

A computable and continuous metric on isometry classes of high-dimensional periodic sequences

Vitaliy Kurlin  

Department of Computer Science, University of Liverpool, Liverpool, United Kingdom

Abstract

This paper introduces a metric that continuously quantifies the similarity between high-dimensional periodic sequences considered up to natural equivalences maintaining inter-point distances. This metric problem is motivated by periodic time series and point sets that model real periodic structures with noise. Most past advances focused on finite sets or simple periodic lattices. The key novelty is the continuity of the new metric under perturbations that can change the minimum period. For any sequences with at most m points within their periods, the metric is computed in time $O(m^3)$.

2012 ACM Subject Classification Theory of computation \rightarrow Computational geometry

Keywords and phrases periodic sequence, isometry, invariant, classification, metric, continuity

1 Motivations, problem statement, and overview of new results

We start from the simplest concepts in \mathbb{R} , which will be later extended to higher dimensions.

► **Definition 1.1** (periodic sequences in \mathbb{R}). A *periodic sequence* $S = \{p_1, \dots, p_m\} + l\mathbb{Z}$ is defined by a finite *motif* of points p_1, \dots, p_m in a period interval $[0, l)$ of a length $l > 0$ as the infinite sequence $p(i + mj) = p_i + jl$ indexed by $i + mj$, where $i = 1, \dots, m$ and $j \in \mathbb{Z}$. ▲

Any periodic sequence $S \subset \mathbb{R}$ is infinite in both directions, though all results below can be adapted to 1-directional sequences. A period interval $[0, l)$ excludes the endpoint l , which is equivalent to 0 by a shift by the period l , so any point $p_i \in [0, l)$ is counted once.

The set of half-integers is the periodic sequence $\{0, \frac{1}{2}\} + \mathbb{Z}$, which can be also defined as $\{0\} + \frac{1}{2}\mathbb{Z}$. A period interval $[0, l)$ is *minimal* for a given sequence $S = \{p_1, \dots, p_m\} + l\mathbb{Z}$ if S can not be represented by a shorter period interval. The points p_1, \dots, p_m are naturally ordered in $[0, l)$. Fixing a minimum period interval $[0, l)$ at the origin in the Euclidean line \mathbb{R} resolves ambiguity only for fixed sequences but our notion of origin is often relative.

In practice, it is natural to consider the two sequences $\{0, 1\} + 3\mathbb{Z}$ and $\{0, 2\} + 3\mathbb{Z}$ equivalent because they both consist of pairs of points at a distance 1 translated with period 3. Any translation in \mathbb{R} is a 1-dimensional *rigid motion* or orientation-preserving isometry.

In general, an *isometry* is any map that maintains all inter-point distances, hence includes all reflections $t \mapsto 2a - t$ around a fixed center $a \in \mathbb{R}$. If we also allow uniform scaling, we get two more equivalence relations: *similarity* (an affine map $t \mapsto at + b$ for $a \neq 0$ and any $b \in \mathbb{R}$) and *orientation-preserving similarity* (an affine map $t \mapsto at + b$ for $a > 0$).

Studying periodic sequences up to rigid motion has important practical motivations in solid crystalline materials whose structures are determined through diffraction patterns in a rigid form. Since many early known crystals were highly symmetric, they were traditionally classified by symmetry groups. This weaker equivalence relation produces 219 types in \mathbb{R}^3 or 230 if mirror images are distinguished up to rigid motion. However, more than 1.1M synthesized materials in the Cambridge Structural Database [34] require finer classifications.

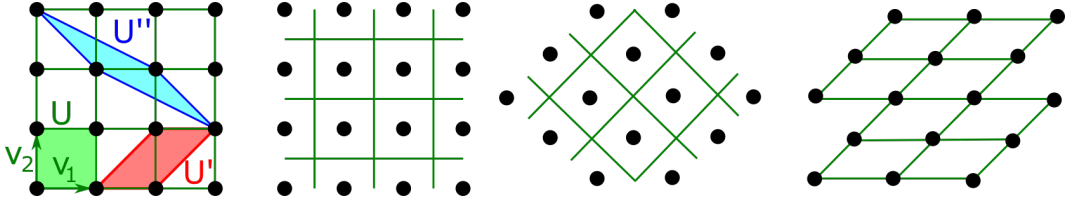
The rigid motion is the strongest equivalence of periodic crystals that remain unchanged under translations and rotations. Equivalence classes of periodic crystals up to rigid motion form an infinite continuous space, whose metric geometry was actively explored since 2020.

Because of noise in measurements, real applications need more than a binary answer to the isometry equivalence problem. Indeed, even if the same material is analyzed twice at different temperatures (or other physical conditions), the resulting structures will have slightly different positions of atomic centers. Hence we need to continuously quantify the similarity between nearly identical periodic structures. This problem of recognizing near-duplicates is the key obstacle in Crystal Structure Prediction, which tries to predict real materials by simulating millions of hypothetical structures. Since every prediction is obtained as an approximation of some local energy minimum, many predictions accumulate around the same minimum but are often represented by incomparable data with different periods and motifs.

Mathematical foundations for resolving the practical challenges above are based on a metric between isometry classes of periodic point sets, which are formally defined below.

► **Definition 1.2** (a lattice Λ and a periodic point set S in \mathbb{R}^n). A *lattice* $\Lambda \subset \mathbb{R}^n$ is the infinite set $\{\sum_{i=1}^n c_i v_i \mid c_i \in \mathbb{Z}\}$ of all integer linear combinations of a linear basis v_1, \dots, v_n in \mathbb{R}^n . A basis of Λ spans a parallelepiped $U(v_1, \dots, v_n) = \{\sum_{i=1}^n c_i v_i \mid 0 \leq c_i < 1\}$ called a *unit cell*. For a basis v_1, \dots, v_n of a lattice Λ and finite *motif* of points $p_1, \dots, p_m \in U(v_1, \dots, v_n)$, the *periodic point set* S is the Minkowski sum $S = M + \Lambda = \{u + v \mid u \in M, v \in \Lambda\} \subset \mathbb{R}^n$. ▲

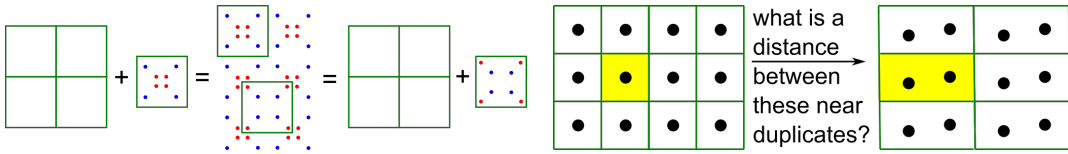
For $n = 1$, a periodic point set becomes a periodic sequence whose period interval $[0, l)$ is a unit cell. A periodic point set can be visualized as a finite union of lattice images $\Lambda + p$ obtained from a given lattice $\Lambda \subset \mathbb{R}^n$ by shifting the origin to each motif point $p \in M \subset U$.



■ **Figure 1** The same square lattices (up to isometry) can be generated by infinitely many bases.

Any lattice can be generated by many bases, see Fig. 1. Even if we fix a basis, different motifs can lead to periodic point sets that differ by translation. This ambiguity motivates us to study metrics on equivalence classes of periodic point sets up to rigid motion or isometry.

Fig. 2 (right) shows that ever-present atomic vibrations can discontinuously affect all discrete invariants such as symmetry groups, and even the minimal cell doubles in volume.



■ **Figure 2 Left:** any periodic point set (up to isometry) can be given by many sums $\Lambda + M$ in Definition 1.2. **Right:** only a continuous metric can quantify the similarity of nearly identical sets.

These perturbations can be measured by the maximum deviation of atoms from their initial positions. The maximum deviation can be small while taking a sum over infinitely many perturbed points is often infinite. This deviation is defined below as the bottleneck distance by considering bijections between atoms that can be displaced but cannot vanish.

► **Definition 1.3** (bottleneck distance d_B). For any finite or periodic point sets $S, Q \subset \mathbb{R}^n$, the *bottleneck distance* $d_B(S, Q) = \inf_{h: S \rightarrow Q} \sup_{p \in S} |p - h(p)|$ is minimized over all bijections $h: S \rightarrow Q$,

where the Euclidean norm of a vector $v = (x_1, \dots, x_m) \in \mathbb{R}^m$ is $|v| = \sqrt{\sum_{i=1}^m x_i^2}$. ▲

Since the bottleneck distance d_B involves the minimization over infinitely many bijections and points, d_B is not computable in practice but can be efficiently used as an upper bound for computable metrics on periodic point sets, see condition (e) in Problem 1.4 below.

Since a periodic point set S with m motif points in a unit cell has an input length of $O(m)$, we define the *size* of S as $|S| = m$. Despite the recent progress in dimensions $n > 1$ discussed in section 2, the following problem remained open even for the dimension $n = 1$.

► **Problem 1.4** (continuous metric on isometry classes of periodic point sets). Find a metric d satisfying all metric axioms and the two practically important conditions (d,e) below.

- (a) *first axiom* : $d(S, Q) = 0$ if and only if periodic point sets $S \cong Q$ are isometric in \mathbb{R}^n ;
- (b) *symmetry axiom* : $d(S, Q) = d(Q, S)$ for any periodic point sets $S, Q \subset \mathbb{R}^n$;
- (c) *triangle inequality* : $d(S, T) \leq d(S, Q) + d(Q, T)$ for any periodic point sets $S, Q, T \subset \mathbb{R}^n$;
- (d) *computability* : $d(S, Q)$ can be exactly computed in a polynomial time in $\max\{|S|, |Q|\}$.
- (e) *continuity* : $d(S, Q) \leq C d_B(S, Q)$ for a fixed constant C and any sets $S, Q \subset \mathbb{R}^n$. ▲

Problem 1.4 was stated for isometry, which can be replaced by closely related equivalences such as rigid motion and similarity. The first axiom in 1.4(a) resolves the isometry problem. To detect an isometry $S \cong Q$, it suffices to check if the distance vanishes: $d(S, Q) = 0$.

The first three conditions in Problem 1.4(a,b,c) imply the positivity $d \geq 0$ of any metric.

Continuity condition 1.4(d) is non-trivial already in dimension $n = 1$ because nearly identical periodic sequences might have very different periods. All the conditions of Problem 1.4 were not fulfilled by all past attempts. Any satisfactory solution should work for the simplest case $n = 1$, which remained open and will be resolved by Theorems 3.6, 4.1, 4.4.

The proposed solution to Problem 1.4 for $n = 1$ will be extended to higher dimensions for general 1-periodic sequences with values in \mathbb{R}^{n-1} , motivated by multivariate time series [18].

► **Definition 1.5** (high-dimensional 1-periodic sequences). Let \vec{e}_1 be the unit vector along the first coordinate axis in the product $\mathbb{R} \times \mathbb{R}^{n-1}$. A *high-dimensional 1-periodic sequence* $S = \{p_1, \dots, p_m\} + l\vec{e}_1\mathbb{Z}$ is defined by a finite *motif* of points p_1, \dots, p_m in a period slice $[0, l) \times \mathbb{R}^{n-1}$ of a width $l > 0$ as the infinite sequence $p(i + mj) = p_i + jl$ indexed by $i + mj$, where $i = 1, \dots, m$ and $j \in \mathbb{Z}$. We assume that the motif points p_1, \dots, p_m have distinct *time projections* $t(p_1), \dots, t(p_m) \in \mathbb{R}$ and *value projections* $v(p_1), \dots, v(p_m)$ in \mathbb{R}^{n-1} . ▲

A 1-dimensional periodic sequence in Definition 1.1 is the simplest case $n = 1$ of Definition 1.5 when value projections are empty. Main Theorems 4.1 and 4.4 solve Problem 1.4 for the first non-trivial case $n = 1$. Further Theorems 5.4 and 5.7 solve the analog of Problem 1.4 for high-dimensional sequences $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$ considered up to a product of isometries $f \times g$ of $\mathbb{R} \times \mathbb{R}^{n-1}$ so that f translates the time projection $t(S)$ within \mathbb{R} while g is an isometry of \mathbb{R}^{n-1} applied to the value projection $v(S)$ under $v: \mathbb{R} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$.

2 A review of the past work on isometry classifications and metrics

For a finite set $S \subset \mathbb{R}^n$, the distribution of all pairwise distances is a complete invariant in general position [7] meaning that almost any finite set can be uniquely reconstructed up to isometry from the set of all pairwise distances. The non-isometric 4-point sets in Fig. 3 are a classical counter-example to the completeness of this distribution of all pairwise distances.

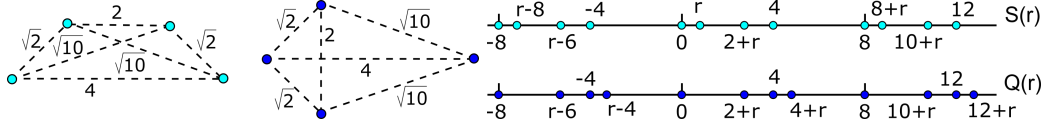


Figure 3 Left: sets $K = \{(\pm 2, 0), (\pm 1, 1)\}$ and $T = \{(\pm 2, 0), (-1, \pm 1)\}$ can not be distinguished by pairwise distances $\sqrt{2}, \sqrt{2}, 2, \sqrt{10}, \sqrt{10}, 4$. Right: sequences $S(r) = \{0, r, 2+r, 4\} + 8\mathbb{Z}$ and $Q(r) = \{0, 2+r, 4, 4+r\} + 8\mathbb{Z}$ for $0 < r \leq 1$ have the same Patterson function [29, p. 197, Fig. 2].

The isometry classification of finite point sets was algorithmically resolved by [1, Theorem 1] saying that the existence of an isometry between two m -point sets in \mathbb{R}^n can be checked in time $O(m^{n-2} \log m)$. The algorithm from [8] checks if two finite sets of m points are isometric in time $O(m^{\lceil n/3 \rceil} \log m)$. The latest advance is the $O(m \log m)$ algorithm in \mathbb{R}^4 [21], see other significant results on matching bounded rigid shapes in [33, 37, 26, 16]. The Euclidean Distance Geometry [24] studies the related problem of uniquely embedding (up to isometry of \mathbb{R}^n) an abstract graph whose straight-line edges must have specified lengths. Mémoli's seminal work on *distributions of distances* [27], also known as *shape distributions* [28, 6, 19, 25, 30], for bounded metric spaces is closest to the new invariants of sequences.

Patterson [29] was probably the first to systematically study periodic point sets. He visualised any periodic sequence $S = \{p_1, \dots, p_m\} + l\mathbb{Z} \subset \mathbb{R}$ by its simpler version in a circle of the length l but described its isometry classes by the more complicated *distance array* defined as the anti-symmetric $m \times m$ matrix of differences $p_i - p_j$ for $i, j \in \{1, \dots, m\}$.

Much later Grunbaum considered rational-valued periodic sequences represented by complex numbers on the unit circle and proved [20, Theorem 4] that the r -th order invariants (combinations of r -factor products of complex numbers) up to $r = 6$ are enough to distinguish all such sequences up to translation, though no metric was considered.

The periodic sequences $S(r), Q(r)$ in Fig. 3 emerged as first examples with identical infinite distributions of distances (or diffraction patterns). They are distinguished by recent Pointwise Distance Distributions [35] but not by the simpler Average Minimum Distances [36]. However, the latter invariants distinguish the even more interesting periodic sequences $S_{15} = \{0, 1, 3, 4, 5, 7, 9, 10, 12\} + 15\mathbb{Z}$ and $Q_{15} = \{0, 1, 3, 4, 6, 8, 9, 12, 14\} + 15\mathbb{Z}$. These sets have identical density functions [2, Example 11], which form an infinite sequence of invariant functions [17] depending on a variable radius. The density functions were fully described for periodic sequences by [4, Theorems 5, 7] but didn't lead to a metric due to incompleteness.

Appendix A reviews the recent isoset invariant for any periodic point sets in \mathbb{R}^n . This isoset reduces the isometry classification of any periodic point sets in \mathbb{R}^n to the simpler equivalence for finite sets up to rotations around a fixed center. The continuous metric on isosets [2, section 7] requires minimizations over infinitely many rotations for any dimension $n \geq 2$. For all dimensions $n \geq 1$, this metric also depends on a stable radius α . The above disadvantages left Problem 1.4 open even for periodic sequences in dimension $n = 1$.

3 Elastic metrics on spaces of isometry classes of periodic sequences

This section introduces a simple distance-based invariant, which can be expanded to both Patterson's distance array and isoset. The completeness of these smallest distance lists (SDL) is proved up to four equivalences in Proposition 3.2. The hardest challenge will be to introduce a metric on these SDL invariants to satisfy all conditions of Problem 1.4.

► **Definition 3.1** (smallest distance lists SDL° and SDL). For a periodic sequence $S = \{p_1, \dots, p_m\} + l\mathbb{Z} \subset \mathbb{R}$, let $d_i = p_{i+1} - p_i$ be the distance between successive points of S for $i = 1, \dots, m$, where $p_{m+1} = p_1 + l$. Consider the lexicographic order on ordered lists so that $(d_1, \dots, d_m) < (d'_1, \dots, d'_m)$ if $d_1 = d'_1, \dots, d_i = d'_i$ for some $0 \leq i < m$, where $i = 0$ means the empty set of identities, and $d_{i+1} < d'_{i+1}$. The oriented *smallest distance list* $\text{SDL}^\circ(S)$ is the lexicographically smallest list obtained from (d_1, \dots, d_m) by cyclic permutations. The unoriented *smallest distance list* $\text{SDL}(S)$ is the lexicographically smallest list obtained from (d_1, d_2, \dots, d_m) and the reversed list $(d_m, d_{m-1}, \dots, d_1)$ by cyclic permutations. ▲

The periodic sequences $S_2 = \{0, 1\} + 3\mathbb{Z}$ and $3 - S_2 = \{0, 2\} + 3\mathbb{Z}$ have the same $\text{SDL}^\circ = (1, 2) = \text{SDL}$. The periodic sequences $S_3 = \{0, 1, 3\} + 6\mathbb{Z}$ and $6 - S_3 = \{0, 3, 5\} + 6\mathbb{Z}$ have $\text{SDL}(S_3) = (1, 2, 3) = \text{SDL}(6 - S_3)$ and $\text{SDL}^\circ(S_3) = (1, 2, 3) \neq (1, 3, 2) = \text{SDL}^\circ(6 - S_3)$.

► **Proposition 3.2** (classifications of periodic sequences). For any periodic sequence $S = \{p_1, \dots, p_m\} + l\mathbb{Z} \subset \mathbb{R}$ with a minimum period $l > 0$, the smallest distance lists $\text{SDL}^\circ(S)$, $\text{SDL}(S)$ are complete invariants up to translation and isometry, respectively. If we scale $\text{SDL}(S)$, $\text{SDL}^\circ(S)$ so that the sum of distances is 1, the *normalized distance lists* $\text{NDL}(S)$, $\text{NDL}^\circ(S)$ are complete up to similarity and orientation-preserving similarity, respectively. ■

Proof. Any translation in \mathbb{R} preserves the cyclic order of the motif points p_i and their inter-point distances $d_i = p_{i+1} - p_i$. Up to translation in \mathbb{R} , the ordered distance list (d_1, \dots, d_m) can change only by cyclic permutation, so $\text{SDL}(S)$ is invariant. Any reflection preserves all distances but reverses their order to $(d_m, d_{m-1}, \dots, d_1)$ for a fixed orientation of \mathbb{R} .

The completeness of $\text{SDL}^\circ(S)$ and $\text{SDL}(S)$ follows by reconstructing the points $p_i = \sum_{j=1}^{i-1} d_j$, where the first point $p_1 = 0$ is at the origin, up to translation and isometry, respectively.

Finally, normalizing the distance lists $\text{SDL}^\circ(S)$ and $\text{SDL}(S)$ to make the sum of distances equal to 1 is equivalent to scaling a periodic sequence to make its period equal to 1. ◀

The naive subtraction of distance lists reveals discontinuity of the component-wise comparison. For example, the periodic sequence $S_0 = \{0, 1, 3, 4\} + 7\mathbb{Z}$ has perturbations $S_\varepsilon^\pm = \{0, 1 \pm \varepsilon, 3 \pm \varepsilon, 4\} + 7\mathbb{Z}$ for any small $\varepsilon > 0$. The smallest distance lists $\text{SDL}^\circ(S_\varepsilon^-) = (1 - \varepsilon, 2, 1 + \varepsilon, 3)$ and $\text{SDL}^\circ(S_\varepsilon^+) = (1 - \varepsilon, 3, 1 + \varepsilon, 2)$ have the large component-wise difference $(0, 1, 0, -1)$ because the minimum distance $1 - \varepsilon$ is followed by different distances $2 < 3$ in the nearly identical $S_\varepsilon^\pm = \{0, 1 \pm \varepsilon, 3 \pm \varepsilon, 4\} + 7\mathbb{Z}$ for any $\varepsilon > 0$. This discontinuity will be resolved by minimising over cyclic permutations but there is one more obstacle below.

It seems natural to always reduce a period of $S = \{p_1, \dots, p_m\} + l\mathbb{Z} \subset \mathbb{R}$ to a minimum positive value $l > 0$. However, the list $\text{SDL}(S) = (d_1, \dots, d_m)$ of a fixed size m cannot be directly used for comparing sequences that have different numbers of motif points.

► **Definition 3.3** (multiple kS of a periodic sequence S). For any integer $k > 1$ and a periodic sequence $S = \{p_1, \dots, p_m\} + l\mathbb{Z} \subset \mathbb{R}$, define the *multiple* periodic sequence as $kS = \{p_1 + jl, \dots, p_m + jl\}_{j=0}^{k-1} + kl\mathbb{Z}$ with km motif points in the period interval $[0, kl)$. ▲

Definition 3.4 will introduce two elastic metrics comparing any periodic sequences S, Q through their multiples from Definition 3.3 with the same number m of motif points.

For any ordered list (d_1, \dots, d_m) of m distances, let $C^o(m)$ denote the group of m cyclic permutations that are iterations of $(d_1, d_2, \dots, d_m) \mapsto (d_2, \dots, d_m, d_1)$. Let $C(m) \supset C^o(m)$ denote the larger group of $2m$ permutations also including the reversed cyclic permutations composed of the reversion $(d_1, d_2, \dots, d_m) \mapsto (d_m, \dots, d_2, d_1)$ with all $\sigma \in C^o(m)$. Denote the action of a permutation σ on a vector $v = (d_1, d_2, \dots, d_m)$ by $\sigma(v) = (d_{\sigma(1)}, \dots, d_{\sigma(m)})$.

► **Definition 3.4** (elastic metrics Elm^o and Elm). For any periodic sequences $S, Q \subset \mathbb{R}$, let $m = \text{lcm}(|S|, |Q|)$ be the lowest common multiple of their motif sizes. Consider the multiple sequences $\frac{m}{|S|}S$ and $\frac{m}{|Q|}Q$ from Definition 3.3, which have the same number m of motif points within the period intervals extended by the integer factors $\frac{m}{|S|}$ and $\frac{m}{|Q|}$, respectively.

The *oriented elastic metric* $\text{Elm}^o(S, Q) = \min_{\sigma \in C^o(m)} \|\text{SDL}^o(\frac{m}{|S|}S) - \sigma(\text{SDL}^o(\frac{m}{|Q|}Q))\|_\infty$ is minimized over all m cyclic permutations $\sigma \in C^o(m)$ of m ordered distances.

The (unoriented) *elastic metric* $\text{Elm}(S, Q) = \min_{\sigma \in C(m)} \|\text{SDL}(\frac{m}{|S|}S) - \sigma(\text{SDL}(\frac{m}{|Q|}Q))\|_\infty$ is minimized over all $2m$ cyclic permutations σ from the group $C(m)$. ▲

The sequences $S_3 = \{0, 1, 3\} + 6\mathbb{Z}$ and $6 - S_3 = \{0, 3, 5\} + 6\mathbb{Z}$ have the same number $m = 3$ of motif points. Since $\text{SDL}(S_3) = (1, 2, 3) = \text{SDL}(6 - S_3)$, the unoriented Elm distance is $\text{Elm}(S_3, 6 - S_3) = 0$. Actually, S_3 and $6 - S_3$ are related by reflection, which also follows from the first metric axiom in Theorem 3.6. Since $\text{SDL}^o(S_3) = (1, 2, 3) \neq (1, 3, 2) = \text{SDL}^o(6 - S_3)$, the oriented Elm distance is $\text{Elm}^o(S_3, 6 - S_3) = \|(1, 2, 3) - (1, 3, 2)\|_\infty = 1$, which is the minimum over all cyclic permutations of $(1, 3, 2)$, for example, $\|(1, 2, 3) - (3, 2, 1)\|_\infty = 2$.

To compute the oriented Elm distance between the periodic sequences $S_2 = \{0, 1\} + 3\mathbb{Z}$ and $S_3 = \{0, 1, 3\} + 6\mathbb{Z}$, we consider $3S_2 = \{0, 1, 3, 4, 6, 7\} + 9\mathbb{Z}$ and $2S_3 = \{0, 1, 3, 6, 7, 9\} + 12\mathbb{Z}$. Then $\text{SDL}^o(3S_2) = (1, 2, 1, 2, 1, 2)$ and $\text{SDL}^o(2S_3) = (1, 2, 3, 1, 2, 3)$. The component-wise comparison gives $\|(1, 2, 1, 2, 1, 2) - (1, 2, 3, 1, 2, 3)\|_\infty = 2$, but shifting any list by one position gives the smaller $\text{Elm}^o(S_2, S_3) = \|(2, 1, 2, 1, 2, 1) - (1, 2, 3, 1, 2, 3)\|_\infty = 1$ by Definition 3.4.

Definition 3.4 used any (not necessarily minimal) periods of sequences S, Q to define the elastic metrics. Lemma 3.5 shows that $\text{Elm}^o(S, Q)$ and $\text{Elm}(S, Q)$ are independent of a period and motivates the term *elastic* for metrics on sequences with extendable periods.

► **Lemma 3.5** (elastic metrics for multiples kS). For any periodic sequences $S, Q \subset \mathbb{R}$, the functions $\text{Elm}^o(kS, Q), \text{Elm}(kS, Q)$ from Definition 3.4 are independent of a factor $k \geq 1$. ■

Proof. To prove that $\text{Elm}^o(kS, Q) = \text{Elm}^o(S, Q)$, let $m = \text{lcm}(|S|, |Q|)$ and $m' = \text{lcm}(k|S|, |Q|)$ be the lowest common multiples. The proof for the unoriented metric is similar.

By Definition 3.4, the elastic metric $\text{Elm}^o(S, Q)$ is computed by comparing the sequences $\frac{m}{|S|}S$ and $\frac{m}{|Q|}Q$. Since $\hat{m} = \frac{m'}{m}$ is integer, the metric $\text{Elm}^o(kS, Q)$ uses the pair $\frac{m'}{|S|}S = \hat{m}\frac{m}{|S|}S$ and $\frac{m'}{|Q|}Q = \hat{m}\frac{m}{|Q|}Q$. The pairs for $\text{Elm}^o(S, Q)$ and $\text{Elm}^o(kS, Q)$ differ by the factor of \hat{m} .

The smallest distance lists $\text{SDL}^o(\hat{m}\frac{m}{|S|}S)$ and $\text{SDL}^o(\hat{m}\frac{m}{|Q|}Q)$ are obtained by concatenating \hat{m} copies of $\text{SDL}^o(\frac{m}{|S|}S)$ and $\text{SDL}^o(\frac{m}{|Q|}Q)$, respectively. For any cyclic permutation $\sigma \in C(\hat{m}m)$, the difference vector $\text{SDL}^o(\hat{m}\frac{m}{|S|}S) - \sigma(\text{SDL}^o(\hat{m}\frac{m}{|Q|}Q))$ is also a concatenation of \hat{m} copies of $\text{SDL}^o(\frac{m}{|S|}S) - \tau(\text{SDL}^o(\frac{m}{|Q|}Q))$, where the cyclic permutation $\tau \in C(m)$ was obtained by restricting $\sigma \in C(\hat{m}m)$ to the first block of m components, which is repeated \hat{m} times in both long vectors. Since the Minkowski norm $\|v\|_\infty$ remains unchanged concatenation of several copies of v , each resulting norm is the same, so $\text{Elm}^o(kS, Q) = \text{Elm}^o(S, Q)$. ◀

The proof of Lemma 3.5 used the fact the Minkowski norm $\|v\|_\infty$ remains unchanged when any vector $v = (x_1, \dots, x_m)$ is concatenated with copies of v . Other Minkowski norms $\|v\|_q = (\sum_{i=1}^m |x_i|^q)^{1/q}$ are not invariant under this transformation for $q \in [1, +\infty)$.

► **Theorem 3.6** (axioms for elastic metrics). $\text{Elm}^\circ, \text{Elm}$ from Definition 3.4 satisfy the metric axioms for any periodic sequences up to translation and isometry in \mathbb{R} , respectively. ■

Proof. To check the axioms for Elm° , we use that the Minkowski norm $\|v\|_\infty$ defines a metric satisfying all axioms in Problem 1.4(a,b,c). The proof for Elm is absolutely similar.

The first axiom requires that $\text{Elm}^\circ(S, Q) = 0$ if and only if S, Q are related by translation. The part *if* follows from the invariance of $\text{SDL}^\circ(S)$ in Lemma 3.2: if S, Q are related by translation then $\text{SDL}(S) = \text{SDL}(Q)$. The converse part *only if* assumes that $\text{Elm}^\circ(S, Q) = 0$. For $m = \text{lcm}(|S|, |Q|)$, there is a cyclic permutation $\sigma \in C^o(m)$ such that $\|\text{SDL}^\circ(\frac{m}{|S|}S) - \sigma(\text{SDL}^\circ(\frac{m}{|Q|}Q))\|_\infty = 0$. Then the oriented smallest distance lists $\text{SDL}^\circ(\frac{m}{|S|}S), \text{SDL}^\circ(\frac{m}{|Q|}Q)$ coincide up to cyclic permutation. The completeness of SDL° in Lemma 3.2 implies that the multiples $\frac{m}{|S|}S, \frac{m}{|Q|}Q$ are related by translation, then so are the original sequences S, Q .

To check the symmetry axiom $\text{Elm}^\circ(S, Q) = \text{Elm}^\circ(Q, S)$, let $m = \text{lcm}(|S|, |Q|)$. Any cyclic permutation $\sigma \in C^o(m)$ has the inverse σ^{-1} in the same group. The vectors $\text{SDL}^\circ(\frac{m}{|S|}S) - \sigma(\text{SDL}^\circ(\frac{m}{|Q|}Q))$ and $\text{SDL}^\circ(\frac{m}{|Q|}Q) - \sigma^{-1}(\text{SDL}^\circ(\frac{m}{|S|}S)) = -(\sigma^{-1}(\text{SDL}^\circ(\frac{m}{|S|}S)) - \text{SDL}^\circ(\frac{m}{|Q|}Q))$ differ by the cyclic permutation σ and sign, hence have the same Minkowski norm. Then the norm minimized over all cyclic permutations is the same, so $\text{Elm}^\circ(S, Q) = \text{Elm}^\circ(Q, S)$.

To check the triangle inequality for three periodic sequences S, Q, T , let $m = \text{lcm}(|S|, |Q|, |T|)$ be the lowest common multiple of their motif sizes. By Lemma 3.5, the function Elm can be computed by using any multiple of a periodic sequence. Hence we can use the periodic sequences $\frac{m}{|S|}S, \frac{m}{|Q|}Q, \frac{m}{|T|}T$, which all have the same number m of points in their motifs. Only for simplicity, we denote these multiple sequences by the same letters S, Q, T below.

In Definition 3.4, let $\sigma \in C^o(m)$ and $\tau \in C^o(m)$ be cyclic permutations that minimize $\text{Elm}^\circ(S, Q) = \|\text{SDL}^\circ(S) - \sigma(\text{SDL}^\circ(Q))\|_\infty$ and $\text{Elm}^\circ(Q, T) = \|\text{SDL}^\circ(Q) - \tau(\text{SDL}^\circ(T))\|_\infty$, respectively. We use the composition $\sigma \circ \tau$ and the triangle inequality for $\|v\|_\infty$: $\text{Elm}^\circ(S, T) \leq \|\text{SDL}^\circ(S) - \sigma \circ \tau(\text{SDL}^\circ(T))\|_\infty = \|\text{SDL}^\circ(S) - \sigma(\text{SDL}^\circ(Q)) + \sigma(\text{SDL}^\circ(Q) - \tau(\text{SDL}^\circ(T)))\|_\infty \leq \|\text{SDL}^\circ(S) - \sigma(\text{SDL}^\circ(Q))\|_\infty + \|\sigma(\text{SDL}^\circ(Q) - \tau(\text{SDL}^\circ(T)))\|_\infty = \text{Elm}^\circ(S, Q) + \text{Elm}^\circ(Q, T)$. ◀

4 Computability and continuity of elastic metrics under perturbations

For the elastic metrics on periodic sequences in \mathbb{R} , Theorem 3.6 fulfilled conditions (a,b,c) of Problem 1.4. The remaining conditions (d,e) hold due to Theorems 4.1 and 4.4.

Any periodic sequence $S = \{p_1, \dots, p_m\} + \mathbb{Z}$ can be given by its m motif points and period l . Hence the input size for all time complexities below is $|S| = m$.

► **Theorem 4.1** (time complexity of elastic metrics). For any periodic sequences $S, Q \subset \mathbb{R}$, let $m = \text{lcm}(|S|, |Q|)$ be the lowest common multiple of their motif sizes. The elastic metrics $\text{Elm}^\circ(S, Q)$ and $\text{Elm}(S, Q)$ from Definition 3.4 can be computed in time $O(m^2)$. ■

Proof. For ordered motif points $p_1, \dots, p_{|S|}$ within a period interval $[0, l)$ of S , in time $O(|S|)$, compute the list $\text{DL}(S)$ of distances $d_i = p_{i+1} - p_i$ for $i = 1, \dots, |S|$, where $p_{|S|+1} = p_1 + l$. The same distance $\text{DL}(Q)$ list for the periodic sequence Q similarly needs $O(|Q|)$ time.

By Definition 3.4 of elastic metrics, find the lowest common multiple $m = \text{lcm}(|S|, |Q|)$. The sequences $\frac{m}{|S|}S$ and $\frac{m}{|Q|}Q$ have distance lists $\text{DL}(\frac{m}{|S|}S), \text{DL}(\frac{m}{|Q|}Q)$ obtained by concatenating $\frac{m}{|S|}, \frac{m}{|Q|}$ copies of $\text{DL}(S), \text{DL}(Q)$, respectively, which uses only $O(m)$ time.

To compute $\text{Elm}^o(S, Q)$ and $\text{Elm}(S, Q)$ by Definition 3.4, we can use the above lists $\text{DL}(\frac{m}{|S|}S), \text{DL}(\frac{m}{|Q|}Q)$ without finding the smallest (up to cyclic permutations) distance lists because the elastic metrics are minimized over cyclic permutations anyway.

For each of $O(m)$ cyclic permutations σ for both metrics Elm^o and Elm , we find the minimum norm $\|\text{DL}(\frac{m}{|S|}S) - \sigma(\text{DL}(\frac{m}{|Q|}Q))\|_\infty$ in time $O(m)$, so the overall time is $O(m^2)$. ◀

Examples 4.2 and 4.3 show below that the bottleneck distance $d_B(S, Q)$ can be infinite or discontinuous even if S, Q are nearly identical lattices. Then Theorem 4.4 will prove that the elastic metrics are continuous under perturbations of any periodic sequences.

► **Example 4.2** (infinite bottleneck). We show that $S = \mathbb{Z}$ and $Q = (1 + \delta)\mathbb{Z}$ for any $\delta > 0$ have $d_B(S, Q) = +\infty$. Assuming that $d_B(S, Q)$ is finite, consider an interval $[-N, N] \subset \mathbb{R}$ containing $2N + 1$ points of S . If there is a bijection $g : S \rightarrow Q$ such that $|p - g(p)| \leq d_B$ for all points $p \in S$, then the image of $2N + 1$ points $S \cap [-N, N]$ under g should be within the interval $[-N - d_B, N + d_B]$. The last interval contains only $1 + \frac{2(N+d_B)}{1+\delta}$ points, which is smaller than $1 + 2N$ when $\frac{N+d_B}{1+\delta} < N$, $d_B < \delta N$, which is a contradiction for $N > \frac{d_B}{\delta}$. ■

Any 1-dimensional lattice $l\mathbb{Z}$ has $\text{SDL}^o(l\mathbb{Z}) = (l)$ consisting of a single distance l . Then any lattices $l\mathbb{Z}$ and $l'\mathbb{Z}$ have the elastic metrics $\text{Elm}(l\mathbb{Z}, l'\mathbb{Z}) = |l - l'| = \text{Elm}^o(l\mathbb{Z}, l'\mathbb{Z})$. In particular, \mathbb{Z} and $(1 + \delta)\mathbb{Z}$ in Example 4.2 have the small distance δ in the elastic metrics.

If we consider only periodic point sets $S, Q \subset \mathbb{R}^n$ with the same density (or unit cells of the same size), the bottleneck distance $d_B(S, Q) = \inf_{g:S \rightarrow Q} \sup_{a \in S} |a - g(a)|$ takes finite values and becomes a well-defined wobbling distance [11], which is unfortunately discontinuous.

► **Example 4.3** (discontinuous bottleneck). Slightly perturb the basis $(1, 0), (0, 1)$ of the integer lattice \mathbb{Z}^2 to the basis vectors $(1, 0), (\varepsilon, 1)$ of the new lattice Λ . We prove that $d_B(\Lambda, \mathbb{Z}^2) \geq \frac{1}{2}$ for any $\varepsilon > 0$. Map \mathbb{R}^2 by \mathbb{Z}^2 -translations to the unit square $[0, 1]^2$ with identified opposite sides (a torus). Then \mathbb{Z}^2 maps to one point represented by the corners of the square $[0, 1]^2$. The perturbed lattice Λ maps to the sequence of points $\{k\varepsilon \pmod{1}\}_{k=0}^{+\infty} \times \{0, 1\}$ in the horizontal edges. If $d_B(\Lambda, \mathbb{Z}^2) = r < \frac{1}{2}$, then all above points should be covered by the closed disks of the radius r centered at the corners of $[0, 1]^2$. For $0 < \varepsilon < \frac{1}{2} - r$, we can find a point $k\varepsilon$ that is between $r, 1 - r$, hence not covered by these disks, so $d_B(\Lambda, \mathbb{Z}^2) \geq \frac{1}{2}$. ■

► **Theorem 4.4** (continuity of elastic metrics). The elastic metrics from Definition 3.4 satisfy the continuity $\text{Elm}(S, Q) \leq \text{Elm}^o(S, Q) \leq 2d_B(S, Q)$ for any periodic sequences $S, Q \subset \mathbb{R}^n$ with a minimum inter-point distance more than $2d_B(S, Q)$. ■

Proof. By Lemma 3.5 one can replace S, Q by their multiples without changing the elastic metrics. If we need to consider multiples of S, Q to compute $\text{Elm}^o(S, Q)$ by Definition 3.4, we continue using the same symbols S, Q for simplicity because the bottleneck distance d_B also satisfies Lemma 3.5. So we can assume that both $S = \{p_1, \dots, p_m\} + l\mathbb{Z}$ and $Q = \{q_1, \dots, q_m\} + l'\mathbb{Z}$ have m motif points, though their periods l, l' can differ.

In Definition 1.3 of the bottleneck distance d_B , let $\varepsilon = d_B(S, Q)$ and $h : S \rightarrow Q$ be an optimal bijection such that $|p - h(p)| \leq \varepsilon$ for any point $p \in S$. When we move any point $p_i \in S$ to its bijective image $h(p_i) \in Q$ by ε smaller than a half of all inter-point distances,

any successive points $p_i, p_{i+1} \in S$ remain successive without colliding with other points of S in the open interval (p_i, p_{i+1}) for $i = 1, \dots, m$, where $p_{m+1} = p_1 + l$.

By cyclically renumbering the points of Q , one can now assume that $h(p_i) = q_i$ for $i = 1, \dots, m$. Since the points in each pair p_i, q_i are ε -close, every distance $d_i(S) = p_{i+1} - p_i$ differs from $d_i(Q) = q_{i+1} - q_i$ by at most 2ε . Hence the corresponding coordinates of the distance lists $(d_1(S), \dots, d_m(S))$ and $(d_1(Q), \dots, d_m(Q))$ differ by at most 2ε . By minimizing over cyclic permutations in Definition 3.4, we can get only a smaller Minkowski norm $\|v\|_\infty$ for the difference v of the lists above. Hence $\text{Elm}^o(S, Q) \leq 2\varepsilon$.

The inequality $\text{Elm}(S, Q) \leq \text{Elm}^o(S, Q)$ follows by Definition 3.4 because $\text{Elm}(S, Q)$ is minimized over the larger group of permutations in comparison with $\text{Elm}^o(S, Q)$. ◀

5 Metrics on isometry classes of high-dimensional periodic sequences

This section solves the analog of Problem 1.4 for high-dimensional periodic sequences $S \subset \mathbb{R} \times \mathbb{R}^{n-1}$ from Definition 1.5. Such a sequence models a multivariate time series [5] with values in \mathbb{R}^{n-1} . We focus on a natural equivalence $f \times g$, where f is a translation in \mathbb{R} and g is any isometry in \mathbb{R}^{n-1} . All results can be adapted to other cases, for example, including reflections in \mathbb{R} or restricting to orientation-preserving isometry in \mathbb{R}^{n-1} .

► **Definition 5.1** (cyclic distance matrix CDM). Let $Q = \{p_1, \dots, p_m\}$ be an ordered sequence of m points in \mathbb{R}^{n-1} . In the *cyclic distance matrix* $\text{CDM}(Q)$ of sizes $(m-1) \times m$, each element d_{ij} is the Euclidean distance $|p_j - p_{i+j}|$ for $i \in \{1, \dots, m-1\}$ and $j \in \{1, \dots, m\}$, where all indices are considered modulo m , for example, $p_{m+1} = p_1$. Any cyclic permutation $\sigma \in C^o(m)$ of indices acts on the matrix $\text{CDM}(Q)$ by cyclically shifting its m columns. ■

Any set Q of $m = 3$ ordered distinct points in \mathbb{R}^2 with pairwise distances a, b, c has $\text{CDM} = \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}$. For any $n \geq 1$, CDM has $\frac{n(n-1)}{2}$ repeated distances but has cyclically permutable columns. Proposition 5.2 can be folklore but we couldn't find a reference.

► **Proposition 5.2** (a complete isometry classification of finite ordered sequences). Any ordered sequences $S, Q \subset \mathbb{R}^{n-1}$ of m points are related by isometry of \mathbb{R}^{n-1} preserving the order of points if and only if their cyclic distance matrices are equal: $\text{CDM}(S) = \text{CDM}(Q)$. ■

Proof. Any isometry preserving the order of points makes the matrices $\text{CDM}(S) = \text{CDM}(Q)$ equal element-wise. Conversely, given a matrix $\text{CDM}(Q)$ from Definition 5.1, one can uniquely reconstruct $Q = \{p_1, \dots, p_m\}$ up to isometry in \mathbb{R}^{n-1} as follows. Fix the point p_1 at the origin of \mathbb{R}^{n-1} . The distance $d_{11} = |p_1 - p_2|$ in $\text{CDM}(Q)$ allows us to fix p_2 in the (positive) x_1 -coordinate axis of \mathbb{R}^{n-1} , while we can rotate the whole set Q around the x_1 -axis.

The distances $d_{12} = |p_1 - p_3|$ and $d_{21} = |p_2 - p_3|$ determine p_3 in the (x_1, x_2) -plane of \mathbb{R}^{n-1} and we can isometrically map Q by $f \in \text{O}(\mathbb{R}^{n-1})$ preserving this plane. If Q is not generic and the triangle inequality for d_{11}, d_{12}, d_{21} degenerates, so p_1, p_2, p_3 are in one line, the position of p_3 in the x_1 -axis is unique and we can rotate Q around the x_1 -axis.

We continue using further distances to similarly fix further points p_1, \dots, p_n until an initial subsequence p_1, \dots, p_k affinely generates \mathbb{R}^{n-1} . After that, any other point p_j for $j = k+1, \dots, m$ is uniquely determined in \mathbb{R}^{n-1} by n distances to the previous n points. If the full sequence Q affinely generates a smaller subspace $\mathbb{R}^k \subset \mathbb{R}^{n-1}$ for $k < n-1$, the previous construction determines Q up to isometry of \mathbb{R}^k , hence up to isometry of \mathbb{R}^{n-1} . ◀

► **Definition 5.3** (time-value invariant TVI). Let $S = \{p_1, \dots, p_m\} + l\vec{e}_1\mathbb{Z}$ be any high-dimensional periodic sequence in $\mathbb{R} \times \mathbb{R}^{n-1}$. The *time projection* $t : \mathbb{R} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ gives the 1-dimensional periodic sequence $t(S)$ with the *distance list* $\text{DL}^\circ(t(S)) = (d_1, \dots, d_m)$ of $d_i = t(p_{i+1}) - t(p_i)$ ordered as motif points for $i = 1, \dots, m$, where $t(p_{m+1}) = t(p_1) + l$. The *value projection* $v : \mathbb{R} \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$ gives the ordered *value sequence* $v(S)$ of m points $v(p_1), \dots, v(p_m) \in \mathbb{R}^{n-1}$. The time-value invariant is the pair $\text{TVI}(S) = (\text{DL}^\circ(t(S)), \text{CDM}(v(S)))$ considered up to cyclic permutations $\sigma \in C^\circ(m)$ simultaneously acting on the distance list $\text{DL}^\circ(t(S))$ and the cyclic distance matrix $\text{CDM}(v(S))$. ▲

The equivalence of high-dimensional sequences is considered in the sense of Definition 1.5.

► **Theorem 5.4** (classification of high-dimensional sequences). Any high-dimensional periodic sequences $S, Q \subset \mathbb{R} \times \mathbb{R}^{n-1}$ with m motif points are equivalent in the sense of Definition 1.5 if and only if there is a cyclic permutation $\sigma \in C^\circ(m)$ such that $\sigma(\text{TVI}(S)) = \text{TVI}(Q)$. ■

Proof. By Definition 1.5 any equivalence $f \times g$ acts by translations f on the time projection, hence preserves the distance list $\text{DL}^\circ(t(S))$ up to cyclic permutation. For a fixed time projection of m ordered points, the isometry g preserves the cyclic distance matrix $\text{CDM}(v(S))$. Given the invariant $\text{TVI}(S) = (\text{DL}^\circ(t(S)), \text{CDM}(v(S)))$, we reconstruct the periodic sequence $t(S)$ up to translation in \mathbb{R} by Proposition 3.2 and the ordered set $v(S)$ up to isometry in \mathbb{R}^{n-1} by Proposition 5.2. Then S is determined up to equivalence by $t(S)$ and $v(S)$. ◀

For any high-dimensional periodic sequence $S = \{p_1, \dots, p_m\} + l\vec{e}_1\mathbb{Z}$ in $\mathbb{R} \times \mathbb{R}^{n-1}$ and integer $k \geq 2$, its *multiple* is $kS = \{p_1 + jl\vec{e}_1, \dots, p_m + jl\vec{e}_1\}_{j=0}^{k-1} + kl\vec{e}_1\mathbb{Z}$. For any matrix M , the Minkowski norm $\|M\|_\infty$ is the maximum absolute value of its real elements.

Definition 5.5 extends the elastic metric Elm° to high-dimensional periodic sequences by taking the maximum with the Minkowski metric on cyclic distance matrices CDM of value projections. Instead of the maximum, one can combine two metrics in many other ways.

► **Definition 5.5** (elastic metric for high-dimensional sequences). For any generic periodic sequences $S, Q \subset \mathbb{R} \times \mathbb{R}^{n-1}$, let $m = \text{lcm}(|S|, |Q|)$ be the lowest common multiple of their motif sizes. Consider the multiple sequences $\frac{m}{|S|}S$ and $\frac{m}{|Q|}Q$, which have the same number m of motif points. The *oriented elastic metric* $\text{Elm}^\circ(S, Q) = \min_{\sigma \in C^\circ(m)} \max\{d_t, d_v\}$ is minimized over all m permutations $\sigma \in C^\circ(m)$, where $d_t = \|\text{DL}^\circ(\frac{m}{|S|}S) - \sigma(\text{DL}^\circ(\frac{m}{|Q|}Q))\|_\infty$, $d_v = \|\text{CDM}(v(\frac{m}{|S|}S)) - \sigma(\text{CDM}(v(\frac{m}{|Q|}Q)))\|_\infty$, σ acts on CDM by cyclic shifts of columns. ▲

Almost all properties in Theorem 5.7 for high-dimensional elastic metrics will easily follow from similar properties in dimension $n = 1$. The only extra step is the continuity of the cyclic distance matrix CDM with respect to the maximum deviation of points below.

► **Lemma 5.6** (continuity of CDM). Let ordered sets $S = \{p_1, \dots, p_m\}$ and $Q = \{q_1, \dots, q_m\}$ have a maximum deviation $\varepsilon = \max_{i=1, \dots, m} |p_i - q_i|$. Then $\|\text{CDM}(S) - \text{CDM}(Q)\|_\infty \leq 2\varepsilon$. ■

Proof. The triangle inequality for the Euclidean distance implies that the corresponding elements of the cyclic distance matrices from Definition 5.1 differ by at most 2ε as follows: $|d_{ij}(S) - d_{ij}(Q)| = |p_j - p_{i+j} - (q_j - q_{i+j})| \leq |p_i - q_j| + |p_{i+j} - q_{i+j}| \leq 2\varepsilon$. ◀

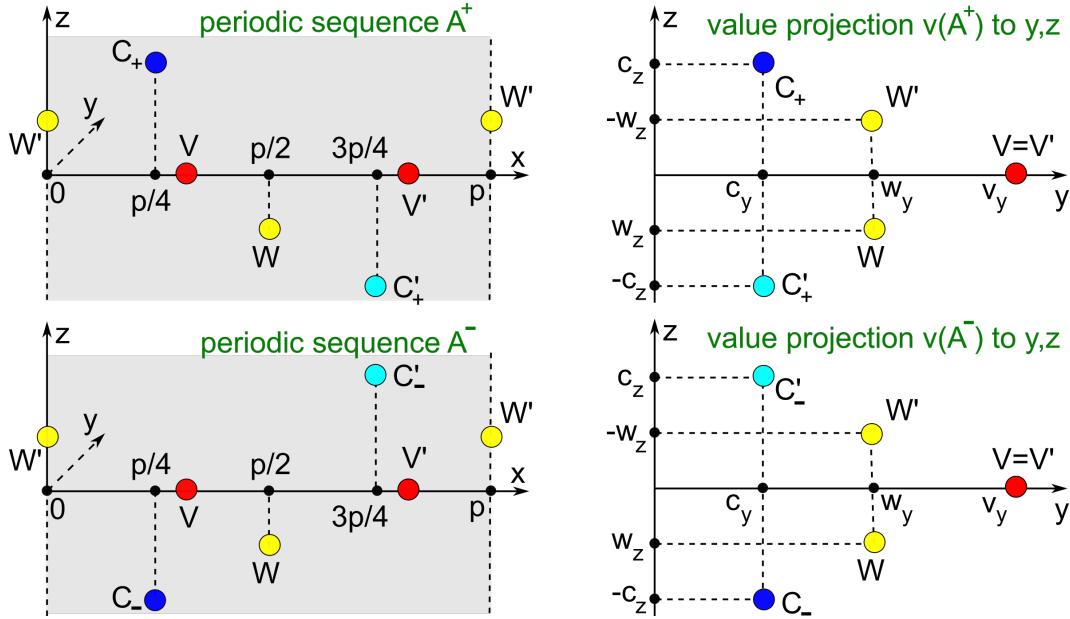
► **Theorem 5.7** (solution to Problem 1.4 for high-dimensional periodic sequences). For any periodic sequences $S, Q \subset \mathbb{R} \times \mathbb{R}^{n-1}$, let $m = \text{lcm}(|S|, |Q|)$ be the lowest common multiple of their motif sizes. The oriented elastic metric $\text{Elm}^\circ(S, Q)$ can be computed in time $O(m^3)$, satisfies all metric axioms, and continuity $\text{Elm}^\circ(S, Q) \leq 2d_B(S, Q)$ in the bottleneck distance when $2d_B(S, Q)$ is smaller than a minimum distance between all points of $t(S), t(Q)$. ▲

Proof. All arguments follow the proofs for the case $n = 1$. For the time complexity, the extra factor $m \geq n$ emerges because the cyclic distance matrices $\text{CDM}(S), \text{CDM}(Q)$ have at most $m \times m$ distances. Computing the Minkowski norm of the matrix difference needs $O(m^2)$ operations for each of m cyclic permutations to find the metric Elm° by Definition 5.5.

To check the metric axioms, we start from the analog of Lemma 3.5 implying that the elastic metric Elm° is independent of a non-minimum period also for high-dimensional sequences. The first axiom follows from Propositions 3.2 and 5.2 because $\text{Elm}^\circ(S, Q) = 0$ implies that there is a cyclic permutation $\sigma \in C^\circ(m)$ matching $\text{DL}^\circ(S)$ with $\text{DL}^\circ(Q)$ and $\text{CDM}(S)$ with $\text{CDM}(Q)$, hence making the sequences S, Q equivalent by Definition 1.5.

The symmetry axiom easily follows similarly to the proof of Theorem 3.6 because any cyclic permutation $\sigma \in C^\circ(m)$ has its inverse σ^{-1} in the same group. To prove the triangle inequality, we find multiples of S, Q, T that have the same number m of motif points. We continue denoting these sequences by S, Q, T for simplicity. Let $\sigma, \tau \in C^\circ(m)$ be cyclic permutations that minimize $\text{Elm}^\circ(S, Q)$ and $\text{Elm}^\circ(Q, T)$, respectively, in Definition 5.5. Then $d_t = \|\text{DL}^\circ(S) - \sigma \circ \tau(\text{DL}^\circ(T))\|_\infty \leq \|\text{DL}^\circ(S) - \sigma(\text{DL}^\circ(Q))\|_\infty + \|\sigma(\text{DL}^\circ(Q) - \tau(\text{DL}^\circ(T)))\|_\infty \leq \text{Elm}^\circ(S, Q) + \text{Elm}^\circ(Q, T)$. Since the above inequalities hold for d_v after replacing DL° by CDM , we get $\text{Elm}^\circ(S, T) = \max\{d_t, d_v\} \leq \text{Elm}^\circ(S, Q) + \text{Elm}^\circ(Q, T)$. The final inequality $\text{Elm}^\circ(S, Q) \leq 2d_B(S, Q)$ directly follows from Theorem 4.4 and Lemma 5.6. \blacktriangleleft

► **Example 5.8** (hardest periodic sequences). The latest counter-example [31, Fig. 2] to completeness of all past distance-based invariants is the pair of the sequences $A^\pm \subset \mathbb{R} \times \mathbb{R}^2$ with a period p and 6 motif points: $A^+ = \{W', C_+, V, W, C'_+, V'\} + p\vec{e}_1\mathbb{Z}$ and $A^- = \{W', C_-, V, W, C'_-, V'\} + p\vec{e}_1\mathbb{Z}$, where $V = (v_x, v_y, 0)$, $W = (\frac{p}{2}, w_y, w_z)$, $C_\pm = (\frac{p}{4}, c_y, \pm c_z)$.



■ **Figure 4** These periodic sequences $A^\pm \subset \mathbb{R} \times \mathbb{R}^2$ from [31, Fig. 2] have identical past invariants.

Any point with a dash is obtained by $g(x, y, z) = (x + \frac{p}{2}, y, -z)$. The time projections are identical: $t(A^\pm) = (0, \frac{p}{4}, v_x, \frac{p}{2}, \frac{3p}{4}, \frac{p}{2} + v_x)$. Assuming that $v_x \in (\frac{p}{4}, \frac{p}{2})$ as in Fig. 4, the distance lists are $\text{DL}^\circ(t(A^\pm)) = (\frac{p}{4}, v_x - \frac{p}{4}, \frac{p}{2} - v_x, \frac{p}{4}, v_x - \frac{p}{4}, \frac{p}{2} - v_x)$. The ordered value projections are $v(A^\pm) = \{(w_y, -w_z), (c_y, \pm c_z), (v_y, 0), (w_y, w_z), (c_y, \mp c_z), (v_y, 0)\}$. Then

$$\text{CDM}(A^+) = \begin{pmatrix} d_{11} & d_{12} & d_{21} & d_{11} & d_{12} & d_{21} \\ d_{21} & d_{22} & d_{12} & d_{21} & d_{22} & d_{12} \\ 2|w_z| & 2|c_z| & 0 & 2|w_z| & 2|c_z| & 0 \end{pmatrix} \neq$$

$$\text{CDM}(A^-) = \begin{pmatrix} d_{22} & d_{12} & d_{21} & d_{22} & d_{12} & d_{21} \\ d_{21} & d_{11} & d_{12} & d_{21} & d_{11} & d_{12} \\ 2|w_z| & 2|c_z| & 0 & 2|w_z| & 2|c_z| & 0 \end{pmatrix}, \text{ where}$$

$$d_{11} = \sqrt{(w_y - c_y)^2 + (w_z + c_z)^2}, \quad d_{12} = \sqrt{(c_y - v_y)^2 + c_z^2},$$

$d_{22} = \sqrt{(w_y - c_y)^2 + (w_z - c_z)^2}, \quad d_{21} = \sqrt{(w_y - v_y)^2 + w_z^2}$. The matrix difference has the Minkowski norm $\|\text{CDM}(A^+) - \text{CDM}(A^-)\|_\infty = |d_{11} - d_{22}| > 0$ unless $c_z = 0$ or $w_z = 0$. If $c_z = 0$, A^\pm are identical. If $w_z = 0$, then A^\pm are isometric by $g(x, y, z) = (x + \frac{p}{2}, y, -z)$.

If both $c_z, w_z \neq 0$, then $\text{Elm}^o(A^+, A^-)$ is obtained by minimizing over 6 cyclic permutations $\sigma \in C^o(6)$. The trivial permutation and the shift by 3 positions give $|d_{11} - d_{12}|$. Any other permutation gives $d_t = \max\{v_x - \frac{p}{4}, \frac{p}{2} - v_x\}$ from comparing $\text{DL}^o(t(A^+))$ with $\sigma(\text{DL}^o(t(A^-)))$ and $d_v = \max\{|a - b|\}$ maximized for all pairs $a, b \in \{d_{11}, d_{12}, d_{21}, d_{22}\}$. In all cases, the elastic metric is positive: $\text{Elm}^o(A^+, A^-) \geq |d_{11} - d_{22}| > 0$. Hence the time-value invariant TVI from Definition 5.3 distinguished this challenging pair $A^+ \not\cong A^-$. ■

In conclusion, the key contribution is the introduction of easily computable and continuous elastic metrics in Definition 3.4 and 5.5 on periodic sequences considered up to several equivalences. Main Theorems 3.6, 4.1, 4.4, 5.4, 5.7 resolved Problem 1.4 for periodic sequences in all dimensions. Even the case $n = 1$ was open and required several new ideas such as using the specific Minkowski metric $\|v\|_\infty$, which respects multiple periods.

Problem 1.4 remains open for dimensions $n \geq 2$ because the last two conditions on exact computability and continuity are the major obstacle for all attempts. For instance, Voronoi domains are combinatorially unstable for finite sets and even for 2-dimensional lattices.

Example 5.8 illustrates how easily one can compute the elastic metric and distinguish any high-dimensional periodic sequences by their complete invariant TVI as expected by Theorem 5.4. The recent developments in Periodic Geometry include continuous maps of Lattice Isometry Spaces in dimension two [23, 9] and three [22, 10], and applications to materials science [32, 38]. The latest ultra-fast and generically complete Pointwise Distance Distributions [35] justified the Crystal Isometry Principle (CRISP) saying that all real periodic crystals live in a common space of isometry classes of periodic point sets continuously parameterised by their complete invariants such as isosets [3].

This research was supported by the £3.5M EPSRC grant ‘Application-driven Topological Data Analysis’ (2018-2023, EP/R018472/1), the £10M Leverhulme Research Centre for Functional Materials Design, and the Royal Academy of Engineering Fellowship ‘Data Science for Next Generation Engineering of Solid Crystalline Materials’ (2021-2023, IF2122/186).

References

- 1 H Alt, K Mehlhorn, H Wagener, and E Welzl. Congruence, similarity, and symmetries of geometric objects. *Discrete & Comp. Geometry*, 3:237–256, 1988.
- 2 Olga Anosova and Vitaliy Kurlin. Introduction to periodic geometry and topology. 2021. URL: <https://arxiv.org/abs/2103.02749>.
- 3 Olga Anosova and Vitaliy Kurlin. An isometry classification of periodic point sets. In *LNCS Proceedings of Discrete Geometry and Mathematical Morphology*, 2021. URL: https://link.springer.com/chapter/10.1007/978-3-030-76657-3_16.

- 4 Olga Anosova and Vitaliy Kurlin. Density functions of periodic sequences. *arxiv:2205.02226*, 2022. URL: <http://kurlin.org/projects/periodic-geometry-topology/densities1D.pdf>.
- 5 Kim Bale, Paul Chapman, Nick Barraclough, Jon Purdy, Nizamettin Aydin, and Paul Dark. Kaleidomaps: A new technique for the visualization of multivariate time-series data. *Information Visualization*, 6(2):155–167, 2007.
- 6 Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape contexts. *Transactions PAMI*, 24(4):509–522, 2002.
- 7 M. Boutin and G. Kemper. On reconstructing n-point configurations from the distribution of distances or areas. *Advances in Applied Mathematics*, 32(4):709–735, 2004.
- 8 Peter Brass and Christian Knauer. Testing the congruence of d-dimensional point sets. In *Proceedings of SoCG*, pages 310–314, 2000.
- 9 Matthew Bright, Andrew I Cooper, and Vitaliy Kurlin. Geographic-style maps for 2-dimensional lattices. *arxiv:2109.10885 (early draft)*, 2021. URL: <http://kurlin.org/projects/periodic-geometry-topology/lattices2Dmap.pdf>.
- 10 Matthew Bright, Andrew I Cooper, and Vitaliy Kurlin. Welcome to a continuous world of 3-dimensional lattices. *arxiv:2109.11538 (early draft)*, 2021. URL: <http://kurlin.org/projects/periodic-geometry-topology/lattices3Dmap.pdf>.
- 11 Hans-Georg Carstens, Walter A Deuber, Wolfgang Thumser, and Elke Koppenrade. Geometrical bijections in discrete lattices. *Combinatorics, Probability and Computing*, 8(1-2):109–129, 1999.
- 12 B.N. Delone, N.P. Dolbilin, M.I. Shtogrin, and R.V. Galiulin. A local criterion for regularity of a system of points. In *Dokl. Akad. Nauk SSSR*, volume 227, pages 19–21, 1976.
- 13 N Dolbilin and M Bouniaev. Regular t-bonded systems in \mathbb{R}^3 . *European Journal of Combinatorics*, 80:89–101, 2019.
- 14 N. Dolbilin, J Lagarias, and M. Senechal. Multiregular point systems. *Discrete & Computational Geometry*, 20(4):477–498, 1998.
- 15 NP Dolbilin. Delone sets in \mathbb{R}^3 with $2r$ -regularity conditions. *Proceedings of the Steklov Institute of Mathematics*, 302:161–185, 2018.
- 16 N Dym and S Kovalsky. Linearly converging quasi branch and bound algorithms for global rigid registration. In *Proceedings ICCV*, pages 1628–1636, 2019.
- 17 Herbert Edelsbrunner, Teresa Heiss, Vitaliy Kurlin, Philip Smith, and Mathijs Wintraecken. The density fingerprint of a periodic point set. In *SoCG*, volume 189, pages 32:1–32:16, 2021.
- 18 Philip Hans Franses and Richard Paap. *Periodic time series models*. Oxford Univ. Press, 2004.
- 19 Cosmin Grigorescu and Nicolai Petkov. Distance sets for shape filters and shape recognition. *IEEE transactions on image processing*, 12(10):1274–1286, 2003.
- 20 F. Grünbaum and C. Moore. The use of higher-order invariants in the determination of generalized patterson cyclotomic sets. *Acta Cryst. A*, 51:310–323, 1995.
- 21 Heuna Kim and Günter Rote. Congruence testing of point sets in 4 dimensions. *arXiv:1603.07269*, 2016.
- 22 Vitaliy Kurlin. A complete isometry classification of 3-dimensional lattices. *arxiv:2201.10543*, 2022. URL: <http://kurlin.org/projects/periodic-geometry-topology/lattices3Dmaths.pdf>.
- 23 Vitaliy Kurlin. Mathematics of 2-dimensional lattices. *arxiv:2201.05150 (early draft)*, 2022. URL: <http://kurlin.org/projects/periodic-geometry-topology/lattices2Dmaths.pdf>.
- 24 Leo Liberti and Carlile Lavor. *Euclidean distance geometry: an introduction*. Springer, 2017.
- 25 Siddharth Manay, Daniel Cremers, Byung-Woo Hong, Anthony J Yezzi, and Stefano Soatto. Integral invariants for shape matching. *IEEE Transactions on pattern analysis and machine intelligence*, 28(10):1602–1618, 2006.
- 26 Haggai Maron, Nadav Dym, Itay Kezurer, Shahar Kovalsky, and Yaron Lipman. Point registration via efficient convex relaxation. *ACM Trans. on Graphics*, 35(4):1–12, 2016.

- 27 Facundo Mémoli. Gromov–wasserstein distances and the metric approach to object matching. *Foundations of Comp. Mathematics*, 11(4):417–487, 2011.
- 28 Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Shape distributions. *ACM Transactions on Graphics (TOG)*, 21(4):807–832, 2002.
- 29 A Patterson. Ambiguities in the x-ray analysis of crystal structures. *Phys. Rev.*, 65:195, 1944.
- 30 Helmut Pottmann, Johannes Wallner, Qi-Xing Huang, and Yong-Liang Yang. Integral invariants for robust geometry processing. *Comp. Aided Geom. Design*, 26(1):37–60, 2009.
- 31 Sergey N Pozdnyakov and Michele Ceriotti. Incompleteness of graph convolutional neural networks for points clouds in three dimensions. *arXiv:2201.07136*, 2022.
- 32 Jakob Ropers, Marco M Mosca, Olga Anosova, Vitaliy Kurlin, and Andrew I Cooper. Fast predictions of lattice energies by continuous isometry invariants of crystal structures. In *Proceedings of DACOMSIN: Data and Computation for Materials Science and Innovation*, 2021. URL: <https://arxiv.org/abs/2108.07233>.
- 33 Mauro R Ruggeri and Dietmar Saupe. Isometry-invariant matching of point set surfaces. In *3DOR*, pages 17–24. Citeseer, 2008.
- 34 S Ward and G Sadiq. Introduction to the cambridge structural database—a wealth of knowledge gained from a million structures. *CrystEngComm*, 22(43):7143–7144, 2020.
- 35 Daniel Widdowson and Vitaliy Kurlin. Pointwise distance distributions. *arXiv:2108.04798*, 2021. URL: <http://kurlin.org/projects/periodic-geometry-topology/PDD.pdf>.
- 36 Daniel Widdowson, Marco Mosca, Angeles Pulido, Vitaliy Kurlin, and Andrew Cooper. Average minimum distances of periodic point sets - fundamental invariants for mapping all periodic crystals. *MATCH Communications in Mathematical and in Computer Chemistry*, 87:529–559, 2022. doi:10.46793/match.87-3.529W.
- 37 Jiaolong Yang, Hongdong Li, Dylan Campbell, and Yunde Jia. Go-ICP: A globally optimal solution to 3D ICP point-set registration. *IEEE transactions on pattern analysis and machine intelligence*, 38(11):2241–2254, 2015.
- 38 Qiang Zhu, Jay Johal, Daniel Widdowson, Zhongfu Pang, Boyu Li, Christopher Kane, Vitaliy Kurlin, Graeme Day, Marc Little, and Andrew Cooper. Analogy powered by prediction and structural invariants: Computationally-led discovery of a mesoporous hydrogen-bonded organic cage crystal. *J Amer Chem Soc*, 2022, to appear.

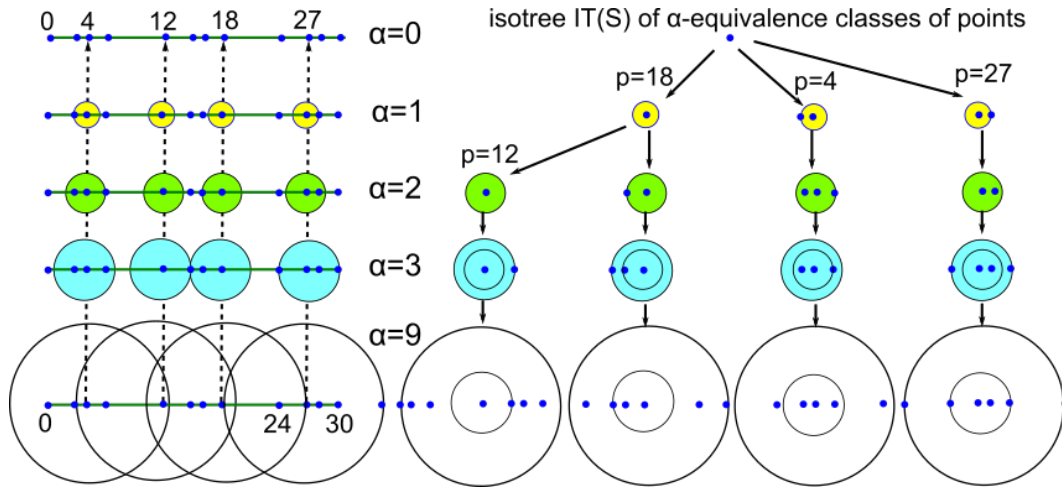
A Appendix: a continuous metric on complete invariant isosets

This appendix simplifies the complete invariant isoset from [3] to the case of a periodic sequence in \mathbb{R} . All underlying concepts were inspired by the seminal work of Delone, Dolbilin and co-authors [12, 14, 15, 13]. Consider any periodic sequences $S = \{p_1, \dots, p_m\} + l\mathbb{Z}$ only up to translation in \mathbb{R} . For any point $p \in S$ and a radius $\alpha \geq 0$, the α -cluster $C(S, p; \alpha) = S \cap [p - \alpha, p + \alpha]$ consists of all points of S at a distance at most α from p .

Two points $p, q \in S$ are called α -equivalent if their α -clusters coincide after the centers are matched by translation: $C(S, p; \alpha) - p = C(S, q; \alpha) - q$. For any fixed radius $\alpha \geq 0$, the α -partition $P(S; \alpha)$ is the splitting of S into α -equivalence classes of points, see Fig. 5.

If α is smaller than the minimum inter-point distance $\min_{i=1, \dots, m} d_i$ of S , the α -partition $P(S; \alpha)$ consists of a single class whose α -cluster consisting of only the central point. When the radius α is growing, α -equivalence classes can split into smaller classes for larger $\alpha' > \alpha$ up to a maximum of m classes, where m is the number of motif points. All α -partitions $P(S; \alpha)$ are organized into the *isotree* $IT(S)$ whose nodes represent α -equivalence classes.

The *bridge* length $\beta(S) = \max_{i=1, \dots, m} d_i$ is the maximum distance between successive points of S . A radius $\alpha \geq \beta(S)$ is called *stable* if the α -partition stabilizes in the sense that $P(S; \alpha) = P(S; \alpha - \beta)$. Then [2, Lemma 23] proves that all stable radii form the interval



■ **Figure 5** The periodic sequence $S_4 = \{0, 3, 4, 6\} + 12\mathbb{Z}$ whose α -clusters (intervals shown by colored disks for easier visibility) with radii $\alpha = 0, 1, 2, 3, 9$ represent points in the isotree $IT(S_4)$.

$[\alpha(S), +\infty)$, where $\alpha(S)$ is the minimum stable radius of S . Also, [2, Lemma 18] gives the easy upper bound, which simplifies for a periodic sequence S to $\alpha(S) \leq \beta(S) + l$, where l is a minimum period of S . This upper bound is achieved for any 1-dimensional lattice $S = l\mathbb{Z}$.

► **Definition A.1** (isotree $I(S; \alpha)$ of a periodic sequence). For any periodic sequence $S = \{p_1, \dots, p_m\} + l\mathbb{Z}$ and a radius $\alpha \geq 0$, each α -equivalence class from $P(S; \alpha)$ consisting of (say) k points from $\{p_1, \dots, p_m\}$ can be associated with the α -cluster $\xi = C(S, p; \alpha) - p$ shifted to the origin of \mathbb{R} . Define the *weight* of this centered α -cluster ξ as $w = k/m$. The *isotree* $I(S; \alpha)$ is the unordered set of all centered α -clusters $(\xi; w)$ with weights. ■

The isotree $I(S; \alpha)$ without weights can be viewed as a set of points in the isotree $IT(S)$ at the radius α . The size of $I(S; \alpha)$ equals the number $|P(S; \alpha)|$ of α -equivalence classes in the α -partition. Though $I(S; \alpha)$ depends on α because α -clusters grow in α , [3, Theorem 10] distinguishes all periodic point sets $S, Q \subset \mathbb{R}^n$ up to isometry by comparing their isotrees at a common (maximum) stable radius α of S, Q .

Any centered α -cluster representing a point p in a periodic sequence S can be considered as a list of vectors or signed distances from p to its neighbors within S up to the radius α . Hence the isotree extended Patterson's idea of neighbor distances to higher dimensions.