On the parameterized complexity of d-dimensional point set pattern matching *

Sergio Cabello† Panos Giannopoulos‡§ Christian Knauer¶

Abstract

Deciding whether two n-point sets $A, B \in \mathbb{R}^d$ are congruent is a fundamental problem in geometric pattern matching. When the dimension d is unbounded, the problem is equivalent to graph isomorphism and is conjectured to be in FPT.

When $|A| = m < |B| = n$, the problem becomes that of deciding whether A is congruent to a subset of B and is known to be NP-complete. We show that point subset congruence, with d as a parameter, is W[1]-hard, and that it cannot be solved in $O(mn^{o(d)})$-time, unless SNP $\subset$ DTIME$(2^{o(n)})$. This shows that, unless FPT = W[1], the problem of finding an isometry of A that minimizes its directed Hausdorff distance, or its Earth Mover's Distance, to B, is not in FPT.

Keywords: Computational Complexity, Computational Geometry, Fixed Parameter Tractability, Geometric Point Set Matching, Congruence, Unbounded Dimension.

1 Introduction

Geometric pattern matching has been a topic of considerable research in computational geometry with applications in computer vision, and is usually modeled as the following optimization problem: given two sets A and B of geometric primitives, an appropriate distance measure, and a transformation group, find a transformation of A that minimizes its distance to B; see the survey by Alt and Guibas [2]. Typical geometric primitives include points, segments,

*A preliminary version of this paper appeared in the 2nd International Workshop on Parameterized and Exact Computation, volume 4169 of Lecture Notes in Computer Science, p.p. 175–183, 2006.

†Department of Mathematics, IMFM and FMF, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia, sergio.cabello@fmf.uni-lj.si; partially supported by the EC Sixth Framework Programme under a Marie Curie Intra-European Fellowship and by the Slovenian Research Agency, project J1-7218.

‡Corresponding author.

§Humboldt-Universität zu Berlin, Institut für Informatik, Unter den Linden 6, D-10099 Berlin, Germany, panos@informatik.hu-berlin.de; partially supported by the DFG project "Parametrische Komplexitätstheorie", no. GR 1492/7-2.

¶Institut für Informatik, Freie Universität Berlin, Takustraße 9, D-14195 Berlin, Germany, Christian.Knauer@inf.fu-berlin.de.

disks, while typical transformations include isometries, and scaling, that is, combinations of translations, rotations, reflection, and scaling.

For sets of points, several distance measures have been extensively studied in this framework, such as, the bottleneck distance [3, 12, 14], the (directed) Hausdorff distance [9, 10, 16] and the Earth Mover's Distance [7, 11]. The case of the bottleneck distance with respect to isometries leads to the fundamental decision problem of whether A is congruent to B or to a subset of B. A formal definition will be given shortly.

The complexity of these two problems for point sets in unbounded dimensions has been already studied within the classical complexity theory: the former is graph-isomorphism-hard and the latter is NP-complete. In this paper we study subset congruence and related problems from the parameterized complexity point of view, with the dimension as the parameter.

Parameterized complexity theory measures the complexity of hard algorithmic problems in terms of parameters in addition to the problem input size. A problem of input size n and a non-negative integer parameter k is *fixed-parameter tractable* if it can be solved by an algorithm that runs in $O(f(k) \cdot n^c)$ time, where f is a computable function depending only on k and c is a constant independent of k. The class of all fixed-parameter tractable problems is denoted by FPT. Additionally, an infinite hierarchy of classes, the W-hierarchy, has been introduced for establishing fixed-parameter intractability: a parameterized problem that is hard for some level of the W-hierarchy, e.g., W[1], is not in FPT (under standard complexity theoretic assumptions). For an introduction to the field, the reader is referred to the textbook of Flum and Grohe [15].

Several important geometric optimization problems have been recently shown to be fixed-parameter tractable with respect to parameters related to properties of the input objects. For example, Deineko et al. [13] showed that the Euclidean TSP is in FPT when parameterized with the number of the points inside the convex hull of the input point set. However, there is only a handful of results regarding the fixed-parameter intractability of certain hard geometric problems, and these concern only standard parameterizations [15] of optimization problems, i.e., where the parameter measures the size of the solution; see, for example, Marx [17] for such results on geometric graph problems. The dimension of the input objects is a "structural" parameter important to many (apparently) intractable geometric problems, and, thus, a good candidate for parameterized complexity.

Preliminaries. For a point $a \in \mathbb{R}^d$, let $a(r)$ denote its rth component. The origin is denoted by o. For two points $a, b \in \mathbb{R}^d$, let $||a - b|| = (\sum_{r=1}^{d}(a(r) - b(r))^2)^{1/2}$ be their Euclidean distance. A map $\mu : \mathbb{R}^d \to \mathbb{R}^d$ is an *isometry* if it preserves distances, that is, if $||\mu(a) - \mu(b)|| = ||a - b||$ for all $a, b \in \mathbb{R}^d$.

Let $A = \{a_1, \ldots, a_m\}$ and $B = \{b_1, \ldots, b_n\}$ be point sets in $\mathbb{R}^d$ with $m \leqslant n$; for simplicity, we will sometimes assume the obvious ordering on the elements of sets defined in this way, e.g., a_1 is the first point in A and so on. We use the notation $\mu(A) = \{\mu(a_1), \ldots, \mu(a_m)\}$.

Sets A and B are said to be *congruent* if there is an isometry μ for which $\mu(A) = B$.

We denote by CONGRUENCE the problem of finding whether two sets A and B are congruent. When $|A| < |B|$, the problem becomes the one of finding whether A is congruent to a subset of B, and is denoted by SUBSET-CONGRUENCE. We are interested in the parameterized version of these problems, referred to as p-CONGRUENCE and p-SUBSET-CONGRUENCE, with the dimension d being the input parameter.

Related work. There is a variety of algorithms that solve CONGRUENCE in $O(n \log n)$ time for $d = 2, 3$; see Alt et al. [3] and references therein. In higher dimensions, the currently best time bound is $O(n^{\lceil d/3 \rceil} \log n)$ [5]; Akutsu [1] claims that a bound of $O(n^{d/4+O(1)})$ (randomized) is possible, but he gives no direct proof. It is conjectured [5] that p-CONGRUENCE is in FPT. Moreover, for unbounded dimension, CONGRUENCE is polynomially equivalent to graph isomorphism [1, 19].

Braß [4] shows that SUBSET-CONGRUENCE can be solved in $O(mn^{4/3} \log n)$ time for $d = 2$ and in $O(mn^{7/4} \log n \, \beta(n))$ time for $d = 3$ (randomized, where $\beta(n)$ is an extremely slow growing function.) In higher dimensions, the fastest known algorithm runs in $O(mn^d)$ time [12], and the problem is NP-complete when d is unbounded [1]. It is an open question whether the high-dimensional bound can be improved [4]; see also Braß and Pach [6] for a survey on computational and combinatorial problems related to geometric patterns.

Model of computation. We assume the standard Turing machine model of computation. The coordinates of the points used in our reduction are rational, with denominators and numerators bounded by a polynomial in n.

Results. We show that p-SUBSET-CONGRUENCE is W[1]-hard. Our reduction from p-CLIQUE, with d being linear in the size of the clique, shows in addition that an $O(mn^{o(d)})$-time algorithm for the former problem exists only if SNP $\subset$ DTIME$(2^{o(n)})$. Moreover, for $|A| < |B|$ and any point set distance D for which $D(A, B) = 0$ if and only if $A \subset B$, e.g., directed Hausdorff distance, Earth Mover's Distance, our hardness result implies that minimizing D under isometries is not in FPT (unless FPT = W[1]).

2 Parameterized subset congruence

First, it is easy to see that p-SUBSET-CONGRUENCE is in W[P]. An accepting certificate of size $O(d \log n)$ for a non-deterministic Turing machine can be given by guessing a bijection between a d-point subset of A and a d-point subset of B, with $\log n$ bits needed to represent each point. Then, one can check in polynomial time whether the bijection is an isometry and, if yes, whether this isometry maps the rest of the points of A to points of B.

We denote by p-CLIQUE the parameterized problem of finding a clique in a given graph, where the size of the clique is the parameter. In the following, we reduce p-CLIQUE, which is known to be W[1]-complete [15], to p-SUBSET-CONGRUENCE.

Theorem 1 p-SUBSET-CONGRUENCE *is* W[1]*-hard.*

Proof: Let k be the size of the clique being looked for in a graph $G(\{1,\ldots,n\},E)$. We construct two point sets A and B in $\mathbb{R}^{2k}$ with the property that A is congruent to a subset of B if and only if G has a k-clique.

We first construct a set $\mathcal{L} = \{L_i | i = 1,\ldots,k\}$ of point sets L_i - referred to as *level* sets. Each level set lies on a two-dimensional plane in $\mathbb{R}^{2k}$, with all k planes being pairwise orthogonal, and contains n points that lie on a unit circle centered at the origin o. The property we exploit is that any two points from different circles are at the same distance, and therefore they are metrically indistinguishable. Furthermore, each of the k circles can rotate around the origin independently of the others, which corresponds to the k choices that have to be made for a k-clique. The construction has to be tuned so that only this type of rigid motions are relevant. A detailed description of a level set is given below.

Level set $L_i = \{l_{i,j} \in \mathbb{R}^{2k} | j = 1,\ldots,n\}$ is constructed as follows. First, we have $l_{i,j}(r) = 0$ for $r \neq 2i-1, 2i$ and $\sum_{r=1}^{2k} l_{i,j}^2(r) = 1$; this implies that $||l_{i,j} - l_{i',j'}|| = \sqrt{2}$ when $i \neq i'$. It will be convenient to choose the points on the unit circle from a short arc such that the distance between any two points is at most $\sqrt{2}/6$, that is, $||l_{i,j} - l_{i,j'}|| < \sqrt{2}/6$ for every i, j and j'; see Fig. 1. This is done according to the following lemma.

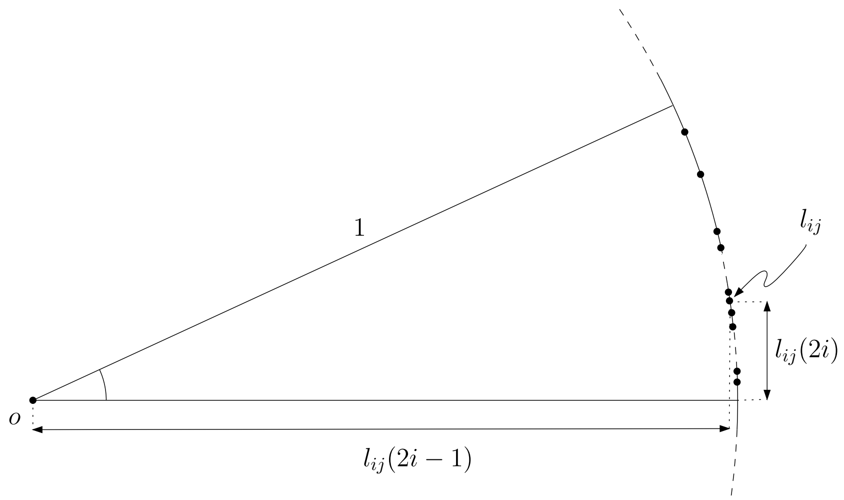

Figure 1: The points of each level set lie on a circular arc of length $\sqrt{2}/6$.

Lemma 2 *For any $n > 0$, there exist n distinct points on the unit circle such that they all have rational coordinates with the numerators and denominators bounded by a polynomial in n, and the distance between any two points is at most $\sqrt{2}/6$.*

Proof: Each point $p_i, i = 1, \ldots, n$, has rational non-zero coordinates generated by Pythagorean triples α_i, β_i and γ_i:

$$p_i(1) = (\beta_i/\gamma_i) \quad \text{and} \quad p_i(2) = (\alpha_i/\gamma_i)$$

with

$$\alpha_i = 2(8+i), \quad \beta_i = (8+i)^2 - 1 \quad \text{and} \quad \gamma_i = (8+i)^2 + 1.$$

Defined this way, all points satisfy $p_i^2(1) + p_i^2(2) = 1$ and have distinct coordinates because the sequence $p_i(2)$ is monotonically decreasing. Moreover, since

$$\arcsin(p_i(2)) \leqslant \arcsin(p_1(2)) = \arcsin(9/41) < \sqrt{2}/6,$$

all points lie on an arc of length $\sqrt{2}/6$, and so, any two of them can be at most $\sqrt{2}/6$ apart. $\square$

The point set $A \subset \mathbb{R}^{2k}$ is constructed as follows. Define $P_A = \{l_{i,1} \mid i = 1, \ldots, k\}$, which has one point per level set. Note that there is nothing special about $l_{i,1}$ here, we could have chosen any other point per level. The set A consists of the points of P_A and the middle-points between them, that is,

$$\begin{aligned} A = \{l_{1,1}, l_{2,1}, \ldots, l_{k,1}\} \, \cup \, \{&\frac{l_{1,1}+l_{2,1}}{2}, \frac{l_{1,1}+l_{3,1}}{2}, \ldots, \frac{l_{1,1}+l_{k,1}}{2}\} \\ \cup \, \{&\frac{l_{2,1}+l_{3,1}}{2}, \ldots, \frac{l_{2,1}+l_{k,1}}{2}\} \\ &\ddots \\ \cup \, \{&\frac{l_{k-1,1}+l_{k,1}}{2}\}. \end{aligned}$$

The set A consists of $k + \binom{k}{2}$ elements, and it has four distinct inter-point distances: $\sqrt{2}$, $\sqrt{3}/2$, $\sqrt{2}/2$ and 1. It is important to note that two points of A are at distance $\sqrt{2}$ if and only if they both belong to P_A.

The point set $B \subset \mathbb{R}^{2k}$ is constructed as follows. Define $P_B = \bigcup_i L_i$, that is, P_b contains all the points of the level sets. For a pair $j, j' \in \{1, \ldots, n\}$, $j \neq j'$, define $B_{j,j'}$ to be the $2\binom{k}{2}$ points

$$B_{j,j'} = \{\frac{l_{i,j} + l_{i',j'}}{2} \mid i, i' = 1, \ldots, k, \ i \neq i'\}.$$

We set $B = P_B \cup \left(\bigcup_{jj' \in E} B_{j,j'}\right)$. Note that if $\{j, j'\} \neq \{k, k'\}$, then $B_{j,j'}$ and $B_{k,k'}$ do not have any point in common. This also means that $\frac{l_{i,j}+l_{i',j'}}{2} \in B$ for some i, i', if and only if $jj' \in E$. Therefore, the set B consists of $kn + 2|E|\binom{k}{2}$ elements. The set of inter-point distances in B has a slightly more complex structure than the set of inter-point distances in A, but it holds that two points of B are at distance $\sqrt{2}$ if and only if they belong to different levels in P_B. Any other pair of points of B is strictly less than $\sqrt{2}$ apart.

The construction of A and B is now complete. It remains to show that G has a k-clique if and only if A is congruent to a subset of B. For the first implication, assume that G has a k-clique with vertices $\{j_1, j_2, j_3 \ldots, j_k\}$. We have to show that there is an isometry μ such that $\mu(A) \subseteq B$. The point sets $\{l_{i,j_i} \mid i = 1, \ldots, k\} \subset B$ and P_A are isometric because the distance between any two of their points is $\sqrt{2}$. Consider an isometry μ that brings $l_{i,1} \in P_A$ over $l_{i,j_i} \in B$, for $i = 1, \ldots, k$. It is clear that $\mu(P_A) \subset B$. Consider any other point p of $A \setminus P_A$, which is of the form $\frac{l_{i,1}+l_{i',1}}{2}$. Since μ is an isometry, and therefore a linear mapping, we have

$$\mu(p) = \frac{\mu(l_{i,1}) + \mu(l_{i',1})}{2} = \frac{l_{i,j_i} + l_{i',j_{i'}}}{2},$$

which is an element of $B_{j_i,j_{i'}} \subset B$ because $j_i j_{i'} \in E$. We conclude that $\mu(p) \in B$, and therefore $\mu(A) \subseteq B$.

Conversely, assume that there is an isometry μ such that $\mu(A) \subseteq B$. We have to show that G has a k-clique. Consider the k points $\mu(P_A) \subset B$. Their pairwise distance is $\sqrt{2}$, and therefore, all points of $\mu(P_A)$ must belong to P_B, and, in particular, there must be one in each of the k levels. Let $l_{1,j_1}, \ldots, l_{k,j_k}$ be the points of $\mu(P_A)$. We claim that $j_1, \ldots, j_k$ is a clique in G. Indeed, the middle-point between any two points of P_A is part of A, and therefore the middle-point between any two points of $\mu(P_A)$ is part of $\mu(A) \subseteq B$. This means that the middle-point between any two points of $l_{1,j_1}, \ldots, l_{k,j_k}$ is in B, which means that every edge between the vertices $j_1, \ldots j_k$ is in G. $\square$

Since in the above fpt-reduction $d = 2k$, an $O(mn^{o(d)})$-time algorithm for p-SUBSET-CONGRUENCE implies an $O(n^{o(k)})$-time algorithm for p-CLIQUE, which in turn implies that SNP $\subset$ DTIME$(2^{o(n)})$ [8].

Corollary 3 p-SUBSET-CONGRUENCE *can be solved in* $O(mn^{o(d)})$ *time, only if* SNP $\subset$ DTIME$(2^{o(n)})$.

Consider a point set distance D for which $D(A, B) = 0$ if and only if $A \subset B$ for every $A, B \subset \mathbb{R}^d$ with $|A| < |B|$; this is a desired property for any distance that is used to find small patterns into larger ones, e.g., directed Hausdorff distance, Earth Mover's Distance. Then, $\min_\mu D(\mu(A), B) = 0$ if and only if $\mu'(A) \subset B$ for some isometry μ'. Hence, we have the following.

Corollary 4 *Given two point sets* $A, B \in \mathbb{R}^d$*, with* $|A| < |B|$*, and a distance* D *for which* $D(A, B) = 0$ *if and only if* $A \subset B$*, the problem of minimizing* D *under isometries, when d is part of the input, is not in FPT, unless* FPT=W[1].

The above also implies that one cannot even approximate $\min_\mu D(\mu(A), B))$ in FPT-time with respect to d, since any such approximation algorithm must return the actual minimum, i.e., the value 0, in the case where $\mu'(A) \subset B$ for some isometry μ'.

3 Concluding remarks

We have studied the parameterized complexity of a fundamental point set pattern matching problem with respect to the dimension. We proved that subset congruence is $W[1]$-hard, which also implies that minimizing under isometries the directed Hausdorff distance, or the Earth Mover's Distance between two point sets in unbounded dimension is not in FPT (unless FPT=W[1]).

There are quite a few other geometric optimization problems whose complexity due to unbounded dimension has been studied; see, for example, Megiddo [18]. We believe that for such problems the dimension is an interesting parameter to be studied within the framework of parameterized complexity theory.

References

[1] T. Akutsu. On determining the congruence of point sets in d-dimensions. *Comput. Geom. Theory Appl.*, 9(4):247–256, 1998.

[2] H. Alt and L.J. Guibas. Discrete geometric shapes: Matching, interpolation, and approximation. In J.R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 121–153. Elsevier Science Publishers B.V. North-Holland, Amsterdam, 1999.

[3] H. Alt, K. Mehlhorn, H. Wagener, and E. Welzl. Congruence, similarity, and symmetries of geometric objects. *Discrete Comput. Geom.*, 3:237–256, 1988.

[4] P. Braß. Exact point pattern matching and the number of congruent triangles in a three-dimensional point set. In *Proc. 8th Annu. European Sympos. Algorithms*, volume 1879 of *Lecture Notes in Computer Science*, pages 112–119. Springer-Verlag, 2000.

[5] P. Braß and C. Knauer. Testing the congruence of d-dimensional point sets. *Internat. J. of Comput. Geom. Appl.*, 12(1):115–124, 2002.

[6] P. Braß and J. Pach. Problems and results on geometric patterns. In D. Avis, A. Hertz, and O. Marcotte, editors, *Graph Theory and Combinatorial Optimization*, Gerad 25th Anniversary, pages 17–36. Springer-Verlag, 2005.

[7] S. Cabello, P. Giannopoulos, C. Knauer, and G. Rote. Matching point sets with respect to the Earth Mover's Distance. In *Proc. 13th Annu. European Sympos. Algorithms*, volume 3669 of *Lecture Notes in Computer Science*, pages 520–531. Springer-Verlag, 2005.

[8] J. Chen, B. Chor, M. Fellows, X. Huang, D. Juedes, I. A. Kanj, and G. Xia. Tight lower bounds for certain parameterized NP-hard problems. *Information and Computation*, 201(2):216–231, 2005.

[9] L.P. Chew, D. Dor, A. Efrat, and K. Kedem. Geometric pattern matching in d-dimensional space. *Discrete Comput. Geom.*, 21, 1999.

[10] L.P. Chew, M. Goodrich, D. P. Huttenlocher, K. Kedem, J. M. Kleinberg, and D. Kravets. Geometric pattern matching under Euclidean motion. *Comput. Geom. Theory Appl.*, 7:113–124, 1997.

[11] S.D. Cohen and L.J. Guibas. The Earth Mover's Distance under transformation sets. In *Proc. 7th IEEE Internat. Conf. Computer Vision*, pages 173–187, 1999.

[12] P.J. de Rezende and D.T. Lee. Point set matching in d-dimensions. *Algorithmica*, 13:387–404, 1995.

[13] V.G. Deineko, M. Hoffmann, Y. Okamoto, and G.J. Woeginger. The traveling salesman problem with few inner points. *Oper. Res. Lett.*, 34:106–110, 2006.

[14] A. Efrat, A. Itai, and M.J. Katz. Geometry helps in bottleneck matching and related problems. *Algorithmica*, 31:1–28, 2001.

[15] J. Flum and M. Grohe. *Parameterized Complexity Theory*, volume XIV of *Texts in Theoretical Computer Science. An EATCS Series.* Springer-Verlag, 2006.

[16] D.P. Huttenlocher, K. Kedem, and M. Sharir. The upper envelope of Voronoi surfaces and its applications. *Discrete Comput. Geom.*, 9:267–291, 1993.

[17] D. Marx. Efficient approximation schemes for geometric problems? In *Proc. 13th Annu. European Sympos. Algorithms*, volume 3669 of *Lecture Notes in Computer Science*, pages 448–459. Springer-Verlag, 2005.

[18] N. Megiddo. On the complexity of some geometric problems in unbounded dimension. *J. Symb. Comput.*, 10(3-4):327–334, 1990.

[19] C. Papadimitriou and S. Safra. The complexity of low-distortion embeddings between point sets. In *Proc. 16th Annu. ACM-SIAM Sympos. Discrete Algorithms*, pages 112–118, 2005.