geometrical considerations: if $p = 2$ then $(p-1)n/p$ is half way around the plate, and one cannot balance the plate using pegs strictly contained in half of the plate; moreover if a balanced configuration uses holes $0 \dots n/2$ it must not place any pegs in holes $1 \dots n/2 - 1$, and it must place the same number of pegs in holes 0 and $n/2$, hence uniqueness.

Theorem 1 implies Conjecture 1 for all cases when n has three or fewer primes in its factorization since $2/p > 1/q_1$ and $2/p > 1/q_1 + 1/q_2$ by virtue of the fact that p is the smallest prime dividing n . The case when n is divisible by exactly two primes p, q already follows from a result of deBruijn [1], who showed that any polynomial with nonnegative coefficients divisible $\Phi_n(x)$ can then be written as $f(x)(1 + x^{n/p} + \dots + x^{(p-1)n/p}) + g(x)(1 + x^{n/q} + \dots + x^{(q-1)n/q})$ for some polynomials $f(x), g(x)$ with nonnegative coefficients.

Because $\Phi_{np}(x) = \Phi_n(x^p)$ when p is a prime dividing n , determining the lowest degree polynomial with nonnegative coefficients divisible by $\Phi_n(x)$ is equivalent to determining the lowest degree polynomial with nonnegative coefficients divisible by $\Phi_{n'}(x)$ where n' is the squarefree part of n (indeed the above identity implies that if $f(x)$ is any polynomial divisible by $\Phi_{n'}(x)$ then $f(x^{n/n'})$ is divisible by $\Phi_n(x)$, and that any polynomial divisible by $\Phi_n(x)$ remains divisible by $\Phi_n(x)$ after removing all terms whose exponents are not $0 \bmod n/n'$ (or not $z \bmod n/n'$ for any z)). This is why neither Conjecture 1 nor Theorem 1 take into account the multiplicity with which primes divide n .

We note that Conjecture 1 would have the following type of non-intuitive consequence: if n , say, is the product of all primes between 10^6 and 10^{100} , then a nonempty balanced

configuration of pegs could still not leave empty any section of size $1/10^6$ -th the length of the total circumference of the plate. One can extend this idea to plates with infinitely many holes: fix a prime p , and assume that we can place pegs at any position along the circumference of the plate whose polar angle is rational multiple of 2π where the denominator of the rational coefficient is not divisible by any prime smaller than p (this amounts to considering vanishing sums using roots of unity whose orders are not divisible by any prime smaller than p); then if Conjecture 1 is true, no nonempty balanced configuration can leave empty a section of length $1/p$ of the total circumference. This follows because any balanced configuration is finite, and one can take n to be the lcm of the orders of the roots of unity used.

While the question of “debunking” Conjecture 1 is of great interest, the remainder of the paper mostly focuses on our actual result, Theorem 1 and its proof (we do briefly re-explain our pessimism at the end of Section 3). The methods used for the proof of Theorem 1 are quite geometric and of potential independent interest.

2 Methods

For the rest of the paper n denotes a squarefree integer > 1 . It will be convenient to write the prime factorization of n as $pq_1 \cdots q_k$ where p is the smallest prime dividing n and $q_1, \dots, q_k$ are the remaining primes dividing n . We may assume $k \geq 1$ since Conjecture 1 is true for $n = p$. We put $P = n/p$, $Q_1 = n/q_1, \dots, Q_k = n/q_k$.

Let $V_n \subseteq \mathbb{R}^n$ be the vector space consisting of all n -tuples $(a_0, \dots, a_{n-1})$ such that $\Phi_n(x) | a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$. We index vectors starting at 0 rather than at 1 (i.e., write $(x_0, \dots, x_{n-1})$ rather than $(x_1, \dots, x_n)$). For convenience we partition the set of indices $\{0, 1, \dots, n-1\}$ into the following sets $\mathcal{F}_0, \dots, \mathcal{F}_{p-1}$:

$$\begin{aligned} \mathcal{H} &= \{0, P, 2P, \dots, n-P\} \\ \mathcal{F}_0 &= \{1, 2, \dots, P-1\} \\ \mathcal{F}_1 &= \{P+1, P+2, \dots, 2P-1\} \\ &\vdots \\ \mathcal{F}_{p-1} &= \{(p-1)P+1, (p-1)P+2, \dots, n-1\}. \end{aligned}$$

We also let $\mathcal{F}^+ = \mathcal{F}_0 \cup \cdots \cup \mathcal{F}_{p-2}$ and $\mathcal{F}^- = \mathcal{F}_{p-1}$. For a set of indices $\mathcal{A} \subseteq \{0, 1, \dots, n-1\}$ and a vector $v = (v_0, \dots, v_{n-1}) \in \mathbb{R}^n$, we say v is 0 on $\mathcal{A}$ (resp. is positive on $\mathcal{A}$, etc) if $v_i = 0$ for $i \in \mathcal{A}$ (resp. $v_i > 0$ for $i \in \mathcal{A}$, etc).

Let $1_{\mathcal{H}} \in \mathbb{R}^n$ be the incidence vector of $\mathcal{H}$; this is also the coefficient vector the polynomial

$$1 + x^{n/p} + \cdots + x^{(p-1)n/p} = 1 + x^P + \cdots + x^{(p-1)P}.$$

Thus $1_{\mathcal{H}} \in V_n$. Conjecture 1 is equivalent to the statement that the only nonnegative vectors in V_n that are 0 on $\mathcal{F}^- = \mathcal{F}_{p-1}$ are scalar multiples of $1_{\mathcal{H}}$.

Let $\langle u, v \rangle$ denote the usual dot product of vectors in $\mathbb{R}^n$ and let $V_n^\perp = \{u \in \mathbb{R}^n : \langle u, v \rangle = 0 \text{ for all } v \in V_n\}$ be the orthogonal complement of V_n in $\mathbb{R}^n$. Our results are based on the following lemma.

Lemma 1. *Conjecture 1 holds for (a squarefree) n if and only if there exists a vector $w \in V_n^\perp$ such that w is positive on $\mathcal{F}^+$ and zero on $\mathcal{H}$.*

The proof of Lemma 1 requires the following elementary proposition, which is a consequence of the Farkas lemma for systems of linear inequalities. We recall that the Farkas lemma (in one of its many incarnations) states that a linear system $Ax = b$, $x \geq 0$ has no solution if and only if there exists a y such that $y^T A \geq 0$ and $y^T b < 0$. (For background, see e.g. [9]; by writing $x \geq 0$ we mean that each coordinate of x is nonnegative.)

Proposition 1. *Let $A \in \mathbb{R}^{m \times n}$ be a real matrix and let $\mathcal{J} \subseteq \{0, \dots, n-1\}$. Then there is no $x = (x_0, \dots, x_{n-1}) \in \mathbb{R}^n$ such that $x \geq 0$, $Ax = 0$ and such that $x_j > 0$ for at least one $j \in \mathcal{J}$ if and only if there exists a vector $y \in \mathbb{R}^m$ such that $y^T A \geq 0$ and such that $(y^T A)_j > 0$ for all $j \in \mathcal{J}$.*

Proof. If y exists then x cannot exist, since otherwise $0 > (y^T A)x = y^T(Ax) = 0$.

For the other direction, assume x does not exist. Then for each $j \in \mathcal{J}$ the system $Ax = 0$, $x_j = 1$, $x \geq 0$ has no solution; by the Farkas lemma, this implies that for each $j \in \mathcal{J}$ there is a vector $(y^j, z^j) \in \mathbb{R}^m \times \mathbb{R} = \mathbb{R}^{m+1}$ such that $(y^j)^T A + z^j e_j \geq 0$ (where e_j is the j -th unit vector) and such that $z^j < 0$. In particular, $(y^j)^T A \geq 0$ and $((y^j)^T A)_j > 0$, so adding the vectors y^j , $j \in \mathcal{J}$ gives the desired y . $\square$

Proof of Lemma 1. Assume first that w exists. Let $v \in V_n$ be a nonnegative vector that is 0 on $\mathcal{F}^-$. Then $\langle v, w \rangle = 0$ implies that v is 0 on $\mathcal{F}^+$. So v 's support is contained in $\mathcal{H}$, and v must be a scalar multiple $1_{\mathcal{H}}$ by the irreducibility¹ of $\Phi_p(x) = 1 + x + \dots + x^{p-1}$, as desired.

For the other direction, assume that Conjecture 1 holds for n . Let B be a matrix whose rows span $V_n^\perp$, and let A be obtained by truncating the last $P-1$ columns of B . We note the rows of A are vectors in $\mathbb{R}^{n-P+1} = \mathbb{R}^{\mathcal{H} \cup \mathcal{F}^+}$. There is no vector $x \in \mathbb{R}^{\mathcal{H} \cup \mathcal{F}^+}$, $x \geq 0$, such that $Ax = 0$ and such that $x_j > 0$ for some $j \in \mathcal{F}^+$ (otherwise extending x to $\mathbb{R}^n$ with 0's would produce a vector $\bar{x} \geq 0$ nonzero on $\mathcal{F}^+$ such that $B\bar{x} = 0$ and $\bar{x}$ is zero on $\mathcal{F}^-$, contradicting Conjecture 1). By Proposition 1 this implies there is a vector y such that $y^T A \geq 0$ and such that $y^T A$ is positive on $\mathcal{F}^+$. Moreover, $y^T A$ must be 0 on $\mathcal{H}$ because $(y^T A)1_{\mathcal{H}} = y^T(A1_{\mathcal{H}}) = 0$ (where we view $1_{\mathcal{H}}$ as a vector in $\mathbb{R}^{\mathcal{H} \cup \mathcal{F}^+}$). Thus $y^T B$ is a vector in $V_n^\perp$ that is positive on $\mathcal{F}^+$ and zero on $\mathcal{H}$, as desired. $\square$

We call the vector w satisfying the conditions of Lemma 1 a *certificate*. We point out the obvious fact that Lemma 1 gives a means of disproving Conjecture 1 without actually exhibiting a counterexample polynomial divisible by $\Phi_n(x)$ (namely, it suffices to show that the constraints on w determine an empty polytope in $\mathbb{R}^n$). However, from a computational

¹More precisely, if the polynomial associated to v is $a_0 + a_1 x^P + \dots + a_{n-1} x^{(p-1)P}$, then we note that $e^{2\pi i/p}$ is a root of $a_0 + a_1 x + \dots + a_{n-1} x^{p-1}$, from which it follows that $\Phi(x) \mid a_0 + a_1 x + \dots + a_{n-1} x^{p-1}$.

standpoint, this method is not necessarily faster than directly seeking a counterexample polynomial (which reduces to showing non-emptiness of a certain polytope in $\mathbb{R}^n$, or in $\mathbb{R}^{\mathcal{H} \cup \mathcal{F}^+}$).

We note that if one is only interested in showing that the smallest degree of a polynomial with nonnegative coefficients divisible by $\Phi_n(x)$ is at least $(p-1)n/p = n - P$ (without proving uniqueness), then a weaker type of certificate suffices, as given by the following lemma.

Lemma 2. *V_n (n squarefree) contains no nonzero, nonnegative vector that is zero on $\mathcal{F}^- \cup \{n - P\}$ if and only if there exists a $w \in V_n^\perp$ such that w is positive outside $\mathcal{F}^- \cup \{n - P\}$.*

The proof of Lemma 2 is similar to the proof of Lemma 1. We omit it because we will not use Lemma 2.

To apply Lemma 1 (by finding a certificate) we need some understanding of $V_n^\perp$. Thankfully $V_n^\perp$ has a relatively simple structure (first investigated by Rédei [8] and deBruijn [1] and also later by Conway and Jones [2]), outlined below.

Let $G_0(x) = 1 + x^P + \dots + x^{(p-1)P}$ and let $G_i(x) = 1 + x^{Q_i} + \dots + x^{(q_i-1)Q_i}$ for $i \geq 1$ (recall that $Q_i = n/q_i$). It is easy to check by examining the roots of $G_0(x), \dots, G_k(x)$ that $\Phi_n(x) = \gcd(G_0(x), \dots, G_k(x))$. Therefore V_n is spanned by the coefficient vectors of all the translates² of degree at most $n - 1$ of the polynomials $G_0(x), \dots, G_k(x)$. (As a sidenote, since the vanishing sums of roots of unity corresponding to $G_0(x), \dots, G_k(x)$ are regular q -gons with $q = p, q_1, \dots, q_k$ it follows from the above observation that every vanishing sum of n -th roots of unity can be obtained as a linear combination of regular q -gons with $q|n$. If any vanishing sum of n -th roots of unity with nonnegative coefficients could be obtained as a *nonnegative* linear combination of regular q -gons then Conjecture 1 would follow immediately, since the q -gon with the lowest highest power of ζ_n is the p -gon. This is true when n has only 2 prime factors by deBruijn's Theorem [1] but is not the case in general. Schoenberg [11] pointed out the first counterexample (for $n = 30$) and others have followed suit [2, 5–7, 10].)

We now identify vectors in V_n with $(k+1)$ -dimensional arrays of real numbers of size $p \times q_1 \times \dots \times q_k$ using the Chinese Remainder Theorem (CRT). Under this mapping, the array entry with coordinate $(t_0, \dots, t_k)$ —where array coordinates are indexed from 0, like vector indices—is mapped to the vector entry with index $t_0P + t_1Q_1 + \dots + t_kQ_k \bmod n$ (Figs. 1 and 4 show such maps for $n = 5 \cdot 7$ and $n = 3 \cdot 5 \cdot 7$), and we say that $t_0P + t_1Q_1 + \dots + t_kQ_k \bmod n$ is the *index* of the entry $(t_0, \dots, t_k)$. This mapping is a bijection, so there is a one-to-one correspondence between vectors in $\mathbb{R}^n$ and arrays of size $p \times q_1 \times \dots \times q_k$.

Our reason for representing vectors in V_n as arrays is that the generators for V_n described above—the coefficient vectors of the translates of $G_0(x), \dots, G_k(x)$ —correspond to particularly simple arrays. The coefficient vector of $G_0(x) = 1 + x^P + \dots + x^{(p-1)P}$ maps to the 0-1 array whose nonzero entries are $\{(h, 0, \dots, 0) : 0 \leq h \leq p-1\}$. Thus the array corresponding to $G_0(x)$ consists of a “line of 1’s” in the direction of the first

²By *translate* of a polynomial $p(x)$ we mean a polynomial of the form $x^k p(x)$.

28	33	3	8	13	18	23
21	26	31	1	6	11	16
14	19	24	29	34	4	9
7	12	17	22	27	32	2
0	5	10	15	20	25	30

28	33	3	8	$\mathcal{F}_{113}$	$\mathcal{F}_2$	$\mathcal{F}_3$
21	26	31	1	$\mathcal{F}_0$	6	11
$\mathcal{H}$	19	24	29	$\mathcal{F}_{434}$	4	9
7	12	17	22	$\mathcal{F}_3$	27	32
0	$\mathcal{F}_0$	$\mathcal{F}_1$	15	$\mathcal{F}_2$	20	25
						30

Figure 1: A partition of the 5×7 array based on the indices of the entries; entries of index $0 \bmod P = 7$ are in $\mathcal{H}$, entries of index $> iP$ and $< (i + 1)P$ are in $\mathcal{F}_i$.

coordinate axis. The translates of $G_0(x)$ of degree $\leq n - 1$ map to all the parallel lines of 1's in that direction. Similarly the translates of $G_1(x), \dots, G_k(x)$ map to all lines of 1's in the remaining coordinate directions, so that V_n is precisely the set of arrays that are generated as linear combinations of lines of 1's in all the coordinate directions.

We will refer to the lines of 1's as *fibers* (to be perfectly clear: for an array of size $a_1 \times \dots \times a_k$, a fiber is a 0-1 array whose support is $\{(x_1, \dots, h, \dots, x_k) : 0 \leq h \leq a_i - 1\}$ where h is in the i -th coordinate place and where $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k$ are constant, for some i). In other papers [3, 10] we have dubbed arrays (of any size) which, like V_n , are linear combinations of fibers, as *cyclotomic arrays*. Cyclotomic arrays have a number of interesting combinatorial properties. For example, the sum of the entries of any nonnegative, integer-valued cyclotomic array is always a nonnegative integer linear combination of its sidelengths [3], generalizing a property of vanishing sums of roots of unity [5].

Knowing that V_n is spanned by fibers makes clear the structure of $V_n^\perp$: points in $V_n^\perp$ are exactly those arrays of size $p \times q_1 \times \dots \times p_k$ whose column sums are zero for all columns in every direction, since the column sums of an array are zero if and only if the dot product of that array with all fibers is zero. We call an array with zero column sums in every direction a *zero-sum* array. Our blueprint for proving Theorem 1 is thus to exhibit a zero-sum array w (the “certificate”) that is positive on entries whose index is in $\mathcal{F}^+$ and zero on entries whose index is in $\mathcal{H}$. In other words, w is positive on entries of index less than $n - P$ except for entries whose index is $0 \bmod P$, on which w is zero.

3 Low-Dimensional Examples

In this section we use the strategy outlined above to prove Conjecture 1 for the two-prime case $n = 35$ and the three-prime case $n = 105$. We start with $n = 35$. Since 35 is a product of two primes this case already follows from deBruijn’s theorem [1], but the point is to illustrate our method.

The left of Fig. 1 shows the mapping from arrays of size 5×7 to $\mathbb{R}^{35}$ given by our application of the CRT. A “layer” of the array refers to a horizontal row of 7 cells. On

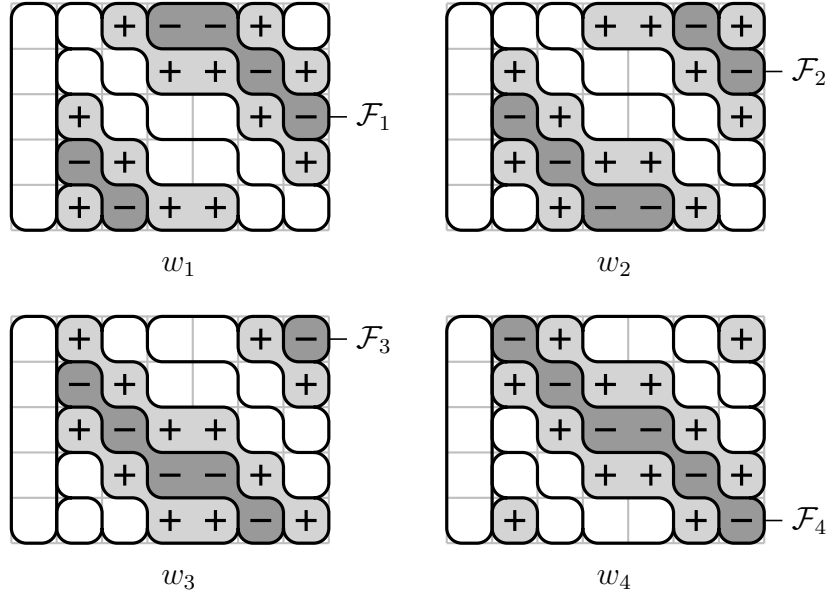


Figure 2: The positive and negative supports of w_1, w_2, w_3, w_4 ; entries with $+$'s are positive, entries with $-$'s are negative, blank entries are 0.

the right of Fig. 1 the array cells have been partitioned into $\mathcal{H}, \mathcal{F}_0, \dots, \mathcal{F}_4$ according to their index, using the same definitions for $\mathcal{H}, \mathcal{F}_0, \dots, \mathcal{F}_{p-1}$ as in Section 2. We also define $\mathcal{F}^+ = \mathcal{F}_0 \cup \dots \cup \mathcal{F}_{p-2} = \mathcal{F}_0 \cup \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ and $\mathcal{F}^- = \mathcal{F}_4$.

We wish, thus, to find a 5×7 zero-sum array w that is zero on $\mathcal{H}$ and positive on $\mathcal{F}^+$. Note that since every column besides the first contains exactly one entry from $\mathcal{F}^- = \mathcal{F}_4$, w must be negative on $\mathcal{F}^-$ in order for column sums to be zero. The $\mathcal{F}_i$'s are translates of each other because going up a layer adds $P \bmod n$ to the index of an entry. As a consequence the $\mathcal{F}_i$'s induce an identical partition of each layer.

We construct our array w as a linear combination of 5×7 zero-sum arrays w_1, w_2, w_3, w_4 such that w_i is negative on $\mathcal{F}_i$, positive on $\mathcal{F}_{i-1} \cup \mathcal{F}_{i+1}$ and 0 elsewhere (indices referring to elements in the set $\{0, 1, \dots, 4\}$ are taken mod 5). The signs of $w_1, \dots, w_4$ are sketched in Fig. 2. Then if $\alpha_2, \alpha_3, \alpha_4$ are taken sufficiently quickly increasing the array $w = w_1 + \alpha_2 w_2 + \alpha_3 w_3 + \alpha_4 w_4$ will be zero-sum, positive on $\mathcal{F}_0 \cup \dots \cup \mathcal{F}_3$ and 0 on $\mathcal{H}$, as desired. Note w_2, w_3, w_4 can be obtained from w_1 by shifting w_1 up by respectively 1, 2 or 3 layers cyclically, so it suffices to construct w_1 .

We build w_1 as the sum of 5 zero-sum arrays $a_0, \dots, a_4$ where a_i is negative on the intersection of layers $i, i+1$ with $\mathcal{F}_1$ and positive on the intersection of layers $i, i+1$ with respectively $\mathcal{F}_0$ and $\mathcal{F}_2$, and 0 elsewhere. The sign patterns for $a_0, \dots, a_4$ are sketched in Fig. 3. Such arrays $a_0, \dots, a_4$ obviously exist. Setting $w_1 = a_0 + \dots + a_4$ we obtain a zero-sum array that is negative on $\mathcal{F}_1$, positive on $\mathcal{F}_0 \cup \mathcal{F}_2$ and 0 elsewhere, as desired. This completes the construction.

Now let $n = 105 = 3 \cdot 5 \cdot 7$. In this case $p = 3, q_1 = 5, q_2 = 7$ and $P = n/p = 35, Q_1 = n/q_1 = 21, Q_2 = n/q_2 = 15$. Fig. 4 shows the CRT mapping from the $3 \times 5 \times 7$ array to

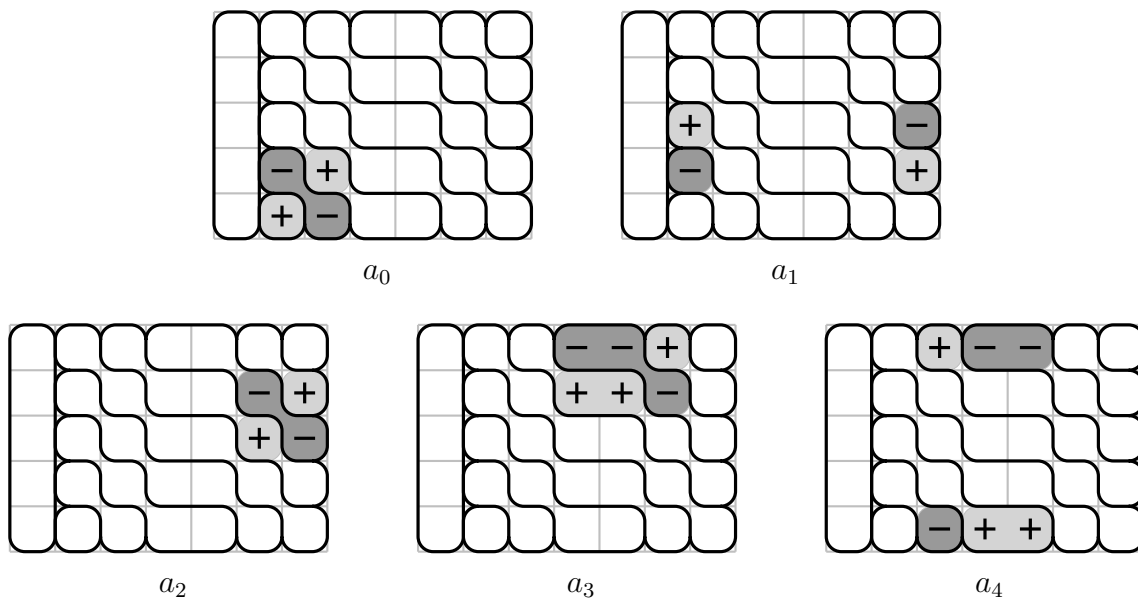


Figure 3: The positive and negative supports of $a_0, \dots, a_4$.

84	99	9	24	39	54	69
63	78	93	3	18	33	48
42	57	72	87	102	12	27
21	36	51	66	81	96	6
0	15	30	45	60	75	90

14	29	44	59	74	89	104
98	8	23	38	53	68	83
77	92	2	17	32	47	62
56	71	86	101	11	26	41
35	50	65	80	95	5	20

49	64	79	94	4	19	34
28	43	58	73	88	103	13
7	22	37	52	67	82	97
91	1	16	31	46	61	76
70	85	100	10	25	40	55

Figure 4: The mapping from the $3 \times 5 \times 7$ array to $\mathbb{Z}_{105}$ given by the CRT.

84	99	9	24	39	$\mathcal{F}_1$	69
63	78	93	3	$\mathcal{F}_0$	33	48
42	57	72	$\mathcal{F}_2$	102	12	27
21	36	$\mathcal{F}_1$	66	81	96	6
$\mathcal{H}$	$\mathcal{F}_0$	30	45	60	75	90

14	29	44	59	74	$\mathcal{F}_2$	104
98	8	23	38	$\mathcal{F}_1$	68	83
77	92	2	$\mathcal{F}_0$	32	47	62
56	71	$\mathcal{F}_2$	101	11	26	41
$\mathcal{H}$	$\mathcal{F}_1$	65	80	95	5	20

49	64	79	94	4	$\mathcal{F}_0$	34
28	43	58	73	$\mathcal{F}_2$	103	13
7	22	37	$\mathcal{F}_1$	67	82	97
91	1	$\mathcal{F}_0$	31	46	61	76
$\mathcal{H}$	$\mathcal{F}_2$	100	10	25	40	55

Figure 5: Partitioning the $3 \times 5 \times 7$ array.

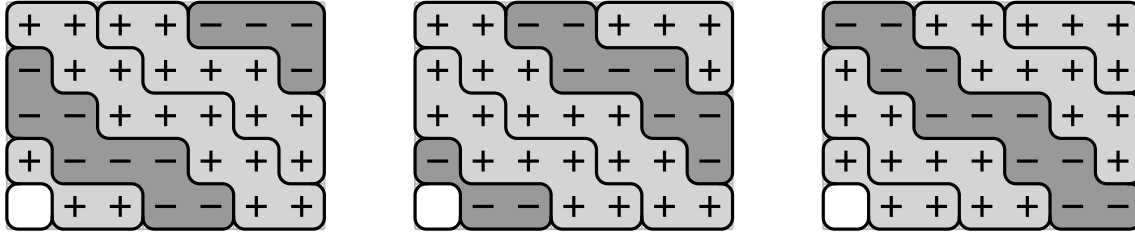


Figure 6: The positive and negative supports of w_1 .

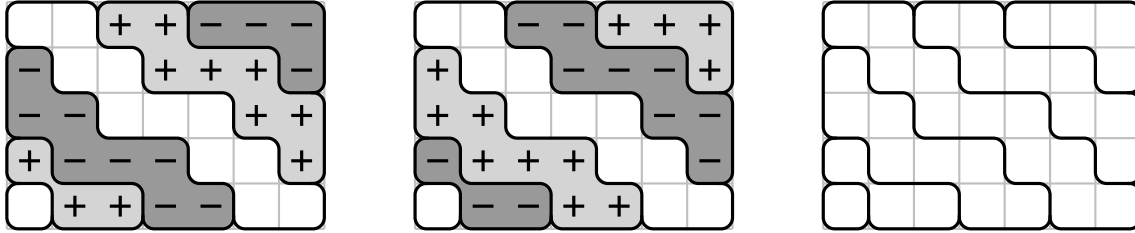


Figure 7: The positive and negative supports of a_0 . The leftmost layer gives the positive and negative supports of the 5×7 array u_0 .

$\mathbb{Z}_{105}$. A “layer” of the array again means a layer perpendicular to the smallest coordinate direction, so there are 3 layers each consisting of an array of size 5×7 . Subscripts referring to values in the sets $\mathbb{Z}_3 = \{0, 1, 2\}$, $\mathbb{Z}_5 = \{0, 1, \dots, 4\}$ or $\mathbb{Z}_7 = \{0, 1, \dots, 6\}$ are taken mod 3, 5 and 7 respectively.

The partition of the array cells into the sets $\mathcal{H}, \mathcal{F}_0, \dots, \mathcal{F}_{p-1}$ is shown in Fig. 5. Note that the $\mathcal{F}_i$ ’s again induce an identical partitioning of each layer, for the same reason that going up one layer increases the index of an entry by $P \bmod n$. Again, we wish to construct a zero-sum array w which is 0 on $\mathcal{H}$ and positive on $\mathcal{F}^+ = \mathcal{F}_0 \cup \dots \cup \mathcal{F}_{p-1} = \mathcal{F}_0 \cup \mathcal{F}_1$. We can note that if w exists it will be negative on $\mathcal{F}^- = \mathcal{F}_2$ for the column sums of length 3 to be zero.

We construct w as a linear combination of $p - 1 = 2$ zero-sum arrays w_1, w_2 where w_i is negative on $\mathcal{F}_i$ and positive on $\mathcal{F}_{i-1} \cup \mathcal{F}_{i+1}$ and zero elsewhere. The sign pattern of w_1 is sketched in Fig. 6. Then if we take $\alpha > 0$ sufficiently large $w = w_1 + \alpha w_2$ will be a zero-sum array positive on $\mathcal{F}_0 \cup \mathcal{F}_1$, negative on $\mathcal{F}_2$ and 0 on $\mathcal{H}$, as desired. Note that w_2 can be obtained by shifting w_1 up by one layer cyclically so it suffices to construct w_1 . (This example with $p = 3$ is a little unfortunate since one could directly take $w = w_2$; a better example would have had $p = 5$, but would have been harder to draw.)

We build w_1 as the sum of $p = 3$ zero-sum arrays a_0, a_1, a_2 where a_i is negative on the intersection of layers $i, i + 1$ with $\mathcal{F}_1$, positive on the intersection of layers $i, i + 1$ with respectively $\mathcal{F}_0$ and $\mathcal{F}_2$, and 0 elsewhere. The sign pattern for a_0 is sketched in Fig. 7. The sum $a_0 + a_1 + a_2$ is negative on $\mathcal{F}_1$, positive on $\mathcal{F}_0 \cup \mathcal{F}_2$ and 0 elsewhere (namely on $\mathcal{H}$), as desired. The reader can compare Figs. 3 and 7. In particular, note that the support of a_i is each time contained in two layers of opposite sign.

Because a_i is a zero-sum array the layers i and $i + 1$ of a_i must be negative of one another (other layers being 0). Thus to determine a_i we only need to determine the i -th layer of a_i , which must be positive on $\mathcal{F}_0$, negative on $\mathcal{F}_1$, zero elsewhere and have zero column sums in the two coordinate directions. We illustrate how to do this for a_0 , in which case the problem is to construct a 5×7 zero-sum array u_0 having the positive and negative supports of the leftmost 5×7 array in Fig. 7.

We now describe the construction of u_0 , sketched in Fig. 8. We found it useful to “unroll” the torus $\mathbb{Z}_5 \times \mathbb{Z}_7$ onto $\mathbb{Z} \times \mathbb{Z}$. Let $\mathcal{G}_0 = \{(j_1, j_2) \in \mathbb{Z}_5 \times \mathbb{Z}_7 : (0, j_1, j_2) \in \mathcal{F}_0\}$, $\mathcal{G}_1 = \{(j_1, j_2) \in \mathbb{Z}_5 \times \mathbb{Z}_7 : (0, j_1, j_2) \in \mathcal{F}_1\}$. (Thus $\mathcal{G}_0$ and $\mathcal{G}_1$ are respectively the positive and negative supports of u_0 .) Let $z = (Q_1, Q_2) = (21, 15) \in \mathbb{Z}^2$ and let $\mathcal{G}'_0 = \{x \in \mathbb{Z}^2 : 0 < \langle x, z \rangle < P\}$, $\mathcal{G}'_1 = \{x \in \mathbb{Z}^2 : P < \langle x, z \rangle < 2P\}$. Let $\rho : \mathbb{Z}^2 \rightarrow \mathbb{Z}_5 \times \mathbb{Z}_7$ be the projection map $\rho(j_1, j_2) = (j_1 \bmod 5, j_2 \bmod 7)$. It is easy to see that $\rho(\mathcal{G}'_0) = \mathcal{G}_0$, $\rho(\mathcal{G}'_1) = \mathcal{G}_1$ from the definition of $\mathcal{F}_0, \mathcal{F}_1$.

Note that $\rho(x) = \rho(y) \implies \langle x, z \rangle - \langle y, z \rangle \equiv 0 \pmod n$, but if $x, y \in \mathcal{G}'_0 \cup \mathcal{G}'_1$ then $|\langle x, y \rangle - \langle y, z \rangle| < 2P \leq n$ so $\rho(x) = \rho(y) \implies \langle x, z \rangle = \langle y, z \rangle$ for $x, y \in \mathcal{G}'_0 \cup \mathcal{G}'_1$. Thus if $f : \mathbb{Z}^2 \rightarrow \mathbb{R}$ is a function with support $\mathcal{G}'_0 \cup \mathcal{G}'_1$ such that $\langle x, z \rangle = \langle y, z \rangle \implies f(x) = f(y)$ there is a well-defined 5×7 array u_f with support $\mathcal{G}_0 \cup \mathcal{G}_1$ whose $\rho(x)$ -th entry has value $f(x)$ for any $x \in \mathcal{G}'_0 \cup \mathcal{G}'_1$. A function f such that $\langle x, z \rangle = \langle y, z \rangle \implies f(x) = f(y)$ will be called *periodic*.

A *line of integers* is a subset of $\mathbb{Z}^2$ of the form $\{(x, y) : x \in \mathbb{Z}\}$ for some $y \in \mathbb{Z}$ or $\{(x, y) : y \in \mathbb{Z}\}$ for some $x \in \mathbb{Z}$. We say that f is *zero-sum* if its support is finite on every line of integers and sums to zero on every line of integers. It is clear that if f is periodic and has support $\mathcal{G}'_0 \cup \mathcal{G}'_1$ then u_f is zero-sum if f is zero-sum. Thus the construction of u_0 reduces to exhibiting a periodic zero-sum function f on $\mathbb{Z}^2$ that is positive on $\mathcal{G}'_0$, negative on $\mathcal{G}'_1$ and zero elsewhere.

Fig. 8 shows our method for constructing such a function f . First we consider $\mathbb{Z}^2$ as a subset of $\mathbb{R}^2$ the natural way. The square $[i - \frac{1}{2}, i + \frac{1}{2}] \times [j - \frac{1}{2}, j + \frac{1}{2}]$ centered at a point $(i, j) \in \mathbb{Z}^2$ is called the *cell around* (i, j) or *cell centered at* (i, j) , and if $\mathcal{A}$ is any subset of $\mathbb{Z}^2$ we write $\overline{\mathcal{A}}$ for the union of all cells centered at points in $\mathcal{A}$. Let L_2 be the line $\{(x, y) : \langle (x, y), z \rangle = P\} \subseteq \mathbb{R}^2$ and let L_1, L_3 be two translates of L_2 equidistant from L_2 such that (i) L_1 and L_3 are contained in $\overline{\mathcal{G}'_0 \cup \mathcal{G}'_1}$, and (ii) the region between L_1 and L_2 (resp. L_3 and L_2) intersects every cell in $\overline{\mathcal{G}'_0}$ (resp. $\overline{\mathcal{G}'_1}$) in a nonzero area. The existence of L_1, L_2, L_3 is clear from Fig. 8 and is formally proved in the next section.

Let R^+ be the region between L_1 and L_2 and R^- the region between L_2 and L_3 . We define

$$f(x) = \text{vol}(\overline{x} \cap R^+) - \text{vol}(\overline{x} \cap R^-)$$

where $\overline{x} = \overline{\{x\}}$ is the cell around $x \in \mathbb{Z}^2$ and $\text{vol}(\cdot)$ is two-dimensional area. Then the support of f is contained in $\mathcal{G}'_0 \cup \mathcal{G}'_1$ because cells not in $\overline{\mathcal{G}'_0 \cup \mathcal{G}'_1}$ have no area in $R^+ \cup R^-$, and f is zero-sum because if $\mathcal{L}$ is any line of integers,

$$\sum_{x \in \mathcal{L}} f(x) = \text{vol}(\overline{\mathcal{L}} \cap R^+) - \text{vol}(\overline{\mathcal{L}} \cap R^-) = 0$$

as R^+, R^- have equal thickness. Moreover $f(x)$ is positive on $\mathcal{G}'_0$ and negative on $\mathcal{G}'_1$,

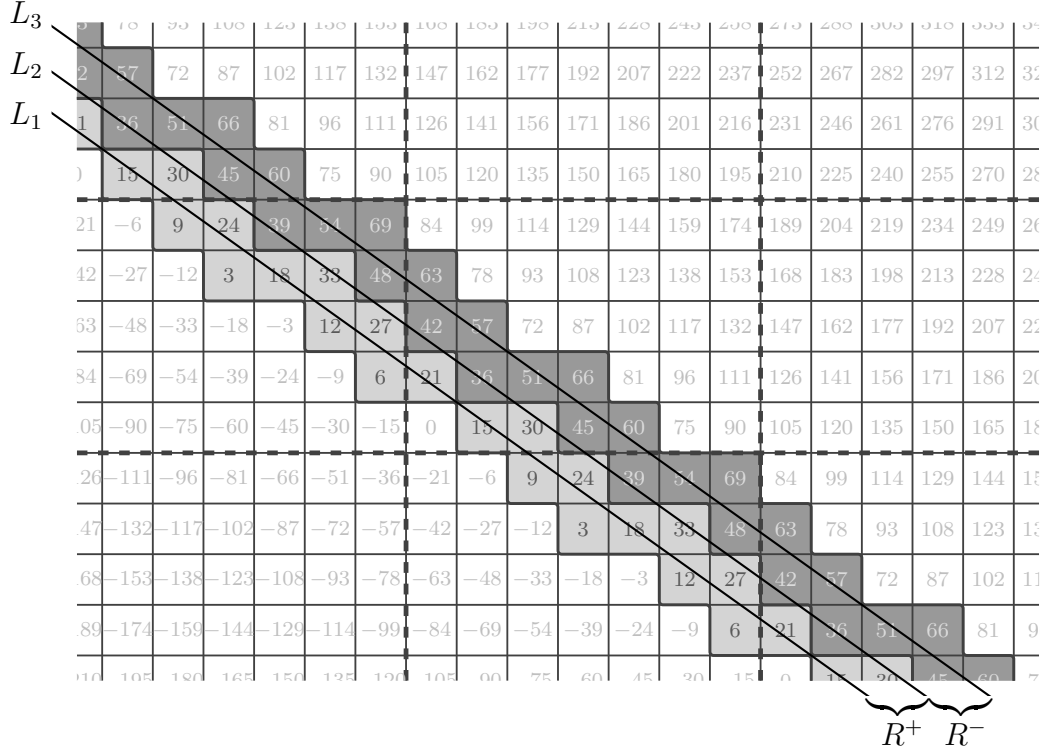


Figure 8: Construction of a zero-sum function f on $\mathbb{Z}^2$ with positive support $\mathcal{G}'_0$ and negative support $\mathcal{G}'_1$. The points in $\mathbb{Z}^2$ (not shown) have unit squares centered around them (shown). The x and y axes are flipped to match the orientation of Fig. 5. The point $(i, j) \in \mathbb{Z}^2$ is labeled with the number $\langle (i, j), (Q_1, Q_2) \rangle = 21i + 15j$, shown in its square; $\mathcal{G}'_0$ is the set of points whose label is between 0 and $P = 35$ and $\mathcal{G}'_1$ is the set of points whose label is between P and $2P$. The area $\overline{\mathcal{G}'_0}$, in light grey, is the set of cells around points in $\mathcal{G}'_0$, and the area $\overline{\mathcal{G}'_1}$, in dark grey, is the set of cells around points in $\mathcal{G}'_1$. Three parallel equidistant lines L_1, L_2, L_3 contained in $\overline{\mathcal{G}'_0} \cup \overline{\mathcal{G}'_1}$ are drawn, with the property that each cell in $\overline{\mathcal{G}'_0}$ has a larger area of intersection with the region R^+ between L_1 and L_2 than with the region R^- between L_2 and L_3 , and each cell in $\overline{\mathcal{G}'_1}$ has a larger area of intersection with R^- than with R^+ . The value of f at $x \in \mathbb{Z}^2$ is then defined to be $f(x) = \text{vol}(\overline{x} \cap R^+) - \text{vol}(\overline{x} \cap R^-)$, where $\overline{x}$ is the square (cell) centered at x and $\text{vol}(\cdot)$ is area.

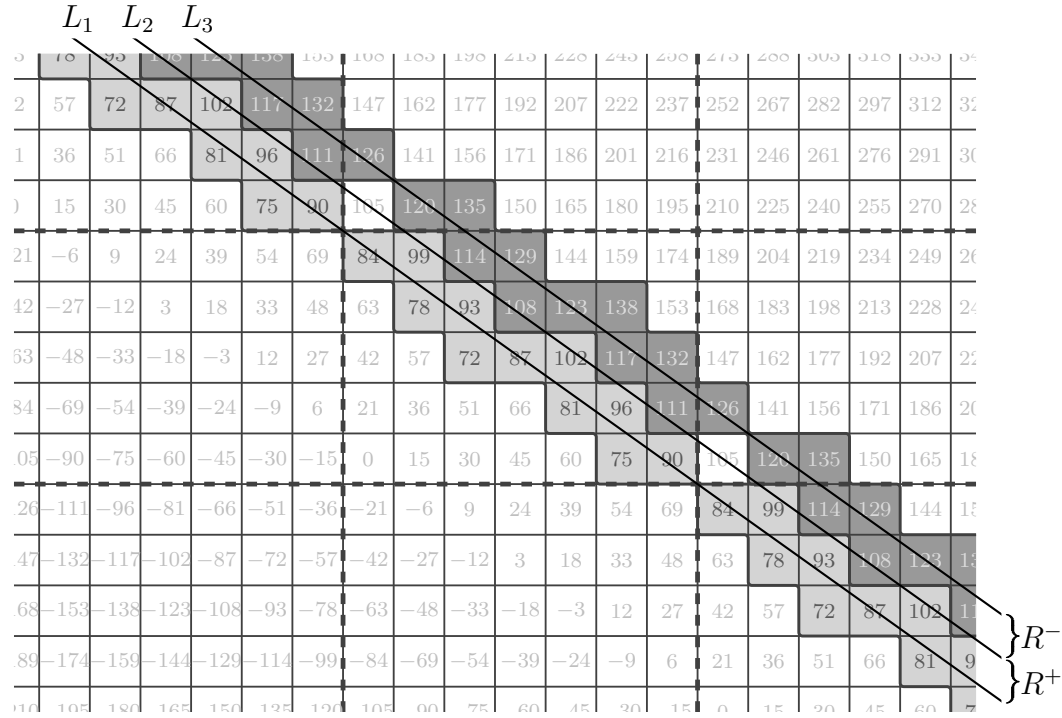


Figure 9: Construction of a zero-sum function f on $\mathbb{Z}^2$ with positive support $\mathcal{G}'_2 = \{(i, j) : 2P < iQ_1 + jQ_2 < 3P\}$ (cells $\overline{\mathcal{G}'_2}$ in light grey) and negative support $\mathcal{G}'_3 = \{(i, j) : 3P < iQ_1 + jQ_2 < 4P\}$ (cells $\overline{\mathcal{G}'_3}$ in dark grey). The construction is similar to that of Fig. 8 except for the presence of blank cells intersecting R^+ , R^- (corresponding to points (i, j) with $iQ_1 + jQ_2 = 3P = 105$). Because L_2 passes through the center of these cells, $f(x) = \text{vol}(\overline{x} \cap R^+) - \text{vol}(\overline{x} \cap R^-)$ still has support $\mathcal{G}'_2 \cup \mathcal{G}'_3$.

since points in $\mathcal{G}'_0$ are below L_2 and points in $\mathcal{G}'_1$ and since every cell in $\overline{\mathcal{G}'_0} \cup \overline{\mathcal{G}'_1}$ intersects $R^+ \cup R^-$ in a nonzero area (point (ii) above). Finally f is periodic since L_1, L_2, L_3 are perpendicular to z .

This concludes the construction of u_0 , and hence of a_0 . (One can also construct u_0 by other, non-geometric approaches, but the geometric method is the only one we found to generalize to higher dimensions (higher numbers of primes).)

The remainder of the construction rests on making 5×7 zero sum arrays u_1 and u_2 whose positive and negative supports are respectively $\mathcal{G}_1$ and $\mathcal{G}_2$ (for u_1) and $\mathcal{G}_2$ and $\mathcal{G}_0$ (for u_2), where $\mathcal{G}_2 = \{(j_1, j_2) \in \mathbb{Z}_5 \times \mathbb{Z}_7 : (0, j_1, j_2) \in \mathcal{F}_2\}$. The arrays u_1, u_2 can be found with the same method as u_0 , though the construction of u_2 presents a difference which is worth mentioning. Fig. 9 shows the analog of Fig. 8 for the construction of u_2 ; here $\mathcal{G}'_2 = \{x \in \mathbb{Z}^2 : 2P < \langle x, z \rangle < 3P\}$, so $\rho(\mathcal{G}'_2) = \mathcal{G}_2$, and $\mathcal{G}'_3 = \{x \in \mathbb{Z}^2 : 3P < \langle x, z \rangle < 4P\}$ is a translate of $\mathcal{G}'_0$ by $(q_1, 0)$ (or $(0, q_2)$), so $\rho(\mathcal{G}'_3) = \rho(\mathcal{G}'_0) = \mathcal{G}_0$ (we cannot use $\mathcal{G}'_0$ and $\mathcal{G}'_2$ because $\overline{\mathcal{G}'_0}$ and $\overline{\mathcal{G}'_2}$ are non-contiguous regions). In this case the region $\overline{\mathcal{G}'_2} \cup \overline{\mathcal{G}'_3}$ is

not simply connected. The “embedded” blank cells correspond to integer points on the line $L_2 = \{x : \langle x, z \rangle = 3P = n\}$, which are preimages of $(0, 0) \in \mathbb{Z}_5 \times \mathbb{Z}_7$ under ρ (by contrast, the lines $\langle x, z \rangle = P$ and $\langle x, z \rangle = 2P$ contain no integer points, as the gcd of the coordinates in z does not divide P or $2P$) (also note that $(0, 0) \in \mathbb{Z}_5 \times \mathbb{Z}_7$ is the set $\{(j_1, j_2) \in \mathbb{Z}_5 \times \mathbb{Z}_7 : (0, j_1, j_2) \in \mathcal{H}\}$). However f remains zero on the integer points in L_2 since the cells centered at these points have equal areas in R^+ and R^- , thus maintaining the correctness of the support.

This concludes the discussion of the cases $n = 35$ and $n = 105$. In higher dimensions (i.e. for n with more prime factors) an obstacle which arises is that the shaded region is too “thin” to contain a hyperplane; then L_1 , L_2 and L_3 (which are hyperplanes in the general case) cannot be embedded in the shaded region, and the whole approach breaks down. In fact, the condition $2/p > 1/q_1 + \cdots + 1/q_k$ from Theorem 1 guarantees that the shaded region is “thick enough” to accomodate the sandwich of hyperplanes L_1 , L_2 , L_3 . Moreover, in the general case and when n is a product of many large distinct primes, the region $\mathcal{F}_{p-1}$ itself becomes a precariously thin slice of the whole array, and the union of the cells in $\mathcal{F}_{p-1}$ becomes very far from containing a hyperplane. This causes us to be pessimistic that the certificate w always exists, and hence that Conjecture 1 holds. This pessimism is reinforced by the fact that the constraints on w for the case $n = pq_1 \cdots q_k$ are a strict subset of those for the case $n = pq_1 \cdots q_k q_{k+1}$, where $q_{k+1} > p$ can be any prime distinct from $q_1, \dots, q_k$. (The latter fact is not surprising considering that Conjecture 1 for $n = pq_1 \cdots q_k$ is implied by Conjecture 1 for $n = pq_1 \cdots q_{k+1}$.) See also the comments after Problem 1 below. Appending primes *ad infinitum* also leads us to consider the existence of an “infinite certificate” for a given prime p . Such an infinite certificate would be, more precisely, an infinite array w_p of size $p \times q_1 \times q_2 \times q_3 \times \cdots$ where $q_1, q_2, q_3, \dots$ enumerate the primes greater than p , where for each k the restriction of w_p to entries that have ℓ -th coordinate 0 for $\ell > k+1$ constitutes a certificate for the case $n = pq_1 \cdots q_k$. For example, we know *finite* certificates³ of all sizes exist for $p = 2$ by Theorem 1, but it is not obvious whether these finite certificates can be extended to an infinite array w_2 as just described. We leave this as another interesting open problem.

4 The General Case

Take $n = pq_1 \cdots q_k$ with $p < q_1, \dots, q_k$. As before we work with arrays of size $p \times q_1 \times \cdots \times q_k$ where the index of entry $(j_0, j_1, \dots, j_k)$ is $j_0P + j_1Q_1 + \cdots + j_kQ_k \bmod n$, where $P = n/p$ and $Q_i = n/q_i$. As before, $\mathcal{H}$ is the set of entries of index 0 mod P and $\mathcal{F}_i$ is the set of entries whose index is greater than iP but less than $(i+1)P$, $0 \leq i \leq p-1$. By Lemma 1 it is sufficient to construct a zero-sum array w of size $p \times q_1 \times \cdots \times q_k$ which is zero on $\mathcal{H}$ and positive on $\mathcal{F}^+ = \mathcal{F}_0 \cup \cdots \cup \mathcal{F}_{p-2}$. A “layer” of the array is an array of size $q_1 \times \cdots \times q_k$ obtained by fixing a value for the first coordinate (so there are p layers). We

³We know these finite certificates exist because Theorem 1 is true for even n by physical considerations; it would indeed be interesting to find combinatorial and/or algebraic constructions (or merely proofs of existence) for these certificates.

note that the sets $\mathcal{F}_i$ induce identical partitions of every layer because the index of an entry increases by $P \bmod n$ as we go up one layer, i.e. if an entry is in $\mathcal{F}_i$ then the entry whose first coordinate is one greater is in $\mathcal{F}_{i+1}$ (as usual indices referring to numbers in $\{0, \dots, p-1\}$ are taken mod p).

We construct w as a linear combination of zero-sum arrays $w_1, \dots, w_{p-1}$ such that each w_i is negative on $\mathcal{F}_i$, positive on $\mathcal{F}_{i+1} \cup \mathcal{F}_{i-1}$ and zero elsewhere. Then if $\alpha_2, \dots, \alpha_{p-1} > 0$ are taken sufficiently quickly increasing $w = w_1 + \alpha_2 w_2 + \dots + \alpha_{p-1} w_{p-1}$ is negative only on $\mathcal{F}_{p-1}$ and positive everywhere else except on $\mathcal{H}$, where it is 0. Because $(j_0, j_1, \dots, j_k) \in \mathcal{F}_i \implies (j_0 + 1, j_1, \dots, j_k) \in \mathcal{F}_{i+1}$ it is sufficient to construct the array w_1 , as $w_2, \dots, w_{p-1}$ can be obtained from w_1 by translation.

We construct w_1 as a sum of zero-sum arrays $a_0, \dots, a_{p-1}$ where a_i is negative on the intersection of layers $i, i+1$ with $\mathcal{F}_1$ and positive on the intersection of layers $i, i+1$ with respectively $\mathcal{F}_0$ and $\mathcal{F}_2$. Let $\mathcal{G}_i = \{(j_1, \dots, j_k) : (0, j_1, \dots, j_k) \in \mathcal{F}_i\}$. Considering layer i as a standalone $q_1 \times \dots \times q_k$ array, the positive support of a_i in layer i is $\{(j_1, \dots, j_k) : (i, j_1, \dots, j_k) \in \mathcal{F}_0\} = \{(j_1, \dots, j_k) : (0, j_1, \dots, j_k) \in \mathcal{F}_{-i} = \mathcal{F}_{p-i}\} = \mathcal{G}_{p-i}$ and the negative support of a_i in layer i is $\{(j_1, \dots, j_k) : (i, j_1, \dots, j_k) \in \mathcal{F}_1\} = \{(j_1, \dots, j_k) : (0, j_1, \dots, j_k) \in \mathcal{F}_{1-i} = \mathcal{F}_{p-i+1}\} = \mathcal{G}_{p-i+1}$. Therefore in order to construct a_i it is sufficient to construct a zero-sum $q_1 \times \dots \times q_k$ array u_{p-i} which is positive on $\mathcal{G}_{p-i}$, negative on $\mathcal{G}_{p-i+1}$ and zero elsewhere. Then a_i can be obtained by placing u_{p-i} at layer i and $-u_{p-i}$ at layer $i+1$. Thus the construction of w_1 reduces to the construction of zero-sum arrays $u_0, \dots, u_{p-1}$ of size $q_1 \times \dots \times q_k$ such that u_i is positive on $\mathcal{G}_i$, negative on $\mathcal{G}_{i+1}$ and zero elsewhere.

The problem of constructing the u_i 's can be rephrased in $\mathbb{Z}^k$. Let $z = (Q_1, \dots, Q_k) \in \mathbb{Z}^k$ and let $\mathcal{G}'_i = \{x \in \mathbb{Z}^k : iP < \langle x, z \rangle < (i+1)P\}$ for $0 \leq i \leq p$. Let $\rho : \mathbb{Z}^k \rightarrow \mathbb{Z}_{q_1} \times \dots \times \mathbb{Z}_{q_k}$ be the projection defined by $\rho(x_1, \dots, x_k) = (x_1 \bmod q_1, \dots, x_k \bmod q_k)$. As before, $\rho(\mathcal{G}'_i) = \mathcal{G}_i$ (and $\rho(\mathcal{G}'_p) = \mathcal{G}_0$). Moreover $\rho(x) = \rho(y) \implies \langle x, z \rangle - \langle y, z \rangle \equiv 0 \bmod n$, but if $x, y \in \mathcal{G}'_i \cup \mathcal{G}'_{i+1}$ then $|\langle x, y \rangle - \langle y, z \rangle| < 2P \leq n$ so $\rho(x) = \rho(y) \implies \langle x, z \rangle = \langle y, z \rangle$ for $x, y \in \mathcal{G}'_i \cup \mathcal{G}'_{i+1}$. Thus if $f : \mathbb{Z}^k \rightarrow \mathbb{R}$ is a function with support $\mathcal{G}'_i \cup \mathcal{G}'_{i+1}$ such that $\langle x, z \rangle = \langle y, z \rangle \implies f(x) = f(y)$ there is a well-defined $q_1 \times \dots \times q_k$ array u_f whose $(j_1, \dots, j_k)$ -th entry has value $f(x)$ for any $x \in \mathcal{G}'_i \cup \mathcal{G}'_{i+1}$ with $\rho(x) = (j_1, \dots, j_k)$ or has value 0 if $(j_1, \dots, j_k) \notin \rho(\mathcal{G}'_i \cup \mathcal{G}'_{i+1}) = \mathcal{G}_i \cup \mathcal{G}_{i+1}$, for all $0 \leq i \leq p-1$.

A *line of integers* is a set of the form $\{(c_1, \dots, c_{j-1}, x_j, c_{j+1}, \dots, c_k) : x_j \in \mathbb{Z}\}$ where the c_i 's are integer constants, for any $j \in \{1, \dots, k\}$. A function f on $\mathbb{Z}^k$ is a *zero-sum* function if its support is finite on every line of integers and sums to zero on every line of integers. If $f_i : \mathbb{Z}^k \rightarrow \mathbb{R}$ is a zero-sum function that is positive on $\mathcal{G}'_i$, negative on $\mathcal{G}'_{i+1}$, zero elsewhere and such that $\langle x, z \rangle = \langle y, z \rangle \implies f_i(x) = f_i(y)$ then it is straightforward to check that u_{f_i} is a zero-sum array that is positive on $\mathcal{G}_i$, negative on $\mathcal{G}_{i+1}$ and zero elsewhere and one may therefore take $u_i = u_{f_i}$. Conjecture 1 thus reduces to an instance of the following more general problem:

Problem 1. Let $P > Q_1, \dots, Q_k$ be positive rationals and let $z = (Q_1, \dots, Q_k)$. Let $h \in \mathbb{R}$ and let $\mathcal{G}^+ = \{x \in \mathbb{Z}^k : h - P < \langle x, z \rangle < h\}$, $\mathcal{G}^- = \{x \in \mathbb{Z}^k : h < \langle x, z \rangle < h + P\}$. Find a zero-sum function f on $\mathbb{Z}^k$ such that $\langle x, z \rangle = \langle y, z \rangle \implies f(x) = f(y)$ and that is positive on $\mathcal{G}^+$, negative on $\mathcal{G}^-$ and zero elsewhere.

Note that if some line of integers contains a point in $\mathcal{G}^+$ but no point in $\mathcal{G}^-$ or vice-versa then Problem 1 has no solution. The condition $P > Q_1, \dots, Q_k$, however, precludes this, as translating x by one unit in the i -th coordinate direction increases $\langle x, z \rangle$ by $Q_i < P$. We also note that an argument using averaging and compactness shows that the condition $\langle x, z \rangle = \langle y, z \rangle \implies f(x) = f(y)$ is without loss of generality: if f exists without this condition, there also exists an f that satisfies this condition. Hence this condition can be included for free.

Problem 1 is trivial for $k = 1$ and can be solved by combinatorial methods for $k = 2$, but is seemingly very hard for $k \geq 3$. Moreover since all the constraints of a k -dimensional instance of the problem can be strictly contained⁴ in a $(k + 1)$ -dimensional instance of the problem, there seems to be no “good reason” why Problem 1 should always have a solution, and indeed it does not. Small sets of parameters for which Problem 1 doesn’t have a solution are, for example, obtained by putting $k = 3$, $h = 0$, $P = n/p$, $Q_1 = n/q_1$, $Q_2 = n/q_2$, $Q_3 = n/q_3$, $n = pq_1q_2q_3$ with either

$$(p, q_1, q_2, q_3) = (6, 7, 8, 9)$$

or

$$(p, q_2, q_2, q_3) = (11, 13, 17, 19).$$

The latter setting of parameters is in fact derived from the smallest value of n not covered by Theorem 1, being $n = 46189 = 11 \cdot 13 \cdot 17 \cdot 19$. These negative results were obtained by using an exact rational LP solver, in this case the program CDD written by Komei Fukuda [4]. (Checking emptiness of the polytope arising from the parameters $(p, q_2, q_2, q_3) = (11, 13, 17, 19)$ took us 14.5 hours on a laptop; for $(p, q_1, q_2, q_3) = (6, 7, 8, 9)$, the computation takes around two minutes.) We note the negative result for $n = 46189$ does not disprove Conjecture 1 for $n = 46189$ since our reduction from the certificate-finding problem to Problem 1 is not if-and-only-if.

We will show that Problem 1 has a solution when $2P > Q_1 + \dots + Q_k$, which is equivalent to the condition $2/p > 1/q_1 + \dots + 1/q_k$ when $P = n/p$ and $Q_j = n/q_j$. Theorem 1 follows as a corollary. Our solution is a direct extension of the method Figs. 8, 9 to higher dimensions. However, the existence of the analogues of the lines L_1, L_2, L_3 needs to be carefully proved in the k -dimensional case.

Theorem 2. *Problem 1 has a solution when $2P > Q_1 + \dots + Q_k$.*

Proof. Let $\bar{x} = [x_1 - \frac{1}{2}, x_1 + \frac{1}{2}] \times \dots \times [x_k - \frac{1}{2}, x_k + \frac{1}{2}]$ for any $x = (x_1, \dots, x_k) \in \mathbb{Z}^k$ (the “cell around x ” or “cell centered at x ”) and let $\bar{\mathcal{S}} = \bigcup_{x \in \mathcal{S}} \bar{x}$ for any set $\mathcal{S} \subseteq \mathbb{Z}^k$. Let $\mathcal{I} = \{x \in \mathbb{Z}^k : \langle x, z \rangle = h\}$, which may be empty. We will exhibit three parallel hyperplanes H_1, H_2, H_3 in $\mathbb{R}^k$ with H_2 equidistant to H_1 and H_3 such that, if R^+ is the region between H_1 and H_2 and R^- is the region between H_2 and H_3 ,

- (i) $\text{vol}(\bar{x} \cap R^+) > \text{vol}(\bar{x} \cap R^-)$ for all $x \in \mathcal{G}^+$

⁴This is indeed reminiscent of Conjecture 1 itself. (We are alluding to the fact that the case $n = m$ of Conjecture 1 is implied by the case $n = mq$ for any prime q larger than the smallest prime dividing m .)

- (ii) $\text{vol}(\bar{x} \cap R^+) < \text{vol}(\bar{x} \cap R^-)$ for all $x \in \mathcal{G}^-$
- (iii) $\text{vol}(\bar{x} \cap R^+) = \text{vol}(\bar{x} \cap R^-)$ for all $x \in \mathcal{I}$
- (iv) $\text{vol}(\bar{x} \cap R^+) = 0, \text{vol}(\bar{x} \cap R^-) = 0$ for all $x \notin \mathcal{G}^+ \cup \mathcal{G}^- \cup \mathcal{I}$

where ‘vol’ is k -dimensional volume. If (i)–(iv) hold it is clear that the function $f(x) = \text{vol}(\bar{x} \cap R^+) - \text{vol}(\bar{x} \cap R^-)$ is periodic, zero-sum, and has the required sign and support (periodicity follows from the fact that H_1, H_2, H_3 are necessarily normal to z).

Let $\mathbf{1} = (1, \dots, 1) \in \mathbb{Z}^k$. Let $\alpha = \sup\{\langle y + \frac{1}{2}\mathbf{1}, z \rangle : y \in \mathbb{Z}^k, \langle y, z \rangle \leq h - P\}$, $\beta = \inf\{\langle y - \frac{1}{2}\mathbf{1}, z \rangle : y \in \mathbb{Z}^k, \langle y, z \rangle \geq h + P\}$. Since z is rational there are some $y_\alpha, y_\beta \in \mathbb{Z}^k$ such that $\alpha = \langle y_\alpha + \frac{1}{2}\mathbf{1}, z \rangle$, $\beta = \langle y_\beta - \frac{1}{2}\mathbf{1}, z \rangle$. Now $2P = (h + P) - (h - P) \leq \langle y_\beta, z \rangle - \langle y_\alpha, z \rangle = \langle y_\beta - \frac{1}{2}\mathbf{1}, z \rangle - \langle y_\alpha + \frac{1}{2}\mathbf{1}, z \rangle + \langle \mathbf{1}, z \rangle = \beta - \alpha + Q_1 + \dots + Q_k < \beta - \alpha + 2P$ so $\beta - \alpha > 0$.

We distinguish between the two cases $\mathcal{I} = \emptyset$ and $\mathcal{I} \neq \emptyset$. Assume first that $\mathcal{I} \neq \emptyset$, so there exists an integer point y_0 such that $\langle y_0, x \rangle = h$. Then $\langle 2y_0 - y_\beta, z \rangle = 2h - \langle y_\beta, z \rangle \leq h - P$ and $2y_0 - y_\beta \in \mathbb{Z}^k$ so $\langle y_\alpha, z \rangle \geq 2h - \langle y_\beta, z \rangle$. Also $\langle 2y_0 - y_\alpha, z \rangle = 2h - \langle y_\alpha, z \rangle \geq h + P$ so $\langle y_\beta, z \rangle \leq 2h - \langle y_\alpha, z \rangle$. Therefore $h = \frac{1}{2}(\langle y_\beta, z \rangle + \langle y_\alpha, z \rangle) = \frac{1}{2}(\beta + \alpha)$.

Let $H_1 = \{x : \langle x, z \rangle = \alpha\}$, $H_2 = \{x : \langle x, z \rangle = h\}$, $H_3 = \{x : \langle x, z \rangle = \beta\}$. Then $R^+ = \{x : \alpha \leq \langle x, z \rangle \leq h\}$ and $R^- = \{x : h \leq \langle x, z \rangle \leq \beta\}$. The regions R^-, R^+ are nonempty since $\beta - \alpha > 0$ and $h = \frac{1}{2}(\alpha + \beta)$. If $x \in \mathcal{G}^+$ then $\text{vol}(\bar{x} \cap \mathcal{G}^+) > 0$ since $\langle x + \frac{1}{2}\mathbf{1}, z \rangle = \langle y_\alpha + \frac{1}{2}\mathbf{1}, z \rangle + \langle x, z \rangle - \langle y_\alpha, z \rangle > \alpha$ (the last inequality follows because $\langle x, z \rangle > h - P$ and $\langle y_\alpha, z \rangle \leq h - P$) and since $\langle x - \frac{1}{2}\mathbf{1}, z \rangle < \langle x, z \rangle < h$. Moreover $\text{vol}(\bar{x} \cap R^+) > \text{vol}(\bar{x} \cap R^-)$ because $\langle x, z \rangle < h$, by convexity and central symmetry of the k -dimensional cube. Symmetrically, $\text{vol}(\bar{x} \cap R^-) > \text{vol}(\bar{x} \cap R^+)$ for all $x \in \mathcal{G}^-$, so conditions (i) and (ii) hold. Condition (iii) is obviously satisfied since $\mathcal{I} \subseteq H_2$. Finally if $x \notin \mathcal{G}^+ \cup \mathcal{G}^- \cup \mathcal{I}$ then $\langle x, z \rangle \leq h - P$ or $\langle x, z \rangle \geq h + P$; in the former case $\langle x + \frac{1}{2}\mathbf{1}, z \rangle \leq \alpha$ by definition of α so $\text{vol}(\bar{x} \cap (R^+ \cup R^-)) = 0$, and in the latter case $\langle x - \frac{1}{2}\mathbf{1}, z \rangle \geq \beta$ by definition of β , so $\text{vol}(\bar{x} \cap (R^+ \cup R^-)) = 0$, proving (iv).

Assume now that $\mathcal{I} = \emptyset$. Put $h' = (\beta + \alpha)/2$. Note $h' = (\langle y_\beta, z \rangle + \langle y_\alpha, z \rangle)/2$. We first show that if $y \in \mathcal{G}^+$ then $\langle y, z \rangle \leq h'$ and that if $y \in \mathcal{G}^-$ then $\langle y, z \rangle \geq h'$. Take any $y_0 \in \mathbb{Z}^k$ with $\langle y_0, z \rangle \leq h$. Let $y_1 = 2y_0 - y_\beta$. Then $y_1 \in \mathbb{Z}^k$ and $\langle y_1, z \rangle = 2\langle y_0, z \rangle - \langle y_\beta, z \rangle \leq h - P$ so $\langle y_1, z \rangle \leq \langle y_\alpha, z \rangle$. On the other hand $y_0 = (y_\beta + y_1)/2$ so $\langle y_0, z \rangle = (\langle y_\beta, z \rangle + \langle y_1, z \rangle)/2 \leq h'$ since $h' = (\langle y_\beta, z \rangle + \langle y_\alpha, z \rangle)/2$. Therefore $\langle y, z \rangle \leq h \implies \langle y, z \rangle \leq h'$ for all $y \in \mathbb{Z}^k$ and symmetrically $\langle y, z \rangle \geq h \implies \langle y, z \rangle \geq h'$ for all $y \in \mathbb{Z}^k$. In particular, $y \in \mathcal{G}^+ \implies \langle y, z \rangle \leq h'$ and $y \in \mathcal{G}^- \implies \langle y, z \rangle \geq h'$.

If neither $\mathcal{G}^+$ or $\mathcal{G}^-$ have points on the hyperplane $\langle x, z \rangle = h'$ then we may proceed as in the first case after replacing H_2 with the hyperplane $\{x : \langle x, z \rangle = h'\}$, so we may assume that either $\mathcal{G}^+$ or $\mathcal{G}^-$ —say $\mathcal{G}^+$ —has points on the hyperplane $\langle x, z \rangle = h'$. Then $\langle x, z \rangle > h'$ for all $x \in \mathcal{G}^-$ since otherwise we would have $\langle y_a, z \rangle = \langle y_b, z \rangle$ for some $y_a \in \mathcal{G}^+$ and $y_b \in \mathcal{G}^-$, a contradiction. Moreover because z is rational there exists some $\delta > 0$ such that $\langle x, z \rangle > h' + \delta$ for all $x \in \mathcal{G}^-$.

Define $R_\varepsilon^+ = \{x \in \mathbb{R}^k : \alpha + 2\varepsilon \leq \langle x, z \rangle \leq h' + \varepsilon\}$, $R_\varepsilon^- = \{x \in \mathbb{R}^k : h' + \varepsilon \leq \langle x, z \rangle \leq \beta\}$ for any $\varepsilon > 0$. Computing as above, we can see that $\text{vol}(\bar{x} \cap R_0^+) > 0$ for all $x \in \mathcal{G}^+$ and $\text{vol}(\bar{x} \cap R_0^-) > 0$ for all $x \in \mathcal{G}^-$, and that $\text{vol}(\bar{x} \cap (R_0^+ \cup R_0^-)) = 0$ for all $x \notin \mathcal{G}^+ \cup \mathcal{G}^-$. Moreover because z is rational there exists some $\gamma > 0$ such that $x \in \mathcal{G}^+ \implies \text{vol}(\bar{x} \cap R_\varepsilon^+) > 0$, $x \in \mathcal{G}^- \implies \text{vol}(\bar{x} \cap R_\varepsilon^-) > 0$ for all $\varepsilon < \gamma$. Let $\varepsilon_0 > 0$

be smaller than δ and γ . Then $x \in \mathcal{G}^+ \implies \langle x, z \rangle \leq h' \implies \langle x, z \rangle < h' + \varepsilon_0$ and $x \in \mathcal{G}^- \implies \langle x, z \rangle > h + \varepsilon_0$ because $\varepsilon_0 < \delta$, so $\text{vol}(\bar{x} \cap R_{\varepsilon_0}^+) > \text{vol}(\bar{x} \cap R_{\varepsilon_0}^-)$ for all $x \in \mathcal{G}^+$ and $\text{vol}(\bar{x} \cap R_{\varepsilon_0}^+) < \text{vol}(\bar{x} \cap R_{\varepsilon_0}^-)$ for all $x \in \mathcal{G}^-$, fulfilling conditions (i) and (ii) with $R^+ = R_{\varepsilon_0}^+$, $R^- = R_{\varepsilon_0}^-$. Condition (iii) is trivially satisfied since $\mathcal{I} = \emptyset$ and condition (iv) is satisfied because $R_{\varepsilon_0}^+ \cup R_{\varepsilon_0}^- \subseteq R_0^+ \cup R_0^-$. $\square$

Corollary 1. *Theorem 1.*

Acknowledgements. I thank Izabella Laba and Phillip Rogaway for moral and financial support. I would also like to thank the anonymous referee for his or her comments.

References

- [1] N.G. de Bruijn, On the factorization of cyclic groups, *Indag. Math.*, **15** (1953), 370–377
- [2] J.H. Conway and A.J. Jones, Trigonometric diophantine equations (On vanishing sums of roots of unity), *Acta Arithmetica*, **30** (1976), 229–240
- [3] D. Coppersmith and J.P. Steinberger, On the entry sum of cyclotomic arrays, *INTEGERS: the Electronic Journal of Combinatorial and Additive Number Theory*, **6** (2006)
- [4] K. Fukuda, CDD. Available at http://www.ifor.math.ethz.ch/~fukuda/cdd_home.
- [5] T.Y. Lam and K.H. Leung, On vanishing sums of roots of unity, *J. Algebra*, **224** no. 1 (2000), 91–109
- [6] Henry B. Mann, On linear relations between roots of unity, *Mathematika*, **12** Part 2 no. 24 (1965), 107–117
- [7] B. Poonen and M. Rubinstein, The number of intersection points made by the diagonals of a regular polygon, *SIAM J. on Disc. Math.*, **11** no. 1 (1998), 133–156
- [8] L. Rédei, Über das Kresiteilungspolynom, *Acta Math. Acad. Sci. Hungar.* **5** (1954), 27–28
- [9] A. Schrijver “Theory of Linear and Integer Programming”, Wiley-Interscience, 1986
- [10] J.P. Steinberger, Minimal vanishing sums of roots of unity with large coefficients, *Proceedings of the London Mathematical Society*, **97** no. 3 (2008), 689–717.
- [11] I.J. Schoenberg, A note on the cyclotomic polynomial, *Mathematika*, **11** (1964), 131–136