Kissing Polytopes in Dimension 3

Antoine Deza, Zhongyuan Liu & Lionel Pournin

Taylor & Francis
Taylor & Francis Group

Check for updates

Kissing Polytopes in Dimension 3

Antoine Deza[a], Zhongyuan Liu[a], and Lionel Pournin[b]

[a]McMaster University, Hamilton, Ontario, Canada; [b]Université Paris 13, Villetaneuse, France

ABSTRACT

It is shown that the smallest possible distance between two disjoint lattice polytopes contained in $[0, k]^3$ is

$$\frac{1}{\sqrt{2(2k^2 - 4k + 5)(2k^2 - 2k + 1)}}$$

for every integer k at least 4. The proof relies on modeling this as a minimization problem over a subset of the lattice points in the hypercube $[-k, k]^9$. A characterization of this subset allows to reduce the problem to computing the roots of a finite number of degree at most 4 polynomials, using symbolic computation.

KEYWORDS

lattice polytopes; facial distance; integer optimization; distances in lattices

1. Introduction

A polytope is the convex hull of finitely many points from $\mathbb{R}^d$ and in the special case when these points form a subset of $\mathbb{Z}^d$, this object is referred to as a *lattice polytope* or sometimes as an integral polytope. Lattice polytopes appear in areas as diverse as optimization [3, 9, 11, 12, 15], combinatorics [2, 5, 8, 10, 16, 19], or algebraic topology [17, 21–23]. They are integer multiples of rational polytopes and as such, they can serve as convex body approximations for computational procedures [6]. The stopping criterion for certain such procedures depends on how close two such disjoint lattice polytopes can be, as for example in von Neumann's alternating projections algorithm that decides whether two convex bodies are disjoint [7, 27]. Without any additional constraint, the distance between two disjoint lattice polytopes P and Q can be arbitrarily small but such constraints arise in practice. One can ask for example for how close can two disjoint lattice polytopes be, provided that the combined size of their binary encoding is bounded by a constant [7, 14, 25].

A similar, combinatorial constraint is to require that P and Q are contained in the hypercube $[0, k]^d$ where k is a fixed positive integer. Throughout the article, we will refer to such polytopes as *lattice (d, k)-polytopes*. Since, there is only finitely many pairs of disjoint lattice (d, k)-polytopes, the smallest possible distance $\varepsilon(d, k)$ is well defined and one can ask for its value. We call *kissing polytopes* two lattice (d, k)-polytopes whose distance is exactly $\varepsilon(d, k)$: even though they do not touch, they cannot get any closer. Lower and upper bounds on $\varepsilon(d, k)$ that are almost matching as d goes to infinity have been given in [14] and a formula for $\varepsilon(2, k)$ in [13] along with the exact value of $\varepsilon(d, k)$ when d and k are sufficiently small for the computations to be tractable.

One may alternatively ask, also under relevant constraints, for how flat a lattice polytope can be or for how close one of its faces can be from its other vertices. These questions arise from continuous optimization or combinatorial problems [1, 4, 18, 20, 24].

We extend the ideas and techniques from [13] to the 3-dimensional situation and establish the following formula for $\varepsilon(3, k)$.

Theorem 1.1. *If k is not equal to 3, then*

$$\varepsilon(3, k) = \frac{1}{\sqrt{2(2k^2 - 4k + 5)(2k^2 - 2k + 1)}}. \tag{1}$$

All the known values of $\varepsilon(d, k)$ are reported in Table 1 and one can see that (1) does not hold when k is equal to 3. In the case when k is at least 6 we will also show that, up to the symmetries of the cube, $\varepsilon(3, k)$ is uniquely achieved by the pair $P^\star$ and $Q^\star$ of line segments such that the vertices of $P^\star$ are the lattice points $(k, 2, 1)$ and $(0, k-1, k)$ while $Q^\star$ has for its extremities, the origin of $\mathbb{R}^3$ and the lattice point $(k-1, k, k)$. These two line segments are depicted in Figure 1 when k is equal to 2 and when k is at least 4. The figure also shows pairs of line segments that achieve $\varepsilon(3, 1)$ and $\varepsilon(3, 3)$.

Table 1. The known values of $1/\varepsilon(d,k)$.

d	k			
	1	2	3	$k \geq 4$
2	$\sqrt{2}$	$\sqrt{5}$	$\sqrt{13}$	$\sqrt{(k-1)^2 + k^2}$
3	$\sqrt{6}$	$5\sqrt{2}$	$\sqrt{299}$	$\sqrt{2(2k^2 - 4k + 5)(2k^2 - 2k + 1)}$
4	$3\sqrt{2}$	$2\sqrt{113}$	$11\sqrt{71}$	
5	$\sqrt{58}$			
6	$\sqrt{202}$			

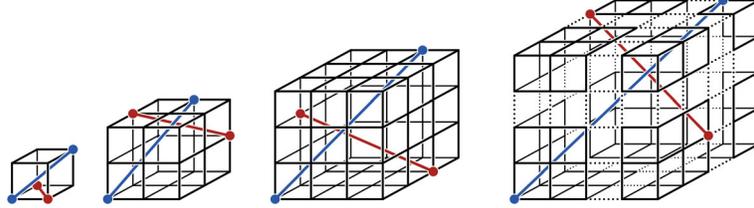The formula shown in bold is provided by Theorem 1.1.

Figure 1. Pairs of lattice $(3,k)$-segments that achieve $\varepsilon(3,k)$ for k equal to 1, 2, 3, and at least 4 (from left to right).

It is shown in [14] that $\varepsilon(d,k)$ is always achieved as the distance between two lattice (d,k)-simplices whose dimensions sum to $d-1$. In the 2-dimensional case, it therefore suffices to consider a point and a line segment. In the 3-dimensional case however, we need to consider both the distance between a point and a triangle and the distance between two line segments. As an intermediate step to proving Theorem 1.1, we will show that the former case can be ignored.

Theorem 1.2. *Consider a lattice point P contained in $[0,k]^3$ and a lattice $(3,k)$-triangle Q. If P is not contained in Q, then*

$$d(P,Q) > \varepsilon(3,k).$$

As a preliminary to the proof of these theorems, we describe in Section 2 the optimization model from [13] that allows to provide lower bounds on $\varepsilon(d,k)$. We then use that model in Section 3 to prove Theorem 1.2 and in Section 4 to prove Theorem 1.1 by reducing the problem to studying a finite number of polynomial expressions using symbolic computation.

2. A discrete optimization model

Consider two lattice (d,k)-simplices P and Q whose dimensions sum to $d-1$. In [13], the distance between P and Q is lower bounded as follows. Denote by p^0 to p^n the vertices of P and by q^0 to q^m the vertices of Q. Consider the $d \times (d-1)$ matrix A whose jth column is $p^j - p^0$ when j is at most n and $q^{j-n} - q^0$ otherwise. Further consider the vector b equal to $q^0 - p^0$. It is shown in [13] (see Lemma 2 therein) that if $A^t A$ is nonsingular, then

$$d\big(\text{aff}(P), \text{aff}(Q)\big) = \|A(A^t A)^{-1} A^t b - b\| \tag{2}$$

where $\text{aff}(P)$ and $\text{aff}(Q)$ denote the affine hulls of P and Q. Hence the distance of P and Q is at least the right-hand side of (2).

Remark 2.1. Observe that A and b are by no means canonical since exchanging P and Q or relabeling their vertices will result in a different matrix A and a different vector b. One can also consider P and Q up to the symmetries of the hypercube, which amounts to permuting or negating a subset of the rows of A and performing the same transformation on the coordinates of b. Any such operation does not change (2) or the determinant of $A^t A$. As remarked in [13] the same is true when a subset of the columns of A is negated.

This construction allows to provide a lower bound on $\varepsilon(d,k)$. Indeed, building on results from [14] it is shown in [13] (see Lemma 2 and Proposition 4 therein) that P and Q can be chosen in such a way that the distance between these simplices is precisely $\varepsilon(d,k)$ while $A^t A$ is nonsingular for any matrix A built from P and Q. It is then immediate that $\varepsilon(d,k)$ is at least the right-hand side of (2) when A and b correspond to such a pair of lattice simplices.

Let us now assume from now on that d is equal to 3. In that case, A is a 3×2 matrix with integer coefficients and b is a vector from $\mathbb{Z}^3$. Moreover, the absolute value of the coefficients of A and b is at most k because they are differences of two nonnegative numbers at most k. Now recall that P and Q are simplices whose dimensions sum to 2. These two simplices are therefore either two line segments or a point and a triangle. In the latter case, we can assume without loss of generality that P is the point while Q is the triangle and by the above construction, the first column of A is $q^1 - q^0$ and its second column is $q^2 - q^0$. In the former case, the first column of A is $p^1 - p^0$ and the second one is $q^1 - q^0$. In both cases, we turn the matrix A and the vector b into the lattice point x contained in the

hypercube $[-k, k]^9$ whose coordinates x_1 to x_6 are obtained by identification from the coefficients of A as

$$\begin{bmatrix} x_1 & x_4 \\ x_2 & x_5 \\ x_3 & x_6 \end{bmatrix} = A \qquad (3)$$

and whose coordinates x_7 to x_9 from those of b as

$$\begin{bmatrix} x_7 \\ x_8 \\ x_9 \end{bmatrix} = b. \qquad (4)$$

In the sequel, we denote by $\mathcal{X}(k)$ the set of all the lattice points x contained in $[-k, k]^9$ that can be obtained as we have just described from a pair of disjoint lattice $(3, k)$-simplices P and Q whose dimensions sum to 2.

Let us now consider an arbitrary 3×2 matrix A and an arbitrary vector b in $\mathbb{R}^3$, both with integer coefficients of absolute value at most k. Here, we no longer assume that A and b are obtained from a pair of lattice polytopes but can still associate to them a lattice point x from $[-k, k]^9$ via (3) and (4). In that case, the determinant of $A^t A$ is equal to $g(x)$ where

$$g(x) = (x_1 x_5 - x_2 x_4)^2 + (x_1 x_6 - x_3 x_4)^2 + (x_2 x_6 - x_3 x_5)^2. \qquad (5)$$

Note that this expression for the determinant of $A^t A$ can be recovered from the Cauchy–Binet formula [26, Example 10.31]. It is observed in [13] that, when $A^t A$ is nonsingular or equivalently when $g(x)$ is not equal to 0, the right-hand side of (2) is

$$\|A(A^t A)^{-1} A^t b - b\| = \frac{|f(x)|}{\sqrt{g(x)}} \qquad (6)$$

where f is the function of x defined as

$$f(x) = x_1(x_6 x_8 - x_5 x_9) + x_2(x_4 x_9 - x_6 x_7) + x_3(x_5 x_7 - x_4 x_8). \qquad (7)$$

The above mentioned results of [13, 14] can be collected into the following theorem in the case when d is equal to 3.

Theorem 2.2. *For every positive integer k, there exists a lattice point x in $\mathcal{X}(k)$ such that $f(x)$ is nonzero, $g(x)$ is positive, and*

$$\varepsilon(3, k) \geq \frac{|f(x)|}{\sqrt{g(x)}}.$$

Let us conclude the section with an upper bound on $\varepsilon(3, k)$ obtained from the two line segments $P^\star$ and $Q^\star$ described in the introduction. They depend on k and can be defined for any integer k at least 2. It is observed in [13] that, for any such k,

$$d(P^\star, Q^\star) = \frac{1}{\sqrt{2(2k^2 - 4k + 5)(2k^2 - 2k + 1)}}$$

and since $P^\star$ and $Q^\star$ are lattice $(3, k)$-polytopes,

$$\varepsilon(3, k) \leq \frac{1}{\sqrt{2(2k^2 - 4k + 5)(2k^2 - 2k + 1)}}. \qquad (8)$$

Note that when k is at most 8, the values of $\varepsilon(3, k)$ are reported in [13]. In particular, $\varepsilon(3, 1)$ is $1/\sqrt{6}$ and therefore, (8) still holds (with equality) when k is equal to 1. For this reason we shall use this inequality for all positive k in the sequel. The values of $\varepsilon(3, k)$ reported in [13] further show that this bound is always sharp when k is at most 8 but different from 3. We shall see that this upper bound is in fact also sharp for every integer k greater than 8.

3. The case of a point and a triangle

The goal of this section is to prove that the distance between a lattice point P contained in $[0, k]^3$ and a lattice $(3, k)$-triangle Q that does not contain this point is greater than $\varepsilon(3, k)$ when k is at least 8. It follows in particular that, for any such value of k, it suffices to consider pairs of lattice $(3, k)$-segments in order to compute $\varepsilon(3, k)$. We will consider two separate cases depending on whether P belongs to the affine hull of Q or not. Indeed, when it does not, it follows from [13, Lemma 2] and [13, Proposition 4] that either the distance between P and Q is greater than $\varepsilon(3, k)$ or for any matrix A and vector b obtained from them as explained in Section 2, $A^t A$ is nonsingular. In the latter case, we can use (2) and (6) in order to lower bound the distance of P and Q, where x is the lattice point in $[-k, k]^9$ obtained from A and b via (3) and (4). The proof will make use of the following straightforward statement, established by computing the roots of $5k^4 - 24k^3 + 40k^2 - 28k + 10$.

Proposition 3.1. *For every positive integer k,*

$$\frac{1}{\sqrt{3k^2}} > \frac{1}{\sqrt{2(2k^2 - 4k + 5)(2k^2 - 2k + 1)}}.$$

We first prove the following.

Lemma 3.2. *Consider a lattice point P contained in the cube $[0, k]^3$ and a lattice $(3, k)$-triangle Q. If P is not contained in the affine hull of Q, then the distance between P and Q is greater than $\varepsilon(3, k)$.*

Proof. Denote by q^0, q^1, and q^2 the vertices of Q. Up to the symmetries of the cube $[0, k]^3$, we can assume that all the coordinates of $q^2 - q^0$ are nonnegative. We will also consider the matrix A and the vector b obtained from P, q^0, q^1, and q^2 as explained in Section 2: the first column of A is $q^1 - q^0$ and its second column $q^2 - q^0$ while b is equal to $q^0 - P$. Assume that P does not belong to the affine hull of Q and, for contradiction, that the distance between P and Q is equal to $\varepsilon(3, k)$. In that case, according to [13, Lemma 2] and [13, Proposition 4] the matrix $A^t A$ is nonsingular. Equivalently, $g(x)$ is positive where x is the lattice point in $[-k, k]^9$ whose coordinates are given by

$$x_i = \begin{cases} q_i^1 - q_i^0 & \text{if } 1 \leq i \leq 3, \\ q_{i-3}^2 - q_{i-3}^0 & \text{if } 4 \leq i \leq 6, \\ q_{i-6}^0 - P_{i-6} & \text{if } 7 \leq i \leq 9. \end{cases}$$

By (2) and (6), the absolute value of $f(x)$ is at least 1 because P does not belong to the affine hull of Q and we obtain

$$d(P, Q) \geq \frac{1}{\sqrt{g(x)}}.$$

According to (5), $g(x)$ is a sum of three squares, each of the form

$$(x_i x_{j+3} - x_j x_{i+3})^2$$

where i and j are distinct and at most 3. We will prove that any such square is at most k^4 and therefore that $g(x)$ is at most $3k^4$.

Since all the coordinates of the vector $q^2 - q^0$ are nonnegative and at most k, so are x_{i+3} and x_{j+3}. We consider different cases depending on the signs of x_i and x_j. First if x_i and x_j are both nonnegative or both non-positive, then the products $x_i x_{j+3}$ and $x_j x_{i+3}$ cannot have opposite signs. As a consequence,

$$(x_i x_{j+3} - x_j x_{i+3})^2 \leq \max\{|x_i x_{j+3}|, |x_j x_{i+3}|\}^2$$

and since x belongs to $[-k, k]^9$, the considered square is at most k^4. Now if x_i is positive and x_j is negative, then

$$(x_i x_{j+3} - x_j x_{i+3})^2 \leq (x_{j+3} - x_j)^2 \max\{x_i, x_{i+3}\}^2. \tag{9}$$

However, recall that x_j is equal to the difference $q_j^1 - q_j^0$ and x_{j+3} to the difference $q_j^2 - q_j^0$. As a consequence, $x_{j+3} - x_j$ is equal to $q_j^2 - q_j^1$ and it follows that the absolute value of this difference is at most k. Hence, according to (9) the considered square is at most k^4 again. Finally if x_i is negative and x_j is positive then

$$(x_i x_{j+3} - x_j x_{i+3})^2 \leq (x_i - x_{i+3})^2 \max\{x_j, x_{j+3}\}^2$$

and the same argument (where i and j are exchanged) proves that the considered square is at most k^4 as well. It follows that

$$d(P, Q) \geq \frac{1}{\sqrt{3k^2}}.$$

By (8) and Proposition 3.1, the distance between P and Q is then greater than $\varepsilon(3, k)$, contradicting our assumption. □

There remains to treat the case when P is contained in the affine hull of Q. This is a consequence of [14, Theorem 5.1] that states that $\varepsilon(d, k)$ is a decreasing function of d for every fixed k. In particular, $\varepsilon(2, k)$ is greater than $\varepsilon(3, k)$.

Lemma 3.3. *Consider a lattice point P contained in the cube $[0, k]^3$ and a lattice $(3, k)$-triangle Q. If P is contained in the affine hull of Q, then the distance between P and Q is greater than $\varepsilon(3, k)$.*

Proof. Consider a nonzero normal vector a to the affine hull of Q. This vector has at least one nonzero coordinate and, up to permuting the coordinates of $\mathbb{R}^3$, it can be assumed without loss of generality that this coordinate is the third one. Let us identify $\mathbb{R}^2$ with the plane spanned by the first two coordinates of $\mathbb{R}^3$. Since the third coordinate of a is nonzero, the orthogonal projection $\pi : \text{aff}(Q) \to \mathbb{R}^2$ is a bijection. In addition π sends a lattice $(3, k)$-polytope contained in the affine hull of Q to a lattice $(2, k)$-polytope. Hence, $\pi(P)$ and $\pi(Q)$ are two disjoint lattice $(2, k)$-polytopes and

$$d\big(\pi(P), \pi(Q)\big) \geq \varepsilon(2, k). \tag{10}$$

As the distance of two points cannot be less than the distance of their images by π, it folows from (10) that,

$$d(P, Q) \geq \varepsilon(2, k).$$

According to [14, Theorem 5.1], $\varepsilon(2, k)$ is greater than $\varepsilon(3, k)$ and so is the distance between P and Q. □

Observe that Theorem 1.2 is an immediate consequence of Lemmas 3.2 and 3.3. It follows in particular from Theorem 1.2 that $\varepsilon(3, k)$ is necessarily achieved as the distance between two lattice $(3, k)$-segments.

4. The case of two line segments

We have shown in Section 2 that $\varepsilon(3, k)$ can only be achieved as the distance between two lattice $(3, k)$-segments. This allows us to refine Theorem 2.2 into the following statement where $\mathcal{Y}(k)$ denotes the subset of the lattice points x in the hypercube $[-k, k]^9$ such that x_1 is non-positive, x_2 to x_6 are nonnegative, and for every integer i satisfying $1 \leq i \leq 3$,

$$\begin{cases} |x_i - x_{i+6}| \leq k, \\ |x_{i+3} + x_{i+6}| \leq k, \\ |x_i - x_{i+3} - x_{i+6}| \leq k. \end{cases} \tag{11}$$

As will be apparent from the proof of that statement, all the points x in $\mathcal{X}(k)$ that correspond to a pair P and Q of lattice $(3, k)$-segments via the construction of Section 2 satisfy (11). It can be proven that the converse is also true, all the points x satisfying (11) correspond to such a pair P and Q, but we will not make use of that property in the sequel.

Theorem 4.1. *For every positive integer k, there exists a lattice point x in $\mathcal{X}(k) \cap \mathcal{Y}(k)$ such that $f(x)$ is nonzero, $g(x)$ is positive, and*

$$\varepsilon(3, k) \geq \frac{|f(x)|}{\sqrt{g(x)}}. \tag{12}$$

Proof. According to [14, Theorem 5.2], there exists two lattice $(3, k)$-polytopes P and Q whose dimensions sum to 2, whose affine hulls are disjoint, and whose distance is equal to $\varepsilon(3, k)$. It follows from Theorem 1.2 that both P and Q are line segments. Denote by p^0 and p^1 the vertices of P and by q^0 and q^1 those of Q. Up to the symmetries of the cube $[0, k]^3$, we can assume that the coordinates of $q^1 - q^0$ are all nonnegative: if $q_i^1 - q_i^0$ is negative, it suffices to replace P and Q by their symmetric with respect to the plane

$$\left\{ x \in \mathbb{R}^2 : x_i = \frac{k}{2} \right\}.$$

We can also assume that $p^1 - p^0$ has at least two nonnegative coordinates by exchanging p^0 and p^1 if needed and that

$$p_1^1 - p_1^0 \leq p_2^1 - p_2^0 \leq p_3^1 - p_3^0$$

by permuting the coordinates of $\mathbb{R}^3$ appropriately. Note that any such permutation does not disturb the non-negativity of the coordinates of $q^1 - q^0$. Consider the matrix A whose first column is $p^1 - p^0$ and whose second column is $q^1 - q^0$ and denote by b the vector $q^0 - p^0$. Further denote by x the lattice point in $\mathcal{X}(k)$ obtained from A and b via (3) and (4). As the distance of P and Q is $\varepsilon(3, k)$ and their affine hulls are disjoint, it follows from [13, Lemma 2] and [13, Proposition 4] that $A^t A$ is nonsingular. Since $g(x)$ is the determinant of $A^t A$, it must be nonzero.

According to (2) and (6), $f(x)$ is nonzero as well because the affine hulls of P and Q are disjoint. Moreover,

$$\varepsilon(3, k) \geq \frac{|f(x)|}{\sqrt{g(x)}}. \tag{13}$$

Recall that the last two coordinates of the vector $p^1 - p^0$ and all the coordinates of the vector $q^1 - q^0$ are nonnegative or, equivalently, that x_2 to x_6 are nonnegative. We can show using an argument similar to that in the proof of Lemma 3.2 that x_1 must be negative. Indeed, assume for contradiction that x_1 is nonnegative. In that case, the right-hand side of (5) is a sum of three squares, each of a difference of two nonnegative terms. As a consequence,

$$g(x) \leq \max\{x_1 x_5, x_2 x_4\}^2 + \max\{x_1 x_6, x_2 x_4\}^2 + \max\{x_2 x_6, x_3 x_5\}^2.$$

However, as x belongs to $[-k,k]^9$, it follows that $g(x)$ is at most $3k^4$ and since $f(x)$ is a nonzero integer, this and (13) imply that $\varepsilon(3,k)$ is at least $1/(\sqrt{3}k^2)$. Together with (8) and Proposition 3.1, this results in a contradiction. There remains to show that the point x satisfies (11) for every integer i such that $1 \le i \le 3$. Consider such an integer i and recall that

$$\begin{cases} x_i = p_i^1 - p_i^0, \\ x_{i+3} = q_i^1 - q_i^0, \\ x_{i+6} = q_i^0 - p_i^0. \end{cases}$$

As an immediate consequence,

$$\begin{cases} x_i - x_{i+6} = p_i^1 - q_i^0, \\ x_{i+3} + x_{i+6} = q_i^1 - p_i^0, \\ x_i - x_{i+3} - x_{i+6} = p_i^1 - q_i^1, \end{cases}$$

and since p^0, p^1, q^0, and q^1 belong to $[0,k]^3$, this shows that x satisfies (11). $\qquad\square$

Now denote by $\mathcal{Z}(k)$ the set of the points x in the hypercube $[-k,k]^9$ (but not necessarily in $\mathbb{Z}^9$) such that x_1 is non-positive, x_2 to x_6 are nonnegative, and for every integer i satisfying $1 \le i \le 3$, the absolute values of x_i and x_{i+3} cannot both be equal to k. We can prove the following property for all the points contained in $\mathcal{Y}(k)$ provided that k is large enough.

Theorem 4.2. *then a point x in $\mathcal{B}$ satisfies $f(x)$ is nonzero, $g(x)$ is positive. If k is at least 5 and x satisfies (12), then the absolute value of $f(x)$ is equal to 1 and x belongs to $\mathcal{Z}(k)$.*

Proof. Assume that k is at least 5. We begin by showing that the absolute value of $f(x)$ is equal to 1. Assume for contradiction that this is not the case. As $f(x)$ is a nonzero integer, its absolute value must then be at least 2. However, the absolute values of the coordinates of x are all at most k and it follows from (5) that $g(x)$ is at most $12k^4$. Therefore, (12) implies

$$\frac{|f(x)|}{\sqrt{g(x)}} \ge \frac{1}{\sqrt{3}k^2}$$

which together with (12) and Proposition 3.1 contradicts (8).

Now consider an integer i such that $1 \le i \le 3$. We will show that $|x_i|$ and $|x_{i+3}|$ cannot both be equal to k. Assume, again for contradiction, that these two absolute values are equal to k and let us first show that x_{i+6} must be an integer multiple of k as well. Since x belongs to $\mathcal{Y}(k)$ its first coordinate is non-positive and the next five are nonnegative. Hence, if i is equal to 1, then x_i is equal to $-k$ and x_{i+3} to k. Therefore, according to the third inequality from (11), $2k+x_{i+6}$ is at most k and it follows that x_{i+6} is necessarily equal to $-k$. If however, i is equal to 2 or 3, then x_i and x_{i+3} are both equal to k. In that case, by the first inequality from (11), x_{i+6} cannot be negative and according to the second inequality, it cannot be positive, which shows that it must be equal to 0. As a consequence, x_i, x_{i+3}, and x_{i+6} all are integer multiples of k. However observe that (7) can be rewritten the three following ways:

$$\begin{cases} f(x) = x_1(x_6x_8 - x_5x_9) + x_4(x_2x_9 - x_3x_8) + x_7(x_3x_5 - x_2x_6), \\ f(x) = x_2(x_4x_9 - x_6x_7) + x_5(x_3x_7 - x_1x_9) + x_8(x_1x_6 - x_3x_4), \\ f(x) = x_3(x_5x_7 - x_4x_8) + x_6(x_1x_8 - x_2x_7) + x_9(x_2x_4 - x_1x_5). \end{cases}$$

Since x_i, x_{i+3}, and x_{i+6} are integer multiples of k, the ith equality among these three shows that $f(x)$ is also an integer multiple of k. Since $f(x)$ is not equal to zero and k is not equal to 1, this implies that the absolute value of $f(x)$ cannot be equal to 1 and we reach a contradiction. As a consequence, the absolute values of x_i and x_{i+3} cannot both be equal to k. $\qquad\square$

According to Theorem 4.2, when k is not too small it suffices to consider the points x in $\mathcal{Y}(k)$ such that the absolute value of $f(x)$ is equal to 1 in order to lower bound $\varepsilon(3,k)$, which given the right-hand side of (12) amounts to maximize $g(x)$. Using this, we will further restrict the search space to a set of lattice points that does not depend on k. In particular, we will prove that the search for a lower bound on $\varepsilon(3,k)$ via Theorem 4.1 can be restricted to the points x in $\mathcal{Y}(k)$ such that $h(x)$ is at least $6k-5$ where

$$h(x) = -x_1 + x_2 + x_3 + x_4 + x_5 + x_6.$$

Consider the set $\mathcal{B}$ of the points x in $\mathbb{N}^9$ whose first six coordinates sum to 6, whose last three coordinates are equal to zero, while $x_i + x_{i+3}$ is at least 1 when i satisfies $1 \le i \le 3$. Observe that $\mathcal{B}$ is a finite set. When k at least 6, we can embed $\mathcal{B}$ as a subset of $\mathcal{Z}(k) \cap \mathbb{Z}^9$ by using the affine map $\phi_k : \mathbb{R}^9 \to \mathbb{R}^9$ such that the ith coordinate of $\phi_k(x)$ is given by

$$[\phi_k(x)]_i = \begin{cases} -k + x_i \text{ if } i \text{ is equal to 1 or 7}, \\ k - x_i \text{ if } i \text{ is at least 2 and at most 6}, \\ x_i \text{ if } i \text{ is equal to 8 or 9}. \end{cases}$$

Note in particular that $h(x)$ is equal to $6k-6$ for every point x in $\phi_k(\mathcal{B})$.

Theorem 4.3. *Consider a lattice point z in $\mathcal{Z}(k)$ such that $h(z)$ is at most $6k - 6$. If k is at least 6, then a point x in $\mathcal{B}$ satisfies*

$$g \circ \phi_k(x) \geq g(z).$$

Proof. Assume that k is at least 6. First observe that $g(z)$ and $h(z)$ do not depend on the last three coordinates of z and we can therefore assume without loss of generality that the seventh coordinate of z is $-k$ and that its two last coordinates are both equal to k. The proof is by induction on $h(z)$. By construction, ϕ_k sends $\mathcal{B}$ to the subset of the points in $\mathcal{Z}(k) \cap \mathbb{Z}^9$ whose image by h is $6k - 6$, whose seventh coordinate is $-k$ and whose last two coordinates are equal to k. Hence, if $h(z)$ is equal to $6k - 6$, then z belongs to $\phi_k(\mathcal{B})$ and it suffices to take for x the pre-image of z by ϕ_k. Now assume that $h(z)$ is at most $6k - 7$. Since z is a lattice point in $\mathcal{Z}(k)$, $-z_1 + z_4$ is at most $2k - 1$. However, for every point x in $\mathcal{Z}(k)$,

$$\frac{\partial g}{\partial x_1}(x) = 2x_1(x_5^2 + x_6^2) - 2x_4(x_2x_5 + x_3x_6)$$

is always non-positive and

$$\frac{\partial g}{\partial x_4}(z) = 2x_4(x_2^2 + x_3^2) - 2x_1(x_2x_5 + x_3x_6)$$

is always nonnegative. If $-z_1 + z_4$ is less than $2k - 1$, decreasing by 1 the first coordinate of z or increasing by 1 its fourth coordinate results in a point z' in $\mathcal{Z}(k)$. By the sign of the above derivatives on the line segment with extremities z and z', the image of z' by g is at least $g(z)$. Moreover, $h(z')$ is greater than $h(z)$ by 1. As a consequence, the result follows by induction. Now assume that $-z_1 + z_4$ is equal to $2k - 1$ and recall that $h(z)$ is less than $6k - 6$. Therefore, $z_2, z_3, z_5,$ and z_6 cannot all be at least $k - 1$. Assume that z_2 is less than $k - 1$. Increasing by 1 the second coordinate of z results in a lattice point z' that is still contained in $\mathcal{Z}(k)$. However,

$$\frac{\partial g}{\partial x_2}(x) = 2x_2(x_4^2 + x_6^2) - 2x_5(x_1x_4 + x_3x_6)$$

for every point x in $\mathcal{Z}(k)$. As $-z_1 + z_4$ is equal to $2k - 1$, every point x that belongs to the line segment with extremities z and z' is such that either x_1 is equal to $1 - k$ and x_4 to k or x_1 is equal to $-k$ and x_4 to $k - 1$. In particular, the product x_1x_4 is necessarily equal to $-k(k - 1)$. However, the product x_3x_6 is at most $k(k - 1)$ on that line segment and the above partial derivative is therefore nonnegative. It follows that $g(z')$ is at least $g(z)$ and since $h(z')$ is greater than $h(z)$ by 1, the result follows by induction.

Given that, for every point x in $\mathcal{Z}(k)$,

$$\frac{\partial g}{\partial x_3}(x) = 2x_3(x_4^2 + x_5^2) - 2x_6(x_1x_4 + x_2x_5),$$

$$\frac{\partial g}{\partial x_5}(x) = 2x_5(x_1^2 + x_3^2) - 2x_2(x_1x_4 + x_3x_6), \text{ and}$$

$$\frac{\partial g}{\partial x_6}(x) = 2x_6(x_1^2 + x_2^2) - 2x_3(x_1x_4 + x_2x_5),$$

the same argument by induction shows that the result also holds when $-z_1 + z_4$ is equal to $2k - 1$ and $z_3, z_5,$ or z_6 is less than $k - 1$. $\square$

Observe that for each point x in $\mathcal{B}$, $g \circ \phi_k(x)$ is a polynomial function of k of degree at most 4. Hence, according to Theorem 4.3, maximizing $g(x)$ over the points in $\mathcal{Z}(k)$ whose image by h is at most $6k - 6$ amounts to compare the values in k of a fixed number (that does not depend on k) of degree at most 4 polynomials. Observe that the square of the denominator of the right-hand side of (8) is the degree 4 polynomial $8k^4 - 24k^3 + 40k^2 - 28k + 10$. It turns out that this polynomial is always greater than $g \circ \phi_k(x)$ when k is at least 6. This can be checked using symbolic computation. Indeed, the roots of

$$8k^4 - 24k^3 + 40k^2 - 28k + 10 - g \circ \phi_k(x) \tag{14}$$

can be explicitly determined for each point x in $\mathcal{B}$ as well as its sign when k is equal to 6. The computations show in particular that the largest real root of (14) when x ranges over $\mathcal{B}$ is less than 6 and that this polynomial is always positive when k is equal to 6. These computational checks can be collected into the the following proposition.

Proposition 4.4. *For every integer k at least 6 and every point x in $\mathcal{B}$,*

$$g \circ \phi_k(x) < 8k^4 - 24k^3 + 40k^2 - 28k + 10.$$

Table 2. The eight lattice points x in $\mathcal{A}$ such that for some integer k at least 6 both the assertion (i) and the assertion (ii) in the statement of Proposition 4.6 hold.

$(0,1,3,1,0,0,0,-1,-2)$	$(0,1,3,1,0,0,0,1,0,-1)$
$(0,3,1,1,0,0,0,-2,-1)$	$(0,3,1,1,0,0,0,1,-1,0)$
$(1,0,0,0,1,3,0,1,2)$	$(1,0,0,0,1,3,1,0,1)$
$(1,0,0,0,3,1,0,2,1)$	$(1,0,0,0,3,1,1,1,0)$

Denote by $\mathcal{A}$ the set of the points x in $\mathbb{N}^7 \times \mathbb{Z}^2$ whose first six coordinates sum to at most 5, whose last three coordinates satisfy

$$\begin{cases} x_7 \leq x_1 + x_4, \\ -x_2 \leq x_8 \leq x_5, \\ -x_3 \leq x_9 \leq x_6, \end{cases} \tag{15}$$

and such that $x_i + x_{i+3}$ is at least 1 when $1 \leq i \leq 3$. Again, $\mathcal{A}$ is a finite set. Moreover, $\phi_k(\mathcal{A})$ contains all the points x in $\mathcal{Y}(k) \cap \mathcal{Z}(k)$ such that $h(x)$ is at least $6k - 5$. Note in particular that for any point x in $\mathbb{N}^7 \times \mathbb{Z}^2$, if the coordinates of $\phi_k(x)$ satisfy (11), then the coordinates of x satisfy (15). Combining Theorems 4.1–4.3 with Proposition 4.4 makes it possible to provide, when k is at least 6, a lower bound on $\varepsilon(3, k)$ that only depends on $f \circ \phi_k(x)$ and $g \circ \phi_k(x)$ where x ranges over $\mathcal{A}$.

Theorem 4.5. *For every integer k at least 6, there exists a point x in $\mathcal{A}$ such that $f \circ \phi_k(x)$ is equal to 1, $g \circ \phi_k(x)$ is positive, and*

$$\varepsilon(3, k) \geq \frac{1}{\sqrt{g \circ \phi_k(x)}}.$$

Proof. Assume that k is at least 6. According to Theorems 4.1 and 4.2, there exists a point z in $\mathcal{Y}(k) \cap \mathcal{Z}(k)$ such that $|f(z)|$ is equal to 1, $g(z)$ is positive, and $\varepsilon(3, k)$ satisfies

$$\varepsilon(3, k) \geq \frac{1}{\sqrt{g(z)}}. \tag{16}$$

It follows from Theorem 4.3 and Proposition 4.4 that $h(z)$ is at least $6k - 5$. Indeed, otherwise, these two results would imply that $g(z)$ is less than the square of the denominator in the right-hand side of (8). In that case, (16) would contradict (8).

As a consequence, z is a point in $\mathcal{Y}(k) \cap \mathcal{Z}(k)$ such that $h(z)$ is at least $6k - 5$ and therefore, this point is contained in $\phi_k(\mathcal{A})$. Taking, for x the preimage of z by ϕ_k completes the proof. □

For every point x in $\mathcal{A}$, both $f \circ \phi_k(x)$ and $g \circ \phi_k(x)$ are polynomial functions of k, the former being of degree at most 3 and the latter of degree at most 4. By Theorem 4.5, one can obtain a lower bound on $\varepsilon(3, k)$ when k is at least 6 by computing these two polynomials for every point x in $\mathcal{A}$ and by checking whether the first one is equal to 1 or to -1 for certain values of k and, among the points x such that this property holds, to pick the one for which the value in k of the second polynomial is maximal. This requires solving a finite number of polynomial equations of degree at most 4. We then obtain the following by symbolic computation.

Proposition 4.6. *All the points x in $\mathcal{A}$ such that*

(i) $|f \circ \phi_k(x)|$ is equal to 1 and
(ii) $g \circ \phi_k(x)$ is not less than $8k^4 - 24k^3 + 40k^2 - 28k + 10$

for some integer k at least 6 correspond, up to the transformations described in Remark 2.1 to the pair $P^\star$ and $Q^\star$ from Section 2.

There are precisely eight points x in $\mathcal{A}$ that satisfy the assertions (i) and (ii) in the statement of Proposition 4.6 for some integer k at least 6. These points are reported in Table 2 as vectors of coordinates. Since these points all correspond to the segments $P^\star$ and $Q^\star$ up to the transformations described in Remark 2.1, they must satisfy the assertions (i) and (ii) in the statement of Proposition 4.6 for every integer k at least 6 and not just for some of these integers. By this observation, Theorem 1.1 follows from Theorem 4.5 and Proposition 4.6. In particular, according to the values of $\varepsilon(3, 1)$, $\varepsilon(3, 2)$, $\varepsilon(3, 4)$ and $\varepsilon(3, 5)$ reported in [13], the theorem indeed holds when k is equal to 1, 2, 4, or 5 even though these four values are not covered by Theorem 4.5 and Proposition 4.6.

References

[1] Alon, N., Vũ, V. H. (1997). Anti-Hadamard matrices, coin weighing, threshold gates, and indecomposable hypergraphs. *J. Comb. Theory Ser. A* 79: 133–160.
[2] Bárány, I., Vershik, A. M. (1991). On the number of convex lattice polytopes. *Geom. Funct. Anal.* 2: 381–393.
[3] Barbato, M., Grappe, R., Lacroix, M., Lancini, E. (2023). Box-total dual integrality and edge-connectivity. *Math. Program.* 197(1): 307–336.
[4] Beck, A., Shtern, S. (2017). Linearly convergent away-step conditional gradient for non-strongly convex functions. *Math. Program.* 164: 1–27.
[5] Beck, M., Pixton, D. (2003). The Ehrhart polynomial of the Birkhoff polytope. *Discrete Comput. Geom.* 30: 623–637.
[6] Beck, M., Robins, S. (2015). *Computing the Continuous Discretely*. New York: Springer.
[7] Braun, G., Pokutta, S., Weismantel, R. (2022). Alternating linear minimization: revisiting von Neumann's alternating projections. arXiv:2212.02933.

[8] Cardinal, J., Pournin, L. (2024). The expansion of half-integral polytopes. *Michigan Math. J.*, to appear.
[9] Cardinal, J., Steiner, R. (2024). Inapproximability of shortest paths on perfect matching polytopes. *Math. Program.*, to appear.
[10] David, J., Pournin, L., Rakotonarivo, R. (2020). Elementary moves on lattice polytopes. *J. Comb. Theory Ser. A* 172: 105200.
[11] Del Pia, A., Michini, C. (2016). On the diameter of lattice polytopes. *Discrete Comput. Geom.* 55: 681–687.
[12] Del Pia, A., Michini, C. (2022). Short simplex paths in lattice polytopes. *Discrete Comput. Geom.* 67: 503–524.
[13] Deza, A., Liu, Z., Pournin, L. (2025). Small kissing polytopes. *Vietnam J. Math.* 53(4):901–913.
[14] Deza, A., Onn, S., Pokutta, S., Pournin, L. (2024). Kissing polytopes. *SIAM J. Discrete Math.* 38(4): 2643–2664.
[15] Deza, A., Pournin, L. (2018). Improved bounds on the diameter of lattice polytopes. *Acta Mathematica Hungarica* 154(2): 457–469.
[16] Deza, A., Pournin, L. (2022). Primitive point packing. *Mathematika* 68(3): 979–1007.
[17] Fulton, W. (1993). *Introduction to Toric Varieties*. Princeton, NJ: Princeton University Press.
[18] Gutman, D. H., Peña, J. F. (2021). The condition number of a function relative to a set. *Math. Program.* 188: 255–294.
[19] Knauer, K., Martínez-Sandoval, L., Alfonsín, J. L. (2018). On lattice path matroid polytopes: integer points and Ehrhart polynomial. *Discrete Comput. Geom.* 60: 698–719.
[20] Lacoste-Julien, S., Jaggi, M. (2015). On the global linear convergence of Frank–Wolfe optimization variants. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS)*, pp. 496–504.
[21] Laplante-Anfossi, G. (2022). The diagonal of the operahedra. *Adv. Math.* 405: 108494.
[22] Loday, J.-L. (2004). Realization of the Stasheff polytope. *Archiv der Math.* 83(3): 267–278.
[23] Masuda, N., Thomas, H., Tonks, A., Vallette, B. (2021). The diagonal of the associahedra. Journal de l'École Polytechnique 8: 121–146.
[24] Peña, J. F. (2019). Generalized conditional subgradient and generalized mirror descent: duality, convergence, and symmetry. arXiv:1903.00459.
[25] Schrijver, A. (1998). *Theory of Linear and Integer Programming*. New York: Wiley.
[26] Shafarevich, I. R., Remizov, A. O. (2013). *Linear Algebra and Geometry*. Berlin: Springer.
[27] von Neumann, J. (1949). On rings of operators. Reduction theory. *Ann. Math.* 50(2): 401–485.