ISSN 1560-3547, Regular and Chaotic Dynamics, 2007, Vol. 12, No. 6, pp. 736–745. © Pleiades Publishing, Ltd., 2007.

ARTICLES

Integral Medial Axis and the Distance Between Closest Points

W. van der Kallen[1*]

[1] *Mathematics Institute, University of Utrecht*
P.O. Box 80.010, 3508 TA Utrecht, The Netherlands

Received March 29, 2007; accepted July 18, 2007

Abstract—In dimension two we prove an inequality that implies a desirable property of the integral medial axis as defined by Hesselink in [1]. In dimension three we conjecture a similar inequality.

MSC2000 numbers: 11Y65, 94A08, 68U05, 52C05

DOI: 10.1134/S1560354707060147

Key words: integral medial axis, continued fraction

1. THE PROBLEM

Let $\vec{v}$ be a nonzero vector in the plane $\mathbb{R}^2$ and let c be a constant. They determine the half plane

$$H = \{\ \vec{x} \in \mathbb{R}^2 \mid \vec{x} \cdot \vec{v} \geqslant c\ \}.$$

Consider the integer lattice $\mathbb{Z}^2$ in the plane. For $P \in \mathbb{R}^2$ we put

$$\mathcal{FT}(P) = \{X \in H \cap \mathbb{Z}^2 \mid ||P - X|| \text{ is minimal}\}.$$

(As explained in [1], $\mathcal{FT}$ stands for feature transform.)

In Fig. 1 the half plane H is shaded and P, Q are lattice points at distance one from each other. The transform $\mathcal{FT}(P)$ of P in the picture consists of two points, one of which is denoted X. And the transform $\mathcal{FT}(Q)$ of the point Q in the picture consists of a single point denoted Y. Although P, Q have distance one, the distance between their respective "closest points" X and Y is 5. So that distance is much larger for the "pixelized" $H \cap \mathbb{Z}^2$ than it would be for its continuous counterpart H.

Our result is

Theorem 1. *Let $P, Q \in \mathbb{Z}^2$ with $||P - Q|| \leqslant 1$. If $X \in \mathcal{FT}(P)$, $Y \in \mathcal{FT}(Q)$, then*

$$||X - Y||^2 \leqslant 1 + 2(P - Q) \cdot (X - Y) + ||P - X + Q - Y||. \tag{1}$$

This confirms a relevant case of a conjecture of Hesselink with intended applications in image processing [1]. Hesselink proposed to define the integral medial axis (IMA) of a "background" $B \subset \mathbb{Z}^n$ more or less as follows. For $P \in \mathbb{Z}^n$ put

$$\mathcal{FT}(P) = \{X \in B \mid ||P - X|| \text{ is minimal}\}.$$

Let P, $Q \in \mathbb{Z}^n$ with $||P - Q|| = 1$. Choose $X \in \mathcal{FT}(P)$, $Y \in \mathcal{FT}(Q)$, and put the edge $[P, Q]$ in the integral medial axis if

$$||X - Y||^2 > 1 + 2(P - Q) \cdot (X - Y) + ||P - X + Q - Y||. \tag{2}$$

This provides a discrete analogue of a medial axis. Recall that the medial axis of a closed region $R \subset \mathbb{R}^n$ consists of those P for which there is more than one point in R that is closest to P. For reasonable R the integral medial axis of the pixelized version $B = R \cap \mathbb{Z}^n$ should look like a discretization of the medial axis of R.

[*] E-mail: vdkallen@math.uu.nl

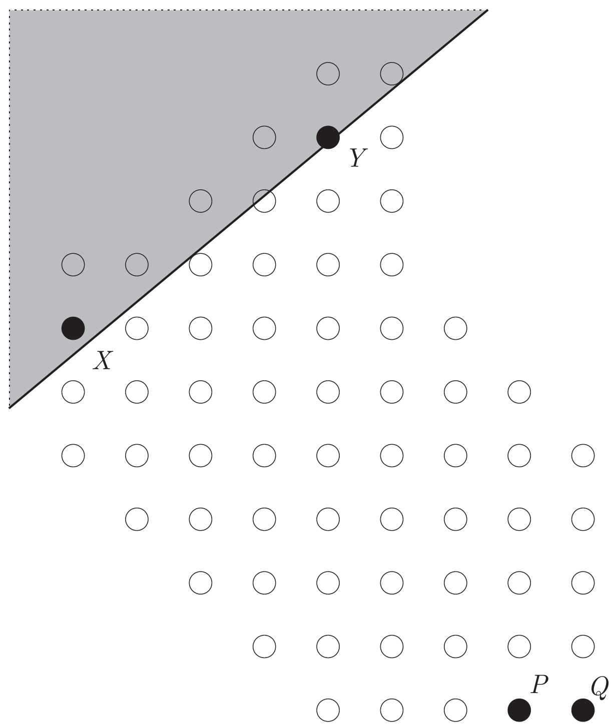

Fig. 1.

Compare inequality (2) with the more naive proposal that would put $[P, Q]$ in the integral medial axis if $||X - Y|| > 1$. This naive proposal leads to annoying artefacts. These are already visible when the background consists of the lattice points in a half space H. In that case one would like the integral medial axis to be empty. Our theorem says that this wish is fulfilled in dimension two if one uses (2).

In dimension three we will see that (2) needs to be modified a little. But in dimension three we only have experimental results. Actual proofs are beyond reach for now. We therefore return to dimension two.

Our method of proof uses the continued fraction algorithm (for the slope of the line through X and Y) to find a point Z in $H \cap \mathbb{Z}^2$, with the angle $\angle XZY$ obtuse, so that the triangle ΔXYZ is "very thin". Here we call a triangle thin if its circumscribed circle is large with respect to the sides of the triangle. Think of X, Y as being given before P, Q. If the triangle is very thin, then linear programming will show that P, Q are forced to positions that make $||P - X + Q - Y||$ large. There remain a few cases where the triangle is not thin enough, but they are few and are easily dealt with, again with linear programming.

We have included many details. Unfortunately this hides the simplicity of the ideas a bit.

1.1. Remark

The condition $||P - Q|| \leqslant 1$ may not be dropped in the theorem: Let m be a positive integer and consider the half space

$$H = \{ (x_1, x_2) \mid -x_1 + 4mx_2 \geqslant 0 \},$$

with $P = (1 + 2m, -2m(1 + m))$, $Q = (4m, 1 - 8m^2)$, $X = (0, 0)$, $Y = (4m, 1)$. Then $X \in \mathcal{FT}(P)$ and $Y \in \mathcal{FT}(Q)$ but

$$||X - Y||^2 - 2(P - Q) \cdot (X - Y) - ||P - X + Q - Y||$$

grows with m like $2m^2$. In particular, it is not bounded.

I thank Johnny Edwards for explaining to me where one must look in a search for counter examples to the version with the condition $||P - Q|| \leqslant 1$ dropped. He also replaces $||P - X + Q - Y||$ with the length of the component of $P - X + Q - Y$ perpendicular to $X - Y$. His analysis showed that in a counterexample with $a := |(X - Y) \cdot (1, 0)|$ at least $b := |(X - Y) \cdot (0, 1)|$, one must have a congruent to -1 modulo b, or $a = b + 1$. By now he has much more precise results for arbitrary $||P - Q||$. See [2].

2. OBSERVATIONS AND SIMPLIFICATIONS

If $X = Y$, then the result (1) is clear. So we further assume

$$X \neq Y. \tag{3}$$

We say that two points are on the *same side* of a given line if they lie off the line and on the same side, but also if at least one of the two points lies on the line. With this language we have that P, X are on the same side of the perpendicular bisector of the segment $[X, Y]$. (Otherwise Y would be closer to P than X.) Similarly Q, Y are on the same side of the perpendicular bisector of $[X, Y]$. Therefore

$$2(P - Q) \cdot (X - Y) \geqslant 0. \tag{4}$$

Suppose $Y = Q$. As $X \neq Y$ we must have $P \neq Q$. So $||P - Q|| = 1$. Now $||P - X|| \leqslant ||P - Y|| = ||P - Q|| = 1$. If $P = X$, then (1) is clear again. In any case, the conditions $Y = Q$, $||P - Q|| = 1$, $||P - X|| \leqslant 1$ leave very few configurations for the lattice points P, Q, X, Y. For each of them (1) is easily checked. The case $P = X$ is similar. So we may further assume

$$Y \neq Q \text{ and } P \neq X. \tag{5}$$

So from now on P, Q are outside H.

If the segment $[X, Y]$ contains another lattice point Z, then we see that

$$2(P - Q) \cdot (X - Y) \geqslant ||X - Y||^2,$$

because P and X are on the same side of the perpendicular bisector of $[X, Z]$ and Q, Y are on the same side of the perpendicular bisector of $[Z, Y]$. So we may further assume

$$[X, Y] \cap \mathbb{Z}^2 \text{ consists of } X \text{ and } Y \text{ only.} \tag{6}$$

In other words, $X - Y$ is a primitive lattice vector.

Lemma 1. *There is a square $\mathcal{S}$ with side $[X, Y]$ that lies inside H. Here we mean by $\mathcal{S}$ the convex set, not just its boundary.*

Proof. Indeed let one square with side $[X, Y]$ have corners X, Y, Y_1, X_1 and let the other have corners X, Y, Y_2, X_2. Both H and its complement are convex. Say Y_1 is in the complement. If Y_2 is also in the complement, then so is the entire segment $[Y_1, Y_2]$. But $[Y_1, Y_2]$ intersects $[X, Y]$, which lies in H, so this is absurd. Similarly X_2 cannot also be in the complement. We conclude that the square XYY_2X_2 lies in H. □

2.1. Standard Position

From now on we forget H and consider the following situation. We are given lattice points P, Q, X, Y and a square $\mathcal{S}$ with side $[X, Y]$ so that no point of $\mathcal{S} \cap \mathbb{Z}^2$ is strictly closer to P than X is, and no point of $\mathcal{S} \cap \mathbb{Z}^2$ is strictly closer to Q than Y is. We are still trying to prove (1) under the assumption $||P - Q|| \leqslant 1$ together with the further simplifying assumptions that have been introduced above.

If α is an isometry of the plane which leaves the lattice invariant, then we may replace P, Q, X, Y and the square $\mathcal{S}$ with their images under α. Using this, one may arrange that $[X, Y]$ has one end at the origin, the other end in the region $\{ (a, b) \mid 0 \leqslant b \leqslant a \}$. And by turning about the midpoint of $[X, Y]$ if necessary, one may also arrange that $\mathcal{S}$ lies above $[X, Y]$. Rename the points, if necessary, so that X is the origin. From now on we assume all this has been done, so that

$$X = (0, 0), \quad Y = (y_1, y_2) \text{ with } 0 \leqslant y_2 \leqslant y_1 \tag{7}$$

and

$$\mathcal{S} \text{ has corners } (0, 0),\ (y_1, y_2),\ (y_1 - y_2, y_1 + y_2),\ (-y_2, y_1). \tag{8}$$

Observe that (6) implies

$$\gcd(y_1, y_2) = 1. \tag{9}$$

3. SPECIAL CASES

We will say that we are in the *generic case* unless $Y = (3, 2)$, or $Y = (4, 3)$, or $y_2 \leqslant 1$. These exceptional cases we now treat separately.

3.1. The case Y=(3,2)

As $(1, 1)$ lies in $\mathcal{S}$, we must have $||P|| = ||P - X|| \leqslant ||P - (1, 1)||$. So $P = (p_1, p_2)$ satisfies

$$(p_1, p_2) \cdot (1, 1) \leqslant \frac{1}{2}(1, 1) \cdot (1, 1) = 1.$$

Similarly $3p_1 + 2p_2 \leqslant \frac{13}{2}$ because $Y \in \mathcal{S}$. But $(p_1, p_2) \cdot (3, 2)$ is an integer, so we get $3p_1 + 2p_2 \leqslant 6$. And for Q we get in the same fashion that $2q_1 + q_2 \geqslant 6$, $3q_1 + 2q_2 \geqslant 7$. Now recall $(P - Q).(X - Y)$ is nonnegative. As $||P - Q|| \leqslant 1$, there are three possibilities: Q equals P or $P + (0, 1)$ or $P + (1, 0)$. Clearly $Q = P$ does not happen.

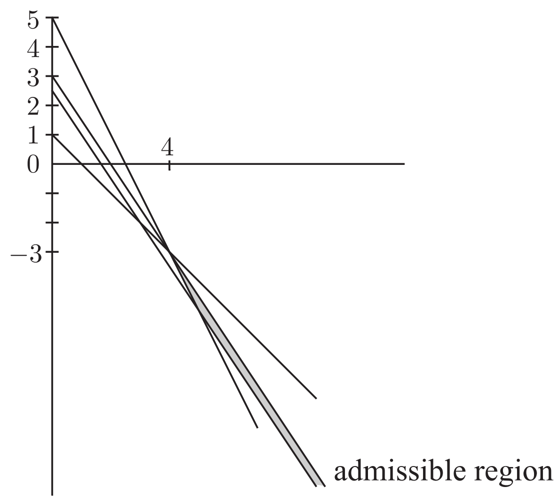

Fig. 2.

Let $Q = P + (0, 1)$. By drawing the picture (see Fig. 2) of the four linear inequalities $p_1 + p_2 \leqslant 1$, $3p_1 + 2p_2 \leqslant 6$, $2p_1 + (p_2 + 1) \geqslant 6$, $3p_1 + 2(p_2 + 1) \geqslant 7$, we get $p_1 \geqslant 4$ and $p_2 \leqslant -3$. We will refer to such reasoning with a picture of linear inequalities as “linear programming in the plane”.

We conclude that the right hand side of (1), which equals $1 + 4 + ||(p_1, p_2) - (0, 0) + (p_1, p_2 + 1) - (3, 2)|| = 5 + \sqrt{(2p_1 - 3)^2 + (2p_2 - 1)^2}$, is at least $5 + \sqrt{5^2 + 7^2}$, and thus at least the left hand side, which is $3^2 + 2^2$. Inequality (1) follows.

Finally if $Q = P + (1, 0)$, then linear programming in the plane yields $p_1 \geqslant 3$ and $p_2 \leqslant -2$. Inequality (1) follows again.

3.2. The case Y=(4,3)

Again $(1, 1)$ lies in $\mathcal{S}$. Arguing as in the previous case we find $p_1 + p_2 \leqslant 1$, $4p_1 + 3p_2 \leqslant 12$, $3q_1 + 2q_2 \geqslant 12$, $4q_1 + 3q_2 \geqslant 13$. Recall there are three possibilities: Q equals P or $P + (0, 1)$ or $P + (1, 0)$. Clearly $Q = P$ does not happen.

Let $Q = P + (0, 1)$. By linear programming in the plane we get $p_1 \geqslant 8$ and $p_2 \leqslant -7$. Therefore the right hand side of (1), which equals $1 + 6 + ||(p_1, p_2) - (0, 0) + (p_1, p_2 + 1) - (4, 3)|| = 7 + \sqrt{(2p_1 - 4)^2 + (2p_2 - 2)^2}$, is at least $7 + \sqrt{12^2 + 16^2}$, and thus at least the left hand side, which is $4^2 + 3^2$. Inequality (1) follows.

Finally if $Q = P + (1, 0)$, then linear programming in the plane yields $p_1 \geqslant 7$ and $p_2 \leqslant -6$. Inequality (1) follows again.

3.3. When the Second Coordinate of Y is at Most One.

If $y_2 = 0$, then $y_1 = 1$ and (1) is obvious. So we may further assume $y_2 = 1$, so that $y_1 \geqslant 1$. Write $m = y_1$. We now exploit as in the previous case that $(m-1, 1)$ lies in $\mathcal{S}$. As before we have three possibilities: Q equals P or $P + (0, 1)$ or $P + (1, 0)$.

We get the inequalities $(m-1)p_1 + p_2 \leqslant (m^2 - 2m + 2)/2$, $mp_1 + p_2 \leqslant (m^2+1)/2$, $q_1 \geqslant m$, $mq_1 + q_2 \geqslant (m^2+1)/2$.

If $Q = P$, then linear programming in the plane gives $p_1 \geqslant m$, $p_2 \leqslant (1-m^2)/2$ and the right hand side of (1) is at least $1 + \sqrt{m^2 + (m^2)^2}$. Clearly this exceeds the left hand side $m^2 + 1$.

If $Q = P + (1, 0)$, then it follows from linear programming in the plane that $p_1 \geqslant m - 1$, $p_2 \leqslant (2m - m^2)/2$ and the right hand side of (1) is at least $1 + 2m + \sqrt{(m-1)^2 + (m-1)^4} \geqslant 1 + 2m + \sqrt{(m-1)^4} = m^2 + 2$. This clearly exceeds the left hand side $m^2 + 1$.

If $Q = P + (0, 1)$, then it follows from linear programming in the plane that $p_1 \geqslant m$ and $p_2 \leqslant -(m^2-1)/2$. Then the right hand side of (1) is at least $1 + 2 + \sqrt{m^2 + (m^2-1)^2}$. This obviously exceeds the left hand side $m^2 + 1$.

4. THE GENERIC CASE

From now on we consider the generic case, with

$$y_2 \geqslant 2, \tag{10}$$

$$Y \neq (3, 2), \quad Y \neq (4, 3). \tag{11}$$

4.1. The Arc Height of a Triangle

Let ΔABC be a triangle with $\angle ACB$ obtuse. Let D be the midpoint of $[A, B]$. The perpendicular bisector of $[A, B]$ and the circumscribed circle of ΔABC intersect in two points. Let E be the intersection point that is closest to D. We call $||D - E||$ the *arc height* of the triangle.

Lemma 2. *The arc height equals*

$$\frac{||A - B|| \, \mathrm{Area}(\Delta ABC)}{||C - A|| \, ||B - C|| + (C - A) \cdot (B - C)}.$$

Proof. Let F be the projection of C onto $[A, B]$. Choose coordinates so that $A = (-1, 0)$, $B = (1, 0)$. Say $C = (a, b)$. Note that $a^2 + b^2 < 1$. Then the center M of the circumscribed circle of triangle ΔABC is $M = (0, \frac{a^2+b^2-1}{2b})$. One checks that

$$2||A - M|| \, ||C - F|| = ||C - A|| \, ||B - C||,$$

$$2||D - M|| \, ||C - F|| = (C - A) \cdot (B - C),$$

$$||C - A||^2 ||B - C||^2 - \Big((C - A) \cdot (B - C)\Big)^2 = ||C - F||^2 ||B - A||^2.$$

So if $h = ||A - M|| - ||D - M||$ is the arc height, then

$$h\Big(||C - A|| \, ||B - C|| + (C - A) \cdot (B - C)\Big)$$

$$= \tfrac{1}{2}||C - F||^{-1}\left(||C - A||^2||B - C||^2 - \Big((C - A) \cdot (B - C)\Big)^2\right)$$

$$= \tfrac{1}{2}||C - F|| \, ||B - A||^2 = ||A - B|| \, \mathrm{Area}(\Delta ABC),$$

whence the result. □

4.2. The Geometric Continued Fraction Algorithm

Recall that $2 \leqslant y_2 \leqslant y_1$ and $\gcd(y_1, y_2) = 1$. So we must actually have

$$2 \leqslant y_2 < y_1. \tag{12}$$

We wish to apply the continued fraction algorithm to y_2/y_1. We now recall the geometric incarnation of the continued fraction algorithm [3], [4, Appendix]. For further explanations see these references and also the next subsection. Put $\vec{v}_0 = (0,1)$ and $\vec{v}_1 = (1,0)$. Given $\vec{v}_i$, $\vec{v}_{i+1}$ with $\vec{v}_{i+1} \neq Y$ one chooses μ to be the smallest integer so that $\vec{v}_i$ and $\vec{v}_i + (\mu+1)\vec{v}_{i+1}$ lie strictly on different sides of the line through $X = (0,0)$ and $Y = (y_1, y_2)$. One has $\mu \geqslant 1$. Put $\vec{v}_{i+2} = \vec{v}_i + \mu\vec{v}_{i+1}$. If $\vec{v}_{i+2} = Y$, stop. Otherwise repeat. All $\vec{v}_i$ are lattice points within the polygon with vertices $(0,0)$, $(0,1)$, Y, $(1,0)$. The process stops. Say it stops with $\vec{v}_{n+2} = Y$. This defines n. For $i \geqslant 1$ the angle between $\vec{v}_i$ and $\vec{v}_{i+1}$ is sharp and $||\vec{v}_{i+1}|| > ||\vec{v}_i||$. Observe that $|\det(\vec{v}_j, \vec{v}_{j+1})| = 1$ for all $0 \leqslant j \leqslant n+1$. In particular, the triangle with vertices X, Y, $\vec{v}_{n+1}$ has area $\frac{1}{2}$.

4.3. Discussion of the Algorithm

From the determinant condition one sees that each $\vec{v}_i$ is primitive. So if $\vec{v}_{i+1} \neq Y$, then $\vec{v}_{i+1}$ does not lie on the line $\mathcal{M}$ through $X = (0,0)$ and Y. It lies strictly on one side of $\mathcal{M}$ and $\vec{v}_i + \vec{v}_{i+1}$ lies strictly on the other side. Thus if $\vec{v}_{i+1} \neq Y$, we can do one more step of the algorithm and get $\vec{v}_{i+2} = \vec{v}_i + \mu\vec{v}_{i+1}$ with $\mu \geqslant 1$. Moreover, the determinant condition also tells us that $\vec{v}_{i+2}$, $\vec{v}_{i+1}$ generate the integral lattice, so Y cannot be in the interior of the triangle with vertices $(0,0)$, $\vec{v}_{i+1}$, $\vec{v}_{i+1} + \vec{v}_{i+2}$. It follows that $\vec{v}_{i+2}$ lies in the polygon with vertices $(0,0)$, $\vec{v}_i$, Y, $\vec{v}_{i+1}$. By induction it lies in the polygon with vertices $(0,0)$, $(0,1)$, Y, $(1,0)$. So the integer $||\vec{v}_{i+2}||^2$ is at most $||Y||^2$. On the other hand $||\vec{v}_{j+1}|| > ||\vec{v}_j||$ for $j \geqslant 1$, so the algorithm stops.

4.4. Remark

The continued fraction algorithm is famous for its fast convergence. In particular we expect the angle between $\vec{v}_n$ and $\vec{v}_{n+1}$ to be very small. This makes that the arc height of the triangle with vertices X, Y, $\vec{v}_{n+1}$ will be small. That is the key idea. A small arc height will push the center of the circumscribed circle of the triangle far away, and by linear programming that will make that the term $||P - X + Q - Y||$ in (1) dominates. All we need to do is make this precise and see when this argument fails. We will see that the failures are covered by the special cases treated above. So we use continued fraction theory for insight in orders of magnitude of the various terms. As the continued fraction theory that we need is very elementary, everything that we prove with it can also be proved without mentioning continued fractions. For instance, one may describe $\vec{v}_{n+1}$ as one of the two lattice points Z so that the angle $\angle XZY$ is obtuse and triangle ΔXYZ has area $1/2$. We do not really need the continued fraction algorithm to find such Z, but by constructing it in this particular way we understand more about it.

Lemma 3. *The triangle with vertices X, Y, $\vec{v}_{n+1}$ has arc height at most $\frac{1}{2}(-5+\sqrt{30})$.*

Proof. Put $\omega = 2(-5+\sqrt{30})$. We have $Y = \vec{v}_n + \mu\vec{v}_{n+1}$ with $\mu \geqslant 1$. As $2 \leqslant y_2 < y_1$, we must have $n \geqslant 1$. Put $\vec{b} = \vec{v}_n$, $\vec{c} = \vec{v}_{n+1}$, $b = ||\vec{b}||$, $c = ||\vec{c}||$. Note that $c > b \geqslant 1$. The triangle with vertices $(0,0)$, $\vec{b}$, $\vec{c}$ has area $\frac{1}{2}$, so $\vec{b}\cdot\vec{c} = bc\sqrt{1 - b^{-2}c^{-2}} = \sqrt{b^2c^2 - 1}$. The triangle of the lemma also has area $\frac{1}{2}$, so what we have to show is that

$$\frac{2||Y - X||}{c\,||\vec{b} + (\mu-1)\vec{c}|| + \vec{c}\cdot\left(\vec{b} + (\mu-1)\vec{c}\right)} \leqslant \omega. \tag{13}$$

This may be checked directly for $Y = (5,4)$, $Y = (6,5)$, $Y = (7,6)$, $Y = (8,7)$.

The angle between $\vec{c}$ and $\vec{b} + (\mu-1)\vec{c}$ is at least as sharp as the angle between $\vec{c}$, $\vec{b}$, so its cosine is larger. Therefore the denominator in (13) is at least

$$c\,||\vec{b} + (\mu-1)\vec{c}||\left(1 + \sqrt{1 - b^{-2}c^{-2}}\right).$$

The numerator in (13) is at most $2||\vec{b}+(\mu-1)\vec{c}||+2c$ and it suffices to show that

$$\frac{2}{c\left(1+\sqrt{1-b^{-2}c^{-2}}\right)}+\frac{2}{||\vec{b}+(\mu-1)\vec{c}||\left(1+\sqrt{1-b^{-2}c^{-2}}\right)}\leqslant\omega. \tag{14}$$

We may rewrite the second term in (14) as

$$\frac{2}{\sqrt{b^2+(\mu-1)^2c^2+2(\mu-1)\sqrt{b^2c^2-1}}\left(1+\sqrt{1-b^{-2}c^{-2}}\right)}.$$

Thus both fractions in (14) may be viewed as functions of the variables b, c, μ and they are descending in each variable, as long as b, c, μ are at least one. Therefore it is routine to check (14) in each of the following cases

- $\mu\geqslant 8$, $c\geqslant\sqrt{2}$, $b\geqslant 1$.
- $\mu\geqslant 2$, $c\geqslant\sqrt{5}$, $b\geqslant 1$.
- $\mu\geqslant 1$, $c\geqslant\sqrt{5}$, $b\geqslant\sqrt{5}$.

But these cases cover most possibilities. Indeed suppose we are in a case that is not covered. Then $b<\sqrt{5}$, as otherwise $c>b\geqslant\sqrt{5}$, $\mu\geqslant 1$. The vector $\vec{b}$ is primitive, so if $b<\sqrt{5}$, then b is 1 or $\sqrt{2}$. As $c>b$, we then have $c=\sqrt{2}$ or $c\geqslant\sqrt{5}$.

If $b=1$, then $n=1$, $\vec{b}=(1,0)$, $\vec{c}=(a,1)$ for some $a\geqslant 1$. If $a=1$, then $c=\sqrt{2}$. We have $\mu=y_2\geqslant 2$. And we may assume $\mu<8$, as $\mu\geqslant 8$ is covered. But $\mu=2$ gives $Y=(3,2)$, which has been excluded. Similarly $\mu=3$ leads to the exclude case $Y=(4,3)$. And $\mu=4$, $\mu=5$, $\mu=6$, $\mu=7$ give the cases $Y=(5,4)$, $Y=(6,5)$, $Y=(7,6)$, $Y=(8,7)$ discussed above. Next let $a\geqslant 2$, so that $c\geqslant\sqrt{5}$. We still have $\mu=y_2\geqslant 2$ and this case is covered.

We turn to $b=\sqrt{2}$. Then $n=2$, $\vec{b}=(1,1)$, $\vec{c}=(1+a,a)$ for some $a\geqslant 1$. So $c\geqslant\sqrt{5}$, and $\mu\geqslant 2$ is covered. Take $\mu=1$. We have $Y=(2+a,1+a)$. But that does not make sense. Such Y is reached in the previous step of the algorithm. □

Lemma 4. *There is a point $Z\in\mathcal{S}\cap\mathbb{Z}^2$ so that $\angle XZY$ is obtuse and so that the triangle ΔXYZ has arc height at most $\frac{1}{2}(-5+\sqrt{30})$. In particular, the arc height is less than $\frac{1}{4}$.*

Proof. If in the previous lemma $\vec{v}_{n+1}\in\mathcal{S}$, take $Z=\vec{v}_{n+1}$. If not, take $Z=Y-\vec{v}_{n+1}$ and observe this yields a triangle that is congruent to the previous one. □

4.5. Some More Points and Lines

The last lemma shows that we have a "very thin" triangle. We wish to show that linear programming now forces P, Q to positions that make $||P-X+Q-Y||$ large enough. Actually we will use a little more geometry than just linear programming. That way we have to treat fewer cases separately. The reader will need to draw an elaborate picture. We now discuss its ingredients. Compare also the (tilted) picture in subsection 4.7 below. Recall from 2.1 that we already have X, Y, $\mathcal{S}$, P, Q and recall that we think of X, Y, $\mathcal{S}$ as given, while P, Q may wander.

Fix Z as in the last lemma. Note that Z lies in the intersection of $\mathcal{S}$ with the interior of the circle with diameter $[X,Y]$. Let M be the center of the circumscribed circle $\mathcal{C}$ of triangle ΔXYZ. So M lies below the line $\mathcal{M}$ through X and Y. Let C be the midpoint of $[X,Y]$. Let $\mathcal{L}$ be the perpendicular bisector of $[X,Y]$ and let D be its upper intersection with $\mathcal{C}$, so that $||C-D||$ is the arc height of triangle ΔXYZ. One has from the power of C with respect to $\mathcal{C}$ that

$$||C-Y||^2=||C-D||\ (2||C-M||+||C-D||)$$

or

$$||C-M||=\frac{||C-Y||^2-||C-D||^2}{2||C-D||} \tag{15}$$

Let A be the projection of P onto $\mathcal{L}$ and let B be the projection of Q onto $\mathcal{L}$. Let α be the angle $\angle XYD$. Then 2α is the angle $\angle DMY$. Let β be the angle $\angle PMD$ and let γ be the angle $\angle QMD$.

As $2 \leqslant y_2 < y_1$ and $\gcd(y_1, y_2) = 1$ and $Y \neq (3,2)$, $Y \neq (4,3)$, we have

$$||X - Y|| > 5, \tag{16}$$

so $||X - Y|| > 2||P - Q||$. Of the four corners of $\mathcal{S}$ the corner X is closest to P and Y is closest to Q. From these facts it easily follows that

$$P, Q \text{ lie below } \mathcal{M}. \tag{17}$$

It is also clear that

$$||B - A|| \leqslant 1, \tag{18}$$

because $||P - Q|| \leqslant 1$. The vector $P - X + Q - Y$ is at least as long as its projection along $\mathcal{L}$, so

$$||P - X + Q - Y|| \geqslant ||C - A|| + ||C - B||. \tag{19}$$

4.6. The Low Case

Suppose A and/or B lies below M (or equals M). If P, Q both lie on $\mathcal{L}$, then in fact both $A = P$ and $B = Q$ lie below M, because otherwise P would be on the wrong side of the perpendicular bisector of $[X, Z]$, or Q on the wrong side of the perpendicular bisector of $[Z, Y]$. (These bisectors meet in M.) So then $||C - A|| + ||C - B|| \geqslant 2||C - M||$. And if P, Q do not both lie on $\mathcal{L}$, then $2(P - Q) \cdot (X - Y) \geqslant 2$ and $||C - A|| + ||C - B|| \geqslant 2||C - M|| - ||A - B|| \geqslant 2||C - M|| - 1$. So in either case the right hand side of (1) is at least $1 + 2||C - M||$, which equals

$$1 + \frac{||C - Y||^2 - ||C - D||^2}{||C - D||}.$$

Now recall that the arc height $||C - D||$ is at most $\frac{1}{4}$. So

$$1 + \frac{||C - Y||^2 - ||C - D||^2}{||C - D||} \geqslant 4||C - Y||^2 = ||X - Y||^2,$$

as required.

4.7. The Main Case

We are left with the case that both A and B lie on $[C, M]$ as in the unrealistic picture below (Fig. 3). It is tilted with respect to the coordinate axes. More importantly, $||C - D||$ is too large with respect to $||P - Q||$ and $||P - Q||$ is too large with respect to $||X - Y||$. Correcting these defects will drive P, Q way down.

The perpendicular bisectors of $[X, Z]$ and $[Z, Y]$ meet at M under an angle 2α. We know that X, P lie on the same side of the first bisector and Q, Y lie on the same side of the second. It follows that

$$\beta + \gamma \geqslant 2\alpha. \tag{20}$$

The function tan is convex on the interval from 0 to $\pi/2$, so we conclude

$$\frac{\tan\beta + \tan\gamma}{2} \geqslant \tan(\frac{\beta + \gamma}{2}) \geqslant \tan\alpha. \tag{21}$$

We have to show that

$$h_1 = ||X - Y||^2 - 1 - 2(P - Q) \cdot (X - Y) - ||P - X + Q - Y||$$

is nonpositive.

Write $a = \min(||A - M||, ||B - M||)$, $b = \max(||A - M||, ||B - M||)$, so that $b - a = ||B - A|| \leqslant 1$. Then

$$2(P - Q) \cdot (X - Y) = 4||C - Y||(||A - P|| + ||B - Q||) \geqslant 4a||C - Y||(\tan\beta + \tan\gamma),$$

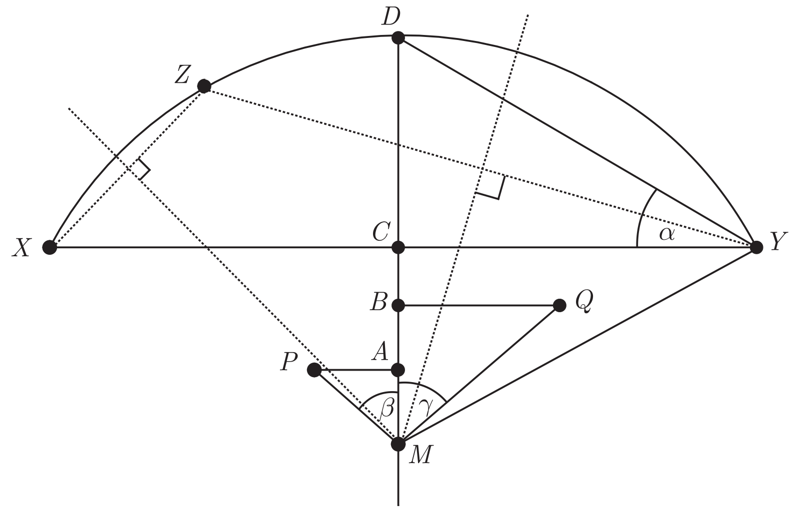

Fig. 3.

which is at least $8a||C-Y||\tan\alpha = 8a||C-D||$. So $h_1 \leqslant h_2$, with

$$h_2 = 4||C-Y||^2 - 1 - 8a||C-D|| - (2||C-M|| - a - b).$$

Using $b \leqslant a+1$, and (15), we get $h_2 \leqslant h_3$, with

$$h_3 = 4||C-Y||^2 - 8a||C-D|| - \frac{||C-Y||^2 - ||C-D||^2}{||C-D||} + 2a.$$

Now $\beta+\gamma \geqslant 2\alpha$, so either $\beta \geqslant \alpha$ and $a \leqslant ||A-P||/\tan\alpha \leqslant 1/\tan\alpha = ||C-Y||/||C-D||$, or $\gamma \geqslant \alpha$ and $a \leqslant ||B-Q||/\tan\alpha \leqslant 1/\tan\alpha = ||C-Y||/||C-D||$. So $a \leqslant ||C-Y||/||C-D||$ in either case. As $||C-D|| < \frac{1}{4}$, this yields $h_3 \leqslant h_4$, where

$$h_4 = ||C-Y||^2(4 - \frac{1}{||C-D||}) - 2||C-Y||(4 - \frac{1}{||C-D||}) + ||C-D||.$$

Remains to show that $h_4 \leqslant 0$.

Now for fixed $||C-D||$ with $0 < ||C-D|| < 1/4$, the function

$$f(x) = x^2(4 - \frac{1}{||C-D||}) - 2x(4 - \frac{1}{||C-D||}) + ||C-D||$$

is concave with $f(0) > 0$. As $||C-Y|| = ||X-Y||/2 > 5/2$, it suffices to show that $f(5/2) \leqslant 0$. Now

$$f(5/2) = \frac{5}{4}(4 - \frac{1}{||C-D||}) + ||C-D||.$$

So consider

$$g(x) = \frac{5}{4}(4x-1) + x^2$$

and recall that $||C-D|| \leqslant \frac{1}{2}(-5+\sqrt{30})$. The function $g(x)$ is convex with $g(0) < 0$ and $g(\frac{1}{2}(-5+\sqrt{30})) = 0$, so $f(5/2) \leqslant 0$ indeed. □

4.8. Remark on Sharpening

One may show in the same fashion that in fact the sharper estimate

$$||X-Y||^2 \leqslant 2 - \sqrt{2} + 2(P-Q)\cdot(X-Y) + ||P-X+Q-Y||$$

holds under the conditions of the theorem.

This sharper estimate is an equality when $X = (0,0)$, $Y = (1,1)$, $P = Q = (1,0)$. So the constant $2-\sqrt{2}$ can not be lowered further.

To get the improved result one shows that in lemma 3 the arc height is actually at most $\frac{\sqrt{82}}{20+14\sqrt{2}}$, treating the cases $Y = (9,8)$, $Y = (10,9)$, $Y = (11,10)$ like the case $Y = (8,7)$. (This to compensate for the fact that $\mu \geqslant 8$, $c \geqslant \sqrt{2}$, $b \geqslant 1$ becomes $\mu \geqslant 11$, $c \geqslant \sqrt{2}$, $b \geqslant 1$.) Instead of $||X-Y|| > 5$ one uses $||X-Y|| \geqslant \sqrt{29}$. The other modifications are then straightforward.

4.9. Remarks on Dimension Three

Consider the problem in dimension three. Now H is a closed half space. Say $X = (0,0,0)$, $Y \in \mathbb{Z}^n$ with $||Y-X|| > 5$. As always we assume P, $Q \in \mathbb{Z}^n$ are such that $X \in \mathcal{FT}(P)$ and $Y \in \mathcal{FT}(Q)$. Suppose one can find $Z \in H \cap \mathbb{Z}^3$ so that $\angle XZY$ is obtuse and so that triangle ΔXZY has arc height at most $\frac{1}{2}(-5+\sqrt{30})$, in the plane $\mathcal{P}$ through X, Y, Z. Then inequality (1) holds. To see this, argue in the manner above with the the projections of P, Q onto $\mathcal{P}$. More generally, it will suffice to have an arc height $h < 1/4$ so that

$$||X-Y||(1-4h) + h^2 + ||X-Y||^2(-1+4h)/4 \leqslant 0.$$

However, although this applies often, it may not apply sufficiently often. It seems better to look for $Z, T \in H \cap \mathbb{Z}^3$, so that the plane through X, Y, Z separates T from the center of the circumscribed sphere of the simplex $XYZT$ and the radius of this circumscribed sphere is sufficiently larger than $||X-Y||^2$. There should be a three dimensional analogue of lemma 3 that shows such Z, T exist in the typical case.

On the other hand, there are some examples in dimension three where inequality (1) fails:

Consider the half space

$$H = \{\ (x_1,x_2,x_3) \mid 2(1-7-7^2)x_1 - x_2 + (7+1)x_3 \geqslant 0\ \},$$

with $P = (2*7^2,1,0)$, $Q = (2*7^2,2,0)$, $X = (0,0,0)$, $Y = (1,2,2*7)$. Then $X \in \mathcal{FT}(P)$ and $Y \in \mathcal{FT}(Q)$ but $||X-Y||^2 - 2(P-Q)\cdot(X-Y) - ||P-X+Q-Y|| = 197 - \sqrt{38222} \approx 1.49552$. In this example the reader may wish to replace the number 7 with something larger.

Next consider the half space

$$H = \{\ (x_1,x_2,x_3) \mid (1-2*7^2)x_1 - x_2 + 7x_3 \geqslant 0\ \},$$

with $P = Q = (2*7^2,1,0)$, $X = (0,0,0)$, $Y = (1,1,2*7)$. Then $X \in \mathcal{FT}(P)$ and $Y \in \mathcal{FT}(Q)$ but $||X-Y||^2 - 2(P-Q)\cdot(X-Y) - ||P-X+Q-Y|| = 198 - \sqrt{38222} \approx 2.49552$.

So in dimension three the constant needs to be larger than one. We do not know yet if that is all that needs to be changed, but it looks like it. From extensive random sampling we expect that the constant under the condition $||P-Q|| = 1$ may be taken to be $3/2$ while under the condition $||P-Q|| = 0$ one must take it to be $5/2$.

5. CONCLUSION

In dimension two we have proved inequality (1). In dimension three we conjecture that if X, Y, P, $Q \in \mathbb{Z}^3$ with $||P-Q|| = 1$, $X \in \mathcal{FT}(P)$, $Y \in \mathcal{FT}(Q)$, then

$$||X-Y||^2 \leqslant 1.5 + 2(P-Q)\cdot(X-Y) + ||P-X+Q-Y||. \tag{22}$$

REFERENCES

1. Hesselink W.H., On Quadratic Pruning of IMA, 2005, unpublished manuscript `http://www.cs.rug.nl/~wim/pub/whh346.pdf`
2. Edwards, J., Closest Points in Lattices, 2006 `http://www.math.uu.nl/people/edwards/lattice.pdf`
3. Continued Fraction from MathWorld `http://mathworld.wolfram.com/ContinuedFraction.html`
4. Fowler, D.H., Ratio in Early Greek Mathematics, *Bulletin (New Series) of the American Mathematical Society*, 1979, vol. 1, pp. 807–846.